The Genetic Organization & Transcriptional Regulation of *Shigella* Virulence Genes

NATASHA WEATHERSPOON-GRIFFIN, MICHAEL A. PICKER & HELEN J. WING

Running Title Transcriptional control of virulence in *Shigella*

Abstract

Central to bacterial pathogenicity is the precise and coordinated control of virulence gene expression in response to environmental cues encountered in the human host. This chapter focuses on the transcriptional regulation of *Shigella* virulence genes encoded by the large virulence plasmid, pINV, found in almost all *Shigella* species. We describe the silencing of pINV-encoded virulence genes by the chromosomally encoded nucleoid structuring protein H-NS, which serves as the backdrop for all other regulatory events on this plasmid. We then describe the four-tiered virulence regulatory cascade emphasizing the environmental stimuli, transcriptional regulators and mechanisms that govern virulence gene expression at each tier. Finally, we review other regulatory inputs that are received from the *Shigella* chromosome and their role in the modulation of virulence gene expression. We begin our discussion by describing the genetic organization of *Shigella*, its relatedness to *Escherichia coli* and the different forms of pINV carried by *Shigella* species to highlight how the regulatory cascades controlling *Shigella* virulence are, in large part, conserved across species. Overall, this chapter reveals that *Shigella* is a fascinating model for the study of virulence gene regulation, which promotes our understanding of *Shigella* pathogenesis and mechanisms of virulence gene regulation in other bacterial pathogens.

I. Introduction

Shigella species embark on a remarkable journey to reach the colonic epithelium, their preferred site of invasion in their human and primate hosts. *Shigella* infections are spread by the fecal-oral route and are generally acquired directly by ingestion of contaminated food or water, or indirectly via person-to-person contact. As Shigellae pass through the mouth, stomach and upper and lower intestines, they encounter antibacterial compounds that form part of our immune system and a wide variety of environmental conditions that threaten bacterial survival. Despite these challenges, reports indicate that the infectious dose of *Shigella* is extremely low (ID_{50} of 10-500; (DuPont et al., 1989)) compared to other enteric pathogens (eg. $ID_{50}>10^5$ for *Salmonella* spp.), demonstrating that *Shigella* species are particularly adept at surviving their journey through the gastrointestinal tract and causing an infection. This observation strongly suggests that *Shigella* strains continually monitor their environment and alter their gene expression to allow them to tolerate, resist and/or exploit conditions encountered in the human body, and indeed all evidence suggests that this is the case.

While it is challenging to devise experiments that accurately measure gene expression within the human host (reviewed in (Marteyn et al., 2012)), careful research has identified some of the most important cues that are likely sensed *in vivo* and how they are responded to by modulation of gene expression. Some of these cues include changes in temperature (Maurelli et al., 1984), osmolarity (Porter and Dorman, 1994) and pH (Nakayama and Watanabe, 1995), the change in freely available iron and oxygen (Marteyn et al., 2010; Murphy and Payne, 2007) and even the activation of the *Shigella* type III secretion system once direct contact with epithelial cells has been made (Demers et al., 1998; Kane et al., 2002; Mavris et al., 2002a). It is reasonable to expect that the bacterial response to these cues directly correlates with the temporo-spatial expression of specific virulence factors required for infection. And, although difficult to assess, it is even conceivable that the order and duration of these cues may help shape the physiology of *Shigella* so that this bacterium is optimally programmed for its life within the human host.

If we are to more thoroughly understand the molecular pathogenesis of *Shigella* spp., a more complete understanding of the environmental cues and genetic networks controlling gene expression in these pathogens is required. In this chapter, our primary focus is the regulatory events that control the transcription of virulence genes carried by the large virulence plasmids found in *Shigella* species. We describe regulatory processes that have allowed the acquisition and maintenance of the virulence plasmid, the regulatory cascade that triggers virulence gene expression on the large plasmid and additional inputs received via some chromosomally encoded regulators. But, we first describe the genetic organization of *Shigella* species, so that the reader can better appreciate how our discussion relates to the different strains and species of *Shigella*.

II. Classification & genetic organization of *Shigella*

In the 1940's, *Shigella* strains were classified as a unique genus because of their pathogenicity and then divided into four species: (1) *S. dysenteriae;* (2) *S. boydii*; (3) *S. sonnei;* and (4) *S. flexneri* (Bensted, 1956; Ewing, 1949, 1986). As early as 1972, however, it was estimated that *Shigella* and *Escherichia coli* were

taxonomically indistinguishable at the species level based on DNA hybridization studies (Brenner et al., 1972). More recently, a variety of studies have confirmed that all *Shigella* strains, except possibly *S. boydii* serotype 13, are close phylogenetic relatives of *E. coli* (Ochman et al., 1983; Pupo et al., 1997; Pupo et al., 2000; Rolland et al., 1998). Consequently, the *Shigella* genus is considered a classic example of a "taxon in disguise," meaning that, although originally classified as a separate genus, its evolutionary history reveals that it evolved from another taxonomic unit (Rediers et al., 2004), in this case *E. coli,* which makes *E. coli* paraphyletic to *Shigella* (Lan and Reeves, 2002).

The chromosome of the *Shigella* spp. is approximately the same size (4.3-4.9 Mb) as the *E. coli* K-12 chromosome (4.6 Mb) and is typically more than 97% identical at the DNA sequence level. Genome sequencing, however, has revealed several unique genetic islands on the *Shigella* chromosome, as well as several *E. coli* genes missing from the *Shigella* chromosome (described in depth for *S. flexneri* 2a; (Wei et al., 2003)). Additional analyses reveal that the closest relatives to the four *Shigella* "species" are the enteroinvasive *E. coli* (EIEC) (Lan et al., 2004; Lan and Reeves, 2002; Pupo et al., 1997; Pupo et al., 2000). The primary reason for the clustering of EIEC with the *Shigella* species is the presence of a common 220 kb plasmid (Lan and Reeves, 2002). This large virulence plasmid is essential for the invasive phenotype of these bacteria, and has collectively been given the name pINV, although specific names are given to pINV plasmids residing in different strains (eg. pWR100 of *S. flexneri* serotype 5, pCP301 of *S. flexneri* 2a and pSS120 of *S. sonnei*). Two major forms of the virulence plasmid, referred to as pINV A and pINV B, have been identified based on the DNA sequence of three genes encoded by pINV (*ipgD, mxiC* and *mxiA*) (Lan and Reeves, 2002). Interestingly, the distribution of the two forms correlates well with the variation in chromosomal housekeeping genes (reviewed in (Lan et al., 2004; Lan and Reeves, 2002)). All *Shigella* spp. contain either form of the plasmid, with the exception of *S. dysenteriae* serotype 1, which contains a recombinant plasmid with a divergent sequence not corresponding to these two forms, and the *S. boydii* serotype 13 strain, which does not contain a plasmid (Table 1). All but two EIEC strains analyzed so far carry pINV A (Lan et al., 2004; Pupo et al., 2000).

When the coding regions (those not associated with insertion sequences) of pINV were compared across three *S. flexneri* strains carrying either pINV A or pINV B, numerous polymorphisms were found in many of the virulence genes encoding effector proteins (Lan et al., 2003). These included genes encoding proteins secreted by the *Shigella* type III secretion system (eg. *ipaA, ipaH9.8*, and *ipaD*) and those that enable *Shigella* to move by actin-based motility within the host cell cytoplasm (eg. *icsA* and *virK*) (Lan et al., 2003). The polymorphisms found in these genes frequently led to amino acid substitutions when pINV A was compared to pINV B, but were silent when found on separate examples of the same form of plasmid (Lan et al., 2003). Significantly though, the genes encoding transcriptional regulators encoded by pINV (*virF, virB, mxiE* and *ipgC*) were found to display far fewer polymorphisms between pINV A and pINV B, and those that were identified were silent mutations (Lan et al., 2003). This strong conservation of the virulence gene regulators strongly suggests that they function in a similar manner in the vast majority of *Shigella* strains regardless of whether they are encoded by pINV A or pINV B. Although, it should be remembered that polymorphisms in DNA binding sites recognized by these regulators may lead to different patterns of gene regulation from one strain to another.

In addition to pINV, some *Shigella* strains bear additional plasmids. For instance, during the sequencing of *S. flexneri* serotype 2a genome, strain 2457T, two additional small multi-copy plasmids and a 165 kb plasmid called pSf-R27 were found (Wei et al., 2003). The largest plasmid, which resembled the R27 plasmid typically associated with *S. enterica* serovar Typhi, has not been found in any other *S. flexneri* strains sequenced to date, raising the possibility that this plasmid was either lost from all other *S. flexneri* strains or was acquired shortly after this particular strain was isolated (Wei et al., 2003). Later in this chapter, the impact of a regulator (Sfh) encoded by this plasmid will be discussed in relation to its regulation of virulence genes in 2457T (Beloin et al., 2003).

Despite the presence of pINV, virulence-associated proteins in *Shigella* spp. are also encoded by its chromosome. In general, these genes reside within pathogenicity islands (SHI-1, SHI-2, SHI-3, SHI-O, and SRL) and encode proteins required for adaptation and maintenance within a host (eg. iron acquisition and antibiotic resistance). In contrast, the genes located on pINV typically encode proteins required for invasion of host cells and intercellular spread (reviewed in (Schroeder and Hilbi, 2008)). Since efficient invasion and infection of *Shigella* spp. requires the timely and coordinated expression of pINV genes, a complex regulatory network involving regulators encoded by both the chromosome and the virulence plasmid has evolved to control the transcription of pINV genes.

The acquisition of any plasmid by a bacterial strain poses a problem of regulation (discussed in (Dorman and Porter, 1998)) because the newly acquired plasmid genes need to be woven into the existing regulatory networks found within the bacterial genome. In the case of the acquisition of pINV, it appears that the genes encoding regulators specific to pINV were also encoded by the plasmid. But surprisingly, the role of at least some of these regulators appears to have been the relief of transcriptional silencing (repression) mediated by chromosomally encoded nucleoid structuring proteins, which allowed the acquisition and subsequent incorporation of plasmid genes into the already present transcriptional control networks. Of course with additional time and co-evolution of the plasmid and the chromosome, additional chromosomal inputs are likely to have evolved to help control the expression of the *Shigella* virulence plasmid genes, as is observed today.

Having discussed the genetic composition of *Shigella* strains, our description of the transcriptional regulation of virulence genes can begin. Owing to the complexity of the regulatory system in *Shigella*, this chapter will focus solely on the transcriptional regulation of virulence genes encoded by pINV. We will: (1) describe the silencing of genes on the pINV plasmid by chromosomally encoded and accessory plasmid-encoded nucleoid structuring proteins; (2) provide an overview of the regulatory cascade that triggers virulence gene expression on pINV; (3) describe each of the pINV-encoded regulators that belong to this regulatory cascade, including events that trigger their expression, the proteins themselves and the effect that each has on their immediate downstream targets; and finally (4) discuss additional regulatory inputs made by chromosomally encoded regulators.

III. Transcriptional silencing of virulence genes encoded by pINV.

The virulence genes encoded by pINV remain transcriptionally silent under non-physiological conditions (osmolarities lower than physiological, pH below 7.4 and temperatures below 37°C) (Hromockyj et al.,

1992; Maurelli et al., 1984; Maurelli and Sansonetti, 1988; Porter and Dorman, 1994, 1997a). This is advantageous for *Shigella* because inappropriate expression of these genes has been shown to reduce fitness of the organism *ex vivo* (Dorman, 2007; Schuch and Maurelli, 1997). In *Shigella,* like many other enteric bacteria (Lucchini et al., 2006; Navarre et al., 2005), a nucleoid structuring protein called H-NS (histone-like nucleoid structuring protein) prevents the expression of genes on pINV in a process that has been coined xenogeneic silencing ((Hromockyj et al., 1992; Maurelli and Sansonetti, 1988) & reviewed in (Stoebel et al., 2008)). As mentioned before, this silencing is thought to have played an important role in the acquisition and maintenance of pINV (Stoebel et al., 2008), but also provides the context in which all other virulence gene regulation on pINV occurs .

Transcriptional silencing of virulence genes on pINV occurs at multiple levels within the regulatory cascade controlling *Shigella* virulence. Under non-physiological conditions, H-NS directly represses *virF* and *virB,* which encode key transcriptional factors required for the upregulation of pINV genes (see below; (Adler et al., 1989; Falconi et al., 1998; Prosseda et al., 2004; Prosseda et al., 1998; Tobe et al., 1991; Tobe et al., 1993), but H-NS also directly represses genes that lie within the VirF and VirB regulons (Beloin and Dorman, 2003; Hromockyj et al., 1992; Porter and Dorman, 1994; Prosseda et al., 2004; Wing et al., 2004). The intimate role that H-NS plays in virulence gene regulation in *Shigella* is further demonstrated by the observation that VirB and, in some cases, VirF function to counteract H-NS-mediated repression rather than functioning as classical activators of transcription (Le Gall et al., 2005; Stoebel et al., 2008; Wing et al., 2004). For this reason, we start our discussion of the transcriptional regulation of virulence genes on pINV by describing the H-NS protein and what is known about its silencing of gene expression. Since *Shigella* strains harbor one (StpA) or sometimes two H-NS paralogues (StpA and Sfh), for the sake of completeness these proteins are also discussed because under certain circumstances these protein can influence virulence gene regulation in *Shigella* (Beloin et al., 2003; Deighan et al., 2003), and it is possible that these proteins play a direct, yet so far poorly understood, role in the transcriptional silencing of virulence genes residing on pINV.

H-NS

H-NS (also referred to as H1 and H1a) was first described in the early 1970s (Cukier-Kahn et al., 1972; Jacquet et al., 1971). Its ability to modulate DNA supercoiling and condense DNA *in vivo* and *in vitro* in a manner similar to eukaryotic histones (Spassky et al., 1984; Spurio et al., 1992) led to its name – the histone-like nucleoid structuring protein, H-NS (Lammi et al., 1984). H-NS, however, does not share sequence homology with histones, so more recently, it has been referred to as a nucleoid associated or nucleoid structuring protein (NAP or NSP) (Atlung and Ingmer, 1997).

H-NS is commonly found throughout the Enterobacteriaceae (Tendeng and Bertin, 2003). While its characterization has primarily been done in *E. coli* K-12 strains, the H-NS protein encoded by *Shigella* spp. is 100% identical at the amino acid level to that produced by *E. coli*. Furthermore, the *hns* genetic locus is also very highly conserved (99.9%; from the beginning of the *galU* gene to the end of the *tdk* gene, which flank *hns*), suggesting that the expression profiles of *hns* in *E. coli* and *Shigella* are similar. H-NS is a 15 kDa, chromosomally encoded protein that is highly expressed in exponential growth (Deighan et al., 2003; Free and Dorman, 1995), accumulating to approximately 20,000 monomers per cell in

stationary phase (Spassky et al., 1984). H-NS can bind to RNA, but it is primarily a DNA binding protein with a preference for curved DNA, which is commonly AT-rich (Brescia et al., 2004; Dame et al., 2001; Prosseda et al., 2004; Tolstorukov et al., 2005; Yamada et al., 1990). In the γ-Proteobacteria, AT-rich DNA sequences are often a hallmark of foreign genes acquired by lateral or horizontal gene transfer (Ochman et al., 2000). Consequently, the newly acquired AT-rich virulence genes located either within chromosomal pathogenicity islands and/or on virulence plasmids of many important pathogens, including *Shigella* (reviewed in (Stoebel et al., 2008)), are often negatively regulated or silenced by H-NS. While a high affinity DNA binding site for H-NS has been proposed (Bouffartigues et al., 2007; Lang et al., 2007), this site cannot always be found in the vicinity of genes repressed by H-NS (Basta et al., 2013; Gordon et al., 2011). This is most likely because H-NS binds to remotely located, high affinity sites and subsequently spreads along the DNA via interactions with the minor groove to occupy the promoter region (Fang and Rimsky, 2008; Gordon et al., 2011).

The H-NS protein consists of two structured domains which are separated by a flexible linker (Dorman, 2004; Tendeng and Bertin, 2003). The N-terminal domain contains the major dimerization domain and a region required for oligomerization, while the purified C-terminal domain confers DNA binding (Arold et al., 2010; Dorman et al., 1999; Sette et al., 2009; Williams and Rimsky, 1997; Williams et al., 1996; Wolf et al., 2006). While H-NS is primarily a dimer in solution (Bloch et al., 2003), its ability to form higher order oligomers on DNA (Arold et al., 2010; Badaut et al., 2002; Esposito et al., 2002; Renzoni et al., 2001), has the potential to lead to the formation of a variety of H-NS:DNA complexes (Amit et al., 2003; Dame et al., 2000; Lim et al., 2014; Lim et al., 2012; Maurer et al., 2009).

So far, two DNA binding modes of H-NS have been implicated in the silencing of transcription (Figure 1) – a rigid nucleoprotein filament which stiffens the DNA (Amit et al., 2003; Lim et al., 2012) and a DNA-bridging mode (Dame et al., 2005; Dame et al., 2000) where multiple H-NS dimers bind to and crosslink two distal regions of DNA, a process which is proposed to facilitate plectoneme formation (Arold et al., 2010; Rimsky and Travers, 2011). The switch between the nucleoprotein filament and the DNA bridging molecule is poorly understood *in vivo,* but can be controlled by the concentration of divalent cations *in vitro* (Liu et al., 2010). Classically nucleoprotein filaments have been proposed to prevent transcription by either blocking the access of RNA polymerase to essential promoter elements or functioning as a roadblock in the path of an elongating RNA polymerase (reviewed in (Lim et al., 2012); Figure 1a & b). In contrast, DNA bridging by H-NS has classically been proposed to either occlude or trap RNA polymerase at or from target gene promoters ((Dame et al., 2002; Lang et al., 2007); Figure 1d & e). Examples of each of these DNA binding modes impacting gene silencing and DNA packaging have been described (Dame et al., 2002; Lang et al., 2007; Lim et al., 2012; Walthers et al., 2011). More recently, however, a more sophisticated view of gene silencing by these two DNA binding modes has been proposed (Lim et al., 2014). In the first, the formation of a rigid H-NS:DNA filament downstream of an active promoter is proposed to increase the level of torque within the DNA helix, causing RNA polymerase to stall during elongation ((Lim et al., 2014); Figure 1c)). In the second, the diffusion of supercoils generated during transcription elongation (Liu and Wang, 1987) is blocked by plectonemes (Ma et al., 2013) that form as a result of H-NS:DNA bridging, which in turn, leads to the transcriptional silencing of a nearby promoter (Figure 1f) because RNA polymerase is inhibited by the newly supercoiled state of the promoter. These

scenarios are interesting because if levels of DNA supercoiling can be indirectly modulated or constrained by H-NS, it is tempting to speculate that transcription factors and/or environmental factors that counter these changes may allow the derepression of H-NS-repressed virulence genes.

