8-2010

Transcriptional regulation of the Shigella flexneri icsp Promoter: Silencing and anti-silencing by H-NS and VirB

Dustin Harrison
University of Nevada, Las Vegas

TRANSCRIPTIONAL REGULATION OF THE *SHIGELLA FLEXNERI ICSP* PROMOTER: SILENCING AND ANTI-SILENCING BY H-NS AND VIRB

by

Lieutenant Dustin John Harrison
Medical Service Corps
U.S. Navy

Bachelor of Science
Marycrest International University
1997

Master of Science
University of Wyoming
2000

A dissertation submitted in partial fulfillment of
the requirements for the

Doctor of Philosophy in Biological Sciences
School of Life Sciences
College of Science

Graduate College
University of Nevada, Las Vegas
August 2010

UNIVERSITY OF NEVADA LAS VEGAS

THE GRADUATE COLLEGE

We recommend that the dissertation prepared under our supervision by

Lieutenant Dustin John Harrison

entitled

Transcriptional Regulation of the *Shigella flexneri icsp* Promoter: Silencing and Anti-silencing by H-NS and VirB

be accepted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Biological Sciences
School of Life Sciences

Helen Wing, Committee Chair

Dennis Bazylinski, Committee Member

Eduardo Robleto, Committee Member

Marshall Monteville, Outside Committee Member

Patricia Cruz, Graduate Faculty Representative

Ronald Smith, Ph. D., Vice President for Research and Graduate Studies
and Dean of the Graduate College

August 2010

ABSTRACT

Transcriptional Regulation of the *Shigella flexneri* promoter: Silencing and Anti-silencing by H-NS and VirB

by

Lieutenant Dustin John Harrison
Medical Service Corps
U.S. Navy

Dr. Helen J Wing, Examination Committee Chair
Assistant Professor of Science
University of Nevada, Las Vegas

Shigella species are gram-negative intracellular pathogens that cause bacillary dysentery in humans. Many genes required for virulence of *Shigella* are carried on a large 230 kb plasmid and many of these are under the transcriptional control of the histone-like nucleoid structuring protein (H-NS) and by the major virulence regulator VirB. At the non-permissive temperature of 30^{0}C, H-NS represses transcription, while at 37^{0}C VirB alleviates this repression. This mechanism of gene regulation has been coined "silencing/anti-silencing" and is commonly found in many important bacterial pathogens including *Salmonella* spp. and *Yersinia* spp. The *icsP* gene, encoded by the *Shigella* virulence plasmid, is positively controlled by VirB and negatively by H-NS. The *icsP* gene encodes an outermembrane protease respsonsible for maintaining the tight polar cap of the actin polymerization protein IcsA, which is involved in the inter-, and intracellular spread of *Shigella*.

Our work has revealed that sequences located over 1 kb upstream of the *icsP* annotated transcription start site (+1) are needed for the VirB-dependent regulation of the

icsP promoter. Using site directed mutagenesis we identified two DNA sequences that are required for the VirB dependent regulation of the icsP promoter, both of which display good matches to the reported consensus VirB binding site. We demonstrate that sequences located upstream of position -665 relative to the annotated transcription start site are needed for complete H-NS-mediated silencing of the *icsP* promoter.

Using electrophoretic mobility shift assays (EMSAs), and DNase I footprint analysis we show that VirB and H-NS bind directly to DNA sequences located both upstream of -665 and downstream of -213 by and that these regions are predicted to display high levels of intrinsic curvature. While we demonstrate that VirB functions to solely de-repress the *icsP* promoter, our EMSA data indicate that VirB and H-NS are capable of binding to the full upstream intergenic region (1232 bp) of the *icsP* gene simultaneously. These *in vitro* data suggest that VirB may function to relocate H-NS bound to the region immediately upstream of the *icsP* gene, rather than displacing H-NS from the DNA.

Our data also suggest that disruption of single regions of predicted curvature, contained within the upstream intergenic region of *icsP*, does not have an effect on either the ability of H-NS to silence the promoter or the ability of VirB to alleviate the H-NS induced repression.

Taken together our data suggest that remote regulation of the *icsP* promoter requires promoter proximal sequences that act in concert with upstream sequences. Our findings raise the possibility that other bacterial promoters may be regulated by DNA-binding proteins binding to remote DNA sequence elements. My work improves our understanding of transcriptional silencing and anti-silencing, a regulatory event that controls the expression of many virulence genes in many important enteric pathogens.

Harrison, Dustin John, is a LT, MSC, USN, Microbiologist, NAVMED MPT&E.

"The views expressed in this article are those of the author and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, nor the U.S. Government."

ACKNOWLEDGMENTS

I would like to like the United States Navy, and the Navy Medicine Manpower, Personnel, Training & Education Command (NM MPT&E) for the opportunity to pursue a PhD and the funding for the last three years.

I would like to thank my advisor, Dr. Helen Wing for her guidance and knowledge throughout these past three years. I would also like to thank my committee Dr. Eduardo Robleto, Dr. Dennis Bazylinski, and Dr. Patricia Cruz for their much needed advice and support. I would like to give a special thanks to CDR Marshall Monteville, PhD, for his inspiration, support, and guidance. I would also like to thank Dr. Martin Roop and Dr. Eric Anderson, from East Carolina University, for all their help and useful discussions, especially with the gel shifts.

I would like to thank the members of the Wing Lab both past and present, including: Eun-Hae Kim, Maria Castellanos, Chris Hensley, Lia Africa, Jensara Clay, Krystle Pew, and everyone else I have forgotten. In particular I would like to thank Stephanie Labahn for all her help and support. A very special acknowledgment goes to Amber Howerton who has been a great friend and who has helped me through this degree.

I want to thank my family even though they have hardly heard from me or seen me in the last three years I couldn't have done it without their support.

And most importantly I want to thank my beautiful wife Sara without whom I don't think I would have finished this. I love you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The choreography of cellular processes is quite possibly one of the least understood, and most interesting, challenges facing molecular biologists. These processes, e.g. transcription, translation, replication, trafficking, and the myriad of mechanisms that control and regulate them are central to understanding life, not only in eukaryotes but prokaryotes as well. The regulation of just one of these processes, transcription, contains multiple pathways designed to effectively control gene expression by down-regulating the amount of gene products produced. At any one time, an organism only expresses a fraction of its total potential, the rest is "turned off" (Yarmolinsky, 2000). There are two ways in which a gene may be "turned off." One way is repression, which is defined as the binding of a protein at a promoter specific binding-site, which interferes locally with RNA polymerase (RNAP) function (Yarmolinsky, 2000). The other way is by a process termed transcriptional silencing, which can be defined as a process that renders appreciable regions of DNA inaccessible to proteins such as RNA polymerase, DNA modifying enzymes, etc., to which it would otherwise be able (Dorman *et al*, 1999, Yarmolinsky, 2000). However, the distinction between repression and silencing seems arbitrary and it is often difficult to distinguish the two, e.g. a silencing protein can interfere with RNAP function but perhaps the specific binding site is not yet known. Transcriptional regulation is a critical process that allows prokaryotic cells to quickly adapt to constantly changing environmental conditions, nutrient concentrations, stress, and antibiotics in order to exploite its desired niche.

1.1 Transcriptional silencing

Gram-negative bacteria, like *E. coli*, contain at least 12 distinct types of nucleoid-associated proteins, each with its own expression pattern and DNA-binding differences (Ali Azam *et al*, 1999, Dorman, 2004). These nucleoid-associated proteins (NAPs) are direct DNA binding proteins that often have global influences on transcriptional expression. Many NAPs are involved in down-regulating transcription while others are involved in up-regulating transcription, and still others have the ability to do both. They are often referred to as histone-like, however, this is based on the functional similarity to eukaryotic histones and is not based on sequence or amino-acid similarity (Dorman, 2004). Like eukaryotic histones, many of these proteins have been shown to condense DNA *in vitro* and *in vivo.* Although many types of NAPs exist the most extensively studied is the histone-like nucleoid structuring protein, H-NS, its paralogue StpA, the factor for inversion stimulation (Fis), the heat-unstable protein (HU), and the integration host factor protein (IHF). My research has focused mainly on H-NS and its regulation of the *Shigella flexneri* gene *icsP*.

The histone-like nucleoid structuring protein

H-NS is a small, 15 kDa, chromosomally encoded, protein with an intracellular concentration of ~20 μM (Ali Azam *et al*., 1999, Bouffartigues *et al*, 2007, Spassky *et al*, 1984). The protein is expressed at a relatively constant level throughout the growth cycle (Dersch *et al*, 1993). H-NS functions as a global repressor of gene transcription and is responsible for the organization and higher-order structure of the bacterial chromosome (Ali Azam *et al*., 1999, Dorman, 2004). The concept of transcriptional silencing by H-NS arises from the observations that certain operons in *E. coli* and *Salmonella* are under

the negative control of this protein, with the important H-NS binding regions being located outside of the canonical promoter sequences (Goransson *et al*, 1990, Higgins *et al*, 1988). It was shown that the presence of H-NS formed a barrier to gene expression that was not easily overcome by either point mutations or small insertions into the regions to which H-NS was shown to bind (Schnetz, 1995, Yarmolinsky, 2000). Furthermore it was suggested that silencing occurred as a result of H-NS interacting with sequences located both upstream and downstream of the genes studied, and then interacting with itself to form looped structures (Schnetz, 1995). Also, it has been proposed that binding of H-NS to DNA affects the superhelicity of bacterial promoters, thereby modulating transcription. H-NS has also been proposed to be an environmental sensing molecule modulating gene expression in response to changes in DNA superhelicity as a result of temperature, pH, osmolarity, and growth phase (Atlung & Ingmer, 1997, Fang & Rimsky, 2008, Dorman, 2004,). While evidence suggests that the relationship between H-NS and changes in environmental conditions is complex, it is clear that H-NS is not simply a temperature or osmolarity sensor (Fang & Rimsky, 2008).

While numerous examples of H-NS induced silencing exist, it is unclear if H-NS can act directly as a transcriptional activator. Accounts of H-NS functioning as an activator have been described, for example, in *E. coli* H-NS has been shown to positively affect flagellum biosynthesis (Bertin *et al*, 1994, Soutourina *et al*, 1999) and H-NS has been shown to act as a positive regulator of pectate-lyase synthesis in the plant pathogen *Erwinia chrysanthemi* (Nasser & Reverchon, 2002). However in both instances, H-NS is functioning indirectly as an activator by directly repressing a repressor, thus eliciting a net positive effect.

Interaction of H-NS with DNA

The ability of H-NS to function at the multitude of promoters as it does, arises from the fact that H-NS does seem to bind to just one consensus binding sequence, but rather H-NS displays binding affinity for a particular DNA structure; curved DNA (Dame *et al.,* 2000, Lang *et al.,* 2007, Bouffatrigues *et al.,* 2007, Dorman, 2004, Spassky *et al*., 1984, Zhang *et al*, 1996, Zuber *et al*, 1994). Curved DNA is commonly found in promoter regions regulated by H-NS (Yamada, 1990, Bracco *et al.,* 1989, Jauregi *et al.,* 2003). The intrinsic curvature of promoters can be specified by a variety of DNA sequences (Pedersen *et al.,* 2000, Barbic *et al.,* 2003a), which allows DNA binding proteins that recognize such structures the ability to interact with a variety of promoters, but are often caused by AT rich regions of DNA. DNase I footprint analysis of H-NS binding has demonstrated H-NS interaction with AT rich DNA time and again (Dame *et al*, 2005, Dame *et al*, 2006, De la Cruz *et al*, 2007, De la Cruz *et al*, 2009, Dorman *et al*, 1999, Dorman & Deighan, 2003, Dorman, 2004, Dorman, 2009, Porter & Dorman, 1994, Rimsky *et al*, 2001, Rimsky, 2004, Shi *et al*, 2004, Stoebel *et al*, 2008, Williams *et al*, 1996, Turner & Dorman, 2007). New insights into the transcriptional regulation by H-NS have yielded a 10 base pair (bp) consensus binding site (Fig. 1) at the osmoregulated *E. coli proU* promoter, which regulates an osmoprotectant uptake locus (Lang *et al*, 2007). This site is found in two positions downstream of the *proU* promoter, termed the negative regulatory element (NRE). This binding motif has been proposed to be a nucleation site for H-NS

binding. Upon binding by H-NS, these high-affinity sites are thought to act synergistically with lower affinity sites located in the promoter which allows the formation of a specific nucleoprotein complex that results in the repression of transcription (Fang & Rimsky, 2008, Bouffartigues *et al.,* 2007, Lang *et al.,* 2007). While promoters such as the *E. coli bgl* promoter contains 10 potential such binding sites, others like the *E. coli nir* promoter contain less (Fang & Rimsky 2008), and still some promoters contain no such sequence yet bind H-NS *in vitro*. Whether the high-affinity binding site is present or not, it is clear that upon H-NS binding to DNA, it spreads, or oligomerizes, along the DNA, coating it.

Structure of H-NS

To understand how H-NS oligomerizes on DNA one must first understand its structure. While the complete crystal structure of H-NS has not been resolved, proteomic analysis (primarily Nuclear Magnetic Resonance [NMR]) has revealed some interesting findings regarding H-NS. The H-NS protein is 137 amino acids in length and is comprised of three structural components: an N-terminal oligomerization domain; a C-terminal DNA binding domain; and a flexible linker that connects the two (Rimsky 2004, Dorman 2003, Dorman 2004). A 46-residue segment of the N-terminal domain was found to be required for dimerization (Bloch *et al*, 2003). NMR analysis of the *E. coli* H-NS protein showed that the resulting structure was an intertwined anti-parallel structure that resembled a "handshake" (Bloch *et al*, 2003). Interestingly, the same structure from the corresponding region of the protein from *Salmonella typhimurium* showed a parallel arrangement (Esposito *et al.,* 2002). It should be noted that both proteins have identical amino-acid sequences, and remains unclear as why both proteins adopt different

structural configurations (Esposito *et al.,* 2002, Rimsky, 2004, Dorman, 2000). The N-terminal structure of VicH, the H-NS-like protein from *Vibrio cholerae*, shows similar structural features to that derived from *E. coli*, which highlights the conserved function of this region (Cerdan *et al*, 2003).

Unlike differences in the N-terminal domain, the C-terminal nucleic acid binding domain remains highly conserved among members of the H-NS family (Tendeng & Bertin, 2003). Mutations in the C-terminal, DNA binding, domain do not impair the proteins ability to dimerize, but instead impairs the ability to recognize curved DNA (Bloch *et al.,* 2003). This indicates a role for the C-terminal domain in DNA binding but shows that DNA binding activity is dependent upon the N-terminal domain, which allows dimerization, initiating the DNA binding event (Dorman, 2004, Ueguchi *et al*, 1996). Experiments using modified H-NS proteins lacking oligomerization properties lack the ability to bind DNA, and therefore lack the ability to induce repression. This supports the findings that oligomerization of the protein is required for full activity of H-NS (Rimsky *et al.,* 2001, Badaut *et al.,* 2002). In solution H-NS forms a dimer, although higher oligomeric states have been observed dependening upon the concentration of H-NS (Smyth *et al*, 2000, Rimsky, 2004). The higher-order oligomerization, which is an essential function of H-NS, is a function of the flexible linker region of the protein (Bloch *et al.,* 2003). Disruption of the flexible linker region impairs the ability of H-NS to form higher-order oligomers, implicating this region in higher-order structure formation (Esposito *et al.,* 2002).

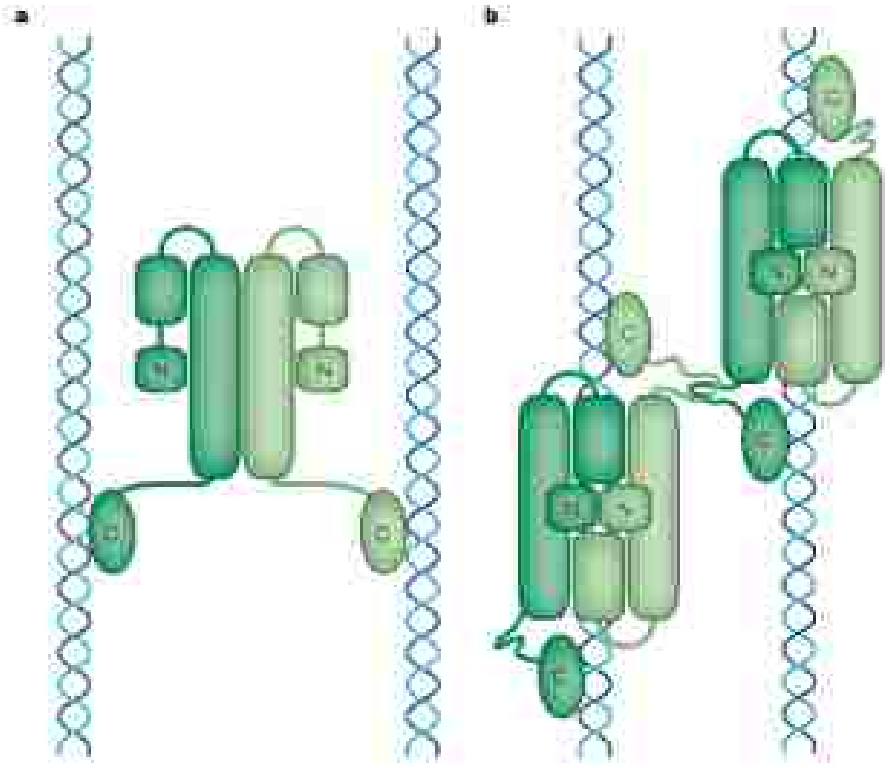

Fig. 2. From Dorman, 2004, showing two views of H-NS-mediated DNA bridging. a) H-NS dimer interacts simultaneously with two DNA duplexes or different parts of the same duplex through its C-terminal domains, based on the model by Esposito *et al.* (2002) for the *Salmonella* protein showing the N terminus lies parallel to the C terminus. b) Two H-NS dimers, each bound to separate DNA duplexes (or separate parts of the same duplex) interact with each other via their extended flexible linker regions. It is proposed that extension of the linker region is brought about by simultaneous involvement of the N-terminal and C-terminal domains in DNA binding. This model is based on analysis of the *E. coli* protein (Rimsky, 2004). The promoters within the dimers are shown in dark green and light green. N = amino-terminus, C = carboxy-terminus.

H-NS silences transcription by altering DNA topology

Based on the structural data available two proposals have been made regarding the mechanism of H-NS function. The first proposal, based on the *Salmonella typhimurium* H-NS protein, suggests a role for head-to-tail dimers, where the unstructured flexible linker region allows the cross-linking of separate DNA duplexes or different parts of the same duplex (Esposito *et al.,* 2002, Dorman 2004) (Fig. 2a). The second proposal, based on the *E. coli* H-NS protein, suggests that the N-terminal and C-terminal interaction with

DNA imposes a more stable structure on the flexible linker, which allows it to interact with the flexible linker region of other bound H-NS molecules which results in a higher-order, cross-linked, nucleoprotein complex being formed (Esposito *et al.,* 2002, Dorman 2004)(Fig. 2b). Atomic force microscopic analysis supports this second proposal that H-NS can cross-link DNA in this manner (Dame *et al.,* 2000).

Either mechanism could contribute to the H-NS condensation of DNA. H-NS has also been shown to bind at a region of curvature to which it nucleates, followed by oligomerization along the DNA (Amit *et al*, 2003). In this model of transcriptional silencing the curved sequence to which H-NS binds is located upstream or downstream of the promoter, which acts as a nucleating site (Rimsky *et al.,* 2001). The initial protein:DNA interaction at a bent region induces an intramolecular bridge. The formation of the intramolecular bridge further stimulates H-NS mediated bridging of the flanking sites resulting in oligomerization (Dame *et al.,* 2000). H-NS oligomerization along the DNA in either direction from nucleating sites is not in conflict with the previously described modes of action for H-NS, as additional higher-order nucleoprotein complexes could form between multiple regions of DNA bound by H-NS separated by un-bound, intervening DNA (Kim & Wang, 1999). Interaction of two sites of H-NS bound DNA has been shown to trap RNA

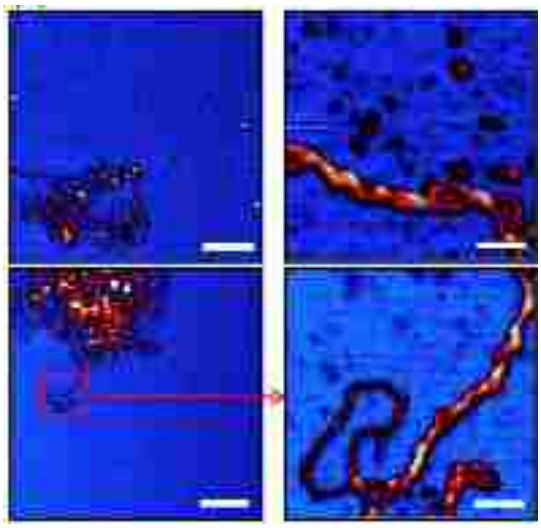

Fig. 3. From Mauer *et al.,* 2009. H-NS induced plectonemes. H-NS bound to lambda DNA at a ratio of 1 H-NS molecule per 1539 bp. The two DNA duplexes in the H-NS "filament" adopt a tightly interwound plectonemic form.

polymerase at the promoter, preventing transcription (Dame *et al.,* 2002). Therefore, the regulation of promoter activity is much more complicated involving the arrangement and multiplicity of nucleation and propagation sites (Rimsky, 2004), followed by bridging of two DNA helices and subsequently, condensation (Dame *et al.,* 2000, Maurer *et al*, 2009).

The alignment of the H-NS dimers effectively "zippers" the DNA, constraining negative supercoils, thus changing DNA topology, which results in transcriptional silencing (Rimsky *et al.*, 2001, Tupper *et al*, 1994). Nucleoid-associated proteins binding to DNA can alter the topology of the DNA into two basic types of coils: either plectonemic (interwound) or toroidal (solenoidal) (Maurer *et al.,* 2009). Binding by H-NS results in plectonemic (interwound) structures that are capable of covering at least 1.65 kb (Maurer *et al.,* 2009) (Fig. 3). Atomic force microscopic analysis of H-NS bound DNA, showed that H-NS-DNA complexes were organized into consistent structures and that the periodicity of H-NS binding induced structures was approximately every 60 bp of DNA (winding two DNA duplexes together, onto themselves) (Maurer *et al.,* 2009). The heat-unstable protein (HU), another NAP, that exists as an octamer, and which has been shown to antagonize H-NS, has been measured to bind to approximately 64-68 bp (Broyles & Pettijohn, 1986). This suggests that one conformation (e.g. plectonemic) could be easily converted into another (e.g. toroidal) simply by reorganizing the nucleoprotein complex by the interaction of a competing protein. This explains how a competing DNA binding protein might function to relieve H-NS induced repression, at a mechanistic level. These tight plectonemic structures stabilized by H-NS are consistent with gene silencing and repression (Maurer *et al.,* 2009), therefore, the promoters of the

regulated genes should then be located at the nodes where they would remain accessible to transcriptional machinery, however, direct evidence for this is lacking (Rimsky, 2004). In these topologically closed domains, the plectonemic conformations of the DNA can be easily overcome by other NAPs or RNA polymerase (RNAP), as the force with which H-NS constrains DNA is relatively weak, ~ 7pN while the force generated by RNAP is up to 25 pN (Dame *et al.,* 2006, Maurer *et al.,* 2009). The ease with which these structures can be overcome ultimately results in transcription.

A plectonemic structure constrained by H-NS, and a simple conversion to a toroidal or open conformation, by another DNA binding protein, or a change in DNA superhelicity subsequent to changes in osmolarity or temperature (Adler *et al*, 1989), satisfies the requirements that the DNA i) remain accessible to transcriptional machinery and ii) maintains the bacterial nucleoid by compacting the DNA (Dame *et al.,* 2006, Maurer *et al.,* 2009). We are beginning to build a more complete mechanistic picture of the way in which silencing is achieved.

H-NS interaction with other nucleoid-associated proteins

The oligomerization domain of H-NS is not only important for H-NS:H-NS interactions, but also facilitates binding with heterologous proteins that share the same domain (Dorman, 2004). Studies have shown that H-NS can oligomerize with up to five different protein species in the H-NS family (Rimsky, 2004, Dorman, 2004). Oligomerization by H-NS to other related proteins raise the possibility that the activity of these heteromers is altered resulting in additional levels of complexity in gene regulation ((Dorman & Deighan, 2003), Rimsky, 2004). The most common of these other heterologous proteins is the paralogue StpA (Rimsky, 2004, Dorman, 2004). StpA has

been described as a molecular back up for H-NS, however, it does not always provide this function (Dorman & Deighan, 2003). StpA shares many properties with H-NS including, 52% homology, the DNA binding domain, oligomerization domain and a flexible linker connecting the two (Zhang *et al.,* 1996, Ueguchi *et al.,* 1996, Williams *et al.*, 1996, Spurio *et al*, 1997, Dorman *et al*., 2000, Smyth e*t al*., 2000). Each protein can inhibit its own expression as well as the expression of the other protein (cross and autoregulation) (Zhang *et al.,* 1996). It has been proposed that StpA acts as a molecular adapter facilitating targeting of H-NS to DNA by multiple protein:protein interactions (Free *et al*, 2001). Some bacteria express more than one member of the H-NS family, an example is S*higella flexneri* that expresses 3 members: H-NS, StpA, and Sfh, and all three proteins can form oligomeric complexes with each other (Deighan *et. al.,* 2003). Of the H-NS-like proteins studied thus far, all have the ability to form protein:DNA bridges (Dame *et al.,* 2005). The resulting complexes in all theses previously described situations leads to transcriptional silencing.

H-NS has also been shown to interact with other, more distantly related proteins. H-NS interaction has been described with the: FliG flagellar motor protein, Hfq which binds RNA and modulates translation of mRNA, Hha a thermo-osmoregulator of the *E. coli* α-haemolysin toxin, and the *Yersinia enterocolitica* Hha homologue YmoA (Donato & Kawula, 1998, Dorman, 2004, Nogueira & Springer, 2000, Ellison & Miller, 2006a, Ellison & Miller, 2006b, Stoebel *et al.,* 2008).

Together, these findings suggest that protein-protein interactions of this nature are widespread in gram-negative bacteria. Interestingly, the genes that encode H-NS-like proteins and the Hha-like proteins are found on plasmids acquired by horizontal gene

transfer (Dorman, 2004). One hypothesis for why a plasmid would carry its own H-NS-like protein is that it appears to protect the recipient cell from the deleterious effects of H-NS titration by AT-rich sequences and allows the cell to tolerate the plasmid (Doyle *et al*, 2007).

H-NS silences genes acquired by horizontal gene transfer

Horizontal gene transfer has been described as one of the primary sources of genetic diversity in microorganisms (Jain *et al*, 2002) and has been most extensively studied in the human bacterial pathogens where many virulence factors have been acquired in this manner (Hacker & Kaper, 2000). Acquiring genes through horizontal transfer raises some very important regulatory issues, e.g. will the new gene be incorporated immediately into existing regulatory circuits? Will the gene operate independently? Foreign sequences are more likely to decrease the fitness of the host rather than increasing it (Navarre *et al*, 2006). Therefore, incoming, newly acquired sequences must be brought under the control of existing cellular regulators, and it is the proteins that show the highest degree of promiscuity for promoters that are best adapted to this function, e.g. H-NS (Dorman, 2004, Navarre *et al.,* 2006). The silencing of newly acquired genes by H-NS is thought to protect the bacterium from potentially deleterious effects of transcribing genes that may compromise their fitness (Navarre *et al.,* 2006). Many examples exist in the literature of H-NS mediated silencing of genes acquired by horizontal gene transfer, for example the large, 230 kb, virulence plasmid of *Shigella* and enteroinvasive *E. coli* which carries many of the genes required for virulence, like a type III secretion system (t3ss) and genes required for invasion and adhesion (Adler e*t al.*, 1989, Beloin & Dorman, 2003, Beloin *et al*, 2003, Hromockyj & Maurelli, 1989,

Hromockyj *et al*, 1992, Porter & Dorman, 1994). Similarly H-NS has been shown to be a repressor of the virulence genes associated with the *Salmonella* SPI-1 pathogenicity island, the virulence regulon in *Vibrio cholerae*, the invasion gene *invA* of *Yersinia*, and the locus of enterocyte effacement (LEE) pathogenicity island of enteropathogenic *E. coli* (Tupper *et al.*, 1994, Porter & Dorman, 1994, Heroven *et al*, 2004, Cerdan *et al.*, 2003, Nye & Taylor, 2003, Navarre *et al.,* 2006). It is interesting to note that foreign sequences inherited by horizontal gene transfer are rich in adenine and thymine (AT-rich) compared to that of their host. AT-rich sequences are more intrinsically curved (Zuber *et al*, 1994), providing an explanation of how these regulatory circuits have evolved using H-NS, and H-NS-like proteins as silencers of transcription (Lucchini *et al*, 2006, Oshima *et al*, 2006, Navarre *et al*., 2006). Although this explains how foreign genes are selectively targeted for silencing, the question still remains as to how H-NS silenced genes are to be integrated into existing regulatory networks and expressed under appropriate conditions.

1.2 Transcriptional anti-silencing

Since transcriptional silencing, is a process that renders appreciable regions of DNA inaccessible to proteins such as RNA polymerase (RNAP), DNA modifying enzymes, etc., to which it would otherwise be able (Dorman et al., 1999, Yarmolinsky, 2000), the mechanism that counters this makes previously inaccessible regions of DNA accessible and has been coined anti-silencing. Anti-silencing differs from true activation in that RNAP is not actively recruited to the DNA, with few exceptions e.g. ToxT of *V. cholerae* (Nye *et al*, 2000). Anti-silencing can occur through a protein independent, or a protein dependent mechanism and often a combination of the two is required for optimal

performance. Therefore, in a protein dependent anti-silencing mechanism, the anti-silencing protein can be any DNA binding protein that antagonizes, interferes with, or alters, the action of a silencing protein, e.g. H-NS. Protein dependent anti-silencing mechanisms generally occur in 3 ways (however, subtle variations in these themes exist, and a anti-silencing proteins may function through more than one mechanism): through direct competition with the silencing protein (this can include blocking the silencing proteins polymerization), through altering the DNA in a way that transcription is allowed to occur, and by de-stabilizing the repressive, silenced, DNA structure.

Protein independent anti-silencing

The simplest form of anti-silencing is a protein-independent mechanism which relies on the static curvature of the DNA. The dimerization and bridging characteristic of H-NS makes the protein uniquely sensitive to DNA structure. The static curvature of DNA can facilitate H-NS cross-linking, but the curvature is also sensitive to environmental conditions (De la Cruz *et al*, 2009, Falconi *et al*, 1998, Prosseda *et al*, 2004, Stoebel *et al*, 2008). Any condition that could alter the DNA structure would then antagonize the H-NS induced bridging. Conditions that are known to alter the curvature of DNA are temperature and osmolarity. An increase in temperature can reduce the degree of curvature thereby reducing the ability of H-NS to constrain the bridged structure (Rohde *et al*, 1999, Ussery *et al*, 1999), as observed in the case of the promoter of the *virF* virulence gene in *Shigella flexneri* (Prosseda *et al.,* 2004). Likewise a change in osmolarity can also affect DNA curvature. The *E. coli proU* promoter, which is negatively regulated by H-NS, is sensitive to changes in DNA superhelicity following osmotic shock, which effects H-NS binding to its regulatory regions (Bouffartigues *et al.,*

2007). Environmentally induced changes in DNA curvature might represent the most fundamental mechanism of transcriptional anti-silencing, but more complex mechanisms of anti-silencing exist in the literature.

Many horizontally acquired elements encode genes for protein antagonists of H-NS, including Ler, Lrp, Fis, IHF, HU etc., and many of them share degenerate target specificities, as H-NS does, making them ideal regulators involved in either, competition for binding sites in DNA, or remodeling the architecture of the DNA that has a positive effect on transcription (Dorman, 2004, Fang & Rimsky, 2008, Navarre *et al*, 2007, Stoebel *et al*., 2008).

Protein dependent anti-silencing: direct competition

The protein dependent mechanism underlying transcriptional silencing/anti-silencing has been proposed to be a straightforward molecular antagonism involving direct competition between H-NS and other sequence specific DNA binding proteins for their overlapping cognate DNA binding sequences (Fang & Rimsky, 2008, Navarre *et al*., 2007). Examples of anti-silencing being a straightforward antagonism exist at a variety of bacterial promoters, e.g. SlyA counters H-NS silencing of the hemolysin gene *hlyE* in *E. coli* through direct competition, similarly RovA from *Yersinia* counters H-NS silencing of the invasin *inv*, and the urease gene activator UreR of *Proteus mirabilis* displaces H-NS by directly competing for essential binding sites (Coker *et al*, 2000, Ellison & Miller, 2006a, Ellison & Miller, 2006b, Westermark *et al*, 2000, Heroven *et al*, 2004, Heroven *et al*, 2007, Lithgow *et al*, 2007, Poore & Mobley, 2003). The ToxT transcription factor from *Vibrio cholerae* antagonizes H-NS at two promoters by direct competition (Nye *et al.,* 2000). In *Salmonella,* many of the genes regulated by SlyA are

likely to have been acquired by horizontal gene transfer, and all but one of these is repressed by H-NS (Navarre *et al.,* 2005, Navarre *et al.,* 2006). A mutual antagonism exists between SlyA and H-NS where one protein is able to displace the other depending on relative concentrations of the two proteins. This could represent one mechanism for the re-establishment of H-NS mediated repression by allowing both proteins to utilize similar mechanisms, allows fine-tuning of transcription in the cell (Lithgow *et al.,* 2007). This mechanism may be more common among silencing and anti-silencing proteins and is unlikely restricted to SlyA and H-NS. Still another mechanism of counter-silencing by direct competition has been observed where two proteins of *E. coli,* PapB and CRP, appear to work in tandem to antagonize H-NS mediated silencing of the *pap* promoters (Forsman *et al*, 1992).

