The *Salmonella enterica* Regulatory System SsrA-SsrB

Characterization of the *Salmonella enterica* Two-Component Regulatory System SsrA-SsrB and the SsrB Regulon

By

David T Mulder, BSc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree

Doctor of Philosophy

McMaster University Doctor of Philosophy (2014), Hamilton, Ontario, Canada

TITLE: Characterization of the *Salmonella enterica* Two-Component Regulatory System SsrA-SsrB and the SsrB Regulon

AUTHOR: David T Mulder, BSc.

SUPERVISOR: Dr. Brian K Coombes, PhD

NUMBER OF PAGES: 267

Abstract

Salmonella enterica is an intracellular bacterial pathogen of humans and the causative agent of the acute gastrointestinal disease, salmonellosis, and the chronic systemic infection, typhoid fever. Sensor proteins convert environmental signals, including signals detected within the host environment, into biochemical signals to control cellular responses. It has been previously established that the two component regulatory system SsrA-SsrB, consisting of the integral membrane sensor kinase protein SsrA and the cytoplasmic DNA-binding response regulator SsrB are essential for regulation of bacterial factors during systemic intracellular infection. The first chapter of this thesis describes characterization of the sensor kinase SsrA. The structure of the periplasmic sensor domain is modeled and evidence is presented that it is involved in enhancing signaling activity in response to environmental acidification encountered within the intracellular environment. A mechanism whereby protonation of histidine residues within this region in response to acidification drives conformational strain and thereby signaling is proposed. The second chapter describes identification of the DNA-binding motif of SsrB within regulated promoters as well as its regulon. Integration of experimental data with comparative genomics data resulted in identification of the palindromic heptameric DNA recognition motif of SsrB as well as identification of novel SsrB-regulated promoters. In addition, a DNA microarray analysis is described wherein the complete SsrB regulon is identified. Finally, the third chapter describes regulatory input of SsrB to the *S. enterica* type VI secretion system. This chapter also describes the contribution of this system to

systemic dissemination of *S. enterica* during host infection. Altogether, these data advance understanding of how *Salmonella* controls factors essential for disease in response to the host environment during infection.

Acknowledgements

This work is the culmination of twelve years of post-secondary education and is made possible by the amazing learning environments within the labs of Fiona Brinkman at Simon Fraser University and Brian Coombes at McMaster University. I am indebted to them for their support over the years and the opportunities that they have made available to me.

I am eternally greatful to my friends and coworkers in the Coombes lab who showed me the ropes and made my time at McMaster educational and fun. Thank you Ana Tomljenovic-Berube, Colin Cooper, Suzanne Osborne, Sarah Reid-Yu, Brian Tuinema, Sarah Allison and Bushra Ilyas. I also thank Ana Pilar and Joe McPhee, who showed great patience in teaching methods, discussing science and showing kindness in tolerating my constant questions.

Finally, I wish to thank departmental staff who make everything possible behind the scenes, particularly Lisa Kush. Lastly, I thank my family and friends for their unwavering love and support during the stresses of graduate school and concurrent life challenges.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Acronyms

Amp	Ampicilin
ASP	Acid Shock Protein(s)
ATR	Acid Tolerence Response system
B1H	Bacterial One-Hybrid
BLAST	Basic Local Alignment Search Tool
BMAA	*Burkholderia mallei* ATCC 23344
Cam/CM	Chloramphenicol
CDS	Coding Sequence
CFU	Colony Forming Units
ChIP	Chromatin ImmunoPrecipitation
CI	Competitive Infection
COG	Cluster of Orthologous Groups
DHp	Dimerization and Histidine phosphorylation
EMSA	ElectroMobility Shift Assay
FACS	Fluorescence Activated Cell Sorting
FLAG	FLAG octapeptide protein tag
GALT	Gut-Associated Lymphoid Tissue
Gent	Gentamycin
HA	Hemagglutinin epitope
Hcp	Haemolysin co-regulated protein
HeLa	Human epithelial carcinoma cell line
HEp-2	Human epithelial carcinoma cell line
HisKA	Histidine Kinase
IGR	Intergenic Region
iNTS	Invasive Non-Typhoidal *Salmonella* disease
Kan/KN	Kanamycin
LB	Lysogeny Broth / Luria Broth
LPM	Low-Phosphate Low-Magnesium Medium
LPS	Lipopolysaccharide
M-cell	Microfold-cell
NADPH	Nictoinamide Adenine Dinucleotide Phosphate
NAP	Nucleoid-Associated DNA-binding Protein
NLR	NOD-Like Receptor
NMR	Nuclear Magnetic Resonance Spectroscopy of Proteins
NM	minimal glucose medium lacking histidine
NOS2	Nitric Oxide Synthase 2
NRAMP1	Natural Resistance-Associated Macrophage Protein 1
OD	Optical Density
PAMP	Pathogen Associated Molecular Pattern(s)
PAS	Per-Arnt-Sim
PDC	PhoQ-DcuS-CitA
PWM	Position Weight Matrix

RAW 264.7	Mouse leukaemic monocyte macrophage cell line
Rcf	Relative centrifugal force
Rhs	Recombination HotSpot
RLU	Relative Light Unit
ROD	Regions of Difference
RPM	Revolutions Per Minute
SCI	*Salmonella enterica* Centisome 7 Genomic Island
SCV	*Salmonella*-Containing Vacuole
SM	Streptomycin
SPI	*Salmonella* Pathogenicity Islands
SSR	*Sodalis* Symbiosis Region
STM	*Salmonella enterica* serovar Typhimurium
T3SS	Type III Secretion System
T4SS	Type IV Secretion System
T6SS	Type VI Secretion System
TLR	Toll-Like Receptor
TM/TMR	Transmembrane Region
Tss3	T6SS-3 of *Burkholderia* spp.
VrgS	Valine-glycine repeat protein family member
WT	Wild Type

Chapter I – Introduction

Chapter I – Introduction

Salmonella enterica is a foodborne intracellular pathogen

Salmonella enterica is a human-adapted pathogen disseminated by the oral-fecal route of transmission that causes outbreaks of disease associated with severe morbidity and mortality across the world. In the western world, *S. enterica* serovars Enteritidis and Typhimurium, among other less prevalent serovars, are associated with acute gastroenteritis known as salmonellosis, typified by gastrointestinal inflammation and diarrhea and generally occur as point outbreaks following contamination of food products (1). Infection with *S. enterica* serovars Typhi and Paratyphi cause the systemic disease typhoid but such outbreaks have been eliminated in the developed world following improved water treatment infrastructure. This disease is typified by onset of a fever of increasing severity over the course of multiple weeks with infrequent shedding of the bacterium followed by severe complications in some patients due to intestinal perforation, or high burdens of bacteria causing intestinal hemmoraging at Peyer's patches and enlargement of the spleen and liver (2, 3). Outcomes of acute disease include elimination of the infection, long-term asymptomatic carriage or death (4). Systemic typhoid remains a major concern in developing nations and in the developed world pursuant to infrastructure-disrupting disasters.

Interest in *Salmonella* pathogenesis has surged following a spate of foodborne outbreaks in North America and the emergence of the highly virulent serovar DT104, which is difficult to treat due to its resistance to ampicillin, chloramphenicol,

streptomycin, sulfonamides and tetracycline (5). Recently, invasive non-typhoidal *Salmonella* disease (iNTS) has emrged in immunocompromized patients in Africa (6, 5). While risk factors for iNTS are primarly associated with poor nutrition and co-infections such as malaria and HIV, the persistence of these outbreaks plus their poor control and ineffective treatment due to insufficient funding and infrastructure continues to provide a selective environment for rapid pathogen evolution and to act as a reservoir for spread to other nations.

In addition to human pathogens, the *Salmonella* genus includes other pathogens and commensals with varying host tropism and differential pathogenicity that facilitate stable reservoirs of infection within and between organisms including plants (7, 8). The wide host tropism of *Salmonella* serovars permits the use of animal models including chickens, cattle and mice to study aspects of bacterial pathogenesis and mechanisms of infection (9). In particular, these *in vivo* models have focused on the roles of horizontally acquired virulence genes that have enabled *Salmonella* to survive and replicate within the intracellular host environment.

Gene Acquisition at Salmonella Pathogenicity Islands

The host environment presents a major hurdle to intracellular persistence and different pathogens employ unique strategies to manipulate this environment and resist host responses during infection (10, 11). Gene acquisition has provided the tools that facilitated progression from a putative ancestral extracellular bacterium to an intracellular pathogenic lifestyle (12). These acquisitions have enabled host cell control and resistance

to the challenging host extracellular and intracellular environments that include hydrolytic activity and acidic pH of the stomach; activity of hydrolases, bile salts and antimicrobial peptides in the gastrointestinal tract; and antimicrobial activity of the innate immune system (13). Therefore, the genetic basis for virulence amongst the *Salmonellae* is predominantly due to a subset of acquired genes that have been integrated by regulatory evolution to overcome these challenges (14, 15).

The genetic repertoire of a bacterium can be considered to be those genes necessary for basic survival and replication within a permissive environment (essential genes), plus a second set of genes, non-essential for growth *in vitro* under permissive conditions (metabolic genes, antibiotic resistance genes, adhesion genes and invasion genes), that expand access of the organism to additional niches (16). This latter set includes acquired genes that are accumulated over time under selection by way of nucleic acid transduction, bacteriophage transduction and bacterial conjugation. These integration processes often target specific regions of the genome such as tRNA loci as a result of homologous recombination leading to clusters of genes with a distinct signature known as pathogenicity islands, or *Salmonella* Pathogenicity Islands (SPI) in *Salmonella* (12, 17). These regions are enriched in virulence genes (18) which contribute to or are essential for disease in animal or cell culture models of infection (19–21).

To date, over 17 SPI loci have been named as a result of having a quantifiable contribution to pathogenesis (22). These regions are not found in all serovars of *S. enterica* and their presence or absence has been used to construct the evolutionary history of the divergence of the *Salmonella* lineage from the common ancestor shared with *E.*

coli (23). Oldest acquisition events are well-conserved within the lineage and the most recent acquisitions cluster in closely related serovars (24). The first acquisition event which may have demarcated the start of the *Salmonella* lineage by opening up the intracellular niche of the mammalian small intestine for exploitation was that of SPI-1 (25). A second acquisition event of SPI-2 led to the branching of the *Salmonella* lineage, and only the *Salmonella bongori* species, commensal of reptiles, lacks this second locus (26). A recent acquisition was that of SPI-6 within the *S. enterica* subspecies *enterica*, which has been acquired at multiple times in different locations in different subspecies (27, 28). SPI-1, SPI-2 and SPI-6 encode major virulence systems and other SPI represent smaller genomic islands that encode accessory proteins for these and other virulence pathways (22).

Not all laterally acquired genetic information is useful to the bacterium. Genomes represent a snapshot of an organism in time and mobile genetic elements containing mobilized DNA are constantly in flux at genomic tRNA hot spots due to the preference of bacteriophages for low-GC content and repetitive sequences (27, 29). The specificity of the global nucleoid-associated protein H-NS for these AT-rich genomic islands results in regulatory silencing through oligomeric protein-DNA interactions of H-NS which restricts access to these regions by transcription factors and RNA polymerase (30). Dissocation of these silencing complexes is observed at host body temperatures allowing expression of these regions during infection on which natural selection can act (31). As a result, useful genes are likely maintained by positive selection and deleterious genes are likely lost by negative selection. As evidenced by the differential characteristics of young

and old acquired SPI loci, the surrounding co-acquired genetic information is progressively lost due to random recombination events leaving either small or large densely packed virulence loci (32). Over time, this genetic information is likely integrated in to the global *Salmonella* virulence programme by cis-regulatory evolution resulting in contribution of these acquired genes to aspects of host infection (33).

Salmonella Infection of the Host

During host infection, *S. enterica* targets to the crypts of the small intestine where it uses flagella and adhesins to transit the mucosal layer and adhere to enterocytes (34). Deeper penetration of the crypts is prevented by the activity of Paneth cells which secrete antimicrobial factors including lysozyme and cationic peptide defensins that challenge pathogen membrane integrity (35, 36). Following adherence, bacteria use type 3 protein secretion systems to translocate proteins in to the host cell cytosol. These translocated proteins induce actin cytoskeleton remodeling ultimately leading to uptake of the bacteria in to vacuolar compartments known as the *Salmonella* Containing Vacuole (SCV) (37, 38). Recognition of bacteria Pathogen Associated Molecular Patterns (PAMP) such as flagellin by intracellular NOD-like host cell receptors, and LPS and flagellin by toll-like receptors TLR4 and TLR5 (39), leads to pro-inflammatory cell death of infected cells and immune cell infiltration which compromise the gut barrier leading to acute diarrheal episodes and which is resolved in immunocompetent individuals following immune response (1).

Typhoidal serovars of *S. enterica* similarly target to the crypts of the small intestine during the systemic infection typhoid. Acute gastroenteritis may occur following localized enterocyte infection; however, infection of the microfold (M-) cells within the Peyer's patches permits access to underlying immune tissue (40). M cells are specialized epithelial cells that sample and transfer luminal antigens by transcytosis to basolateral-associated lymphoid tissue enriched in antigen presenting cells. Here, *S. enterica* is taken up by macrophages via macropinocytosis into what becomes the SCV (41). *Salmonella* is also taken up by dendritic cells directly from the luminal environment, and by neutrophils at the sites of infection; however, these cell types are less permissive of survival and dissemination than macrophages (42). Following phagocytic uptake, some macrophages remain within the Gut Associated Lymphoid Tissue (GALT) Peyer's patches, while others migrate through the lympathic system via mesenteric lymph nodes where *Salmonella* persists (43). Further dissemination of infected phagocytes through the lympathic system leads to productive foci of infection in the spleen and liver; however, infected phagocytes can also disseminate through the blood stream to systemic sites (44). Interestingly, *S. enterica* alters the migration of host phagocytes, which facilitates persistence (4, 44). End stages of infection result in caspase-1 dependent pro-inflammatory death of immune cells, termed pyroptosis, following intracellular NOD-like receptor (NLR) recognition of intracellular PAMPs (45).

The Host Intracellular Environment

Upon internalization of *S. enterica,* a race between the bacterium and the phagocytic host cell begins. The outcome is either establishment of an antimicrobial phagosome leading to bacterial lysis and death, or interruption of this process and establishment of a vacuolar environment permissive of bacterial survival and replication (37, 46). The vacuolar environment is initially composed of extracellular fluid surrounded by a membrane derived from the host cell plasma membrane. The vacuole then begins a rapid process of maturation by fusion with early endosomes, late endosomes then lysosomes to ultimately form the phagolysosome (47). These fusion events deliver host proteins carried in the membrane of endosomes or the contents of the endosome itself. The phagosome functions primarily in recycling and degredation of soluble and membrane components through fusion with endosomes and budding of vesicles for recovery or targeting for destruction. Formation of the phagolysosome occurs pursuant to lysosomal fusion at which point the vacuole is capable of bactericidal activity (47).

One of the first measurable changes during vacuolar maturation is acidification as a result of vacuolar ATPase activity. The pH of the vacuolar environment changes from 7.4 to 4.5 over time, decreasing with each successive fusion step as more of the ATPase is recruited to the vacuolar membrane (47, 48). Acidification of the vacuole results in a bacteriostatic environment, but functions primarily to generate conditions that activate hydrolytic pro-enzymes, to generate a proton gradient that favors production of electronegative reactive oxygen and nitrogen species by other proteins through charge balancing and generation of H^+, and to counter-transport ions that are permissive for bacterial growth (47). Interestingly, it has been shown that Toll-like receptors (TLR), a

class of PAMP receptor signaling is essential for full acidification of SCV and that this acidification is necessary for *S. enterica* virulence (49). While these highly acidic conditions discourage growth of most bacteria, *S. enterica* can grow well under acidic conditions *in vitro* and persists in acidic environments *in vivo* as a result of its acid tolerance response (ATR) system (13, 50). In *S. enterica* this system is an inducible adaptive response, and acid exposure in the stomach may prime the bacterium for survival within the intracellular environment (51). Indeed, exposure to an acidic environment is necessary for full virulence potential of *S. enterica* in mice (52). Adaptation to acid exposure involves activation of global regulators including PhoQ-PhoP, OmpR, the stress-associated alternative sigma factor RpoS and other specific regulators (13). Activation of these regulators drive expression of inducible lysine and arginine decarboxylase systems that maintain neutral intracellular pH, activation of Acid Shock Proteins (ASP) that function in acid stress repair, and modification of membrane fatty acid composition in order to reduce oxidative damage.

While acidity is increasing during phagosomal maturation, fusions with late endosomes recruit NRAMP1 which actively transports divalent cations such as Fe^{2+}, Zn^{2+} and Mn^{2+} out of the phagosome (47). Upon iron limitation, *S. enterica* secretes a high affinity iron siderophore, enterobactin, the uptake of which is mediated by its high affinity transporters IroN and FepA (53). NRAMP1 activity leads to upregulation of lipocalin-2 by the host cell which blocks iron scavenging by the bacterium through binding of enterobactin (54). Establishment of an environment depleted of divalent cations is important for enhancing the activity of positively charged defensin peptides that

are recruited through fusion of the phagosome with lysosomes and specialized granules in some phagocytes. These peptides permeabilize the bacterial membranes through interaction with the negatively charged membrane components, establishing pores leading to lysis (55). The activity of these peptides is further enhanced by a collection of soluble hydrolases that target surface exposed lipids, carbohydrates and proteins on the bacterial cell surface. *S. enterica* resists the activity of antimicrobial peptides through acylation of lipid A, part of its lipopolysaccharide (LPS) in order to mask the negative charge and discourage defensin interaction at the membrane (56). Modification of lipid A is primarily regulated by the two component systems PhoQ-PhoP, activated directly by defensin activity, and PmrB-PmrA, activated by the low iron conditions of the phagosome generated by NRAMP1 (57).

While most bacteria are eliminated by defensin and hydrolase activity in the low pH bacteriostatic environment, professional intracellular pathogens require more drastic host response measures. To this end the NOX2 NADPH oxidase and the NOS2 inducible nitric oxide synthase in phagocytes generate a bactericidal environment (47). These proteins generate reactive oxygen and nitrogen species which directly attack bacterial components leading to oxidative damage and lysis of the bacterium (58). *S. enterica* resists reactive species using the enzymes catalyase and superoxide dismutase, as well as small proteins and small molecules including glutaredoxin, thioredoxin and glutathione in a process regulated by the sensors OxyR and SoxRS (59). Interestingly, it has been shown that reactive nitrogen species can directly activate some intracellular regulators,

suggesting that oxidative damage may not only be a consequence, but a cue detected at multiple levels by the bacterium for virulence gene regulation (60).

Countering of antibacterial strategies employed by the host require the bacterium to sense and respond to specific signals in this host environment (47). As an intracellular pathogen, *S. enterica* uses these cues for the regulation of virulence. Loss of some of these cues paradoxically reduces the ability of *S. enterica* to survive and proliferate in cells (49). Virulence factor expression for survival within the intracellular environment is therefore dependent on the unique signature of the intracellular environment detected by two-component regulatory systems.

Two-Component Regulatory System Sensor Kinases

The ability of *S. enterica* to occupy many niches can be attributed to the extensive set of acquired genes within SPI loci and elsewhere in the genome. However, this genetic information must be appropriately expressed in response to environmental and physiological cues within these niches (14). Spatiotemporal gene expression is made possible by a hierarchy of regulatory proteins which exhibit exquisite control of gene expression, particularly as *S. enterica* transitions from a cell-associated to intracellular environment to avoid host immune recognition, modulate motility, modify surface characteristics, coordinate adhesion and deliver effector proteins (11, 34, 61–65).

Perception of disparate environmental signals requires translation into a common biochemical signal that can control processes in the cell including DNA replication, RNA transcription or protein translation. In bacteria, a significant proportion of signal

transduction relies on two-component regulatory systems, in which a sensor protein converts conformational changes from signal detection into a phosphorylation cascade that activates a response regulator (66, 67).

Two-component signal transduction systems are present in all domains of life and are widespread amongst bacteria, with some having hundreds of such systems (68). These systems consist of sensor proteins that detect diverse types of signals including ligands, light and osmolarity in the cytoplasm, membrane or periplasm by transduction of conformational changes in variable N-terminal sensing domains into phosphorylation of conserved C-terminal domains (69). Molecular recognition between the sensor kinase and response protein and phosphatase activity of the former for the latter ensures that signaling specifity is high with little cross-talk amongst potentially hundreds of active TCRS in a cell (70, 71). Alone, these systems provide binary control of downstream processes but can be combined to fine-tune control at the single gene, global regulon, whole-cell or multicellular levels (72).

Sensor domains detect different types of stimulus including specific ligand concentrations and non-ligand stressors such as osmotic stress or pH (69). Ligand-based activation of sensor domains is much better understood, however all mechanisms of sensing ultimately induce conformational change of the sensor domain in order to transduce a signal across the membrane (73). Sensor kinases with periplasmic sensing domains are integral membrane proteins and the most common class (73). These sensing domains can be either all α-helical or mixed α/β structures and commonly sense extracellular ligands. Sensor kinases with cytoplasmic sensing domains are either soluble

or integral membrane proteins and sense intracellular conditions such as oxygen concentrations, redox state or light and can be grouped in to three different structural families (73). Sensor kinases with integral membrane sensing domains such as the osmotic stress sensor EnvZ are the least common and are typified by multiple trans-membrane α-helices with short loop regions between these domains that sense by way of an unknown mechanism (73).

The most common sensing domain organization is the PhoQ-DcuS-CitA (PDC) domain of periplasmic sensor kinases. This PDC fold structural organization is a sub-class of the Per-Arnt-Sim (PAS) domain of cytoplasmic sensors (73). This domain consists of a long N-terminal α-helix (cap), followed by a shorter α-helix, and then an anti-parallel five stranded β-sheet and terminated with an α-helix. Modified PDC-domain sensors are common, in which a second PDC domain has been inserted within the α-helical cap to form a double-PDC domain, which has a lengthened α-helical cap on which the membrane-distal and membrane-proximal β-sheet scaffold are hung (73). This structural organization loosely resembles the number "18" extending from the membrane, where the number "1" is the α-helical cap and the number "8" represents the two β-sheet scaffolds. The periplasmic PDC domain region is flanked by trans-membrane α-helices, which form a coiled-coil within the membrane. Typically sensor kinases function as dimers, and the two transmembrane regions combine to form a dimeric four-helical trans-membrane bundle (73).

In the case of ligand detection, the β-sheet scaffold of the membrane distal domain forms the ligand interaction site. This pocket is conserved in chemical character and key

residues, while ligand specificity is determined by major and minor ligand binding loops that drape over the surface of this site (74). Ligand interaction is hypothesized to drive conformational changes in the α-helical regions that are propagated through the membrane-proximal PDC domain and across the membrane by way of the trans-membrane bundle to the cytoplasmic domains (73). Interestingly, the function of the membrane-proximal domain is unclear in many sensors. At a minimum, it may simply increase sensitivity of the membrane distal domain by extending it away from the membrane and other membrane-associated components into the periplasm, or in the case of the sensor kinase PhoQ, it may provide an additional sensor platform. Recent work has shown that sensor kinases detect more than a single cue and PhoQ has evolved sensitivity to multiple signals including pH, phospholipids, ion concentrations and antimicrobial peptides (75–77). Antimicrobial peptides are the primary signal *in vivo*, and the mechanism of sensing relies on the disruption of a divalent cation bridge between negatively charged phospholipids and the membrane-proximal domain. In any case, for all PDC-fold sensors, the anchored nature of the N-terminal α-helical cap likely enables the generation of strain which permits phosphorylation of the downstream cytoplasmic kinase domain (78).

The cytoplasmic domains of sensor kinases are well conserved and consist of a HAMP (present in Histidine kinases, Adenylate cyclases, Methyl accepting proteins and Phosphatases) linker domain and a transmitter domain, which itself consists of a dimerization and histidine phosphorylation (DHp or HisKA) domain and a catalytic ATPase domain (79). Propagated strain from the sensor domain through the

transmembrane and linker domains leads to exposure of the histidine residue normally buried in the DHp domain to which the gamma-phosphate of ATP is transferred from the ATPase domain of the opposing monomer (69). This phosphoryl group is then transferred to an aspartyl residue of a downstream receiver domain. This receiver domain can be part of the sensor kinase itself in the case of a hybrid histidine kinase, on a separate intermediate signaling protein or on a terminal response regulatory protein (80). In the first two cases, the receiver domain is fused with an additional phosphotransfer domain for terminal transfer of the phosphoryl group to a downstream response regulatory protein.

Two-Component Regulatory System Response Regulators

Response regulators control aspects of nearly all cellular processes of bacteria (72). They typically consist of a receiver domain that act as a phosphorylation-dependent switch to control activity of a downstream response domain (81). The most common response domain is a DNA-binding domain for activation or inhibition of gene transcription (72, 80). As a result, environmental signals can be directly linked to fundamental control of the cell through a sensor and a regulator, and with the two acting together as an amplifier.