Although it is clear that H-NS plays a major role in the silencing of *Shigella* virulence genes encoded by pINV, the complete picture of where H-NS binds on pINV, how its binding pattern is remodeled, what triggers this remodeling and how this impacts gene expression remains poorly understood. Furthermore, even though six mechanisms of H-NS-mediated repression have now been proposed ((Dame et al., 2002; Lang et al., 2007; Lim et al., 2012); Figure 1a-f), for the most part it remains unclear which mechanisms occur where, or how the resulting repression is relieved to facilitate virulence gene expression in *Shigella*. These topics are the focus of ongoing investigations.

Impact of H-NS paralogues StpA & Sfh

Studies of H-NS-mediated repression of genes encoded by pINV is complicated by the fact that all *Shigella* strains carry the chromosomally encoded, paralogous gene *stpA* (Zhang and Belfort, 1992; Zhang et al., 1996) and that *S. flexneri* strain 2457T encodes another *hns* paralogue called *sfh* on an R27-like accessory plasmid found exclusively in this strain (Beloin et al., 2003; Deighan et al., 2003). Both StpA and Sfh share significant amino acid homology with H-NS (58% and 59% identity, respectively) (Beloin et al., 2003). The domain structure of these three proteins is very similar (Deighan et al., 2003) and to some extent, they appear to share overlapping functional activities with H-NS (discussed more below) (Free et al., 1998; Johansson et al., 1998; Williams et al., 1996; Zhang and Belfort, 1992; Zhang et al., 1996). In addition, these proteins not only auto-repress their own promoters, but also repress the promoters of each other (Deighan et al., 2003). Taken together, these observations have led some to propose that StpA and Sfh function, at least in part, as a back-up system for H-NS (Beloin et al., 2003; Wolf et al., 2006; Zhang et al., 1996).

The negative cross-regulation and the apparent functional redundancy of H-NS and its paralogues have the potential to impact studies of *Shigella* virulence gene expression under certain circumstances. Clean deletions of *hns* can lead to the upregulation of *stpA* and *sfh* (Deighan et al., 2003), which encode proteins that can functionally substitute for H-NS (Beloin et al., 2003; Deighan et al., 2003), thereby masking the effect of the *hns* deletion. To circumvent some of these problems, dominant negative alleles, rather than clean deletions of *hns* are commonly used by those studying virulence gene expression in *Shigella* (Williams et al., 1996; Wing et al., 2004; Wolf et al., 2006; Yamada et al., 1991). This approach relies on the observation that H-NS, StpA and Sfh not only form homodimers, but can also heterodimerize with each other (Beloin et al., 2003; Deighan et al., 2003). The most commonly used dominant negative forms of H-NS or its paralogues are truncated proteins which consist solely of the N-terminal dimerization domain of the protein (Williams et al., 1996; Wing et al., 2004; Yamada et al., 1991). These proteins retain their ability to heterodimerize with the other H-NS paralogues, thereby significantly hindering the ability of the other protein(s) to bind to DNA and functionally substitute for H-NS. The result is a significant loss of the paralogues' ability to repress transcription. Although it is common for cells bearing dominant negative alleles of *hns* to grow more slowly than wild type cells, their growth is far less impacted than cells that lack all *hns* paralogues (Beloin et al., 2003), which are

prohibitively sick. Therefore, even though dominant negative alleles of *hns* can aid studies of H-NS-mediated repression of virulence genes in *Shigella,* it is imperative to interpret data collected carefully and to fully understand the nature of the *hns* mutant in use. Similarly, the tendency for this family of proteins to form heterodimers means that additional care needs to be taken when using inducible systems to express *hns* or its paralogues at super-physiological concentrations, since heterodimer formation may lead to anomalous and/or non-physiological results.

Although functional similarities between the H-NS paralogues exist, these proteins are not completely redundant with one another. For instance, the expression patterns of the three paralogues differ from one another (Deighan et al., 2003; Free and Dorman, 1995, 1997). While H-NS levels remains fairly constant throughout growth in rich broth, StpA levels increase during mid-exponential growth phase and then drop precipitously, whereas Sfh levels increase by 2.5-fold as the cells enter stationary phase (Deighan et al., 2003). The relative abundance of each of these proteins in the cell has the potential to influence the gene regulation mediated by these NSPs, although this possibility has yet to be explored. In addition, at certain genetic loci the regulatory activities of H-NS and StpA differ from one another (Deighan et al., 2000; Suzuki et al., 1996). For example, both StpA and H-NS regulate OmpF porin expression by downregulating the sRNA *micF*, however H-NS negatively regulates the transcription of *micF,* whereas StpA primarily destabilizes *micF* RNA (Deighan et al., 2000). Thus, it is clear that these proteins can serve independent and distinct roles as regulators, which makes sense given their widespread conservation in the genomes of *Shigella* and other enteric species. Finally, since all three NSPs have been shown to form heteromeric complexes with each other (Deighan et al., 2003), it is possible that each of the different heterodimer combinations have different activities, further increasing the potential impact that NSPs play in regulation of *Shigella* virulence genes (Cusick and Belfort, 1998; Deighan et al., 2003; Dorman et al., 1999; Free et al., 1998; Williams et al., 1996). Clearly, there is still much to be learned about the impact that H-NS, StpA and Sfh have on the virulence genes encoded by pINV, but despite this, we do know that a combination of environmental cues (Falconi et al., 1998; Maurelli et al., 1984; Maurelli and Sansonetti, 1988; Porter and Dorman, 1994) and the regulatory cascade controlling virulence gene expression on pINV is required to overcome transcriptional silencing mediated by H-NS ((Beloin and Dorman, 2003; Dorman and Porter, 1998); described next; Figure 2).

IV. Transcriptional control mediated by the pINV-encoded regulators

a. An overview of the pINV-encoded regulatory cascade

The regulatory cascade that controls virulence gene expression on pINV is encoded by pINV and is hierarchical in nature (depicted in Figure 2). One of the major roles that this regulatory cascade plays is to counteract transcriptional silencing of pINV genes mediated by H-NS. Each level of the hierarchy (referred to as “tiers” in this discussion) responds to a set of environmental cues encountered in the human host (Dorman and Porter, 1998; Le Gall et al., 2005). It has been proposed that each tier is required to prevent the untimely expression of virulence determinants, which is metabolically expensive and may negatively influence the competitive fitness of the organism (Lucchini et al., 2006; Porter and Dorman, 1997a).

In tier 1, the cascade begins with activation of the *virF* promoter (described in more detail in later sections (Adler et al., 1989; Falconi et al., 1998; Falconi et al., 2001; Prosseda et al., 2004; Prosseda et al., 1998)). As determined by northern analyses, expression of *virF* can be induced when grown in standard LB medium (Miller, 1972) by the following physiological cues: temperature (37°C, when compared to cells grown at 30°C), pH (7.4, when compared to cells grown at pH 6.0) or osmolarity (physiological i.e. LB, when compared to cells grown in LO, a lower osmolarity growth medium); with maximal expression occurring when each of these environmental cues are present (Porter and Dorman, 1997a).

In tier 2, the VirF protein binds its downstream targets to upregulate the transcription of *virB* (Jost and Adler, 1993; Tobe et al., 1991; Tobe et al., 1993) and *icsA* ((Bernardini et al., 1989; Lett et al., 1989; Sakai et al., 1988; Tran et al., 2011); Figure 2). Unlike *virF, virB* expression can be induced by 37°C and physiological osmolarity, but is not influenced by changes in pH, based on northern analyses (Porter and Dorman, 1997a).

In tier 3, the tier 2 gene product VirB binds to multiple downstream targets (tier 3) to counteract H-NS mediated transcriptional repression of various virulence-associated genes ((Adler et al., 1989; Dorman et al., 2001; Tobe et al., 1991; Turner and Dorman, 2007; Uchiya et al., 1995; Wing et al., 2004); Figure 2)). Many of the genes regulated by VirB lie within the 30 kb *ipa-mxi-spa* pathogenicity island while others are found elsewhere on pINV (Castellanos et al., 2009; Santapaola et al., 2002; Uchiya et al., 1995; Wing et al., 2004). Contrary to *virF* and *virB*, tier 3 genes in the *ipa-mxi-spa* locus are only thermally induced (Porter and Dorman, 1997a); they neither respond to changes in pH nor osmolarity, based on northern analyses (Porter and Dorman, 1997a). The VirB protein is also involved in a positive regulatory feedback loop at both the *virF* and *virB* promoters ((Kane and Dorman, 2012); Figure 2). This finding demonstrates that VirB is not only capable of directly regulating genes in tier 3 and indirectly regulating genes in tier 4 via its control of *mxiE* expression, but is capable of upregulating the first two tiers of the regulatory cascade too.

In tier 4, MxiE and its co-activator, IpgC, regulate the transcription of genes encoding additional type III effector proteins and other post-invasion functions (Kane et al., 2002; Le Gall et al., 2005; Mavris et al., 2002a; Mavris et al., 2002b). Interestingly, not all tier 4 genes are located on pINV; some are located on the *Shigella* chromosome (Bongrand et al., 2012; Le Gall et al., 2005; Mavris et al., 2002b). Unlike tiers 1-3, which respond to environmental signals, tier 4 genes are expressed once the type III secretion system is activated (Demers et al., 1998; Mavris et al., 2002a; Mavris et al., 2002b). Since *mxiE* mutants are still invasive in tissue culture experiments (Kane et al., 2002), MxiE-regulated genes appear to encode a set of secreted proteins that are needed for post-invasion events related to virulence, including those that manipulate the host cell to promote bacterial survival within the host cell cytoplasm (Sperandio et al., 2008). Two sub-classes of MxiE-regulated genes have been discovered; those regulated by MxiE alone and those regulated by VirB under conditions of non-secretion and further upregulated by MxiE under conditions of secretion ((Le Gall et al., 2005); Figure 2). Thus, based on the hierarchy presented in this chapter, the latter group can be considered tier 3 genes that receive additional regulatory inputs from the tier 4 regulator.

From this overview, it is clear that the four tiers of the regulatory cascade in *Shigella* have distinct levels of stringent and complex regulatory control. In a 1997 study where the first 3 tiers were compared to one another (Porter and Dorman, 1997a), tier 3 promoters were found to be more tightly controlled than tier 2, and tier 2 promoters were more tightly controlled than tier 1. The varying degrees of stringent control were found to manifest in multiple ways: (1) basal level expression, tier 1 > tier 2 > tier 3; (2) level of thermal induction, tier 3 (induced 100-fold) > tier 2 (induced 10-fold) > tier 1 (induced 2-fold); and (3) response to multiple environmental (discussed above), tier 1 > tier 2 > tier 3 ((Porter and Dorman, 1997a); Figure 2). Tier 4 expression levels were not measured in these studies as MxiE was not discovered or described until later (Kane et al., 2002; Mavris et al., 2002b). Nevertheless, it is clear that MxiE-regulated genes are also prone to stringent control due to their distal position in the regulatory cascade, their induction by a single cue (i.e., type III secretion activation) and the dual regulation of at least a subset of these genes by VirB.

At first glance, the position of VirB in the hierarchy under the control of VirF suggests a less significant role for VirB (Figure 2). However, it has been argued that this position allows VirB to serve as a regulatory checkpoint to ensure that appropriate conditions are present prior to committing to full-scale expression of key virulence factors (Dorman, 2009). As mentioned earlier, VirB positively regulates both the *virF* and *virB* promoters (Kane and Dorman, 2012) and co-regulates a subset of MxiE-regulated promoters ((Le Gall et al., 2005); Figure 2), which indicates that VirB acts as a regulatory sentinel at every level; integrating additional inputs to determine absolute levels of expression within the regulatory cascade. These findings elevate the importance of VirB in the cascade and somewhat diminish the role of VirF, which was originally coined the "master regulator" (Porter and Dorman, 1997a), even though the *virB* promoter remains completely dependent on VirF for its activity (Tobe et al., 1993).

b. Regulatory control & function of the pINV-encoded regulators

From the overview presented above, it is clear that VirF, VirB and MxiE are key virulence gene regulators in *Shigella* spp. that are required for transcription of genes encoded by pINV. Moreover, in their absence, cells are avirulent in animal models due to the lack of downstream virulence factors (Kane et al., 2002; Sakai et al., 1986; Sasakawa et al., 1988). Thus, systems are in place to ensure the timely and robust expression of these key regulators. We will next discuss these key virulence regulators in the context of how their expression is triggered, how each regulator functions and the direct effect that each regulator has on its regulated genes.

i/ VirF

Events that trigger the expression of *virF*

Activation of the *virF* promoter, and hence the regulatory cascade, is primarily triggered by a temperature of 37°C, although moderate osmotic stress (equivalent to physiological saline) and a pH of 7.4 are also needed for maximal activation (Falconi et al., 1998; Porter and Dorman, 1997a). The temperature dependency of the *virF* promoter relies on i) H-NS and ii) a temperature dependent "hinge" region located within the promoter region itself (Falconi et al., 1998; Prosseda et al., 2004; Prosseda et

al., 1998; Ulissi et al., 2014). At temperatures below 32°C, the bending and the rotational orientation of two HNS binding sites centered at -250 and -1 relative to transcription start site (TSS) of *virF* is optimized for H-NS binding and repression of the promoter (Ulissi et al., 2014). Both sites must be occupied and the template DNA must be supercoiled for full transcriptional repression to occur, although H-NS binding occurs regardless of whether the DNA is supercoiled or not (Falconi et al., 1998). At temperatures higher than 32°C, *in vitro* experiments demonstrate that the angle of the bend in the DNA decreases and its center (found halfway between the two HNS binding sites at 32°C) moves downstream (Prosseda et al., 2004). These events are proposed to favor the removal of H-NS and facilitate the binding of FIS (factor for inversion stimulation) to sites that overlap the H-NS binding sites (Falconi et al., 2001; Prosseda et al., 2004). At 37°C, FIS activates the *virF* promoter (conditions that prevent H-NS-mediated repression), but at the critical temperature of 32°C, FIS is seen to partially counteract H-NS-mediated repression (Falconi et al., 2001).

The transcriptional regulators VirF and VirB can also modulate the activity of the *virF* promoter. VirF autorepresses the *virF* promoter (Porter and Dorman, 1997a), although it is unclear whether this effect is direct or indirect, while VirB boosts the initial activation of *virF* transcription upon thermal upshift to 37°C (Kane and Dorman, 2012). At first glance the regulation of the *virF* promoter by the downstream regulator VirB seems counterintuitive – would VirB even be present if the *virF* promoter hasn't fired? However, low levels of VirB protein have been detected even at non-permissive temperatures in wild type *Shigella* (Beloin et al., 2002; Porter and Dorman, 1997a). So, it is possible that the VirB-dependent effect on the *virF* promoter occurs at very low concentrations of VirB, although this has not been formally tested. It is also interesting to consider how VirB functions to boost *virF* transcription upon thermal upshift. The VirB binding sites that are proposed to be involved in the VirB-dependent upregulation of the *virF* promoter (Kane and Dorman, 2012) are located midway between the two H-NS binding sites and in close proximity to the FIS sites (described above; (Falconi et al., 2001; Prosseda et al., 2004)). Therefore, the role that VirB plays in enhancing *virF* transcription upon thermal induction may be similar to the role that FIS plays at this promoter: one of partially counteracting H-NS-mediated repression (Falconi et al., 2001), perhaps by stabilizing the new, thermally derepressed promoter region.

The *virF* promoter can be further modulated by another chromosomally encoded NSP, integration host factor (IHF) and by other chromosomally encoded transcriptional regulators including CpxR, Fur, Cra, CsrA, and ArcA (see section V). IHF increases *virF* promoter activity by 2-fold in both the logarithmic and the early stationary phases of growth in a temperature-independent manner. Surprisingly, the IHF binding site implicated in mediating this effect is located immediately downstream of the TSS (+45 to +57) (Porter and Dorman, 1997b). In addition, *virF* expression can also be controlled at the post-transcription level. tRNA modification mediated by *tgt* and *miaA* gene products (tRNA guanine transglycosylase and tRNA prenyl transferase, respectively) is required for efficient translation of *virF* mRNA (Durand and Bjork, 2003; Durand et al., 1997; Durand et al., 2000; Durand et al., 1994; Hurt et al., 2007). Thus, it is clear that although temperature plays a major role in the upregulation the *virF* promoter, a multitude of environmental conditions can ultimately influence the amount of *virF* expression and hence VirF protein produced.