Protein dependent anti-silencing: alteration of topology

In contrast, at some bacterial promoters the anti-silencing proteins function to alter the DNA topology in a way that no longer allows H-NS to constrain transcription. For example, LeuO counters silencing of the quiescent porin *ompS1* in *Salmonella* by altering the topology of the *ompS1* promoter region (De la Cruz *et al*, 2007, De la Cruz *et al*, 2009) and in *Shigella flexneri,* VirB antagonizes H-NS by affecting the structure of the *icsB* promoter (Turner & Dorman, 2007). The mechanism by which this is achieved involves protein binding to specific binding sites followed by propagation of the protein along the DNA with associated wrapping of the DNA. H-NS is no longer able to maintain the repression complex, which allows RNAP to access the DNA and initiate transcription (Beloin & Dorman, 2003, De la Cruz *et al.,* 2009). Interestingly, the SlyA protein of *Salmonella*, which directly competes with H-NS at the *hlyE* promoter,

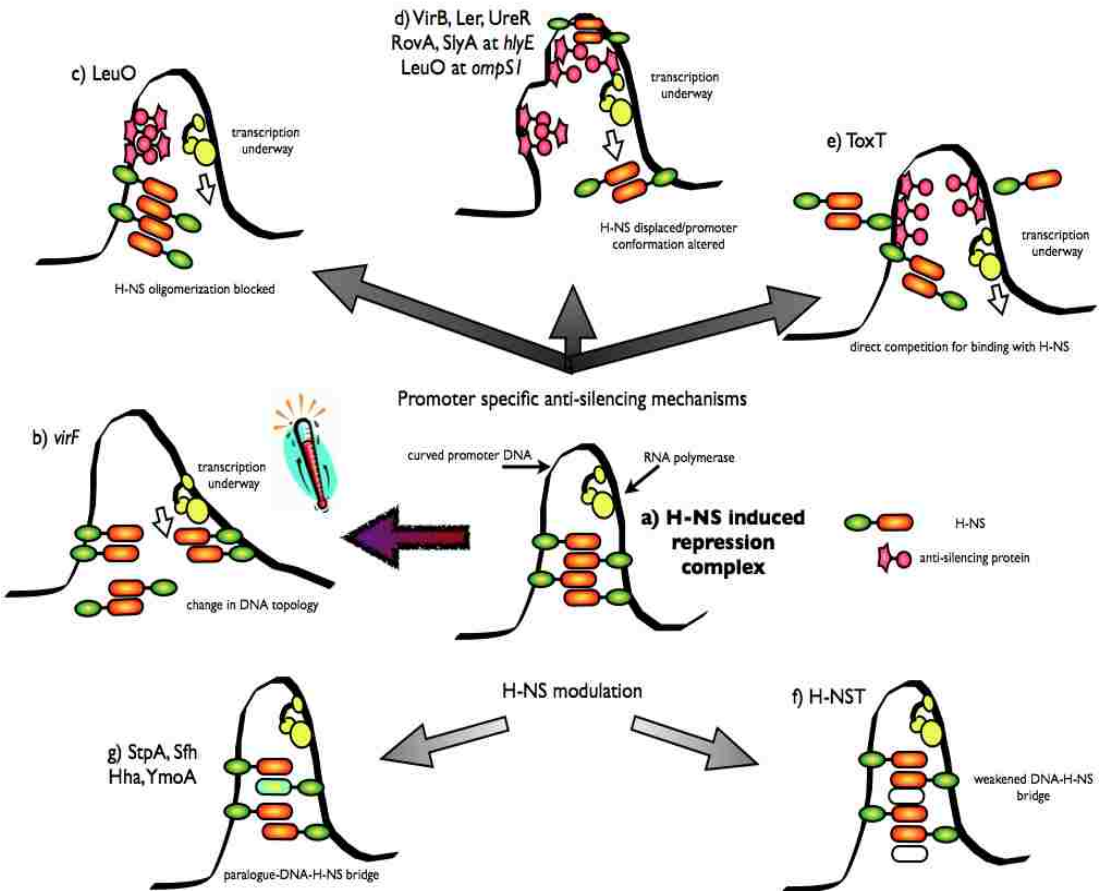

Fig. 4. Adapted from Stoebel *et al.*, 2008. Illustration of the ways in which a repressive H-NS promoter complex (a) can be modified to permit transcription. At certain promoters (b), specific de-repression occurs via a protein-independent mechanism involving environmentally induced changes in DNA bending or supercoiling (colored arrow). Many promoters are de-repressed via promoter-specific DNA-binding proteins (black arrows), which can work by preventing H-NS oligomerization along the DNA (c), displacing H-NS, or modifying the conformation of the promoter, or both while some H-NS remains bound (d), or displacing H-NS, by direct competition for H-NS binding regions (e). Other de-repression mechanisms are general for many H-NS repressed promoters (grey arrows). These include weakening the H-NS-promoter complex (H-NST)(f), or expression of an H-NS paralogue (discussed in "silencing;" see above) that may enhance repression or alter its sensitivity to other de-repression mechanisms dependent upon the exact properties of the protein-DNA complex (g).

functions to remodel the architecture of the DNA at the *ugtL* and *pagC* promoters (Kong *et al*, 2008, Perez *et al*, 2008, Shi *et al*, 2004). LeuO has also been shown to block H-NS propagation by binding between the H-NS upstream and downstream binding sites, functioning as a boundary element (Chen & Wu, 2005, Chen *et al*, 2005). These results show that a single H-NS antagonist can function via two distinct mechanisms: by direct

competition, and remodeling the nucleoprotein complex (Stoebel *et al.,* 2008). The common feature that is shared by the H-NS antagonizing proteins is that most, if not all, bind to AT-rich sequences, which may explain why these proteins have evolved to function as anti-silencers (Navarre *et al.,* 2006).

Protein dependent anti-silencing: destabilization

A third mechanism of counter-silencing, that is not often described, is the use of truncated H-NS molecules (H-NST). The genes encoding these molecules have been found in pathogenicity islands of pathogenic *E. coli* strains (Williamson & Free, 2005). H-NST molecules show homology to the H-NS oligomerization domain, but lack nucleic acid binding and linker regions (Willamson & Free, 2005). As opposed to the sequence-specific proteins directly competing with H-NS for binding to DNA, or remodeling the nucleoprotein complex, H-NST molecules interfere with H-NS oligomerization and is thought to weaken the ability of H-NS to form bridges by replacing full length H-NS. However, the biologic significance of H-NST molecules remains unclear and examples have not been demonstrated outside of the pathogenic *E. coli* strains (Navarre *et al.,* 2006, Stoebel *et al.,* 2008, Williamson & Free, 2005).

1.3 *Shigella*, the bacterial pathogen

Shigella species are gram-negative intracellular pathogens that cause bacillary dysentery in humans. *Shigella* infections are responsible for approximately 164 million infections each year resulting in 1.1 million deaths, most of these occur in developing countries and in children less than 5 years old (Kotloff *et al*, 1999, Lee *et al*, 2005, Li *et al*, 2009). The genus is divided into 4 species, *Shigella flexneri, Shigella boydii, Shigella*

sonnei, and *Shigella dysenteriae.* The species are further divided into serotypes based on biochemical differences and variations in their O-antigens, such that, *Shigella flexneri* is further divided into 13 serotypes. *Shigella flexneri* is endemic in most developing countries and causes more mortality than any of the other *Shigella* species (Bennish & Wojtyniak, 1991). Of the 13 serotypes of *S. flexneri* serotypes 1b, 2a, 3a, 4a, and 6 are predominant in developing countries, while 2a is predominant in industrialized countries including the United States (Kotloff *et al*., 1999).

Shigella **pathogenicity**

The mechanism of pathogenicity of *Shigella flexneri* is based on the capacity of the organism, once ingested, to reach the colonic mucosa and to invade the colonic epithelial cells. Invasion of the epithelial cells leads to intracellular bacterial multiplication, spread to adjacent cells, cell death, and eventually inflammation and ulceration of the colonic mucosa (Hale, 1991). This pathogenicity is a complex phenomenon, which relies on the coordinated expression of a suite of genes, many of which are found on the large (230 kbp) virulence plasmid. The structural genes required for invasion, including a type III secretion system (t3ss), and the intracellular spread of *Shigella* in host cells are encoded by separate regions of the virulence plasmid: the *ipa-mxi-spa* locus and the *icsA gene* (Dorman & Porter, 1998). The large virulence plasmid carries genes encoding key virulence gene regulators. Further regulators of virulence gene expression in *Shigella* are located on the chromosome: the plasmid-located positive transcriptional regulators *virF* and *virB* and the chromosomally encoded negative regulator *hns* (Adler *et al.,* 1989, Dagberg & Uhlin, 1992).

The regulatory cascade controlling *Shigella* virulence

The VirF and VirB proteins act in a cascade to activate virulence gene promoters in response to environmental stimuli that indicate the bacterium is in the appropriate location in the host. Temperature is a key regulatory factor of virulence gene transcription, since virulence gene transcription is repressed at temperatures below 37 °C. H-NS has been shown to direct the temperature-regulated expression of virulence genes by repressing their transcription during growth at 30 °C (Dorman *et al*., 1990, Hromockyj *et al*., 1992). At the non-permissive temperature of 30 °C, H-NS binds to two sites within the *virF* promoter (Fig.5). The two regions of bound H-NS interact in the formation of a nucleoprotein complex that results in repression of *virF* transcription. An increase in temperature to 37°C, antagonizes the DNA topology resulting in the loss of contact between the H-NS regions and this is thought to de-stabilize the nucleo-repressive complex resulting in de-repression and subsequent transcription of the *virF* gene (Falconi *et al*., 1998, Prosseda *et al*., 2004).

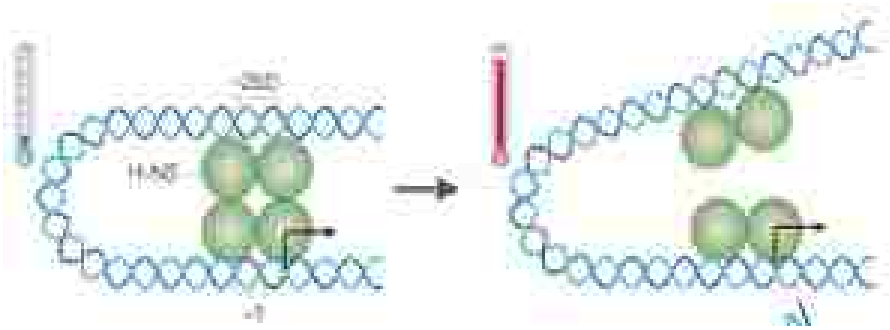

Fig. 5. From Dorman, 2004. H-NS mediated repression of the *virF* virulence-gene promoter in *Shigella flexneri.* At the S. *flexneri virF* promoter, there are two binding sites for the H-NS protein centered at -1 and -250 with respect to the transcriptional start site, which is shown with an angled arrow. Formation of the repression complex at low temperatures requires looping of the intervening DNA. The formation of a DNA topology that is conducive to repression is antagonized by increased temperature. The resulting loss of contact between the bound H-NS proteins is thought to result in the instability of the nucleoprotein complex at the *virF* promoter, leading to derepression of transcription (Falconi *et al.*, 1998)

The major transcriptional activator VirF

Once expressed, VirF, an AraC-like transcription activator, binds directly to the *virB* promoter. Tobe *et al*. (1995) have shown that VirF is present *in vivo* at both permissive and non-permissive temperatures (although transcription of *virF* at 30 °C is one quarter of that at 37 °C at which maximal transcription occurs) and that it can that it can bind *in vitro* to the *virB* promoter at both temperatures. Therefore, the thermal activation signal is not just production of VirF. These findings suggest that an increase in temperature results in a change in DNA supercoiling that is responsible for the VirF dependent activation of the *virB* promoter. Porter & Dorman (1997) also showed that *virB* DNA topology was affected by changes in osmolarity and pH resulting in increased negative supercoiling. Therefore, the alteration in DNA topology facilitates a productive interaction between bound VirF and RNA polymerase and also raises the possibility that a change in supercoiling may promote VirF oligomerization on the DNA, which results in the antagonism of H-NS leading to transcription of *virB* (Fig. 5)(Dorman & Porter, 1998).

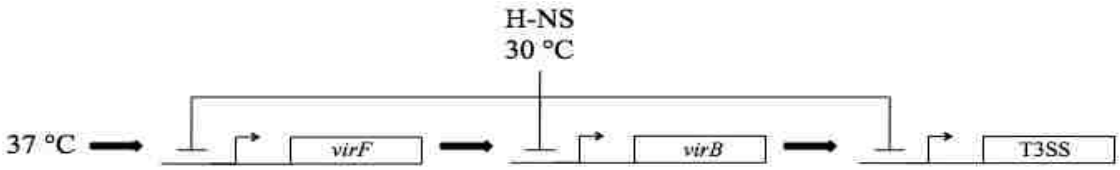

Fig. 6. The regulatory cascade controlling *Shigella* virulence. At the non-permissive temp of 30 °C H-NS causes repression of the major transcription factors *virF* and *virB*. Upon a shift in temperature to the permissive temp of 37 °C a change in DNA topology relieves the H-NS induced repression of *virF*. VirF antagonizes the H-NS induced repression of the *virB* promoter and subsequently VirB alleviates the H-NS dependent repression of many virulence plasmid genes including those of the invasion locus; the type III secretion system shown here is representative.

Subsequently, VirB alleviates H-NS dependent repression of many virulence plasmid genes, including those of the invasion locus *ipa*, *mxi*, and *spa; virA, ospB* and *phoN2* (Beloin & Dorman, 2003, Berlutti *et al*, 1998, Maurelli *et al*, 1984a, Maurelli *et al*, 1984b, Maurelli & Sansonetti, 1988, Nicoletti *et al*, 2008, Prosseda *et al*, 1998, Sasakawa *et al*, 1993, Tobe *et al*, 1991, Tobe *et al*, 1993, Tobe *et al*, 1995, Dorman, 2006, Hromockyj *et al*., 1992, Porter & Dorman, 1994, Porter *et al*., 1998, Prosseda *et al*., 1998).

The major transcriptional activator VirB: an unusual transcriptional regulator

The VirB protein was first identified through transposon mutagenesis of the large virulence plasmid, when it was shown to be essential for the expression of almost all the structural virulence genes (Watanabe *et al*, 1990). VirB is an unusual candidate for a transcriptional activator sharing no homology to previously described conventional transcriptional activators. Instead, VirB shares considerable homology with a family of plasmid partition proteins, including ParB and SopB, proteins that are involved in the plasmid partition and maintenance of stable plasmid copy number in P1/P7 and F plasmids, respectively (Fig.

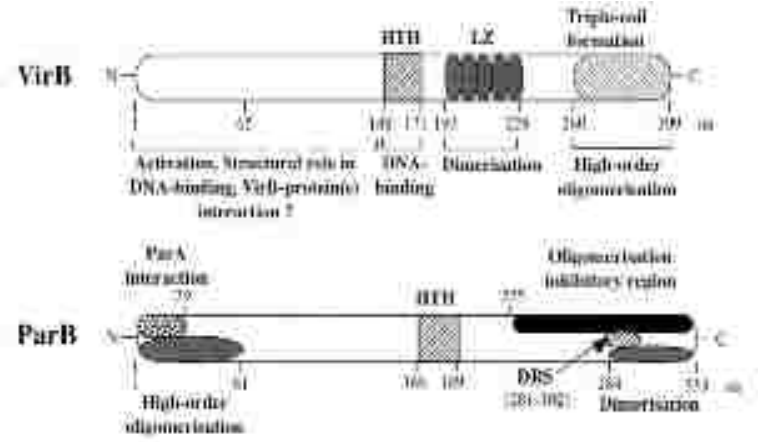

Fig. 7. From Beloin *et al.*,2002. Model of VirBs functional and structural domains. A structural model of VirB deduced from experiments performed in Beloin *et al.*, 2002. Below is the structural model of the ParB protein based on the work of Surtees & Funnell, 1999,2001. In ParB, *DRS*, represents "discriminator recognition sequence," implicated in the DNA binding activity of ParB.

7)(Watanabe *et al.*, 1990, Abeles *et al*, 1985, Bignell & Thomas, 2001). VirB is small, 32 kDa, and basic and has been shown to form dimers, trimers, and higher order oligomers i*n vitro* and *in vivo* (Beloin *et al*, 2002). VirB can oligomerize independent of DNA binding and occurs through two domains, a leucine zipper, and a C-terminal domain predicted to form a triple coil structure (Beloin *et al.,* 2002).

VirB shows homology to the partition proteins SopB/ParB VirB interacts with DNA *in vivo* and is dependent on the presence of the N-terminus and the integrity of the helix-turn-helix (HTH) binding motif. The HTH was found to be essential for structural gene activation (Beloin *et al.,* 2002). The HTH DNA-binding motif in VirB is 80% identical to the HTH in the homologous ParB protein, which mediates the binding of ParB to *parS* which leads to plasmid partitioning (Beloin *et al.,* 2002, Lobocka & Yarmolinsky, 1996, Surtees & Funnell, 2001).

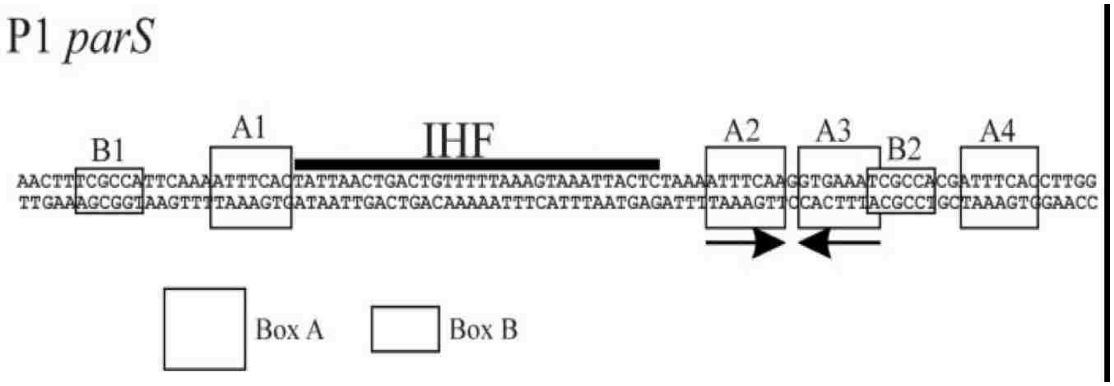

Fig. 8. From Surtees & Funnell, 2001. The *parS* site. *tall* and *short* boxes outline the box A and box B sequences respectively, and the black line shows the IHF binding site.

Beloin *et al.* (2002) showed that the leucine zipper (LZ) motif was essential for VirB to form oligomers, and that oligomerization was essential for the VirB dependent activation of a *mxiC-lacZ* fusion. Complete loss of the LZ resulted in complete loss of the ability of VirB to activate gene expression and reduces its ability to bind DNA

(Beloin *et al.,* 2002). However, the LZ cannot promote protein:protein interactions in isolation, instead it requires the C-terminal domain (Beloin *et al.,* 2002). Therefore, the LZ most likely functions to promote oligomerization, which is then stabilized by the C-terminal coiled-coil structure, making both regions mutually dependent on one another (Boss *et al*, 1999, Beloin *et al.,* 2002). Also, the HTH motif was found to be essential in DNA binding but did not contribute significantly to VirB oligomerization (Beloin *et al.,* 2002, (McKenna *et al*, 2003). The work of Beloin *et al.*(2002), and later McKenna *et al.* (2003), demonstrated that VirB is a direct DNA binding protein that interacted specifically with the promoters of the virulence genes its known to regulate. They also demonstrated the necessity of VirB oligomerization in the activation of the structural genes, a situation that resembles the homologous ParB and SopB functions (Beloin *et al.,* 2002, McKenna *et al.,* 2003). ParB oligomerization, which ultimately leads to protein:DNA bridges via protein:DNA and protein:protein interactions, is an essential step in the formation of the partition complex to effectively segregate plasmids (Bouet *et al*, 2000, Schumacher & Funnell, 2005, Schumacher *et al*, 2007). Even though VirB shows homology to ParB and SopB there is no data to support that VirB is involved in plasmid partitioning or maintenance of the

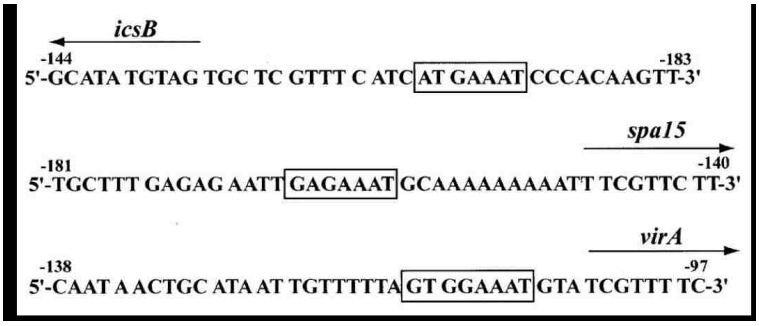

Fig. 9. From Taniya *et al.*, 2003. Comparison of the putative InvE binding regions in *S. sonnei*. Beloin *et al.,* (2002) revealed that InvE (VirB) binds to the regions upstream of these genes (*icsB*, *spa-15*, and *virA*) specifically. The numbers indicate the positions of these regions (translation start site of each gene as nucleotide +1). The boxes indicate the position of the ParB BoxA-like sequence. The solid arrows indicate the direction of each gene transcription.

modern virulence plasmid (McKenna *et al.,* 2003, Harrison *et al.,* unpub. data). While there is structural homology between VirB and ParB/SopB, there are functional differences. Firstly, ParB represses transcription (by spreading along the DNA for great distances eg. 2 kb)(Lobocka & Yarmolinsky, 1996, Rodionov *et al.,* 1999), while VirB activates transcription by alleviating transcriptional repression (Beloin *et al.,* 2002). Secondly, ParB binds to a well defined *parS* site of plasmid P1 (Fig. 8) which consists of 4 box A sequences (heptameric) and 2 box B sequences (hexameric) imperfect repeats, separated by an integration host factor (IHF) binding site (Surtees & Funnell, 2001), in contrast VirB binds to a sequence that is homologous to the box A motif in the *parS* site, and this site in the *S. sonnei icsB* promoter is essential for the expression of the *ipa* operon. Based on a previous study in *S. flexneri* (Beloin *et al.,* 2002) which showed VirB (InvE) binding to various promoter sequences, Taniya *et al.* (2003) also found this 7 bp ParB box A-like sequence [5'-(A/G)(A/T)G(G)AAAT-3'] in 2 additional promoter sequences (Fig. 9). In fact this sequence was shown by gel shift assay to be required for InvE binding (Taniya, *et al.,* 2003). This work was further corroborated by Turner & Dorman (2007) which showed that VirB bound to a regulatory region within the divergently transcribed *S. flexneri icsB-ipgD* promoter that contained a near-perfect, inverted repeat, which is homologous to the *parS* box A sequence (labeled box 1 and box 2) (Fig. 10). However, in Turner & Dorman (2007), only the promoter distal binding site (box 2) was found to be critical for the VirB-mediated activation of the promoter. Interestingly, the *icsB* regulatory region shows DNA sequence homology to the *parS* site of plasmid P1 and P7, and furthermore, *icsB* shares a dependency, as does *parS*, on the IHF protein (IHF has been shown to bind and enhance activity of the *icsB* promoter; IHF

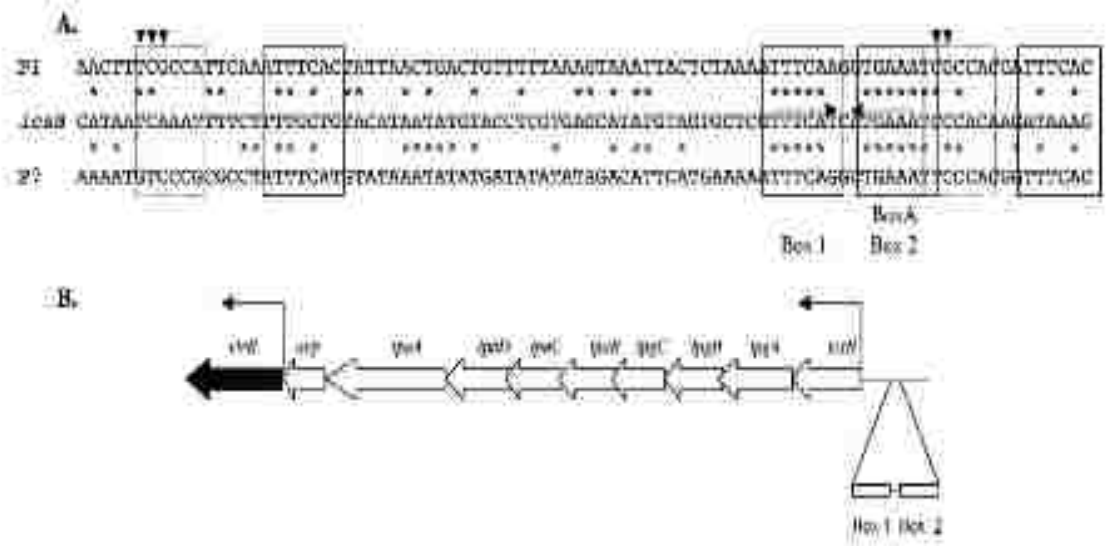

Fig. 10. From Turner & Dorman, 2007. Comparison of *icsB* and *parS*. A) alignment of the *icsB* regulatory region with the *parS* sequences of phages/plasmids P1 and P7. The DNA sequence of the promoter-distal portion of the *icsB* regulatory region that contains boxes 1 and 2 is aligned with the *parS* sequences of phages/plasmids P1 and P7. The converging horizontal arrows show the inverted repeats associated with boxes 1 and 2. The four heptameric and two hexameric *parS* motifs involved in ParB protein interaction are boxed by solid- and dotted-line rectangles, respectively. Downward-pointing arrowheads indicate residues within the hexamers that allow ParB proteins to distinguish different *parS* sequences. The asterisks indicate residues that are conserved between the *icsB* regulatory region and the *parS* sequences. B) Genetic map of the portion of the large virulence plasmid showing the relative locations of the *virB* gene and the regulatory sequences of the i*csB-ipg-ipa-acp* operon. The angled arrows represent promoters. The relative positions of the box 1 and box 2 motifs upstream of the *icsB* promoter are shown. The diagram is not drawn to scale.

is required for full activation of *virF* and *virB*) (Turner & Dorman, 2007, Porter & Dorman, 1997). Together, these findings suggest that VirB is likely to be specific for the promoters that contain this box 2 element, or at least close matches to it. The significance of this finding suggests that VirB is a former plasmid partitioning protein that has been conscripted to perform a new function, that of an anti-silencer, antagonizing the H-NS mediated repression of promoters (Turner & Dorman, 2007). Unlike the ParB/*parS* system, the virulence plasmid of *Shigella* does not encode for a partner protein for VirB that is equivalent to ParA. The ParA protein is an accessory ATPase that interacts directly with ParB bound at *parS* and is required for the proper localization of

the ParB-*parS* complexes, although it is not known how this positioning is accomplished (Bouet & Funnell, 1999). Although, it is possible that another protein interacts with VirB, by binding the N-terminal domain of VirB, as proposed in the work of Beloin *et al.* (2002) to date no such co-factor has been found. Future investigations may one day find an accessory protein for VirB.

VirB function

The mechanism by which VirB activates transcription is not completely understood. At the *icsB* promoter, VirB has been reported to bind and oligomerize on DNA to a short sequence [5'-(A/G)(A/T)G(G)AAAT-3'] located upstream of the promoter, and this binding is required for activity of this promoter (Beloin *et al*., 2002, McKenna *et al.*, 2003, Taniya *et al*., 2003). However, this sequences occurs 56 times on pWR100 of *Shigella flexneri*, including within numerous insertion sequences scattered along the plasmid, and many more (more than 1000 times) with one mismatch, suggesting that this sequence alone does not define the VirB target (Le Gall *et al*, 2005). An examination of all the genes under the control of VirB show a lower G+C content, compared to the host, therefore, bending might be an essential part of the VirB recognition motif (it should be noted that not all promoters with low G+C content are under the control of VirB) (Le Gall *et al.,* 2005). It has been shown that the promoters under the positive transcriptional control of VirB are also under the negative transcriptional control of H-NS, and that VirB has been shown to displace H-NS from these promoters (Turner & Dorman, 2007, Beloin *et al.*, 2003). This further strengthens the case for curvature of DNA playing an important role in the VirB dependent activation of the promoters for which it regulates.

These studies, and many more, all point to a finely tuned mechanism of regulation involving the H-NS mediated repression of the virulence system in *S. flexneri*, that would not be permitted to be expressed until the appropriate environmental signals initiate a regulatory cascade involving VirF and VirB, that would limit the wasteful expression of the type III secretion system, its effector proteins and additional factors until authentic contact with host. However, many more questions remain to be answered, e.g.; What are the full repertoire of genes under the control of VirB? How does VirB alleviate H-NS mediated repression of promoters? Does VirB have more than one mechanism, which can be employed to alleviate H-NS repression? Is there a better-defined VirB binding motif? An enhanced understanding of the silencing/anti-silencing mechanisms that control expression of virulence genes in *Shigella,* and many other bacteria, is central to understanding pathogenesis.

1.4 Objectives of this study

The aim of this work is to determine the role that H-NS and VirB play in the regulation of the *icsP* promoter and to examine the molecular mechanism of silencing and anti-silencing. This dissertation is divided into three main chapters encompassing five objectives.

Objective 1: To determine regions of the *icsP* promoter needed for H-NS induced repression and regions required for VirB dependent de-repression. This objective is addressed in chapters 2 and 3.

Objective 2: To demonstrate H-NS and VirB bind directly to the *icsP* promoter. This objective is addressed in chapters 3 and 4.

Objective 3: To determine the sequences to which H-NS and VirB bind. This objective is addressed in chapters 3 and 4.

Objective 4: To examine the effect that *icsP* static DNA curvature plays in the ability of H-NS and VirB to regulate promoter activity. This objective is addressed in addressed in chapter 4.

Objective 5: To propose a mechanism for the H-NS and VirB regulation of the *icsP* promoter. This objective is addressed in all chapters and model is presented in chapter 5.

CHAPTER 2

VIRB ALLEVIATES H-NS REPRESSION OF THE *ICSP* PROMOTER IN *SHIGELLA FLEXNERI* FROM SITES OVER 1 KB UPSTREAM OF THE TRANSCRIPTION START SITE

Abstract

The *icsP* promoter of *Shigella spp.* is repressed by H-NS and derepressed by the VirB. An inverted repeat located between -1144 and -1130 relative to the transcription start is required for VirB dependent de-repression of the *icsP* promoter. The atypical location of this *cis*-acting site relative to the promoter is discussed.

Authors: Maria I. Castellanos[1^], Dustin J. Harrison[1^], Jennifer M. Smith[1⊥], Stephanie K. Labahn[1], Karen M. Levy[1], and Helen J. Wing[1]*

[1]School of Life Sciences, University of Nevada Las Vegas, Las Vegas, NV 89154-4004, USA.

^ These authors contributed equally to this work

Present address: Thomas Jefferson University, Jefferson College of Graduate Studies, Jefferson Alumni Hall, 1020 Locust Street, Philadelphia, PA 19107

This study was supported by the NIH grant P20 RR-016464 from the INBRE Program of the National Center for Research Resources, by the NSF grant 0649267 and by UNLV start-up funds to H.J.W. J.M.S. was a recipient of a NSF REU scholarship (NSF 0649267) in 2007; D.J.H. is sponsored by the U.S. Navy; K.M.L. is a recipient of a Barry Goldwater scholarship.

2.1 Introduction

Shigella species are gram-negative intracellular pathogens that invade cells of the lower intestinal epithelia of humans and primates, causing bacillary dysentery. All *Shigella* species carry a large virulence plasmid, and many genes carried by these plasmids are thermo-regulated. At a non-permissive temperature of 30°C, the global regulator H-NS (histone-like nucleoid structuring protein) represses transcription of many virulence plasmid genes (Beloin & Dorman, 2003, Porter & Dorman, 1994, Hromockyj *et al*, 1992). At the permissive temperature of 37°C, H-NS dependent repression is relieved by temperature induced changes in DNA topology (Falconi *et al.,* 1998, Prosseda *et al.,* 1998), VirF, or its subordinate regulator VirB (InvE) (reviewed in reference Porter & Dorman, 1997). The mechanism that leads to the alleviation of transcriptional repression by H-NS has been coined "anti-silencing." Anti-silencing is thought top play an important role in controlling the expression of genes acquired through horizontal gene transfer and is common in bacterial pathogens in which a variety of transcription factors function to relieve repression by H-NS (reviewed in reference Stoebel *et al.,* 2008).