Receiver domains all feature an aspartate phosphorylation acceptor site and have a well-conserved structure consisting of an alternating $(\beta\text{-}\alpha)_5$ organization forming a five-stranded parallel β-sheet sandwiched between two and three α-helices (82). Three aspartyl residues in the core of the β-sheet are necessary for interaction with a divalent cation, necessary for catalyzing phosphorylation and dephosphorylation, and one has an

additional function of phosphoryl acceptor (80). Phosphorylation of the aspartyl group leads to conformational change in the receiver domain. The phosphoryl group oxygens interact with specific conserved elements of the response domain in order to stabilize an alternative configuration that leads to restructuring of one side of the protein, the α4β5α5 face (83). Changes in conformation at this site result in rearrangements of the response domain between active and inactive states. In the case of transcription factors, the DNA interacting helix-turn-helix motif is made accessible to DNA by phosphorylation-mediated stabilization of the active conformation (81).

Receiver domain phosphorylation is generally controlled by transfer of phosphoryl groups from a sensor kinase and lost by way of autodephosphorylation. However phosphorylation can also occur as a result of self-catalyzed autophosphorylation from phospodonors in the cell (84–86). An additional level of response regulator activity is through control of dephosphorylation. Response regulators are capable of autodephosphorylation but sensor kinases and stand-alone proteins can both function to control the rate of this reaction with phosphatase activity (87–89). Together this balance of receiver domain phosphorylation and dephosphorylation permits tight control of response domain activity. Phosphatase activity of the sensor kinase in the absence of signal-induced activation ensures that response regulators remain inactive, despite any coincidental phosphorylation from alternative donors such as small molecule phosphodonors (70).

Response domains are diverse in structure and function, and share in common only conformational regulation by an upstream receiver domain. Classes are observed that

regulate cell activity at the transcriptional, translational and post-translational levels through interactions with DNA, RNA and proteins while others have been described to have transporter or enzymatic functions (72). Response domains can be confidently assigned to families based on motifs or domain organization, however precise prediction of function is difficult as strong sequence conservation is not predictive of functional conservation, especially for transcription factors where minute differences in response domain structure determine DNA interaction specificity (72).

The most common class of response regulators are DNA-binding transcription factors that function as dimers and which have helix-turn-helix motifs in their response domains (72, 80, 81). These motifs are characterized by a tri-helical bundle in which the second and third helices make contacts with the DNA major groove (90). Upon dimerization of the α4β5α5 face, or more rarely within the response domain itself, the second helices from each bundle end up separated by approximately 34 angstroms to permit interaction with adjacent major grooves. The third helices provide stabilization through general DNA contacts but do not confer specificity, which is determined by the second helix (90). The central region between the two monomers is formed by the rest of the response domain and generally does not contribute to DNA specificity. As a result of this configuration, DNA sequences recognized by such transcription factors are palindromic with centrally variable regions of two to four base pairs. Alternative head-to-tail dimers that recognize tandem DNA repeats have also been described and spacing can be variable due to differences in the location of the dimerization interface (81). In

addition, some response regulators dimerize after recruitment of the individual monomers to the DNA (81).

The target for these activated response regulators are gene promoters, where they can assist or impede in the recruitment of RNA polymerase depending on the relative position of the binding site within the promoter with respect to the sigma factor interaction sites (91). While the structure of DNA-binding response regulators is informative as to their mode of regulation and functional conformation, it is not predictive of the DNA sequence recognized. To identify this sequence, the protein-DNA interactions must be captured, isolated, sequenced and assembled using bioinformatics methods (92).

Integration of Regulatory Signals by the TCRS SsrA-SsrB

In *Salmonella*, two-component regulatory systems consisting of membrane-anchored sensor kinases and their cognate DNA-binding response regulators are major players in the control of virulence (62, 93). Although distinct combinations of regulatory proteins are activated within the lumen of the gut and within the intracellular SCV, both responses involve the two component master regulators of virulence PhoQ-PhoP and EnvZ-OmpR (94). Combinatorial integration of signaling from these and other regulators in the lumen of the gastrointestinal tract leads to activation of the Hil regulatory proteins while signaling in the intracellular SCV ultimately leads to expression of the SsrA-SsrB two component system (62, 93). These components are the primary regulators of the type III secretion systems (T3SS) that translocate proteins from the bacterial cytoplasm across host membrane barriers. An extensive number of other regulators contribute to systemic

dissemination and survival within the host but their regulatory interactions are downstream or indirect (63).

While most data suggest that the SsrA-SsrB system is active only within the intracellular macrophage environment *in vivo*, there is evidence that it is expressed in advance of cellular invasion, possibly to prime the bacteria for rapid response to this environment. This priming may rely on de-repression of higher level mechanisms of transcriptional silencing (95, 96). Nucleoid-associated DNA-binding proteins (NAP) control transcription from these promoters through negative regulation of AT-rich regions such as the T3SS loci (31, 97, 98). This silencing is locally countered by the previously described transcriptional activators (99, 100). Global cues such as oxygen gradients in the gut or temperature differences between the host and external environments may also relieve silencing at these promoters by modulating NAP activity (101, 102).

When considering the regulation of intracellular virulence, integration of upstream regulatory signals from PhoQ and EnvZ by OmpR (activated by EnvZ), and SlyA (activated by PhoQ-PhoP) occur at the *ssrA* promoter (103, 104). A promoter upstream of *ssrB* is positively regulated by PhoP (activated by PhoQ), and OmpR (activated by EnvZ) and negatively regulated by the iron regulators PmrA (activated by PmrB) and Fur (105–108). Regulation is also integrated at the post-transcriptional level by an interaction between the RNA transcript of *ssrA* and PhoP (106). Interestingly, there is evidence that SsrB participates in a feedback loop by interacting at the *ssrA* and *ssrB* promoters (109). It is not unusual for two-component regulatory systems to be autoregulated however

separation of the two-gene operon into transcriptional units by a cryptic promoter suggests the need for finely tuned expression of SsrB.

SsrA is a hybrid histidine kinase, putatively localized to the bacterial inner membrane with a single periplasmic sensor domain, consistent with other PDC domain type sensor kinases. It was first identified as a two-component regulatory protein by alignments with other TCRS and was found to be necessary for T3SS-dependent protein secretion (110). It was later confirmed to be necessary for gene expression from SPI-2 (111, 112) and from other SPI loci in the genome (113). SsrA was activated within the intracellular macrophage environment and *in vitro* through magnesium limitation and phosphate deprivation (114). This activity could be repressed by alkaline pH *in vitro* (115) and activated by acidic pH in a phosphate limiting environment (61). Finally, a recent report has shown that acidification of the vacuolar endosome is dependent on innate immune sensing of the bacterium by the host and that this acidification is necessary for SPI-2 T3SS activation (49). Together, these data suggest that SsrA is sensitive to low pH and low ion concentrations in the extracellular environment, however structural data is absent and the mechanism of sensing is unknown. There has been no reported investigation of the regulatory implications of the additional receiver and phosphotransfer domains (hybrid histidine kinase).

SsrB is a two-domain, helix-turn-helix response regulator belonging to the NarL/FixJ family, the second most common DNA-binding response regulator family in bacteria (72, 116). The structure of the C-terminal response domain has been solved by NMR spectroscopy and comprises a helix-turn-helix motif that is active in the absence of

the N-terminal receiver domain (109, 116). Initially thought to regulate genes only within the T3SS locus in which it is encoded, SsrB was later found to regulate virulence genes encoded within other loci (117). Further regulon characterization solidified the role of SsrB as a global regulatory protein (118). Efforts then focused on identifying the specific sequence of DNA recognized by the SsrB response domain using DNase I footprinting, however no consensus sequence could be identified due to inconsistentcy in footprints (100, 109). The lack of a specific sequence was interpreted to suggest that SsrB functioned to relieve silencing at these promoters by displacing H-NS rather than acting as a traditional transcription factor to recruit RNA polymerase despite its helix-turn-helix structure and familial relationship with other tranditional transcription factors. Therefore this sequence, if any, as well as biochemical analysis of the oligomeric state, remained to be determined.

Protein Translocation Systems Modulate Host Cells

SsrA-SsrB directs expression of the T3SS secretome that forms a multi-protein trans-membrane secretion system. This system confers the ability to manipulate the host environment in order to facilitate replication and survival of *S. enterica* (34). Bacterial pathogens that translocate proteins across the host cell membrane use type 3, 4 and 6 secretion systems (11). *S. enterica* has two T3SS and one T6SS (15, 28). While the *S. enterica* T6SS is not well understood, the two T3SS promote pathogen survival through manipulation of the host cell cytoskeleton, vesicular trafficking, cell-death pathways, and host immune processes (119).

T3SS are hypothesized to have been derived from the flagellar secretion system through the addition of a translocon complex that repurposes this system from secretion of flagellar components to secretion and translocon of protein effectors (120, 121). The T3SS of *S. enterica* are encoded within SPI-1 and SPI-2 and are used to translocate proteins across the host cell membrane and across the intracellular vacuolar membrane, respectively (34). These systems are composed of a basal body in the bacterial membranes, and a translocon pore complex in the host membrane to which the T3SS is connected by an extracellular needle (122). While most apparatus proteins are well characterized in the *Salmonella*, *Escherichia* and *Yersinia* systems, regulation of expression and assembly is less well understood. In the case of the SPI-2 T3SS there are multiple operons of apparatus genes loosely organized by function, but aside from a dependency on SsrA-SsrB in directing their expression, how these operons are temporally transcribed in order to assemble the system is not well understood. In other systems a temporal progression of expression from basal apparatus onwards occurs (123). Similarly, work has only recently begun to understand how the system transitions from apparatus component secretion for assembly to effector protein secretion for translocation, with pH appearing to play an important role in this switch (124).

Following adhesion, the SPI-1 T3SS is essential for invasion of non-phagocytic intestinal epithelial cells through manipulation of the host actin cytoskeleton (125). To induce bacterial mediated internalization, the T3SS delivers effector proteins into the host cell, which act as guanine-nucleotide-exchange factors for host Rho-like G proteins. As a result, localized polymerization of actin filaments leads to ruffling of the host cytoplasmic

membrane and engulfment of the bacterium (38). Numerous other protein effectors are involved in targeting other host G proteins and GTPases leading to increased intestinal permeability, and drive stimulation of pro-inflammatory cytokine release, leading to diarrheal symptoms in the host (126–128). To further increase bacterial numbers, an additional effector inhibits apoptosis of infected cells by inhibition of immune pathway signaling (129). Together, the role of this system and these effectors is to permit invasion of host cells and to drive a pro-inflammatory response that facilitates shedding and transmission of the bacteria through the fecal-oral route.

The SPI-2 T3SS is essential for systemic dissemination during infection by promoting survival and replication within the intracellular environment (34, 130, 131). Secretion of effector proteins by the internalized bacterium diverts the SCV away from normal phagolysosomal maturation (37). While the SPI-2 T3SS also has important roles in avoiding the reactive oxygen and nitrogen species and acquisition of lyososomal hydrolases through lysosome fusion, the effectors involved are unknown (131). However, effectors and mechanisms involved in manipulation of vacuolar membrane, vacuolar trafficking, host cytoskeleton rearrangement and immune signaling are characterized.

First, four effector proteins (SifA, PipB2, SseJ, SopD2) interact with the vacuolar membrane and host proteins to manipulate vacuolar membrane dynamics of the SCV (131). SifA stabilizes the growing vacuolar membrane by inhibiting accumulation of the host molecular motor protein kinesin (132). In the absence of SifA, premature rupture of the SCV occurs leading to effective clearance of the bacteria from the host cell (133, 134). In contrast with the function of SifA, PipB2 is involved in recruitment of kinesin to the

SCV suggesting a temporal control of membrane dynamics (131). The roles of SseJ and SopD2 are not well understood, despite also being involved in establishment and maintenance of the SCV membrane. SseJ appears to be involved in establishing rigidity of the membrane via removal of cholesterol to lipid droplets (131). While the function of SopD2 is unknown, its loss prevents establishment of membrane tubules known as *Salmonella* Induced Filaments (SIFs) that extend radially out from the SCV along host microtubules (135, 136).

Trafficking of the SCV in the host cell is further modified by two additional effector proteins (SseF, SseG) which are necessary for dynein-dependent movement of the SCV towards a juxtanuclear position (137). This brings the SCV in to close proximity to the Golgi network, putatively for vesicle recruitment, and may involve the SIFs (138). In the absence of these effector proteins the juxtanuclear position is not maintained and bacteria fail to replicate (137).

SspH1, SspH2, SlrP, and SseL manipulate host cell ubiquitination and allow the bacterium to save or target proteins for degradation (131). Most importantly, the effector SspH1 inhibits NF-κB pathway mediated proinflammatory cytokine production (139). This strategy is likely important in preventing cellular cytotoxicity of phagocytes and instead helping to promote delayed-onset cytotoxicity (45). While not achieved through manipulation of ubiquitination, the effector SpvC also has an anti-inflammatory function (140, 141).

A final role for the SPI-2 T3SS is in manipulation of the host cell cytoskeleton independent of the SPI-1 T3SS (131). The effector SteC is involved in polymerization of

an actin meshwork adjacent to the SCV (142, 143). The function of this meshwork is unclear; however, two additional effectors (SspH2 and SrfH) are recruited to this formation and a third (SpvB) appears to have a role in its disassembly, suggesting that *Salmonella* may transiently manipulate the host cytoskeleton to perform an unknown function (144).

Type VI Secretion System (T6SS) genes were first identified in *Rhizobium leguminosarum* as being involved in nitrogen fixation and in *Edwardsiella tarda* as virulence factors (145, 146), however these genes were not understood to be part of a secretion system until a study of the association of *Vibrio cholerae* with *Dictyostelium* amoeba (147). Similar systems were subsequently identified computationally in an extensive number of Gram negative proteobacteria (148, 149) and classified into distinct classes based on sequence similarity (150). T6SS appear to function primarily in interbacterial competition, through translocation of proteins across the cell walls of bacteria but have also been shown to target eukaryotic host cells in a similar manner (151, 152). T6SS resemble a hybrid of Type IV secretion system (T4SS) related proteins localized to the membranes and inverted bacteriophage tail-spike machinery that function to deliver effector proteins directly or indirectly associated with apparatus proteins (153, 154). Together this system provides an innovative protein translocation system which is easily adaptable with various effector proteins often delivered as protein fusions with the secreted apparatus proteins (155).

Most serovars of *S. enterica* encode a T6SS; however, these acquisitions appear to be recent and their functions are not well understood (28). The genes encoding the *S.*

enterica T6SS were first identified as contributing to systemic dissemination as part of the centisome 7 genomic island (now SPI-6) (156). Subsets of these genes were then later identified as part of genome-wide transposon mutagenesis screens for virulence genes (19–21). Further work then suggested a role for the SPI-6 T6SS in late-stages of cell infection based on increased T6SS expression in a *ssrB* mutant, and the observation that SPI-2 T3SS activity decreases over the course of cell infection (157). Regulation of the T6SS in *S. enterica* is not well understood, in contrast with *Pseudomonas aeruginosa* where the T6SS regulation is integrated into global virulence regulon by RetS and LadS (158). Given the contribution of individual T6SS-associated genes to virulence, this system, along with the other T3SS, are likely integrated into a spatiotemporal virulence strategy for successful infection, survival and replication during during host cell infection.

Purpose and Goals of the Present Study

While much has been learned about the genetic basis of virulence in *Salmonella*, notably identification of Type III secretion system genes that contribute to host pathogenesis, their regulation as well as the function of the regulators themselves remains poorly defined. Therefore, the purpose of the present study was to gain insight into the function of the two-component regulatory system sensor kinase SsrA and response regulator SsrB at the protein level. Additionally, we sought to better characterize the regulon of SsrB, including the Type VI secretion system within *Salmonella* Pathogenicity Island 6, which had only recently been identified.

Specific goals highlighted in the following three chapters

1. Histidine residues in the periplasmic domain of the SsrA sensor kinase enhance signaling in response to vacuolar acidification.
    - Modeling of the structure of the periplasmic sensing domain of SsrA
    - Establishment of the role for SsrA in SsrB activity
    - Defining histidine residues necessary for periplasmic pH sensing

2. Identification of the regulatory logic controlling *Salmonella* pathoadaptation by the SsrA-SsrB two-component system.
    - Characterization of the complete SsrB regulon by DNA microarray
    - Determination of the genome-wide interaction map for SsrB by Chromatin immunoprecipitation
    - Elucidation of the SsrB DNA recognition motif

3. Type VI secretion system associated gene clusters contribute to pathogenesis of *Salmonella enterica* serovar Typhimurium.
    - Genetic description of the T6SS assisted via comparative genomics
    - Assessment of contribution of core T6SS genes and T6SS-associated genes to systemic dissemination during infection
    - Investigation of T6SS gene expression and regulation by virulence associated regulatory proteins including SsrB

Chapter II - Multiple histidines in the periplasmic domain of the *Salmonella enterica* sensor kinase SsrA enhance signaling in response to extracellular acidification

Chapter II - Co-authorship statement

Chapter II is a manuscript that is currently in peer review for publication as of submission of this thesis.

The manuscript was written by DTM, BKC and PJS.

The following experiments were performed by collaborators other than myself:

(1) Competitive Infection experiments were performed by JBM (Figure 2.6).

(2) Bafilomycin experiments were performed by SAR (Figure 2.6).

(3) Molecular modeling of SsrA was performed by PJS (Figure 2.2).

Multiple histidines in the periplasmic domain of the *Salmonella enterica* sensor kinase SsrA enhance signaling in response to extracellular acidification

Running title: *Salmonella* senses acidic pH through SsrA histidines

David T. Mulder[1], Joseph B. McPhee[1,*], Sarah A. Reid-Yu [1,*], Peter J Stogios[2], Alexei Savchenko[2], Brian K. Coombes [1,3]

[1] Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, ON L8N 3Z5 CANADA

[2] Department of Chemical Engineering and Applied Chemistry, Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario M5G 1L6 CANADA

[3] Michael G. DeGroote Institute for Infectious Disease Research, McMaster University, Hamilton, ON L8N 3Z5 CANADA

[*] these authors contributed equally

Address correspondence to:

Brian K. Coombes, Department of Biochemistry and Biomedical Sciences, McMaster University, 1200 Main St. West, Hamilton, Ontario CANADA, L8S 4K1. Tel.: 905-525-9140; E-mail: coombes@mcmaster.ca

Summary

The two-component regulatory system SsrA-SsrB in *Salmonella enterica* controls expression of a virulence gene program required for intracellular survival in host cells. SsrA signaling is induced within the acidic host vacuole in which the bacteria reside, however the mechanism by which SsrA senses this intracellular environment is unknown. Here, we show that the periplasmic sensor domain of SsrA is enriched in histidine residues that increase SsrA signaling below external pH of 6. While no single histidine accounted for the full acid-responsiveness of SsrA, we localized the acid-responsiveness principally to 5 histidines in the C-terminal end of the periplasmic sensor domain, with input from additional histidines in the N-terminal end of the senor. A sensor mutant lacking critical pH-responsive histidines was defective for acid-promoted activity, yet retained basal activity similar to wild type at neutral pH, indicating that the role of these histidines is to enhance signaling in response to acidification. In support of this finding, a pH-blind mutant was insensitive to the vacuole acidification blocking activity of bafilomycin, and was attenuated for competitive fitness during infection of mice. Our data demonstrate that SsrA contains a histidine-rich periplasmic sensor that enhances signaling in response to the innate host defense of vacuolar acidification.

Introduction

The mammalian pathogen *Salmonella enterica* subspecies *enterica* is adapted to the intracellular vacuolar environment despite the antibacterial activity of host immune cells (Haraga *et al.*, 2008). Following uptake into host cells, antibacterial defenses are initiated as the vacuole lumen acidifies and becomes subject to reactive oxygen stress (Steele-Mortimer, 2008)(159)(158)(157)(158)(157)(156). Survival and replication by *Salmonella* in this niche requires proteins that sense host-derived signals and integrate these with regulatory networks controlling virulence gene expression (Beier & Gross, 2006). Expression of the *Salmonella* Pathogenicity Island 2 (SPI-2) type III secretion system (T3SS-2) helps establish the *Salmonella*-containing vacuole (SCV) and is required for intracellular survival and replication (Steele-Mortimer, 2008, Kuhle & Hensel, 2004).

T3SS-2 expression is initiated by the SsrA sensor kinase and inhibition of phagosomal acidification prevents this activity (Fass & Groisman, 2009, Arpaia *et al.*, 2011). Based on structural studies, signal or ligand specificity of bacterial histidine kinases is typically achieved by the flexible loops of a Per-Arnt-Sim (PAS) domain containing a rigid central beta-sheet with overlaid helical regions (Krell *et al.*, 2010). Modification of the major and minor ligand binding loops allows for sensing of an extensive number of stimulus types, particularly extracellular signals, which are detected by the PhoQ-DcuS-CitA (PDC or PAS_p) domain subclass (Cheung & Hendrickson, 2010). Following ligand detection by the PDC sensor domain, activation proceeds through structural distortions propagated to

the cytoplasmic histidine kinase domain through a piston-like motion of the transmembrane helical bundle (Cheung & Hendrickson, 2010). Autophosphorylation of the histidine kinase domain initiates a phosphorelay that activates a cognate response regulator for targeted gene expression, which for the SsrA sensor, is SsrB in *S.* Typhimurium (Deiwick *et al.*, 1999, Cirillo *et al.*, 1998, Carroll *et al.*, 2009).

Like other tandem PDC fold sensor kinases, the sensor domain of SsrA is situated in the periplasm, oriented by two transmembrane domains (Fass & Groisman, 2009). Little is known about the mechanism of SsrA sensing but it has been known for some time that acidification is a critical cue (Miao *et al.*, 2002). External acidification of Gram-negative bacteria leads to acidification of the periplasm and cytoplasm, with the cytosol recovering to above neutral within 30 seconds and the periplasmic space remaining acidified (Wilks & Slonczewski, 2007). Acidification of the *Salmonella* vacuole is quickly achieved following bacterial uptake into macrophages (Rathman *et al.*, 1996), with an attendant induction of SsrA signaling (Cirillo *et al.*, 1998). Early work indicated that this induction could be recapitulated *in vitro* by limiting magnesium and phosphate ions in acidified minimal medium (Deiwick *et al.*, 1999, Miao *et al.*, 2002, Lober *et al.*, 2006). However subsequent work has since ruled out vacuolar magnesium limitation as a likely *in vivo* cue (Martin-Orozco *et al.*, 2006). Transcription of SPI-2 genes is significantly enhanced upon invasion of host cells, in comparison to the low levels of transcription observed while the bacteria are extracellular (Brown *et al.*, 2005, Osborne & Coombes, 2011). This intracellular transcription is blocked if vacuolar acidification is inhibited by bafilomycin

(Arpaia *et al.*, 2011), indicating a direct link between virulence gene expression and acidification of the external environment.

Salmonella can sense and respond to increasing cytosolic ATP concentrations as a result of increasing environmental (160). However, sensor proteins exposed to the periplasm can sense acidity directly. Side chains of the amino acids aspartate, glutamate and histidine have pK_a values between 4-6 and therefore are protonated at pH levels observed within the acidified vacuole. Protonation of histidine imidazole (p*Ka* ~6) has been shown to have biological consequences leading to changes in protein conformation as reviewed in (Achilonu *et al.*, 2012). Histidine residues are important for signaling activity of PhoQ and PmrB in *S. enterica* and ArsS in *Helicobacter pylori* (Muller *et al.*, 2009, Perez & Groisman, 2007, Prost *et al.*, 2007), and for activity of the diphtheria toxin T domain (Perier *et al.*, 2007). Here we show that the periplasmic sensing domain of SsrA is enriched in histidine residues that enhance signaling in response to extracellular acidification. A strain lacking critical pH-responsive histidines was insensitive to the acidification-blocking activity of bafilomycin A-1 during macrophage infection, and was attenuated in mice during competitive infections against wild type *Salmonella*. Thus, our data demonstrate that pH-sensing histidine residues confer a selective advantage to *Salmonella* in the host environment.