The VirF protein

VirF is a 30 kDa protein (Sakai et al., 1988), which is part of the AraC superfamily of transcriptional regulators (Dorman, 1992; Martin and Rosner, 2001; Savelkoul et al., 1990). This group of regulators shares a conserved C-terminal domain containing two HTH motifs (presumed to be required for DNA binding) (Gallegos et al., 1997), and can be divided into three subgroups depending on the nature of the response signal: a chemical, stress or physical signal (Porter and Dorman, 2002). VirF is part of the latter group because like other virulence gene regulators in this subgroup, the gene encoding this regulator is thermally induced (Dorman, 1992; Jordi et al., 1992; Porter et al., 1998). Another member of this subgroup, Rns, shares a 36% amino acid identity with VirF throughout the length of the protein sequence (Munson, 2013; Porter et al., 1998). The genes encoding these regulators are proposed to have been derived from a common ancestor outside of the Enterobacteriaceae family due to their relatively low G+C content (*virF* 30% and *rns* 28%) (Dorman, 1992). Interestingly, the Rns protein present in enterotoxigenic *E. coli* (ETEC) was able to complement a *virF* null mutation in *S. flexneri* to activate transcription of virulence genes (Porter et al., 1998), suggesting that these regulators recognize similar DNA binding sites (Munson and Scott, 1999) and function similarly to activate transcription (Porter et al., 1998). However VirF was unable to complement a deficient *rns* strain, suggesting there are subtle, but significant differences that are required for VirF-dependent activation (Porter et al., 1998). The mechanism in which VirF activates transcription at its target promoters remains elusive, but evidence suggests that VirF monomers can interact (Porter and Dorman, 2002), which may explain the large and continuous VirF-dependent footprints observed at the *virB* promoter region (Tobe et al., 1993). Oligomerization of VirF also lends support for an oligomerization-driven activation hypothesis, where VirF binds to the promoter and then oligomerizes along the DNA in order to activate transcription (Porter and Dorman, 2002).

The regulatory effect of VirF

To date, VirF-mediated regulation has been characterized at the *icsA* (also referred to as *virG*)*, rnaG, virB* and *virF* promoters (Adler et al., 1989; Jost and Adler, 1993; Sakai et al., 1988; Tobe et al., 1993; Tobe et al., 1995; Tran et al., 2011). IcsA is a 120 kDa outer membrane protein specifically localized to one pole of the bacterium (Goldberg et al., 1993). It mediates the actin tail assembly, which is required for cell-to-cell spread of the bacterium within the human intestinal epithelium, by recruiting host cell actin to the bacterial surface (Bernardini et al., 1989). *In vitro* transcription assays reveal that VirF alleviates H-NS-mediated repression of the *icsA* promoter, but also stimulates transcription of *icsA* directly in the absence of H-NS (Tran et al., 2011). Since *virF* expression increases rapidly above 32°C (described above; (Prosseda et al., 2004)) and the occupancy of the *icsA* promoter by H-NS is lost at 37°C (Tran et al., 2011), it has been proposed that VirF initially functions to alleviate H-NS-mediated repression of the *icsA* promoter, but switches to become an activator of *icsA* transcription once H-NS has been removed from the promoter region (Tran et al., 2011). Regulation of the *icsA* promoter region is further complicated by RnaG, an antisense *cis*-encoded RNA that functions to prematurely terminate *icsA* transcription (Giangrossi et al., 2010; Tran et al., 2011). *In vitro* transcription assays reveal that concentrations of VirF that stimulate the *icsA* promoter also repress the *rnaG* promoter. Therefore, VirF can also indirectly increase *icsA* expression post-transcriptionally by reducing expression of the anti-sense RNA RnaG (Giangrossi et al., 2010; Tran et al., 2011). Importantly, at the *icsA* genetic locus, all sites bound by H-NS

and VirF are located downstream or overlapping of the *icsA* TSS. The regulatory architecture of this promoter therefore demonstrates the importance of considering downstream sequences when studying virulence gene regulation in *Shigella* and its close relatives. It is noteworthy that transcriptional regulation by Rns, the VirF homologue in ETEC, has been reported to occur from downstream binding sites (Munson et al., 2001; Munson and Scott, 2000).

At the *virB* promoter, in contrast to the *icsA* promoter, VirF binds to sequences located upstream of the TSS (between -17 and -117) (Jost and Adler, 1993; Tobe et al., 1993). Here, though, VirF functions to activate the *virB* promoter (mechanism described in more detail below).

ii/ VirB

Events that trigger the expression of *virB*

A temperature shift to 37°C is the primary trigger for *virB* mRNA expression, while moderate osmotic stress (equivalent to physiological saline), but not neutral pH, is required for maximal expression (Porter and Dorman, 1997a). Several lines of evidence suggest that the *virB* promoter may be repressed by H-NS (Beloin and Dorman, 2003; Hromockyj et al., 1992; Maurelli and Sansonetti, 1988; Porter and Dorman, 1997a; Tobe et al., 1993), but few test this directly. Those that do, use *hns* mutants of *Shigella* in which VirF levels are now known to be elevated due to the dysregulation of the *virF* promoter (described above). Since VirF is a known activator of the *virB* promoter, the *hns* mutant background is likely to cause indirect regulatory effects ((Tobe et al., 1991; Tobe et al., 1993); described more below). The best evidence for direct H-NS-mediated repression of the *virB* promoter comes from *in vitro* DNase I footprinting experiments, where H-NS can be seen to occupy a single region within the *virB* promoter from -20 to +20 (relative to the *virB* TSS) (Tobe et al., 1993). This site of occupancy strongly suggests that when H-NS is bound, RNA polymerase is excluded from the promoter and transcription is repressed.

VirF is absolutely required for induction of the *virB* promoter in wild type and *hns* mutant backgrounds (Tobe et al., 1991; Tobe et al., 1993) and occupies the *virB* promoter *in vitro* from -17 to -105 (relative to the *virB* TSS) (Tobe et al., 1993). VirF does not function as a derepressor that displaces H-NS at the *virB* promoter, but plays a positive role in its regulation, as demonstrated by the requirement for VirF in *in vitro* transcription assays (Tobe et al., 1993). Interestingly though, when VirF levels are artificially elevated to wild type levels and above at the non-permissive temperature of 30°C, *virB* transcription does not reach the level seen in wild type cells grown at 37°C. Three observations strongly suggest that this thermo-regulation is mediated by a temperature-dependent change in the supercoiled state of the *virB* promoter region. Firstly, temperature is known to modulate levels of DNA supercoiling (Dorman et al., 1990; Falconi et al., 1998). Secondly, the *virB* promoter is significantly activated in the presence of VirF *in vitro* when placed within an artificially supercoiled DNA molecule, but not a relaxed or linearized molecule (Tobe et al., 1993). Thirdly, the addition of a gyrase inhibitor novobiocin at 37°C, which relaxes supercoiled DNA *in vivo*, drastically reduces transcription of the *virB* promoter to levels similar to those observed at 30°C (Tobe et al., 1993). Thus, unlike the *virF* promoter where thermal induction triggers a change in DNA topology that removes H-NS and allows FIS to bind (discussed above), at the *virB* promoter thermal induction triggers a change in DNA supercoiling which is predicted to promote the

interaction of VirF with RNA polymerase, allowing maximal VirF-dependent activation of the *virB* promoter (Tobe et al., 1993; Tobe et al., 1995).

The *virB* promoter has additional regulatory inputs that work to fine tune its regulatory response. (i) The VirB protein participates in a feedback loop to positively autoregulate its own promoter (Kane and Dorman, 2012). At 30°C, when native *virB* expression level is low, exogenous expression of *virB* can significantly increase the activity of the *virB* promoter to levels similar to those observed in an *hns* null mutant (Kane and Dorman, 2012) by binding to a *cis*-acting binding site within the *virB* promoter (located between -285 and -277 relative to the *virB* TSS) (Kane and Dorman, 2012). During the transition from 30 to 37°C, VirB participates in the full and continuous activation of the *virB* promoter, unlike at the *virF* promoter where VirB only participates in the initial activation (Kane and Dorman, 2012). To date, however, it remains unclear how VirB impacts or is constrained by the regulatory effects mediated by VirF and H-NS at the *virB* promoter. (ii) The integration host factor (IHF, discussed below), is required for the continuous expression of *virB* in stationary phase by binding a region, -171 to -183, located upstream of the essential VirF binding site (Porter and Dorman, 1997b), although the effect is mild (an approximately 2-fold reduction in *virB* expression in the absence of IHF). Notably, exogenous expression of *virB* is able to compensate for the lack of IHF (Porter and Dorman, 1997b). (iii) The iron-responsive *trans*-acting small RNA, RyhB, is able to transcriptionally regulate *virB* independently of VirF when expressed exogenously ((Broach et al., 2012; Murphy and Payne, 2007); see Chapter on Ribo-regulation for further discussion). While RyhB-dependent modulation of *virB* transcription is independent of all factors known to regulate *virB* transcription, DNA sequences in the middle of the *virB* open reading frame appear to be required for this regulation (Broach et al., 2012), which suggests that RyhB-dependent regulation of *virB* transcription is both direct and mechanistically novel.

virB mRNA levels are also post-transcriptionally regulated by the RNA binding protein Hfq (Mitobe et al., 2008, 2009). Under non-physiological temperatures (Mitobe et al., 2008) or osmolarities (Mitobe et al., 2009) *virB* mRNA is relatively unstable when compared to physiological conditions. Deletion of the *hfq* gene increases the stability of *virB* mRNA and consequently increases VirB protein levels and VirB-regulated gene expression under non-physiological conditions (Mitobe et al., 2008, 2009). Therefore, although a description of this post-transcriptional effect may appear misplaced in a chapter that focuses on the transcriptional control of *Shigella* virulence genes, Hfq is clearly involved in the modulation of *virB* mRNA levels in *Shigella.*

The VirB protein

VirB (also called InvE in *S. sonnei*; (Watanabe et al., 1990)) is small (35.4 kDa), basic and encoded by pINV and represents a unique and novel class of transcriptional regulators. It has been extensively characterized as a positive transcriptional regulator of virulence-associated genes (Adler et al., 1989; Le Gall et al., 2005; Watanabe et al., 1990), but does not appear to function as a transcription activator *per se* (Lee et al., 2012). Instead, it functions to counteract H-NS-mediated repression (Stoebel et al., 2008; Wing et al., 2004). VirB shares no homology with currently characterized transcriptional regulators, but shares significant homology with ParB (42.8% amino acid identity over 278 residues) and SopB (30.4% amino acid identity over 168 residues) (Watanabe et al., 1990), which are proteins involved in plasmid

partitioning and maintenance of the P1 and F plasmids, respectively (Abeles et al., 1985; Mori et al., 1986). Despite this similarity, which is especially high within the first two thirds of these proteins, VirB does not retain any plasmid partitioning activity (Buchrieser et al., 2000; Radnedge et al., 1997; Taniya et al., 2003).

Two structural domains have been characterized in the VirB protein: the centrally located DNA binding domain, which contains a HTH motif (residues 148-171) and an oligomerization domain, which is predicted to form a trimeric-coiled coil region at the extreme C-terminus (residues 260-309), (Beloin et al., 2002; Gao et al., 2013). Initially, a leucine zipper dimerization domain was also predicted in VirB (residues 193-228) (Beloin et al., 2002), but structural studies did not support this prediction (Gao et al., 2013). Based on high levels of homology with the N-terminus of ParB and SopB, the VirB N-terminal region is likely to be flexible in solution (Vecchiarelli and Funnell, 2013). This probably explains why structural studies of the N-terminus of VirB are lacking, even though a variety of functions have been proposed for it, including transcriptional activation, structural involvement in DNA binding and protein interactions (Beloin et al., 2002).

The DNA binding domain of VirB is 80% identical to the HTH domain of ParB (Beloin et al., 2002; Surtees and Funnell, 2001), which strongly suggests that the DNA binding properties of these two proteins are similar. Not surprisingly, DNA binding studies for VirB have often paralleled those done with ParB (Beloin et al., 2002; Gao et al., 2013; McKenna et al., 2003; Taniya et al., 2003; Turner and Dorman, 2007; Watanabe et al., 1990). To date, the only established VirB binding site (A/G)(A/T)G(G)AAAT (Taniya et al., 2003), is a composite of three sites (located at the *virA*, *spa15*, and *icsB* promoters) that bear similarity to a part of the ParB binding site *parS* on the P1 plasmid (ATTTCAC, also known as A-boxes) (Radnedge et al., 1996). Importantly, this site does not represent a consensus of all VirB binding sites found on pINV, nor is it an *in vitro*-generated protein binding logo. To complicate matters further, the sites required for the VirB-dependent regulation of virulence genes are sometimes organized as inverted repeats (separated by a single nucleotide) (Castellanos et al., 2009; Turner and Dorman, 2007), and sometimes organized as single heptameric sequences (Kane and Dorman, 2012). This raises questions about which DNA sequences are sufficient for VirB binding, and how VirB initially engages these sequences. Although inverted repeats are often bound by dimeric proteins, recent crystallographic studies of the DNA binding domain of VirB bound to two inverted repeats, suggest that VirB does not need to dimerize before engaging its recognition site (Gao et al., 2013). Instead, the DNA binding domain of VirB, which importantly cannot dimerize/oligomerize, interacts with just one half of the inverted repeat, via the half site known as box 2 (Gao et al., 2013). This is in contrast to ParB, which is demonstrated to occupy both half sites within its recognition sequences in similar crystallographic studies (Schumacher and Funnell, 2005). Interestingly, VirB commonly binds to DNA non-specifically *in vitro* (Harrison, 2010; McKenna et al., 2003; Turner and Dorman, 2007), yet is capable of binding to its targets *in vivo* with specificity (Beloin et al., 2002). This strongly suggests that the binding specificity of VirB is determined by a factor or condition present *in vivo* that is absent *in vitro.* So far, however, the identity of this factor or condition remains elusive. Clearly, there is still much that needs to be learned about how VirB recognizes and binds to its DNA targets.

The oligomerization domain of VirB is found in the C-terminus, which is the least similar region of VirB when compared to ParB and SopB. This region enhances DNA binding both *in vivo* and *in vitro* and is necessary for the oligomerization of VirB in solution, because C-terminally truncated proteins have lost this ability (Beloin et al., 2002; Gao et al., 2013). While ParB also oligomerizes in solution and along DNA (Rodionov et al., 1999; Rodionov and Yarmolinsky, 2004; Surtees and Funnell, 1999), this activity is mediated by its extreme N-terminus (Surtees and Funnell, 1999, 2001), demonstrating that different regions within ParB and VirB confer their shared ability to oligomerization. *In vitro* VirB:DNA complexes appear as smears and large molecular weight complexes at the top of the gel when resolved electrophoretically, rather than discrete bands, which supports oligomerization along DNA (Harrison, 2010; McKenna et al., 2003; Turner and Dorman, 2007). VirB oligomerization in the context of transcriptional derepression appears important, but so far its precise role remains unclear.

Although many questions about VirB-mediated transcriptional derepression remain unanswered, a working model has been proposed. The HTH domain of VirB binds to the *cis*-acting heptameric sequence, which serves as a nucleation point (Beloin et al., 2002; Gao et al., 2013; Kane and Dorman, 2012; Turner and Dorman, 2007). VirB binding introduces a bend in the intrinsically, highly flexible A-tract segment within the binding site, which facilitates the non-specific binding of VirB oligomers along the DNA (Gao et al., 2013). This is a novel feature of VirB because ParB-like proteins are unable to induce bends in the DNA. VirB oligomerization along the DNA strand destabilizes the H-NS-bound DNA complex to allow transcription to proceed (Beloin et al., 2002; Gao et al., 2013; Kane and Dorman, 2012; Turner and Dorman, 2007). Taken together, while it is clear that DNA binding and oligomerization activities of VirB are key for its activity as a transcriptional derepressor of *Shigella* virulence genes, much remains unclear about how this unusual transcriptional regulator counteracts repression mediated by the nucleoid structuring protein H-NS. It is hoped that in depth studies of transcriptional derepression at VirB–regulated promoters will shed light on this non-traditional, yet conserved, regulatory mechanism.

The regulatory effect of VirB

In contrast to VirF, VirB has an extensive regulon. This includes genes of the *ipa-mxi-spa* entry region, genes located outside of this region including *icsP, phoN2, ospB*, *ospF, ospC1*, *ospC2/3/4, virA, ospD1, ospD2, orf13, orf81, orf137, ipaJ, virB* and the genes of the VirF and MxiE regulons, which are indirectly regulated by VirB via its direct control of *virF* and *mxiE* (Adler et al., 1989; Berlutti et al., 1998; Bongrand et al., 2012; Le Gall et al., 2005; Santapaola et al., 2002; Taniya et al., 2003; Uchiya et al., 1995; Watanabe et al., 1990; Wing et al., 2004). Of those genes thought to be directly regulated by VirB, binding of VirB to the promoter region has only been demonstrated for a subset of genes, including *icsB*, *icsP, ospZ* and *virB* (Harrison, 2010; Kane and Dorman, 2012; McKenna et al., 2003; Taniya et al., 2003; Turner and Dorman, 2007).

The best characterized VirB-dependent promoter is P*icsB* (Beloin and Dorman, 2003; McKenna et al., 2003; Taniya et al., 2003; Turner and Dorman, 2007; Watanabe et al., 1990). This promoter is found in the entry region and controls the transcriptional expression of the *ipa* operon, of which *icsB* is the first gene. H-NS occupies a large region within the *icsB* promoter from -110 to +25 relative to the *icsB* TSS (Beloin and Dorman, 2003; Porter and Dorman, 1997b; Turner and Dorman, 2007). When bound, H-NS

inhibits open complex formation (Turner and Dorman, 2007). Based on its DNase I footprint, however, it seems more likely that H-NS blocks the recruitment RNA polymerase to the *icsB* promoter, rather than blocking open complex formation directly. Located immediately upstream of the H-NS bound region, is the site required for VirB-dependent regulation of P*icsB.* While this site (centered at -114 relative to the TSS (Porter and Dorman, 1997b; Turner and Dorman, 2007)) is organized as an inverted repeat, mutagenesis studies reveal that only the promoter-distal half of the inverted repeat (Box 2) is required for the VirB-dependent regulation of the *icsB* promoter (Turner and Dorman, 2007). VirB binding to this site changes the DNA topology within the H-NS-occupied region, as judged by DNase I footprinting. It has been proposed that these changes destabilize the H-NS:DNA complex, leading to the displacement of H-NS from the *icsB* promoter (Turner and Dorman, 2007). Since the VirB protein has no effect on RNA polymerase recruitment or formation of the open transcription complex, VirB can be considered a derepressor rather than a positive transcription factor at the *icsB* promoter (Turner and Dorman, 2007).