The *icsP* (*sopA*) gene is carried on the large virulence plasmid in all *Shigella* species (Egile *et al*, 1997, Shere *et al.,* 1997) and encodes an outer membrane protease, which belongs to the omptin protease family (Hritonenko & Stathopoulous, 2007, Kukkonen & Korhonen, 2004) and cleaves the actin-tail assembly protein IcsA from the surface of *Shigella* (Egile *et al*, 1999, Shere *et al.,* 1997, Steinhauer *et al.,* 1999). Previous studies have revealed that *icsP*, like other Shigella virulence plasmid genes, including *virA, ospB* and *phoN2* and those of the invasion locus *ipa*, *mxi*, and *spa* (Adler *et al*, 1989,

Berlutti *et al*, 1998, Beloin *et al.,* 2002, Sakai *et al.,* 1988, Santapaola *et al.,* 2006, Taniya *et al.,* 2003, Tobe *et al.,* 1991, Uchiya *et al.,* 1995) is repressed by H-NS and de-repressed by VirB (Wing *et al.,* 2004). In this study, we identify sequences upstream of the *icsP* gene necessary for de-repression by VirB, with a view to improve our understanding of the mechanism of transcriptional anti-silencing at the *icsP* promoter.

2.2 Materials and Methods

Bacterial strains, plasmids and media.

The bacterial strains and plasmids used in the present study are listed in Appendix 1. Bacteria were grown routinely at 37°C in Luria-Bertani (LB) broth with aeration or on LB agar (LB broth containing 1.5% [wt/vol] agar). Antibiotics, as needed, were added to achieve the following final concentrations: ampicillin, 100 μg/ml; chloramphenicol, 25 μg/ml; kanamycin, 50 μg/ml; and tetracycline, 12.5 μg/ml. Where appropriate, to ensure that *Shigella* strains had maintained the large virulence plasmid during manipulation, Congo Red binding was tested on Trypticase Soy Agar (TSA) plates containing 0.01% (wt/vol) Congo Red (Sigma Chemical Co., St. Louis, Mo.).

Construction of the P*icsP-lacZ* reporter plasmid and truncated promoter fragments.

The starting point for this work was the P*icsP-lacZ* reporter plasmid pHJW7 (described in Wing *et al*, 2004; Table 1). pHJW7 carries a *Pst*I *Xba*I DNA fragment that contains 1232 bp upstream of the transcription start site of the *icsP* promoter and the first 48 bp of the *icsP* coding region. This fragment is cloned upstream of a translation stop site and a promoterless *lacZ* gene, so that expression of *lacZ* is directly regulated by the *icsP* promoter. To facilitate the construction of a series of P*icsP-lacZ* fusion plasmids, a second *Xba*I restriction site located downstream of the *lacZ* gene in pHJW7 was removed

using the 'mega'-primer PCR method (Perrin & Gilliland, 1990) with oligonucleotides W1016-1018 (Appendix 1). The resulting plasmid, pHJW20 (Table 1), contains a unique *Xba*I site upstream of the *lacZ* gene. *icsP* promoter fragments were PCR amplified from pHJW20 using oligonucleotide W89 in combination with each of the following oligonucleotides: W44, 45, 46, 47, 528, 529 or 530 (Appendix 1). PCR products were digested with *Pst*I and *Xba*I, and were ligated into pHJW20 that had also been digested with *Pst*I and *Xba*I. This created a series of P*icsP-lacZ* fusion plasmids containing promoter fragments lengths of 1056 bp, 893 bp, 665 bp, 351 bp, 254 bp, 150 bp, and 92 bp, respectively. These plasmids were called pJS01, pJS02, pDH01, pJS04, pHJW34, pHJW35 and pHJW36, respectively (Appendix 1).

To create pMIC01, pHJW20 was digested with *Sal*I and *Pst*I, treated with T4 DNA polymerase and self-ligated using the blunt-ended ligation protocol supplied with T4 DNA ligase (Promega). To create pMIC02, a 200 bp fragment was amplified by PCR from the virulence plasmid of the *S. flexneri* strain 2457T using oligonucleotides W81 and W82 (Appendix 1). The product was digested with *Sal*I and *Pst*I and the resulting fragment was ligated with pHJW20 that had also been digested with *Sal*I and *Pst*I. This created a P*icsP-lacZ* fusion carrying 1473 bp of sequence naturally found upstream of the *icsP* transcription start site. The promoterless reporter plasmid pMIC21 was created by removing the entire *icsP* promoter from pHJW20. To do this, pHJW20 was digested with *Pst*I and *Xba*I, the resulting DNA was treated with T4 DNA polymerase and self-ligated using the blunt-ended ligation protocol supplied with T4 DNA ligase (Promega). The sequence of all truncated promoters fragments was confirmed after cloning. All plasmids are described in Appendix 1.

***In silico* analyses of the *icsP* promoter and upstream sequences.**

Curvature predictions of the *icsP* promoter and upstream sequences were performed using the bend.it® program (http://www.icgeb.org/dna/bend_it.html. The curvature is calculated as a vector sum of dinucleotide geometries (roll, tilt and twist angles) using the BEND algorithm (Goodsell & Dickerson, 1994) and expressed as degrees per helical turn (10.5°/helical turn = 1°/bp). Experimentally tested curved motifs produce curvature values of 5–25°/helical turn, whereas straight motifs give values below 5°/helical turn.

To analyze sequences upstream of the *icsP* tss, the program "Clone Manager 9, Basic Edition" (Scientific and Educational Software) was used. Open reading frames (ORFs) greater than 50 amino acids were identified using the "ORF search" feature and the "Open Reading Frame analysis" feature was used to predict whether ORFs were coding regions. The "Open Reading Frame analysis" feature uses Fickett's TESTCODE algorithm (Fickett, 1982) to generate a score that can help to determine whether or not an ORF is likely to be a protein coding region and is based on the observed asymmetry of base utilization within coding regions. ORFs with scores over 0.95 are predicted to be coding regions. To predict whether predicted coding regions identified by our search were linked to possible promoter sequences, the BPROM program (Softberry) was used. This algorithm predicts potential transcription start positions of bacterial genes regulated by sigma70 promoters (major *E. coli* promoter class). Linear discriminant function combines characteristics describing functional motifs and oligonucleotide composition of these sites. BPROM has an accuracy of *E. coli* promoter recognition of about 80%.

Site-directed mutagenesis of upstream putative VirB-binding sites.

The two upstream putative VirB-binding sites were mutated using site-directed mutagenesis by PCR using a two step procedure, as described previously (Lie & Leigh, 2007) using primers listed in Table 2. In the first step, the 5' binding site and the 3' binding site were amplified in two separate PCRs. To amplify the 5' binding site, a reverse oligonucleotide, either mutagenic (W100) or complementary to the wild-type sequence of the putative VirB-binding site (W99), was used with a forward primer (W98). To amplify the 3' binding site a forward primer, either mutagenic (W110) or complementary to the natural sequence of the putative VirB-binding site (W101), was used with a reverse primer (W89). These reactions created four products: a 5' binding site with wild-type sequence; a mutated 5' binding site; a 3' binding site with wild-type sequence and a 3' mutated binding site. All PCR products were gel purified and treated with Promega T4 DNA polymerase (as per manufacturer's instructions) to remove 3' overhanging nucleotides. In the second stage of the procedure, the PCR products were ligated together in three different combinations using Promega T4 DNA ligase (as per manufacturer's instructions): wild-type 5' sequences were ligated to mutagenic 3' sequences, mutated 5' sequences were ligated to wild-type 3' sequences, and mutated 5' sequences were ligated to mutated 3' sequences. Where necessary, 5' phosphates were incorporated during oligonucleotide synthesis (Appendix 1). The resulting ligations were used as template in a second round of PCR using oligonucleotides W89 and W98. This PCR allowed the amplification of correctly oriented promoter fragments. Correctly sized PCR products were then gel purified and digested with *Pst*I and *Bgl*II. The resulting DNA fragments were cloned into a pHJW20 *Pst*I *Bgl*II vector to give pMIC13,

pMIC17 and pMIC18. The sequence of all mutated promoters fragments was confirmed after cloning.

Quantification of *icsP* promoter activity using the P*icsP-lacZ* reporter and derivatives.

To measure promoter activities the P*icsP-lacZ* fusion plasmids described in this work were introduced into *S. flexneri* and *E. coli* strains by electroporation. Activities of the *icsP* promoter constructs were determined by measuring β-galactosidase activity as described previously (Wing *et al*, 2004), using the Miller protocol (Miller, 1972). Overnight cultures were diluted 1:100 and grown for 4 to 5 h in either TSB medium (*S. flexneri*) or LB (*E. coli*) at 37°C, prior to cell lysis. Routinely, β-galactosidase levels were measured in early stationary phase cultures grown from three independent transformants because experiments had shown that *icsP* expression significantly increases under these conditions (data not shown).

2.3 Results and Discussion

Identification of sequences required for VirB dependent regulation of *icsP* promoter

Previous work has demonstrated that VirB regulates an P*icsP-lacZ* fusion integrated into the *icsP* locus on the *Shigella* virulence plasmid (Wing *et al.,* 2004). To identify regions upstream of the *icsP* gene that mediate VirB dependent de-repression, a nested set of promoter deletions was created (Fig. 2A) and cloned into a medium-copy-number *lacZ* reporter pHJW20 (Appendix 1) to replace the existing 1,232 bp *icsP* promoter fragment. This created a set of eight P*icsP-lacZ* fusions whose upstream limits varied

from -1232 to -92 relative to the previously annotated transcription start site of the *icsP* gene (Egile *et al*, 1997) (Fig. 11A; Appendix 1).

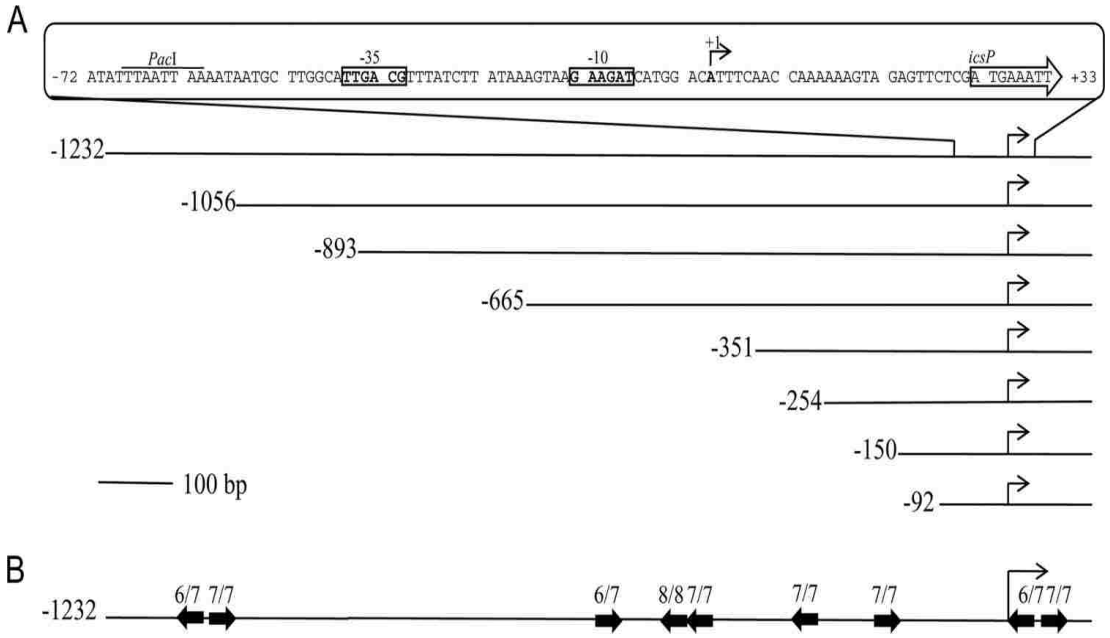

Fig. 11. Promoter elements of the *icsP* promoter and schematic representation of the truncated promoter series. (A) Angled arrows represent the *icsP* transcriptional start site (+1), determined previously (Egile *et al.,* 1997). The -10 and -35 hexamers are boxed and shown in bold. The translational start site is represented by an open arrow. The truncated promoters are drawn to scale and numbers represent the upstream boundary of the *icsP* promoter, relative to the tss (+1). The promoter fragments represented are found in pHJW20, pJS01, pJS02, pDH01, pJS03, pHJW34, pHJW35 and pHJW36, respectively. (B) Solid black arrows represent the relative position and orientation of the nine putative VirB binding sites identified within the promoter fragments used in this work. In each case, the match to the consensus sequence 5'-(A/G)(A/T)G(G)AAAT-3' (Tobe *et al.,* 1993, Watanabe *et al.,* 1990) is given.

Each promoter construct was introduced into wild-type *Shigella* 2457T and a mutant derivative, AWY3 (*virB::*Tn5), and β-galactosidase production was measured using the Miller protocol (Miller, 1972). Of the eight promoter fragments tested, only one displayed a >2-fold increase in the presence of VirB (Fig. 12). Surprisingly, this was the longest promoter fragment (found in pHJW20). The activity of this promoter fragment was 17-fold higher in the presence of VirB than its absence. This increase was unlikely to be caused by sequences in the pACYC184 plasmid backbone, because these

sequences would also influence the activity of the other promoter fragments. Furthermore, the increase in promoter activity was unlikely to be caused by the creation of a new VirB binding site at the boundary of the plasmid backbone and the promoter region, because two additional constructs with altered plasmid-promoter boundaries (pMIC01 and pMIC02 [Table 1]) were found to have similar activities to those displayed by pHJW20.

The simplest interpretation of these data was that DNA sequences located between positions 1232 and 1056 upstream of the *icsP* transcription start site are required for VirB dependent regulation of the *icsP* promoter.

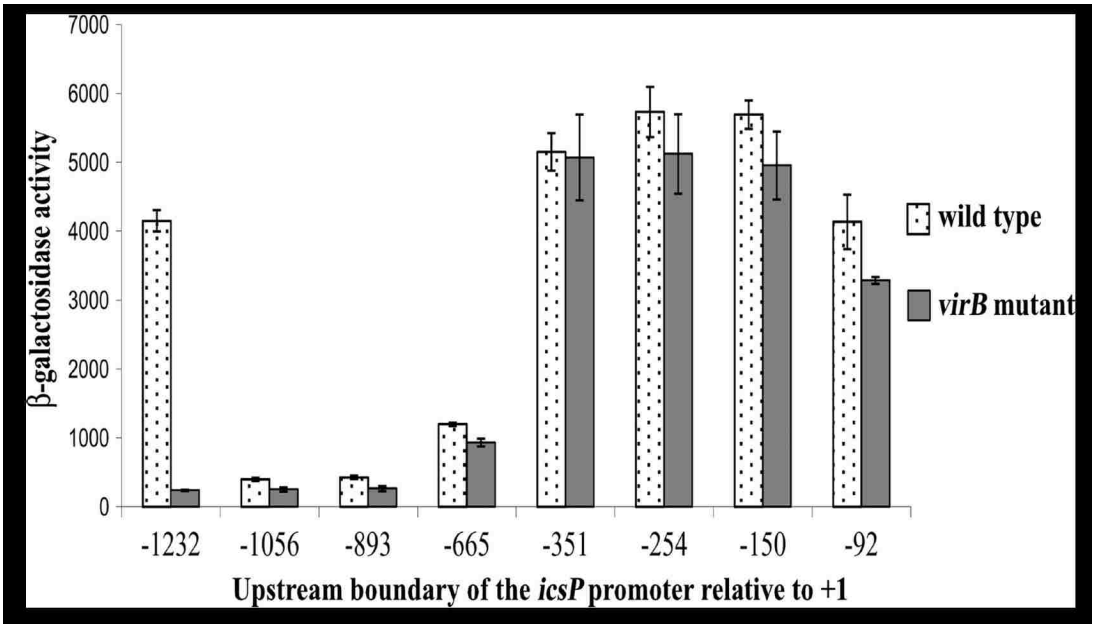

Fig. 12. Activities of the truncated *icsP* promoter series in *Shigella*. Bars indicate β-galactosidase expression of the P*icsP-lacZ* fusions in wild-type *S. flexneri* (2457T) and an isogenic strain lacking *virB* (AWY3). β-galactosidase activities are expressed in Miller units. Assays were run in triplicate and the mean and standard deviations are shown.

Identification and site-directed mutagenesis of putative VirB binding sites responsible for VirB-dependent regulation of the *icsP* promoter.

Previous analysis of the *icsB*, *spa15* and *virA* promoters of *Shigella sonnei* established a consensus binding site for VirB (Taniya *et al*, 2003). Our analysis of sequences upstream of the *icsP* gene identified nine sites with greater than a 6/7 match to the consensus, 5'-(A/G)(A/T)G(G)AAAT-3' (Fig. 12B; Turner & Dorman, 2007). Three of these sites are located between -665 to -351. The location of these sites may explain why small and yet significant increases in VirB dependent promoter activity were observed in wild-type *Shigella flexneri* carrying promoter constructs with upstream boundaries of -665, -883 and -1056 relative to the *icsP* transcription start site (+1) (Fig. 11). Two other putative VirB binding sites are located immediately downstream of the transcription start site. Although the location of these sites may explain the small increase in VirB dependent promoter activity associated with the shortest promoter fragment used in our studies, our data suggest that these sites alone play no significant role in the presence of upstream promoter sequences.

Interestingly, two of the nine sites identified upstream of the *icsP* gene are located between positions -1144 and -1130 and are organized as an inverted repeat (Fig. 11B). Since our truncation analysis indicates that sequences between -1232 and -1063 are essential for a 17-fold increase in promoter activity in the presence of VirB (Fig. 12), we chose to analyze these sites further. Seven base pair substitutions were made in either, the upstream site (box 2), the downstream site (box 1) or both, using a PCR site-directed mutagenesis method described by Lie & Leigh (2007) (Appendix 1; Table 1).

Table 1. Summary of mutations introduced into the two boxes that form the upstream inverted repeat and activities of wild-type and mutated *icsP* promoter fragments

Fragment description	Sequence[a]	β-galactosidase activity [b]	
		virB$^+$	*virB* mutant
WT Box 1 & 2	CGGGGATTTCAGTATGAAATGAAGTA	4291 ± 453	397 ± 12
Mutated Box 1	CGGGGATTTCAGT**CGACCCG**GAAGTA	345 ± 46	378 ± 29
Mutated Box 2	CGGGG**GCCCAGC**TATGAAATGAAGTA	370 ± 65	359 ± 29
Mutated Box 1 & 2	CGGGG**GCCCAGC**T**CGACCCG**GAAGTA	360 ± 57	371 ± 26
Promoterless *lacZ*		346 ± 70	415 ± 33

[a] 5'-3' DNA sequences of the wild-type and mutated boxes that form the upstream inverted repeat. Sequences lie between -1144 and -1130 relative to the annotated tss of *icsP* (+1;(Egile *et al.,* 1997)). Bases in bold indicate those mutated by site-directed mutagenesis.

[b] All promoter fragments were fused to *lacZ* and β-galactosidase activities were measured in wild-type *S. flexneri* (2457T) and the isogenic strain that lacks *virB* (AWY3). The parent cloning vector with a promoterless *lacZ* gene was included as a negative control. β-galactosidase activities are expressed in Miller units. Assays were run in triplicate and the mean and standard deviations are shown.

Each mutated promoter fragment was then introduced into the *lacZ* reporter plasmid pHJW20 to replace the existing wild-type sites the resulting plasmids pMIC13, pMIC17, pMIC18, and pHJW20 and a promoterless control, pMIC21 (Table 1) were introduced into wild-type *Shigella* (2457T) and the *virB* mutant (AWY3) and β-galactosidase levels were measured (Table 1).

Our data revealed that complete mutagenesis of the upstream binding site (box 2), the downstream binding site (box 1) or both, resulted in complete loss of VirB dependent regulation of the *icsP* promoter. Further more, these results were not an artifact of the *lacZ* reporter constructs, because similar patterns of expression were observed when the *icsP* gene was fused to each of our promoter constructs and IcsP levels were measured by Western blotting (data not shown). These data strongly suggest that VirB regulates the *icsP* promoter from sequences located more than 1 kb upstream of the *icsP*

transcription start site. To our knowledge, this is the first evidence that VirB can influence promoter activity from such distal sites.

Conserved sequence and location of the two distal VirB binding sites in all known *Shigella* sequences and in EIEC strain HN280.

To examine how well conserved DNA sequences located upstream of the *icsP* promoter are among other *Shigella* strains, species and other enterics, a 2 kb sequence upstream of the *icsP* gene in *Shigella flexneri* 2457T was subjected to BLAST analysis. All known *Shigella* virulence plasmid sequences and the virulence plasmid of the enteroinvasive *Escherichia coli* (EIEC) strain HN280 contain nearly identical sequences (99-100% identity) over the entire 2 kb sequence upstream of the *icsP* gene. Furthermore, the upstream inverted repeat identified by our studies was 100% identical in all strains and located in exactly the same position relative to the annotated transcription start site identified in *S. flexneri*. These findings strongly suggest that *icsP* genes found in all *Shigella* spp. and the EIEC strain HN280 are likely to be regulated by VirB from a biding site located more than 1 kb upstream of the gene.

VirB is structurally homologous to plasmid partitioning proteins, which can influence transcription from distances of several kilobases.

While it is unusual for transcription factors to influence transcription from distances greater than 200 bp upstream or downstream of the transcription start site in bacteria (Balleza *et al.*, 2009, Collado-Vides *et al*, 1991), some examples exist. For example, the enhancer of the *Bacillus subtilis rocG* gene is located 1.5 kb downstream of the promoter and, beyond the end of the *rocG* coding region (Belitsky *et al.,* 1999) and the two NtrC binding sites required for transcriptional activation of *E. coli* σ^{54}-

regulated *glnA* promoter can still function when placed as far as 3 kb from the promoter (Reitzer & Magasanik, 1986). Furthermore, bacterial plasmid partitioning factors, while not typically considered transcription factors, have also been shown to silence the promoters of genes in the vicinity of their *cis*-acting binding sites from over distances of several kilobase pairs (Kim & Wang, 1999, Rine, 1999, Rodionov *et al*, 1999, Rodionov & Yarmolinsky, 2004). One of these proteins, the P1 ParB protein, displays structural homology to VirB and has bridging capabilities: the ability to interact with other ParB monomers located at binding sites further up or downstream. It is therefore possible that the other seven sites with close matches to the VirB consensus binding site play an important role in bridging by VirB and the resulting DNA topology is central to the alleviation of H-NS dependent repression of the *icsP* promoter, although this needs to be tested.

In summary, although transcription factors typically bind to sequences located within 200 bp upstream or downstream of the transcription start site (Balleza *et al*, 2009, Collado-Vides *et al*, 1991), here we provide strong evidence that VirB has the capacity to alleviate H-NS dependent repression of the *icsP* promoter from sites located over 1 kb upstream of the transcription start site. This raises two important questions. (i) Are other *Shigella* virulence plasmid genes regulated from remote VirB binding sites? (ii) Is it common for transcriptional anti-silencing mechanisms to employ distal regulator binding sites? Future studies will address these questions and elucidate the molecular mechanism of H-NS dependent repression and VirB dependent de-repression of the *icsP* promoter.

CHAPTER 3

TRANSCRIPTIONAL SILENCING AND ANTI-SILENCING OF THE *ICSP* PROMOTER OF *SHIGELLA FLEXNERI*: REMOTE REGULATION BY H-NS AND VIRB

Abstract

In the bacterial pathogen *Shigella flexneri,* many genes encoded by the large virulence plasmid are repressed by the nucleoid structuring protein H-NS and de-repressed by the major virulence gene activator VirB. One example is the *icsP* gene, which encodes an outer membrane protease. In this study, we demonstrate 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), we show 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. These data support a model where H-NS docked at remote sites may act in concert with H-NS docked at promoter proximal sequences to mediate H-NS-dependent repression of the *icsP* promoter. While we again demonstrate that VirB functions solely to de-repress the promoter, our EMSA data indicate that VirB and H-NS are capable of binding simultaneously to the full intergenic region (1232 bp) that lies immediately upstream of the *icsP* gene. This suggests that VirB may function to relocate H-NS, rather than causing H-NS to dissociate from the DNA upstream of the *icsP* gene. Our findings contribute to our understanding of transcriptional silencing and anti-silencing mechanisms that control virulence gene expression in *Shigella* and other enteric

pathogens and raise the possibility that other bacterial promoters may be regulated from remote transcription factor binding sites.

Authors: Dustin J. Harrison, Karen M. Levy, Amber J. Howerton, Christopher T. Hensley and Helen J. Wing

We thank Dr. Eric Anderson and Dr. Martin Roop from East Carolina University for all their help and guidance with the EMSAs, Stephanie Labahn for technical support and Drs. Eduardo Robleto, Dennis Bazylinski and Patricia Cruz for critical reading of the manuscript.

3.1 Introduction

Shigella species are gram-negative intracellular pathogens that cause bacillary dysentery in humans. *Shigella* infections are responsible for approximately 164 million infections each year resulting in 1.1 million deaths, most of these occur in developing countries and in children less than 5 years old (Kotloff *et al*., 1999, Lee *et al*., 2005, Li *et al*., 2009). All four *Shigella* species carry a large (230 kbp) virulence plasmid. Many genes encoded by these plasmids are under the transcriptional control of the histone-like nucleoid structuring protein (H-NS) and the major virulence gene activator VirB (Adler *et al*., 1989, Berlutti *et al*., 1998).

H-NS is a nucleoid structuring protein that is well characterized as a silencer of bacterial transcription (Ali Azam *et al*., 1999, Dorman, 2004). In enteric pathogens, H-NS silences the transcription of many virulence genes, which have been acquired by horizontal gene transfer (Dorman, 2009a). Silencing of newly acquired genes by H-NS is thought to protect the bacterium from potentially deleterious effects of transcribing genes that may compromise their fitness (Navarre *et al*., 2006) until they can be integrated into

existing regulatory networks (Navarre *et al*., 2007, Dorman, 2009b). Among the enterics, horizontally acquired genes are frequently AT rich. Such sequences typically display high levels of intrinsic curvature. Previous investigations indicated that H-NS is a direct, non-specific DNA binding protein with an affinity for DNA with high intrinsic curvature (Spassky *et al*., 1984, Zhang *et al*., 1996, Zuber *et al*., 1994, Williams *et al*., 1996, Williams & Rimsky, 1997, Yamada *et al*., 1991, Zuber *et al*., 1994) which may explain, at least in part, why H-NS is the protein of choice for transcriptional silencing in these bacteria. More recent work at the *E. coli proU* promoter (Lang *et al*., 2007) has led to the description of a high affinity H-NS binding site, however it remains unclear whether this sequence is found in all H-NS regulated promoters, or whether other sequences can contribute to H-NS binding. Regardless, DNase I protection assays have demonstrated that H-NS oligomerizes along AT rich sections of bacterial promoters (Rimsky *et al*., 2001, Turner & Dorman, 2007). This can lead to the formation of H-NS-DNA-H-NS bridges, which have been visualized using atomic force microscopy as looped DNA structures constrained by H-NS (Maurer *et al.,* 2009). These structures may involve long stretches of DNA, extending over 1 kb (Dame *et al*., 2000, Dame *et al*., 2005, Dame *et al*., 2006, Maurer *et al*., 2009), and have been documented as playing important roles in promoter silencing (Bouffartigues *et al*., 2007, De la Cruz *et al*., 2009, Dorman, 2004, Dorman, 2009a, Fang & Rimsky, 2008, Navarre *et al*., 2006, Stoebel *et al*., 2008). Disruption of these H-NS:DNA complexes by other DNA-binding proteins can allow transcriptional de-repression; a process coined anti-silencing (Fang & Rimsky, 2008, Navarre *et al*., 2007, Stoebel *et al*., 2008).

The virulence gene activator VirB is encoded by the *Shigella* virulence plasmid and its expression is regulated by another transcription factor VirF (Adler *et al*., 1989). At the non-permissive temperature of 30 ˚C many genes encoded by the virulence plasmid are repressed by H-NS (Beloin & Dorman, 2003, Hromockyj *et al*., 1992, Porter & Dorman, 1994). Upon a switch to 37 °C, H-NS dissociates from the *virF* promoter leading to the production of the virulence gene regulator, VirF (Adler *et al*., 1989). VirF relieves H-NS-dependent repression of the v*irB* promoter (Tobe *et al*., 1993, Watanabe *et al*., 1990) and subsequently VirB alleviates H-NS dependent repression of many virulence plasmid genes, including *virA, icsP, ospB, phoN2* and those encoding the type III secretion system that mediate host cell invasion; *ipa*, *mxi* and *spa* (Beloin & Dorman, 2003, Berlutti *et al*., 1998, Dorman, 2006, Hromockyj *et al*., 1992, Maurelli *et al*., 1984a, Maurelli *et al*., 1984b, Nicoletti *et al*., 2008, Porter & Dorman, 1994, Porter *et al*., 1998, Prosseda *et al*., 1998, Sasakawa *et al*., 1993, Tobe *et al*., 1991, Tobe *et al*., 1993, Tobe *et al*., 1995). Consequently, transcriptional de-repression of virulence genes by VirF and VirB is central to the pathogenicity of *Shigella.* Since VirF and VirB both function to antagonize H-NS mediated repression of virulence genes, both proteins can be considered transcriptional anti-silencing proteins.

The mechanism underlying transcriptional silencing/anti-silencing has been proposed to be a straightforward molecular antagonism involving competition between the silencing protein and the anti-silencing protein for their overlapping cognate DNA binding sequences, which are usually located in the immediate vicinity of the promoter. Typically this competition results in the dissociation of the silencing molecule from the DNA (Fang & Rimsky, 2008, Navarre *et al*., 2007). Examples of this kind of

straightforward transcriptional silencing/anti-silencing exist at a variety of bacterial promoters, e.g. SlyA counters H-NS silencing of the hemolysin gene *hlyE* in *E. coli* through direct competition for the DNA, RovA from *Yersinia* counters silencing of the invasin *inv,* and the urease gene activator UreR of *Proteus mirabilis* causes H-NS to dissociate from DNA located between the divergently transcribed *ureR* and *ureD* genes (Coker *et al*., 2000, Ellison & Miller, 2006a, Ellison & Miller, 2006b, Heroven *et al*., 2004, Heroven *et al*., 2007, Lithgow *et al*., 2007, Poore & Mobley, 2003). In contrast, at some bacterial promoters the anti-silencing proteins function to alter the DNA topology in a way that no longer allows H-NS to constrain transcription. For example, LeuO counters silencing of the quiescent porin *ompS1* in *Salmonella* by altering the topology of the *ompS1* promoter region (De la Cruz *et al*., 2007, De la Cruz *et al*., 2009) and in *Shigella flexneri,* VirB antagonizes H-NS by affecting the structure of the *icsB* promoter (Turner & Dorman, 2007).

Our previous work shows that the horizontally acquired, monocistronic *icsP* gene, located on the *Shigella* virulence plasmid, is repressed by H-NS and de-repressed by VirB. We have demonstrated that the role of VirB is to solely alleviate transcriptional repression of the *icsP* promoter, which is mediated by H-NS (Castellanos *et al*, 2009, Wing *et al*., 2004). Surprisingly, two putative VirB binding sites, organized as an inverted repeat and located over 1 kb upstream of the annotated transcription start site are indispensible for the de-repression of the promoter (Castellanos *et al*., 2009). These sites lie in a designated 1.2 kb intergenic region (Jin *et al*, 2002); Reference Sequence: NC_004851.1). Currently, it remains unclear which DNA sequences are required for the H-NS dependent repression of *icsP*. Furthermore, although our previous work strongly

suggests VirB binds directly to this upstream intergenic region, this has never been categorically demonstrated.

In this study, we further examine the role that H-NS and VirB play in the regulation of the *icsP* promoter in an attempt to discern the molecular mechanism responsible. Our results suggest that both H-NS and VirB regulate the activity of the *icsP* promoter from remotely located sequences. We propose a model to explain how silencing and anti-silencing of the *icsP* promoter is achieved by DNA binding-proteins binding to remote sequence elements. We envisage that an improved understanding of H-NS and VirB dependent regulation of the *icsP* promoter will provide insight into bacterial regulation of transcription from remote sites.

3.2 Materials and Methods

Bacterial strains and plasmids.

The bacterial strains and plasmids used in the present study are given in Appendix 1. Bacteria were grown routinely at 37 °C in Luria-Bertani broth, with aeration, or on LB agar (LB broth containing 1.5 % [wt/vol] agar). Antibiotics were added at the following final concentrations: ampicillin, 100 μg ml^{-1}; chloramphenicol, 25 μg ml^{-1}; kanamycin, 50 μg ml^{-1}; and tetracycline, 12.5 μg ml^{-1}.

Construction of the P*icsP-lacZ* reporter plasmid and truncated promoter fragments.

The P*icsP-lacZ* reporter plasmid pHJW20 was used as the basis for this work (Castellanos *et al*., 2009, Wing *et al*., 2004); Table 1). pHJW20 carries 1232 bp upstream of the annotated transcription start site of the *icsP* promoter, the first 48 bp of the *icsP* coding region cloned upstream of a translation stop site, and a promoterless *lacZ* gene, so that expression of *lacZ* is directly regulated by the *icsP* promoter. Promoter

fragments of 1056 bp, 893 bp, 665 bp, 351 bp, 254 bp, 150 bp, and 92 bp, were PCR amplified and cloned into pHJW20 to replace the 1232 bp *icsP* promoter (Castellanos *et al*., 2009). This created a nested set of P*icsP-lacZ* fusion plasmids that could be used to identify regions of the *icsP* promoter needed for H-NS and VirB-dependent regulation. All cloned promoters have been sequenced and are identical to the sequence reported for pCP301 (*S. flexneri* serotype 2a: NCBI Reference Sequence: NC_004851.1). All plasmids are described in Appendix 1.