Results

The SsrA sensor domain has a predicted double PDC domain structure - The periplasmic domain of SsrA in *S. enterica* subspecies *enterica* is closely related to orthologs in other subspecies and is distantly related to orthologs in *Photorhabdus asymbiotica*, *Sodalis spp.*, *Yersinia spp.*, and *Pantoea stewartii* as identified in BlastP queries (Figure 2.1**A**). Attempts to purify any quantity of soluble SsrA protein for structural studies were unsuccessful and so we pursued alternative informatics methods using standard modeling approaches. Secondary structure of the periplasmic sensor domain of SsrA was predicted using the Phyre2 and PSIPRED servers, with the top-scoring hit from Phyre2 (PDB 3LIC, (Zhang & Hendrickson, 2010)) used for further analysis. Fold recognition searches, secondary structure predictions, and alignments with other family 1 sensor domains predict an architecture similar to other tandem-PDC domain sensor kinases where the membrane-distal periplasmic PDC domain is involved in ligand recognition (Figure 2.1**B**). In the SsrA model, a long α-helix (44-85) and two PDC domains (94-209, 213-281) are bounded by two transmembrane domains (13-42, 292-316) (Figure 2.2**A**). PDC domains contain an N-terminal α-helix, two β-strands, a linker region, and ends with three β-strands that form a β2-β1-β4-β3 topology (Zhang & Hendrickson, 2010) (Figure 2.2**B**). The long N-terminal amphipathic α-helix is predicted to form a platform on which the two PDC domains would lie, stabilized by hydrophobic interactions (I49, I50, L53, L56, I61 and L63 with PDC domain 2 [membrane-proximal]; A71, A75, L78 and M79 with PDC domain 1 [membrane-distal]; F68 with W177, W188,

W226 and W274 in the hydrophobic core formed between all three domains) (Figure 2.2**C**). In the model, the remaining hydrophobic amino acids (V55, M59, V62, Y80 and L84) in the platform protrude away from the PDC domains. These exposed residues suggest dimeric interaction sites along the length of helix 1 in a manner similar to other family 1 structures (Zhang & Hendrickson, 2010).

Sequence alignments of the 250 amino acid periplasmic region of SsrA with its orthologs revealed 48 conserved residues that we mapped on to the structural model for downstream validation. Sequence conservation was strongest in the β-strands, while the linker regions (α2 in predicted PDC domain 1, α2 in predicted PDC domain 2) were least conserved. Sequence conservation was weaker in the predicted membrane-distal PDC domain 1 compared to the membrane-proximal PDC domain 2 (Figure 2.1**B**). To better understand this sequence variation we scored conservation between non-*Salmonella* ortholog sequences as well as orthologs within subspecies of *Salmonella* (little variation was observed amongst serovars of *S. enterica* subsp. *enterica)*. Eleven residues conserved in the non-*Salmonella* ortholog set were not conserved in the *Salmonella* group, with eight of these in the membrane-distal PDC domain suggesting selection on this domain (Table 2.1**)**. The Y155 and F120 residues, important for forming the general structure of the ligand recognition site in PDC 1 domains of HK1 family sensor kinases (Zhang & Hendrickson, 2010), were conserved. As expected, the major and minor ligand binding loop sequences that determine ligand specificity were less well conserved. Four residues in the minor ligand binding loop were conserved across both sets. In the major ligand

binding loop, no residues were conserved across both sets, while two and three residues were conserved within *Salmonella* respectively (Table 2.1). Mapping the conservation scores onto secondary structure predictions suggested residues that may be involved in recognizing a *Salmonella*-specific ligand, particularly N142 and E143 in the major loop and Y182 in the minor loop (Figure 2.2**D**, Table 2.1).

Table 2.1. Conserved residues identified for SsrA in *Salmonella* and non-*Salmonella* ortholog alignments.

The table groups all residues within the periplasmic sensing domain in to groups (as shown with a horizontal line) based on level of conservation of the individual residues across all *Salmonella* orthologs (*Salmonella*) or across all non-*Salmonella* orthologs (non-*Salmonella).* The conservation value is a fraction of the number of orthologs that share residue identity at that position with SsrA. Location with respect to HK1 family is determined by sequence alignment (74). Order in each group is based on residue position.

Table 2.1.

Residue	Position	*Salmonella*	Non-*Salmonella*	Domain	Location
S	41	1	1	TMR	TMR
L	53	1	1	Platform	A1
L	63	1	1	Platform	A1
N	65	1	1	Platform	A1
R	67	1	1	Platform	A1
F	68	1	1	Platform	A1
A	71	1	1	Platform	A1
E	72	1	1	Platform	A1
R	73	1	1	Platform	A1
D	74	1	1	Platform	A1
A	75	1	1	Platform	A1
L	78	1	1	Platform	A1
S	98	1	1	PDC1	A2
C	108	1	1	PDC1	A2-B1 loop
P	110	1	1	PDC1	A2-B1 loop
F	120	1	1	PDC1	B1 (HK1 family ligand contact)
F	136	1	1	PDC1	B2
I	137	1	1	PDC1	B2
Y	155	1	1	PDC1	A3
P	167	1	1	PDC1	A3-B3 loop
H	172	1	1	PDC1	A3-B3 loop
W	177	1	1	PDC1	B3
P	180	1	1	PDC1	B3-B4 loop (HK1 family Ligand Minor Loop)
E	181	1	1	PDC1	B3-B4 loop (HK1 family Ligand Minor Loop)
G	187	1	1	PDC1	B3-B4 loop (HK1 family Ligand Minor Loop)
W	188	1	1	PDC1	B3-B4 loop (HK1 family Ligand Minor Loop)
A	193	1	1	PDC1	B4
L	211	1	1	PDC2	A4
I	223	1	1	PDC2	A4-B6 loop
W	226	1	1	PDC2	B6
D	228	1	1	PDC2	B6
L	234	1	1	PDC2	B7
P	235	1	1	PDC2	B7
F	236	1	1	PDC2	B7
L	247	1	1	PDC2	A5
L	252	1	1	PDC2	A5-B8 loop
G	255	1	1	PDC2	B8
W	256	1	1	PDC2	B8
Q	257	1	1	PDC2	B8
P	260	1	1	PDC2	B8-B9 loop
G	261	1	1	PDC2	B9
L	265	1	1	PDC2	B9
R	266	1	1	PDC2	B9
L	269	1	1	PDC2	B9

Residue	Position	*Salmonella*	Non-*Salmonella*	Domain	Location
G	271	1	1	PDC2	B9-B10 loop
P	272	1	1	PDC2	B9-B10 loop
W	274	1	1	PDC2	B10
L	279	1	1	PDC2	B10
P	281	1	1	PDC2	B10
Q	122	0.9	1	PDC1	B1
D	134	0.5	1	PDC1	B1-B2 loop
S	135	0.7	1	PDC1	B2
S	145	0.9	1	PDC1	B2' (HK1 family Ligand Major Loop)
D	154	0.5	1	PDC1	A3
G	185	0.6	1	PDC1	B3-B4 loop (HK1 family Ligand Minor Loop)
V	194	0.9	1	PDC1	B4
K	207	0.5	1	PDC1	B5
D	210	0.9	1	PDC2	A4
L	227	0.9	1	PDC2	B6
G	273	0.9	1	PDC2	B10
M	42	1	0.5	TMR	TMR
Q	44	1	0.5	Platform	A1
H	48	1	0.16	Platform	A1
I	49	1	0.66	Platform	A1
D	52	1	0.16	Platform	A1
S	54	1	0.66	Platform	A1
N	60	1	0	Platform	A1
I	61	1	0	Platform	A1
Q	66	1	0	Platform	A1
E	69	1	0.83	Platform	A1
E	70	1	0	Platform	A1
Y	80	1	0	Platform	A1
Q	81	1	0	Platform	A1
C	82	1	0	Platform	A1
A	85	1	0	Platform	A1
T	86	1	0	Platform	A1-A2 loop
H	90	1	0	Platform	A1-A2 loop
R	99	1	0.83	PDC1	A2
H	100	1	0	PDC1	A2
L	101	1	0.16	PDC1	A2
N	107	1	0.33	PDC1	A2
T	109	1	0.5	PDC1	A2-B1 loop
L	112	1	0	PDC1	A2-B1 loop
K	116	1	0	PDC1	B1
I	126	1	0	PDC1	B1-B2 loop
N	139	1	0	PDC1	B2
N	142	1	0	PDC1	B2-B2' loop (HK1 family Ligand Major Loop)
E	143	1	0	PDC1	B2-B2' loop (HK1 family Ligand Major Loop)
L	146	1	0.83	PDC1	B2' (HK1 family Ligand Major Loop)
T	157	1	0.16	PDC1	A3

Residue	Position	*Salmonella*	Non-*Salmonella*	Domain	Location
L	158	1	0.16	PDC1	A3
Q	159	1	0.33	PDC1	A3
L	168	1	0.16	PDC1	A3-B3 loop
P	170	1	0.5	PDC1	A3-B3 loop
F	175	1	0	PDC1	B3
Y	176	1	0.16	PDC1	B3
S	178	1	0	PDC1	B3
Y	182	1	0.5	PDC1	B3-B4 loop (HK1 family Ligand Minor Loop)
I	183	1	0.66	PDC1	B3-B4 loop (HK1 family Ligand Minor Loop)
S	191	1	0.66	PDC1	B4
V	192	1	0.5	PDC1	B4
Q	197	1	0	PDC1	B4-B5 loop
G	199	1	0.16	PDC1	B4-B5 loop
V	200	1	0.5	PDC1	B4-B5 loop
F	201	1	0	PDC1	B5
L	208	1	0.83	PDC1	B5-A4 loop
L	219	1	0.33	PDC2	A4-B6 loop
D	220	1	0.5	PDC2	A4-B6 loop
D	221	1	0	PDC2	A4-B6 loop
S	222	1	0	PDC2	A4-B6 loop
V	225	1	0	PDC2	B6
Q	229	1	0.16	PDC2	B6-B7 loop
N	230	1	0.16	PDC2	B6-B7 loop
N	231	1	0.83	PDC2	B6-B7 loop
L	233	1	0.16	PDC2	B7
S	237	1	0.83	PDC2	B7-A5 loop
Y	238	1	0	PDC2	B7-A5 loop
Q	241	1	0.16	PDC2	A5
K	242	1	0	PDC2	A5
I	243	1	0.16	PDC2	A5
R	244	1	0	PDC2	A5
T	245	1	0.16	PDC2	A5
E	248	1	0	PDC2	A5-B8 loop
V	250	1	0.16	PDC2	A5-B8 loop
T	251	1	0.33	PDC2	A5-B8 loop
H	253	1	0.5	PDC2	A5-B8 loop
D	254	1	0.83	PDC2	B8
Q	258	1	0.5	PDC2	B8
I	259	1	0.66	PDC2	B8-B9 loop
F	262	1	0	PDC2	B9
L	263	1	0.5	PDC2	B9
I	264	1	0	PDC2	B9
T	267	1	0.33	PDC2	B9
H	270	1	0	PDC2	B9-B10 loop
V	277	1	0.83	PDC2	B10
Y	280	1	0.66	PDC2	B10-A6 loop

Residue	Position	*Salmonella*	Non-*Salmonella*	Domain	Location
Y	282	1	0	PDC2	B10-A6 loop
N	284	1	0	PDC2	B10-A6 loop
L	285	1	0.16	PDC2	A6
L	290	1	0.66	PDC2	A6
K	291	1	0	PDC2	A6
V	43	0.9	0.5	TMR	TMR
K	45	0.6	0.5	Platform	A1
R	46	0.9	0.83	Platform	A1
Q	47	0.9	0.33	Platform	A1
I	50	0.6	0.33	Platform	A1
E	51	0.6	0.83	Platform	A1
V	55	0.2	0	Platform	A1
L	56	0.5	0.16	Platform	A1
S	57	0.7	0.83	Platform	A1
E	58	0.8	0	Platform	A1
M	59	0.9	0	Platform	A1
V	62	0.7	0	Platform	A1
S	64	0.5	0.5	Platform	A1
K	76	0.9	0	Platform	A1
N	77	0.6	0	Platform	A1
M	79	0.8	0	Platform	A1
S	83	0.9	0	Platform	A1
L	84	0.6	0.83	Platform	A1
E	87	0.2	0	Platform	A1-A2 loop
I	88	0.9	0	Platform	A1-A2 loop
H	89	0.5	0	Platform	A1-A2 loop
N	91	0.9	0	Platform	A1-A2 loop
D	92	0.9	0.33	Platform	A1-A2 loop
I	93	0.4	0.16	Platform	A1-A2 loop
F	94	0.6	0	PDC1	A2
P	95	0.9	0	PDC1	A2
E	96	0.2	0	PDC1	A2
V	97	0.3	0	PDC1	A2
S	102	0.6	0	PDC1	A2
V	103	0.9	0.16	PDC1	A2
G	104	0.9	0	PDC1	A2
P	105	0.9	0.33	PDC1	A2
S	106	0.5	0.33	PDC1	A2
T	111	0.8	0	PDC1	A2-B1 loop
N	113	0.7	0	PDC1	A2-B1 loop
G	114	0.5	0	PDC1	A2-B1 loop
E	115	0.2	0.33	PDC1	A2-B1 loop
H	117	0.5	0	PDC1	B1
R	118	0.7	0.33	PDC1	B1
L	119	0.5	0	PDC1	B1
L	121	0.9	0	PDC1	B1

Residue	Position	*Salmonella*	Non-*Salmonella*	Domain	Location
S	123	0.6	0	PDC1	B1-B2 loop
S	124	0.2	0	PDC1	B1-B2 loop
D	125	0.9	0	PDC1	B1-B2 loop
D	127	0.9	0	PDC1	B1-B2 loop
E	128	0.8	0	PDC1	B1-B2 loop
K	129	0.2	0	PDC1	B1-B2 loop
S	130	0.6	0	PDC1	B1-B2 loop
F	131	0.5	0	PDC1	B1-B2 loop
R	132	0.3	0	PDC1	B1-B2 loop
R	133	0.4	0	PDC1	B1-B2 loop
L	138	0.7	0.83	PDC1	B2
H	140	0.9	0	PDC1	B2
K	141	0.5	0	PDC1	B2
I	144	0.9	0.5	PDC1	B2' (HK1 family Ligand Major Loop)
L	147	0.7	0.83	PDC1	B2' (HK1 family Ligand Major Loop)
S	148	0.4	0	PDC1	B2' (HK1 family Ligand Major Loop)
T	149	0.7	0	PDC1	B2'-A3 loop (HK1 family Ligand Major Loop)
D	150	0.6	0	PDC1	B2'-A3 loop (HK1 family Ligand Major Loop)
N	151	0.4	0	PDC1	B2'-A3 loop (HK1 family Ligand Major Loop)
P	152	0.6	0	PDC1	B2'-A3 loop (HK1 family Ligand Major Loop)
S	153	0.9	0	PDC1	B2'-A3 loop (HK1 family Ligand Major Loop)
S	156	0.8	0	PDC1	A3
P	160	0.2	0	PDC1	A3
L	161	0.2	0.16	PDC1	A3
T	162	0.6	0	PDC1	A3
R	163	0.5	0	PDC1	A3
K	164	0.5	0	PDC1	A3
S	165	0.2	0	PDC1	A3
F	166	0.9	0.83	PDC1	A3
Y	169	0.7	0	PDC1	A3-B3 loop
T	171	0.9	0.5	PDC1	A3-B3 loop
A	173	0.9	0	PDC1	A3-B3 loop
G	174	0.9	0	PDC1	B3
E	179	0.5	0	PDC1	B3-B4 loop (HK1 family Ligand Minor Loop)
N	184	0.8	0	PDC1	B3-B4 loop (HK1 family Ligand Minor Loop)
K	186	0.9	0	PDC1	B3-B4 loop (HK1 family Ligand Minor Loop)
H	189	0.2	0	PDC1	B3-B4 loop (HK1 family Ligand Minor Loop)
A	190	0.2	0	PDC1	B3-B4 loop (HK1 family Ligand Minor Loop)
A	195	0.5	0.66	PDC1	B4
D	196	0.9	0.16	PDC1	B4-B5 loop
Q	198	0.9	0	PDC1	B4-B5 loop
F	202	0.8	0	PDC1	B5
E	203	0	0	PDC1	B5
V	204	0.9	0	PDC1	B5
T	205	0.9	0.66	PDC1	B5
V	206	0.9	0.83	PDC1	B5

Residue	Position	*Salmonella*	Non-*Salmonella*	Domain	Location
P	209	0.2	0	PDC1	B5-A4 loop
I	212	0.9	0	PDC2	A4
T	213	0.9	0	PDC2	A4
K	214	0.2	0	PDC2	A4
S	215	0.8	0	PDC2	A4-B6 loop
H	216	0.8	0.66	PDC2	A4-B6 loop
L	217	0.5	0	PDC2	A4-B6 loop
P	218	0.2	0	PDC2	A4-B6 loop
R	224	0.2	0	PDC2	B6
H	232	0.7	0	PDC2	B6-B7 loop
I	239	0.3	0	PDC2	B7-A5 loop
P	240	0.5	0.33	PDC2	B7-A5 loop
Q	246	0.2	0	PDC2	A5
N	249	0.9	0.33	PDC2	A5-B8 loop
T	268	0.9	0	PDC2	B9
S	275	0.7	0	PDC2	B10
L	276	0.5	0	PDC2	B10
T	278	0.5	0	PDC2	B10
G	283	0.9	0	PDC2	B10-A6 loop
H	286	0.2	0	PDC2	A6
N	287	0.2	0	PDC2	A6
R	288	0.9	0.83	PDC2	A6
I	289	0.8	0.16	PDC2	A6

Figure 2.1. Comparison of SsrA to related orthologous proteins

(**A**) The periplasmic domain of SsrA in *S. enterica* subspecies *enterica* is closely related to orthologs in other subspecies and is distantly related to that of orthologs in *Photorhabdus asymbiotica*, *Sodalis spp.*, *Yersinia spp.*, and *Pantoea stewartii*. Represented as a neighbor-joining tree without distance corrections. (**B**) Multiple sequence alignment of SsrA orthologs. PDC domains are indicated with green and blue boxes. Secondary structure elements indicated above the alignment. Residues are colored according to a heat map: blue for least conserved to red for most conserved. Position of histidine residues in *Salmonella enterica Typhimurium LT2* SsrA indicated with dashed boxes and numbered above alignment. (TM: transmembrane region).

Figure 2.1.hesis

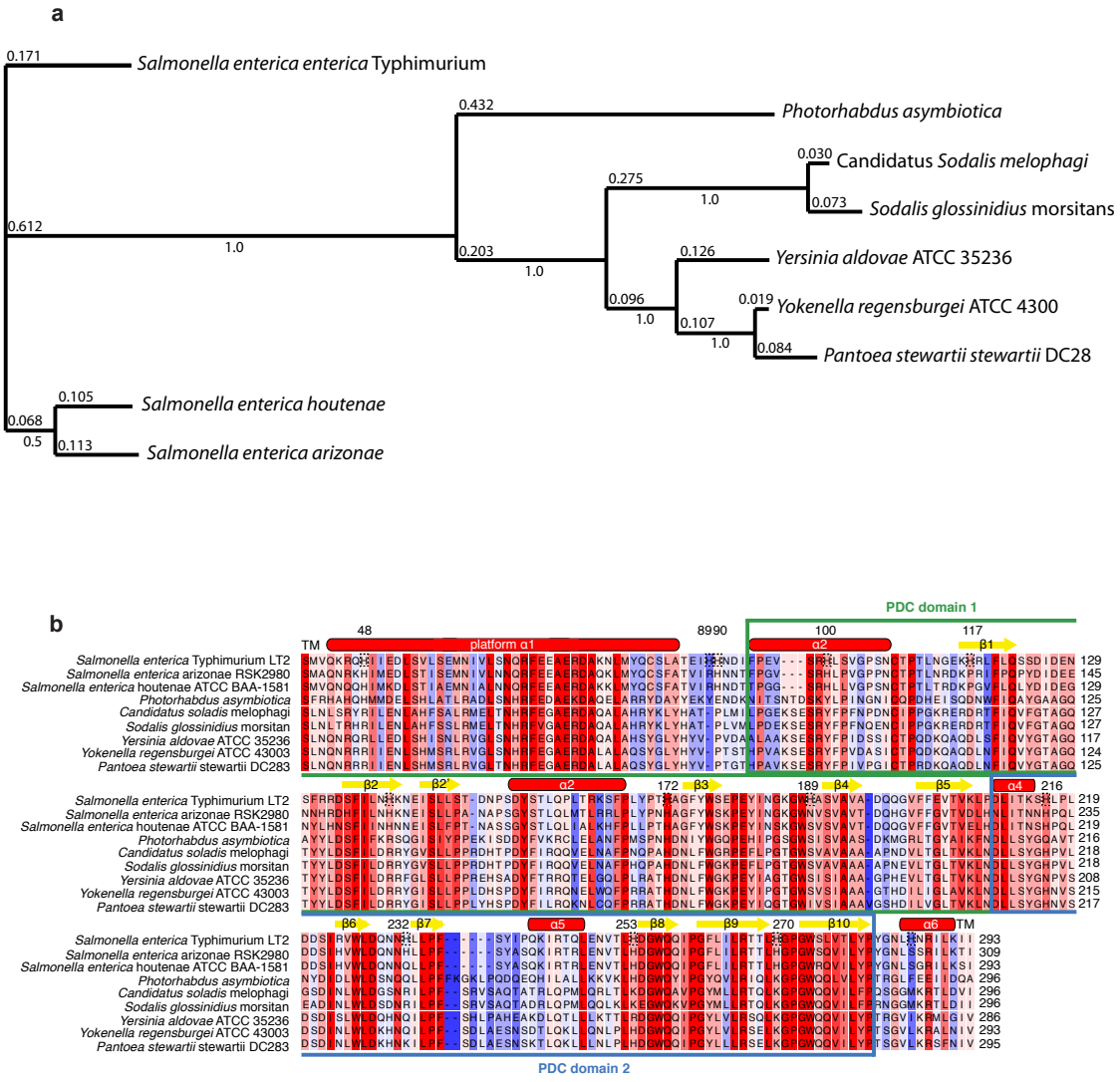
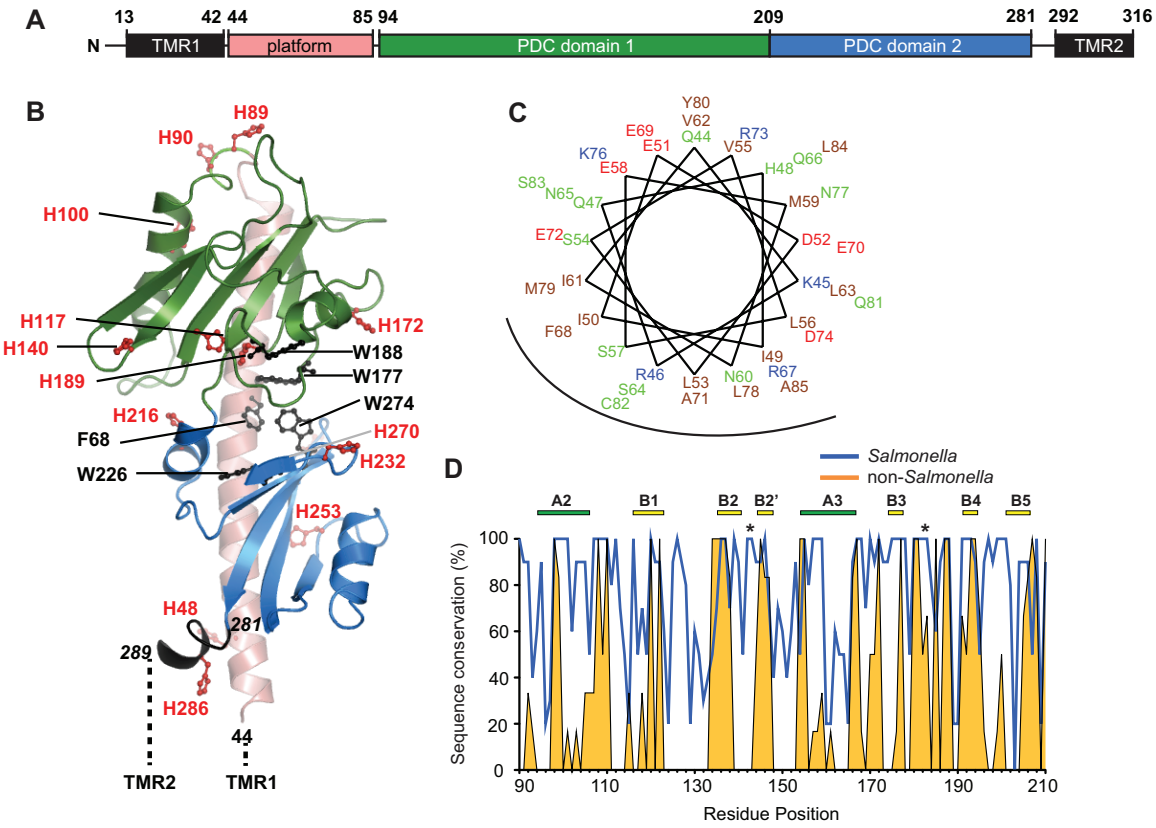


Figure 2.2. Analysis of the SsrA sensor domain.