The *icsP* promoter is located on the opposite side of pINV relative to the entry region. It controls the expression of a monocistronic gene that encodes the outer membrane protease IcsP, which modulates the amount of the actin tail assembly protein IcsA that is associated with the bacterial surface (Egile et al., 1997; Shere et al., 1997; Steinhauer et al., 1999). At P*icsP*, VirB also functions to counteract H-NS-mediated repression of transcription, which remains in effect even after a shift to 37°C (Wing et al., 2004). But here, in contrast to the situation at the *icsB* promoter, both halves of an inverted repeat, made up of two heptamers each closely matching the characterized VirB binding site (Taniya et al., 2003), are required for VirB-dependent regulation (Castellanos et al., 2009). Importantly, these sites are located more than 1 kb upstream of the *icsP* TSS (from position -1144 to -1130, relative to the TSS) (Castellanos et al., 2009), which is striking because traditionally bacterial transcription factors were thought to mediate their effects from sites located within 200 bp of the TSS (Balleza et al., 2009; Collado-Vides et al., 1991). Consequently, this example demonstrates that VirB is capable of mediating long range effects on transcription and raises the possibility that VirB may control the transcription of other pINV-encoded virulence genes from previously overlooked remote sites. More recently, the VirB binding sites found to regulate the *icsP* promoter were also found to regulate transcription from a second *icsP* promoter (Hensley et al., 2011) as well as the divergent *ospZ* promoter (Basta et al., 2013). Interestingly, the VirB binding sites lie closer to the beginning of the *ospZ* gene (425 bp upstream of ATG) than the *icsP* gene (1164 bp upstream of ATG) (Basta et al., 2013). Even so, the regulatory effect mediated by VirB is much stronger on *icsP* than on *ospZ* (Basta et al., 2013; Hensley et al., 2011). Based on 5' truncation analysis of both *icsP* and *ospZ* promoters*,* the region required for H-NS-mediated repression of the *icsP* and *ospZ* promoters seems to lie immediately downstream of the VirB binding site in the direction of the *icsP* gene located between -893 and -351, relative to the primary *icsP* TSS (Basta et al., 2013; Castellanos et al., 2009; Harrison, 2010). Therefore, unlike the *icsB* promoter, H-NS-mediated repression of *icsP/ospZ* locus require sequences located upstream of the promoter elements. However, similar to the situation at the *icsB* promoter, H-NS-mediated repression of *icsP* and *ospZ* does appear to be mediated by sequences that lie immediately adjacent to the region occupied by VirB.

As described previously, at the *virB* promoter, which is located immediately adjacent and downstream of the *ipa* operon, H-NS binds between -20 and +20 (Tobe et al., 1993) and VirB mediates its effect from

a single heptameric sequence rather than an inverted repeat (Kane and Dorman, 2012). This site displays a 7/8 match to the characterized VirB binding site and is centered at position –281.5 relative to the TSS of *virB*. Protection of this site by VirB can be seen by DNase I footprinting. As described above, VirB binding to this site, as cultures are switched from 30 to 37°C, is reported to relieve H-NS-mediated repression of P*virB* more rapidly and in a more sustained manner than is caused by the change in temperature alone (Kane and Dorman, 2012).

Based on these examples, it is clear that there is significant diversity in the architecture of VirB-regulated promoters, not only in the position of the VirB binding site relative to the promoter elements and regions bound by H-NS, but also within the VirB binding sites themselves. Fifty six occurrences of the VirB binding site found at P*icsB* (5'-ATTTCAT-3') have been found on pINV taken from S. *flexneri* serotype 5a (pWR100) and over 1000 sites were found with a single mismatch (Le Gall et al., 2005). Based on these findings, it seems highly unlikely that the characterized VirB binding site is a strong predictor of VirB-regulated targets. Nevertheless, one common theme is that when VirB-mediated regulation is established, the function of the VirB protein appears to be conserved: to counteract H-NS-mediated repression, even though the mechanism driving VirB–dependent derepression remains poorly understood.

As the tier 3 regulator of virulence gene expression on pINV, VirB also regulates the transcription of the genes encoding the tier 4 regulators, MxiE and IpgC, which are encoded by the 30 kb entry region. As described in the next section, MxiE and IpgC regulate the transcription of genes required for post-invasion activities (discussed below); triggering the expression of a unique set of genes (tier 4) as well enhancing the expression of some VirB-regulated genes (tier 3) during conditions of secretion.

iii/ Regulatory control mediated by the pINV-encoded regulators, MxiE and IpgC

Events that trigger the expression of mxiE, ipgC, ospD1 & spa15

Transcription of tier 4 genes, those expressed in response to activation of the TTSS, relies on a complex circuitry involving tier 3 gene products: the AraC-like transcriptional activator MxiE, its co-activator IpgC, the anti-MxiE activator OspD1, and OspD1 chaperone Spa15 (Parsot et al., 2005). Each of these proteins is encoded by a gene located on pINV. The *mxiE* gene lies within the *mxi* operon of the 30 kb entry region (which is controlled by the P*ipgD*), *ipgC* is located within the divergently transcribed *ipa* operon (which is controlled by P*icsB*), *spa15* is the first gene of the *spa* operon and *ospD1* is encoded by a monocistronic gene located outside of the entry region (residues 20964-21641 on pWR100). As expected by their tier 3 status, *mxiE, ipgC, spa15* and *ospD1* are all maximally expressed at 37°C in a VirB-dependent manner (Le Gall et al., 2005). Importantly though, the amount of MxiE protein made by *Shigella* is determined by regulatory events that occur during transcription elongation of the *mxiE* locus (Penno et al., 2005).

The *mxiE* locus contains two overlapping open reading frames (ORF), *mxiEa* and *mxiEb*, which are out of frame with one another by -1. Transcriptional slippage of RNA polymerase on a run of 9 T residues found

in this region can place these two ORFs in the same frame, allowing the full-length *mxiE* gene to be transcribed (Penno et al., 2005). While *mxiEa* (codons 1-43) is often expressed without *mxiEb,* the small MxiEa peptide does not appear to function in the regulatory circuit controlling expression of tier 4 genes (Penno et al., 2005). In contrast, *mxiEb* (codons 44-251) is only ever transcribed in the context of full-length *mxiE,* but is essential because it encodes its DNA binding determinants for MxiE (Penno et al., 2005). Although conditions that influence frameshifting efficiency have the potential to regulate production of MxiE, a small increase in the amount of MxiE is unlikely to significantly impact expression of MxiE-regulated genes. This is because the activity of MxiE is so tightly controlled by the availability of its co-activator IpgC (described below) (Mavris et al., 2002a). If, however, transcriptional slippage was prevented under conditions that trigger the activation of the TTS apparatus, MxiE production would fail and MxiE-regulated genes would not be expressed (Penno et al., 2005). Thus, specific environmental conditions that hinder transcriptional slippage by RNA polymerase during the transcription of *mxiE* have the potential to negatively influence MxiE-regulated gene expression under conditions of secretion. To date, the precise nature of these conditions, if they exist, remains elusive.

The MxiE, IpgC, OspD1 and Spa15 proteins and their interactions

The capacity of MxiE to function as a transcriptional activator is controlled by the secretion status of *Shigella* (Figure 3). Key players in the modulation of MxiE-regulatory activity are i) the MxiE co-activator IpgC, which also serves as the chaperone for the TTSS translocators IpaB and IpaC (Menard et al., 1994; Page et al., 2001) (Mavris et al., 2002a), ii) the MxiE anti-activator OspD1 and iii) the OspD1 chaperone Spa15 (Marteyn et al., 2012; Page et al., 2002). As the TTSS is activated, IpaB and IpaC dissociate from IpgC and OspD1 dissociates from both MxiE and its chaperone Spa15 (Parsot et al., 2005), allowing secretion of IpaB, IpaC and OspD1. As a consequence, levels of liberated MxiE and IpgC rise in the bacterial cytoplasm, where these two proteins are believed to associate with each other to trigger transcriptional activation of MxiE-regulated genes (Parsot et al., 2005). In support of this, MxiE and IpgC have been shown to associate in the cytoplasm of *E. coli* after co-expression of their genes (Pilonieta and Munson, 2008). Furthermore, MxiE and IpgC are sufficient to activate transcription of target promoters in the absence of other virulence plasmid encoded factors (Mavris et al., 2002a; Penno et al., 2005). So far, however, binding of MxiE or even MxiE:IpgC complexes to DNA has not been demonstrated and the molecular details of the interaction between MxiE and its co-activator IpgC remain poorly understood, although it is predicted that the N-terminal part of MxiE, encoded by *mxiE*a, is involved (Penno et al., 2005), based on similarity to the SicA/InvF system in *Salmonella enterica* (Darwin and Miller, 2001). While additional work is required to fully understand the molecular details of the regulatory control mediated by MxiE and IpgC, the sophisticated use of an apparently defective transcriptional activator, a co-activator that doubles as a TTSS chaperone, an anti-activator and its chaperone to control the transcription of MxiE-regulated genes is both evolutionarily impressive and elaborate way to tie the activity of the TTSS to the expression of a second wave of TTSS effector proteins.

The regulatory effect of MxiE and IpgC

Although *mxiE* and *ipgC* are transcribed prior to secretion-specific conditions, most of its regulon members are only transcribed once the TTSS is activated ((Demers et al., 1998; Kane et al., 2002; Mavris

et al., 2002b); Figure 3). During infection, this activation occurs once host cell contact has been made by *Shigella;* but the addition of Congo red dye to growth media or the use of *ipaB* mutants, in which type III secretion is dysregulated (Kane et al., 2002; Mavris et al., 2002a; Mavris et al., 2002b) can "artificially" trigger the activation the type III secretion system and hence the expression of the MxiE regulon. By combining these approaches, the full extent of the MxiE regulon has been characterized ((Bongrand et al., 2012; Buchrieser et al., 2000; Demers et al., 1998; Kane et al., 2002; Mavris et al., 2002a; Mavris et al., 2002b); Figure 2). Thirteen MxiE-regulated genes are encoded by the large virulence plasmid pINV: *ospB, phoN2, ospC1, ospD3, ospE1, ospE2, ospF, ospG, virA, ipaH1.4, ipaH4.5, ipaH7.8 and ipaH9.8* (Bongrand et al., 2012; Le Gall et al., 2005); the remaining MxiE-regulated genes are the chromosomally encoded *ipaH* genes, the number of which varies from strain to strain (Bongrand et al., 2012; Mavris et al., 2002b).

The DNA binding site required for MxiE-mediated activation is 17 bp in length and typically overlaps the -35 region of MxiE-regulated promoters. This site designated the MxiE box (5'-GTATCGTTTTTTTAnAG-3') was originally identified by promoter mapping and deletion analysis of three MxiE-regulated promoters (*ipaH9.8, virA* and *ospC1*) (Mavris et al., 2002b), but similar sites (with up to 3 base pair mismatches) have since been found upstream of all of the designated MxiE-regulated genes (Bongrand et al., 2012). Interestingly, the identification and requirement of MxiE boxes upstream of each one of the chromosomally encoded *ipaH* genes (Bongrand et al., 2012), demonstrates that chromosomal loci can indeed fall under the transcriptional control of a pINV-encoded regulator, raising the possibility that other chromosomal genes may be pINV-regulated too.

By comparing the expression of the MxiE-regulated genes in a variety of *Shigella* mutant backgrounds, two sub-classes of MxiE-regulated genes were discovered (Le Gall et al., 2005): i) those regulated by MxiE alone, and ii) those directly regulated by VirB under conditions of non-secretion and further upregulated by MxiE under conditions of secretion (Figure 2). It has been proposed that this hierarchy allows VirB-regulated primary effector proteins to be secreted immediately after the TTSS is activated, and then a subset of these proteins (including *ospB*, *phoN2*, *ospF*, *ospC1* and *virA*) to be further upregulated by MxiE. Thus, MxiE can upregulate expression of effectors initially expressed by VirB, and also regulate a unique set of genes (the true tier 4 genes) in a second wave of effector secretion. The organization of genes into one of these subclasses allows the precise and deliberate secretion of proteins into the host cell cytoplasm, so that the host cell can be manipulated in an exquisite and coordinated manner by these *Shigella* encoded MxiE-regulated effectors.

V. Transcriptional control mediated by other chromosomally encoded regulators

As we have seen, the expression of virulence genes encoded by pINV relies on a suite of transcriptional regulators that are also encoded by this large virulence plasmid (discussed in section IV). Evolutionarily this makes sense; the horizontally-acquired virulence genes were acquired along with their cognate transcriptional regulators. But, regulatory cross-talk between pINV and the chromosome is evident in *Shigella*. For instance, the chromosomally encoded *ipaH* genes, which encode type III translocators, are

directly regulated by the pINV-encoded transcriptional regulator MxiE (Bongrand et al., 2012). Furthermore, several chromosomally encoded transcriptional regulators have been shown to directly bind to promoters located on the virulence plasmid to regulate their expression (Figure 4). In this section, we will focus on the regulatory inputs made by chromosomally encoded transcription regulators, first focusing on those that bind directly to pINV promoters, including FIS, IHF, CpxR, FNR and PhoP (H-NS and its homologs have been previously discussed above) and concluding with those whose influence, thus far, appears to be indirect, such as Fur, CsrA, ArcA and OmpR (Figure 4).

Chromosomally encoded regulators with a direct role in the transcriptional regulation of virulence genes located on pINV

FIS

FIS, the factor for inversion stimulation, is characterized as a nucleoid structuring protein, like H-NS. In *E. coli,* FIS contributes to the overall cell physiology by participating in many essential processes and by binding to hundreds of DNA targets (discussed in (Browning et al., 2010)). Consequently, FIS has been given the title "global modulator of metabolism" (Schneider et al., 1999). Although the primary role of FIS is to inhibit transcription of nonessential genes during exponential growth in *E. coli* and *Salmonella enterica,* it does have the capability to activate gene transcription in *E. coli* (Opel et al., 2004; Wilson et al., 2001). In *Shigella*, FIS facilitates the temperature-dependent activation of P*virF* to positively influence *virF* expression (discussed previously; Figure 4). After temperature-induced (37°C) changes in the DNA topology has removed H-NS and exposed a FIS binding site, FIS binds to the *virF* promoter to further upregulate transcription (Falconi et al., 2001). Additionally, during the transition between non-permissive (30°C) to permissive (37°C) temperatures, FIS either (1) competes with H-NS for binding sites at P*virF*, since 2 of the 4 identified FIS binding sites overlap with one of the two H-NS binding sites, or (2) forms an intermediate complex with H-NS to promote the transition from a repressed promoter state to an active state (Falconi et al., 2001). Therefore, while FIS directly regulates transcription from the pINV-located *virF* promoter and likely contributes to its robust expression during infection, it is not absolutely essential for *virF* expression. It remains unclear if FIS regulates other genes on pINV at this time.

IHF

The integration host factor, IHF, is another example of a nucleoid structuring protein that influences transcription of pINV genes. The alpha and beta subunits of this heterodimeric protein are encoded by the *ihfA* and *ihfB* genes, respectively. In *E. coli*, the primary role of IHF is to sharply bend DNA (more than 160°) (Rice et al., 1996) to cooperatively work with other NSPs or transcription factors to facilitate gene transcription, DNA recombination or DNA replication. Moreover, IHF can contribute to promoter activation by directly interacting with the alpha subunit of RNA polymerase (Goosen and van de Putte, 1995). Evidence suggests IHF binds directly to both tier 1 (P*virF*) and tier 2 promoters (P*virB* and P*icsA*) on pINV (Figure 4), based on the formation of IHF-dependent protein DNA complexes in *in vitro* binding assays from crude extracts, and the presence of putative IHF binding sites in the DNA surrounding these promoters [*virF* +45 to +57; *icsA* +109 to +122; *virB* -171 to -183, relative to the respective TSS] (Porter and Dorman, 1997b). The role of IHF at these promoters appears to be one of modulation rather than essentiality, but its regulatory function appears to vary; being an activator at P*virF* and P*icsA* during

exponential growth phase and an alleviator of H-NS-mediated repression at P*virB* during stationary phase (Porter and Dorman, 1997b).

CpxR

The CpxR/CpxA two component system is a signal transducing pathway that is initiated in response to extracytoplasmic stress, including, but not limited to, cell envelope perturbations, misfolded proteins and changes in pH (recently reviewed in (Raivio, 2014)). As a two component system, the presence of a specific signal is recognized by the sensor kinase CpxA, which phosphorylates its cognate response regulator CpxR, which in turn regulates the expression of several genes needed to overcome the stress. In *Shigella,* the CpxR/CpxA system was first implicated in the modulation of virulence gene expression when an *E. coli cpxA* mutant led to the pH-dependent dysregulation of a *virF'-'lacZ* translational fusion, at both pH 7.4 and 6.0, but particularly at pH 6.0 (Nakayama and Watanabe, 1995). A subsequent study, established CpxR as an essential activator of *virF* expression and demonstrated that CpxR binds to sequences located between -103 to -37 relative to the TSS of the *virF* gene (Nakayama and Watanabe, 1998). In this study, however, disparate phenotypes were associated with the *cpxR* and *cpxA* mutants; the *cpxR* mutant abolished *virF* expression, while the *cpxA* mutant retained comparatively high levels of *virF* expression, regardless of pH. Two explanations were offered (Nakayama and Watanabe, 1998). First, in the absence of *cpxA,* CpxR may have become phosphorylated by another phosphate donor. Second, phosphorylated CpxR may not be the only factor that contributes to *virF* expression *in vivo,* which we now know to be true (Falconi et al., 2001; Kane and Dorman, 2012). A later study, characterized another *cpxA* mutant, which had a transposon insertion closer to the 5' end of the *cpxA* gene (Mitobe et al., 2005), than the originally described mutant (Nakayama and Watanabe, 1995). Strikingly, this *cpxA* mutant did not show a difference in *virF* expression profiles compared to wild type at neutral pH, despite this mutation causing a down-regulation of the TTSS in *Shigella* (Mitobe et al., 2005). Instead, this *cpxA* mutant exhibited a decrease in the posttranscriptional processing of *virB* mRNA (InvE) (Mitobe et al., 2005), which suggest that CpxA mediates post-transcriptional regulation of *virB* via an unidentified factor. Based on these three studies (Mitobe et al., 2005; Nakayama and Watanabe, 1995; Nakayama and Watanabe, 1998), it is clear that the chromosomally encoded CpxR/CpxA system plays an important role in the activation of the *Shigella* virulence regulatory cascade (Figure 4): CpxR functions as a direct and essential activator of *virF* gene expression (Nakayama and Watanabe, 1998), while CpxA is involved in the pH-dependent modulation of *virF* expression and the post-transcriptional modulation of *virB* mRNA levels (Mitobe et al., 2005; Nakayama and Watanabe, 1995).