Quantification of *icsP* promoter activity using the P*icsP-lacZ* reporter and derivatives.

Activity of the *icsP* promoter was determined by measuring β-galactosidase activity as described previously (Wing *et al*., 2004) using the Miller protocol (Miller, 1972), in strains carrying pHJW20 or derivatives. Freshly transformed cells were back-diluted 1:100 in LB and grown at 37 °C for 4 to 5 h to early stationary phase prior to cell lysis, because *icsP* expression is maximal under these conditions (Wing *et al.,* 2004). When appropriate, to examine the effect of expressing *virB* from the pBAD vectors, cells were instead back-diluted 1:100 in 5 ml of LB medium containing 0.2% (wt/vol) glucose and, after 4 to 5 h were harvested, washed with an equivalent volume of LB medium and diluted 5-fold into LB medium containing either 0.08% (wt/vol) L-arabinose. Cultures were then grown for an additional 2 h before being harvested. All assays were performed in triplicate on three separate occasions.

***In silico* analyses of the *icsP* promoter and upstream sequences.**

The MUTACURVE server (http://132.248.32.45/cgi-bin/mutacurv.cgi) was used to predict regions of intrinsic DNA curvature. This program evaluates the amplitude of

intrinsic curvature for every nucleotide in a given sequence using the algorithm of Goodsell & Dickerson (1994), with the addition of Satchwell's contribution matrices for rotational and spatial displacements (Satchwell *et al*., 1986). These algorithms have been routinely used by other investigators to identify putative HNS binding sites in DNA (Beloin & Dorman, 2003, De la Cruz *et al*., 2009, Flores-Valdez *et al*, 2003,Will & Frost, 2006).

Construction of inducible plasmids and purification of the VirB and H-NS proteins.

The *virB* gene was digested from pATM324 using *Hin*dIII and *Nco*I restriction enzymes and gel purified. The primers W38 and W39 were used to amplify the *virB* gene which was then ligated into pQE-60 (Qiagen) before electroporation into *E. coli* M15 pREP4 creating a C-terminal His-tagged VirB in an inducible plasmid, pAJH01. A similar strategy was used to create a C-terminal His-tagged H-NS in an inducible plasmid, pCTH01. Briefly, the primers W134 and W137 (Appendix 1) were used to amplify the *hns* gene from *Shigella flexneri* serotype 2a with *Nco*I and *Bgl*II sites for cloning into pQE-60.

His-tagged-VirB and His-tagged-H-NS proteins were overproduced from plasmid pAJH01, and pCTH01 respectively. Proteins were expressed in the *E. coli* strain M15 carrying the plasmid pREP4. The expression of the C-terminally His-tagged proteins were induced in 500 ml cultures growing exponentially with 1mM IPTG (isopropyl-β-thiogalactopyranoside). Two hours post-induction, the cells were harvested and frozen at 80 °C overnight. The cell pellet was thawed on ice and resuspended in lysis buffer (40 mM Tris-base [pH 8.93], and 80 mM NaCl). Cells were lysed by sonication, and cellular debris was pelleted by centrifugation at 10 000 x g at 4 °C. Cell lysates were applied to

Ni-NTA columns (Qiagen) pre-equilibrated with equilibration buffer (40 mM Tris-base [pH 8.93], 80 mM NaCl, 10 mM imidazole). The columns were then washed with 10 bed volumes of wash buffer (equilibration buffer with 20 mM imidazole) prior to elution of the proteins by the addition of 2.5 bed volume of elution buffer (equilibration buffer with 250 mM imidazole). Eluates were analyzed by SDS-PAGE followed by Coomassie staining. Purified proteins were dialyzed in equilibration buffer overnight at room temperature. Protein concentrations were determined using Bradford Assays. The hexa-his tag was not found interfere with VirB expression or activity, because His-tagged VirB was observed to restore IcsP expression to wild type levels in a *strain* lacking *virB in vivo* (data not shown). Furthermore, previous studies using an identical His-tag H-NS fusion protein, produced in a manner similar to that described above, was shown to retain normal function of H-NS in assays (Williams & Rimsky, 1997).

Electrophoretic Mobility Shift Assays (EMSAs).

The upstream intergenic region of the *icsP* promoter (1232 bp) was used as a template to amplify six, overlapping smaller fragments, ranging from 252 bp to 358 bp by PCR (Appendix 1). A positive control promoter fragment, *icsB*, was amplified by PCR using the *Shigella flexneri* virulence plasmid as a template, a negative control fragment, *pstS*, was amplified from the *E. coli* genome (Appendix 1) and a non-specific competitor, a 147 bp fragment of DNA from pACYC184 was amplified from pACYC184 by PCR. The full-length upstream intergenic region of the *icsP* promoter (*icsP'*) was also used in EMSA (generated using primers W63 and W149), as well as a 1011 bp negative control fragment of the *E. coli pstS* gene (*pstS'*) (Appendix 1). 100 nM of each DNA fragment was incubated with increasing concentrations of purified protein, VirB, H-NS, or both (0

to 920 nM) for 30 minutes at 30 °C for both VirB and H-NS, in a 20 µl reaction mixture containing 40 mM Tris-base (pH 8.93), 80 mM NaCl, 1 mM EDTA, 1 mM DTT, 200 ng ml^{-1} bovine serum albumin (BSA), 200 nM of competitor p184 DNA and 10 % glycerol. DNA concentrations were measured using the NanoView (GE). The protein-DNA complexes were resolved by electrophoresis in 1.5 % TBE agarose gels for 4 h at room temperature at 60 V constant. Following electrophoresis, the agarose gels were stained with ethidium bromide (1 µg ml^{-1}) and visualized using a Typhoon 9410 (Amersham) variable mode imager. EMSAs were performed a minimum of three times and representative results are shown.

3.3 Results

Identification of regions necessary for the H-NS dependent repression of the *icsP* promoter

Although we have demonstrated that DNA sequences located over 1 kb upstream of the annotated *icsP* transcription start site are required for VirB-dependent de-repression

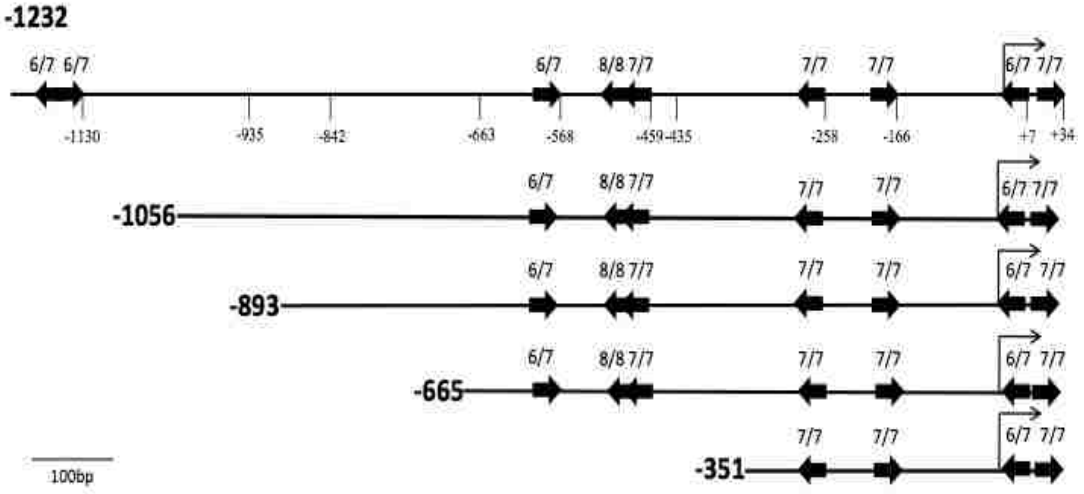

Fig. 13. Schematic of the 1.2 kb region located upstream of the *icsP* gene. Relative positions and orientation of putative VirB-binding sites are shown (Castellanos *et al*., 2009). Fragments of different lengths represents regions DNA cloned upstream of our *lacZ* reporter in our truncation series. Coordinates are given in bp and relative to the annotated transcription start site (characterized by Egile *et al*., 1997).

of the *icsP* promoter (Castellanos *et al*., 2009), little is known about the sequences required for H-NS-dependent repression of the *icsP* promoter. To identify regions of the *icsP* promoter required for H-NS-dependent repression we used a nested set of P*icsP-lacZ* fusion plasmids (constructed previously in Castellanos *et al*., 2009). The promoters contained within this P*icsP-lacZ* fusion series share the same downstream boundary, but their upstream boundaries differ, ranging from -1232 to -351 with respect to the annotated transcription start site (+1) (Egile *et al*., 1997) (Fig.13).

Each reporter plasmid was transformed into the *E. coli* strain MC4100 and an isogenic strain lacking *hns* (MC4100 *hns)*. These strains were used to avoid interference arising from the fact that in *Shigella virB* is directly regulated by H-NS and the fact that *Shigella hns* mutants are notoriously unstable (Maurelli, personal comm.). A similar strategy has been used by others to study the regulation of other *Shigella* promoters by H-NS (Beloin & Dorman, 2003). A plasmid carrying an inducible copy of the *virB* gene was next introduced into each of the *E. coli* strains (neither strain expresses *virB* naturally) so that the effect of H-NS on the *icsP* promoter could be measured in the presence and absence of VirB.

The results of the β-galactosidase assay are shown in Fig. 14. In the wild type background, the four longest promoter fragments had significantly lower activity than in the *hns* mutant. These data suggest sequences upstream of -351 are sufficient for H-NS-dependent repression, although sequences upstream of -665 are required for full repression. Induction of VirB in the wild-type background increased the activity of the longest promoter by 10 fold. This is consistent with data obtained in a *Shigella* background (Castellanos *et al*., 2009) and demonstrates that VirB alleviates H-NS

dependent repression of the *icsP* promoter when sequences located over 1 kb upstream of the annotated transcription start site are present. In the *hns* mutant background, all constructs displayed activity higher than 1000 units, regardless of the whether *virB* was induced or not. This indicates that H-NS mediates repression of the four longest constructs in our promoter series and that VirB functions solely to relieve H-NS dependent repression at the longest construct. Similar observations were made in previous studies (Wing *et al*., 2004, Castellanos *et al*., 2009).

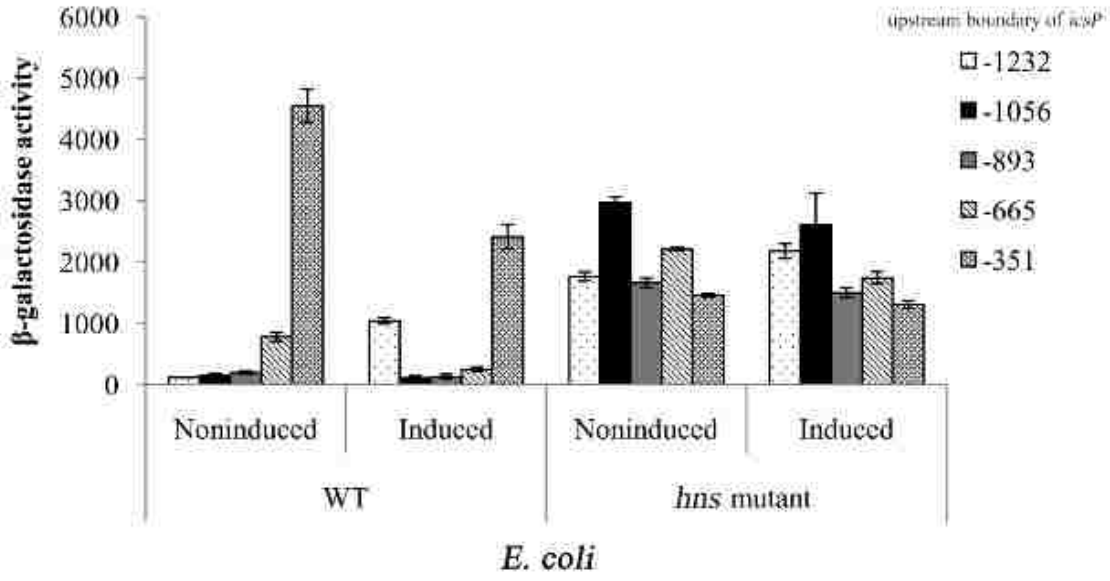

Fig. 14. Activities of the truncated *icsP* promoter fragments in the presence and absence of H-NS and /or VirB. Constructs are labeled according to the upstream boundary from the transcription start site. Bars indicate β-galactosidase expression of the P*icsP-lacZ* fusions in wild type *E. coli* MC4100 and an isogenic strain lacking *hns* (MC4100 *hns*::Kn^r) under inducing and non-inducing conditions for *virB* which is included on an inducible plasmid. β -galactosidase activities are expressed in Miller units. Assays were run in triplicate and error bars represent the mean and standard deviations.

Surprisingly, induction of VirB in the wild type background has a repressive effect on the activity of the two shortest promoter constructs examined, -665 bp and -351 bp. It is not clear whether VirB mediates this effect directly, although putative VirB binding sites have been identified within these fragments (Fig. 4; Castellanos *et al*., 2009).

Furthermore, in the context of the shortest promoter construct (-351), the typical role of H-NS and VirB is reversed; VirB still functions to antagonize H-NS, but now functions to alleviate H-NS-dependent activation. These data demonstrate that H-NS and VirB are true molecular antagonists, which, under the right circumstances, can reverse their roles. This finding provides valuable insight into the molecular mechanisms of transcriptional silencing and anti-silencing in bacteria. *In vitro* transcription assays will be performed to test the hypothesis that H-NS increases the activity of the short promoter constructs.

Taken together, these data demonstrate that sequences located upstream of -351 are needed for H-NS-dependent repression of the *icsP* promoter. In the absence of these upstream sequences, sequences downstream of 351 permit the H-NS-dependent activation of the *icsP* promoter. Although H-NS-dependent repression and activation of these truncated *icsP* promoter fragments is likely mediated by H-NS binding directly to DNA sequences located upstream of -351 and downstream of -351, respectively, this needed testing.

H-NS binds to two discrete regions upstream of the *icsP* gene

In order to determine whether H-NS binds to DNA located upstream of the *icsP* gene, electrophoretic mobility shift assays (EMSAs) were

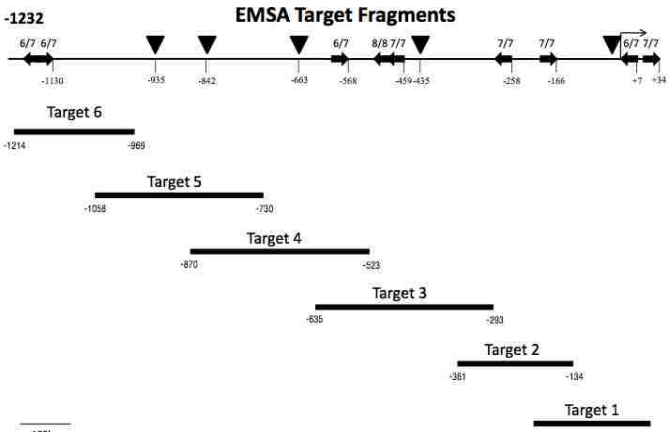

Fig. 15. Schematic showing the 6 target fragments used in the EMSAs. Coordinates of each target fragment are given relative to the transcription start site. Inverted triangles denote the position of a peak of high intrinsic curvature predicted using the MUTACURVE program.

used. The full intergenic region that lies upstream of the *icsP* gene (1232 bp) was divided up into six near equal length fragments and each of these fragments were used in EMSAs with purified His-tagged H-NS (Fig. 15).

For these experiments, a DNA fragment containing sequences that lie between the divergent *icsB* and *ipgD* genes of *Shigella flexneri* was used as a positive control (*icsB*), because H-NS has been shown to bind to this promoter region by others (Turner & Dorman, 2007) and an internal fragment of the *E .coli pstS* gene was used as a negative control (*pstS*). Each DNA fragment was incubated with increasing concentrations of purified His-tagged H-NS in the presence of a 147 bp fragment, which had been amplified from the plasmid pACYC184 and which served as non-specific competitor DNA (200 nM). DNA and DNA:protein complexes were then resolved by agarose gel electrophoresis.

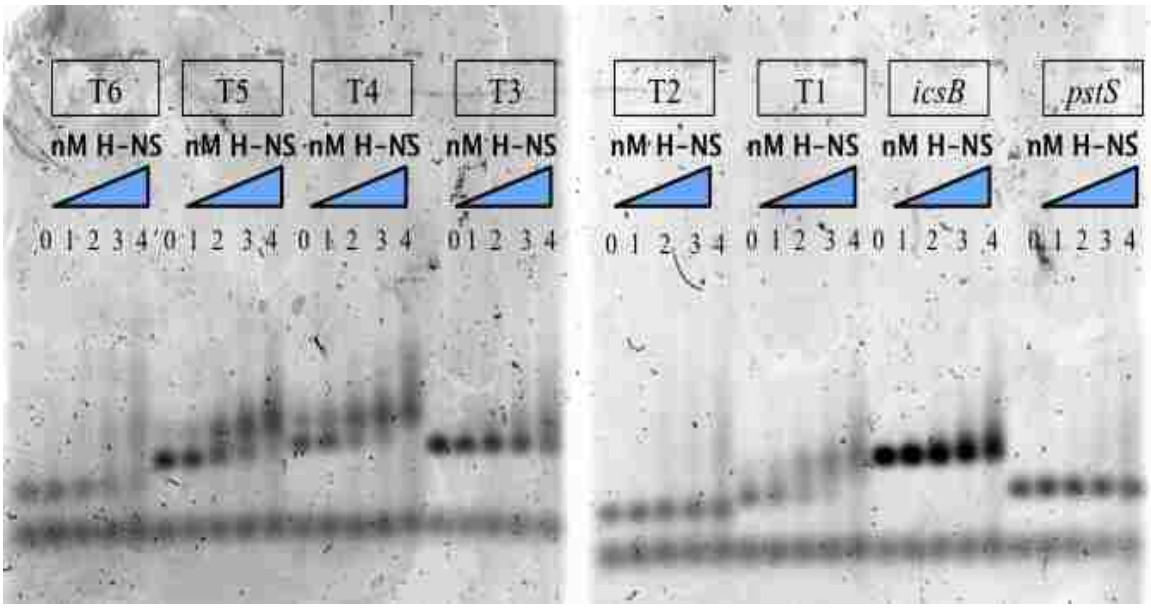

Fig. 16. H-NS binding to discrete regions within the *icsP* upstream intergenic region. EMSAs were conducted using 100 nM of each fragment shown in Fig. 3 and increasing concentrations of purified His-tagged H-NS. The concentrations of purified protein as well as the target fragment used are given above each lane. In all cases, the lower band is 200 nM of competitor DNA generated from pACYC184. Positive control, and negative controls are labeled *icsB* and *pstS* respectively. 0=no protein, 1=120 nM, 2=230 nM, 3=460 nM,4=920 nM protein.

Based on previous *in vitro* studies with H-NS and H-NS-regulated promoters, and due to the ability of H-NS to oligomerize along DNA, EMSAs performed with H-NS often do not appear as discrete shifted bands, but are more diffuse (Azam & Ishihama, 1999, Doyle *et al*., 2007). Despite this, we fully expected that some of the six DNA fragments would harbor regions that would interact with H-NS at lower concentrations than others. Our results show that H-NS binds to target fragments, T6, T5, T4, T3 and T1, with T5 and T4 shifting at the lowest concentration of protein, 120 nM (Fig. 16). The shift of T1 is apparent after incubation with 230 nM of protein while T6 and T3 only shift after incubation with H-NS concentrations of 460 nM. In contrast, T2 fails to shift in the presence of H-NS even at high concentrations of H-NS protein (920 nM). In each case, H-NS binding to the *icsP* promoter is specific as evidenced by the uniform presence of free non-specific competitor DNA in all lanes.

The DNA fragments that shift with the lowest concentration of H-NS protein, T5 and T4, contain sequences located between -1058 to -523 relative to the transcription start site. This region is present in all constructs that displayed full H-NS-dependent repression of the *icsP* promoter in our previous *in vivo* experiments (Fig 15; promoters with upstream boundaries of -1056, -893, and -665 all contain sequences contained within target T5 and T4). These data support our hypothesis that H-NS directly binds to the *icsP* upstream intergenic region to mediate H-NS dependent repression. Furthermore, T1 also shifted with relatively low concentration of H-NS. This fragment contains sequences located between 213 and +24 relative to the transcription start site. The binding of H-NS to this region of DNA may explain the H-NS-dependent increase in *icsP* promoter activity observed at the shortest promoter construct in our *in vivo* assays, although this

still needs to be tested. Also, future studies using DNase I footprint analyses, described in Chapter 4, will enhance our understanding of the nature of H-NS binding to these DNA target sequences.

Regions bound by H-NS are predicted to display high levels of intrinsic curvature

H-NS is well documented to bind preferentially to DNA that displays high intrinsic curvature (Dame *et al*., 2000, Dorman, 2006, Rimsky *et al*., 2001). To determine whether regions lying upstream of the *icsP* gene are likely to be intrinsically curved, *in silico* analysis of the *icsP* promoter was performed using the MUTACURVE program. Based on these analyses, five regions were found to display curvature values ≥7°/helical turn. Since experimentally tested curved motifs produce curvature values of 5–25°/helical turn, whereas straight motifs give values below 5°/helical turn, our findings are consistent with these five regions being curved and potentially binding H-NS.

Four of the regions predicted to be intrinsically curved lie upstream of -351 (at positions -935, -842, -663, -435; data not shown; inverted triangles Fig. 15). In our *in vivo* assays this DNA was required for the H-NS dependent repression of the *icsP* promoter (Fig. 15). Furthermore, H-NS bound specifically to this DNA in our EMSAs (Fig. 16). A fifth region of predicted curvature was found to lie around the transcription start site itself. This may explain why H-NS interacts directly with target 1 in our EMSAs (Fig. 16) and why this region mediates H-NS dependent activation of the *icsP* promoter in our β-galactosidase assays (Fig 14).

The finding that H-NS binds specifically to regions upstream of the *icsP* gene that are predicted to be intrinsically curved is consistent with the previously reports that H-NS display affinity for DNA sequences with high intrinsic curvature (Azam & Ishihama,

1999, Badaut *et al*, 2002, Beloin & Dorman, 2003, Berlutti *et al*., 1998, Bloch *et al*., 2003, Bouffartigues *et al*., 2007, Corcoran & Dorman, 2009, Dorman & Kane, 2009, Noom *et al*, 2007, Dame *et al*., 2000, Dame *et al*., 2005, Dame *et al*., 2006, De la Cruz *et al*., 2007, De la Cruz *et al*., 2009, Rimsky *et al*., 2001, Rimsky, 2004, Turner & Dorman, 2007). Although, more recent work at the *proU* promoter of *E. coli* has identified a 10 bp sequence which binds H-NS with high-affinity (Lang *et al*., 2007), a rudimentary search for such a sequence, within the upstream intergenic region of *icsP* to 100 bp downstream of the tss, yielded no good matches.

Preliminary experiments designed to test the curvature of these 5 regions, by electrophoresis in acrylamide gels at room temperature and 4 °C (curved DNA displays anomalous migration at 4 °C), suggest that these regions are indeed curved (*k values* all greater than 1 [*k value* = ratio of the apparent size of a DNA fragment/ true size of a DNA fragment; $k \geq 1$ = curved DNA]), however, this needs repeated (data not shown; De la Cruz *et al.,* 2009, Prosseda *et al.,* 2004).

VirB binds to the intergenic region that lies upstream of the *icsP* gene

Our previous studies revealed that sequences located 1 kb away from the annotated transcription start site are absolutely required for the VirB dependent regulation of *icsP* activity (Castellanos *et al.*, 2009). It remains unclear, however whether VirB binds directly to the *icsP* promoter or mediates is effect indirectly through another VirB-regulated protein. To test this, EMSAs using purified His-tagged VirB protein were used.

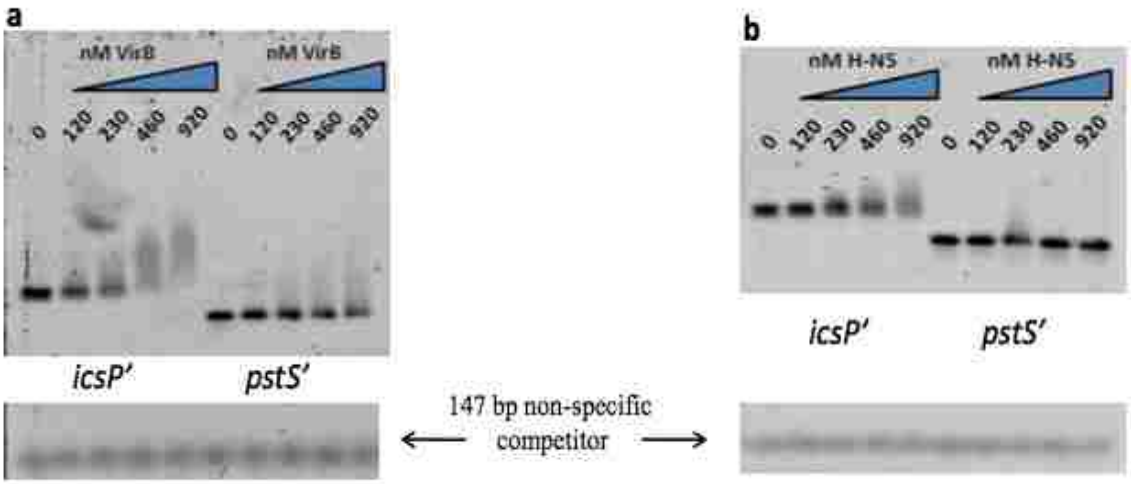

Fig. 17. VirB and H-NS binding of the full length *icsP* upstream intergenic region. EMSA results of the long, >1 kb, DNA fragments. a) shows the shift the full length *icsP* upstream intergenic region (*icsP'*), compared to the negative control, the 1011 bp fragment of the *E. coli pstS* (*pstS'*), b) shows H-NS the shift of the full length *icsP* upstream intergenic region (*icsP'*), compared to the negative control. The concentrations (nM) of purified protein are given above each lane with the identity of the target given below. In all cases the lower band represents 200 nM of non-specific competitor DNA derived from pACYC184.

The full upstream intergenic region of the *icsP* promoter (*icsP'*; 1232 bp) was incubated with increasing concentrations of purified His-tagged VirB. Again, an internal fragment of the *E. coli pstS* gene was used as a negative control for these experiments, but a longer fragment (1011 bp) was used to mirror the increased length of the fragment of interest. As previously described, a 147 bp fragment amplified from the plasmid pACYC184 (p184; 200 nM) was included in each reaction as a non-specific competitor DNA and DNA:protein complexes were resolved by agarose gel electrophoresis.

Our data show that a DNA fragment amplified from the entire intergenic region (1232 bp) located upstream of the *icsP* gene, was retarded by 230 nM of VirB, with complete retardation after incubation with 460 nM of VirB (Fig. 17a). The binding of VirB to this DNA was specific at concentrations between 230-460 nM, because VirB does not bind to the non-specific competitor DNA p184 nor the negative control fragment *pstS'* at these

concentrations. Some non-specific binding to the *pstS'* fragment was observed at VirB concentrations above 460 nM, however.

Although the VirB-dependent shift of the upstream intergenic region appears as a smear in these EMSAs, this appears to be a characteristic of VirB:DNA interaction *in vitro,* because Turner & Dorman (2007) reported similar observations when studying VirB:DNA interactions with the *Shigella icsB* promoter by EMSAs. Furthermore, although the magnitude of the shift is not as dramatic as commonly seen in EMSAs using DNA fragment of approximately 100-300 bp, it should be remembered that the DNA probe in these experiments much larger. It is likely that this accounts for the moderate shift observed. Regardless, there is a specific loss of free *icsP'* probe DNA with increasing concentrations of protein. We therefore conclude that VirB binds directly to the 1.2 kb region that lies immediately upstream of the *icsP* promoter, and this binding is likely required for the VirB-dependent regulation of the *icsP* promoter.

H-NS does not bind to discrete regions upstream of the *icsP* gene cooperatively

Both our *in vivo* and *in vitro* data strongly suggests that H-NS binds to two discrete regions upstream of the *icsP* gene (upstream of -351 and downstream of -213). We next wanted to determine whether H-NS binds cooperatively to these regions when these regions are juxtaposed in the context of the whole upstream intergenic region. If so, we predicted that lower concentrations of H-NS would be required to bind the whole region upstream of the *icsP* gene than fragments of it.

To test this the full length upstream intergenic region (1232 bp; *icsP'*) and the negative control fragment (1011 bp internal to the *pstS* gene; *pstS*') were incubated with increasing concentrations of purified His-tagged H-NS protein and then resolved by

agarose gel electrophoresis. Our results show that again H-NS binds to the upstream intergenic region and not to the negative control or the non-specific competitor (Fig. 17b). The H-NS dependent shift of the *icsP'* fragment begins at 230 nM of protein with near complete retardation of the DNA probe by 920 nM. Furthermore the shift observed with the full upstream intergenic region, occurs at approximately the same concentration (230 nM) as the shift observed when some of the short DNA fragments were incubated with H-NS (Fig. 16). These data suggest that H-NS binds to the full length *icsP* upstream intergenic region as efficiently as it binds to some of the smaller DNA fragments used in our previous EMSA experiment, but not better. This indicates that H-NS binding to the entire intergenic region upstream of the *icsP* gene does not occur cooperatively, at least *in vitro*.

H-NS and VirB bind simultaneously to DNA located upstream of the *icsP* gene *in vitro*

Our current understanding of the mechanism of silencing/anti-silencing at the *icsP* promoter involves H-NS binding to regions of DNA with high intrinsic curvature and VirB antagonizing H-NS-DNA-H-NS bridges, allowing transcription to occur. This understanding is based on work conducted by others, which has examined the interaction of H-NS with other DNA binding proteins (De la Cruz *et al*., 2007, Hulbert & Taylor, 2002, Huo *et al*, 2009). What remains unclear about this mechanism of transcriptional regulation is whether VirB causes H-NS to dissociate from the DNA, or whether VirB simply displaces H-NS, resulting in both proteins binding simultaneously to the DNA. To shed light on the mechanism of transcriptional silencing and anti-silencing at the *icsP*

promoter, we chose to study whether VirB and H-NS interacts simultaneously with the DNA using EMSAs.

The full upstream intergenic region (*icsP'*; 1232 bp) and our negative control (1011 bp; *pstS'*) were first incubated with VirB alone (460 nM) or H-NS alone (1.8 μM). In the second part of the experiment, the *icsP'* DNA fragment and the negative control *pstS'* was pre-incubated with 1.8 μM of H-NS before the addition of increasing concentrations of VirB (120, 240 & 460 nM).

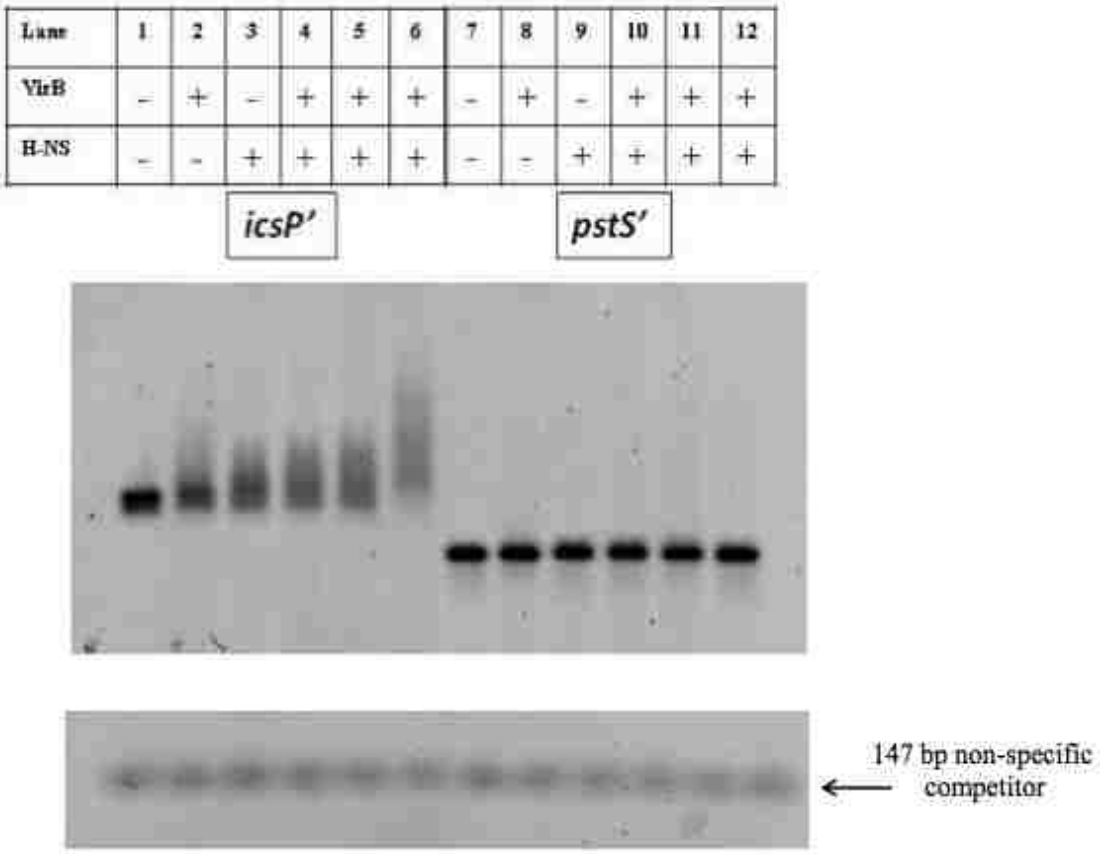

Fig. 18. VirB and H-NS binding of the *icsP* upstream intergenic region *in vitro*. EMSA results using the full length *icsP'* and *pstS'* are shown. Lanes 2 and 8 contain 460 nM purified VirB. Lanes 3 and 9 contain 1.8 μM purified H-NS. Lanes 4-6 and 10-12 were pre-loaded with 1.8 μM H-NS before increasing concentrations of VirB, 120 nM, 230 nM, and 460 nM, were added. The lower band is 200 nM of competitor DNA generated from pACYC184.