(**A**) Domain architecture of the periplasmic sensor domain of SsrA (TMR: transmembrane region). Amino acid positions for domain boundaries are indicated. (**B**) Model of the periplasmic sensor domain of SsrA predicted using Phyre2 and PSIPRED servers and based on top scoring hit from Phyre2 (PDB 3LIC, (Zhang & Hendrickson, 2010)). Backbone is colored according to domain architecture in (*A*). Sticks are shown for conserved aromatic residues in the hydrophobic core (black) and for histidine residues (red). Termini of PDC domains are indicated in black italics with dashes indicating links to TMR regions. (**C**) Helical wheel projection of the platform α-helix (residues 44-85) reveals general amphipathic character (brown, hydrophobic residues; green, polar; blue, basic; red, acidic). Semi-circle indicates conserved hydrophobic face. (**D**) Overlay of alignment information between SsrA and *Salmonella* orthologs (blue) and non-*Salmonella* orthologs (orange) for the PDC 1 domain. Structural information is indicated above (α-helix, green; β-sheet, yellow; ligand binding loops, grey). Regions of high conservation limited to *Salmonellae* in PDC-1 and PDC-2 in putative ligand major and minor binding loops are indicated with an asterisk and described in Table 2.1.

Figure 2.2.


A
N
13
42 44
85 94
209
281 292
316
TMR1
platform
PDC domain 1
PDC domain 2
TMR2
B
H89
H90
H100
H117
H140
H189
H172
W188
W177
W274
H216
F68
W226
H270
H232
H253
H48
281
289
H286
44
TMR2
TMR1
C
Y80
V62
Q44
E69
E51
R73
V55
K76
E58
L84
Q66
H48
S83
N65
Q47
M59
N77
E72
S54
D52
E70
M79
I61
K45
L63
Q81
F68
I50
L56
D74
S57
I49
R46
R67
S64
L53
N60
A85
C82
A71
L78
D
Salmonella
non-Salmonella
A2
B1
B2
B2'
A3
B3
B4
B5
Sequence conservation (%)
100
80
60
40
20
0
90
110
130
150
170
190
210
Residue Position

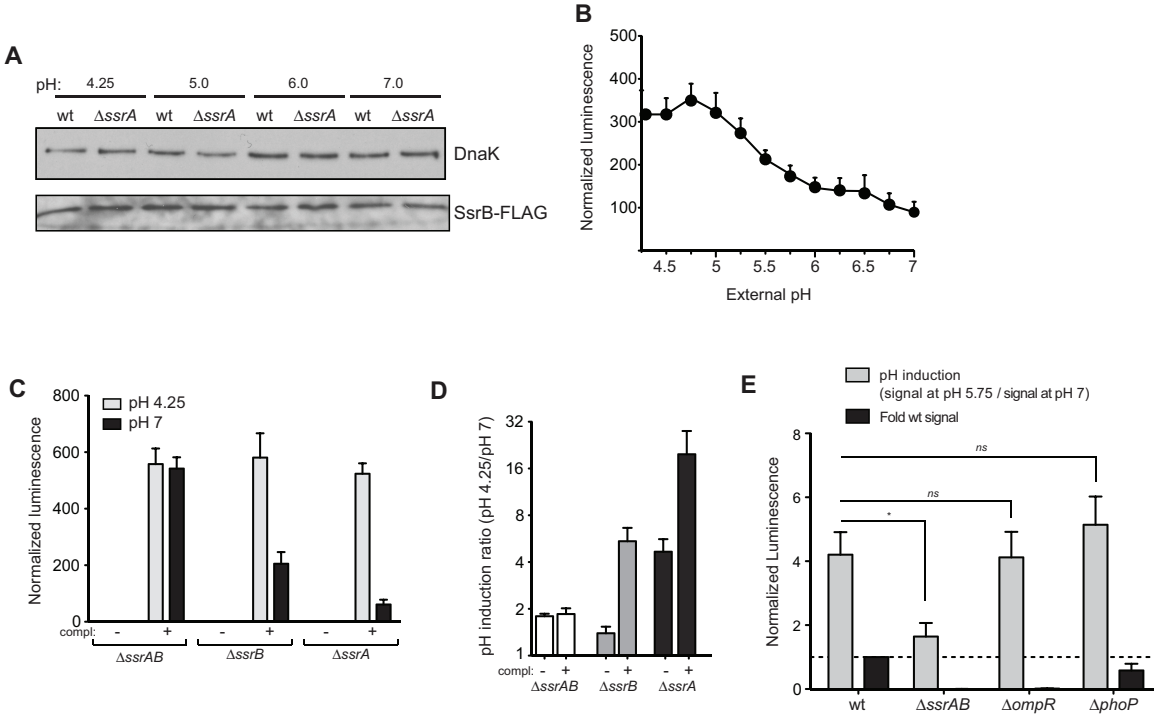
Response to external acidification is integrated through SsrA – Activation of SsrB by SsrA is low under alkaline conditions and high in acidic conditions (Deiwick *et al.*, 1999, Miao *et al.*, 2002). To investigate this phenotype we generated a two-plasmid complementation and reporter system that reports the level of SsrA activity and variants thereof. A Δ*ssrA* strain was generated that kept the upstream promoter and downstream *ssrB* gene intact and which did not affect SsrB expression (Figure 2.3**A**). This mutant was transformed with pP*sseA-lux* that reports SsrB-dependent transcription with the production of bioluminescence, and with a low copy complementation plasmid encoding either wild type SsrA or mutant variants of SsrA expressed under control of its native promoter. The *sseA* promoter was selected based on its tight dependence on SsrA-SsrB for transcription, unlike other promoters with more complex regulatory inputs (Osborne & Coombes, 2011). To verify that this system responds to acidification as expected, we inoculated log-phase cells into the inducing media LPM at a pH range between 4.25 and 7 and measured luminescence over time. During exponential growth, luminescence increased as pH decreased (Figure 2.3**B**), with a 300-fold increase in activity at pH 4.25 over the Δ*ssrA* empty plasmid complementation strain. We confirmed that this signal was dependent on SsrA by measuring luminescence from Δ*ssrA,* Δ*ssrB* and Δ*ssrA*Δ*ssrB* mutant strains and their complemented variants (Fig **3C** and **3D**). The highest level of acid-promoted activity (pH induction) was observed upon SsrA complementation where activity was ~20 fold higher at pH 4.25 than at pH 7, confirming that this phenotype is dependent on SsrA. The sensor kinase PhoQ (76) contributes to post-transcriptional regulation to SsrA-SsrB (Fass & Groisman, 2009, Yoon *et al.*, 2009). To determine

whether other sensor kinases that are activated in the host were involved in this phenotype, we measured reporter activities in Δ*ompR* or Δ*phoQ* mutants. While the overall signal magnitude was decreased in these mutants, as expected, the acid-promoted activity of SsrA was intact in both mutant backgrounds (Figure 2.3**E)** in keeping with previously reported data (Miao *et al.*, 2002).

Figure 2.3. SsrA exhibits acid-promoted activity.

(A) Expression levels of chromosomal SsrB-FLAG from wild type and Δ*ssrA* mutants at indicated time points and pH. (**B**) Activity of an SsrA-SsrB-dependent *sseA-lux* transcriptional reporter was measured over a pH range of 4.25-7. Luminescence was normalized to culture optical density and is shown as signal change over a Δ*ssrA* mutant. Data are means with standard error from three separate experiments. (**C**) The response to acidification is SsrA-dependent and can be complemented. Luminescence was normalized to culture OD and is shown as signal change over a Δ*ssrAB* mutant. Data are the means with standard error from three separate experiments. (**D**) pH induction measured as a ratio of activity at pH 4.25 over pH 7 with (+) and without (-) complementation (compl.). Data are the means with standard errors normalized to culture OD from three experiments. (**E**) Reporter activity in *ssrAB, ompR and phoP* mutant backgrounds of *Salmonella* SL1344. Data are the means with standard errors normalized to culture OD from three experiments.

Figure 2.3.


A
pH: 4.25 5.0 6.0 7.0
wt ΔssrA wt ΔssrA wt ΔssrA wt ΔssrA
DnaK
SsrB-FLAG
B
Normalized luminescence
External pH
C
pH 4.25
pH 7
Normalized luminescence
compl:
ΔssrAB
ΔssrB
ΔssrA
D
pH induction ratio (pH 4.25/pH 7)
E
pH induction
(signal at pH 5.75 / signal at pH 7)
Fold wt signal
Normalized Luminescence
wt
ΔssrAB
ΔompR
ΔphoP

Conserved residues in the SsrA sensor domain are not involved in pH-dependent signaling. To validate the structural model established at the outset of this work, we generated a series of mutations in the broadly conserved residues and measured reporter activity at acidic and neutral pH, and calculated pH induction ratios. We were particularly interested in residues that when mutated, had no effect on the low-level signaling at pH 7 (suggesting correct localization of SsrA and its ability to interact with SsrB), but which compromised the acid-promoted activity. The amphipathic platform α-helix is predicted to form a hydrophobic core between the PDC domains using the conserved residue F68. $SsrA^{F68A}$ was strongly reduced in signal intensity at both pH conditions (Figure 2.4**A**), but was still pH-responsive. Disruption of charged residues in this same α-helix had opposing effects: $SsrA^{D74A}$ had increased signal while $SsrA^{E72A}$ had decreased signal. Mutants $SsrA^{R67A}$, $SsrA^{E72A}$, $SsrA^{C82A}$ had decreased signaling at pH 7, however no mutants in this region were impaired for acid-promoted activity (Figure 2.4**B**).

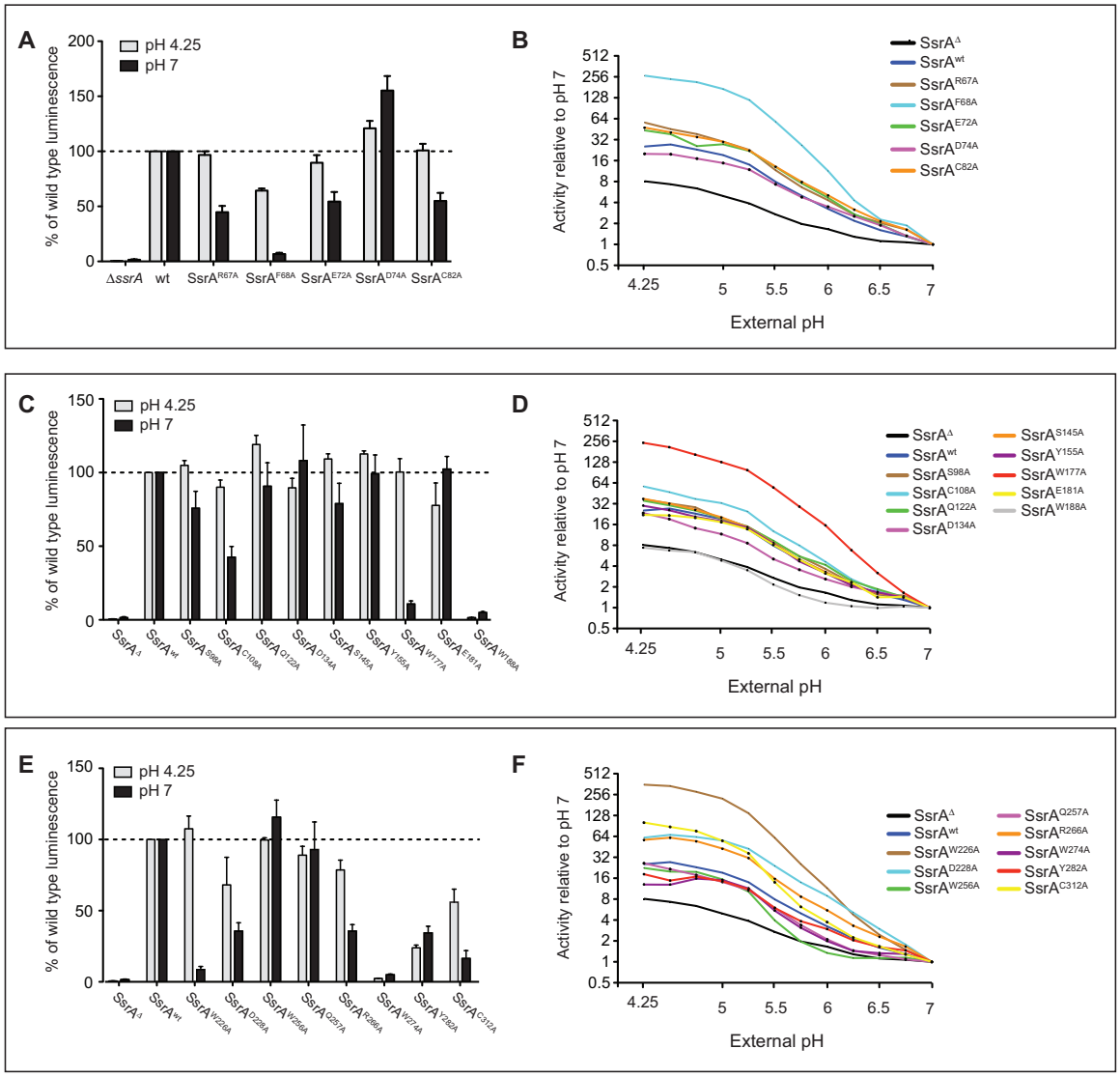
In the membrane-distal PDC domain, mutation of tryptophan residues weakened ($SsrA^{W177A}$) or abolished signaling ($SsrA^{W188A}$) at pH 7, consistent with their predicted position near the hydrophobic core (Figure 2.4**C**). Interestingly, $SsrA^{W177A}$ had wild type signaling at pH 4.25 suggesting that it is dispensable in this activated state. $SsrA^{S98A}$ and $SsrA^{C108A}$ had slightly reduced signaling at pH 7. Aside from the tryptophan mutants, none of the mutants in the membrane-distal PDC domain had impaired pH-responsiveness (Figure 2.4**D**). Mutation of tryptophan residues in the membrane-proximal PDC domain had major effects on signaling (Figure 2.4**E**). Disruption of tryptophan

residues W226 and W274 both disrupted signaling at pH 7 but SsrA^{W226A} had wild type levels of activity at pH 4.25 indicating that it is also dispensable in this activated state. Mutants SsrA^{D228A} and SsrA^{R266A} had decreased levels of activity under both conditions, but pH-responsiveness was intact when baseline signaling was not abolished (Figure 2.4**F**). These experiments indicated that conserved residues in the SsrA sensor domain, while important structurally, are not directly involved in SsrA's ability to respond to external acidification.

Figure 2.4. Mutagenesis of conserved residues in the SsrA sensor domain.

Amino acids conserved in the periplasmic domain between SsrA and orthologs were mapped to the structural model, which were distributed in the putative N-terminal helix, (**A-B**); the PDC1 domain, (**C-D**); and the PDC2 domain, (**E-F**). Reporter activity was measured over the range of pH 4.25 to pH 7.00 in LPM media and displayed relative to wild type or as pH induction where the readings for each strain are normalized its pH 7.00 value. Data are from cultures normalized to OD and are the means with standard errors from three experiments.

Figure 2.4.


A
% of wild type luminescence
pH 4.25
pH 7
ΔssrA
wt
SsrA^R67A
SsrA^F68A
SsrA^E72A
SsrA^D74A
SsrA^C82A
B
Activity relative to pH 7
External pH
SsrA^Δ
SsrA^wt
SsrA^R67A
SsrA^F68A
SsrA^E72A
SsrA^D74A
SsrA^C82A
C
D
SsrA^S98A
SsrA^C108A
SsrA^Q122A
SsrA^D134A
SsrA^S145A
SsrA^Y155A
SsrA^W177A
SsrA^E181A
SsrA^W188A
E
F
SsrA^W226A
SsrA^D228A
SsrA^W256A
SsrA^Q257A
SsrA^R266A
SsrA^W274A
SsrA^Y282A
SsrA^C312A

The periplasmic domain of SsrA is enriched with histidine residues – Since conserved residues were not necessary for acid-promoted activity we hypothesized that it may instead be a *Salmonella* specific adaptation conferred by residues unique to this group, in which histidine figures prominently. Given that histidine can drive reversible pH-sensitive protein conformational changes at ~ pH 6, we focused our investigation on histidine residues in the SsrA periplasmic domain of which there are 13 (~5% frequency). A computational analysis of periplasmic histidine frequency in all members of the Cluster of Orthologous Groups (COG) T (signal transduction) family (>400 amino acids, n=50) in *S.* Typhimurium revealed SsrA to be the most histidine-enriched member of this COG (4.81%, Mean: 1.94% st. dev: 1.03%), and among non-*Salmonella* SsrA orthologs *(Sodalis spp.* 2.7-3.5%, *P. asymbiotica* 3.9%, *P. stewartii* 3.9%, *Yersinia spp*. 3.9%), suggesting that histidine enrichment in SsrA may be a unique selective adaptation.

Periplasmic histidine residues are required for acid-promoted activity of SsrA – Below pH 6, the imidazole ring of histidine is protonated, which has been shown in other proteins to drive conformational changes (Achilonu *et al.*, 2012). When mapped on the structural model, most of the 13 histidines in the SsrA periplasmic domain lay within putative loop regions, particularly in the C-terminal end of the membrane-proximal domain where 6 of 7 histidines are in predicted loops (Figure 2.5**A**). We substituted each histidine with alanine and measured the signaling ability of each mutant at low and high pH (Figure 2.5**B**) and calculated pH induction ratio (Figure 2.5**C**). No single histidine mutant was defective for acid-promoted activity. However we observed a large defect in

signaling under both pH conditions for the SsrA^{H48A} mutant and an increased activity at neutral pH for the SsrA^{H286A} mutant, which we inferred to mean they have structural roles based on their immediate proximity to TMRs. Mutants SsrA^{H100A}, SsrA^{H172A}, SsrA^{H189A}, and SsrA^{H270A} had slightly decreased activity at pH 7, whereas mutants SsrA^{H172A}, SsrA^{H270A} had decreased activities at pH 4.25.

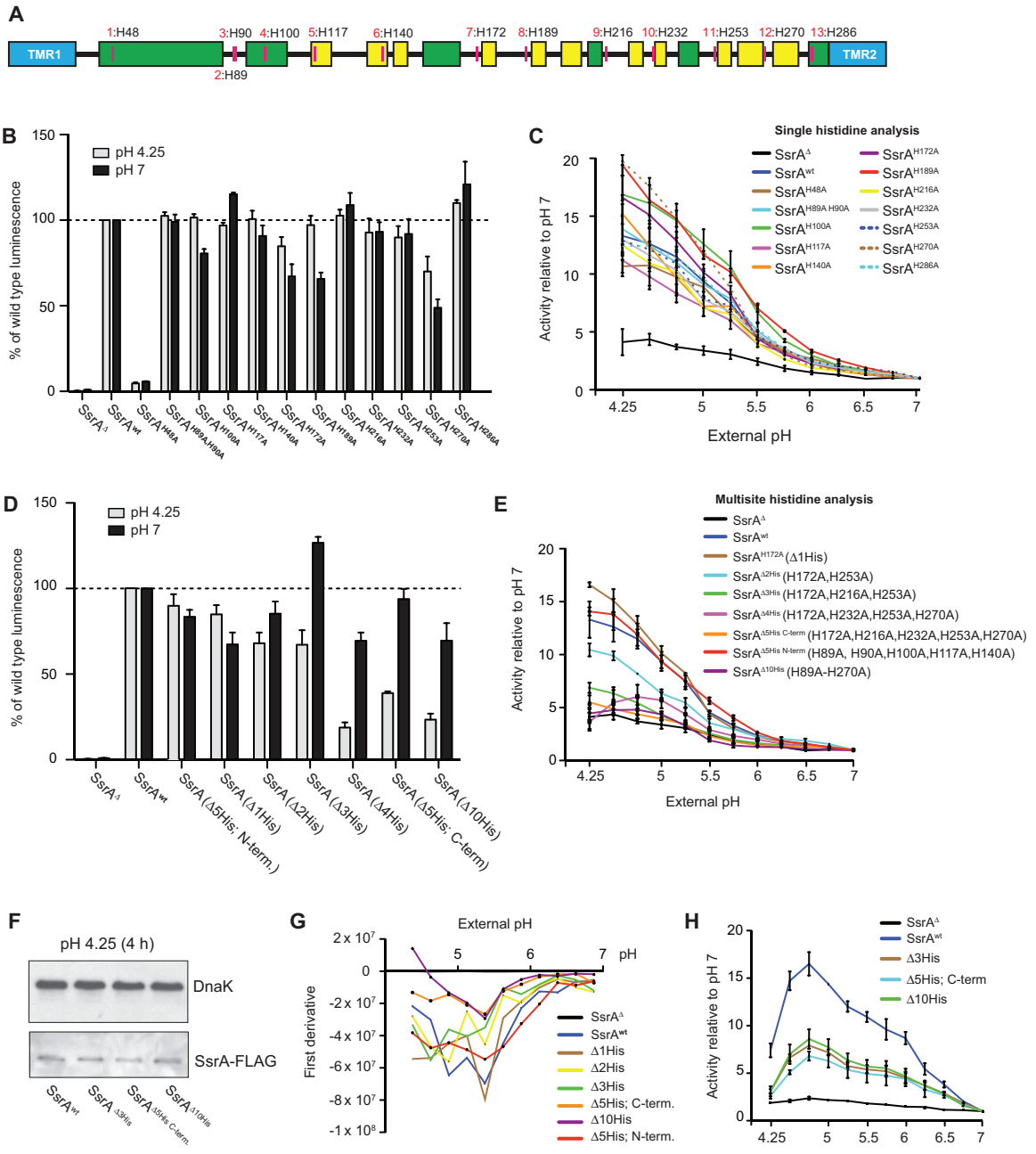
Since no single histidine was responsible for the full acid responsiveness of SsrA, we hypothesized that the concerted action of multiple histidines may be involved, as is the case for the diphtheria toxin T domain (Perier *et al.*, 2007). We generated compounded mutants for histidine and measured their activity at neutral and acidic pH (Figure 2.5**D**) and pH induction ratios (Figure 2.5**E**). Histidine residues H48, H286 and H189 were excluded in this analysis because we concluded from other experiments that they likely have core structural roles. A compound mutant of the membrane-distal histidine residues (SsrA$^{\Delta 2\text{-}6His}$) did not differ significantly from wild type while a compound mutant of membrane-proximal histidine residues (SsrA$^{\Delta 7,9\text{-}12His}$) was defective for acid-promoted signaling and had acid induction ratio *in vitro* similar to an *ssrA* deletion. A combined mutant incorporating all ten of these residues (SsrA$^{\Delta 10His}$) was statistically similar to SsrA$^{\Delta 7,9\text{-}12His}$, consistent with the previous data showing that histidines 2 to 6, while contributing additively to the pH response, are not the major pH-responsive residues. The activity of all multisite histidine mutants remained above 75% of wild type at neutral pH, indicating that these mutations preferentially blunted the acid-promoted activity. Western blot analysis of the various SsrA histidine mutants confirmed that these substitutions did

not greatly affect SsrA levels (Figure 2.5**F**). A first-derivatives analysis of these data showed a peak activity change at pH 5.5, consistent with the pK_a of ~ 6 for histidine imidazole and that the membrane-proximal histidines are required for this response (Figure 2.5**G**). A similar defect in pH response was observed for these mutants when expressed under control of the native *ssrA* promoter following chromosomal replacement (Figure 2.5**H**).

Figure 2.5. Mutagenesis of histidine residues in the SsrA sensor domain.

(**A**) The positions of the thirteen histidines in the periplasmic domain are indicated with respect to the secondary structure predictions. (**B**) SsrA reporter activity expressed as fraction of wild type from strains expressing wild type SsrA, and the indicated single $SsrA^{His}$-mutations at pH 4.25 and pH 7. (**C**) Single histidine to alanine mutants over the range of pH 4.25-7. Luminescence is expressed as the change in signal over that at pH 7. (**D**) SsrA reporter data from compound histidine mutants. Reporter activity expressed as fraction of wild type from strains expressing wild type SsrA, and the compounded $SsrA^{His}$-mutations at pH 4.25 and pH 7. (**E**) SsrA reporter data from compound histidine mutants over the range of pH 4.25-7. Luminescence is expressed as the change in signal over that at pH 7. (**F**) Expression levels of SsrA:FLAG and mutant variants from the low copy plasmid pWSK129 under control of the endogenous PssrA promoter during exponential growth at pH 4.25. (**G**) SsrA reporter activity from compound mutants represented as a first derivative to show differences in response between pH 5-6. (**H**) SsrA reporter activity from chromosomal *ssrA* replacement mutants the range of pH 4.25 to pH 7.00 with each mutant reading normalized to its pH 7.00 value. Data are normalized to culture OD and are the means with standard errors from three experiments.

Figure 2.5.