FNR

The anaerobic transcriptional regulator, FNR (fumarate and nitrate reduction), is required for the invasive phenotype of *Shigella* in anaerobic environments, such as those encountered in the host gastrointestinal (GI) tract lumen (Marteyn et al., 2010). In fact, cells grown anaerobically exhibit an increase in cell entry into epithelial cells when compared to cells grown aerobically (Marteyn et al., 2010). This is because *Shigella* becomes "primed" during anoxic conditions via FNR. FNR directly represses the promoters of the tier 3 genes *spa32* and *spa33* ((Marteyn et al., 2010); Figure 4). Spa32 controls the production of proteins that mediate TTSS needle length and the selection of substrates for secretion (Magdalena et al., 2002; Tamano et al., 2002) while dysregulation of Spa33 blocks Ipa

secretion (Schuch and Maurelli, 2001). Therefore, FNR-mediated repression of *spa32* and *spa33* causes TTSS needle length to increase, which is predicted to facilitate contact with host cells during infection, and prevents type III secretion, thus allowing effectors to accumulate in the bacterial cytoplasm (Marteyn et al., 2010). It is proposed that as the "primed" *Shigella* approaches the epithelial layer of the lower GI tract, there is sufficient oxygen to inactivate FNR, allowing type III secretion to occur from the longer and more numerous needles in contact with host cells, ultimately leading to enhanced host cell entry. These proposed events are congruent with multiple lines of investigation, including microelectrode monitoring of the surface of epithelial cells that have recently been moved to an anaerobic cabinet, oxygen reporters used within the rabbit ileal loop model of infection and SEM imaging of *Shigella* taken from aerobic or anaerobic cultures (Marteyn et al., 2010). In summary, the chromosomally encoded transcription factor FNR transcriptionally regulates the pINV-located genes, *spa32* and *spa33,* to facilitate the priming of bacterial cells in anoxic conditions, like those found in the lower intestine, so that *Shigella* is primed and ready for invasion of the colonic epithelium.

PhoP

In many Gram-negative pathogens, the PhoP/PhoQ two component regulatory system modulates the expression of genes essential for virulence. This system is thought to be activated intracellularly when PhoQ senses the low magnesium environment of the phagosome. In *Shigella, phoP* mutants are more susceptible to killing by polymorphonuclear leukocytes (PMNs) and cationic antimicrobial peptides, which play important roles in late stage infections (Moss et al., 2000), but the role of PhoP in the invasion and intercellular spread of *Shigella* remains somewhat ambiguous (Cai et al., 2011; Moss et al., 2000; Sidik et al., 2014). Although the genes responsible for the phenotypes observed in the original Moss et al. study were not identified, at least one operon on pINV, which is comprised of four genes (*shf/pgdA, rfbU/wabB, virK, msbB2*), has subsequently been shown to be regulated by the PhoP/Q system ((Goldman et al., 2008); Figure 4). Upstream of this operon and overlapping the -35 promoter element lies a putative PhoP binding site, mutations in which lead to diminished PhoP-dependent promoter activity (Goldman et al., 2008). Furthermore, mutations in the last two genes of the operon, *msbB2* and *virK,* generate phenotypes that are consistent with other studies of *phoP* mutants (Cai et al., 2011; Goldman et al., 2008; Moss et al., 2000; Sidik et al., 2014). Briefly, the *msbB2* gene*,* encodes a myristoyl transferase enzyme that catalyzes the last step in the synthesis of the lipid A moiety of LPS. Expression of *msbB2* is detected in low magnesium ion conditions, but not in the presence of high magnesium ion concentration (Goldman et al., 2008). Under low magnesium conditions, MsbB2 functions to increase the overall hexa-acylation of lipid A (Goldman et al., 2008), which is proposed to confer increased resistance to antimicrobial peptides: a phenotype observed by Moss et al. (Moss et al., 2000). On the other hand, the upstream gene, *virK,* was initially implicated in both intracellular and intercellular spread because a *virK*::Tn10 mutation led to a decrease in *Shigella*-associated IcsA (Nakata et al., 1992), the protein required for actin-based motility. Subsequent analyses revealed that this mutation caused an increase in the amount of IcsP, an outer membrane protease that functions to proteolytically cleave the IcsA determinant for actin-based motility from the bacterial surface (Wing et al., 2005). More recently, *virK* has also been implicated in relieving bacterial envelope stress that occurs during growth at 37°C (Sidik et al., 2014). Although the precise activity of VirK remains unclear at this time, it is likely that all the phenotypes associated with *virK* mutants will be explained once it is

discovered (Nakata et al., 1992; Sidik et al., 2014; Wing et al., 2005). In summary, the chromosomally encoded PhoP/Q system is essential for *Shigella* virulence because it plays an important role in the regulation of pINV genes encoding traits such as actin-based intracellular and intercellular spread, resistance to antimicrobial peptides and PMN-mediated killing. Whether PhoP mediates control over these virulence traits solely via its direct regulation of the *shf/pgdA, rfbU/wabB, virK, msbB2* operon, or whether other PhoP-regulated virulence loci on pINV are directly or indirectly involved, still needs to be established.

Chromosomally encoded regulators with an indirect role in the expression of genes located on pINV

Fur

In the presence of iron, the global transcriptional repressor of bacterial iron acquisition, Fur (Ferric uptake regulator), complexes with iron and binds target DNA sites to repress transcription of iron-responsive genes. Meanwhile, in iron-limiting conditions, Fur-mediated repression is alleviated, which permits the transcription of iron-responsive genes. In *Shigella* and *E. coli*, Fur represses expression of the sRNA RyhB in the presence of iron (Massé and Gottesman, 2002; Oglesby et al., 2005). Strikingly, in *Shigella,* production of RyhB not only decreases mRNA levels of targets known to be regulated by RyhB in *E. coli,* but it also decreases *virB* mRNA levels, presumably by RyhB binding to a complementary sequence within the *virB* coding region (discussed above; (Broach et al., 2012; Murphy and Payne, 2007). Consequently, RyhB production leads to the downregulation of tier 3 VirB-dependent loci, including *ipaABCDE, ipgABCE, mxiACDEJL, virA* and *icsP* (Africa et al., 2011; Murphy and Payne, 2007). Thus, under iron-limiting conditions, which are likely encountered within the intracellular compartment of host cells, the expression of *icsP* (Africa et al., 2011), and presumably other VirB-regulated genes, is downregulated (Figure 4). Therefore, transcriptional repression of P*ryhB* by Fur, the chromosomally encoded iron responsive transcription factor, indirectly leads to the upregulation of the pINV-encoded virulence regulator gene *virB* and its downstream targets on pINV, demonstrating yet another layer of a stringent control of tier 3 promoters.

CsrA

The carbon storage regulator CsrA has been shown to positively influence the expression of *virF* and *virB* ((Gore and Payne, 2010); Figure 4). Although this glycolysis activator/gluconeogenesis repressor has been demonstrated as a RNA binding protein in *E. coli* (Baker et al., 2002; Dubey et al., 2003; Liu et al., 1995; Romeo, 1998; Sabnis et al., 1995), the mechanism in which it modulates *virF* and *virB* expression is unclear. What is known is that a *csrA* null mutation results in an inability of cells to attach and invade Henle cells, likely due to the reduction of *virF* and *virB* expression, which directly correlates to a reduction of invasion proteins (Gore and Payne, 2010). Meanwhile, the opposing carbon metabolism regulator Cra, which serves as a glycolysis repressor/gluconeogenesis activator (Saier and Ramseier, 1996), has an opposite effect regarding cell invasion. *cra* mutants are able to attach and invade more Henle cells when compared to wild type (Gore and Payne, 2010). These observations suggest that *Shigella* invasion and virulence is intimately connected to glycolysis and carbon metabolism via the regulators controlling these metabolic pathways (also discussed in (Eisenreich et al., 2010)). Indeed, these conclusions were supported by a recent proteomic survey in which high levels of proteins involved in glycolysis were found in bacteria localized within Henle cells (Pieper et al., 2013). Therefore, although

the role of carbon metabolism regulators on pINV-encoded virulence genes needs further investigation, there is undoubtedly a connection between carbon metabolism and virulence gene expression.

ArcA

The presence of an additional anaerobic transcriptional regulator ArcA has the potential to add another layer of complexity onto the existing virulence regulatory circuit. As previously discussed, cells are primed during anoxic growth via FNR-dependent repression of *spa32* and *spa33,* which prepares bacterial cells for invasion by inhibiting the secretion of effector proteins and by increasing the length and number of the type III secretion system needles (Marteyn et al., 2010). Furthermore, ArcA can bind to the *fur* promoter and repress its transcription during anaerobiosis (Boulette and Payne, 2007) which has the potential to increase production of RyhB (described above and (Massé and Gottesman, 2002; Oglesby et al., 2005)) and decrease *virB* transcription (Broach et al., 2012; Murphy and Payne, 2007). Therefore, while ArcA may indirectly influence expression of pINV-encoded virulence genes (Figure 4), it is apparent that it plays a role in fine-tuning and shaping the overall physiology of the bacterium to optimize infection and invasion strategies.

OmpR

The OmpR/EnvZ two component system, encoded by the *ompB* locus of *Shigella,* has been demonstrated to facilitate expression of tier 3 virulence genes of the entry region in response to changes in osmolarity and is required for *Shigella* virulence (Bernardini et al., 1990). Subsequent studies, however, demonstrated that constitutive expression of the OmpR-dependent *ompC* gene, which encodes an outer membrane porin, bypassed the need for OmpR/EnvZ and restored virulence phenotypes to *Shigella* (Bernardini et al., 1993). As such, the role of OmpR/EnvZ in the regulation of genes of the entry region appears to be indirect, but exactly how the modulation of OmpC by OmpR/EnvZ impacts virulence gene expression has yet to be characterized.

VI. Conclusions

The regulation of *Shigella* virulence genes is a broad topic, which can be viewed succinctly as the mechanism by which *Shigella* adapts to environmental conditions encountered 'en route' through its human host. The continual sensing of the surrounding environment by *Shigella* that are in-transit triggers a whole range of bacterial responses that occur at every level of gene expression (transcriptional, post-transcriptional, translational and post-translational) to ensure that *Shigella* is both competitive and successful within its human host. In this chapter, the primary focus has been the transcriptional regulation of virulence genes carried by the large virulence plasmid pINV, which is found in almost all pathogenic *Shigella* strains (with the exception of *S. dysenteriae* serotype 1 and *S. boydii* serotype 13, see section II). This emphasis has allowed us to highlight the environmental cues, regulatory hierarchies, transcriptional regulators and the regulatory mechanisms that modulate the first step of virulence gene expression in *Shigella.* Moreover, it has demonstrated that transcription of the virulence genes carried by pINV is orchestrated by so much more than simple transcription factor:DNA interactions. To emphasize some of this complexity, we conclude by reviewing the central themes of

virulence gene expression on pINV, namely transcriptional silencing and anti-silencing, which commonly controls virulence gene expression in other bacterial pathogens. We also highlight some of the regulatory mechanisms that allow virulence gene expression in *Shigella* to be tied to cues encountered within the host environment, and briefly discuss some of the more unusual and intriguing features of virulence gene expression on pINV, which merit further investigation. Finally, we describe the challenge of identifying where and when specific genes are expressed in the human host, as well as some recent advances that may significantly improve our understanding of this topic. This section will cumulatively reveal what an excellent model system *Shigella* is for the study of transcriptional regulation of virulence genes, but also demonstrate, more generally, what a fascinating organism it is for the study of regulatory mechanisms controlling bacterial transcription.

As we have seen, the virulence genes carried by pINV are transcriptionally silenced by the nucleoid structuring protein, H-NS. The H-NS-mediated transcriptional silencing of horizontally acquired genes on large plasmids, like pINV, is relatively common among the Enterobacteriaceae. Indeed, it has been proposed that the silencing of these foreign genes (xenogeneic silencing) was central in the evolution of *Shigella* species and other enteric pathogens because it allowed the acquisition and maintenance of large plasmids like pINV in many bacterial species (Stoebel et al., 2008). Transcriptional silencing occurs at ambient temperatures (below 37°C), when H-NS nucleoprotein complexes form to organize large DNA molecules like pINV. This renders the DNA inaccessible to sequence-specific DNA binding proteins, like RNA polymerase, that are needed for transcription to proceed. Although many questions remain about the extent and mechanism of H-NS-mediated transcriptional repression on pINV, it is clear that the transcriptional silencing of pINV is central to the genetic circuitry and cellular physiology of *Shigella.* Transcriptional silencing by H-NS also provides the context in which almost all other virulence gene regulation expression occurs on pINV. Consequently, *Shigella* is an excellent model system for studying mechanisms of transcriptional repression by H-NS, as well as mechanisms of transcriptional derepression. In *Shigella*, the major protein involved in transcriptional derepression of virulence genes is the anomalous transcriptional regulator VirB. This protein appears to have been co-opted from its former role as a plasmid partitioning protein to its new role as a transcriptional regulator. Although the array of transcriptional derepressors employed by bacterial pathogens is impressive (Stoebel et al., 2008), evidence suggests that the various mechanisms of transcriptional derepression may display some common features (see text below). Hence, studies of VirB-mediated transcriptional derepression in *Shigella* are likely to promote our understanding of virulence gene expression in *Shigella* and provide valuable insight into mechanisms of transcriptional derepression mediated by other DNA binding proteins in other bacteria.

The *Shigella* system exemplifies how the transcriptional regulation of virulence gene expression can be exquisitely tied to environmental conditions and/or cues encountered in the human host. Examples of this can be seen throughout *Shigella's* journey through its host. While the initial expression of virulence genes carried by pINV can be enhanced by physiological pH and osmolarities found in the human body (Porter and Dorman, 1994, 1997a), undoubtedly the primary cue that triggers the first three tiers of the regulatory cascade controlling virulence is an upshift to 37°C (Hromockyj et al., 1992; Maurelli et al., 1984; Maurelli and Sansonetti, 1988; Porter and Dorman, 1994, 1997a). This increase in temperature

triggers a variety of regulatory events and mechanisms at each of the first 3 levels of the cascade that vary in their complexity. First, as seen at the *virF* promoter, a shift to 37°C triggers a change in DNA curvature that is sufficient to destabilize H-NS:DNA complexes that mediate H-NS-dependent repression of the *virF* gene ((Falconi et al., 1998; Falconi et al., 2001; Prosseda et al., 2004); Figure 5). This simple and yet elegant way of tying transcriptional derepression to a thermal upshift relies on nothing more than the DNA molecule itself, and yet it triggers the production of VirF, the first regulator of the transcriptional cascade. Second, at the *virB* promoter, even though VirF is essential for the activation of the *virB* promoter, the genome-wide increase in DNA supercoiling that occurs in response to a temperature of 37°C, maximizes transcription of the *virB* gene ((Tobe et al., 1991; Tobe et al., 1993; Tobe et al., 1995); Figure 5). Hence, a thermally induced change in DNA topology again leads to the upregulation of a gene within the regulatory cascade, but here the change in DNA topology is modulated indirectly via the essential enzyme DNA gyrase, which ultimately facilitates activation of the *virB* promoter by VirF. Third, the expression of the *virB* gene in response to a thermal upshift ultimately allows VirB to bind to its DNA recognition site at tier 3 promoters, where it relieves H-NS-mediated repression ((Castellanos et al., 2009; Stoebel et al., 2008; Turner and Dorman, 2007; Wing et al., 2004); Figure 5). Although the mechanistic details of VirB-mediated transcriptional derepression remain elusive, transcription derepression also appears to be accompanied by a change in DNA topology, at least at the *icsB* promoter (Turner and Dorman, 2007). These three examples demonstrate that *Shigella* has evolved a variety of mechanisms to tie virulence gene expression to human body temperature. Regardless, all three examples involve the modulation of DNA topology, although the importance of this in the context of transcriptional derepression by VirB remains unclear at this time.

Another example of environmental conditions impacting the transcription of virulence genes occurs in the anaerobic environment of the lower intestine and allows the “anoxic priming” of *Shigella* for cell entry (Marteyn et al., 2010). In this case, FNR-mediated repression of two genes found in the entry region of pINV increases the number and length of type III secretion needles, which facilitates contact with epithelial cells, and promotes the build-up of type III effector proteins in the bacterial cytosol, so they are ready for immediate released once host cell contact is made. Perhaps the best example of a cue that triggers the transcription of virulence genes in the host environment, however, occurs during the upregulation of tier 4 genes (Kane et al., 2002; Le Gall et al., 2005; Mavris et al., 2002a), where contact with host epithelial cells activates the *Shigella* type III secretion system. This allows the effector proteins IpaB and IpaC to be secreted from the cell and liberates their chaperone, IpgC, so that it can now function as a co-regulator with the otherwise incapacitated MxiE transcription factor. This elegant system allows the precise and coordinated expression of tier 4 genes to occur only after the type III secretion system is active and the primary wave of effectors have been secreted into the host. Clearly then, *Shigella* has evolved a variety of mechanisms to tie virulence gene expression to environmental cues encountered within its host. This ultimately allows virulence genes to be expressed in a deliberate and controlled manner, so that virulence gene products are synthesized and ready for use when needed by the bacterial pathogen.