As predicted, when VirB or H-NS were incubated independently with the DNA fragments (Fig. 18; lanes 2 & 8, 3 & 9), the *icsP'* fragment alone was retarded in the gel matrix indicating the formation of specific nucleoprotein complex. This is consistent with our previous observations (Fig. 17a & b). Interestingly though, when DNA fragments were pre-loaded with H-NS and then incubated with 460 nM of VirB (Fig. 9, lanes 6 & 9), DNA-protein complexes formed on the *icsP'* fragment that migrate slower than complexes formed by each protein individually. These data suggest that under these experimental conditions, VirB and H-NS simultaneously bind to the *icsP'* DNA fragment. This could occur if VirB is displacing H-NS along the DNA or reorganizing the nucleoprotein complex, rather than causing the dissociation of H-NS from the DNA. Although it is likely that both proteins have remained associated with the DNA, future experiments will detect whether both protein species are present in the higher molecular weight complex by using mass spectroscopy, and immunodetection.

3.4 Discussion

Nucleoid structuring proteins play important roles in gene silencing in many bacterial species (Azam & Ishihama, 1999). Gene silencing in bacteria usually involves the formation of a nucleoprotein complex that renders the DNA inaccessible to DNA-binding proteins that are required for transcription (Azam & Ishihama, 1999, Dorman & Deighan, 2003, Yarmolinsky, 2000). Transcription factors that alleviate this silencing are central to bacterial physiology and regulate a variety of bacterial processes including virulence gene expression, biofilm formation, and bacterial adaptation to stress (Stoebel *et al*., 2009). The mechanisms of silencing and anti-silencing, however, are still being elucidated.

While it has been established that H-NS is a global repressor of horizontally acquired genes what remains elusive is a complete picture of how H-NS represses transcription and how it is alleviated by other transcription factors. In this work, we have focused on the regulation of the *icsP* gene encoded by the *Shigella* virulence plasmid. We describe regions necessary for the H-NS dependent repression of the *icsP* promoter and regions responsible for the binding of the global repressor H-NS and the major virulence factor VirB.

Our data show that H-NS represses *icsP* promoter activity and that VirB alleviates this repression (Castellanos *et al*., 2009). Through deletion analysis of the *icsP* upstream intergenic region and *in vivo* assays, we have identified that sequences upstream of -665 (with respect to the *icsP* annotated transcription start site [tss]) are needed for full repression of activity the promoter, although sequences between -665 and -351 appear sufficient for partial repression. Although it is unusual for H-NS to impart its effects from so far upstream of the transcription start site, our truncation analysis of the *icsP* upstream intergenic region has shed light on how this remote regulation might occur.

We observed that the removal of DNA sequences upstream of -351 resulted in an H-NS-dependent increase in promoter activity. The simplest interpretation of these data is that H-NS interacts directly with sequences downstream of -351 leading to an increase in the activity of the promoter. This is corroborated by our EMSA data, which shows specific binding of H-NS to sequences between –213 and +24 with relatively high affinity. Taken together, our data strongly suggests that H-NS binds to two discrete regions (upstream of -351 and downstream of -213) that lie upstream of the *icsP* gene and

that these two regions are separated by a region with little to no affinity for HNS (regions contained within T2 in Fig. 13).

H-NS has been demonstrated to oligomerize along DNA (Rimsky *et al*., 2001, Turner & Dorman, 2007), and form H-NS-DNA-H-NS bridges, which are described as looped DNA structures constrained by H-NS (Maurer *et al.,* 2009, Stoebel *et al*., 2009). Based on this work, it is tantalizing to imagine H-NS docked to upstream sites interacting with H-NS docked to downstream sites to mediate repression of the *icsP* promoter. Although, it is unusual for "classical" transcription factors to influence transcription from DNA binding sites located over 500 bp away, atomic force microscopy has revealed that H-NS: DNA: HNS bridges may involve long stretches of DNA, extending over 1 kb (Maurer *et al*., 2009), and these structures have been documented to play important roles in promoter silencing (Dame *et al*., 2000, Dame *et al*., 2005, Dame *et al*., 2006). Although we have no evidence to support the involvement of downstream sites in the H-NS mediated repression of the *icsP* promoter so far, this is currently being tested.

H-NS is well documented as displaying a binding preference for regions of DNA that are AT rich (Dame *et al*., 2005, Dame *et al*., 2006, De la Cruz *et al*., 2009, Dorman, 2009a, Dorman, 2009c, Porter & Dorman, 1994, Rimsky *et al*., 2001, Shin *et al*., 2005, Stoebel *et al*., 2008, Turner & Dorman, 2007, Williams *et al*., 1996, Williams & Rimsky, 1997, Zhang *et al*., 1996). These sequences typically display high levels of intrinsic curvature (Zuber *et al*., 1994). The work presented here demonstrates that regions upstream of the *icsP* gene, which bind H-NS, coincide with regions predicted to be intrinsically curved, as determined by the MUTACURVE server (http://132.248.32.45/cgi-bin/mutacurv.cgi). Although more recent work has revealed the

presence of a high affinity binding sequence for H-NS at the *E. coli proU* promoter (Bouffartigues *et al.* 2007, Lang *et al*., 2007) this 10 bp sequence was not found within the full-length fragment, which constitutes the intergenic region upstream of the *icsP* gene. This suggests that H-NS recognizes sequences other than those proposed by Bouffartigues *et al.* (2007) and Lang *et al.*, (2007). It is possible that future investigations, using DNase I footprinting, may identify other sequences with high affinity for H-NS, which would initiate H-NS binding and oligomerization along the DNA.

Our previous studies revealed that VirB functions to derepress the *icsP* promoter and that two VirB-binding sites organized as inverted repeat and located over 1 kb upstream of the annotated *icsP* transcription start site are required for VirB-mediated derepression (Castellanos *et al*., 2009). This work is corroborated by the β-galactosidase data presented here. Using EMSAs we demonstrate that VirB binds to the full upstream intergenic region of the *icsP* promoter. This is the first indication that VirB interacts directly with the *icsP* promoter, and we hypothesize that this binding is important for de-repression of the full-length promoter.

While our previous work strongly suggests that VirB interacts directly with the two VirB-binding sites located over 1 kb upstream of the *icsP* tss (Castellanos *et al*., 2009), our *in vivo* assays suggests that VirB may also interact with promoter proximal sequences. In these experiments VirB was found to antagonize H-NS dependent "activation" of the *icsP* promoter in the shortest construct tested. Although, this effect has the potential to be mediated by a VirB-regulated protein, the simplest interpretation of these results is that VirB interacts directly with sequence downstream of -351 to

mediate this effect. Four putative VirB binding sites have been identified in this region, although none of these are organized as inverted repeats. It remains unclear whether any of these, or other sites, mediates this VirB-dependent repression of the short fragments or, indeed, what they play in derepression of promoter in the context of the full upstream intergenic region. Future investigations will assess the role that these downstream putative VirB-binding sites play in the regulation of the promoter.

The results obtained with the shortest construct in our truncation series are interesting for another reason; in the context of this construct, both H-NS and VirB reverse their "typical" roles, H-NS leads to an increase in promoter activity and VirB dampens this effect. Although, it is well established that DNA binding proteins can serve dual roles, usually this phenomenon is dependent upon the location of the protein binding site (Corcoran & Dorman, 2009, Browning *et al*, 2000). Our data suggest, at least in the case of the i*csP* upstream intergenic region, the context each protein finds itself in with respect to the transcription start site and to the other protein, may determine whether the protein serves as an activator or repressor. These findings provide valuable insight into the plasticity of the molecular mechanisms that leads to transcriptional silencing and anti-silencing in bacteria.

Based on our findings, we are able to propose a model of transcriptional silencing and anti-silencing at the *icsP* promoter which involves the use of remotely located binding sites. At the non-permissive temperature of 30 ˚C, H-NS binds to upstream and downstream region in the intergenic region that lies upstream of the *icsP* promoter. This allows the formation of an H-NS-DNA-H-NS loops or bridges between promoter distal (upstream of -665) and promoter proximal sequences (downstream of -213). At the

permissive temperature of 37˚C, the temperature at which VirB is maximally expressed, VirB binds to the essential VirB-binding sites organized as an inverted repeat and located at position - 1144 with respect to the tss. Although, the distance over which VirB mediates its effect is unusual, if this protein functions to solely antagonize H-NS-mediated repression of the promoter, VirB may not need to bind to sites located in the vicinity to the *icsP* promoter in order to mediate transcriptional de-repression, eventhough putative binding sites have been found. Regardless of which sites are engaged by VirB, our EMSAs suggest that the interaction of VirB with the full upstream intergenic region does not cause H-NS to dissociate from the DNA. Instead, it seems more likely that VirB functions to reorganizes the nucleoprotein complex, which ultimately leads to the formation of a complex that is permissive for transcription to occur at the *icsP* promoter. This model will be used to frame our future studies of the *icsP* promoter region and transcriptional silencing and anti-silencing mechanisms.

In conclusion, this study raises several questions that are central to virulence gene expression in *Shigella* and other enteric pathogens and which lie at the heart of bacterial physiology. How does VirB function to alleviate H-NS dependent repression of the *icsP* promoter from such distances? Does VirB function from remote sites to regulate the expression of other *Shigella* virulence genes? Can transcription factors found in other bacteria function from remote sites to alleviate transcriptional repression by nucleoid structuring proteins? If so, how widespread is this phenomenon? We anticipate that these questions will be answered as transcriptional silencing and anti-silencing mechanisms in bacteria continue to be studied.

CHAPTER 4

VIRB AND H-NS BIND TO THE *ICSP* PROMOTER OF *SHIGELLA FLEXNERI* AND THE ROLE OF STATIC DNA CURVATURE IN THE REGULATION OF *ICSP*

Abstract

In the bacterial pathogen *Shigella flexneri,* many genes encoded by the large virulence plasmid are repressed by the nucleoid structuring protein H-NS and de-repressed by the major virulence gene activator VirB. One example is the *icsP* gene, which encodes an outer membrane protease. Using electrophoretic mobility shift assays (EMSAs), and DNase I footprinting analysis we show that VirB binds directly to the upstream intergenic region of *icsP* and demonstrate that VirB binds to DNA sequences located promoter distally and proximally. These data support a model where VirB bound to remote sites may act in concert with VirB bound to promoter proximal sequences to mediate VirB dependent de-repression of the *icsP* promoter. We show that the disruption of regions displaying high intrinsic curvature do not have an effect on either the ability of H-NS to repress promoter activity or VirB to alleviate this repression, at least in the context of the full upstream intergenic region. Our findings enhance our understanding of the mechanism of transcriptional silencing and anti-silencing that control virulence gene expression in *Shigella* and raise the possibility that other bacterial promoters may be regulated from remote transcription factor binding sites.

Authors: Dustin J Harrison and Helen J Wing

We thank Dr. Eric Anderson and Dr. Martin Roop from East Carolina University for all their help and guidance with the EMSAs, and Stephanie Labahn for technical support.

4.1 Introduction

Shigella species are gram-negative intracellular pathogens that cause bacillary dysentery in humans. *Shigella* infections are responsible for approximately 164 million infections each year resulting in 1.1 million deaths, most of these occur in developing countries and in children less than 5 years old (Kotloff *et al*., 1999, Lee *et al*., 2005, Li *et al*., 2009). All four *Shigella* species carry a large (230 kbp) virulence plasmid. Many genes encoded by these plasmids are under the transcriptional control of the histone-like nucleoid structuring protein (H-NS) and the major virulence gene activator VirB (Adler *et al*., 1989, Berlutti *et al*., 1998).

H-NS is a nucleoid structuring protein that is well characterized as a silencer of bacterial transcription (Ali Azam *et al*., 1999, Dorman, 2004). H-NS has been shown to oligomerize on DNA (Rimsky *et al*., 2001, Turner & Dorman, 2007) and this oligomerization can lead to the formation of H-NS-DNA-H-NS bridges, which have been visualized using atomic force microscopy as looped DNA structures constrained by H-NS (Maurer *et al.,* 2009). These structures may involve long stretches of DNA, extending over 1 kb (Dame *et al*., 2000, Dame *et al*., 2005, Dame *et al*., 2006, Maurer *et al*., 2009), and have been documented as playing important roles in promoter silencing (Bouffartigues *et al*., 2007, De la Cruz *et al*., 2009, Dorman, 2004, Dorman, 2009a, Fang & Rimsky, 2008, Navarre *et al*., 2006, Stoebel *et al*., 2008). Disruption of these H-NS:DNA complexes by other DNA-binding proteins can allow transcriptional de-repression; a process coined anti-silencing (Fang & Rimsky, 2008, Navarre *et al*., 2007, Stoebel *et al*., 2008).

The virulence gene activator VirB is encoded by the *Shigella* virulence plasmid and its expression is regulated by another transcription factor VirF (Adler *et al*., 1989). At the non-permissive temperature of 30 ˚C many genes encoded by the virulence plasmid are repressed by H-NS (Beloin & Dorman, 2003, Hromockyj *et al*., 1992, Porter & Dorman, 1994). Upon a switch to 37 °C, a change in DNA superhelicity causes H-NS to dissociate from the *virF* promoter leading to the production of the virulence gene regulator, VirF (Adler *et al*., 1989). VirF relieves H-NS-dependent repression of the v*irB* promoter (Tobe *et al*., 1993, Watanabe *et al*., 1990) and subsequently VirB alleviates H-NS dependent repression of many virulence plasmid genes, including *icsA* and those encoding the type III secretion system that mediate host cell invasion; *ipa*, *mxi* and *spa* (Beloin & Dorman, 2003, Berlutti *et al*., 1998, Dorman, 2006, Hromockyj *et al*., 1992, Maurelli *et al*., 1984a, Maurelli *et al*., 1984b, Nicoletti *et al*., 2008, Porter & Dorman, 1994, Porter *et al*., 1998, Prosseda *et al*., 1998, Sasakawa *et al*., 1993, Tobe *et al*., 1991, Tobe *et al*., 1993, Tobe *et al*., 1995). Transcriptional de-repression (anti-silencing) of virulence genes by VirF and VirB is central to the pathogenicity of *Shigella.*

The mechanism underlying transcriptional silencing/anti-silencing has been proposed to be a straightforward molecular antagonism involving competition between the silencing protein and the anti-silencing protein for their overlapping cognate DNA binding sequences, which are usually located in the immediate vicinity of the promoter. Typically this competition results in the dissociation of the silencing molecule from the DNA (Fang & Rimsky, 2008, Navarre *et al*., 2007). In contrast, at some bacterial promoters the anti-silencing proteins function to alter the DNA topology in a way that no longer allows H-NS to constrain transcription. For example, LeuO counters silencing of

the quiescent porin *ompS1* in *Salmonella* by altering the topology of the *ompS1* promoter region (De la Cruz *et al*., 2007, De la Cruz *et al*., 2009) and in *Shigella flexneri,* VirB antagonizes H-NS by affecting the structure of the *icsB* promoter (Turner & Dorman, 2007). Previous analysis of the *icsB*, *spa15* and *virA* promoters of *Shigella sonnei*, and *icsB* from *S. flexneri,* established a consensus binding site for VirB. Our analysis of sequences upstream of the *icsP* gene identified nine sites with greater than a 6/7 match to the consensus, 5'-(A/G)(A/T)G(G)AAAT-3' (Fig. 12B; Taniya *et al*, 2003, Turner & Dorman, 2007).

Genome-wide studies have revealed that intrinsically curved DNA in the promoter regions of a large proportion of many important bacteria may constitute one of the most conserved and primitive forms of transcriptional regulation before the co-evolution of trans-acting factors that recognize bent DNA, e.g. DNA binding proteins (Kozobay-Avraham *et al*, 2004, Prosseda *et al*, 2010). Static curvature of DNA has been shown to activate transcription, by facilitating RNA polymerase (RNAP) binding to promoters, as well as the nucleoid proteins IHF, and Fis which both recognize curved DNA and allow enhanced transcription initiation (Prosseda *et al*, 2004, Prosseda *et al.,* 2010, Falconi *et al*, 1998, Falconi *et al*, 2001, Olivares-Zavaleta *et al*, 2006, Dorman & Deighan, 2003, Luijsterburg *et al*, 2006). On the other hand, the static curvature of DNA can play an indirect role in the transcriptional repression allowing silencers like H-NS to engage DNA to stabilize or enhance a preexisting DNA loop that blocks transcription (Prosseda *et al.,* 2004, Falconi *et al.,* 1998). Therefore, curvature is an important determinant in the regulation of promoter activity. Many related bacterial species including *E. coli*, *Salmonella*, and *Yersinia* have global transcription factors, as well as specific regulators,

that regulate promoters with curved DNA, and we anticipate that the *Shigella icsP* promoter is no exception. In Prosseda *et al.* (2004), De la Cruz *et al.* (2009), and Falconi *et al.* (1998) disruption of a region of curved DNA had significant effects on the ability of transcriptional regulators to exert their effects highlighting the importance that curved sequences have on transcriptional regulation.

Our previous work shows that the *icsP* gene, located on the *Shigella* virulence plasmid, is repressed by H-NS and de-repressed by VirB. We have demonstrated that the role of VirB is to solely alleviate transcriptional repression of the *icsP* promoter mediated by H-NS from sites over 1 kb upstream of the annotated transcription start site (Harrison *et al.,* under revision, Castellanos *et al*., 2009, Wing *et al*., 2004). Surprisingly, two putative VirB binding sites, organized as a perfect inverted repeat, are found in the extreme distal region of the *icsP* upstream intergenic region and are essential for the VirB dependent regulation of *icsP in vivo* (Castellanos *et al.,* 2009). Although our previous work strongly suggests VirB binds directly to this upstream intergenic region, this has never been categorically demonstrated.

In this study, we examine further the role that VirB plays in the regulation of the *icsP* promoter in an attempt to discern the molecular mechanism responsible, as well as examining the effect that regions of intrinsic curvature have on promoter activity in the presence and absence of both H-NS and VirB. Our results suggest that VirB regulates the activity of the *icsP* promoter by binding directly to the upstream intergenic region of *icsP*. Our data also suggest that disruption of a single peak of intrinsic curvature in the context of the full upstream intergenic region does not have an effect on the H-NS mediated repression, or the VirB dependent de-repression of the promoter. We propose a

model to explain how silencing and anti-silencing of the *icsP* promoter is achieved by DNA binding-proteins binding to remote sequence elements. An improved understanding of the H-NS and VirB dependent regulation of the *icsP* promoter will provide insight into bacterial regulation of transcription from remote sites.

4.2 Materials and Methods

Bacterial strains, plasmids and media.

The bacterial strains and plasmids used in the present study are listed in Appendix 1. Bacteria were grown routinely at 37°C in Luria-Bertani (LB) broth with aeration or on LB agar (LB broth containing 1.5% [wt/vol] agar). Antibiotics were added at the following final concentrations: ampicillin, 100 μg/ml; chloramphenicol, 25 μg/ml; kanamycin, 50 μg/ml; and tetracycline, 12.5 μg/ml. Where appropriate, to ensure that *Shigella* strains had maintained the large virulence plasmid during manipulation, Congo red binding was tested on Trypticase soy broth agar plates containing 0.01% (wt/vol) Congo red (Sigma Chemical Co., St. Louis, Mo.).

Construction of the P*icsP-lacZ* reporter plasmid and site directed mutagenesis of P*icsP*.

The P*icsP-lacZ* reporter plasmid pHJW20 was used as the basis for this work (Castellanos *et al*., 2009, Wing *et al*., 2004; Appendix 1). pHJW20 carries 1232 bp upstream of the annotated transcription start site of the *icsP* promoter, the first 48 bp of the *icsP* coding region cloned upstream of a translation stop site, and a promoterless *lacZ* gene, so that expression of *lacZ* is directly regulated by the *icsP* promoter.

Site-directed mutagenesis was performed using the Quick Change Lightning II kit (Stratagene) according to manufacturer instructions. Complimentary oligonucleotides

containing the mutations predicted by MUTACURVE were used to disrupt/restore the curvature of the *icsP* promoter in the template plasmid pNEO1 (Appendix 1). Following mutation, the 1232 bp insert was digested using *Pst*I, *Pac*I and cloned into pDH01 creating 4 plasmids designed to disrupt (D) curvature and 4 plasmids designed to restore (R) curvature around the predicted peaks of highest intrinsic curvature: pMut1 D/R (-842), pMut 2 D/R (-935), pMut 3 D/R (-663), pMut 4 D/R (-435) (Appendix 1; Table 4 and 5). All promoters have been sequenced. All primers and plasmids are described in Appendix 1.

Quantification of *icsP* promoter activity using the P*icsP-lacZ* reporter and derivatives.

Activity of the *icsP* promoter was determined by measuring β-galactosidase activity (Wing *et al*., 2004) by using the Miller protocol (Miller, 1972) in strains carrying pHJW20 or derivatives. Freshly transformed cells were back-diluted 1:100 and grown for 4 to 5 h in either TSB medium (*S. flexneri*) or LB (*E. coli*) at 37°C prior to cell lysis because *icsP* expression significantly increases under these conditions (Wing *et al.,* 2004). Assays were performed in triplicate on three separate occasions.

***In silico* analyses of the *icsP* promoter and upstream sequences.**

Prediction of intrinsic DNA curvature was done using MUTACURVE (http://132.248.32.45/cgi-bin/mutacurv.cgi). The MUTACURVE program evaluates the amplitude of intrinsic curvature for every nucleotide in a given sequence using the algorithm of Goodsell & Dickerson (1994), with the addition of Satchwell's contribution matrices for rotational and spatial displacements (Satchwell *et al*., 1986). These algorithms have been routinely used by other investigators to identify putative H-NS

binding sites in DNA (Beloin & Dorman, 2003, De la Cruz *et al*., 2009, Flores-Valdez *et al*., 2003, Will & Frost, 2006).

Construction of inducible plasmids and purification of the VirB and H-NS proteins.

The *virB* gene was digested from pATM324 using *Hin*dIII and *Nco*I restriction enzymes and gel purified. The primers W38 and W39 were used to amplify the *virB* gene, which was then ligated into pQE-60 (Qiagen) before electroporation into *E. coli* M15 pREP4 creating a C-terminal His-tagged VirB in an inducible plasmid, pAJH01. A similar strategy was used to create a C-terminal His-tagged H-NS in an inducible plasmid, pCTH01. Briefly, the primers W134 and W137 (Appendix 1) were used to amplify the *hns* gene from *Shigella flexneri* serotype 2a with *Nco*I and *Bgl*II sites for cloning into pQE-60.

His-tagged-VirB and His-tagged-H-NS proteins were overproduced from plasmid pAJH01, and pCTH01 respectively. Proteins were expressed in the *E. coli* strain M15 carrying the plasmid pREP4. The expression of C-terminally His-tagged proteins was induced in 500 ml cultures growing exponentially with 1mM IPTG (isopropyl-β-thiogalactopyranoside). Two hours post-induction, the cells were harvested and frozen at 80 °C overnight. The cell pellet was thawed on ice and resuspended in lysis buffer (40 mM Tris-base [pH 8.93], and 80 mM NaCl). Cells were lysed by sonication, and cellular debris was pelleted by centrifugation at 10 000 x g at 4 °C. Cell lysates were applied to Ni-NTA columns (Qiagen) pre-equilibrated with equilibration buffer (40 mM Tris-base [pH 8.93], 80 mM NaCl, 10 mM imidazole). The columns were then washed with 10 bed volumes of wash buffer (equilibration buffer with 20 mM imidazole) prior to elution of the proteins by the addition of 2.5 bed volume of elution buffer (equilibration buffer with

250 mM imidazole). Eluates were analyzed by SDS-PAGE followed by Coomassie staining. Purified proteins were dialyzed in equilibration buffer overnight at room temperature. Protein concentrations were determined using Bradford Assays. The hexa-his tag was not found interfere with VirB expression or activity, because His-tagged VirB was observed to restore IcsP expression to wild type levels in a *strain* lacking *virB in vivo* (data not shown). Furthermore, previous studies using an identical His-tag H-NS fusion protein, produced in a manner similar to that described above, was shown to retain normal function of H-NS in assays (Williams & Rimsky, 1997).

Electrophoretic Mobility Shift Assays (EMSAs).

The upstream intergenic region of the *icsP* promoter (1232 bp) was used as a template to construct six, overlapping smaller fragments, ranging from 252 bp to 358 bp by PCR (Appendix 1). A positive control fragment, *icsB*, was made by PCR using the *Shigella flexneri* virulence plasmid as a template and the negative control, an internal fragment of *pstS*, was made using *E. coli* genomic DNA as the template (Appendix 1). As a non-specific competitor, a 147 bp fragment of DNA from pACYC184 was generated by PCR. The PCR probe fragments were incubated with increasing concentrations of purified protein, VirB (0 to 920 nM) for 30 minutes at 30°C in a 20 µl reaction mixture containing 40 mM Tris-base (pH 8.93), 80 mM NaCl, 1 mM EDTA, 1 mM DTT, 200 ng/ml bovine serum albumin (BSA), 200 nM of competitor p184 DNA and 10% glycerol. Each reaction mixture contained approximately 100 nM of DNA, measured using the NanoView (GE). The protein-DNA complexes were resolved by electrophoresis in 1.5% TBE agarose gels for 4 hours at room temperature at 60 V constant. Following electrophoresis the agarose gels were stained with ethidium bromide (1 µg/ml), and

visualized using a Typhoon 9410 (Amersham) variable mode imager. EMSAs were performed a minimum of three times with representative results shown.

DNase I footprinting.

Fragments of the *icsP* promoter were amplified by PCR (Appendix 1) before being restriction digested using *SalI* and *EcoRI,* and ligated into a pBluescript (Stratagene) vector. The ligated plasmids were transformed by electroporation into *E. coli* DH10B cells and plated to media containing ampicillin (100 μg/ml). Sequence analysis confirmed the presence of the correct *icsP* target sequence. Overnight cultures of DH10B with the holding vector plasmids were grown overnight and back diluted 1/100 into 500 ml cultures of LB containing ampicillin. The cells were harvested by centrifugation and plasmids purified by Qiagen Mega-prep (Qiagen). Aliquots of Mega-preps were then digested with *SalI*, or *EcoR1*, followed by calf intestinal alkaline phosphate (CIAP) digestion to remove the terminal phosphate. Phenol:chloroform extraction removed the CIAP followed by ethanol precipitation. Plasmids were next digested with *SalI*, or *EcoR1*, before being purified by native acrylamide gel electrophoresis. Excised bands were extracted from the acrylamide and a PCR clean-up kit was used (Qiagen) before T4 kinase end labeling with [γ^{32}P]ATP. End labeled products were cleaned using illustra ProbeQuant G-50 Micro Columns (GE Healthcare) to remove unincorporated radionucleotides.

Increasing amounts of VirB or H-NS protein were incubated for 30 min at 30° C with a labeled probe (~200 nM of DNA) in binding buffer (40 mM Tris-base [pH 8.93], 80 mM NaCl, 1 mM EDTA, 1 mM DTT, 200 ng/ml bovine serum albumin [BSA], 25 μg/ml sonicated Herring sperm DNA, 200 nM of competitor p184 DNA, and 10%

glycerol) with protein buffer (40 mM Tris-base [pH 8.93], 80 mM NaCl, 1 mM EDTA, 1 mM DTT, and 10% glycerol) to a final volume of 20 µl. DNase I diluted 1:100 in 4 mM Tris-base (pH 8.93), 8 mM NaCl, 0.5 mM $CaCl_2$, and 40 mM Mg Cl_2, was added to the reaction mixtures at 30° C and incubated for 40 seconds. The addition of stop solution (0.3M sodium acetate, and 5 mM EDTA) stopped the reactions. DNA was extracted with 200 µl of phenol:chloroform followed by ethanol precipitation. DNA was electrophoresed through a 6% polyacrylamide sequencing gel run at 60 Watts constant until the bromophenol blue dye front migrated out of the gel. Following completion of the electrophoresis the gel was dried and exposed to a phosphor-imaging screen and viewed on a Typhoon 9410 variable mode imager. Protected bands were identified by comparison to a DNA sequencing ladder generated using the method of Maxam and Gilbert (Maxam & Gilbert, 1986). Experiments were conducted a minimum of three times with representative images shown.

4.3 Results and Discussion

VirB binds the upstream intergenic region of *icsP*.

Our previous work has demonstrated the VirB dependent regulation of *icsP* and that VirB functions solely to alleviate the H-NS dependent repression of promoter activity (Castellanos *et al*., 2009, Harrison *et al.,* under revision). Harrison *et al.,* (under revision; CH 3) showed VirB binds to the full upstream intergenic region upstream of the *icsP* gene. However, to determine which sequences are required for VirB to bind, we again used electrophoretic mobility shift assays (EMSAs). The full upstream intergenic region that lies upstream of the *icsP* gene (1232 bp) was divided up into six near equal length

fragments and each of these fragments were used in EMSAs with purified His-tagged VirB (Fig. 19).

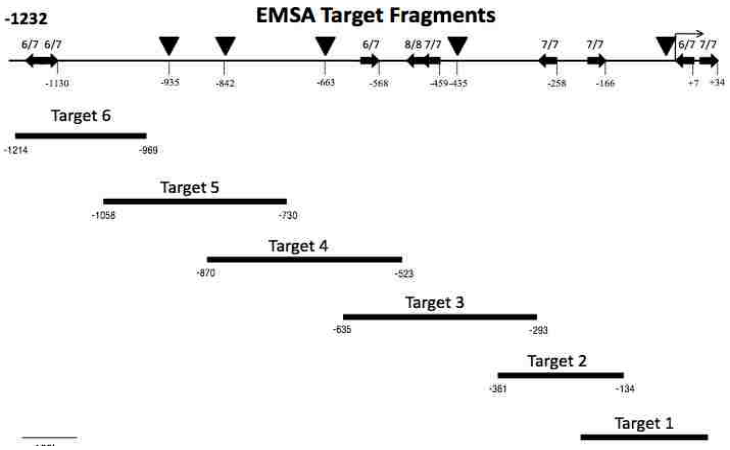

Fig. 19. From Harrison, *et al*., under revision, CH 3, schematic showing the 6 target fragments used in the EMSAs. Coordinates of each target fragment are given relative to the transcription start site. Inverted triangles denote the position of a peak of high intrinsic curvature predicted using the MUTACURVE program.

For these experiments, a DNA fragment containing sequences that lie between the divergent *icsB* and *ipgD* genes of *Shigella flexneri* was used as a positive control (*icsB*), because VirB has been shown to bind to this promoter region by others (Turner & Dorman, 2007) and an internal fragment of the *E .coli pstS* gene was used as a negative control (*pstS*). Each DNA fragment (100 nM) was incubated with increasing concentrations of purified His-tagged VirB in the presence of a 147 bp fragment, which had been amplified from the plasmid pACYC184 and which served as non-specific competitor DNA (200 nM). DNA and DNA:protein complexes were then resolved by agarose gel electrophoresis.