A
TMR1
TMR2
1:H48
2:H89
3:H90
4:H100
5:H117
6:H140
7:H172
8:H189
9:H216
10:H232
11:H253
12:H270
13:H286
B
pH 4.25
pH 7
% of wild type luminescence
C
Single histidine analysis
Activity relative to pH 7
External pH
D
pH 4.25
pH 7
% of wild type luminescence
E
Multisite histidine analysis
Activity relative to pH 7
External pH
F
pH 4.25 (4 h)
DnaK
SsrA-FLAG
G
External pH
First derivative
pH
H
Activity relative to pH 7

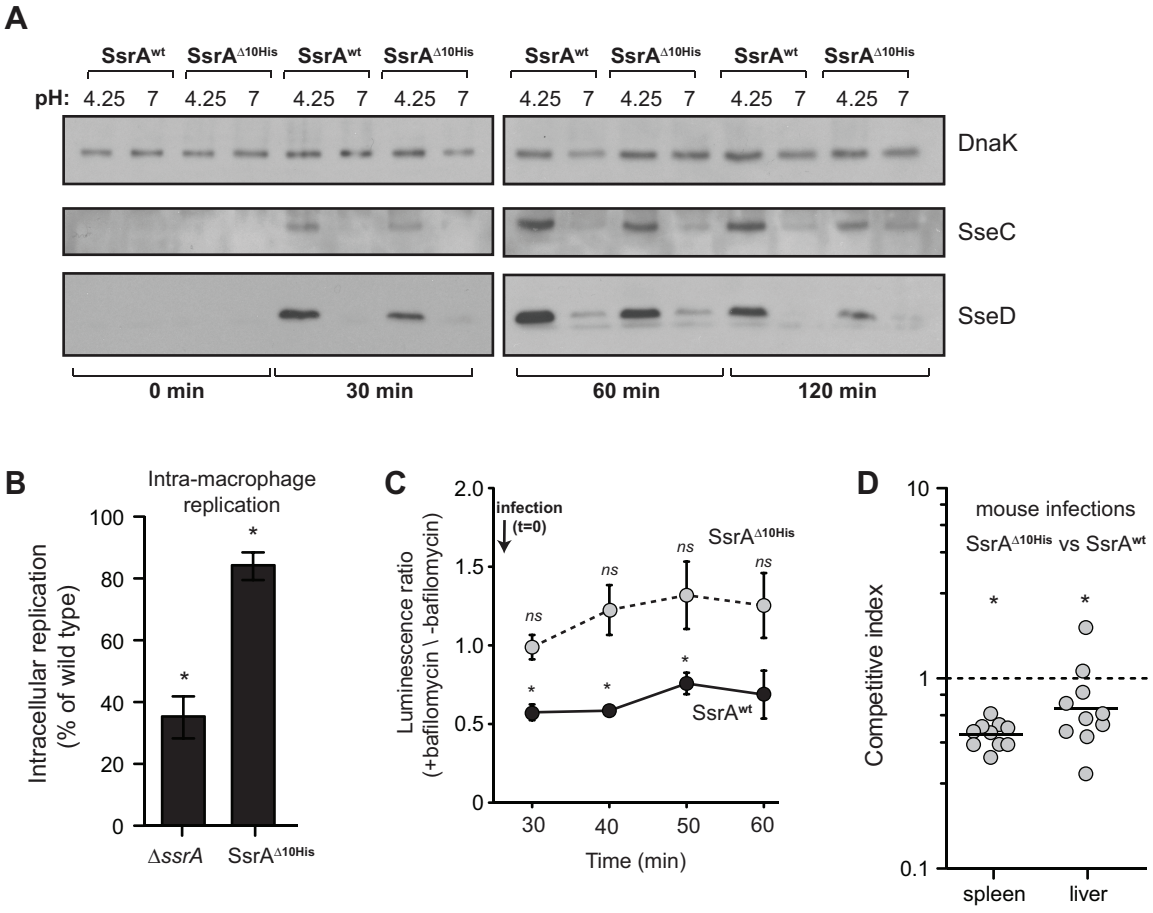
pH-sensitive histidine residues are required for bacterial fitness during infection – To determine whether pH sensing by the histidine-enriched periplasmic domain of SsrA contributes to fitness during infection we replaced *ssrA* on the chromosome with the $SsrA^{\Delta 10His}$ variant (H89A, H90A, H100A, H117A, H140A, H172A, H216A, H232A, H253A, H270A). We first tested whether this mutant had decreased levels of the translocon proteins SseC and SseD. In cells with $SsrA^{wt}$, levels of SseC and SseD increased within 30 min following transfer to acidic minimal medium but remained low under neutral conditions (Figure 2.6**A**). Cells expressing $SsrA^{\Delta 10His}$, while able to produce SseC and SseD, had ~50% reduced levels of these proteins compared to wild type. In keeping with this result, we found that the $SsrA^{\Delta 10His}$ strain was significantly attenuated for intracellular survival in macrophages but not to the same extent as Δ*ssrA* (Figure 2.6**B**). Blocking acidification of the SCV with bafilomycin (Rathman *et al.*, 1996) inhibits gene expression of SsrA-SsrB regulated effectors (Arpaia *et al.*, 2011). Our signaling model predicted that the SsrA histidine mutant would be insensitive to the acidification blocking activity of bafilomycin. In macrophages pre-inhibited with the vacuole acidification-blocking agent bafilomycin A-1, wild type SsrA reported ~50% less activity compared to infection of cells without bafilomycin, confirming that failure to acidify the *Salmonella*-containing vacuole compromised SsrA signaling activity. However, $SsrA^{\Delta 10His}$ was insensitive to bafilomycin and reported equivalently in the presence or absence of the drug (Figure 2.6**C**) indicating that the $SsrA^{\Delta 10His}$ was defective for acid-promoted activity during infection. In support of this, *Salmonella* expressing $SsrA^{\Delta 10His}$ were significantly attenuated in mice compared to wild type following competition infection (Figure 2.6**D**).

These data indicate that SsrA signal amplification achieved by pH-responsive histidines is selective in the host environment.

Figure 2.6. Functional analysis of a pH-blind histidine mutant.

ssrA was replaced on the chromosome with a variant lacking the ten periplasmic histidines that contribute to acid-promoted activity ($SsrA^{\Delta 10His}$). (**A**) Expression of effector proteins SseC and SseD following incubation at indicated time points at indicated pH. (**B**) Intracellular replication, expressed as percent of wild type, in macrophages between 2 and 20 h post infection represented as a fraction of wild type. Data are the means with standard errors from five experiments. (**C**) SsrA reporter activity following intracellular infection of bafilomycin-treated macrophages, or DMSO-treated control macrophages. Data is expressed as bafilomycin-sensitivity of SsrA (luminescence in the presence of bafilomycin/luminescence without bafilomycin). Data are from three experiments. (**D**) Competitive infection of C57BL/6 mice. Mice were infected by intraperitoneal injection with equal numbers of wild type *S.* Typhimurium and the $SsrA^{\Delta 10His}$ strain and the competitive index was calculated after 3 days. Each data point is from one mouse. Data is from two independent experiments.

Figure 2.6.


A
SsrA^wt
SsrA^Δ10His
pH: 4.25 7
DnaK
SseC
SseD
0 min
30 min
60 min
120 min
B
Intra-macrophage replication
Intracellular replication (% of wild type)
ΔssrA
SsrA^Δ10His
C
infection (t=0)
Luminescence ratio (+bafilomycin \ -bafilomycin)
Time (min)
ns
D
mouse infections
SsrA^Δ10His vs SsrA^wt
Competitive index
spleen
liver

Discussion

Changes in *Salmonella* gene expression occur rapidly in the host environment, allowing for adaptive responses to cope with host defenses (Yoon *et al.*, 2009). Vacuolar acidification is an innate immune response in mammals that serves to limit pathogen replication. After ingestion of bacteria, the phagosomal compartment is acidified to pH 5.7-6.0 within 6 minutes (Rathman *et al.*, 1996, Lukacs *et al.*, 1991), thus presenting an early intracellular cue for bacteria to detect. In *Salmonella*, induction of acid tolerance involves the OmpR/EnvZ and PhoP/PhoQ two component regulatory systems (Bang *et al.*, 2002, Riesenberg-Wilmes *et al.*, 1996, Bearson *et al.*, 1998)(161–163)(160–162)(159–161)(158–160)(159–161)(158–160)(157–159), which have regulatory input at the promoters for *ssrA, ssrB,* and some effector genes (Osborne & Coombes, 2011, Yoon *et al.*, 2009). The pH-sensing mechanism we have described for SsrA is independent of these histidine kinases because the reporter we used is dependent entirely on SsrA-SsrB and the phenotype persists in Δ*ompR* and Δ*phoP* mutants.

Our data support a model whereby histidine residues in the membrane-proximal periplasmic sensor domain of SsrA confer acid-promoted signal enhancement. The apparent enrichment of pH-responsive histidines outside of the canonical ligand-binding region (PDC-1) of this family of sensors (Cheung & Hendrickson, 2010, Zhang & Hendrickson, 2010) suggests that SsrA has evolved to integrate pH information in conjunction with another putative ligand(s). The PhoQ sensor in *Salmonella* has also evolved unique modifications in the membrane-proximal PDC domain to integrate

multiple inputs such as antimicrobial peptides, pH and fatty acids (Prost *et al.*, 2007, Viarengo *et al.*, 2013, Bader *et al.*, 2003).

While the genetic regulation of SsrA is well understood, little is known about its structure and function (Fass & Groisman, 2009). We made numerous attempts to purify SsrA and variants thereof for structural characterization, however the protein expressed poorly or was insoluble upon overexpression. Nevertheless, we were able to validate our structural model by showing that the predicted inter-domain hydrophobic regions are essential for function. Of interest were two tryptophan residues (W177 and W226) that are required for SsrA function at neutral pH, but were dispensable for signaling under acidic conditions suggesting two structural orientations of the protein. These residues may be key to understanding the structural changes that accompany the acid-promoted activity of SsrA, and perhaps tandem-PDC sensor kinases in general (Cheung & Hendrickson, 2010, Sevvana *et al.*, 2008). Future work will also focus on residues N142, E143, and Y182 in the putative membrane-distal ligand-binding region, as the conservation of these among *Salmonella* subspecies suggests the possibility of a distinct ligand recognized by this domain in *Salmonella*.

Previous work has shown that some sensor kinases incorporate amino acids with ionizable side chains including aspartate, glutamate and histidine in their signaling mechanism (Muller *et al.*, 2009, Perez & Groisman, 2007, Prost *et al.*, 2007). Localized histidine protonation can lead to changes in protein function by affecting global protein

conformational states (Achilonu *et al.*, 2012). Whereas in other pH-responsive sensor kinases, single histidine residues were shown to promote acid responsiveness, the SsrA mechanism uses multisite histidines that work in a concerted manner for acid-promoted activity. Only one of these histidines is conserved outside of the *Salmonella* subspecies, and which on its own does not significantly contribute to acid sensitivity, suggesting that SsrA histidine enrichment is a *Salmonella*-specific adaptation. Given that individual histidines were additive toward the acid-promoted activity of SsrA, multisite histidine modification may help set activation thresholds for the regulated activity of SsrA during transition to the intracellular environment. This activity was selective in the host environment because *Salmonella* carrying a pH-blind SsrA variant was significantly attenuated for competitive infection with wild type, even though it could still produce modest activity through a pH-independent mechanism. This study represents the first characterization of the mechanism of pH mediated signal enhancement for SsrA and provides insight in to the evolution of this regulatory phenotype.

Experimental Procedures

Growth Medium Compositions. LB media was made as follows: 10 g tryptone, 10 g NaCl, 5 g yeast extract per liter of water. N9 media was made as follows: 0.1 M Tris base, 0.1% casamino acids, 5 mM KCl, 0.5 mM K_2SO_4, 1 mM KH_2PO_4, 5.2 mM glycerol, 200 µM bis-TRIS, 12 µM $MgCl_2$, and pH adjusted to indicated pH at room temperature with HCl. M9 media was made as follows: 0.1% casamino acids, 11.1 mM glucose, 2 mM $MgSO_4$, 1 mM Na_2HPO_4, 4.5 mM KH_2PO_4, 1.72 mM NaCl, 3.72 mM NH_4Cl, 100 µM $CaCl_2$, and pH adjusted to indicated pH at room temperature with NaOH. LPM media was made as follows: 5mM KCl, 7.5mM $(NH_4)_2SO_4$, 0.5mM K_2SO_4, 80mM MES, 38mM glycerol, 0.1% casamino acids, 25uM $MgCl_2$, 337uM $PO4^{3-}$ and adjusted to indicated pH at room temperature with NaOH. When used, antibiotic concentrations were ampicillin 100 µg/mL, kanamycin 50 µg/mL, chloramphenicol 34 µg/mL, streptomycin 50 µg/mL. In test experiments after four hours of bacterial growth in LPM, the pH of the culture remained accurate to within ~0.2 units.

Culturing conditions. Prior to reporter experiments, *Salmonella* cultures were grown overnight in LB with antibiotics at 37°C with shaking, and then sub-cultured 1:50 into LB with antibiotics for 3 h. This sub-culture was done to avoid stationary-phase activity of SsrB (164).

Histidine frequency analysis. Annotation information for all *S*. Typhimurium LT2 (NCBI Genbank accession number NC_003197) proteins were loaded from the .ptt genome file and sequences were loaded from the .faa genome file. One hundred eighty-one proteins classified as COG category T were further analyzed. To eliminate response regulators 76 proteins of 400 amino acids or greater were retained. These proteins were then manually annotated for transmembrane domains using the SMART database (Letunic *et al.*, 2012). The intervening periplasmic regions were scored for histidine frequency.

Cloning. All primers and strains used and generated in this study are listed in Supplementary Tables A1.4 and A1.5. Mutants were generated using lambda Red mutagenesis (Datsenko & Wanner, 2000) and verified by sequencing. Resistance markers were removed by transformation with the helper plasmid pCP20. The wild type complementation construct was generated by cloning ssrA and its promoter from *S*. Typhimurium SL1344 into XbaI/SalI digested pBluescript, subcloning into the low copy expression vector pWSK129, and transforming into S. Typhimurium for expression (Wang & Kushner, 1991). To generate mutant complementation constructs the wild type gene in pBluescript was modified by site-directed mutagenesis, sub-cloned in to pWSK129 and transformed in to *S*. Typhimurium SL1344 for expression. Site-directed mutagenesis primers containing the desired mutation were used to amplify pBluscript:PssrA-ssrA with Pfu polymerase (Fermentas). Amplified DNA was purified on silica columns, digested with DpnI to remove parent plasmid, and transformed into E. coli DH5α. Sequence verified plasmids were transformed into the *S*. Typhimurium SL1344

ΔssrA strain containing the pGEN:PsseA-luxCDABE reporter plasmid and recovered on LB Kn-Amp plates (Lane *et al.*, 2007, Osborne & Coombes, 2011). The alanine mutants, SsrA$^{H172A, H216A, H232A, H253A, H270A}$ and SsrA $^{H89A, H90A, H100A, H117A, H140A, H172A, H216A, H232A, H253A, H270A}$ were synthesized by Genscript and sub-cloned from pUC57 in to pWSK129 as previously described. Chromosomal replacement mutants were generated by first replacing the periplasmic domain coding sequence of ssrA with a kanamycin resistance gene as described above using lambda Red mediated mutagenesis. The P*ssrA-ssrA* insert for the periplasmic domain mutant of choice was then sub-cloned from pWSK129 in to the suicide vector pRE112 and recovered in *E. coli* DH5α λpir (Edwards *et al.*, 1998). The vector was then transformed into the conjugation strain *E. coli* S17-1 λpir and mated with the kanamycin marked, *ssrA* disrupted strain of *S.* Typhimurium. Single recombinants were selected on LB Cm-Kn plates, grown in LB with 5% sucrose without antibiotics, serial diluted and plated on LB containing 5% sucrose. Colonies were replica plated on LB Sm-Kn-Cm plates to isolate Kn-Cm sensitive recombinants that had replaced the chromosomal kanamycin marker. Recombinants were verified by amplifying the P*ssrA-ssrA-ssrB* region and sequencing.

SsrA structural modeling. Domain prediction in SsrA was performed via the CDD (Marchler-Bauer *et al.*, 2013), PFAM (Punta *et al.*, 2012), SMART (Letunic *et al.*, 2012) and Interpro (Hunter *et al.*, 2012) servers. Fold recognition was performed via the Phyre2 (Kelley & Sternberg, 2009) and PSIPRED/GenTHREADER (Buchan *et al.*, 2010) servers. Secondary structure was predicted via the Phyre2 and PSIPRED servers, with the top

scoring hit from Phyre2 (PDB 3LIC, (Zhang & Hendrickson, 2010)) used for further analysis. Transmembrane helices were predicted using the TMHMM server (Kahsay *et al.*, 2005); amino acids scoring over 0.25 were considered part of a transmembrane helix. α-helical wheel projections were generated by modifying the output of the server at http://trimer.tamu.edu/cgi-bin/wheel/wheel.pl. Structure figures were prepared using Pymol (www.pymol.org).

Macrophage cell-based assays. RAW 264.7 macrophages were seeded 16 h prior to infection in 24 well plates at $2x10^5$ cells per well in DMEM + 10% FBS and incubated at 37°C with 5% CO_2. Bacteria grown overnight in LB were pelleted and resuspended in 400 µL DMEM + 10% FBS and 100 µL human serum for 25 minutes at 37°C. Bacteria were then diluted in DMEM + 10% FBS to obtain a multiplicity of infection of 100. One-mL of inoculum was applied to wells and incubated for 30 min for infection. Infected cells were washed three times with PBS and incubated in 1 mL DMEM + 10% FBS + 100 µg/mL gentamicin for 90 min at 37°C. Cells were washed twice with PBS. Half the wells were lysed in lysis buffer (1% Triton X-100, 0.1% SDS, PBS), serial diluted in PBS and plated on LB-Sm plates. Bacteria were counted to represent a 2 h time point. The other half of the wells were returned to the incubator in 1 mL DMEM + 10% FBS + 10 µg/mL gentamicin until 20 h post infection at which point they were washed twice with PBS and lysed in lysis buffer, serial diluted in PBS and plated on LB-Sm plates. Colony forming units were averaged and a 20 h/2 h cfu ratio was calculated to represent fold replication

over the course of the experiment. Experiments were independently performed at least three times with triplicate technical replicates.

Transcriptional reporter assay. RAW 264.7 murine macrophages were seeded 16 h prior to infection at a density of $5x10^4$ cells per well in a 96-well plate in complete RPMI medium (10% FBS, 1% HEPES, 1% L-glutamine, 1% sodium pyruvate, and 1% non-essential amino acids) and grown overnight at 37°C in 5% CO_2 in the presence of 100 ng/mL LPS. Prior to infection, cells were treated with either 1% DMSO or 500 nM bafilomycin A-1 for 30 min. Overnight cultures of *S.* Typhimurium containing the P*sseA-luxCDABE* transcriptional fusion reporter were grown in LB with antibiotics. These cultures were sub-cultured for 3 h prior to infection for 30 min at a multiplicity of infection of 20:1. Infected cells were washed and luminescence was recorded over the first 30 minutes. Cells were lysed, diluted, and plated for cfu determinations.

Mouse infections. All experiments with animals were conducted according to guidelines set by the Canadian Council on Animal Care and were approved by the Animal Review Ethics Board at McMaster University. *S.* Typhimurium (*ushA*:Cm) (Coombes *et al.*, 2005) and the $SsrA^{\Delta 10His}$ strain were competed in infections of C57BL/6 mice. Equal volume of each overnight culture was mixed and the bacteria were resuspended in 100 mM HEPES/0.9% NaCl (pH 8.0) to ~5 X 10^8 cfu/ml. Mice were orogastrically infected with 200 µl of this suspension. For intraperitoneal infections, the mixed bacterial culture was resuspended to 10^7 cfu/ml in PBS and mice were injected intraperitoneally with 100

µl of this suspension. The input ratio of bacteria was determined by serial dilution and replica plating on streptomycin and chloramphenicol. Intraperitoneally infected mice were sacrificed 48 h post-infection while orogastrically infected mice were sacrificed at 72 h post-infection. The cecum, liver and spleen were aseptically removed from each mouse, homogenized in 1 ml of PBS and serial dilutions of each organ were plated on LB-Sm100. Colonies were replica plated to LB-Cm34 and LB-Sm100. The competitive index was expressed as: C.I. = (cfu $SsrA^{\Delta 10His}_{output}$/cfu *ushA*:$Cm_{output}$)/(cfu $SsrA^{\Delta 10His}_{input}$/cfu *ushA*:$Cm_{input}$).

Acknowledgements

We are grateful to Brian Tuinema and other members of the Coombes laboratory for ongoing discussions on this work. D.T.M. was supported by a Frederick Banting and Charles Best Canada Graduate Scholarship. S.A.R. was supported by an Ontario Graduate Scholarship. J.B.M. was supported by a Michael G. DeGroote Fellowship. B.K.C. is the Canada Research Chair in Infectious Disease Pathogenesis. This research was supported by the Canadian Institutes of Health Research (MOP 82704), the Natural Sciences and Engineering Research Council of Canada, and the Canada Foundation for Innovation.

Chapter III – Identification of the regulatory logic controlling *Salmonella* pathoadaptation by the SsrA-SsrB two-component system

Chapter III - Co-authorship statement

Chapter III consists of the following publication:

Tomljenovic, A.M, Mulder, D.T., Whiteside, M.D., Brinkman, F.S.L., Coombes, B.K., (2010) Identification of the regulatory logic controlling *Salmonella* pathoadaptation by the SsrA-SsrB two-component system. PLoS Genet 6(3): e1000875. doi:10.1371/journal.pgen.1000875

The manuscript was written by AMT, BKC and DTM.

The following experiments were performed by collaborators other than myself:

(1) The SsrB DNA microarray experiment was performed by BKC and analyzed by DTM (Figure 3.1).

(2) The SsrB Chromatin Immunoprecipitation on chip (ChIP on chip) experiment was performed by AMT and analyzed by DTM (Figure 3.2).

(3) The SsrB bacterial one-hybrid screen experiment was performed by AMT, and analyzed by MDW and DTM (Figure 3.4).

(4) Alignments with *Sodalis glossinidius* SSR3, integration of data sources and computational identication of the SsrB DNA recognition motif was performed by DTM (Figure 3.4).

(5) SsrB DNA recognition motif sequence scrambling experiment was performed by AMT (Figure 3.6).

(6) Experimental investigation of the *ssaR* promoter was performed by DTM (Figure 3.9).

Identification of the regulatory logic controlling *Salmonella* pathoadaptation by the SsrA-SsrB two-component system

Ana M. Tomljenovic[1, †], David T. Mulder[1, †], Matthew D. Whiteside[2], Fiona S. L. Brinkman[2], Brian K. Coombes[1]

[1] Michael G. DeGroote Institute for Infectious Disease Research and the Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, ON L8N 3Z5

[2] Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, BC, Canada V5A 1S6

[†] These authors contributed equally to this work

Address correspondence to BKC (coombes@mcmaster.ca)

Abstract

Sequence data from the past decade has laid bare the significance of horizontal gene transfer in creating genetic diversity in the bacterial world. Regulatory evolution, in which non-coding DNA is mutated to create new regulatory nodes, also contributes to this diversity to allow niche adaptation and the evolution of pathogenesis. To survive in the host environment, *Salmonella enterica* uses a type III secretion system and effector proteins, which are activated by the SsrA-SsrB two-component system in response to the host environment. To better understand the phenomenon of regulatory evolution in *S. enterica* we defined the SsrB regulon and asked how this transcription factor interacts with the *cis*-regulatory region of target genes. Using ChIP-on-chip, cDNA hybridization and comparative genomics analyses, we describe the SsrB-dependent regulon of ancestral and horizontally acquired genes. Further, we used a genetic screen and computational analyses integrating experimental data from *S. enterica* and sequence data from an orthologous regulatory system in the insect endosymbiont, *Sodalis glossinidius*, to identify the conserved yet flexible palindrome sequence that defines DNA recognition by SsrB. Mutational analysis of a representative promoter validated this palindrome as the minimal architecture needed for regulatory input by SsrB. These data provide a high-resolution map of a regulatory network and the underlying logic enabling pathogen adaptation to a host.

Author Summary

All organisms have a means to control gene expression ensuring correct spatiotemporal deployment of gene products. In bacteria, gene control presents a challenge because one species can reside in multiple niches, requiring them to coordinate gene expression with environmental sensing. Also, widespread acquisition of DNA by horizontal gene transfer demands a mechanism to integrate new genes into existing regulatory circuitry. The environmental awareness issue can be controlled using two-component regulatory systems that connect environmental cues to transcription factor activation, whereas the integration problem can be resolved using DNA regulatory evolution to create new regulatory connections between genes. The evolutionary significance of regulatory evolution for host adaptation is not fully known. We studied the convergence of environmental sensing and genetic networks by examining how the *Salmonella enterica* SsrA-SsrB two-component system, activated in response to host cues, has integrated ancestral and acquired genes into a common regulon. We identified a palindrome as the major element apportioning SsrB on the chromosome. SsrB binding sites have been selected to co-regulate a gene program involved in pathogenic adaptation of *Salmonella* to its host. In addition, our results indicate that promoter architecture emerging from SsrB-dependent regulatory evolution may support both mutualistic and parasitic bacteria-host relationships.