The *Shigella* system also provides plenty of scope for those interested in the more unusual regulatory scenarios and/or processes. **H-NS and its paralogues.** While much is still to be learned about H-NS

mediated transcriptional repression, the presence of one, or sometimes two, H-NS paralogues (StpA & Sfh) in different *Shigella* strains, adds additional complexity to the *Shigella* system, making it worthy of investigation (Beloin et al., 2003; Deighan et al., 2003). While these paralogues appear to act as functional backups for H-NS, at least in the context of virulence gene expression, it remains possible that they serve additional roles either by functioning on their own or by forming complexes with the other H-NS paralogues (Beloin et al., 2003; Deighan et al., 2003). **sRNA-mediated regulation.** While it is firmly established that sRNA-mediated regulation plays a significant role in bacterial gene expression (Waters and Storz, 2009), the extent by which it controls *Shigella* virulence remains unclear. The presence of novel mechanisms of sRNA-mediated regulation in *Shigella* is also possible, based on recent findings regarding RyhB-dependent regulation of *virB* transcription (Africa et al., 2011; Broach et al., 2012; Murphy and Payne, 2007), although further studies are required to support this unusual regulation. ***mxiE* expression**. Although MxiE plays an important role in tying the expression of tier 4 genes to the activation of the type III secretion, the transcriptional slippage that must occur in order for the full-length *mxiE* gene to be transcribed is another fascinating example of transcriptional control in *Shigella* (Penno et al., 2005). While the purpose of this regulation remains unclear, it certainly is both novel and interesting from a gene expression standpoint. **pINV entry region**. Finally, the entry region of pINV, which contains the *ipa-mxi-spa* operons (which appear to range from approximately 7-11 kb in length), raises questions about the transcripts generated from this locus; what is their length, and how is the length of each transcript regulated during the transcription of these extremely long operons?

Throughout this chapter we have presented our current understanding of the transcriptional control of virulence genes on pINV, and highlighted the extent and/or limit of our knowledge to facilitate future research questions. However, the topic of where and when genes of pINV are transcribed in the human host has only been briefly discussed. This is because studies that specifically address this question are relatively rare. Of course, the ideal solution to this problem is to study transcriptional regulation of virulence genes *in situ* within human tissues. Although a lofty goal, recent advances have perhaps made this quest more attainable (Campbell-Valois and Sansonetti, 2014; Campbell-Valois et al., 2014). The use of a fast maturing green fluorescent protein as a transcriptional reporter of the MxiE-regulated *ipaH7.8* gene has revealed the dynamic and temporal regulation of this gene during the course of infection, being initially upregulated upon cell entry and subsequently downregulated when bacteria gain access to the host cell cytoplasm (Campbell-Valois et al., 2014). It is proposed that this pattern of transcription allows the replenishment of bacterial stores of this type III effector so that the cell is prepared for invasion of the neighboring cell (Campbell-Valois et al., 2014). While these studies were primarily done in the colonic epithelial cell line TC7 cells, it seems reasonable to predict that with the right animal model, suitable reporter protein and high resolution imaging system, it would be possible to observe the expression of various virulence genes in real time within animal tissue. This future direction is an exciting prospect for those interested in *Shigella* pathogenicity because it would not only allow us to identify where and when a particular gene is expressed, but may also allow us to determine the sequential expression of *Shigella* virulence genes in the human host. This information has the potential to be extremely powerful for those interested in novel drug design, because if we know that a specific gene product is only expressed in the host cell cytoplasm, the importance of delivering a drug that inhibits the action of this gene product to the host cell cytoplasm becomes abundantly clear.

In summary then, *Shigella* provides an excellent model system in which to study transcriptional silencing and anti-silencing, regulatory processes that not only control virulence gene expression in *Shigella* but also in many other bacterial pathogens. The *Shigella* system also yields an exciting array of transcriptional mechanisms that tie virulence gene expression to environmental conditions and cues encountered in the human host, making it a fascinating organism for the study of regulatory mechanisms controlling bacterial transcription. While the transcription regulation of virulence genes may seem disjointed from bacterial pathogenesis at first glance, in this chapter we have highlighted how the precise and coordinated transcriptional regulation of *Shigella* virulence is essential for the success and pathogenicity of *Shigella* species. Therefore, studies that improve understanding of transcriptional regulation of virulence genes will also undoubtedly advance our knowledge of *Shigella* pathogenesis and its interplay with its human host.

Footnote: For those interested in a broader view of virulence gene regulation, including a description of regulatory events that occur at and beyond the transcriptional level, a recent review entitled "*Shigella*: a model of virulence regulation *in vivo*" describes these events from the perspective of regulation of the T3SS; an essential virulence structure and secretory system in the pathogenicity of *Shigella* (Marteyn et al., 2012). Those specifically interested in the regulatory pathways that control the expression of virulence gene encoded by the *Shigella* chromosome are referred to recent article from Dr. Shelley Payne's laboratory that presents a helpful summary ((Pieper et al., 2013); Fig. S1).

Acknowledgments

We wish to thank Drs. Claude Parsot and Sylvie Rimsky who discussed various aspects of this chapter with us during its preparation. We also thank the editor for the opportunity to present our view of this field. Our research has been and is supported by the NIH grant AI 090573.

BIBLIOGRAPHY

Abeles, A.L., Friedman, S.A., and Austin, S.J. (1985). Partition of unit-copy miniplasmids to daughter cells. III. The DNA sequence and functional organization of the P1 partition region. Journal of molecular biology *185*, 261-272.

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

Africa, L.A., Murphy, E.R., Egan, N.R., Wigley, A.F., and Wing, H.J. (2011). The iron-responsive Fur/RyhB regulatory cascade modulates the *Shigella* outer membrane protease IcsP. Infection and immunity *79*, 4543-4549.

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

Arold, S.T., Leonard, P.G., Parkinson, G.N., and Ladbury, J.E. (2010). H-NS forms a superhelical protein scaffold for DNA condensation. Proceedings of the National Academy of Sciences of the United States of America *107*, 15728-15732.

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

Badaut, C., Williams, R., Arluison, V., Bouffartigues, E., Robert, B., Buc, H., and 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.

Baker, C.S., Morozov, I., Suzuki, K., Romeo, T., and Babitzke, P. (2002). CsrA regulates glycogen biosynthesis by preventing translation of glgC in *Escherichia coli*. Mol Microbiol *44*, 1599-1610.

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

Basta, D.W., Pew, K.L., Immak, J.A., Park, H.S., Picker, M.A., Wigley, A.F., Hensley, C.T., Pearson, J.S., Hartland, E.L., and Wing, H.J. (2013). Characterization of the *ospZ* promoter in *Shigella flexneri* and its regulation by VirB and H-NS. Journal of bacteriology *195*, 2562-2572.

Beloin, C., Deighan, P., Doyle, M., and Dorman, C.J. (2003). *Shigella flexneri* 2a strain 2457T expresses three members of the H-NS-like protein family: characterization of the Sfh protein. Molecular genetics and genomics : MGG *270*, 66-77.

Beloin, C., and 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., and Dorman, C.J. (2002). Molecular Dissection of VirB, a Key Regulator of the Virulence Cascade of *Shigella flexneri*. Journal of Biological Chemistry *277*, 15333-15344.

Bensted, H.J. (1956). Dysentery bacilli-*Shigella*; a brief historical review. Canadian journal of microbiology *2*, 163-174.

Berlutti, F., Casalino, M., Zagaglia, C., Fradiani, P.A., Visca, P., and 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. Infection and immunity *66*, 4957-4964.

Bernardini, M.L., Fontaine, A., and Sansonetti, P.J. (1990). The two-component regulatory system *ompR-envZ* controls the virulence of *Shigella flexneri*. Journal of bacteriology *172*, 6274-6281.

Bernardini, M.L., Mounier, J., d'Hauteville, H., Coquis-Rondon, M., and Sansonetti, P.J. (1989). Identification of *icsA*, a Plasmid Locus of *Shigella flexneri* that Governs Bacterial Intra- and Intercellular

Spread through Interaction with F-Actin. Proceedings of the National Academy of Sciences of the United States of America *86*, 3867-3871.
Bernardini, M.L., Sanna, M.G., Fontaine, A., and Sansonetti, P.J. (1993). OmpC is involved in invasion of epithelial cells by Shigella flexneri. Infection and immunity *61*, 3625-3635.
Bongrand, C., Sansonetti, P.J., and Parsot, C. (2012). Characterization of the promoter, MxiE box and 5' UTR of genes controlled by the activity of the type III secretion apparatus in *Shigella flexneri*. PloS one *7*, e32862.
Bouffartigues, E., Buckle, M., Badaut, C., Travers, A., and Rimsky, S. (2007). H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. Nature structural & molecular biology *14*, 441-448.
Boulette, M.L., and Payne, S.M. (2007). Anaerobic regulation of *Shigella flexneri* virulence: ArcA regulates Fur and iron acquisition genes. Journal of bacteriology *189*, 6957-6967.
Brenner, D.J., Fanning, G.R., Steigerwalt, A.G., Orskov, I., and Orskov, F. (1972). Polynucleotide sequence relatedness among three groups of pathogenic *Escherichia coli* strains. Infection and immunity *6*, 308-315.
Brescia, C.C., Kaw, M.K., and Sledjeski, D.D. (2004). The DNA binding protein H-NS binds to and alters the stability of RNA *in vitro* and *in vivo*. Journal of molecular biology *339*, 505-514.
Broach, W.H., Egan, N., Wing, H.J., Payne, S.M., and Murphy, E.R. (2012). VirF-independent regulation of *Shigella virB* transcription is mediated by the small RNA RyhB. PloS one *7*, e38592.
Browning, D.F., Grainger, D.C., and Busby, S.J. (2010). Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression. Current opinion in microbiology *13*, 773-780.
Buchrieser, C., Glaser, P., Rusniok, C., Nedjari, H., D'Hauteville, H., Kunst, F., Sansonetti, P., and Parsot, C. (2000). The virulence plasmid pWR100 and the repertoire of proteins secreted by the type III secretion apparatus of *Shigella flexneri*. Molecular Microbiology *38*, 760-771.
Cai, X., Zhang, J., Chen, M., Wu, Y., Wang, X., Chen, J., Zhang, J., Shen, X., Qu, D., and Jiang, H. (2011). The effect of the potential PhoQ histidine kinase inhibitors on *Shigella flexneri* virulence. PloS one *6*, e23100.
Campbell-Valois, F.X., and Sansonetti, P.J. (2014). Tracking bacterial pathogens with genetically-encoded reporters. FEBS letters *588*, 2428-2436.
Campbell-Valois, F.X., Schnupf, P., Nigro, G., Sachse, M., Sansonetti, P.J., and Parsot, C. (2014). A fluorescent reporter reveals on/off regulation of the *Shigella* type III secretion apparatus during entry and cell-to-cell spread. Cell host & microbe *15*, 177-189.
Castellanos, M.I., Harrison, D.J., Smith, J.M., Labahn, S.K., Levy, K.M., and 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. Journal of bacteriology *191*, 4047-4050.
Collado-Vides, J., Magasanik, B., and Gralla, J.D. (1991). Control site location and transcriptional regulation in *Escherichia coli*. Microbiological reviews *55*, 371-394.
Cukier-Kahn, R., Jacquet, M., and Gros, F. (1972). Two heat-resistant, low molecular weight proteins from *Escherichia coli* that stimulate DNA-directed RNA synthesis. Proceedings of the National Academy of Sciences of the United States of America *69*, 3643-3647.
Cusick, M.E., and Belfort, M. (1998). Domain structure and RNA annealing activity of the *Escherichia coli* regulatory protein StpA. Mol Microbiol *28*, 847-857.
Dame, R.T., Luijsterburg, M.S., Krin, E., Bertin, P.N., Wagner, R., and Wuite, G.J. (2005). DNA bridging: a property shared among H-NS-like proteins. Journal of bacteriology *187*, 1845-1848.
Dame, R.T., Wyman, C., and Goosen, N. (2000). H-NS mediated compaction of DNA visualised by atomic force microscopy. Nucleic Acids Res *28*, 3504-3510.
Dame, R.T., Wyman, C., and Goosen, N. (2001). Structural basis for preferential binding of H-NS to curved DNA. Biochimie *83*, 231-234.

Dame, R.T., Wyman, C., Wurm, R., Wagner, R., and Goosen, N. (2002). Structural basis for H-NS-mediated trapping of RNA polymerase in the open initiation complex at the *rrnB* P1. J Biol Chem *277*, 2146-2150.
Darwin, K.H., and Miller, V.L. (2001). Type III secretion chaperone-dependent regulation: activation of virulence genes by SicA and InvF in *Salmonella typhimurium*. Embo J *20*, 1850-1862.
Deighan, P., Beloin, C., and Dorman, C.J. (2003). Three-way interactions among the Sfh, StpA and H-NS nucleoid-structuring proteins of *Shigella flexneri* 2a strain 2457T. Mol Microbiol *48*, 1401-1416.
Deighan, P., Free, A., and Dorman, C.J. (2000). A role for the *Escherichia coli* H-NS-like protein StpA in OmpF porin expression through modulation of *micF* RNA stability. Mol Microbiol *38*, 126-139.
Demers, B., Sansonetti, P.J., and Parsot, C. (1998). Induction of type III secretion in *Shigella flexneri* is associated with differential control of transcription of genes encoding secreted proteins. EMBO J *17*, 2894-2903.
Dorman, C.J. (1992). The VirF protein from *Shigella flexneri* is a member of the AraC transcription factor superfamily and is highly homologous to Rns, a positive regulator of virulence genes in enterotoxigenic *Escherichia coli*. Mol Microbiol *6*, 1575.
Dorman, C.J. (2004). H-NS: a universal regulator for a dynamic genome. Nat Rev Microbiol *2*, 391-400.
Dorman, C.J. (2007). H-NS, the genome sentinel. Nat Rev Microbiol *5*, 157-161.
Dorman, C.J. (2009). The virulence plasmids of *Shigella flexneri*. In Microbial megaplasmids (Springer), pp. 151-170.
Dorman, C.J., Bhriain, N.N., and Higgins, C.F. (1990). DNA supercoiling and environmental regulation of virulence gene expression in *Shigella flexneri*. Nature *344*, 789-792.
Dorman, C.J., Hinton, J.C., and Free, A. (1999). Domain organization and oligomerization among H-NS-like nucleoid-associated proteins in bacteria. Trends Microbiol *7*, 124-128.
Dorman, C.J., McKenna, S., and Beloin, C. (2001). Regulation of virulence gene expression in *Shigella flexneri*, a facultative intracellular pathogen. Int J Med Microbiol *291*, 89-96.
Dorman, C.J., and Porter, M.E. (1998). The *Shigella* virulence gene regulatory cascade: a paradigm of bacterial gene control mechanisms. Molecular Microbiology *29*, 677-684.
Dubey, A.K., Baker, C.S., Suzuki, K., Jones, A.D., Pandit, P., Romeo, T., and Babitzke, P. (2003). CsrA regulates translation of the *Escherichia coli* carbon starvation gene, *cstA*, by blocking ribosome access to the *cstA* transcript. Journal of bacteriology *185*, 4450-4460.
DuPont, H.L., Levine, M.M., Hornick, R.B., and Formal, S.B. (1989). Inoculum size in shigellosis and implications for expected mode of transmission. The Journal of infectious diseases *159*, 1126-1128.
Durand, J.M., and Bjork, G.R. (2003). Putrescine or a combination of methionine and arginine restores virulence gene expression in a tRNA modification-deficient mutant of *Shigella flexneri*: a possible role in adaptation of virulence. Mol Microbiol *47*, 519-527.
Durand, J.M., Bjork, G.R., Kuwae, A., Yoshikawa, M., and Sasakawa, C. (1997). The modified nucleoside 2-methylthio-N6-isopentenyladenosine in tRNA of *Shigella flexneri* is required for expression of virulence genes. Journal of bacteriology *179*, 5777-5782.
Durand, J.M., Dagberg, B., Uhlin, B.E., and Bjork, G.R. (2000). Transfer RNA modification, temperature and DNA superhelicity have a common target in the regulatory network of the virulence of *Shigella flexneri*: the expression of the *virF* gene. Mol Microbiol *35*, 924-935.
Durand, J.M., Okada, N., Tobe, T., Watarai, M., Fukuda, I., Suzuki, T., Nakata, N., Komatsu, K., Yoshikawa, M., and Sasakawa, C. (1994). *vacC*, a virulence-associated chromosomal locus of *Shigella flexneri,* is homologous to *tgt*, a gene encoding tRNA-guanine transglycosylase (Tgt) of *Escherichia coli* K-12. Journal of bacteriology *176*, 4627-4634.
Egile, C., d'Hauteville, H., Parsot, C., and Sansonetti, P.J. (1997). SopA, the outer membrane protease responsible for polar localization of IcsA in *Shigella flexneri*. Mol Microbiol *23*, 1063-1073.