Previous *in vitro* studies with VirB and VirB-regulated promoters have shown that the DNA band in EMSAs performed with VirB often do not appear as discrete shifts but are more diffuse which has been proposed to be evidence of protein oligomerization (Turner & Dorman, 2007). Despite this, we anticipated that VirB would interact with with some of the DNA fragments at lower concentrations than others, in particular T6 which contains putative VirB binding sites that were absolutely required for the VirB dependent

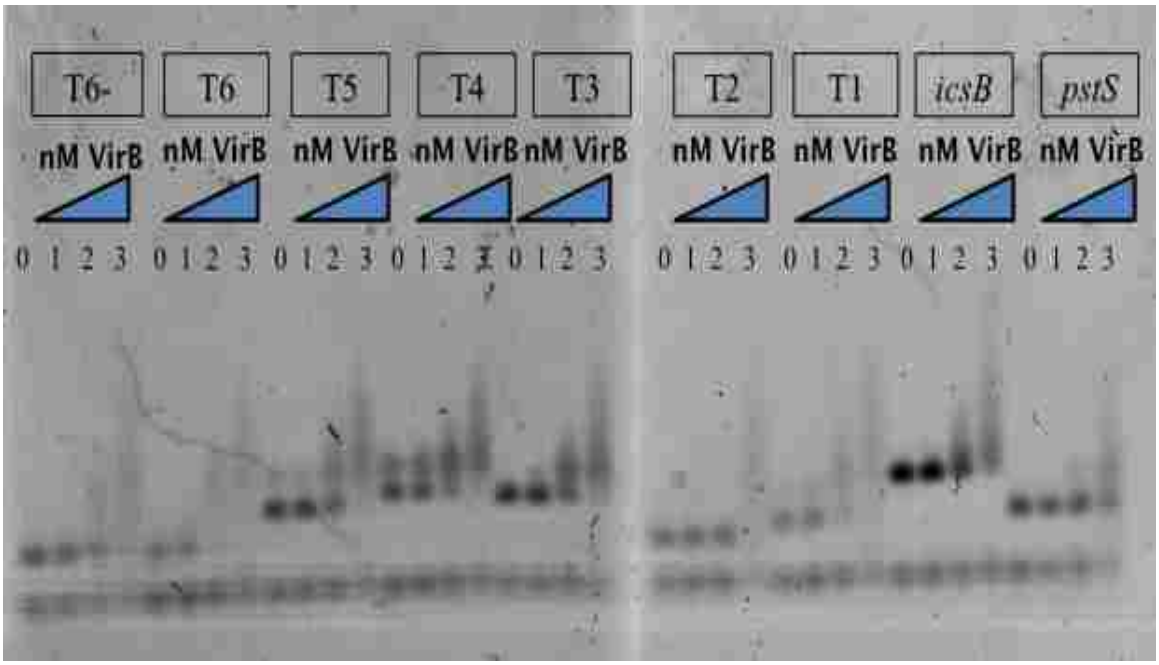

Fig. 20. VirB binding within the *icsP* upstream intergenic region. EMSAs were conducted using 100 nM of each fragment shown in Fig. , as well as T6 - (mutated VirB binding site, from Castellanos *et al*., 2009) and increasing concentrations of purified His-tagged VirB. The concentrations of purified protein as well as the target fragment used are given above each lane. In all cases, the lower band is 200 nM of competitor DNA generated from pACYC184. Positive control, and negative controls are labeled *icsB* and *pstS* respectively. 0=no protein, 1 = 120 nM, 2 =230, 3=460 nM protein.

de-repression of *icsP* promoter activity (Castellanos *et al.,* 2009). Included in this assay is T6- which is identical to T6 except that it contains the mutated VirB binding sites from Castellanos *et al.,* (2009) which were shown to eliminate the VirB dependent activity of the promoter. Our results show that VirB binds to target fragments T6 and T1 beginning at 120 nM of VirB with complete sequestration of free DNA at 230 nM. VirB also binds T6-, T5, T4, and T3 but at a higher concentration of 230 nM VirB. In contrast, T2 only shifts at 460 nM, the highest concentration of VirB used, but it should be noted that even VirB binds to the *pstS* negative control at this concentration. In all but this case, VirB binding to the *icsP* promoter appears to be specific as evidenced by the uniform presence of free non-specific competitor DNA in all lanes and does not shift.

The T6- fragment, which contains the mutated essential VirB binding sites identified in our *in vivo* studies (Table 1), shows a defect in the ability to bind VirB when compared to the wild type T6 (the shift with T6 starts at 120 nM of VirB, while T6- shows a moderate shift with 230 nM VirB). These data suggest VirB may have some affinity to other sequences contained within this fragment, at least *in vitro* and under the conditions used. Nevertheless, these results support our hypothesis that VirB interacts directly with the upstream intergenic region of *icsP* and supports our observation that sequences located over 1 kb upstream of the annotated transcription start site are required for the VirB dependent de-repression of the *icsP* promoter.

Target T1, which contains sequences located between 213 and +24 relative to the transcription start site, also shifted with low concentrations of protein. The binding of VirB to this region of DNA may explain the VirB dependent *decrease* in *icsP* promoter activity observed at the shortest promoter construct in our *in vivo* assays. The binding of VirB to sequences contained within this region may have disrupted an H-NS induced structure that increased promoter activity (Harrison *et al.,* under revision; Fig. 14). Putative VirB binding sites have been found in this region, and these EMSA data, combined with the observations made *in vivo*, suggest VirB binds to these binding sites to antagonize H-NS *in vivo*, although this still needs to be tested. Interestingly, VirB binds to similar fragments of the *icsP* upstream intergenic region as H-NS (Harrison *et al*., under revision).

VirB binds sequences within the intergenic region upstream of the annotated *icsP* transcription start site.

DNase I footprints were used to test the hypothesis that VirB binds to the upstream intergenic region of *icsP*. The same 6 fragments of the *icsP* upstream intergenic region used in the EMSA were incubated with increasing concentrations of purified VirB (250 nM, 500 nM, and 1 μM). We predicted that VirB would clearly bind the DNA fragments identified in the EMSAs, in particular T6, and T1. Consistent with our previous *in vitro* experiments, (Fig. 20), VirB bound to all 6 targets (Fig. 21), however, protection was observed with the lowest concentration of VirB (250 nM, lane 1) in targets T1, T5 and T6 suggesting that VirB binds more efficiently to sequences contained in these fragments (results are summarized in Table 2).

Target T1 shows VirB-dependent protection at a concentration of 500 nM VirB, before the complete protection of DNA at 1 μM. Interestingly, no hypersensitive bands were observed when T1 was incubated with VirB, which would be indicative of a discrete DNA-protein interaction. Instead large regions of DNase I protection are observed in T1, strongly suggesting VirB oligomerizes along the DNA, preventing DNase I access. VirB was observed to bind to sequences in T1 between +10 and -33 with respect to the tss. Contained within this region are two sequences that display good matches to the consensus VirB binding site 5'-(A/G)(A/T)G(G)AAAT-3' (Turner & Dorman, 2007, Taniya *et al.,* 2003). One surrounds the transcription start site (+1) and the other lies between -15 to -21 (although in this case the matching sequence lies in the opposite orientation (3'-5') to the reported consensus).

Table 2. Summary of VirB DNase I footprints

Target	Result
T1	VirB bound at 500 nM; full protection at 1 μM; no hypersensitive sites
T2	Bands fade with 500 nM VirB, but not much more with 1 μM; no hypersensitive sites; VirB bound to 3 regions (-201 to -274; -296 to -337; and upstream of -347)
T3	VirB binds 2 regions; intensity of upper region remains same while lower region changes between 250 μM and 500 μM
T4	Small regions of protection with VirB; no hypersensitive sites
T5	Protection of bands lower in the gel at 250 nM; top arrow fades out (*-829), while bottom arrow fades in (-896; hypersensitive) indicating a protein induced structural distortion
T6	Footprint changes at 1 μM VirB compared to 250 nM and 500 nM, with bands shifting up and fading out

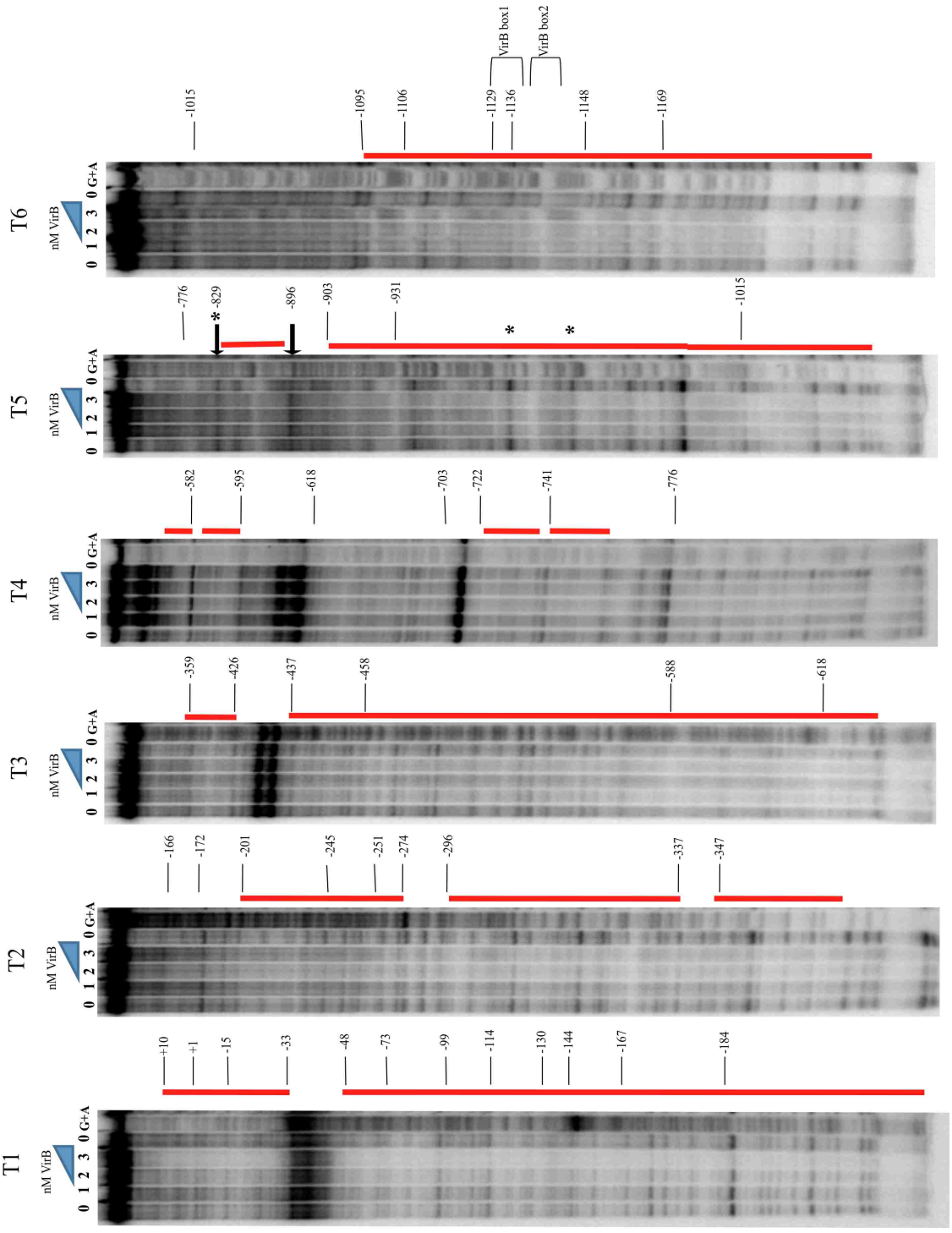

Fig. 21. VirB binds the intergenic region of *icsP*. DNase I footprints were conducted using approx. 200 nM of each fragment shown in Fig.19, and increasing concentrations of purified His-tagged VirB. The concentrations of purified protein as well as the target fragment used are given above each lane. A Maxam-Gilbert G+A ladder is included as well as position of relevant bp given relative to the tss. The most distal (-1130 to -1145) VirB binding sites are given labeled Box 2, Box 1. Red bars indicate regions of protection. * = bands protected by increasing protein concentration. Black arrow -896 indicates hypercleavable site. 1=250 nM VirB; 2=500 nM VirB; 3=1 µM VirB.

Our data strongly suggest that these sites are true VirB binding sites, although site directed mutagenesis of these sites will be needed to confirm this.

A second long region of protection can be observed extending upstream of position -48. This is indicative of oligomerization. Strikingly, the protection observed in T1 is more pronounced compared to the other targets, which could help explain why this target shifts so well with low concentrations of VirB in our EMSAs (Fig. 20).

In target T5, VirB bound at low concentrations (250 nM) to sequences -829 to -896, and to sequences upstream of -903 (sites marked with * indicate specific bases of interest as being protected by VirB at low concentrations). Increasing concentrations of VirB resulted in a single hypercleavable site, which indicative of a protein induced structural distortion (-896). This is consistent with the wrapping of the DNA around the protein, leading to increased sensitivity of the DNA to the DNase I enzyme.

In target T6, VirB bound at the low concentration of 250 nM. Protection of sequences upstream of -1095 can be visualized. Contained within T6 are the two most distal VirB binding sites (-1130 to -1145), which we have shown to be essential in the regulation of *icsP* (Castellanos *et al.,* 2009). The footprint in T6 is unusual in that, at high concentrations of VirB (1μM) it is possible to see protein induced distortion of the DNA, bands are shifted up as well as fading out. The exact significance of this finding is unknown, but the unusual appearance of this footprint suggests VirB interacted with sequences contained within this fragment in a manner that is distinct compared to the other target fragments, this phenomenon was observed regularly.

These data suggest that VirB binds directly to the upstream intergenic region of *icsP* (both promoter distally and promoter proximally), and that VirB oligomerizes along the

DNA, which is consistent with previous finding regarding VirB binding (Turner & Dorman, 2007). Our findings demonstrate that VirB bound to target fragments that were predicted to contain VirB binding sites, as well as sequences that do not contain any matches to the previously described VirB binding consensus sequence (e.g. T5). This raises the possibility that VirB binds to sequences other than its consensus site, at least *in vitro*, and under these conditions. Interestingly, but not surprisingly, the regions displaying VirB binding overlap with those shown to be protected by H-NS (Fig. 22). Overlapping, or similar, protein binding regions is a common feature among silencing and anti-silencing proteins (Navarre *et al.,* 2006). Taken together these data support our model for the regulation of *icsP* involving the interaction of VirB with sequences located both, promoter distally and promoter proximally, which likely displaces H-NS, resulting in a permissive structure for transcription, possibly a loop structure. DNA looping has been established as a mechanism of regulation at promoters where interactions involving upstream and downstream regions bound by protein have been observed, and it is likely that *icsP* is no different (De la Cruz *et al.,* 2009, Prosseda *et al.,* 2004). It should be remembered that these *in vitro* data, with short, linear fragments of DNA, may or may not adequately reflect the action of VirB *in vivo*.

H-NS binds sequences within the intergenic region upstream of the annotated *icsP* transcription start site.

We also predicted that H-NS would bind the regions identified in Harrison *et al.,* (under revision), by EMSAs, in particular, to sites located within targets T5, T4 and T1. Our results demonstrate H-NS bound to all 6 targets (Fig. 22); however, protection was observed with the lowest concentration of H-NS (250 nM, lane 1) in targets T1, T4, T5, and T6 (results are summarized in Table 3).

T1 shows H-NS-dependent protection within 2 regions at a concentration of 250 nM H-NS, before the complete protection of DNA at 1 μM. H-NS bound to sequences in T1 between +10 upstream of the transcription start site and -33 bp downstream of the tss before complete oligomerization and protection of the entire region upstream of -48.

Table 3. Summary of H-NS DNase I footprints

Target	Results
T1	H-NS bound at low concentration (250 nM); no hypersensitive sites
T2	H-NS binds 4 regions; Bands fade at 500 nM H-NS
T3	Bands fade at low concentration; bands fade faster in bottom part of gel
T4	H-NS binds 5 regions at low concentration (250 nM); 1 hypersensitive site upstream of -618 indicating a protein induced structural distortion
T5	H-NS binds at low concentration (250 nM); 2 hypersensitive sites (-829 and -896) indicating a protein induced structural distortion
T6	Protected bands fade out at the bottom of the gel; distortion of protected bands similar to VirB

In T4, H-NS bound to 5 regions within the fragment at low concentrations, corroborating our EMSA data (Fig. 16). There is the presence of a hypercleavable site upstream of -618 indicating a protein induced structural distortion consistent with the wrapping of the DNA around the protein. Two hypercleavable sites are also present in T5.

In target T5, H-NS bound at the low concentration of 250 nM to a region upstream of -903. Surprisingly, the sites (-829 and -896; black arrows in Fig. 21) which were observed in T5 using VirB are again present with the H-NS protein (however with H-NS both are hypercleavable) indicating a protein induced structural distortion leading to increased sensitivity of the DNA to DNase I.

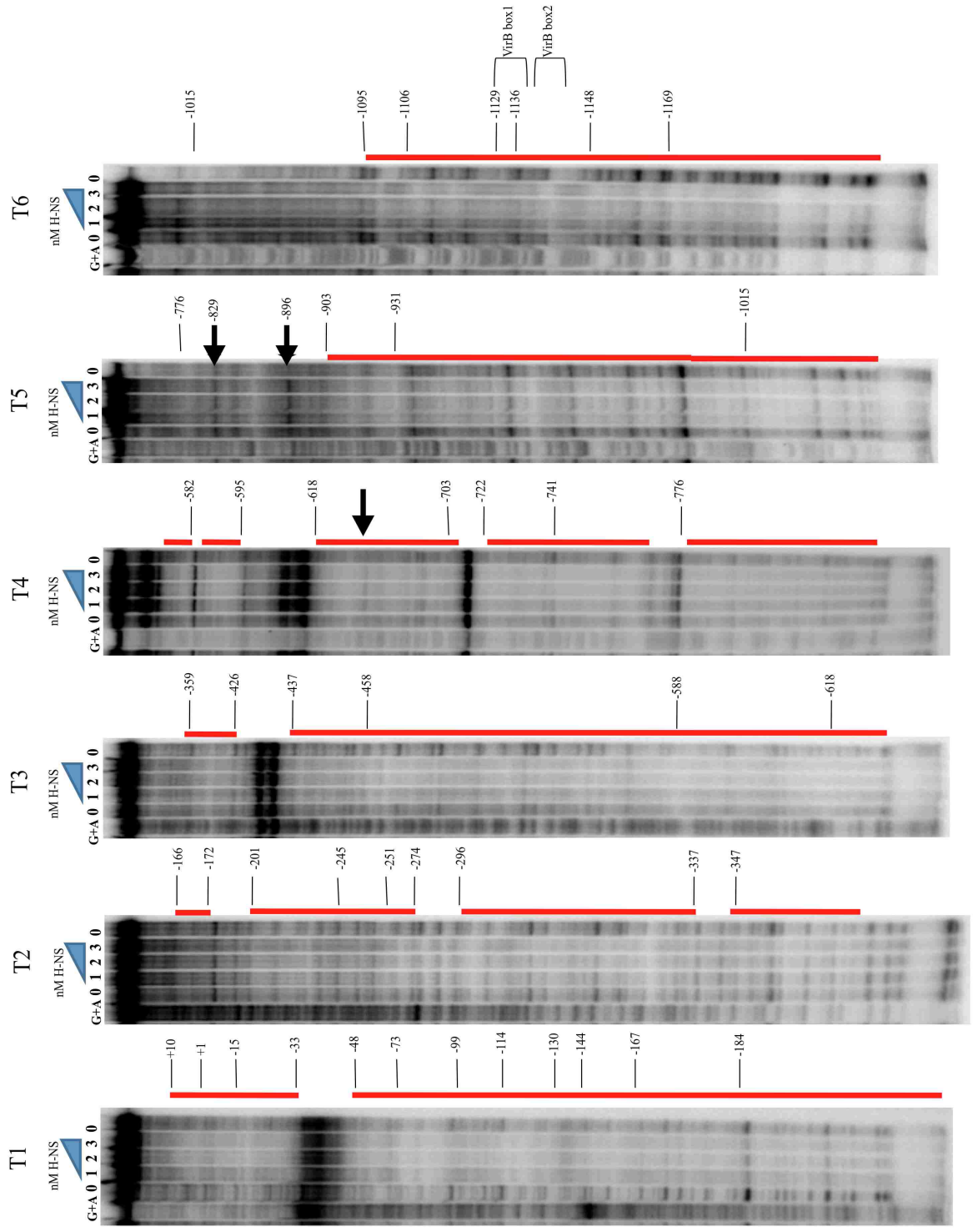

Fig. 22. H-NS binds the upstream intergenic region of *icsP*. DNase I footprints were conducted using approx. 200 nM of each fragment shown in Fig.19, and increasing concentrations of purified His-tagged H-NS. The concentrations of purified protein as well as the target fragment used are given above each lane. A Maxam-Gilbert G+A ladder is included as well as position of relevant bp given relative to the transcription start site. The most distal VirB binding sites (-1130 to -1145) are given labeled Box 2, Box 1. Red bars indicate regions of protection. Black arrows indicate hypercleavable sites. 1=250 nM H-NS; 2=500 nM H-NS; 3=1 µM H-NS.

In target T6 H-NS bound to the same region as did VirB. Again, at high protein concentration (1 μM) a distortion of the DNA is visible, similar to VirB. Despite the striking similarities between VirB and H-NS protection signals in all the targets examined, there are many subtle differences in the footprints with both proteins, e.g. extended region of binding by H-NS in T4 compared to VirB.

In summary, these results demonstrate that H-NS binds directly to the upstream intergenic region of *icsP* and demonstrates the ability of H-NS to oligomerize along DNA, which is consistent with previous findings regarding H-NS function (Bouffartigues *et al*., 2007, De la Cruz *et al*., 2009, Dorman, 2004, Dorman, 2009a, Fang & Rimsky, 2008, Navarre *et al*., 2006, Stoebel *et al*., 2008). The hypercleavable sites in T4 and T5 suggest that H-NS binding produces a conformational change in the DNA that is responsible for the H-NS induced repression. Our data also show that H-NS binds to promoter proximal sequences and further supports our model for the regulation of *icsP* involving DNA looping and interaction with promoter distal and promoter proximal sequences being bound by H-NS (Chapter 3, Discussion). Again H-NS bound to sequences that overlap regions occupied by VirB highlighting the molecular antagonism that exists between the two.

Alteration of curvature does not effect the ability of H-NS to repress promoter activity nor does it effect the ability of VirB to alleviate repression.

H-NS has been shown previously to bind to DNA with high intrinsic curvature and our previous findings demonstrate that both H-NS and VirB bind to similar regions within the *icsP* upstream intergenic region. Using MUTACURVE (http://132.248.32.45/cgi-bin/mutacurv.cgi) to profile and predict the intrinsic curvature of the *icsP* upstream intergenic region, we have found 5 peaks displaying curvature values greater than 5 degrees/helical turn indicating a high degree of curvature (shown in Fig. 23). The advantage of using MUTACURVE is that it allows the *in silico* evaluation

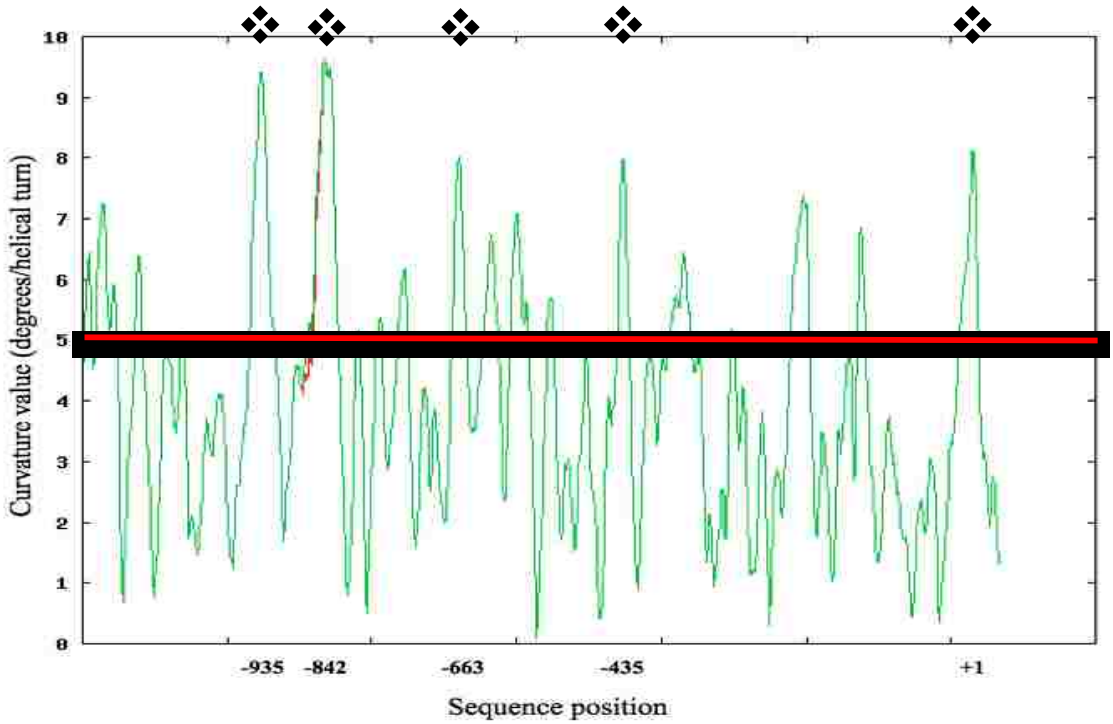

Fig. 23. *In silico* analysis of the static curvature of the DNA in the intergenic region of *icsP*. MUTACURVE predictions of curvature showing the 5 peaks with the highest predicted intrinsic curvature labeled below. >5 degrees/helical turn = curved (red line) ❖ marks the peaks of highest intrinsic curvature. Numbers below give positions of curvature relative to the tss (+1).

of the effect that double point mutations have on the effect of curvature; to either disrupt/reduce or to restore curvature. In order to examine the effect that this curvature

has on the regulation of *icsP* we adopted a strategy similar to that of De la cruz *et al.* (2009), making double point mutations to disrupt/reduce the curvature surrounding the peaks identified previously (Harrison *et al*., under revision; Fig. 19; Fig. 23).

Table 4. Substitutions to disrupt curvature

Postition of curved region	**Predicted curvature value (deg/helix turn) prior to substitution**	**Predicted curvature value (deg/helix turn) following substitution**	**Base pair substitutions responsible**
-842	9.65	4.82	A>C at -876 & C>A at -880
-935	9.36	5.19	G>T at -968 & A>C at -874
-663	8.09	2.67	C>A at -961 & T>G at -696
-435	7.88	2.99	T>G at -454 & A>C at -463

The mutations that were predicted to affect DNA curvature *in silico* were introduced into the wild type *icsP* upstream intergenic region upstream of +1. Each of the predicted peaks of curvature were disrupted separately by site directed mutagenesis before a second round of mutations was introduced to restore the predicted curvature without restoring the DNA to the original, wild type, sequence Table 2 and 3 (the predicted curvature surrounding the tss was not mutated). The mutated sequences were cloned into the *lacZ* reporter plasmid used previously (Appendix 1) and transformed into either *E. coli* (wild type and a strain lacking *hns*) or *Shigella* (wild type and a strain lacking *virB*). We expected that mutation of one of the regions of curvature would affect the ability of H-NS

Table 5. Substitutions to restore curvature

Position of curved region	Predicted curvature value (deg/helix turn) prior to substitution	Predicted curvature value (deg/helix turn) following substitution	Predicted curvature value (deg/helix turn) following substitution	Base pair substitutions responsible
-842	9.65	4.82	7.86	C>T at -876 & A>G at -880
-935	9.36	5.19	7.05	T>C at -968 & C>T at -974
-663	8.09	2.67	6.63	A>G at -691 & G>A at -696
-435	7.88	2.99	5.18	G>A at -454 & C>T at -463

to silence, and VirB to anti-silence, promoter activity, thereby revealing the sequences required for the regulation of the promoter.

Results of the β-galactosidase assay are given in Fig. 24. Surprisingly, comparing wild type (WT) to both, mutated and restored constructs, it is clear that substitutions predicted to disrupt curvature in each of the 4 regions had no effect on either: 1) promoter activity in the presence or absence of H-NS or 2) promoter activity in the presence and absence of VirB.

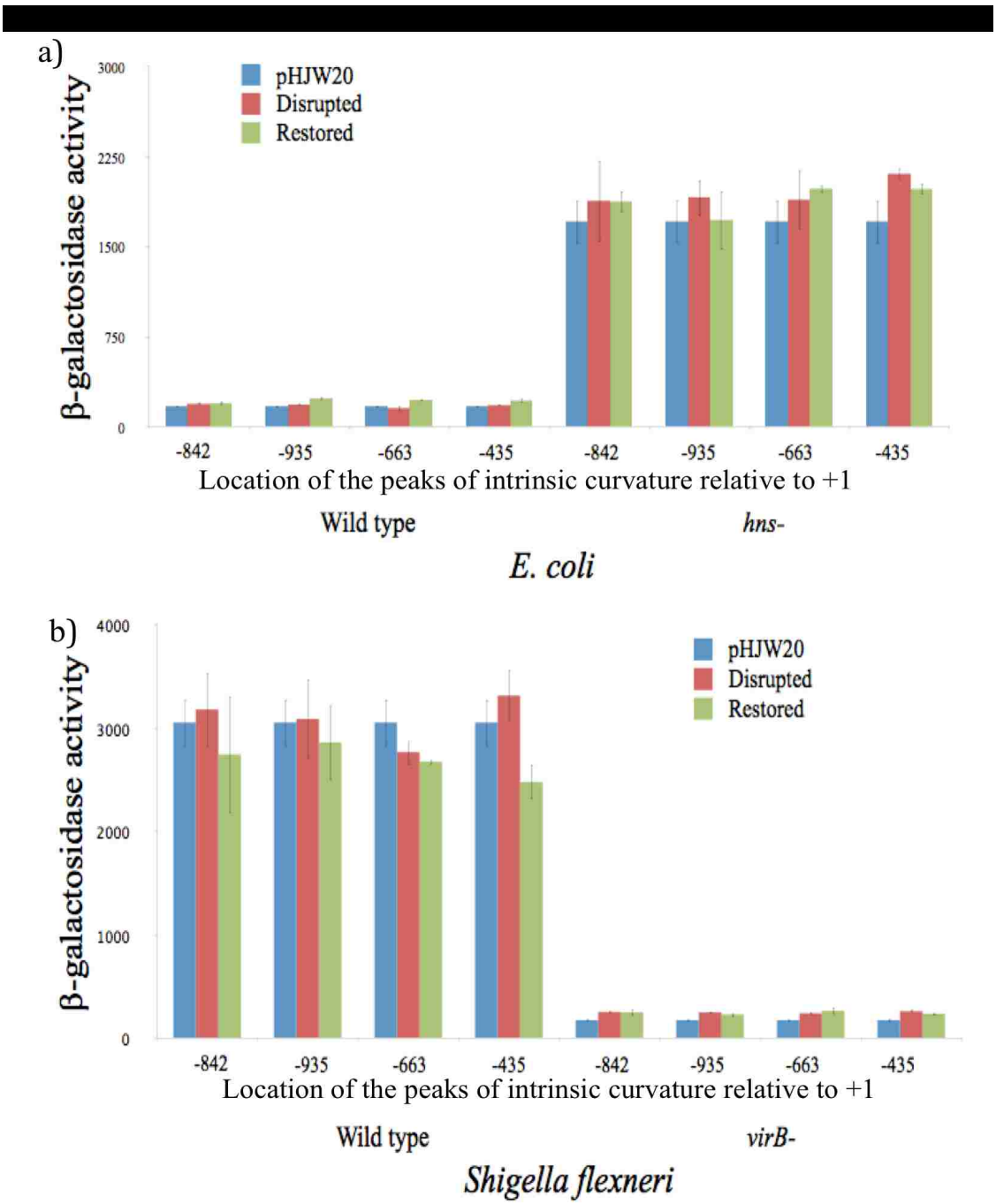

Fig. 24. Activities of the curve disrupted *icsP* promoter fragments. Activities of the curve disrupted promoter fragments in the presence and absence of **a) H-NS** or **b) VirB**. Constructs are labeled according to the predicted position of curvature upstream of the transcription start site in order from highest to lowest predicted curvature. Bars indicate β-galactosidase expression of the P*icsP-lacZ* fusions in either **a)** wild type *E. coli* MC4100 and an isogenic strain lacking *hns* (MC4100 *hns*::Kn^r) **b)** wild type *S. flexneri* 2457T and an isogenic strain lacking *virB* (AWY3 *virB*::Kn^r). WT *icsP* (pHJW20) is included in each lane for comparison. β -galactosidase activities are expressed in Miller units. This is a representative assay and error bars represent the mean and standard deviations.

It is possible that double point mutations intended to disrupt the curvature of one single peak have no effect given the overall length of the upstream intergenic region (1232 bp), where another curved region may compensate for the reduced curvature of an adjoining region simply because of the plasticity of such a large A-T rich fragment of DNA. It is also possible that computer software might not be able to accurately predict the curvature of DNA. We are following up these experiments to test whether these regions are curved using biochemical tests, as described by Prosseda *et al.* (2004).

It is possible that mutations made to disrupt multiple curved regions may have a greater effect on H-NS and VirB function by: i) possibly reducing the ability of the proteins to bind to *icsP*, ii) by reducing the ability of the upstream regions to interact with downstream regions, iii) or a combination of the two. Future studies will address this concern, by disrupting multiple regions of curvature in combination, to elucidate a complete model of transcriptional regulation involving *icsP*, and these studies will frame the basis for future work.

4.4 Conclusion

In conclusion, this study furthers our understanding of the mechanism of transcriptional regulation with the *Shigella icsP* gene, and raises several questions that are central to virulence gene expression in *Shigella* and other enteric pathogens and which lie at the heart of bacterial physiology. How do VirB and H-NS interact in the full context of the upstream intergenic region of *icsP*? In the absence of a consensus sequence, how is H-NS able to recognize DNA? What effect does curvature play on the H-NS dependent silencing and VirB dependent anti-silencing of *icsP*? These questions,

and many more, will need to be answered in order to achieve a complete picture of the mechanism of transcriptional silencing and anti-silencing.

CHAPTER 5

CONCLUSION

In this work I have focused on the regulation of the *icsP* gene, and examined the roles that H-NS and VirB play in the mechanism of silencing and anti-silencing at this promoter. This chapter addresses experiments that focus on five objectives. The first objective was to identify regions of the *icsP* promoter required for H-NS mediated repression and VirB dependent de-repression of the promoter. As a starting point for this work I conducted a promoter deletion analysis. While these experiments do not prove that either protein interacts directly with the *icsP* promoter, they allowed me to identify regions of the promoter needed for the VirB and, to some extent H-NS, dependent regulation of the *icsP* promoter. These regions could then be further interrogated using Electrophoretic Mobility Shift Assays (EMSAs), the second objective. The third objective was to determine the sequences to which H-NS and VirB bound *in vitro*. The fourth objective was to examine the effect that *icsP* static DNA curvature plays in the ability of H-NS and VirB to regulate promoter activity. Finally, the fifth objective was to propose a mechanism for the H-NS and VirB regulation of the *icsP* promoter.