Introduction

Precise gene regulation is crucial to the successful activation and execution of virulence programs for all pathogenic organisms. The acquisition of genes through horizontal gene transfer, a widespread means of bacterial evolution (165), requires a process to integrate new coding sequence into pre-existing regulatory circuitry. Silencing of horizontally-acquired genes by DNA binding proteins like H-NS is one way some incoming genes are initially controlled (30, 31), which can then be subject to regulatory evolution by mutating *cis*-regulatory operator regions to select for optional gene expression. The eventual promoter architecture selected to deploy virulence genes is often modular and should reflect a design that maximizes organismal fitness while limiting fitness trade-offs and antagonistic pleiotropy (166). Both simulated (167) and functional experiments (168, 169) show that mutation of *cis*-regulatory sequences can be rapid, and that plasticity - the degree to which regulatory mutation can perturb the larger gene network - can be well tolerated in bacterial systems.

Promoter architectures that control quantitative traits such as bacterial virulence are, in fact, modular and evolvable (170). For instance, *Salmonella enterica* has a multi-faceted pathogenic strategy fine-tuned by several transcriptional regulators. Intracellular survival and persistence of *Salmonella* requires a type III secretion system (T3SS) encoded in a horizontally-acquired genomic island called *Salmonella* Pathogenicity Island-2 (SPI-2) (110, 171). T3SS are complex secretion machines that deliver bacterial effector proteins directly into host cells through an injectisome during infection (172, 173). Several ancestral regulators control the genes in the SPI-2 genomic island including

the two-component systems EnvZ-OmpR and PhoQ-PhoP, and the regulatory protein SlyA (93). SsrA-SsrB is another two-component regulatory system co-inherited by genetic linkage with the SPI-2 locus that is essential for gene expression in SPI-2 (113, 174, 175). SsrA is a sensor kinase activated in the host environment that phosphorylates the SsrB response regulator to create an active transcription factor needed for spatiotemporal control of virulence genes (93, 95). In the *Salmonella* genus, the SPI-2 genomic island is found only in pathogenic serotypes of *Salmonella enterica* that infect warm-blooded animals and is absent from *Salmonella bongori*, which colonizes cold-blooded animals (176). It is generally accepted that SPI-2 was acquired by *Salmonella enterica* after divergence from *S. bongori*, providing a useful pedigree to study regulatory evolution influenced by SsrB.

We recently demonstrated the evolutionary significance of *cis*-regulatory mutations for pathoadaptation of *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium) to an animal host (33). Our focus was on SsrB because of its broad conservation among the pathogenic Salmonellae and its essentiality for animal infection, suggesting it coordinates fundamental aspects of *Salmonella* pathogenesis beyond the SPI-2 genomic island. In this study we investigated how regulatory evolution assimilates horizontally acquired and ancestral genes into the SsrB regulon on a genome-wide scale using an integrated set of experimental methods. Combining our data with previous biochemical work, along with comparative genomic analyses with an orthologous T3SS-encoding genomic island in the tsetse fly endosymbiont, *Sodalis glossinidius*, we reveal the flexible DNA palindrome that distributes SsrB in the genome to influence

transcriptional activation of the SPI-2 T3SS and almost all of its accessory effector proteins. Our data uncovers the principal SsrB circuitry that appears to have been conserved to support multiple bacterial lifestyles, including parasitic and mutualist symbioses.

Results

Transcriptional profiling of an *ssrB* mutant

To begin to understand regulatory evolution and network expansion of the SsrB response regulator, we profiled the transcriptome of an *ssrB* mutant and compared it to *S.* Typhimurium wild type cells grown in an acidic minimal medium that activates the SsrA-SsrB two-component regulatory system (177). We identified 133 genes that were significantly down-regulated in the *ssrB* mutant [$z < -1.96$] (178) (Table 3.1). This included almost all genes in the SPI-2 genomic island as well as effector genes encoded throughout the genome. Next, we performed a Clusters of Orthologous Groups (COG) analysis (179) on the 118 genes that had an ortholog in the annotated genome of *S.* Typhimurium strain LT2 (15). Among these, 45 genes lacked a functional COG assignment and the 73 remaining genes were distributed among 86 COGs (Figure 3.1). The majority of functions in the latter groups are in transport, secretion, and trafficking of cellular components in addition to protein and membrane modification.

The SsrA-SsrB system was acquired by horizontal gene transfer into the *S. enterica* species after divergence from what is now extant *S. bongori*. As such, *S. bongori* has evolved in the absence of SsrA-SsrB and its regulatory architecture has not been influenced by it. Orthologous genes ancestral to both species but regulated by SsrB in *S. enterica* provide evidence for network expansion and regulatory evolution that we previously showed can be mapped to a single *cis*-input location by using functional and comparative genomics (33). To expand on this, we used a reciprocal BLAST-based analysis and identified 47 orthologs in *S. bongori* among the 133 genes whose

transcription was down-regulated in an *ssrB* mutant (Table 3.1). In Δ*ssrB* cells, the mean fold-change of the orthologous genes (-6.1-fold) was subtler than for the *S. enterica*-specific gene set (mean -21.2-fold), which included the T3SS and associated effector genes. We also determined the distribution of down-regulated genes among genomic islands (180), including prophages, pathogenicity islands (SPI-islands) and additional regions of difference (ROD) between *S. enterica* and *S. bongori*. For this we used a BLAST-based comparison of genome-wide synteny between *S.* Typhimurium and *S. bongori* and identified 50 ROD that included 17 previously reported SPI-islands and prophages. Of the 133 down-regulated genes identified, 56 were present within genomic islands (Table 3.1), with a mean change in gene expression of -25.3-fold in Δ*ssrB* cells.

Table 3.1. Summary of microarray and ChIP-on-chip data

Microarray	
Total significant SsrB-activated genes, *n*	133
Genes with *S. bongori* ortholog, *n*	47
Mean fold change	-15.8 (-6.0*)
Median fold change	-6.5 (-4.5*)
Genes within pathogenicity islands, *n*	45
Genes with *S. bongori* ortholog, *n*	2
Mean fold change	-27.4 (-6.5*)
Median fold change	-13.1 (-6.5*)
Genes within phage islands, *n*	4
Genes with *S. bongori* ortholog, *n*	0
Mean fold change	-20.4 (n/a*)
Median fold change	-9.0 (n/a*)
Genes within *S. bongori* ROD, *n*	7
Genes with *S. bongori* ortholog, *n*	0
Mean fold change	-14.6 (n/a*)
Median fold change	-7.1 (n/a*)
Genes not in genomic islands, *n*	77 [a]
Genes with *S. bongori* ortholog, *n*	45
Mean fold change	-8.9 (-6.0*)
Median fold change	-4.4 (-4.4*)
ChIP-on-chip	
Probes, *n*	42,021
Probe size	60-bp
Average probe separation	57-bp
Genome probe average $\log_2$ signal	-0.02
Standard deviation	0.52
Peaks above 3 st. dev., *n*	256
Peaks within published islands, *n*	47
Peaks within *S. bongori* ROD, *n*	15
Within CDS, *n*	126
Within/Overlapping IGR, *n*	130

n, number

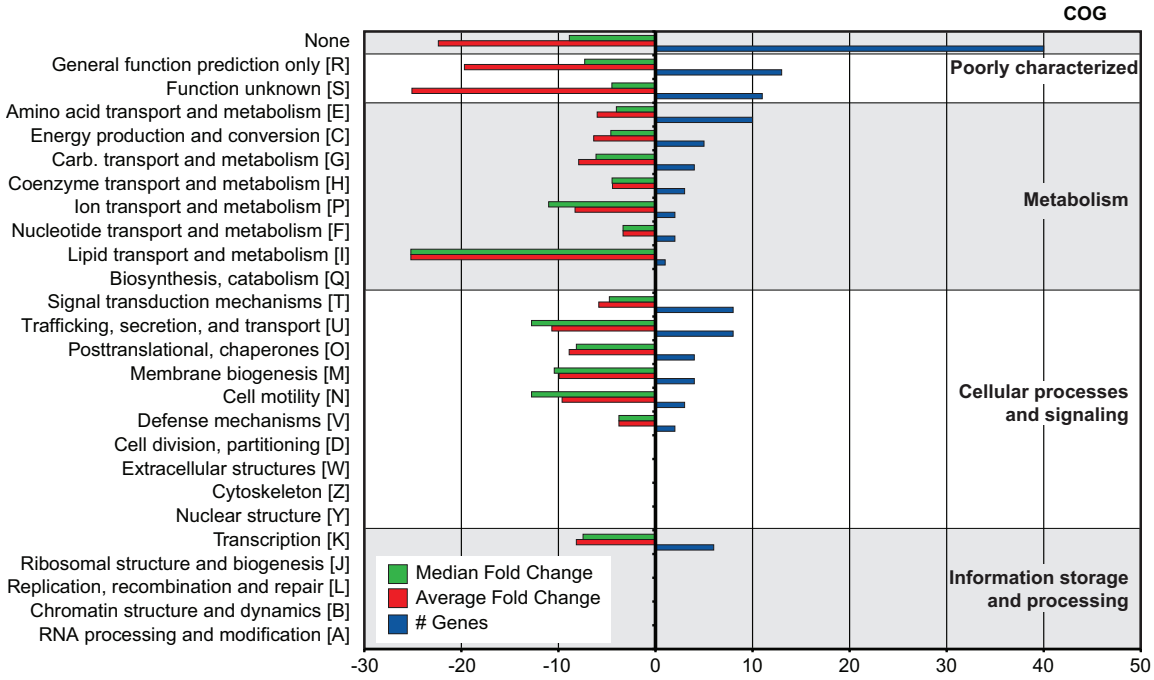
*, (value for genes with an ortholog in *S. bongori*)

[a], includes 21 genes on the microarray that could not be mapped because they were not annotated by the array manufacturer (TIGR).

Figure 3.1. **COG analysis of 133 genes co-regulated with SPI-2.**

COG categories are indicated to the right of the figure and COG sub-categories are ordered in terms of decreasing gene representation (blue) within each category. Median and average fold change values for each sub-category are indicated in green and red, respectively.

Figure 3.1.


COG
None
General function prediction only [R]
Function unknown [S]
Amino acid transport and metabolism [E]
Energy production and conversion [C]
Carb. transport and metabolism [G]
Coenzyme transport and metabolism [H]
Ion transport and metabolism [P]
Nucleotide transport and metabolism [F]
Lipid transport and metabolism [I]
Biosynthesis, catabolism [Q]
Signal transduction mechanisms [T]
Trafficking, secretion, and transport [U]
Posttranslational, chaperones [O]
Membrane biogenesis [M]
Cell motility [N]
Defense mechanisms [V]
Cell division, partitioning [D]
Extracellular structures [W]
Cytoskeleton [Z]
Nuclear structure [Y]
Transcription [K]
Ribosomal structure and biogenesis [J]
Replication, recombination and repair [L]
Chromatin structure and dynamics [B]
RNA processing and modification [A]
Median Fold Change
Average Fold Change
Genes
Poorly characterized
Metabolism
Cellular processes and signaling
Information storage and processing
-30
-20
-10
0
10
20
30
40
50

Genome-wide SsrB interactions

To examine SsrB allocation on the chromosome *in vivo*, we isolated functional SsrB-DNA interactions using chromatin immunoprecipitation and examined the bound DNA by chip analysis (ChIP-on-chip) using an *S.* Typhimurium SL1344 array containing 44,021 probes. With this method, we identified 256 significant interaction peaks distributed throughout the genome that were enriched under SsrB-activating conditions and with interaction scores three standard deviations greater than the mean probe score (Figure 3.2A and Table 3.1). Of these 256 peaks, 126 (49%) occurred within coding regions of genes (CDS) and 130 (51%) were in intergenic regions (IGR). Given the strong influence of SsrB on horizontally acquired genes (Table 3.1), we plotted the ChIP-on-chip data against all genomic islands in *S.* Typhimurium SL1344. From this analysis, 62 of 256 SsrB binding peaks (24.3%) occurred within genomic islands (Figure 3.2B).

SsrB ChIP peaks were observed upstream of previously identified SsrB regulated genes indicating that our ChIP-on-chip data captured functional interactions. To generate a consensus set of SsrB regulated genes, we performed an analysis to identify operons in the *S.* Typhimurium SL1344 genome that encoded at least one gene down-regulated in Δ*ssrB* cells and that possessed an SsrB binding peak in the upstream regulatory region as defined by our ChIP-on-chip analysis. From this, the 133 down-regulated genes mapped to 86 operons, 49 of which had an SsrB interaction upstream or within the first gene of the operon (Table 3.2). This analysis captured all five reported operons in the SPI-2 genomic island in addition to ten operons outside of this island that encode SPI-2 translocated effectors.

Table 3.2. Predicted operons encoding SPI-2 associated genes

Gene Range	Operon Size	Fold Change	First Gene	LT2 Ortholog	Annotation	Probe Location[a]	ChIP Score[b]
SL2601-SL2604	4	-1.45	SL2604	STM2640:*rpoE*		CDS	4.12
SL0991-SL0991	1	-57.81	SL0991	STM1051:*sseI*	Effector	IGR	3.37
SL2256-SL2256	1	-43.50	SL2256	STM2287:*sseL*	Effector	IGR	3.02
SL2217-SL2217	1	-7.07	SL2217	STM2241:*sspH2*	Effector	CDS	2.95
SL1561-SL1561	1	-25.20	SL1561	STM1631:*sseJ*	Effector	CDS	2.82
SL1628-SL1628	1	-51.60	SL1628	STM1698:*steC*	Effector	5'-IGR-overlap	2.48
SL1574-SL1576	3	-3.41	SL1576	STM1646:*ydbH*		CDS	2.40
SL1325-SL1326	2	-9.64	SL1326	STM1392:*ssrA*	SPI-2 gene	IGR	2.31
SL1327-SL1330	4	-18.09	SL1327	STM1393:*ssaB*	SPI-2 gene	IGR	2.31
SL1161-SL1161	1	-17.23	SL1161	STM1224:*sifA*	Effector	IGR	2.26
SL1340-SL1340	1	-10.51	SL1340	STM1406:*ssaG*	SPI-2 gene	CDS	1.93
SL0909-SL0909	1	-15.47	SL0909	STM0972:*sopD2*	Effector	5'-IGR-overlap	1.87
SL1027-SL1027	1	-33.97	SL1027	STM1088:*pipB*	Effector	IGR	1.85
SL1331-SL1336	6	-49.46	SL1331	STM1397:*sseA*	SPI-2 gene	IGR	1.81
SL1353-SL1356	4	-12.90	SL1353	STM1419:*ssaR*	SPI-2 gene	5'-IGR-overlap	1.78
SL1347-SL1352	6	-16.59	SL1347	STM1413:*ssaM*	SPI-2 gene	5'-IGR-overlap	1.76
SL1563-SL1566	4	-3.10	SL1563	STM1633		5'-IGR-overlap	1.76
SL2763-SL2763	1	-195.60	SL2763	STM2780:*pipB2*	Effector	IGR	1.75
SL0083-SL0083	1	-8.71	SL0083	STM0082:*srfN*		IGR	1.54
SL1908-SL1910	3	-2.68	SL1908	STM1979:*fliP*		3'-IGR-overlap	1.35
SL2001-SL2011	11	0.66	SL2011	STM2035:*cbiA*		IGR	1.30
SL4242-SL4245	4	-0.98	SL4242	STM4305.S		IGR	1.28
SL0700-SL0700	1	-7.06	SL0700	STM0719		CDS	1.18
SL3259-SL3261	3	-1.29	SL3261	STM3288:*yhbC*		IGR	1.13
SL1992-SL2000	9	-1.57	SL2000	STM2024:*cbiL*		CDS	0.96
SL2114-SL2114	1	-19.35	SL2114	STM2138:*srcA*	Effector	CDS	0.94
SL1026-SL1026	1	-2.64	SL1026	STM1087:*pipA*	Effector	IGR	0.90
SL1159-SL1160	2	-5.29	SL1160	STM1223:*potC*		CDS	0.86
SL1341-SL1346	6	-42.72	SL1341	STM1407:*ssaH*	SPI-2 gene	3'-IGR-overlap	0.84
SL1785-	1	-11.81	SL1785	STM1856		IGR	0.81

SL1785							
SL0037-SL0038	2	-4.00	SL0037	STM0036		5'-IGR-overlap	0.78
SL1170-SL1171	2	-2.84	SL1171	STM1233:*ycfC*		5'-IGR-overlap	0.74
SL1337-SL1339	3	-8.51	SL1337	STM1403:*sscB*	SPI-2 gene	3'-IGR-overlap	0.72
SL1878-SL1878	1	-7.65	SL1878	STM1949:*yecF*		IGR	0.72
SL1705-SL1705	1	-6.08	SL1705	STM1777:*hemA*		5'-IGR-overlap	0.69
SL2193-SL2196	4	-2.43	SL2193	STM2216:*yejA*		CDS	0.68
SL4478-SL4479	2	-1.83	SL4478	STM4547:*yjjQ*		IGR	0.67
SL1763-SL1763	1	-21.12	SL1763	STM1834:*yebN*		IGR	0.62
SL0758-SL0760	3	-6.10	SL0758	STM0781:*modA*		5'-IGR-overlap	0.60
SL1200-SL1200	1	-10.18	SL1200	STM1265		CDS	0.58
SL1238-SL1242	5	-1.54	SL1238	STM1303:*argD*		CDS	0.56
SL1958-SL1961	4	-26.21	SL1961	-		IGR	0.56
SL2210-SL2211	2	-19.29	SL2211	STM2235		CDS	0.56
SL0701-SL0703	3	7.94	SL0701	STM0719		5'-IGR-overlap	0.55
SL0547-SL0549	3	-5.97	SL0549	STM0559:*rfbI*		IGR	0.55
SL0393-SL0393	1	-3.20	SL0393	STM0398:*phoR*		CDS	0.53
SL1689-SL1692	4	0.78	SL1692	STM1764:*narG*		IGR	0.53
SL0160-SL0160	1	-3.75	SL0160	STM0159		IGR	0.52
SL1373-SL1375	3	-1.29	SL1375	STM1443:*ydhI*		IGR	0.52

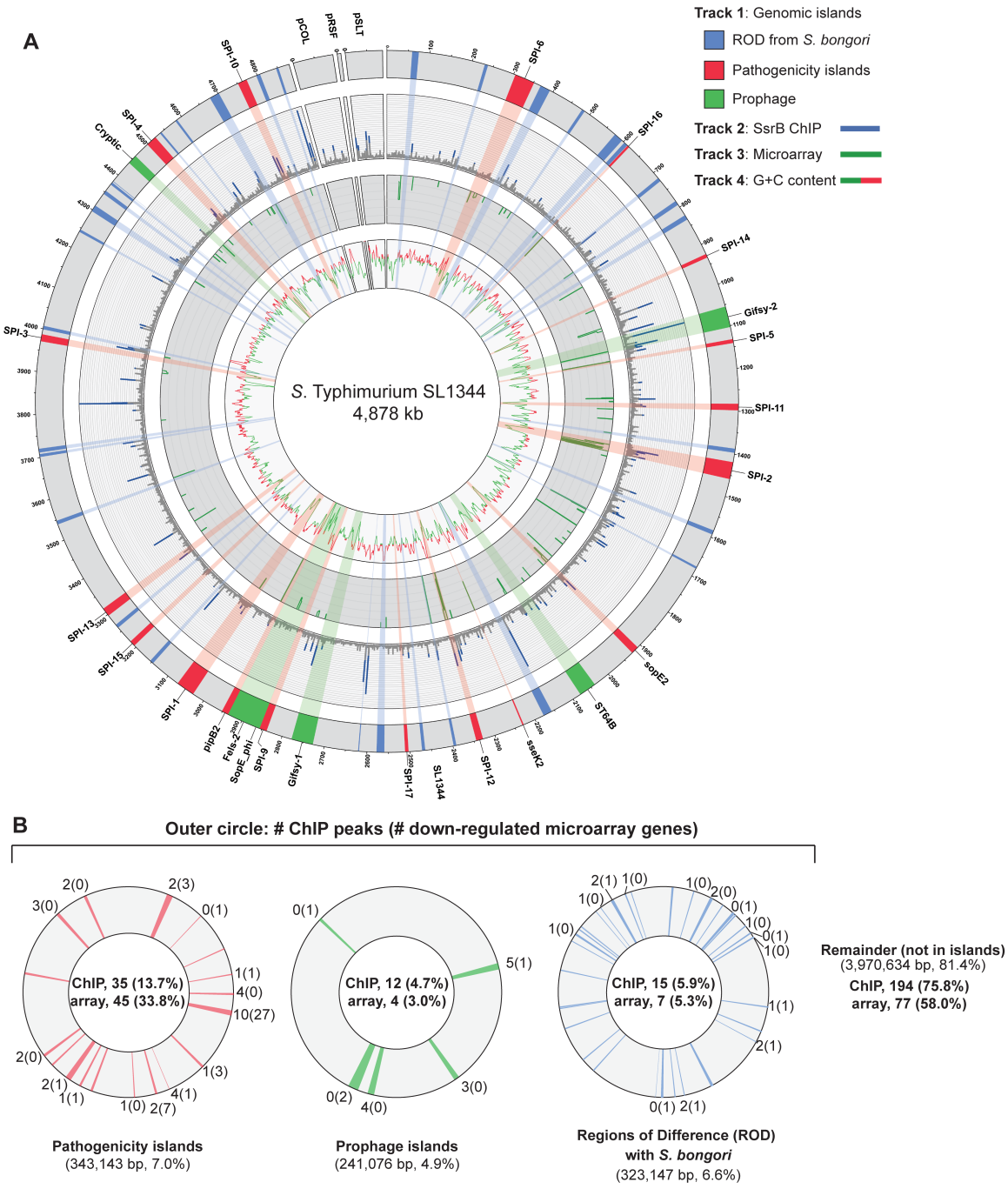
a, CDS (within coding sequence); IGR (in intergenic region); 5'-IGR overlap (probe overlaps IGR and 5' region of CDS)

b, $\log_2$

Figure 3.2. **Genome-wide functional genomics analyses for SsrB.**

(A) ChIP-on-chip and microarray analysis. Genomic islands are indicated and labelled on track 1 (outermost circle), where prophages are in green, pathogenicity islands are in red and regions of difference (ROD) with *S. bongori* are in blue. ChIP-on-chip interaction scores (log_2) for SsrB-activating conditions are plotted on track 2, where peaks with signals greater than one standard deviation from the mean probe signal are indicated in blue. Transcriptional profiling analysis of SsrB-regulated genes is plotted on track 3. Genes repressed more than 5-fold in an *ssrB* mutant are indicated in green. G+C content is plotted on track 4, where values greater than 53% (genome mean) are indicated in red and less than 53% are indicated in green. (B) Summary of gene expression data and SsrB binding data for pathogenicity islands, prophage islands, *S. bongori* ROD and genome sequence not in islands. The number of SsrB binding peaks identified by ChIP-on-chip for each of the individual islands is shown in the outermost track, along with the number of genes down-regulated in the transcriptional profiling analysis (in brackets). The total number of ChIP-on-chip interaction peaks and genes regulated by SsrB are shown in the innermost circles along with the percent of total (in brackets). The total genomic content of each island or conserved regions are shown below each circle, along with the percent of total genome. The remainder of microarray data (not in islands) includes 194 interaction peaks (75.8%) for ChIP-on-chip and 77 genes (58.0%) for transcriptional profiling arrays.

Figure 3.2.