Eisenreich, W., Dandekar, T., Heesemann, J., and Goebel, W. (2010). Carbon metabolism of intracellular bacterial pathogens and possible links to virulence. Nat Rev Microbiol *8*, 401-412.
Esposito, D., Petrovic, A., Harris, R., Ono, S., Eccleston, J.F., Mbabaali, A., Haq, I., Higgins, C.F., Hinton, J.C., Driscoll, P.C*., et al.* (2002). H-NS oligomerization domain structure reveals the mechanism for high order self-association of the intact protein. Journal of molecular biology *324*, 841-850.
Ewing, W.H. (1949). *Shigella* Nomenclature. Journal of bacteriology *57*, 633-638.
Ewing, W.H. (1986). Edward and Ewing's Identification of the Enterobacteriaceae (Elsevier Science Publishers, Amsterdam, The Netherlands).
Falconi, M., Colonna, B., Prosseda, G., Micheli, G., and 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.
Falconi, M., Prosseda, G., Giangrossi, M., Beghetto, E., and 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.
Fang, F.C., and Rimsky, S. (2008). New insights into transcriptional regulation by H-NS. Current opinion in microbiology *11*, 113-120.
Free, A., and Dorman, C.J. (1995). Coupling of *Escherichia coli hns* mRNA levels to DNA synthesis by autoregulation: implications for growth phase control. Mol Microbiol *18*, 101-113.
Free, A., and Dorman, C.J. (1997). The *Escherichia coli stpA* gene is transiently expressed during growth in rich medium and is induced in minimal medium and by stress conditions. Journal of bacteriology *179*, 909-918.
Free, A., Williams, R.M., and Dorman, C.J. (1998). The StpA protein functions as a molecular adapter to mediate repression of the *bgl* operon by truncated H-NS in *Escherichia coli*. Journal of bacteriology *180*, 994-997.
Gallegos, M.T., Schleif, R., Bairoch, A., Hofmann, K., and Ramos, J.L. (1997). Arac/XylS family of transcriptional regulators. Microbiology and molecular biology reviews : MMBR *61*, 393-410.
Gao, X., Zou, T., Mu, Z., Qin, B., Yang, J., Waltersperger, S., Wang, M., Cui, S., and Jin, Q. (2013). Structural insights into VirB-DNA complexes reveal mechanism of transcriptional activation of virulence genes. Nucleic Acids Res.
Giangrossi, M., Prosseda, G., Tran, C.N., Brandi, A., Colonna, B., and Falconi, M. (2010). A novel antisense RNA regulates at transcriptional level the virulence gene *icsA* of *Shigella flexneri*. Nucleic Acids Res *38*, 3362-3375.
Goldberg, M.B., Barzu, O., Parsot, C., and Sansonetti, P.J. (1993). Unipolar localization and ATPase activity of IcsA, a *Shigella flexneri* protein involved in intracellular movement. Infectious agents and disease *2*, 210-211.
Goldman, S.R., Tu, Y., and Goldberg, M.B. (2008). Differential regulation by magnesium of the two MsbB paralogs of *Shigella flexneri*. Journal of bacteriology *190*, 3526-3537.
Goosen, N., and van de Putte, P. (1995). The regulation of transcription initiation by integration host factor. Molecular Microbiology *16*, 1-7.
Gordon, B.R., Li, Y., Cote, A., Weirauch, M.T., Ding, P., Hughes, T.R., Navarre, W.W., Xia, B., and Liu, J. (2011). Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. Proceedings of the National Academy of Sciences of the United States of America *108*, 10690-10695.
Gore, A.L., and Payne, S.M. (2010). CsrA and Cra Influence *Shigella flexneri* Pathogenesis. Infection and immunity *78*, 4674-4682.
Harrison, D.J. (2010). Transcriptional regulation of the *Shigella flexneri icsP* promoter: Silencing and anti-silencing by H-NS and VirB. In School of Life Sciences (Las Vegas, Nevada, USA.: University of Nevada, Las Vegas), pp. 145.

Hensley, C.T., Kamneva, O.K., Levy, K.M., Labahn, S.K., Africa, L.A., and Wing, H.J. (2011). Two promoters and two translation start sites control the expression of the *Shigella flexneri* outer membrane protease IcsP. Arch Microbiol *193*, 263-274.
Hromockyj, A.E., Tucker, S.C., and 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.
Hurt, J.K., Olgen, S., and Garcia, G.A. (2007). Site-specific modification of *Shigella flexneri virF* mRNA by tRNA-guanine transglycosylase *in vitro*. Nucleic Acids Res *35*, 4905-4913.
Jacquet, M., Cukier-Kahn, R., Pla, J., and Gros, F. (1971). A thermostable protein factor acting on *in vitro* DNA transcription. Biochemical and biophysical research communications *45*, 1597-1607.
Johansson, J., Dagberg, B., Richet, E., and Uhlin, B.E. (1998). H-NS and StpA proteins stimulate expression of the maltose regulon in *Escherichia coli*. Journal of bacteriology *180*, 6117-6125.
Jordi, B.J., Dagberg, B., de Haan, L.A., Hamers, A.M., van der Zeijst, B.A., Gaastra, W., and Uhlin, B.E. (1992). The positive regulator CfaD overcomes the repression mediated by histone-like protein H-NS (H1) in the CFA/I fimbrial operon of *Escherichia coli*. EMBO J *11*, 2627-2632.
Jost, B.H., and Adler, B. (1993). Site of transcriptional activation of *virB* on the large plasmid of *Shigella flexneri* 2a by VirF, a member of the AraC family of transcriptional activators. Microb Pathog *14*, 481-488.
Kane, C.D., Schuch, R., Day, W.A., Jr., and Maurelli, A.T. (2002). MxiE regulates intracellular expression of factors secreted by the *Shigella flexneri* 2a type III secretion system. Journal of bacteriology *184*, 4409-4419.
Kane, K.A., and Dorman, C.J. (2012). VirB-Mediated Positive Feedback Control of the Virulence Gene Regulatory Cascade of *Shigella flexneri*. Journal of bacteriology *194*, 5264-5273.
Lammi, M., Paci, M., Pon, C.L., Losso, M.A., Miano, A., Pawlik, R.T., Gianfranceschi, G.L., and Gualerzi, C.O. (1984). Proteins from the prokaryotic nucleoid: biochemical and 1H NMR studies on three bacterial histone-like proteins. Advances in experimental medicine and biology *179*, 467-477.
Lan, R., Alles, M.C., Donohoe, K., Martinez, M.B., and Reeves, P.R. (2004). Molecular evolutionary relationships of enteroinvasive *Escherichia coli* and *Shigella* spp. Infection and immunity *72*, 5080-5088.
Lan, R., and Reeves, P.R. (2002). *Escherichia coli* in disguise: molecular origins of *Shigella*. Microbes and infection / Institut Pasteur *4*, 1125-1132.
Lan, R., Stevenson, G., and Reeves, P.R. (2003). Comparison of two major forms of the *Shigella* virulence plasmid pINV: positive selection is a major force driving the divergence. Infection and immunity *71*, 6298-6306.
Lang, B., Blot, N., Bouffartigues, E., Buckle, M., Geertz, M., Gualerzi, C.O., Mavathur, R., Muskhelishvili, G., Pon, C.L., Rimsky, S*., et al.* (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., and 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, D.J., Minchin, S.D., and Busby, S.J. (2012). Activating transcription in bacteria. Annual review of microbiology *66*, 125-152.
Lett, M.C., Sasakawa, C., Okada, N., Sakai, T., Makino, S., Yamada, M., Komatsu, K., and Yoshikawa, M. (1989). *virG*, a plasmid-coded virulence gene of *Shigella flexneri*: identification of the *virG* protein and determination of the complete coding sequence. Journal of bacteriology *171*, 353-359.
Lim, C.J., Kenney, L.J., and Yan, J. (2014). Single-molecule studies on the mechanical interplay between DNA supercoiling and H-NS DNA architectural properties. Nucleic Acids Res *42*, 8369-8378.
Lim, C.J., Lee, S.Y., Kenney, L.J., and Yan, J. (2012). Nucleoprotein filament formation is the structural basis for bacterial protein H-NS gene silencing. Scientific reports *2*, 509.

Liu, L.F., and Wang, J.C. (1987). Supercoiling of the DNA template during transcription. Proceedings of the National Academy of Sciences of the United States of America *84*, 7024-7027.
Liu, M.Y., Yang, H., and Romeo, T. (1995). The product of the pleiotropic *Escherichia coli gene csrA* modulates glycogen biosynthesis via effects on mRNA stability. Journal of bacteriology *177*, 2663-2672.
Liu, Y., Chen, H., Kenney, L.J., and Yan, J. (2010). A divalent switch drives H-NS/DNA-binding conformations between stiffening and bridging modes. Genes & development *24*, 339-344.
Lucchini, S., Rowley, G., Goldberg, M.D., Hurd, D., Harrison, M., and Hinton, J.C.D. (2006). H-NS Mediates the Silencing of Laterally Acquired Genes in Bacteria. PLoS Pathogens *2*, 0746-0752.
Magdalena, J., Hachani, A., Chamekh, M., Jouihri, N., Gounon, P., Blocker, A., and Allaoui, A. (2002). Spa32 regulates a switch in substrate specificity of the type III secreton of *Shigella flexneri* from needle components to Ipa proteins. Journal of bacteriology *184*, 3433-3441.
Marteyn, B., Gazi, A., and Sansonetti, P. (2012). *Shigella*: a model of virulence regulation *in vivo*. Gut microbes *3*, 104-120.
Marteyn, B., West, N.P., Browning, D.F., Cole, J.A., Shaw, J.G., Palm, F., Mounier, J., Prevost, M.C., Sansonetti, P., and Tang, C.M. (2010). Modulation of *Shigella* virulence in response to available oxygen *in vivo*. Nature *465*, 355-358.
Martin, R.G., and Rosner, J.L. (2001). The AraC transcriptional activators. Current opinion in microbiology *4*, 132-137.
Massé, E., and Gottesman, S. (2002). A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. Proceedings of the National Academy of Sciences *99*, 4620-4625.
Maurelli, A.T., Blackmon, B., and Curtiss, R., 3rd (1984). Temperature-dependent expression of virulence genes in *Shigella* species. Infection and immunity *43*, 195-201.
Maurelli, A.T., and Sansonetti, P.J. (1988). Identification of a Chromosomal Gene Controlling Temperature-Regulated Expression of *Shigella* Virulence. Proceedings of the National Academy of Sciences of the United States of America *85*, 2820-2824.
Maurer, S., Fritz, J., and Muskhelishvili, G. (2009). A systematic *in vitro* study of nucleoprotein complexes formed by bacterial nucleoid-associated proteins revealing novel types of DNA organization. Journal of molecular biology *387*, 1261-1276.
Mavris, M., Page, A.L., Tournebize, R., Demers, B., Sansonetti, P., and Parsot, C. (2002a). Regulation of transcription by the activity of the *Shigella flexneri* type III secretion apparatus. Mol Microbiol *43*, 1543-1553.
Mavris, M., Sansonetti, P.J., and Parsot, C. (2002b). Identification of the *cis*-acting site involved in activation of promoters regulated by activity of the type III secretion apparatus in *Shigella flexneri*. Journal of bacteriology *184*, 6751-6759.
McKenna, S., Beloin, C., and Dorman, C.J. (2003). *In vitro* DNA-binding properties of VirB, the *Shigella flexneri* virulence regulatory protein. FEBS letters *545*, 183-187.
Menard, R., Sansonetti, P., Parsot, C., and Vasselon, T. (1994). Extracellular association and cytoplasmic partitioning of the IpaB and IpaC invasins of *S. flexneri*. Cell *79*, 515-525.
Miller, J. (1972). Experiments in Molecular Genetics. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
Mitobe, J., Arakawa, E., and Watanabe, H. (2005). A sensor of the two-component system CpxA affects expression of the type III secretion system through posttranscriptional processing of InvE. Journal of bacteriology *187*, 107-113.
Mitobe, J., Morita-Ishihara, T., Ishihama, A., and Watanabe, H. (2008). Involvement of RNA-binding protein Hfq in the post-transcriptional regulation of *invE* gene expression in *Shigella sonnei*. J Biol Chem *283*, 5738-5747.

Mitobe, J., Morita-Ishihara, T., Ishihama, A., and Watanabe, H. (2009). Involvement of RNA-binding protein Hfq in the osmotic-response regulation of *invE* gene expression in *Shigella sonnei*. BMC microbiology *9*, 110.
Mori, H., Kondo, A., Ohshima, A., Ogura, T., and Hiraga, S. (1986). Structure and function of the F plasmid genes essential for partitioning. Journal of molecular biology *192*, 1-15.
Moss, J.E., Fisher, P.E., Vick, B., Groisman, E.A., and Zychlinsky, A. (2000). The regulatory protein PhoP controls susceptibility to the host inflammatory response in *Shigella flexneri*. Cellular microbiology *2*, 443-452.
Munson, G.P. (2013). Virulence regulons of enterotoxigenic *Escherichia coli*. Immunologic research *57*, 229-236.
Munson, G.P., Holcomb, L.G., and Scott, J.R. (2001). Novel group of virulence activators within the AraC family that are not restricted to upstream binding sites. Infection and immunity *69*, 186-193.
Munson, G.P., and Scott, J.R. (1999). Binding site recognition by Rns, a virulence regulator in the AraC family. Journal of bacteriology *181*, 2110-2117.
Munson, G.P., and Scott, J.R. (2000). Rns, a virulence regulator within the AraC family, requires binding sites upstream and downstream of its own promoter to function as an activator. Mol Microbiol *36*, 1391-1402.
Murphy, E.R., and Payne, S.M. (2007). RyhB, an iron-responsive small RNA molecule, regulates *Shigella dysenteriae* virulence. Infection and immunity *75*, 3470-3477.
Nakata, N., Sasakawa, C., Okada, N., Tobe, T., Fukuda, I., Suzuki, T., Komatsu, K., and Yoshikawa, M. (1992). Identification and characterization of *virK*, a virulence-associated large plasmid gene essential for intercellular spreading of *Shigella flexneri*. Mol Microbiol *6*, 2387-2395.
Nakayama, S., and Watanabe, H. (1995). Involvement of *cpxA*, a sensor of a two-component regulatory system, in the pH-dependent regulation of expression of *Shigella sonnei virF* gene. Journal of bacteriology *177*, 5062-5069.
Nakayama, S., and Watanabe, H. (1998). Identification of *cpxR* as a positive regulator essential for expression of the *Shigella sonnei virF* gene. Journal of bacteriology *180*, 3522-3528.
Navarre, W.W., Halsey, T.A., Walthers, D., Frye, J., McClelland, M., Potter, J.L., Kenney, L.J., Gunn, J.S., Fang, F.C., and Libby, S.J. (2005). Co-regulation of *Salmonella enterica* genes required for virulence and resistance to antimicrobial peptides by SlyA and PhoP/PhoQ. Mol Microbiol *56*, 492-508.
Ochman, H., Lawrence, J.G., and Groisman, E.A. (2000). Lateral gene transfer and the nature of bacterial innovation. Nature *405*, 299-304.
Ochman, H., Whittam, T.S., Caugant, D.A., and Selander, R.K. (1983). Enzyme polymorphism and genetic population structure in *Escherichia coli* and *Shigella*. Journal of general microbiology *129*, 2715-2726.
Oglesby, A.G., Murphy, E.R., Iyer, V.R., and Payne, S.M. (2005). Fur regulates acid resistance in S*higella flexneri* via RyhB and ydeP. Mol Microbiol *58*, 1354-1367.
Opel, M.L., Aeling, K.A., Holmes, W.M., Johnson, R.C., Benham, C.J., and Hatfield, G.W. (2004). Activation of transcription initiation from a stable RNA promoter by a Fis protein-mediated DNA structural transmission mechanism. Mol Microbiol *53*, 665-674.
Page, A.L., Fromont-Racine, M., Sansonetti, P., Legrain, P., and Parsot, C. (2001). Characterization of the interaction partners of secreted proteins and chaperones of *Shigella flexneri*. Mol Microbiol *42*, 1133-1145.
Page, A.L., Sansonetti, P., and Parsot, C. (2002). Spa15 of *Shigella flexneri*, a third type of chaperone in the type III secretion pathway. Mol Microbiol *43*, 1533-1542.
Parsot, C., Ageron, E., Penno, C., Mavris, M., Jamoussi, K., d'Hauteville, H., Sansonetti, P., and Demers, B. (2005). A secreted anti-activator, OspD1, and its chaperone, Spa15, are involved in the control of transcription by the type III secretion apparatus activity in *Shigella flexneri*. Mol Microbiol *56*, 1627-1635.

Penno, C., Sansonetti, P., and Parsot, C. (2005). Frameshifting by transcriptional slippage is involved in production of MxiE, the transcription activator regulated by the activity of the type III secretion apparatus in *Shigella flexneri*. Mol Microbiol *56*, 204-214.
Pieper, R., Fisher, C.R., Suh, M.J., Huang, S.T., Parmar, P., and Payne, S.M. (2013). Analysis of the proteome of intracellular *Shigella flexneri* reveals pathways important for intracellular growth. Infection and immunity *81*, 4635-4648.
Pilonieta, M.C., and Munson, G.P. (2008). The chaperone IpgC copurifies with the virulence regulator MxiE. Journal of bacteriology *190*, 2249-2251.
Porter, M.E., and Dorman, C.J. (1994). A role for H-NS in the thermo-osmotic regulation of virulence gene expression in *Shigella flexneri*. Journal of bacteriology *176*, 4187-4191.
Porter, M.E., and Dorman, C.J. (1997a). Differential regulation of the plasmid-encoded genes in the *Shigella flexneri* virulence regulon. Molec Gen Genet *256*, 93-103.
Porter, M.E., and Dorman, C.J. (1997b). Positive regulation of *Shigella flexneri* virulence genes by integration host factor. Journal of bacteriology *179*, 6537-6550.
Porter, M.E., and Dorman, C.J. (2002). *In Vivo* DNA-Binding and Oligomerization Properties of the *Shigella flexneri* AraC-Like Transcriptional Regulator VirF as Identified by Random and Site-Specific Mutagenesis. Journal of bacteriology *184*, 531-539.
Porter, M.E., Smith, S.G., and Dorman, C.J. (1998). Two highly related regulatory proteins, *Shigella flexneri* VirF and enterotoxigenic *Escherichia coli* Rns, have common and distinct regulatory properties. FEMS microbiology letters *162*, 303-309.
Prosseda, G., Falconi, M., Giangrossi, M., Gualerzi, C.O., Micheli, G., and 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., and 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.
Pupo, G.M., Karaolis, D.K., Lan, R., and Reeves, P.R. (1997). Evolutionary relationships among pathogenic and nonpathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and *mdh* sequence studies. Infection and immunity *65*, 2685-2692.
Pupo, G.M., Lan, R., and Reeves, P.R. (2000). Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics. Proceedings of the National Academy of Sciences of the United States of America *97*, 10567-10572.
Radnedge, L., Davis, M.A., and Austin, S.J. (1996). P1 and P7 plasmid partition: ParB protein bound to its partition site makes a separate discriminator contact with the DNA that determines species specificity. Embo J *15*, 1155-1162.
Radnedge, L., Davis, M.A., Youngren, B., and Austin, S.J. (1997). Plasmid maintenance functions of the large virulence plasmid of *Shigella flexneri*. Journal of bacteriology *179*, 3670-3675.
Raivio, T.L. (2014). Everything old is new again: an update on current research on the Cpx envelope stress response. Biochimica et biophysica acta *1843*, 1529-1541.
Rediers, H., Vanderleyden, J., and De Mot, R. (2004). *Azotobacter vinelandii*: a Pseudomonas in disguise? Microbiology *150*, 1117-1119.
Renzoni, D., Esposito, D., Pfuhl, M., Hinton, J.C., Higgins, C.F., Driscoll, P.C., and Ladbury, J.E. (2001). Structural characterization of the N-terminal oligomerization domain of the bacterial chromatin-structuring protein, H-NS. Journal of molecular biology *306*, 1127-1137.
Rice, P.A., Yang, S., Mizuuchi, K., and Nash, H.A. (1996). Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. Cell *87*, 1295-1306.
Rimsky, S., and Travers, A. (2011). Pervasive regulation of nucleoid structure and function by nucleoid-associated proteins. Current opinion in microbiology *14*, 136-141.