In Chapter 2, my data show that H-NS represses *icsP* promoter activity and that VirB alleviates this repression from sites located over 1 kb upstream of the annotated transcription start site (Castellanos *et al*., 2009). Through deletion analysis we have also identified a region located promoter distally, -1130 to -1144, that is required for the de-repression of promoter activity. Located in this distal region are two VirB binding sites, organized as an inverted repeat, 5'-CGGGG**ATTTCAG**T**ATGAAAT**GAAGTA-3′(wild type VirB boxes shown here; Table 1), that were identified in an *in silico* analysis, and shown

to be absolutely required for the VirB dependent de-repression of the promoter. Our data revealed that complete mutagenesis of the upstream binding site (box 2), the downstream binding site (box 1) or both, resulted in complete loss of VirB dependent regulation of the *icsP* promoter (Table 1). To our knowledge, this is the first evidence that VirB can influence promoter activity from such distal sites. Although it is unusual for H-NS and VirB to impart its effects from so far upstream of the transcription start site, our truncation analysis of the *icsP* upstream intergenic region has shed light on how this remote regulation might occur, but raises some very important questions: (i) Are other *Shigella* virulence plasmid genes regulated from remote VirB binding sites? (ii) Is it common for transcriptional silencing and anti-silencing mechanisms to employ distal regulator binding sites?

The experiments described in Chapter 3, focus on the regulation of the *icsP* promoter by H-NS. Through deletion analysis of the *icsP* upstream intergenic region and *in vivo* assays, I, and others in the Wing lab, identified that sequences upstream of -665 (with respect to the *icsP* annotated transcription start site [tss]) are needed for full H-NS induced repression of activity the promoter, although sequences between -665 and -351 appear sufficient for partial repression. Furthermore, our data strongly suggests that H-NS binds to two discrete regions (upstream of -351 and downstream of -213) *in vivo*, that lie upstream of the *icsP* gene. This observation is backed up by *in vitro* EMSAs and based on our assays it seems that these two regions are separated by a region with little to no affinity for HNS (regions contained within T2 in Fig. 16). Both regions to which H-NS binds, contain sequences predicted to display high levels of intrinsic curvature. These data support a model where H-NS docked at remote sites may act in concert with H-NS

docked at promoter proximal sequences to mediate H-NS-dependent repression of the *icsP* promoter. While we again demonstrate that VirB functions solely to de-repress the promoter, our EMSA data indicate that VirB and H-NS are capable of binding simultaneously to the full intergenic region (1232 bp) that lies immediately upstream of the *icsP* gene. This suggests that VirB may function to relocate H-NS, rather than causing H-NS to dissociate from the DNA upstream of the *icsP* gene.

In Chapter 4, my work has demonstrated that VirB binds the upstream intergenic region of *icsP*, and I have shown that both VirB and H-NS oligomerize along the DNA. I have also investigated the role that DNA intrinsic curvature plays in the regulation of *icsP*. Using electrophoretic mobility shift assays (EMSAs), and DNase I footprinting analysis I show that VirB binds directly to the upstream intergenic region of *icsP* and demonstrate that VirB and H-NS bind and oligomerize to DNA sequences located promoter distally and proximally. From these results we conclude that both H-NS and VirB interact directly with the *icsP* promoter and confirm our *in vivo* results demonstrating that sequences located over 1 kb upstream of the annotated tss are required for the H-NS dependent silencing and the VirB dependent anti-silencing of the *icsP* promoter. These data support a model where VirB bound to remote sites may act in concert with VirB bound to promoter proximal sequences to mediate VirB dependent anti-silencing of the *icsP* promoter.

In Chapter 4, I also show that the disruption of single regions displaying high intrinsic curvature do not have an effect on *icsP* promoter activity. These regions of predicted curvature were mutated using site directed mutagenesis to disrupt and ultimately restore the curvature. Surprisingly, disruption of these regions of predicted curvature did not

have an effect on either the ability of H-NS to silence the promoter or the ability of VirB to alleviate the H-NS induced repression. Likewise, mutations to restore the predicted curvature had no effect. It is likely that disruption of a single peak within the context of 1.2 kb is not sufficient to elicit an effect.

Taken together, these data have allowed us to propose a model for the regulation of *icsP*. At the non-permissive temperature of 30 ˚C H-NS binds and induces H-NS-DNA-H-NS loops or bridges between promoter distal and promoter proximal sequences. At the permissive temperature of 37 ˚C, the temperature at which VirB is maximally expressed, VirB binds to its consensus site (Taniya *et al*., 2003, Tobe *et al*., 1991, Turner & Dorman, 2007), which is AT rich and frequently lies in AT rich stretches of DNA. These sequences likely are already bound by H-NS. We hypothesize that VirB does not completely remove H-NS from the DNA, but instead binds simultaneously and possibly re-organizes the DNA:protein complex by binding to promoter distal VirB binding sites in conjunction with promoter proximal sequences bound by VirB. The nucleoprotein complex might be stabilized through the interaction of VirB bound to sites located internally. This multi-region DNA:protein interaction then possibly leads to a re-organization of the topology of the upstream intergenic region that is permissive for *icsP* promoter activity (Fig. 25). This model should frame our future investigations.

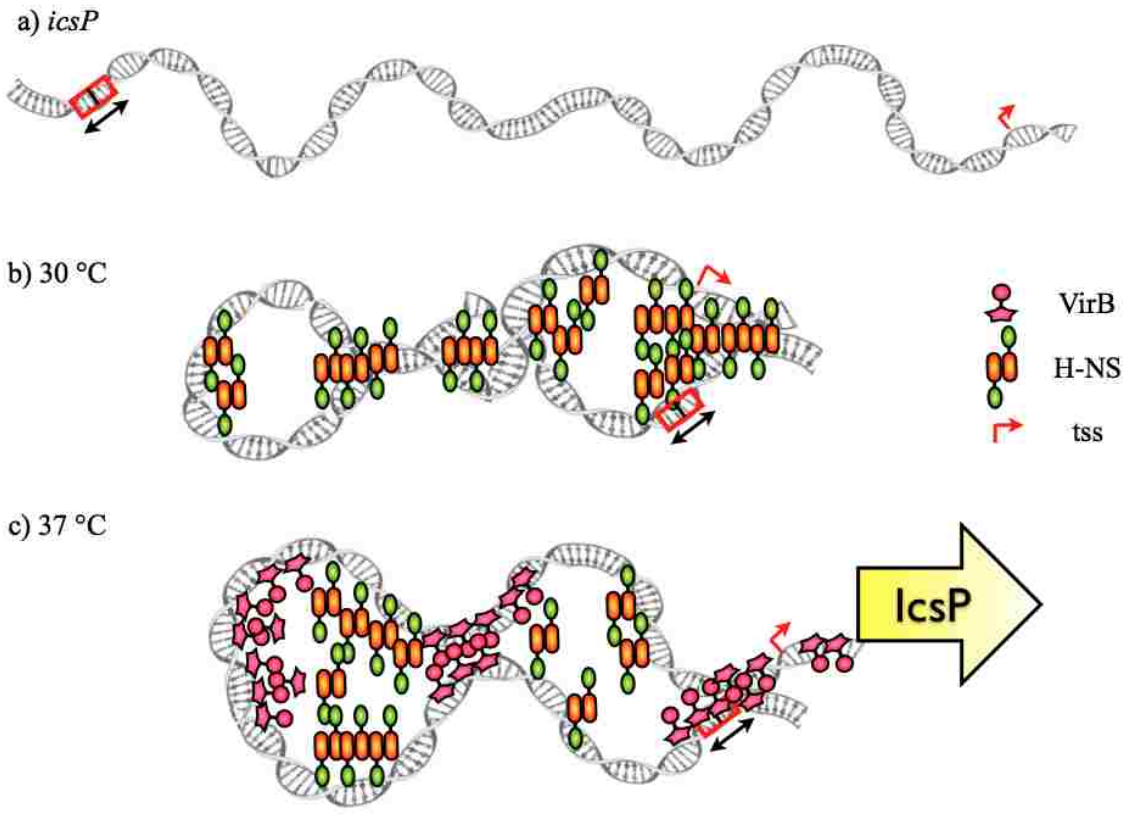

Fig. 25. Proposed model for the regulation of *icsP* by H-NS and VirB. The upstream intergenic region of *icsP* (a) showing the two distal, essential, VirB binding sites (black arrows under the red box). At the non-permissive temp of 30 °C H-NS induces repression of promoter activity by binding to sequences located distally and proximally, resulting in a plectonemic structure stabilized by oligomerization of H-NS (b). Upon a shift in temperature, to the permissive temperature of 37 °C, VirB is expressed and binds to *icsP*. Instead of displacing H-NS, VirB binds simultaneously causing the reorganization in the topology of the promoter, interacting with the 2 most distal VirB binding sites and promoter proximal sites. The resulting structure is stabilized through VirB bound to internal sequences. The altered architecture of the promoter is permissive for transcription (c). Shown is a VirB monomer; H-NS in dimeric form; bent arrow = tss.

While the work containd herein contributes to our knowledge, much more work is required to build a complete picture of transcriptional silencing and anti-silencing at the *icsP* promoter. It would be interesting to asses the contribution of all the VirB binding sites identified in Castellanos *et al.* (2009) in the regulation of the *icsP* promoter, and to examine whether a VirB:DNA:VirB loop is responsible for the observed effects presented in this work. Site directed mutagenesis of combinations of putative VirB binding sites would indicate which sequences are essential, and could be combined with

$KMNO_3$ footprinting to demonstrate a loop(s) is/are being formed. The intergenic region that lies upstream of *icsP* is very AT rich, and typically such sequences have high intrinsic curvature. While our initial results suggest that no single region of predicted curvature is required it could be that disruption of multiple peaks in combination will produce an observable effect. Therefore, it would also be interesting to assess the effect of combinations of disrupted peaks on the ability of H-NS and VirB to regulate promoter activity.

It would be beneficial to investigate the possibility that a multi-protein complex is involved in *icsP* regulation, similar to situations described in other promoters (Browning *et al.,* 2000). The VirB protein is homologous to the P1 protein ParB, and the ParB binding site, *parS,* shows a high degree of similarity to the *icsB* binding site described by Turner & Dorman (2007) and Taniya *et al.* (2003) (Beloin *et al.,* 2002, Taniya *et al.,* 2003, Surtees & Funnell, 2001). We have shown that the essential distal VirB binding site in *icsP* shows sequence homology with both these sites (*parS* and *icsB*). An *in silico* analysis of this region using the Softberry software BPROM, which analyzes bacterial promoters, predicts an integration host factor (IHF) binding site oriented in a similar position to that found in *parS.* In the *parS* system, IHF bends the *parS* site which allows ParB to contact its specific binding motifs located on either side of the bend resulting in a high affinity protein:DNA complex. Therefore, I have made an IHF mutant in *Shigella* and preliminary data suggests the possible involvement of IHF in the regulation of promoter activity and should be followed up (See Appendix 2 for full description). To find other potential candidates, chemical cross-linking of proteins, followed by co-

immunoprecipitation should be able to pull out any proteins associating with VirB, and likewise *icsP*.

Finally, a closer examination of the sequences to which VirB is able to bind might elucidate a more complete description of a VirB binding site. Evidence for this comes from the observation that VirB bound very well to T5 by DNase I footprint analysis, and is worthy of further study.

In conclusion, this work has focused on the regulation of *icsP* and has examined the roles that H-NS and VirB play in the mechanism of silencing and anti-silencing. This work has shed light on the phenomenon of regulation from remote sites, yet it raises many more questions. Does VirB function from remote sites to regulate the expression of other *Shigella* virulence genes? An *in silico* analysis of another H-NS and VirB-regulated promoter located on the *Shigella* virulence plasmid, *icsB-ipgD* (Turner & Dorman, 2007), reveals the presence of additional putative VirB binding sites with the most distal located at -1615, and -1464 upstream of the *ipgD* tss, and one located at -1625

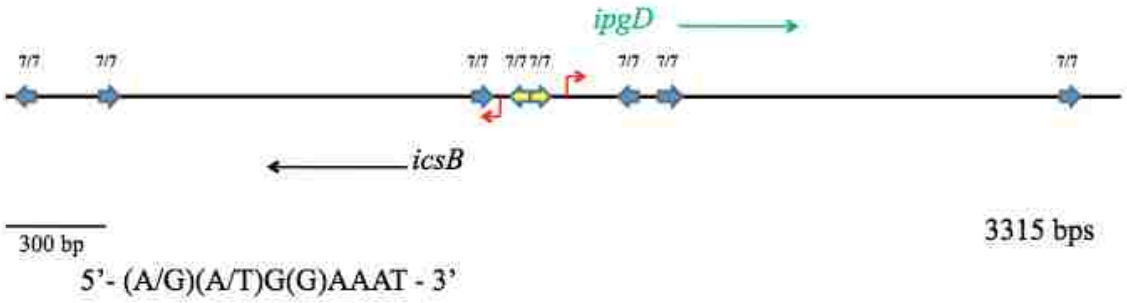

Fig. 26. *In silico* analysis of the *icsB-ipgD* promoter region. *In silico* analysis reveals the presence of additional putative VirB binding sites that show consensus sequence matches to the reported VirB binding site. The binding sites in *icsB* are located at: -1625, -390, and -361 upstream of the tss. The binding sites in *ipgD* are located at: -1615, -1464, and -289 upstream of the tss. Arrows show direction of the binding site 5' - 3.' Binding site sequence is given below the figure. Figure is drawn to scale.

upstream of the *icsB* tss, respectively (Fig. 26). Is it possible that these divergently transcribed genes are regulated from remote sites?

Can transcription factors found in other bacteria function from remote sites to alleviate transcriptional repression by nucleoid structuring proteins? If so, how widespread is this phenomenon? Silencing and anti-silencing mechanisms that control expression of virulence genes in *Shigella* and many other bacteria, is central to understanding pathogenesis. However, these mechanisms are not restricted to the human pathogens but may be more wide spread, we need only look.

APPENDIX 1

BACTERIAL STRAINS, PLASMIDS, AND PRIMERS

Table 1 Bacterial strains and plasmids

Strain or Plasmid	Description[a]	Source or Reference
Strains		
E. coli		
MC4100	*E. coli* strain K-12 with *ara*D and *lacZ* deletion	(Pogliano & Beckwith, 1994)
MC4100 *hns*	MC4100 *hns*::Kn^r	(Yamada *et al*., 1991)
S. flexneri		
2457T	*S. flexneri* serotype 2a	(Formal *et al*., 1958)
AWY3	2457T *virB*::Tn5; Kn^r	(Wing *et al.,* 2004)
Plasmids		
pACYC184	Cloning vector; p15A replicon Tet^r/Cm^r/source of non-specific competitor DNA	(Rose, 1988)
pHJW7	*icsP* promoter region transcriptionally fused to *lacZ* in pACYC184 Cm^r.	(Wing *et al.,* 2004)
pBluescript	Cloning vector	Stratagene
pQE-60	Cloning vector for production of C-terminally His-tagged proteins	Qiagen
pHJW20	pHJW7 lacking *XbaI* site downstream of *lacZ* gene. Carries 1232 bp of native sequence upstream of the *icsP* transcription start site	(Wing *et al*., 2004)
pJS01	pHJW20 carrying 1056 bp of native sequence upstream of the *icsP* transcription start site.	(Castellanos *et al*., 2009)
pJS02	pHJW20 carrying 891 bp of native sequence upstream of the *icsP* transcription start site.	(Castellanos *et al*., 2009)

pDH01	pHJW20 carrying 663 bp of native sequence upstream of the *icsP* transcription start site.	(Castellanos *et al*., 2009)
pJS04	pHJW20 carrying 351 bp of native sequence upstream of the *icsP* transcription start site.	(Castellanos *et al*., 2009)
pHJW34	pHJW20 carrying 254 bp of native sequence upstream of the *icsP* transcription start site.	(Castellanos *et al*., 2009)
pHJW35	pHJW20 carrying 150bp of native sequence upstream of the *icsP* transcription start site.	(Castellanos *et al*., 2009)
pHJW36	pHJW20 carrying 92 bp of native sequence upstream of the *icsP* transcription start site.	(Castellanos *et al*., 2009)
pMIC01	pHJW20 with 33bp deleted between *Sal*I and *Pst*I sites in the multiple cloning site of pACYC184 this changes the context of the upstream *icsP* promoter sequences.	(Castellanos *et al*., 2009)
pMIC02	pHJW20 carrying 1437 bp of native sequence upstream of the *icsP* transcription start site.	(Castellanos *et al*., 2009)
pMIC13	pHJW20 with all 7 bps substituted in promoter proximal upstream VirB binding site	(Castellanos *et al*., 2009)
pMIC17	pHJW20 with all 7 bps substituted in distal upstream VirB binding site	(Castellanos *et al*., 2009)
pMIC18	pHJW20 with all 7bps substituted in both upstream VirB binding sites	(Castellanos *et al*., 2009)
pShfT1	EMSA/DNase I footprint Target#1 in pBluescript	This work
pShfT2	EMSA/DNase I footprint Target#2 in pBluescript	This work
pShfT3	EMSA/DNase I footprint Target#3 in pBluescript	This work
pShfT4	EMSA/DNase I footprint Target#4 in pBluescript	This work

pShfT5	EMSA/DNase I footprint Target#5 in pBluescript	This work
pShfT6	EMSA/DNase I footprint Target#6 in pBluescript	This work
pAJH01	IPTG inducible vector carrying his-*virB*	This work
pCTH01	IPTG inducible vector carrying his-*hns*	This work
pBAD18	L-arabinose inducible pBAD expression vector, pBR *ori;* Amp^r	(Guzman *et al.,* 1995)
pATM324	pBAD18-*virB*; Amp^r	((Schuch *et al*, 1999)
pNEO1	*Pst*I, *Xba*I insert of *icsP* cloned into pBluescript. Carries 1232 bp of native sequence upstream of the *icsP* transcription start site. Used as template for site-directed mutagenesis	This work
pMut1 D	pHJW20 with 2bp substitution to *disrupt* curvature at -842 A>C at -876 & C>A at -880 relative to *icsP* tss.	This work
pMut2 D	pHJW20 with 2bp substitution to *disrupt* curvature at -935 G>T at -968 & A>C at -974 relative to *icsP* tss.	This work
pMut3 D	pHJW20 with 2bp substitution to *disrupt* curvature at -663 C>A at -691 & T>G at -696 relative to *icsP* tss.	This work
pMut4 D	pHJW20 with 2bp substitution to *disrupt* curvature at -435 T>G at -454 & A>C at -463 relative to *icsP* tss.	This work
pMut1 R	pHJW20 with 2bp substitution to *restore* curvature at -842 C>T at -876 & A>G at -880 relative to *icsP* tss.	This work
pMut2 R	pHJW20 with 2bp substitution to *restore* curvature at -935 T>C at -968 & C>T at -974 relative to *icsP* tss.	This work
pMut3 R	pHJW20 with 2bp substitution to *restore* curvature at -663 A>G at -691 & G>A at -696 relative to *icsP* tss.	This work

pMut4 R	pHJW20 with 2bp substitution to *restore* curvature at -435 G>A at -454 & C>T at -463 relative to *icsP* tss.	This work

[a] Amp[r], ampicillin resistance; Tet[r], tetracycline resistance; Cm[r], chloramphenicol resistance; Kn[r], kanamycin resistance

Table 2 Oligonucleotide Primers

Primer	**Sequence[a] 5'-3'**	Description or use[b]
W1016	CAGATAAAATATTTCAAGATTTCAGTGC	Mutation of XbaI binding site downstream of the lacZ gene.
W1017	CCCGCATCCACAGGACGGGTGTGGTCG	Used with W1016 to create 'mega'-primer lacking XbaI site.
W1018	GTGAGAGGGCCGCGGCAAAGCC	Used with 'mega'-primer to create fragment lacking XbaI site.
W530	GGGTACCTGCAGTAAAAGATATGTTCTTGG	Binds to sequences -92 to -75 relative to the icsP tss. Used with W48 to make pHJW36.
W44	GGGTACCTGCAGGCCTCTTTATTATAAGTAAGATCTGGC	Binds to sequences -1056 to -1030 relative to the icsP tss. Used with W48 to make pJS01.
W45	GGGTACCTGCAGAGTGAACTCTCGCCCTATACGGCG	Binds to sequences -891 to -868 relative to the icsP tss. Used with W48 to make pJS02.
W46	GGGTACCTGCAGTTTGGTCACTTTAACTGTATTAGTCGC	Binds to sequences -663 to -637 relative to the icsP tss. Used with W48 to make pDH01.
W47	GGGTACCTGCAGGTTGGTACTGAAAGGCACGTTGGC	Binds to sequences -351 to -328 relative to the icsP tss. Used with W48 to make pJS04.

W528	GGGTACCTGCAGGGCGAGAGAACAAGAAGGTGAGTAG	Binds to sequences -254 to -241 relative to the icsP tss. Used with W48 to make pHJW34.
W529	GGGTACCTGCAGCTTTATTCGCACAATGAGG	Binds to sequences -150 to -132 relative to the icsP tss. Used with W48 to make pHJW35.
W530	GGGTACCTGCAGTAAAAGATATGTTCTTGG	Binds to sequences -92 to -75 relative to the icsP tss. Used with W48 to make pHJW36.
W63	TGAGACGTCGACAGAACTCTACTTTTTTGGTTGAAATGTCC	Binds to sequences -3 to +24. Used to make *icsP* promoter Target#1 Reverse
W144	CCGGGAATTCTTGTCGCGGAATCCTGAAACTATCAGCC	Binds to sequences -227 to -200. Used to make *icsP* promoter Target#1 Forward
W65	TGAGACGTCGACTCATTGTGCGAATAAAGTAACGGGGGC	Binds to sequences -160 to -134. Used to make *icsP* promoter Target#2 Reverse
W145	CCGGGAATTCAGGTTGTCATGTTGTGACTGAAAGGC	Binds to sequences -361 to -336 . Used to make *icsP* promoter Target#2 Forward
W67	TGAGACGTCGACCGATAAGACAATAAATTGTCATCACGGC	Binds to sequences -320 to -293. Used to make *icsP* promoter Target#3 Reverse
W157	CCGGGGATCCTCTGACTCTCGACTTTAAAAGGATGGG	Binds to sequences -635 to -608 . Used to make *icsP* promoter Target#3 Forward

W69	TGAGACGTCGACGGGCACCTCACTTTAGCACTGAAGCC	Binds to sequences -548 to -523 . Used to make *icsP* promoter Target#4 Reverse
W147	CCGGGAATTCGCGTTTTCAAGGATTAGTCCTTAATCGG	Binds to sequences -870 to -843 . Used to make *icsP* promoter Target#4 Forward
W71	TGAGACGTCGACAATAAATATTCTACTAATTTAATGCATCAC	Binds to sequences -759 to -730 . Used to make *icsP* promoter Target#5 Reverse
W148	CCGGGAATTCAAGCCTCTTTATTATAAGTAAGATCTGGC	Binds to sequences -1058 to -1030 . Used to make *icsP* promoter Target#5 Forward
W73	TGAGACGTCGACCAATAAAATGGTTGGTTGAAGGTCGTG	Binds to sequences -995 to -969. Used to make *icsP* promoter Target#6 Reverse
W149	CCGGGAATTCATGCTTGATAACTTAATTGGGGCTCCC	Binds to sequences -1214 to -1188. Used to make *icsP* promoter Target#6 Forward
W81	ATGTTTATGATACCTGCAGG	Binds to sequences -1219 to -1232 relative to the icsP tss. Used to amplify sequences contained in pMIC02.
W82	AGGACTTCCGTCGACCAGGAAATAGTGCGGGCAAC	Binds to sequences -1473 to -1454 relative to the icsP tss. Used to amplify sequences contained in pMIC02.
W89	CAAATTATCTAGAAAAGATCGCAGG	Binds downstream of the transcription start site in icsP gene. Used in the construction of pMIC13, 17 & 18.

W98	CCCGTTCCATGTGCTCGCC	Binds to pACYC184 upstream of cloned sequences. Used in the construction of pMIC13, 17 & 18.
W99	/5Phos/ACTGAAATCCCCGGAAGTATTCAATC	Binds to promoter proximal upstream VirB binding site; wild type sequence. Used in the construction of pMIC17.
W100	/5Phos/A**GCTGGGC**CCCCGGAAGTATTCAATC	Binds close to promoter proximal upstream VirB binding site used to mutate this sequence. Used in the construction of pMIC13 & 18.
W101	/5Phos/ATGAAATGAAGTATATTTAATATACTTTAC	Binds to distal upstream VirB binding site; wild type sequence. Used in the construction of pMIC13.
W110	/5Phos/**CGACCCG**GAAGTATATTTAATATACTTTAC	Binds close to distal upstream VirB binding site used to mutate this sequence. Used in the construction of pMIC17 & 18.
W149	CCGGGAATTCATGCTTGATAACTTAATTGGGGCTCCC	Binds to sequences -1214 to -1188. Used to make icsP promoter Target#6 Forward
W75	CCGGGAATTCTCAATGAAATTGCTAATT	icsB positive control Reverse (Turner & Dorman, 2007)
W76	CCGGGAATTCATGCAATCCCAAATTAGT	icsB positive control Forward (Turner & Dorman, 2007)
W77	TGAGACGTCGACCCACTCTTCGTTCACTTTCGCC	pstS negative control Forward binds internal to translation start site

W78	GTGTCAGTCGACCTCAGGAAGGTCTGTTCCAGTTCCC	pstS negative control Reverse binds internal to translation start site
W108	GTGTCAGTCGACTATTGGTCTTCCACGCAGCGC	pstS' negative control Reverse orf
W109	GTGTCAGTCGACTTATGAAAGTTATGCGTACCACCG	pstS' negative control Forward orf
W38	CAGGGTGTGCCATGGTGGATTTGTGC	Used to clone virB into pQE60
W39	CAGCTAATTAAGCTTAGTGATGGTGATGGTGATGTCCTGAAGAC GATAGATGGC	Used to clone virB into pQE60
W134	TCATTCCCATGGCCGAAGCACTTAAAATTCTGAAC	Used to clone hns into pQE60
W135	TCATTCAGATCTTTGCTTGATCAGGAAATCGTCG	Used to clone hns into pQE60
W112	CGCATTGTTAGATTTCATACACGG	Reverse primer binds downstream of ClaI in pACYC184 used to make non-specific competitor DNA.
W113.1	AAATGTAGCACCTGAAGTCAGCC	Forward primer binds upstream of ClaI in pACYC184 used to make non-specific competitor DNA.
W206	TCCTTGAAAACGCCGTAT**C**GGG**A**GAGAGTTCACTCTATTG	Forward primer to disrupt curvature at -842 relative to the icsP tss
W207	CAATAGAGTGAACTCTC**T**CCC**G**ATACGGCGTTTTCAAGGA	Reverse primer to disrupt curvature at -842 relative to the icsP tss
W208	CCACTACTGAACTATTCTGGGATTTAAAT**T**CAATA**C**AATGGTTG GTTGAAGGTC	Forward primer to disrupt curvature at -935 relative to the icsP tss

W209	GACCTTCAACCAACCATT**G**TATTG**A**ATTTAAATCCCAGAATAGT TCAGTAGTGG	Reverse primer to disrupt curvature at -935 relative to the icsP tss
W210	GTTATCCTCTACACATTTTATCT**A**ATGT**G**TACTTGGACAATTCA TTGGTGC	Forward primer to disrupt curvature at -663 relative to the icsP tss
W211	GCACCAATGAATTGTCCAAGTA**C**ACAT**T**AGATAAAATGTGTAGA GGATAAC	Reverse primer to disrupt curvature at -663 relative to the icsP tss
W212	GTATAGTTTTTCAGATTTTGTTT**G**TCCCGTGA**C**ATATTATGGAA ATGGCAGAAGCA	Forward primer to disrupt curvature at -435 relative to the icsP tss
W213	TGCTTCTGCCATTTCCATAATAT**G**TCACGGGA**C**AAACAAAATCT GAAAAACTATAC	Reverse primer to disrupt curvature at -435 relative to the icsP tss
W214	TCCTTGAAAACGCCGTAT**T**GGG**G**GAGAGTTCACTCTATTG	Forward primer to restore curvature at -842 relative to the icsP tss
W215	CAATAGAGTGAACTCTC**C**CCC**A**ATACGGCGTTTTCAAGGA	Reverse primer to restore curvature at -842 relative to the icsP tss
W216	ACTACTGAACTATTCTGGGATTTAAAT**C**CAATA**T**AATGGTTGGT TGAAGGT	Forward primer to restore curvature at -935 relative to the icsP tss
W217	ACCTTCAACCAACCATT**A**TATTG**G**ATTTAAATCCCAGAATAGTT CAGTAGT	Reverse primer to restore curvature at -935 relative to the icsP tss
W218	GTTATCCTCTACACATTTTATCT**G**ATGT**A**TACTTGGACAATTCA TTGGTGC	Forward primer to restore curvature at -663 relative to the icsP tss

W219	GCACCAATGAATTGTCCAAGTA**T**ACAT**C**AGATAAAATGTGTAGA GGATAAC	Reverse primer to restore curvature at -663 relative to the icsP tss
W220	GTATAGTTTTTCAGATTTTGTTT**A**TCCCGTGA**T**ATATTATGGAA ATGGCAGAAGCA	Forward primer to restore curvature at -435 relative to the icsP tss
W221	TGCTTCTGCCATTTCCATAATAT**A**TCACGGGA**T**AAACAAAATCT GAAAAACTATAC	Reverse primer to restore curvature at -435 relative to the icsP tss

[a] Restriction sites underlined.

[b] tss, transcription start site.

[c] Bold indicates bases mutated by site directed mutagenesis

APPENDIX 2

PRELIMINARY EXAMINATION OF THE INVOLVEMENT OF IHF IN THE REGULATION OF THE *ICSP* PROMOTER

The VirB protein is homologous to the P1 protein ParB, and the ParB binding site, *parS,* shows a high degree of similarity to the *icsB* binding site described by Turner & Dorman (2007) and Taniya *et al.* (2003) (Beloin *et al.,* 2002, Taniya *et al.,* 2003, Surtees & Funnell, 2001). We have shown that the essential distal VirB binding site in *icsP* shows sequence homology with both these sites (*parS* and *icsB*). An *in silico* analysis of this region using the Softberry software BROM, which analyzes bacterial promoters, predicts an integration host factor (IHF) binding site oriented in a similar position to that found in *parS* (Fig. 27).

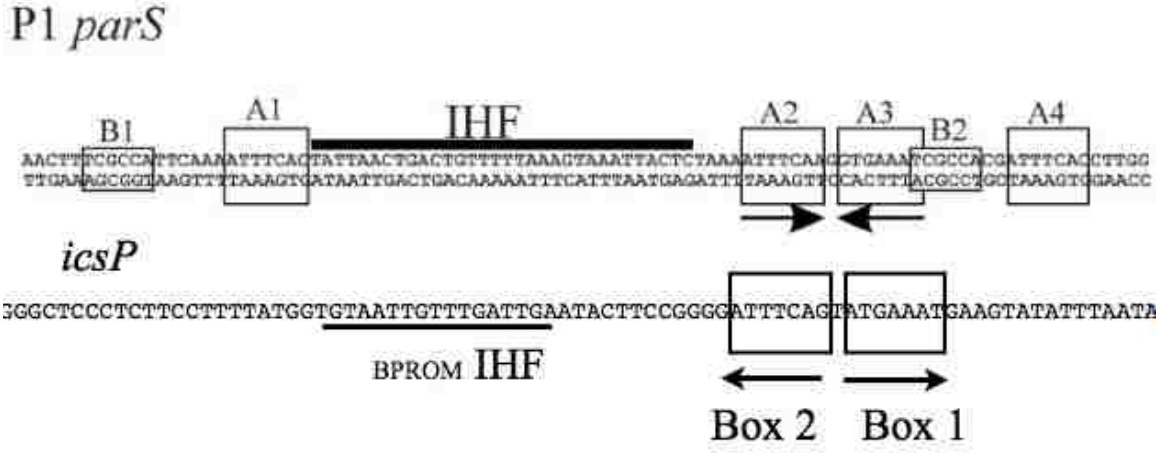

Fig. 27. Comparison of P1 *parS* and *icsP*. Comparison of the P1 *parS* and *icsP* sites showing the ParB binding sites in *parS* (upper) and the VirB binding sites in *icsP* (lower). The IHF binding site is shown for *parS* and the BPROM predicted IHF site is shown in *icsP*. *parS* (Surtees & Funnell, 2001).

Based on this finding, I conducted a preliminary experiment to assess the potential contribution that IHF makes in the regulation of *icsP* promoter. IHF exists as a

heterodimeric sequence specific DNA binding protein and functions as both a structural element in the cell and as a conventional regulator of gene expression. The two subunits of the protein are encoded by the chromosomal genes *himA* and *himB* (also seen in the literature as *ihfA* and *ihfB*) (Funnell, 1988). Its role is to bend DNA up to 180° which facilitates the formation of nucleoprotein complexes. IHF is a positive regulator of both *virF* and *virB* (Porter & Dorman, 1997).