A
Track 1: Genomic islands
ROD from S. bongori
Pathogenicity islands
Prophage
Track 2: SsrB ChIP
Track 3: Microarray
Track 4: G+C content
S. Typhimurium SL1344
4,878 kb
pCOL
pRSF
pSLT
SPI-6
SPI-16
SPI-14
Gifsy-2
SPI-5
SPI-11
SPI-2
sopE2
STM4AB
sseK2
SPI-12
SL1344
SPI-17
Gifsy-1
SPI-9
SopE_phi
Fels-2
pipB2
SPI-1
SPI-15
SPI-13
SPI-3
Cryptic
SPI-4
SPI-10
B
Outer circle: # ChIP peaks (# down-regulated microarray genes)
ChIP, 35 (13.7%)
array, 45 (33.8%)
Pathogenicity islands
(343,143 bp, 7.0%)
ChIP, 12 (4.7%)
array, 4 (3.0%)
Prophage islands
(241,076 bp, 4.9%)
ChIP, 15 (5.9%)
array, 7 (5.3%)
Regions of Difference (ROD)
with S. bongori
(323,147 bp, 6.6%)
Remainder (not in islands)
(3,970,634 bp, 81.4%)
ChIP, 194 (75.8%)
array, 77 (58.0%)

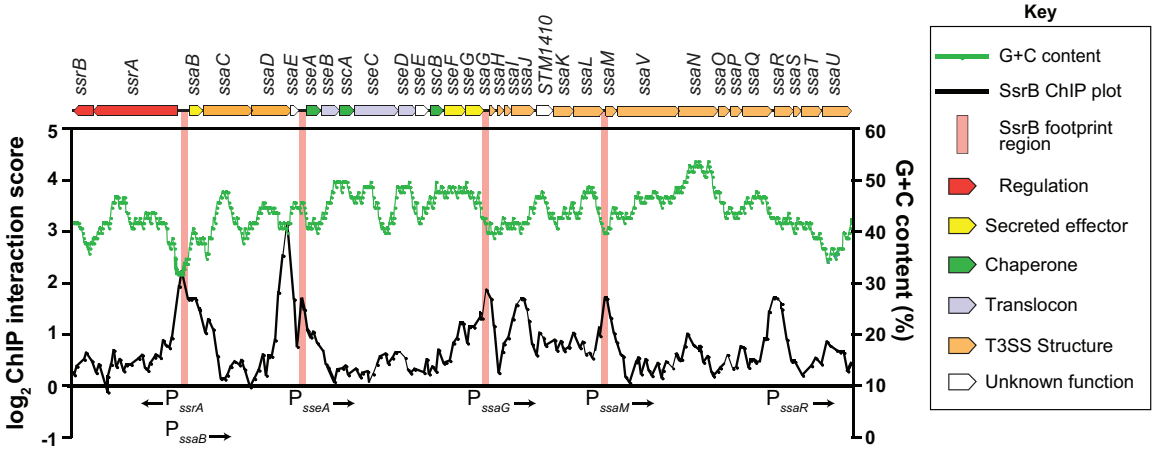
SsrB binds to six promoters within the SPI-2 genomic island

In order to rigorously evaluate our genome-wide functional genomics data, we compared it against traditional biochemical experiments describing SsrB-DNA interactions at the SPI-2 locus. Previous data reported SsrB footprints upstream of 6 genes in SPI-2: *ssrA, ssrB, ssaB, sseA, ssaG,* and *ssaM* (100, 109). Our ChIP-on-chip data showed discrete SsrB binding at all of these promoters except for the promoter reported to be between *ssrA* and *ssrB* (109) (Figure 3.3). We attempted to verify functional activity at this site, but could not using transcriptional fusions (data not shown). Our data also identified three additional SsrB binding peaks upstream of *sseA*, within the CDS of *ssaJ*, and in the IGR upstream of *ssaR* (Figure 3.3). Functional interactions were confirmed for *sseA* and *ssaR* in subsequent reporter experiments described below.

Figure 3.3. **The SsrB binding profile at the SPI-2 locus**.

Plotted are the SsrB ChIP interaction scores (black trace) for all probes within the SPI-2 locus. ChIP-on-chip experiments were performed with *S*. Typhimurium grown in SPI-2 inducing (LPM) and non-inducing (LB) conditions. ChIP score was calculated as LB (control/experimental) / LPM(control/experimental) where LB scores were used to normalize the data. Regions of SPI-2 footprinted by SsrB are highlighted with vertical pink bars. G+C content across the region is plotted as a green trace. Genes in the SPI-2 genomic island (top) are plotted to scale and are coloured according to function (see key).

Figure 3.3.


ssrB
ssrA
ssaB
ssaC
ssaD
ssaE
sseA
sseB
sscA
sseC
sseD
sseE
sscB
sseF
sseG
ssaG
ssaH
ssaJ
STM1410
ssaK
ssaL
ssaM
ssaV
ssaN
ssaO
ssaP
ssaQ
ssaR
ssaS
ssaT
ssaU
log2 ChIP interaction score
G+C content (%)
P ssrA
P ssaB
P sseA
P ssaG
P ssaM
P ssaR
Key
G+C content
SsrB ChIP plot
SsrB footprint region
Regulation
Secreted effector
Chaperone
Translocon
T3SS Structure
Unknown function

Identification of a functional DNA element for SsrB binding

Previous attempts by others to identify a conserved SsrB DNA recognition motif have been unsuccessful. To overcome this, we employed a bacterial one-hybrid screen originally developed to define binding site preferences for eukaryotic transcription factors (181). We fused the DNA binding domain of SsrB (SsrBc) to the α-subunit of RNA polymerase and screened a prey library of ~10^8 DNA molecules previously counter-selected against self-activation (Figure 3.4A). We used the PhoP response regulator from *E. coli* as a control because a DNA recognition sequence for it was known (182). Bait-prey combinations surviving selection on medium lacking histidine were purified, and preys were sequenced and analyzed using the motif-finding program MEME (183). From 189 unique sequences isolated for SsrBc, over 80% contained a degenerate consensus motif, mCCyTA (Figure 3.4B). In control screens with the PhoP-αNTD fusion, the PhoP box sequence (G/T)GTTTA was identified in 11% of sequenced preys (12/109, data not shown) but this sequence was never captured by SsrB-αNTD and vice versa, demonstrating specificity of the bacterial one-hybrid system for prokaryotic regulatory proteins.

Next, we examined our ChIP-on-chip data for the presence of a conserved regulatory motif. We extracted sequence data from the local maxima of the 256 binding peaks and analyzed the sequences with the computational program MDscan (184). Using the highest-ranking probes to generate an initial prediction followed by lower-ranking probes for refinement, this analysis identified motifs that represented either the forward or the reverse complement of the consensus sequence ACmTTA, which shares consensus

with the motif identified in the bacterial one-hybrid screen (Figure 3.4B). We identified variations of this motif within footprinted regions of SsrB-regulated promoters (100, 109), however sequence degeneracy made it difficult to precisely map the input functions.

Identification of a functional DNA palindrome conserved in *S.* Typhimurium SPI-2 and the *Sodalis glossinidius* SSR-3 region

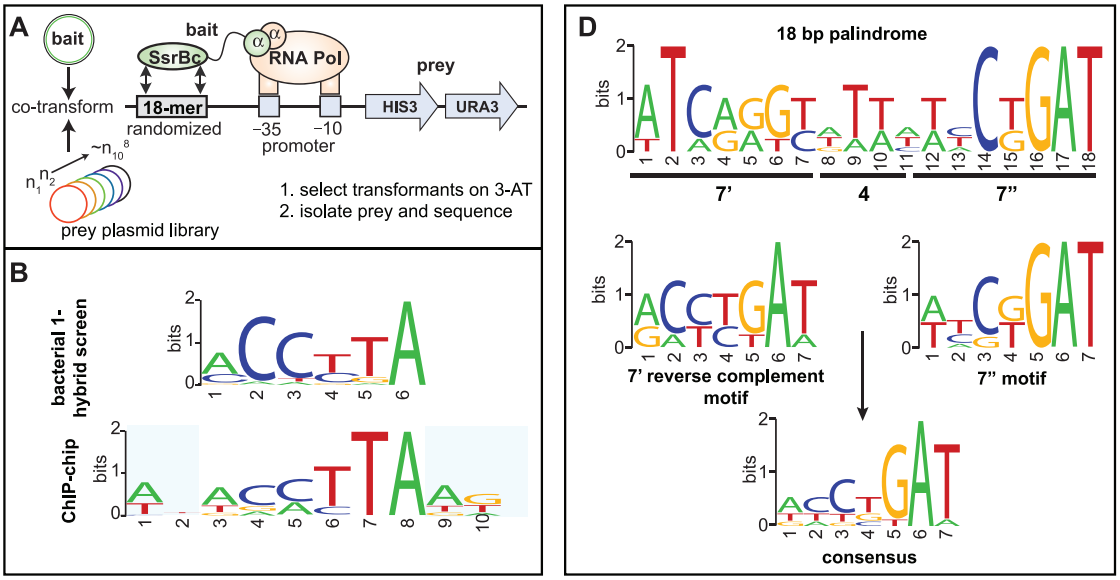
The analysis of regulatory evolution is particularly challenging because it is difficult to distinguish neutral stochastic change from functional divergence. To solve this problem in the context of mapping the SsrB binding element, we used comparative genomics to search for conserved promoter architecture in another organism with a similar genomic island to *Salmonella* SPI-2. The tsetse fly endosymbiont *Sodalis glossinidius* contains the *Sodalis* Symbiosis Region-3 (SSR-3) that is similar in content and synteny with the *S. enterica* SPI-2 locus (185). Gene conservation includes the entire T3SS structural module extending to the regulatory genes *ssrA-ssrB* and all other genes except the effectors *sseF* and *sseG*, which are not present in *Sodalis* SSR-3. We aligned the sequences of the five mapped promoters in SPI-2 with the orthologous SSR-3 regions to identify local conservation. Highly conserved sites within the promoters were restricted to regions previously footprinted by SsrB (100, 109), whereas adjacent sequence showed substantial drift (Figure 3.4C). Within the conserved sites we identified a heptameric sequence in 7-4-7 tail-to-tail architecture that created an 18-bp degenerate palindrome. This palindrome was found in all SPI-2 and SSR-3 T3SS promoters with the exception of the *sseA* promoter that had only one reasonably well-conserved heptamer in the

footprinted region (Figure 3.4C and Figure 3.5). Interestingly, two copies of the palindrome occur upstream of the *ssrA-ssrB* operon in *S.* Typhimurium within the same footprint, and the conservation of either site in *Sodalis* was weak. Evaluation of the heptamer motif in the palindrome showed high similarity to the motifs identified by the bacterial one-hybrid screen and the ChIP-on-chip experiments (Figure 3.4B and 4D), giving us confidence that we had identified the major recognition module for transcriptional input by SsrB. In accord with a previous observation (100), there was not a strict requirement in the spacing between the SsrB binding site and the downstream transcriptional start site.

Figure 3.4. **Identification of a conserved palindrome in *S.* Typhimurium SPI-2 and *S. glossinidius* SSR-3 promoters.**

(A) Schematic representation of the bacterial one-hybrid screen for functional SsrB binding sequences. (B) Motif logos generated independently from the bacterial one-hybrid screen and ChIP-on-chip data show similar but degenerate consensus along 6 nucleotides. (C) Aligned promoter regions from the *S.* Typhimurium SPI-2 island and *S. glossinidius* SSR-3 island containing the putative SsrB DNA recognition palindrome. Coordinates indicate position of the conserved palindrome with respect to the translational start site of the downstream gene. The 7'-4-7" tail-to-tail palindrome sequences are underlined in blue (7' site) and red (7" site), aligned and highlighted in yellow. Asterisks indicate conserved nucleotides between *S.* Typhimurium and *S. glossinidius*. The positions of SsrB footprints protected from DNase1 (100, 109) are shown as black text highlighted with a blue horizontal bar. (D) Consensus motif logos for the palindrome sequences from the SPI-2 promoters shown in C.

Figure 3.4.



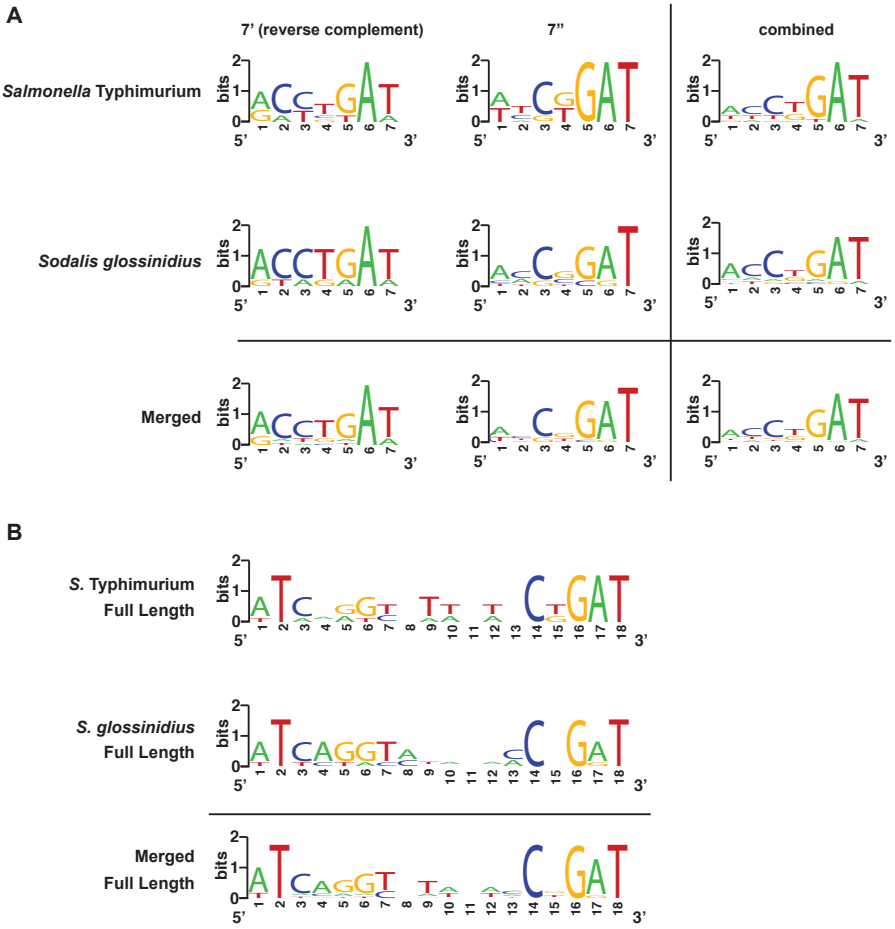
C

Organism	Promoter	Coord.	Sequence alignment (7' 4 7")
S. Typhimurium	*ssrA*	-70..-87	ATTTATCTT-CCTAATGATAAC-ACCATCGATTAATCTTCTGATGAAACTATATGTACTGCGATAGTGATCAAGTGCCAAA
S. glossinidius	*SG1280*	-61..-78	ATACATAAAGCCTCGTGTCATCCATTATTCTATAGCCAACCGGT--AATTTTAAAAAATG-GACATAGCCTGAGTAGCCAT
S. Typhimurium	*ssaB*	-53..-36	AAGAGTATAGTCGAAATAAATGTG-AATCAGGCTTTTTACGGATGTGGTTGTGAGCGAATTTGATAGAAACTCCC--ATTT
S. glossinidius	*SG1281*	-72..-55	ATGTGTCCAATTCCACCCCATGGGTAATCAGGTATTTTACTGATGTACCCTGCTGGGTTTTCAACTGAGTATGACGGAGTT
S. Typhimurium	*sseA*	-65..-48	ACTGCCCAGGTTCAAAATGGCAAGTAAAATAGCCCTAATGGGATAGGCTCTTAGT----TAGCACGTTAATTATCTATCGT
S. glossinidius	*SG1285*	-123..-106	GGCGC-----TACAGGCTG------AAGATATTCTTAATGGCATTTACTTTTGTTATCACAATGCCTCATTGAGTTGCCGT
S. Typhimurium	*ssaG*	-88..-71	AACAACGTGCGCCGGAGTAATCGTT-TTCAGGTATATACCGGATGTTCATTGCTTTCTAAATTTTGCTATGTTGCCAGTAT
S. glossinidius	*SG1292*	-121..-104	GCTAACGT-CGCCGGGCCGGCTGGAATTCAGGTCTACACCGGATGCTGGCAGCGCTGCGG-TGCGGGGATAATCCGCATAT
S. Typhimurium	*ssaM*	-74..-57	CGAAACGCTTCGTGAAGTAAAGATAAATCAGGTTTTATTCTGATACCTGGCTTTCA-ATATTTAGGTAAATTGGCTTTCTG
S. glossinidius	*SG1299*	-74..-57	CGACGTATTGCGGCAGCTGGCGGCAAATCAGGTATAGCCCCGATACCCGCGTCCCGCACCATTTGGGACGCTGTCTGCACT

Figure 3.5. ***S.* Typhimurium SL1344 SPI-2 and *S. glossinidius str. 'morsitans'* SSR-3 motifs.**

Sequence logos of the half-site (7' and 7") and full-length (7'-4-7") palindrome motif identified within the promoter regions of SPI-2 in *S.* Typhimurium and SSR-3 T3SS in *S. glossinidius*. 'Combined' refers to the consensus sequence based on the single left and right heptamer sequences within a given organism. 'Merged' refers to the consensus sequence for individual left and right heptamers from both *Salmonella* and *Sodalis*.

Figure 3.5.


A
7' (reverse complement)
7"
combined
Salmonella Typhimurium
Sodalis glossinidius
Merged
bits
5'
3'
B
S. Typhimurium
Full Length
S. glossinidius
Full Length
Merged
Full Length

SsrB directs transcriptional input using a flexible palindrome architecture

The presence of a conserved palindrome sequence in SPI-2 promoters and in related sequences from the endosymbiont *S. glossinidius* suggested that regulatory input by SsrB was through a palindrome sequence architecture. However, other lines of evidence suggested that the recognition site architecture was flexible in nature: (*i*) our bacterial one-hybrid screen isolated functional single hexamer sequences, (*ii*) the SsrB footprint at the naturally evolved *sseA* promoter within SPI-2 (100) contained only one reasonably well-conserved heptamer, and (*iii*) degenerate or non-ideal palindromes exist in the genome. In order to deconstruct this architecture, we designed a set of experiments to test the palindrome's tolerance to mutation. We chose the *ssaG* promoter for these experiments because 16 of 18 bases were identical between SPI-2 and SSR-3 from *S. glossinidius*, differing only in the 4-bp spacer between heptamers (Figure 3.4C). We mutated the palindrome in a series of transcriptional reporters that were otherwise identical to the evolved *ssaG* promoter (Figure 3.6A) and promoter activity was compared to that of a wild type palindrome sequence. Variants in which the first half-site (7') or second half-site (7") was deleted produced similar transcriptional activity to the wild type palindrome, verifying that a single well-conserved heptamer is sufficient for transcriptional input under these experimental conditions (Figure 3.6B). Deletion of the 4-bp spacer sequence between the heptamers - the most degenerate element of the palindrome - also generated wild type promoter activity. However, the orientation of individual heptamers was essential for transcriptional input since rewiring the palindrome in any head-to-head orientation produced negligible promoter activity. However, if the

two half sites were swapped front-to-back so that they maintained tail-to-tail orientation (construct labelled “Reverse” in Figure 3.6), wild type promoter activity was restored. Precise deletion of the entire 18-bp palindrome lead to ~10% residual activity in wild type cells, which was reduced to less than 1% in an *ssrB* mutant (Figure 3.6C). To determine whether the remaining 10% transcriptional activity was a result of an SsrB-dependent feed-forward mechanism or transcriptional read-through of our chromosomally integrated reporter, we constructed an ectopic deletion reporter. Assessment of reporter activity for this construct in addition to wild type constructs in wild type and *ssrB* mutant backgrounds showed a similar level of activity to the *ssrB* mutant (Figure 3.6D).

The results for the half-site deletion constructs, which retained activity similar to wild type, were unexpected. Therefore, we compared the sequences generated upon mutation against a consensus palindrome matrix generated from all SPI-2 and other identified putative elements. The 7’-4-X, X-4-7” and 7’-X-7” mutations introduced a number of base transitions and transversions never occurring in the matrix, however the modified 7 base pair heptamer retained 4-5 naturally-occurring bases along with the unchanged wild type sequence in the other heptamer (Figure 3.7). The possibility existed that this modified heptamer, although now weaker in consensus, could still be sufficient for recruitment of a functional form of SsrB when paired with the other wild type heptamer. To test this, we created an additional ectopic transcriptional fusion construct in which the left half (7’) of the palindrome was mutated to bases never occurring in the consensus matrix. When tested in promoter activity experiments, this reporter was unable to activate transcription and was identical to the X-X-X mutant construct (Figure 3.6D).

Salmonella SsrB and the *Sodalis* ortholog (*SG1279*) are 69% identical and 81% similar at the amino acid level. All of the critical residues in the dimerization helix and HTH motif required for specific transcriptional activity by SsrB [32] are conserved in the *Sodalis* ortholog (Figure 3.8). To demonstrate a functional role for the palindrome identified in *Sodalis*, we engineered luciferase transcriptional reporters that either contained (7'-4-7") or lacked (X-X-X) the identified palindrome from the *Sodalis SG1292* promoter (*ssaG* ortholog) and transformed them into wild type *S.* Typhimurium and an *ssrB* mutant. The transcriptional activity from a wild type *Sodalis* palindrome sequence was high, but was completely abolished in an *ssrB* mutant and in experiments where only the palindrome sequence was precisely deleted (Figure 3.6D). These experiments demonstrated a functional role for the conserved palindrome in *Sodalis* and the requirement for SsrB for transcriptional activation.

Figure 3.6. **The SsrB binding motif architecture is flexible and conserved**.

(A) The sequences of the wild type SsrB binding motif and mutated permutations from the *ssaG* promoter are shown. The orientation of the 7-bp half sites are indicated, where the blue arrow represents the 7' site and the red arrow represents the 7" site as identified in Figure 4. The orientation of the 4-bp spacer sequence is represented as a black arrow. The colours of the bases are used to distinguish the different deoxyribonucleotides. (B) Transcriptional reporter data for the wild type *cis*-regulatory input (wt) and the seven permutations as defined in A. Promoter activity is shown at 5 h as measured by β-galactosidase assays normalized to promoter activity from the wild type reporter. (C) Comparison of *ssaG* promoter activity between wild type *S*. Typhimurium and an *ssrB* mutant. β-galactosidase assays were conducted for the P*ssaG* X-X-X reporter in both wt and Δ*ssrB* cells. Data represents the promoter activity as a percent of wild type. (D) The conserved palindrome sequence from *Sodalis glossinidius* is functional. Transcriptional reporters that either contained (7'-4-7'') or lacked (X-X-X) the 18-bp palindrome sequence were constructed from evolved orthologous promoters in *Salmonella* (P*ssaG*) and *Sodalis* (P*SG1292*). An additional reporter that mutated the bases in the left heptamer to those never occurring in the consensus matrix (scrambled) was also constructed. Reporters were transformed into wild type *Salmonella* and an *ssrB* mutant and tested for transcriptional activity. Luminescence data was normalized to the OD of the culture. Shown are the means with standard deviation for three separate experiments.

Figure 3.6.