Rodionov, O., Lobocka, M., and Yarmolinsky, M. (1999). Silencing of genes flanking the P1 plasmid centromere. Science *283*, 546-549.
Rodionov, O., and Yarmolinsky, M. (2004). Plasmid partitioning and the spreading of P1 partition protein ParB. Mol Microbiol *52*, 1215-1223.
Rolland, K., Lambert-Zechovsky, N., Picard, B., and Denamur, E. (1998). *Shigella* and enteroinvasive *Escherichia coli* strains are derived from distinct ancestral strains of *E. coli*. Microbiology *144 (Pt 9)*, 2667-2672.
Romeo, T. (1998). Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB. Mol Microbiol *29*, 1321-1330.
Sabnis, N.A., Yang, H., and Romeo, T. (1995). Pleiotropic regulation of central carbohydrate metabolism in *Escherichia coli* via the gene *csrA*. J Biol Chem *270*, 29096-29104.
Saier, M.H., Jr., and Ramseier, T.M. (1996). The catabolite repressor/activator (Cra) protein of enteric bacteria. Journal of bacteriology *178*, 3411-3417.
Sakai, T., Sasakawa, C., Makino, S., Kamata, K., and Yoshikawa, M. (1986). Molecular cloning of a genetic determinant for Congo red binding ability which is essential for the virulence of *Shigella flexneri*. Infection and immunity *51*, 476-482.
Sakai, T., Sasakawa, C., and Yoshikawa, M. (1988). Expression of four virulence antigens of *Shigella flexneri* is positively regulated at the transcriptional level by the 30 kilo Dalton *virF* protein. Molecular Microbiology *2*, 589-597.
Santapaola, D., Casalino, M., Petrucca, A., Presutti, C., Zagaglia, C., Berlutti, F., Colonna, B., and Nicoletti, M. (2002). Enteroinvasive *Escherichia coli* virulence-plasmid-carried apyrase (*apy*) and *ospB* genes are organized as a bicistronic operon and are subject to differential expression. Microbiology *148*, 2519-2529.
Sasakawa, C., Kamata, K., Sakai, T., Makino, S., Yamada, M., Okada, N., and Yoshikawa, M. (1988). Virulence-associated genetic regions comprising 31 kilobases of the 230-kilobase plasmid in *Shigella flexneri* 2a. Journal of bacteriology *170*, 2480-2484.
Savelkoul, P.H., Willshaw, G.A., McConnell, M.M., Smith, H.R., Hamers, A.M., van der Zeijst, B.A., and Gaastra, W. (1990). Expression of CFA/I fimbriae is positively regulated. Microb Pathog *8*, 91-99.
Schneider, R., Travers, A., Kutateladze, T., and Muskhelishvili, G. (1999). A DNA architectural protein couples cellular physiology and DNA topology in *Escherichia coli*. Mol Microbiol *34*, 953-964.
Schroeder, G.N., and Hilbi, H. (2008). Molecular Pathogenesis of *Shigella* spp.: Controlling Host Cell Signaling, Invasion, and Death by Type III Secretion. Clinical Microbiology Reviews *21*, 134-156.
Schuch, R., and Maurelli, A.T. (1997). Virulence plasmid instability in *Shigella flexneri* 2a is induced by virulence gene expression. Infection and immunity *65*, 3686-3692.
Schuch, R., and Maurelli, A.T. (2001). Spa33, a cell surface-associated subunit of the Mxi-Spa type III secretory pathway of *Shigella flexneri*, regulates Ipa protein traffic. Infection and immunity *69*, 2180-2189.
Schumacher, M.A., and Funnell, B.E. (2005). Structures of ParB bound to DNA reveal mechanism of partition complex formation. Nature *438*, 516-519.
Sette, M., Spurio, R., Trotta, E., Brandizi, C., Brandi, A., Pon, C.L., Barbato, G., Boelens, R., and Gualerzi, C.O. (2009). Sequence-specific recognition of DNA by the C-terminal domain of nucleoid-associated protein H-NS. J Biol Chem *284*, 30453-30462.
Shere, K.D., Sallustio, S., Manessis, A., D'Aversa, T.G., and Goldberg, M.B. (1997). Disruption of IcsP, the major *Shigella* protease that cleaves IcsA, accelerates actin-based motility. Mol Microbiol *25*, 451-462.
Sidik, S., Kottwitz, H., Benjamin, J., Ryu, J., Jarrar, A., Garduno, R., and Rohde, J.R. (2014). A *Shigella flexneri* Virulence Plasmid Encoded Factor Controls Production of Outer Membrane Vesicles. G3 *4*, 2493-2503.

Spassky, A., Rimsky, S., Garreau, H., and 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.
Sperandio, B., Regnault, B., Guo, J., Zhang, Z., Stanley, S.L., Jr., Sansonetti, P.J., and Pedron, T. (2008). Virulent *Shigella flexneri* subverts the host innate immune response through manipulation of antimicrobial peptide gene expression. The Journal of experimental medicine *205*, 1121-1132.
Spurio, R., Durrenberger, M., Falconi, M., La Teana, A., Pon, C.L., and Gualerzi, C.O. (1992). Lethal overproduction of the *Escherichia coli* nucleoid protein H-NS: ultramicroscopic and molecular autopsy. Molecular & general genetics : MGG *231*, 201-211.
Steinhauer, J., Agha, R., Pham, T., Varga, A.W., and Goldberg, M.B. (1999). The unipolar *Shigella* surface protein IcsA is targeted directly to the bacterial old pole: IcsP cleavage of IcsA occurs over the entire bacterial surface. Mol Microbiol *32*, 367-377.
Stoebel, D.M., Free, A., and 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., and Funnell, B.E. (1999). P1 ParB domain structure includes two independent multimerization domains. Journal of bacteriology *181*, 5898-5908.
Surtees, J.A., and 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.
Suzuki, T., Ueguchi, C., and Mizuno, T. (1996). H-NS regulates OmpF expression through *micF* antisense RNA in *Escherichia coli*. Journal of bacteriology *178*, 3650-3653.
Tamano, K., Katayama, E., Toyotome, T., and Sasakawa, C. (2002). *Shigella* Spa32 is an essential secretory protein for functional type III secretion machinery and uniformity of its needle length. Journal of bacteriology *184*, 1244-1252.
Taniya, T., Mitobe, J., Nakayama, S., Mingshan, Q., Okuda, K., and 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. Journal of bacteriology *185*, 5158-5165.
Tendeng, C., and Bertin, P.N. (2003). H-NS in Gram-negative bacteria: a family of multifaceted proteins. Trends Microbiol *11*, 511-518.
Tobe, T., Nagai, S., Okada, N., Adler, B., Yoshikawa, M., and 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.
Tobe, T., Yoshikawa, M., Mizuno, T., and Sasakawa, C. (1993). Transcriptional control of the invasion regulatory gene *virB* of *Shigella flexneri*: activation by *virF* and repression by H-NS. Journal of bacteriology *175*, 6142-6149.
Tobe, T., Yoshikawa, M., and Sasakawa, C. (1995). Thermoregulation of *virB* transcription in *Shigella flexneri* by sensing of changes in local DNA superhelicity. Journal of bacteriology *177*, 1094-1097.
Tolstorukov, M.Y., Virnik, K.M., Adhya, S., and Zhurkin, V.B. (2005). A-tract clusters may facilitate DNA packaging in bacterial nucleoid. Nucleic Acids Res *33*, 3907-3918.
Tran, C.N., Giangrossi, M., Prosseda, G., Brandi, A., Di Martino, M.L., Colonna, B., and Falconi, M. (2011). A multifactor regulatory circuit involving H-NS, VirF and an antisense RNA modulates transcription of the virulence gene *icsA* of *Shigella flexneri*. Nucleic Acids Res *39*, 8122-8134.
Turner, E.C., and 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. Journal of bacteriology *189*, 3403-3413.
Uchiya, K., Tobe, T., Komatsu, K., Suzuki, T., Watarai, M., Fukuda, I., Yoshikawa, M., and Sasakawa, C. (1995). Identification of a novel virulence gene, *virA*, on the large plasmid of *Shigella*, involved in invasion and intercellular spreading. Mol Microbiol *17*, 241-250.

Ulissi, U., Fabbretti, A., Sette, M., Giuliodori, A.M., and Spurio, R. (2014). Time-resolved assembly of a nucleoprotein complex between *Shigella flexneri virF* promoter and its transcriptional repressor H-NS. Nucleic Acids Res *42*, 13039-13050.
Vecchiarelli, A.G., and Funnell, B.E. (2013). Probing the N-terminus of ParB using cysteine-scanning mutagenesis and thiol modification. Plasmid *70*, 86-93.
Walthers, D., Li, Y., Liu, Y., Anand, G., Yan, J., and Kenney, L.J. (2011). *Salmonella enterica* response regulator SsrB relieves H-NS silencing by displacing H-NS bound in polymerization mode and directly activates transcription. J Biol Chem *286*, 1895-1902.
Watanabe, H., Arakawa, E., Ito, K., Kato, J., and 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. Journal of bacteriology *172*, 619-629.
Waters, L.S., and Storz, G. (2009). Regulatory RNAs in bacteria. Cell *136*, 615-628.
Wei, J., Goldberg, M.B., Burland, V., Venkatesan, M.M., Deng, W., Fournier, G., Mayhew, G.F., Plunkett, G., 3rd, Rose, D.J., Darling, A*., et al.* (2003). Complete genome sequence and comparative genomics of *Shigella flexneri* serotype 2a strain 2457T. Infection and immunity *71*, 2775-2786.
Williams, R.M., and Rimsky, S. (1997). Molecular aspects of the *E. coli* nucleoid protein, H-NS: a central controller of gene regulatory networks. FEMS microbiology letters *156*, 175-185.
Williams, R.M., Rimsky, S., and Buc, H. (1996). Probing the structure, function, and interactions of the *Escherichia coli* H-NS and StpA proteins by using dominant negative derivatives. Journal of bacteriology *178*, 4335-4343.
Wilson, R.L., Libby, S.J., Freet, A.M., Boddicker, J.D., Fahlen, T.F., and Jones, B.D. (2001). Fis, a DNA nucleoid-associated protein, is involved in *Salmonella typhimurium* SPI-1 invasion gene expression. Mol Microbiol *39*, 79-88.
Wing, H.J., Goldman, S.R., Ally, S., and Goldberg, M.B. (2005). Modulation of an outer membrane protease contributes to the virulence defect of *Shigella flexneri* strains carrying a mutation in the *virK* locus. Infection and immunity *73*, 1217-1220.
Wing, H.J., Yan, A.W., Goldman, S.R., and 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. Journal of bacteriology *186*, 699-705.
Wolf, T., Janzen, W., Blum, C., and Schnetz, K. (2006). Differential dependence of StpA on H-NS in autoregulation of *stpA* and in regulation of *bgl*. Journal of bacteriology *188*, 6728-6738.
Yamada, H., Muramatsu, S., and Mizuno, T. (1990). An *Escherichia coli* protein that preferentially binds to sharply curved DNA. Journal of biochemistry *108*, 420-425.
Yamada, H., Yoshida, T., Tanaka, K., Sasakawa, C., and Mizuno, T. (1991). Molecular analysis of the *Escherichia coli hns* gene encoding a DNA-binding protein, which preferentially recognizes curved DNA sequences. Molecular & general genetics : MGG *230*, 332-336.
Zhang, A., and Belfort, M. (1992). Nucleotide sequence of a newly-identified *Escherichia coli* gene, *stpA*, encoding an H-NS-like protein. Nucleic Acids Res *20*, 6735.
Zhang, A., Rimsky, S., Reaban, M.E., Buc, H., and 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.

[1]Table 1. Clustering of *Shigella* species and serotypes based on the type of virulence plasmid carried.

Cluster	Strain	Virulence plasmid carried (pINV A or B)
***Shigella* cluster 1**	*S. dysenteriae* serotypes 3-7, 9, 11-13	A
	S. boydii serotypes 1-4, 6, 8, 10, 14, 18	A
	S. flexneri serotype 6 & 6a	A
***Shigella* cluster 2**	*S. boydii* serotypes 5, 7, 9, 15-17	A
	S. boydii serotype 11	B
	S. dysenteriae serotype 2	A
***Shigella* cluster 3**	*S. flexneri* 1-5 and subtypes	B
	S. boydii serotype 12	B
Outliers	S. *sonnei*	Either A or B
	S. *dysenteriae* serotypes 8 & 10	Either A or B
	S. *dysenteriae* serotype 1	Distinct form of pINV
	S. *boydii* serotype 13	No plasmid detected

[1]Adapted from phylogenetic trees that were constructed based on the sequence of three genes in the *ipa-mxi-spa* genetic locus on pINV (*ipgD*, *mxiC* and *mxiA*) (Lan and Reeves, 2002).

Figure Legends

Figure 1. Proposed mechanisms of H-NS-mediated transcriptional silencing. Six proposed mechanisms of H-NS-mediated repression are depicted. Each is predicted to occur through one of two distinct H-NS binding modes: filament formation (a-c) or DNA bridging (d-f). These result in either the exclusion of RNAP from the promoter (a, d) or the blockage of transcription elongation (b, c, e, f). RNAP represents RNA polymerase. Small, white circles represent H-NS monomers. Line with arrow represents the transcription start site. Dashed, semicircle line represents a DNA twist. "+" symbol represents positive supercoils.

Figure 2. Overview of the regulatory cascade controlling *Shigella* virulence. The four tiers of the regulatory cascade that control gene expression on pINV are depicted along with relative levels of gene expression and external regulatory inputs from H-NS. Arrows indicate positive regulation while lines with a terminal, perpendicular line indicate negative regulation. Grey arrows indicate the positive feedback input of VirB. Dotted lines represent protein production by translation. Grey box indicates promoters that are regulated by both VirB and MxiE.

Figure 3. Events required for the MxiE-dependent activation of tier 4 promoters. (a) Under conditions in which the type three secretion system (TTSS) is inactive (non-secretion), IpgC serves as a chaperone to associate with the TTSS translocators, IpaB and IpaC, while the MxiE anti-activator, OspD1, is bound to both MxiE and its chaperone, Spa15. In this case, both the activator (MxiE) and co-regulator (IpgC) are sequestered and unable to bind tier 4 promoters to activate transcription. (b) Under conditions in which the TTSS is active (secretion), each protein complex dissociates resulting in secretion of IpaB, IpaC and OspD1. Meanwhile, liberated IpgC and MxiE accumulate within the bacterial cytoplasm, presumably associate and subsequently bind to tier 4 promoters to activate transcription. Open box represents MxiE binding site. Shaded box represents MxiE-dependent gene (tier 4). Line with arrow represents the transcription start site.

Figure 4. Chromosomally encoded regulators that influence expression of *Shigella* virulence genes on pINV. Direct and indirect regulatory inputs are illustrated. The environmental cue triggering the activity of each regulator is given directly below each regulator and the specific gene(s) whose expression is modulated is depicted. Straight lines indicate direct regulation while dashed lines indicate indirect regulation. Arrows indicate positive regulation while lines with a terminal, perpendicular line indicate negative regulation. Question marks indicate unidentified components that may regulate the gene of interest. OmpR/EnvZ has been excluded since it is unclear where the virulence regulatory input is received.

Figure 5. Temperature-dependent events that upregulate the first three tiers of the pINV regulatory cascade. A temperature switch to 37°C upregulates tier 1-3 promoters. At the *virF* promoter (tier 1), a thermally induced change in local DNA structure alleviates H-NS-mediated repression. At the *virB* promoter (tier 2), DNA gyrase increases negative supercoiling of the promoter region, which promotes maximal VirF-dependent transcription activation. At tier 3 promoters, VirB, which is maximally expressed at 37°C, alleviates H-NS-mediated repression through a mechanisms predicted to involve DNA wrapping and VirB oligomerization (Gao et al., 2013). Note: the positive regulatory feedback loops

mediated by VirB at 37°C (Kane and Dorman, 2012) are not shown for simplicity sake. Additional regulatory inputs occur at each tier depicted (see text), but these have been omitted because they are not directly modulated by temperature. White circles represent H-NS monomers. Black hexagon represents VirF. Grey pentagon represents VirB. Line with arrow represents the transcription start site. "-" symbol represents negative supercoils.