Through P1 transduction I introduced a tetracycline resistance cassette into the *himA* gene of *Shigella flexneri* 2a, wild type and an isogenic strain lacking *virB,* to ablate the function of the IHF protein. Similar strategies have been used to disrupt IHF function since IHF exists as a heterodimeric protein; disruption of either gene is sufficient to destroy function (Funnell, 1988). These cells were transformed with our reporter plasmid containing the full upstream intergenic region of *icsP::lacZ* and assayed in the β-galactosidase assay (described in materials and methods, chapters 2, 3 and 4). The results in Fig. 28 suggest that IHF does not play a role in the VirB dependent regulation of *icsP* promoter activity.

In the *parS* system, IHF bends the *parS* site which allows ParB to contact its specific binding motifs located on either side of the bend resulting in a high affinity protein:DNA complex. However, in the absence of IHF, ParB binds specifically, but more weakly to *parS* and only requires the right half of the *parS* site for activity (Vecchiarelli *et al*, 2007). Therefore, it could be possible that in our reporter system the absence of IHF is not sufficient to produce a defect in the ability of VirB to regulate promoter activity, in fact, the data suggests that in the absence of *himA*- the presence of VirB increases promoter activity above what is seen in the wild type. Could this be evidence for IHF

participation in promoter regulation? It is an interesting finding that bears further investigation. It would be interesting to investigate the interaction between VirB, H-NS, and IHF. An *E. coli* MC4100 *himA*- construct has been made, however, the *hns-, himA*- double knock out has been difficult to create as this is likely deleterious to the cell, but the potential information that could be gained from demonstrating a mulit-nucleoprotein complex in the regulation of *icsP* would be worth the effort.

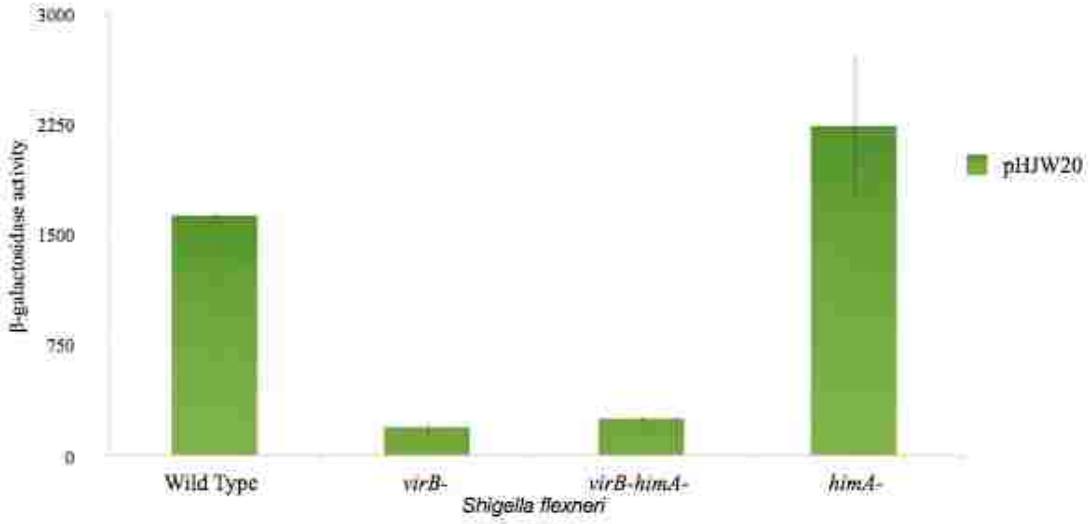

Fig. 28. IHF does not regulate *icsP* promoter activity. β-galactosidase assay of the full upstream intergenic region of *icsP::lacZ* (pHJW20). Bars indicate β-galactosidase expression in the presence and absence of *virB*, *himA* and the double mutant lacking both. β-galactosidase activities are expressed in Miller units and mean and standard deviations are shown. This graph is a representative sample.

BIBLIOGRAPHY

Abeles 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* **185**, 261-272.

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**, 627-635.

Ali Azam T., Iwata A., Nishimura A., Ueda S., Ishihama A. (1999). Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J Bacteriol* **181**, 6361-6370.

Amit R., Oppenheim A. B., Stavans J. (2003). Increased bending rigidity of single DNA molecules by H-NS, a temperature and osmolarity sensor. *Biophys J* **84**, 2467-2473.

Atlung T. & Ingmer H. (1997). H-NS: A modulator of environmentally regulated gene expression. *Mol Microbiol* **24**, 7-17.

Azam T. A. & Ishihama A. (1999). Twelve species of the nucleoid-associated protein from *Escherichia coli*. sequence recognition specificity and DNA binding affinity. *J Biol Chem* **274**, 33105-33113.

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**, 41657-41666.

Balleza E., Lopez-Bojorquez L. N., Martinez-Antonio A., Resendis-Antonio O., Lozada-Chavez I., Balderas-Martinez Y. I., Encarnacion S., Collado-Vides J. (2009). Regulation by transcription factors in bacteria: Beyond description. *FEMS Microbiol Rev* **33**, 133-151.

Barbic A. & Crothers D. M. (2003). Comparison of analyses of DNA curvature. *J Biomol Struct Dyn* **21**, 89-97.

Barbic A., Zimmer D. P., Crothers D. M. (2003). Structural origins of adenine-tract bending. *Proc Natl Acad Sci U S A* **100**, 2369-2373.

Belitsky B. R. & Sonenshein A. L. (1999). An enhancer element located downstream of the major glutamate dehydrogenase gene of *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **96**, 10290-10295.

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**,

825-838.

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**, 15333-15344.

Beloin C., Deighan P., Doyle M., Dorman C. J. (2003). *Shigella flexneri* 2a strain 2457T expresses three members of the H-NS-like protein family: Characterization of the *sfh* protein. *Mol Genet Genomics* **270**, 66-77.

Bennish M. L. & Wojtyniak B. J. (1991). Mortality due to shigellosis: Community and hospital data. *Rev Infect Dis* **13 Suppl 4**, S245-51.

Berlutti F., Casalino M., Zagaglia C., Fradiani P. A., Visca P., Nicoletti M. (1998). Expression of the virulence plasmid-carried *apyrase* gene (*apy*) of enteroinvasive *Escherichia coli* and *Shigella flexneri* is under the control of H-NS and the VirF and VirB regulatory cascade. *Infect Immun* **66**, 4957-4964.

Bertin P., Terao E., Lee E. H., Lejeune P., Colson C., Danchin A., Collatz E. (1994). The H-NS protein is involved in the biogenesis of flagella in *Escherichia coli*. *J Bacteriol* **176**, 5537-5540.

Bignell C. & Thomas C. M. (2001). The bacterial ParA-ParB partitioning proteins. *J Biotechnol* **91**, 1-34.

Bloch V., Yang Y., Margeat E., Chavanieu A., Auge M. T., Robert B., Arold S., Rimsky S., Kochoyan M. (2003). The H-NS dimerization domain defines a new fold contributing to DNA recognition. *Nat Struct Biol* **10**, 212-218.

Boss A., Nussbaum-Shochat A., Amster-Choder O. (1999). Characterization of the dimerization domain in BglG, an RNA-binding transcriptional antiterminator from *Escherichia coli*. *J Bacteriol* **181**, 1755-1766.

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**, 1415-1424.

Bouet J. Y., Surtees J. A., Funnell B. E. (2000). Stoichiometry of P1 plasmid partition complexes. *J Biol Chem* **275**, 8213-8219.

Bouffartigues E., Buckle M., Badaut 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.

Bracco L., Kotlarz D., Kolb A., Diekmann S., Buc H. (1989). Synthetic curved DNA sequences can act as transcriptional activators in *Escherichia coli*. *EMBO J* **8**, 4289-4296.

Browning D. F., Cole J. A., Busby S. J. (2000). Suppression of FNR-dependent transcription activation at the *Escherichia coli nir* promoter by fis, IHF and H-NS: Modulation of transcription initiation by a complex nucleo-protein assembly. *Mol Microbiol* **37**, 1258-1269.

Broyles S. S. & Pettijohn D. E. (1986). Interaction of the *Escherichia coli* HU protein with DNA. evidence for formation of nucleosome-like structures with altered DNA helical pitch. *J Mol Biol* **187**, 47-60.

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**, 4047-4050.

Cerdan R., Bloch V., Yang Y., Bertin P., Dumas C., Rimsky S., Kochoyan M., Arold S. T. (2003). Crystal structure of the N-terminal dimerisation domain of VicH, the H-NS-like protein of vibrio cholerae. *J Mol Biol* **334**, 179-185.

Chen C. C. & Wu H. Y. (2005). LeuO protein delimits the transcriptionally active and repressive domains on the bacterial chromosome. *J Biol Chem* **280**, 15111-15121.

Chen C. C., Chou M. Y., Huang C. H., Majumder A., Wu H. Y. (2005). A cis-spreading nucleoprotein filament is responsible for the gene silencing activity found in the promoter relay mechanism. *J Biol Chem* **280**, 5101-5112.

Coker C., Bakare O. O., Mobley H. L. (2000). H-NS is a repressor of the *Proteus mirabilis* urease transcriptional activator gene *ureR*. *J Bacteriol* **182**, 2649-2653.

Collado-Vides J., Magasanik B., Gralla J. D. (1991). Control site location and transcriptional regulation in *Escherichia coli*. *Microbiol Rev* **55**, 371-394.

Corcoran C. P. & Dorman C. J. (2009). DNA relaxation-dependent phase biasing of the *fim* genetic switch in *Escherichia coli* depends on the interplay of H-NS, IHF and LRP. *Mol Microbiol* .

Dagberg B. & Uhlin B. E. (1992). Regulation of virulence-associated plasmid genes in enteroinvasive *Escherichia coli*. *J Bacteriol* **174**, 7606-7612.

Dame R. T., Noom M. C., Wuite G. J. (2006). Bacterial chromatin organization by H-NS protein unravelled using dual DNA manipulation. *Nature* **444**, 387-390.

Dame R. T., Wyman C., Goosen N. (2000). H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res* **28**, 3504-3510.

Dame R. T., Luijsterburg M. S., Krin E., Bertin P. N., Wagner R., Wuite G. J. (2005). DNA bridging: A property shared among H-NS-like proteins. *J Bacteriol* **187**,

1845-1848.

De la Cruz M. A., Merino E., Oropeza R., Tellez J., Calva E. (2009). The DNA static curvature has a role in the regulation of the *ompS1* porin gene in *Salmonella enterica* serovar Typhi. *Microbiology* **155**, 2127-2136.

De la Cruz M. A., Fernandez-Mora M., Guadarrama C., Flores-Valdez M. A., Bustamante V. H., Vazquez A., Calva E. (2007). LeuO antagonizes H-NS and StpA-dependent repression in *Salmonella enterica ompS1*. *Mol Microbiol* **66**, 727-743.

Dersch P., Schmidt K., Bremer E. (1993). Synthesis of the *Escherichia coli* K-12 nucleoid-associated DNA-binding protein H-NS is subjected to growth-phase control and autoregulation. *Mol Microbiol* **8**, 875-889.

Donato G. M. & Kawula T. H. (1998). Enhanced binding of altered H-NS protein to flagellar rotor protein FliG causes increased flagellar rotational speed and hypermotility in *Escherichia coli*. *J Biol Chem* **273**, 24030-24036.

Dorman C. J. (2009). H-NS and genomic bridge building: Lessons from the human pathogen salmonella typhi. *Microbiology* **155**, 2114-2115.

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

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**, 587-592.

Dorman C. J. & Deighan P. (2003). Regulation of gene expression by histone-like proteins in bacteria. *Curr Opin Genet Dev* **13**, 179-184.

Dorman C. J., Hinton J. C., Free A. (1999). Domain organization and oligomerization among H-NS-like nucleoid-associated proteins in bacteria. *Trends Microbiol* **7**, 124-128.

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**, 251-252.

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**, 1063-1073.

Egile C., Loisel T. P., Laurent V., Li R., Pantaloni D., Sansonetti P. J., Carlier M. F. (1999). Activation of the CDC42 effector N-WASP by the *Shigella flexneri* IcsA protein promotes actin nucleation by Arp2/3 complex and bacterial actin-based motility. *J Cell*

Biol **146**, 1319-1332.

Ellison D. W. & Miller V. L. (2006a). H-NS represses inv transcription in *Yersinia enterocolitica* through competition with RovA and interaction with YmoA. *J Bacteriol* **188**, 5101-5112.

Ellison D. W. & Miller V. L. (2006b). Regulation of virulence by members of the MarR/SlyA family. *Curr Opin Microbiol* **9**, 153-159.

Esposito D., Petrovic A., Harris R., Ono S., Eccleston J. F., Mbabaali A., Haq I., Higgins C. F., Hinton J. C., Driscoll P. C., Ladbury J. E. (2002). H-NS oligomerization domain structure reveals the mechanism for high order self-association of the intact protein. *J Mol Biol* **324**, 841-850.

Falconi M., Prosseda G., Giangrossi M., Beghetto E., Colonna B. (2001). Involvement of FIS in the H-NS-mediated regulation of *virF* gene of *Shigella* and enteroinvasive *Escherichia coli*. *Mol Microbiol* **42**, 439-452.

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**, 7033-7043.

Fang F. C. & Rimsky S. (2008). New insights into transcriptional regulation by H-NS. *Curr Opin Microbiol* **11**, 113-120.

Fickett J. W. (1982). Recognition of protein coding regions in DNA sequences. *Nucleic Acids Res* **10**, 5303-5318.

Flores-Valdez M. A., Puente J. L., Calva E. (2003). Negative osmoregulation of the *Salmonella ompS1* porin gene independently of OmpR in an *hns* background. *J Bacteriol* **185**, 6497-6506.

Forsman K., Sonden B., Goransson M., Uhlin B. E. (1992). Antirepression function in *Escherichia coli* for the cAMP-cAMP receptor protein transcriptional activator. *Proc Natl Acad Sci U S A* **89**, 9880-9884.

Fozo E. M., Hemm M. R., Storz G. (2008). Small toxic proteins and the antisense RNAs that repress them. *Microbiol Mol Biol Rev* **72**, 579-89, Table of Contents.

Free A., Porter M. E., Deighan P., Dorman C. J. (2001). Requirement for the molecular adapter function of StpA at the *Escherichia coli bgl* promoter depends upon the level of truncated H-NS protein. *Mol Microbiol* **42**, 903-917.

Fukuda I., Suzuki T., Munakata H., Hayashi N., Katayama E., Yoshikawa M., Sasakawa C. (1995). Cleavage of *Shigella* surface protein VirG occurs at a specific site,

but the secretion is not essential for intracellular spreading. *J Bacteriol* **177**, 1719-1726.

Funnell B. E. (1988). Participation of *Escherichia coli* integration host factor in the P1 plasmid partition system. *Proc Natl Acad Sci U S A* **85**, 6657-6661.

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

Goransson M., Sonden B., Nilsson P., Dagberg B., Forsman K., Emanuelsson K., Uhlin B. E. (1990). Transcriptional silencing and thermoregulation of gene expression in *Escherichia coli*. *Nature* **344**, 682-685.

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

Hacker J. & Kaper J. B. (2000). Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* **54**, 641-679.

Hale T. L. (1991). Genetic basis of virulence in *Shigella* species. *Microbiol Rev* **55**, 206-224.

Heroven A. K., Bohme K., Tran-Winkler H., Dersch P. (2007). Regulatory elements implicated in the environmental control of *invasin* expression in enteropathogenic *Yersinia*. *Adv Exp Med Biol* **603**, 156-166.

Heroven A. K., Nagel G., Tran H. J., Parr S., Dersch P. (2004). RovA is autoregulated and antagonizes H-NS-mediated silencing of *invasin* and *rovA* expression in *Yersinia pseudotuberculosis*. *Mol Microbiol* **53**, 871-888.

Higgins C. F., Dorman C. J., Stirling D. A., Waddell L., Booth I. R., May G., Bremer E. (1988). A physiological role for DNA supercoiling in the osmotic regulation of gene expression in *S. typhimurium* and *E. coli*. *Cell* **52**, 569-584.

Hritonenko V. & Stathopoulos C. (2007). Omptin proteins: An expanding family of outer membrane proteases in gram-negative enterobacteriaceae. *Mol Membr Biol* **24**, 395-406.

Hromockyj A. E. & Maurelli A. T. (1989). Identification of an *Escherichia coli* gene homologous to *virR*, a regulator of *Shigella* virulence. *J Bacteriol* **171**, 2879-2881.

Hromockyj A. E., Tucker S. C., Maurelli A. T. (1992). Temperature regulation of *Shigella* virulence: Identification of the repressor gene *virR*, an analogue of *hns*, and partial complementation by tyrosyl transfer RNA (tRNA1(tyr)). *Mol Microbiol* **6**, 2113-

2124.

Hulbert R. R. & Taylor R. K. (2002). Mechanism of ToxT-dependent transcriptional activation at the *Vibrio cholerae* t*cpA* promoter. *J Bacteriol* **184**, 5533-5544.

Huo Y. X., Zhang Y. T., Xiao Y., Zhang X., Buck M., Kolb A., Wang Y. P. (2009). IHF-binding sites inhibit DNA loop formation and transcription initiation. *Nucleic Acids Res* **37**, 3878-3886.

Jain R., Rivera M. C., Moore J. E., Lake J. A. (2002). Horizontal gene transfer in microbial genome evolution. *Theor Popul Biol* **61**, 489-495.

Jauregui R., Abreu-Goodger C., Moreno-Hagelsieb G., Collado-Vides J., Merino E. (2003). Conservation of DNA curvature signals in regulatory regions of prokaryotic genes. *Nucleic Acids Res* **31**, 6770-6777.

Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H., Yang J.& other authors. (2002). Genome sequence of *Shigella flexneri* 2a: Insights into pathogenicity through comparison with genomes of *Escherichia coli* K12 and O157. *Nucleic Acids Res* **30**, 4432-4441.

Kim S. K. & Wang J. C. (1999). Gene silencing via protein-mediated subcellular localization of DNA. *Proc Natl Acad Sci U S A* **96**, 8557-8561.

Kong W., Weatherspoon N., Shi Y. (2008). Molecular mechanism for establishment of signal-dependent regulation in the PhoP/PhoQ system. *J Biol Chem* **283**, 16612-16621.

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**, 651-666.

Kozobay-Avraham L., Hosid S., Bolshoy A. (2004). Curvature distribution in prokaryotic genomes. *In Silico Biol* **4**, 361-375.

Kukkonen, M., and T. K. Korhonen. 2004. The omptin family of enterobacterial surface proteases/adhesins: from housekeeping in *Escherichia coli* to systemic spread of *Yersinia pestis*. Int J Med Microbiol **294:**7-14.

Lang B., Blot N., Bouffartigues E., Buckle M., Geertz M., Gualerzi C. O., Mavathur R., Muskhelishvili G., Pon C. L.& other authors. (2007). High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. *Nucleic Acids Res* **35**, 6330-6337.

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**, 951-962.

Lee H., Kotloff K., Chukaserm P., Samosornsuk S., Chompook P., Deen J. L., Von Seidlein L., Clemens J. D., Wanpen C. (2005). Shigellosis remains an important problem in children less than 5 years of age in Thailand. *Epidemiol Infect* **133**, 469-474.

Li Y., Cao B., Liu B., Liu D., Gao Q., Peng X., Wu J., Bastin D. A., Feng L., Wang L. (2009). Molecular detection of all 34 distinct O-antigen forms of *Shigella*. *J Med Microbiol* **58**, 69-81.

Lie T. J. & Leigh J. A. (2007). Genetic screen for regulatory mutations in *Methanococcus maripaludis* and its use in identification of induction-deficient mutants of the euryarchaeal repressor NrpR. *Appl Environ Microbiol* **73**, 6595-6600.

Lithgow J. K., Haider F., Roberts I. S., Green J. (2007). Alternate SlyA and H-NS nucleoprotein complexes control *hlyE* expression in *Escherichia coli* K-12. *Mol Microbiol* **66**, 685-698.

Lobocka M. & Yarmolinsky M. (1996). P1 plasmid partition: A mutational analysis of ParB. *J Mol Biol* **259**, 366-382.

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**, e81.

Luijsterburg M. S., Noom M. C., Wuite G. J., Dame R. T. (2006). The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: A molecular perspective. *J Struct Biol* **156**, 262-272.

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**, 2820-2824.

Maurelli A. T., Blackmon B., Curtiss R.,3rd. (1984a). Loss of pigmentation in *Shigella flexneri* 2a is correlated with loss of virulence and virulence-associated plasmid. *Infect Immun* **43**, 397-401.

Maurelli A. T., Blackmon B., Curtiss R.,3rd. (1984b). Temperature-dependent expression of virulence genes in *Shigella* species. *Infect Immun* **43**, 195-201.

Maurer S., Fritz J., Muskhelishvili G. (2009). A systematic in vitro study of nucleoprotein complexes formed by bacterial nucleoid-associated proteins revealing novel types of DNA organization. *J Mol Biol* **387**, 1261-1276.

McKenna S., Beloin C., Dorman C. J. (2003). In vitro DNA-binding properties of VirB, the *Shigella flexneri* virulence regulatory protein. *FEBS Lett* **545**, 183-187.

Miller J., editor. (1972). Experiments in Molecular Genetics. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Nasser W. & Reverchon S. (2002). H-NS-dependent activation of pectate lyases synthesis in the phytopathogenic bacterium *Erwinia chrysanthemi* is mediated by the PecT repressor. *Mol Microbiol* **43**, 733-748.

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**, 1456-1471.

Navarre W. W., Porwollik S., Wang Y., McClelland M., Rosen H., Libby S. J., Fang F. C. (2006). Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* **313**, 236-238.

Nicoletti M., Santino I., Petrucca A., Del Chierico F., Cannavacciuolo S., Casalino M., Sessa R., Cipriani P. (2008). Evaluation by real-time PCR of the expression of *S. flexneri* virulence-associated genes *ospB* and *phoN2* under different genetical backgrounds. *Int J Immunopathol Pharmacol* **21**, 707-714.

Nogueira T. & Springer M. (2000). Post-transcriptional control by global regulators of gene expression in bacteria. *Curr Opin Microbiol* **3**, 154-158.

Noom M. C., Navarre W. W., Oshima T., Wuite G. J., Dame R. T. (2007). H-NS promotes looped domain formation in the bacterial chromosome. *Curr Biol* **17**, R913-4.

Nye M. B. & Taylor R. K. (2003). *Vibrio cholerae* H-NS domain structure and function with respect to transcriptional repression of ToxR regulon genes reveals differences among H-NS family members. *Mol Microbiol* **50**, 427-444.

Nye M. B., Pfau J. D., Skorupski K., Taylor R. K. (2000). *Vibrio cholerae* H-NS silences virulence gene expression at multiple steps in the ToxR regulatory cascade. *J Bacteriol* **182**, 4295-4303.

Olivares-Zavaleta N., Jauregui R., Merino E. (2006). Genome analysis of *Escherichia coli* promoter sequences evidences that DNA static curvature plays a more important role in gene transcription than has previously been anticipated. *Genomics* **87**, 329-337.

Oshima T., Ishikawa S., Kurokawa K., Aiba H., Ogasawara N. (2006). *Escherichia coli* histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Res* **13**, 141-153.

Pedersen A. G., Jensen L. J., Brunak S., Staerfeldt H. H., Ussery D. W. (2000). A DNA structural atlas for *Escherichia coli*. *J Mol Biol* **299**, 907-930.

Perez J. C., Latifi T., Groisman E. A. (2008). Overcoming H-NS-mediated transcriptional silencing of horizontally acquired genes by the PhoP and SlyA proteins in *Salmonella enterica*. *J Biol Chem* **283**, 10773-10783.

Perrin S. & Gilliland G. (1990). Site-specific mutagenesis using asymmetric polymerase chain reaction and a single mutant primer. *Nucleic Acids Res* **18**, 7433-7438.

Peschel A. (2002). How do bacteria resist human antimicrobial peptides? *Trends Microbiol* **10**, 179-186.

Pogliano J. A. & Beckwith J. (1994). SecD and SecF facilitate protein export in *Escherichia coli*. *EMBO J* **13**, 554-561.

Poore C. A. & Mobley H. L. (2003). Differential regulation of the *Proteus mirabilis* urease gene cluster by UreR and H-NS. *Microbiology* **149**, 3383-3394.

Porter M. E. & Dorman C. J. (1994). A role for H-NS in the thermo-osmotic regulation of virulence gene expression in *Shigella flexneri*. *J Bacteriol* **176**, 4187-4191.

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**, 303-309.

Prosseda G., Mazzola A., Di Martino M. L., Tielker D., Micheli G., Colonna B. (2010). A temperature-induced narrow DNA curvature range sustains the maximum activity of a bacterial promoter *in vitro*. *Biochemistry* **49**, 2778-2785.

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**, 523-537.

Prosseda G., Fradiani P. A., Di Lorenzo M., Falconi M., Micheli G., Casalino M., Nicoletti M., Colonna B. (1998). A role for H-NS in the regulation of the *virF* gene of *Shigella* and enteroinvasive *Escherichia coli*. *Res Microbiol* **149**, 15-25.

Reitzer L. J. & Magasanik B. (1986). Transcription of *glnA in E. coli* is stimulated by activator bound to sites far from the promoter. *Cell* **45**, 785-792.

Rimsky S. (2004). Structure of the histone-like protein H-NS and its role in regulation and genome superstructure. *Curr Opin Microbiol* **7**, 109-114.

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**, 1311-1323.

Rine J. (1999). On the mechanism of silencing in *Escherichia coli*. *Proc Natl Acad Sci U S A* **96**, 8309-8311.

Rodionov O. & Yarmolinsky M. (2004). Plasmid partitioning and the spreading of P1 partition protein ParB. *Mol Microbiol* **52**, 1215-1223.

Rodionov O., Lobocka M., Yarmolinsky M. (1999). Silencing of genes flanking the P1 plasmid centromere. *Science* **283**, 546-549.

Rohde J. R., Luan X. S., Rohde H., Fox J. M., Minnich S. A. (1999). The *Yersinia enterocolitica* pYV virulence plasmid contains multiple intrinsic DNA bends which melt at 37 degrees C. *J Bacteriol* **181**, 4198-4204.

Sakai, T., C. Sasakawa, and M. Yoshikawa. (1988). Expression of four virulence antigens of *Shigella flexneri* is positively regulated at the transcriptional level by the 30 kiloDalton *virF* protein. Mol Microbiol **2:**589-97.

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

Schnetz K. (1995). Silencing of *Escherichia coli bgl* promoter by flanking sequence elements. *EMBO J* **14**, 2545-2550.

Schuch R., Sandlin R. C., Maurelli A. T. (1999). A system for identifying post-invasion functions of invasion genes: Requirements for the *mxi-spa* type III secretion pathway of *Shigella flexneri* in intercellular dissemination. *Mol Microbiol* **34**, 675-689.

Schumacher M. A. & Funnell B. E. (2005). Structures of ParB bound to DNA reveal mechanism of partition complex formation. *Nature* **438**, 516-519.

Schumacher M. A., Mansoor A., Funnell B. E. (2007). Structure of a four-way bridged ParB-DNA complex provides insight into P1 segrosome assembly. *J Biol Chem* **282**, 10456-10464.

Shere, K. D., S. Sallustio, A. Manessis, T. G. D'Aversa, and M. B. Goldberg. (1997). Disruption of IcsP, the major *Shigella* protease that cleaves IcsA, accelerates actin-based motility. Mol Microbiol **25:**451-62.

Shi Y., Latifi T., Cromie M. J., Groisman E. A. (2004). Transcriptional control of the antimicrobial peptide resistance *ugtL* gene by the *Salmonella* PhoP and SlyA regulatory proteins. *J Biol Chem* **279**, 38618-38625.

Smyth C. P., Lundback T., Renzoni D., Siligardi G., Beavil R., Layton M., Sidebotham J. M., Hinton J. C., Driscoll P. C., Higgins C. F., Ladbury J. E. (2000).

Oligomerization of the chromatin-structuring protein H-NS. *Mol Microbiol* **36**, 962-972.

Steinhauer, J., R. Agha, T. Pham, A. W. Varga, and M. B. Goldberg. (1999). The unipolar *Shigella* surface protein IcsA is targeted directly to the bacterial old pole: IcsP cleavage of IcsA occurs over the entire bacterial surface. Mol Microbiol **32:**367-77.

Soutourina O., Kolb A., Krin E., Laurent-Winter C., Rimsky S., Danchin A., Bertin P. (1999). Multiple control of flagellum biosynthesis in *Escherichia coli*: Role of H-NS protein and the cyclic AMP-catabolite activator protein complex in transcription of the *flhDC* master operon. *J Bacteriol* **181**, 7500-7508.

Spassky A., Rimsky S., Garreau H., Buc H. (1984). H1a, an *E. coli* DNA-binding protein which accumulates in stationary phase, strongly compacts DNA in vitro. *Nucleic Acids Res* **12**, 5321-5340.

Spurio R., Falconi M., Brandi A., Pon C. L., Gualerzi C. O. (1997). The oligomeric structure of nucleoid protein H-NS is necessary for recognition of intrinsically curved DNA and for DNA bending. *EMBO J* **16**, 1795-1805.

Stapleton M. R., Norte V. A., Read R. C., Green J. (2002). Interaction of the *Salmonella typhimurium* transcription and virulence factor SlyA with target DNA and identification of members of the SlyA regulon. *J Biol Chem* **277**, 17630-17637.

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**, 2533-2545.

Surtees J. A. & Funnell B. E. (2001). The DNA binding domains of P1 ParB and the architecture of the P1 plasmid partition complex. *J Biol Chem* **276**, 12385-12394.

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**, 5158-5165.

Tendeng C. & Bertin P. N. (2003). H-NS in gram-negative bacteria: A family of multifaceted proteins. *Trends Microbiol* **11**, 511-518.

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**, 1094-1097.

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**, 6142-6149.

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**, 887-893.

Tupper A. E., Owen-Hughes T. A., Ussery D. W., Santos D. S., Ferguson D. J., Sidebotham J. M., Hinton J. C., Higgins C. F. (1994). The chromatin-associated protein H-NS alters DNA topology in vitro. *EMBO J* **13**, 258-268.

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**, 3403-3413.

Ueguchi C., Suzuki T., Yoshida T., Tanaka K., Mizuno T. (1996). Systematic mutational analysis revealing the functional domain organization of *Escherichia coli* nucleoid protein H-NS. *J Mol Biol* **263**, 149-162.

Ussery D. W., Higgins C. F., Bolshoy A. (1999). Environmental influences on DNA curvature. *J Biomol Struct Dyn* **16**, 811-823.

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**, 10944-10952.

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**, 619-629.

Westermark M., Oscarsson J., Mizunoe Y., Urbonaviciene J., Uhlin B. E. (2000). Silencing and activation of ClyA cytotoxin expression in *Escherichia coli*. *J Bacteriol* **182**, 6347-6357.

Will W. R. & Frost L. S. (2006). Characterization of the opposing roles of H-NS and TraJ in transcriptional regulation of the F-plasmid *tra* operon. *J Bacteriol* **188**, 507-514.

Williams R. M. & Rimsky S. (1997). Molecular aspects of the *E. coli* nucleoid protein, H-NS: A central controller of gene regulatory networks. *FEMS Microbiol Lett* **156**, 175-185.

Williams R. M., Rimsky S., Buc H. (1996). Probing the structure, function, and interactions of the *Escherichia coli* H-NS and StpA proteins by using dominant negative derivatives. *J Bacteriol* **178**, 4335-4343.

Williamson H. S. & Free A. (2005). A truncated H-NS-like protein from enteropathogenic *Escherichia coli* acts as an H-NS antagonist. *Mol Microbiol* **55**, 808-

827.

Wing H. J., Yan A. W., Goldman S. R., Goldberg M. B. (2004). Regulation of *IcsP*, the outer membrane protease of the *Shigella* actin tail assembly protein IcsA, by virulence plasmid regulators VirF and VirB. *J Bacteriol* **186**, 699-705.

Yamada H., Muramatsu S., Mizuno T. (1990). An *Escherichia coli* protein that preferentially binds to sharply curved DNA. *J Biochem* **108**, 420-425.

Yarmolinsky M. (2000). Transcriptional silencing in bacteria. *Curr Opin Microbiol* **3**, 138-143.

Zhang A., Rimsky S., Reaban M. E., Buc H., Belfort M. (1996). *Escherichia coli* protein analogs StpA and H-NS: Regulatory loops, similar and disparate effects on nucleic acid dynamics. *EMBO J* **15**, 1340-1349.

Zuber F., Kotlarz D., Rimsky S., Buc H. (1994). Modulated expression of promoters containing upstream curved DNA sequences by the *Escherichia coli* nucleoid protein H-NS. *Mol Microbiol* **12**, 231-240.

VITA

Graduate College
University of Nevada, Las Vegas

Lieutenant Dustin John Harrison
Medical Service Corps
U.S. Navy

Degrees:

Bachelor of Science, Biology, 1997
Marycrest International University

Master of Science, Pathobiology, 2000
University of Wyoming

Publications: 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**, 4047-4050.

Dissertation Title: Transcriptional Regulation of the *Shigella flexneri icsp* Promoter: Silencing and Anti-silencing by H-NS and VirB

Dissertation Examination Committee:

Chairperson, Helen Wing, Ph.D.
Committee Member Dennis Bazylinski, Ph.D.
Committee Member Eduardo Robleto, Ph.D.
Graduate Faculty Representative Patricia Cruz, Ph.D.
Outside Committee Member CDR Marshall Monteville, USN, Ph.D.