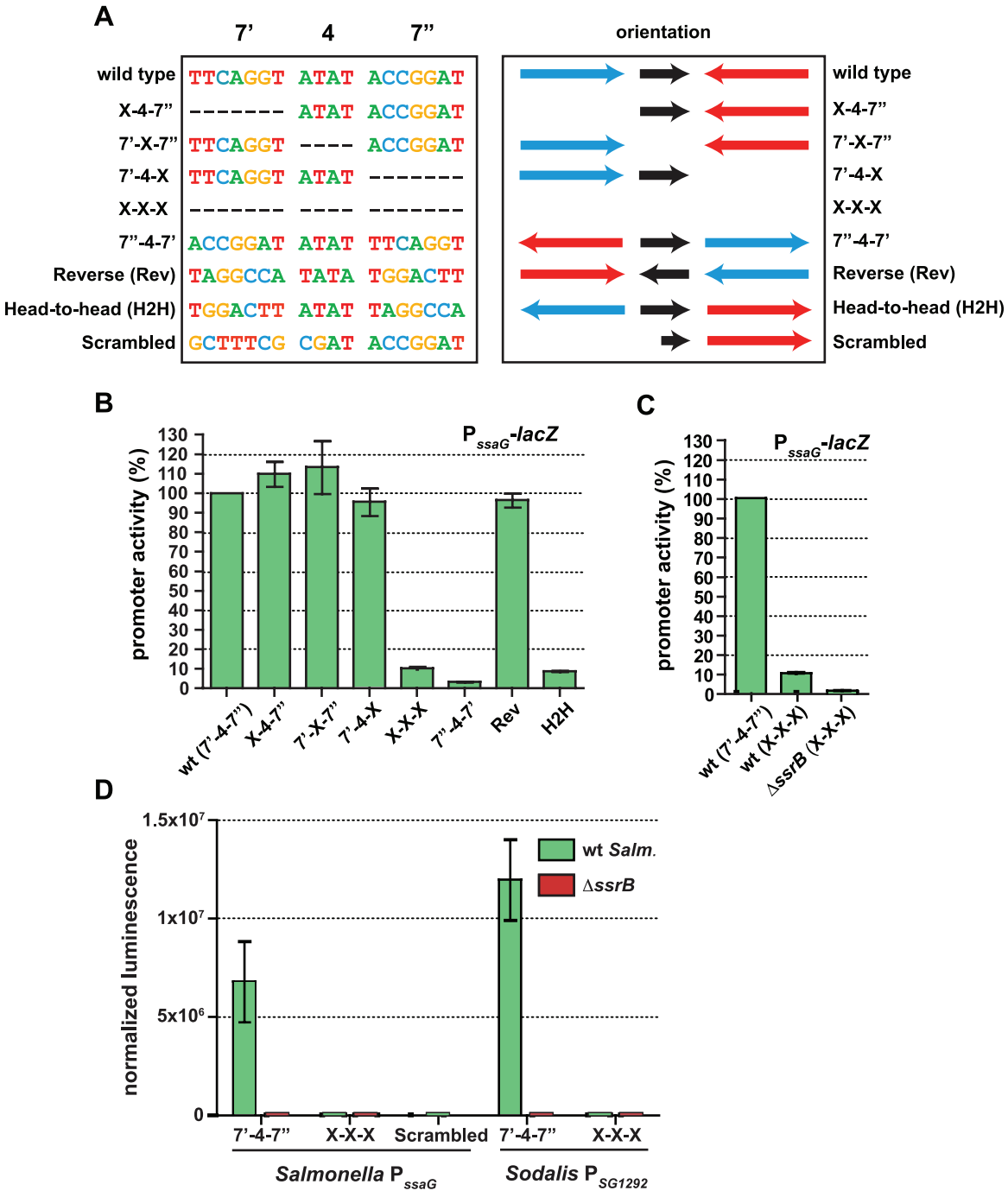
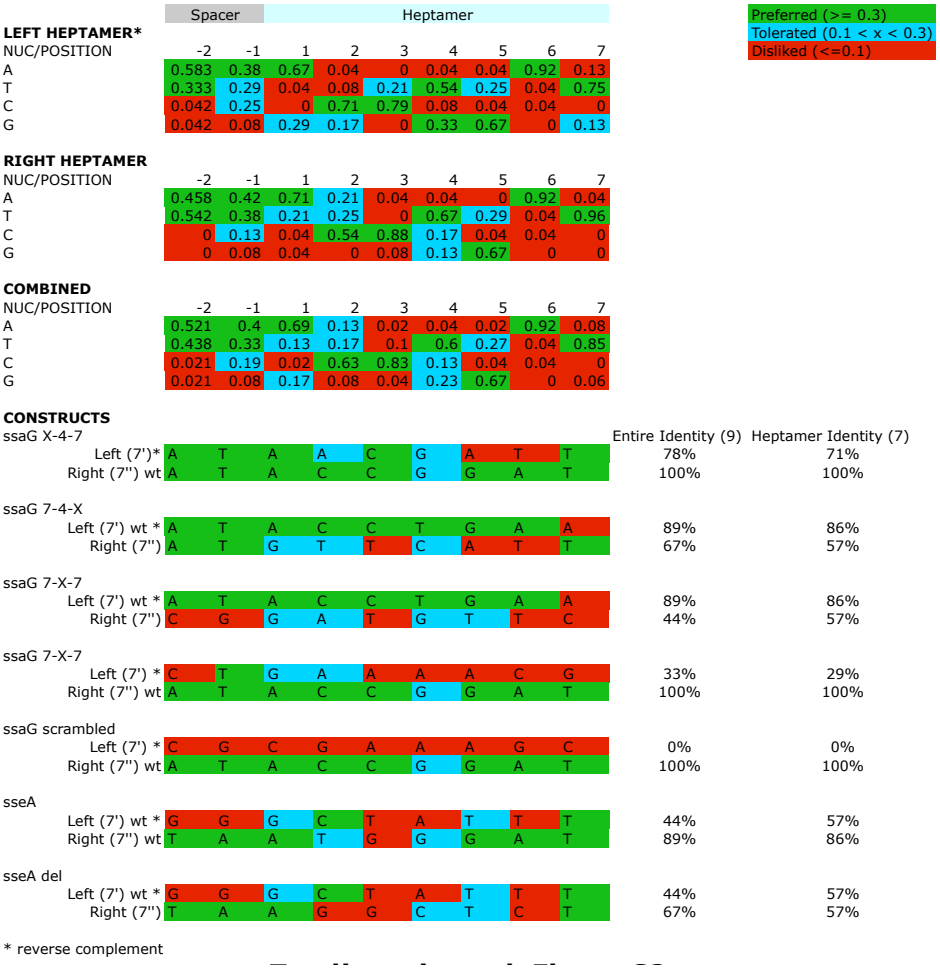

A
7'
4
7"
orientation
wild type TTCAGGT ATAT ACCGGAT
X-4-7" ------- ATAT ACCGGAT
7'-X-7" TTCAGGT ---- ACCGGAT
7'-4-X TTCAGGT ATAT -------
X-X-X ------- ---- -------
7"-4-7' ACCGGAT ATAT TTCAGGT
Reverse (Rev) TAGGCCA TATA TGGACTT
Head-to-head (H2H) TGGACTT ATAT TAGGCCA
Scrambled GCTTTCG CGAT ACCGGAT
wild type
X-4-7"
7'-X-7"
7'-4-X
X-X-X
7"-4-7'
Reverse (Rev)
Head-to-head (H2H)
Scrambled
B
P ssaG -lacZ
promoter activity (%)
wt (7'-4-7")
X-4-7"
7'-X-7"
7'-4-X
X-X-X
7"-4-7'
Rev
H2H
C
P ssaG -lacZ
promoter activity (%)
wt (7'-4-7")
wt (X-X-X)
ΔssrB (X-X-X)
D
normalized luminescence
wt Salm.
ΔssrB
7'-4-7"
X-X-X
Scrambled
7'-4-7"
X-X-X
Salmonella P ssaG
Sodalis P SG1292

Figure 3.7. **Frequency matrix analysis of SsrB palindromes.**

Presented is a frequency matrix of nucleotides occurring at each position of the 18-bp position in the 24 identified palindromes in S. Typhimurium SL1344. The 18-bp palindrome is shown as two 9-bp sequences with positions (-2) and (-1) referring to the spacer and positions 1-7 referring to the heptamer as indicated at the top of the figure. Matrices are shown for the left and right components of the palindrome in addition to a combined matrix.

Preferred, tolerated and disliked nucleotides are assigned based on frequency thresholds of 0.1 and 0.3 and are indicated by green, blue, and red respectively. Sequences for generated constructs are shown at the bottom of the figure and sequence identities with respect to the combined matrix are noted to the right of each construct sequence. The two possible permutations of the mutated palindrome are shown for the *ssaG* 7-X-7 construct.

Figure 3.7.



	Spacer		Heptamer						
LEFT HEPTAMER*									
NUC/POSITION	-2	-1	1	2	3	4	5	6	7
A	0.583	0.38	0.67	0.04	0	0.04	0.04	0.92	0.13
T	0.333	0.29	0.04	0.08	0.21	0.54	0.25	0.04	0.75
C	0.042	0.25	0	0.71	0.79	0.08	0.04	0.04	0
G	0.042	0.08	0.29	0.17	0	0.33	0.67	0	0.13
RIGHT HEPTAMER									
NUC/POSITION	-2	-1	1	2	3	4	5	6	7
A	0.458	0.42	0.71	0.21	0.04	0.04	0	0.92	0.04
T	0.542	0.38	0.21	0.25	0	0.67	0.29	0.04	0.96
C	0	0.13	0.04	0.54	0.88	0.17	0.04	0.04	0
G	0	0.08	0.04	0	0.08	0.13	0.67	0	0
COMBINED									
NUC/POSITION	-2	-1	1	2	3	4	5	6	7
A	0.521	0.4	0.69	0.13	0.02	0.04	0.02	0.92	0.08
T	0.438	0.33	0.13	0.17	0.1	0.6	0.27	0.04	0.85
C	0.021	0.19	0.02	0.63	0.83	0.13	0.04	0.04	0
G	0.021	0.08	0.17	0.08	0.04	0.23	0.67	0	0.06

Preferred (>= 0.3)
Tolerated (0.1 < x < 0.3)
Disliked (<=0.1)

CONSTRUCTS

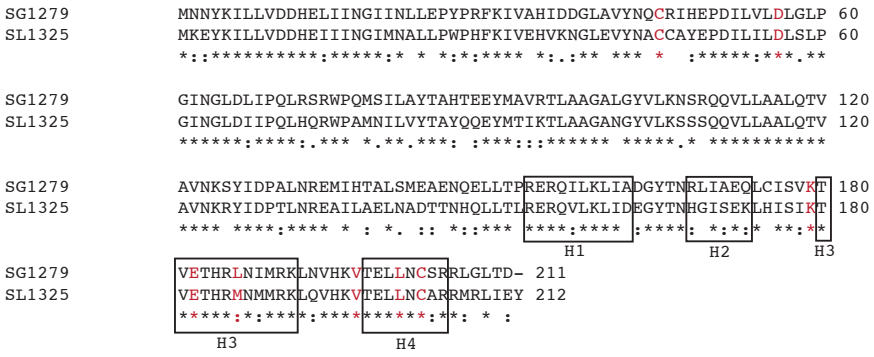
		Entire Identity (9)	Heptamer Identity (7)
ssaG X-4-7			
Left (7')*	A T A A C G A T T	78%	71%
Right (7'') wt	A T A C C G G A T	100%	100%
ssaG 7-4-X			
Left (7') wt *	A T A C C T G A A	89%	86%
Right (7'')	A T G T T C A T T	67%	57%
ssaG 7-X-7			
Left (7') wt *	A T A C C T G A A	89%	86%
Right (7'')	C G G A T G T T C	44%	57%
ssaG 7-X-7			
Left (7') *	C T G A A A A C G	33%	29%
Right (7'') wt	A T A C C G G A T	100%	100%
ssaG scrambled			
Left (7') *	C G C G A A A G C	0%	0%
Right (7'') wt	A T A C C G G A T	100%	100%
sseA			
Left (7') wt *	G G G C T A T T T	44%	57%
Right (7'') wt	T A A T G G G A T	89%	86%
sseA del			
Left (7') wt *	G G G C T A T T T	44%	57%
Right (7'')	T A A G G C T C T	67%	57%

* reverse complement

Figure 3.8. **Alignment of *Salmonella* SsrB and the ortholog from *Sodalis glossinidius*.**

Full length protein sequences of SsrB from *S.* Typhimurium strain SL1344 (SL1325) and the *Sodalis* ortholog (SG1279) were aligned with ClustalW. The helices are labeled according to the structure of SsrBc as described by Carroll et al. 2009. Helices H2 and H3 make up the helix-turn-helix DNA binding motif (H168-K191) and the dimerization helix is H4. The conserved aspartic acid residue (D56) and other residues shown to be important for transcriptional activity by Carroll et al. 2009 are indicated in red.

Figure 3.8.


SG1279 MNNYKILLVDDHELIINGIINLLEPYPRFKIVAHIDDGLAVYNQCRIHEPDILVLDLGLP 60
SL1325 MKEYKILLVDDHEIIINGIMNALLPWPHFKIVEHVKNGLEVYNACCAYEPDILILDLSLP 60
SG1279 GINGLDLIPQLRSRWPQMSILAYTAHTEEYMAVRTLAAGALGYVLKNSRQQVLLAALQTV 120
SL1325 GINGLDIIPQLHQRWPAMNILVYTAYQQEYMTIKTLAAGANGYVLKSSSQQVLLAALQTV 120
SG1279 AVNKSYIDPALNREMIHTALSMEAENQELLTPRERQILKLIADGYTNRLIAEQLCISVKT 180
SL1325 AVNKRYIDPTLNREAILAELNADTTNHQLLTLRERQVLKLIDEGYTNHGISEKLHISIKT 180
H1
H2
H3
SG1279 VETHRLNIMRKLNVHKVTELLNCSRRLGLTD- 211
SL1325 VETHRMNMMRKLQVHKVTELLNCARRMRLIEY 212
H3
H4

Regulatory evolution of the SPI-2 T3SS effector repertoire

The above results identified the conserved, yet flexible, palindrome sequence defining DNA recognition by SsrB. To examine the extent to which regulatory evolution has been selective for this genetic architecture, we created a position weight matrix (PWM) for the five strongest palindrome sites in SPI-2 and the orthologous sites in *Sodalis* SSR-3. We then searched for representative candidates of this motif in the *S.* Typhimurium genome using the simple scoring algorithm MotifLocator (186, 187). This analysis recovered the motifs upstream of *ssaB*, *ssaG*, *ssaM*, and *ssaR* that were used to generate the PWM. The palindrome in the *ssrA* promoter was not used to create the PWM due to its weaker consensus in the left heptamer, however, it was recovered in the computational analysis in a second group of lower-scoring motifs (Figure 3.9A). We identified 24 palindromes co-occurring with ChIP-on-chip peaks upstream of 24 different SsrB-regulated genes or operons. Applying a stringent threshold to the output allowed us to identify two groups - genes with high-scoring upstream palindromes (*ssaB, ssaG, ssaM, ssaR, sopD2*, *sifA*, *sifB*, *sseK2, sseK3, sseL, sseA', steC,* and *srcA*) and those with medium-scoring palindromes (0.7-0.8 threshold; *ssrA, STM1633, sseI, slrP sspH2, pipB, sseJ, pipB2, srfN, sseA* and *steB*) (Figure 3.9A) (*sseA'* denotes the SsrB palindrome sequence upstream of *sseA* that falls within the *ssaE* CDS, while *sseA* refers to the SsrB-footprinted IGR site with only one conserved heptamer defined in Figure 3.4C). Remarkably, this accounted for 17 of 22 SL1344 genome-encoded effectors translocated by the SPI-2-encoded T3SS (exceptions are chromosomal *steA*, *gogB*, and *sseK*, and plasmid-encoded *spvB* and *spvC*). These genes either lacked an upstream ChIP peak

above our 3-standard deviation cut-off (*sseK*) or had such a peak but did not reach statistical significance in our transcriptional profiling experiments (*steA, gogB*).

Our ChIP-on-chip data revealed three additional strong SsrB binding peaks within SPI-2: one in the IGR directly upstream of *ssaR*, a second within the CDS for *ssaJ,* and a third within the CDS for *ssaE* that would be predicted to influence transcription of the downstream effector/chaperone operon beginning with *sseA*. The analysis described above recovered SsrB palindrome sequences at the *sseA'* and *ssaR* locations prompting further validation of these sites. No palindrome was identified for the *ssaJ* interaction peak and so further characterization was not pursued. For the IGR palindrome upstream of *ssaR*, we tested both a chromosome-integrated transcriptional fusion and an autonomous episomal reporter. In wild type cells these reporters were as active as other SPI-2 promoters, whereas promoter activity was abrogated in Δ*ssrB* cells, implicating this IGR as a functional promoter for *ssaR* (Figure 3.9B and Figure 3.10). We next tested the function of the intragenic palindrome within *ssaE* (*sseA'*). For this, we constructed a single-copy transcriptional reporter that either contained (WT P*sseA*) or lacked (P*sseA* del) the single heptamer site in the *sseA* IGR and integrated this reporter into wild type cells and mutants lacking either *ssrB* or the *ssaE* coding sequence that removed the high-scoring intragenic palindrome *sseA'*. These experiments showed that the *sseA'* sequence contributes approximately 75% of transcriptional output at the *sseA* promoter (Figure 3.9C), since deleting the single heptamer in the *sseA* IGR had little effect on transcriptional output in any of the strain backgrounds. These reporter data are in line with the respective binding scores for the ChIP-on-chip interaction peaks (Figure 3.3) and

the sequence similarity for these elements with respect to the consensus palindrome (Figure 3.9A and Figure 3.4D). Together, these data provide compelling evidence for the identity of the DNA recognition element that has been selected through evolution to co-regulate an SsrB-dependent gene program involved in adaptation to a host.

Figure 3.9. **Genome-wide identification of SsrB palindrome sequences.**

(A) A position weight matrix was generated from all naturally-evolved palindromes in SPI-2 and used to search the genome for similar sequences. Palindromes identified in regulatory DNA that co-occurred with ChIP peaks upstream of SsrB-regulated genes were selected, aligned and binned according to those scoring >0.8 against the position weight matrix (top box) and those scoring between 0.7-0.8 (lower box). The left (7') and right (7") heptamers and the 4-bp spacer are displayed as a heat-map to show bases of high conservation (dark blue) from degenerate regions (light blue/white). The genes controlled by these promoters are indicated to the left of the sequences and coloured according to function: structural components of the T3SS (orange), effectors (blue), regulatory elements (red), T3SS chaperones and translocon (green) and SsrB-regulated genes of unknown function (black). (B) The high-scoring palindrome in the *ssaR* IGR is functional. A single-copy *ssaR* reporter was integrated on the chromosome and tested for functional activity in wild type cells and in Δ*ssrB*. Promoter activity is shown as the mean with standard error from three separate experiments. (C) Functional validation of the intragenic high-scoring palindrome in the *ssaE* coding sequence. A single-copy transcriptional reporter that either contained (WT P*sseA*) or lacked (P*sseA* del) the single heptamer site in the *sseA* IGR was integrated on the chromosome in wild type cells, or mutants lacking either *ssrB* or the *ssaE* coding sequence that removed the high-scoring intragenic palindrome (P*sseA'*). Transcriptional activity at 6 h is shown as the mean of triplicate determinations with standard error.

Figure 3.9.

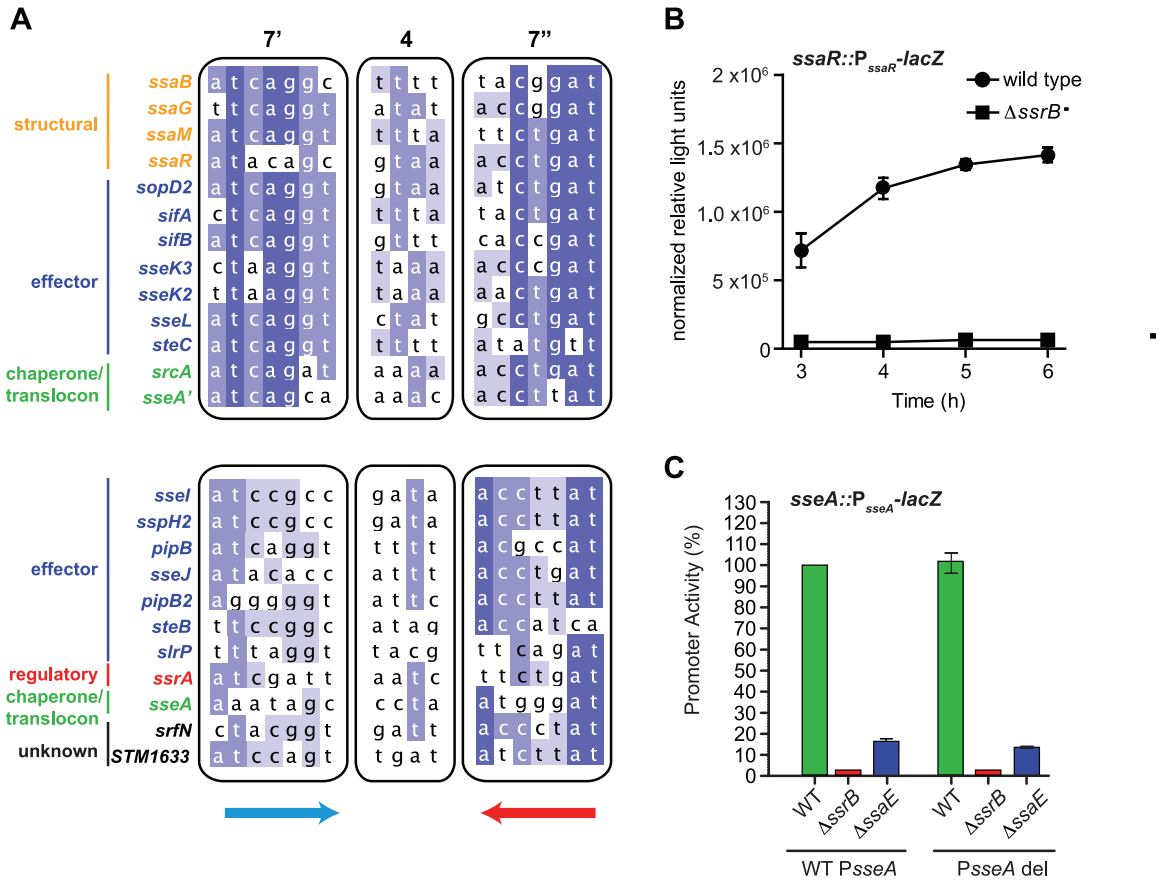
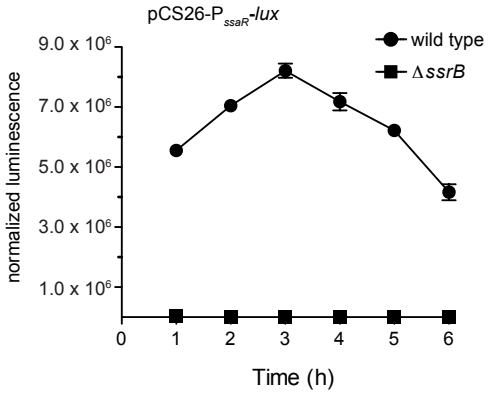

A
7'
4
7"
structural
ssaB
ssaG
ssaM
ssaR
effector
sopD2
sifA
sifB
sseK3
sseK2
sseL
steC
chaperone/ translocon
srcA
sseA'
effector
sseI
sspH2
pipB
sseJ
pipB2
steB
slrP
regulatory
ssrA
chaperone/ translocon
sseA
srfN
unknown
STM1633
B
ssaR::P ssaR -lacZ
wild type
ΔssrB
normalized relative light units
2 x10^6
1.5 x10^6
1.0 x10^6
5 x10^5
0
3
4
5
6
Time (h)
C
sseA::P sseA -lacZ
Promoter Activity (%)
130
120
110
100
90
80
70
60
50
40
30
20
10
0
WT
ΔssrB
ΔssaE
WT
ΔssrB
ΔssaE
WT PsseA
PsseA del

Figure 3.10. **Episomal transcriptional reporter for the *ssaR* palindrome**.

Luminescence production from a P*ssaR-lux* transcriptional fusion indicates an active, SsrB-dependent promoter at this location, in accord with SsrB binding to this location *in vivo*. Shown are data (mean with standard deviation) from triplicate determinations from three separate experiments. Luminescence was normalized to the optical density of the culture.

Figure 3.10.


pCS26-P$_{ssaR}$-*lux*
wild type
Δ*ssrB*
normalized luminescence
9.0 x 10^6
7.0 x 10^6
5.0 x 10^6
3.0 x 10^6
1.0 x 10^6
0
1
2
3
4
5
6
Time (h)

Discussion

Horizontal gene transfer is a well-recognized mechanism of bacterial evolution that gives rise to new phenotypes due to the coordinated expression of novel genetic components (165). A good example of this is acquisition of type III secretion by mutualists and pathogenic bacteria enabling new colonization strategies within a host (188, 189). Evolved changes to regulatory circuitry can also give rise to phenotypic diversity at the species level (33). In both cases, regulatory evolution is required to correctly deploy gene products during infection, yet the extent to which regulatory evolution contributes to pathogenic adaptation is only beginning to be realized (170). The SsrA-SsrB two-component regulatory system in *S. enterica* has been the focus of our efforts to understand the significance of regulatory evolution for pathogenic adaptation. This regulatory system was co-acquired with a T3SS encoded in the SPI-2 pathogenicity island and likely contributed to immediate and gradual phenotypic diversity as new regulatory nodes were explored and acted upon by natural selection.

Extensive work has been reported on the characterization of SsrB dependent genes, including functional evaluation of genes encoded within SPI-2 in addition to genome-wide transcriptional studies (113, 175). In this study we identified genes co-expressed under SsrB-inducing conditions and found those with strong levels of expression localized predominantly to mobile genetic elements, recently acquired genomic islands or other annotated islands. We also identified many weakly co-expressed genes, some of which may represent ancestral *Salmonella* genes recruited into the SsrB regulon like the

previously reported *srfN* (176). Some of these genes may not be directly regulated by SsrB and will require further experimental investigation.

Direct profiling of SsrB-DNA interactions using ChIP-on-chip was used to identify SsrB binding sites in the genome. This analysis identified many interactions which have not been previously described and interaction sites within coding regions of genes which may represent non-canonical functions for SsrB. Other groups have reported the existence of similar numbers of ChIP-on-chip interactions within intragenic regions for other transcription factors (190, 191) suggesting that this phenomenon is not restricted to SsrB. In light of the disparate number of microarray genes in comparison to ChIP-on-chip peaks we attempted to generate a more comprehensive picture of the SsrB regulon by combining these data sets at the operon level. In doing so we believe that the nineteen operons containing differentially expressed genes determined by microarray and containing a ChIP-on-chip peak three standard deviations above the mean captured by this analysis represent the genes directly activated by SsrB (Table 3.2). Those operons having a ChIP-on-chip peak directly upstream in the IGR region encompass the majority of known SsrB regulated genes while those possessing a ChIP peak within the CDS of the first gene may represent non-functional interactions that deserve follow-up experimental investigation.

The ChIP-on-chip data not only provided information on the identities of SsrB-regulated genes but also gave insight as to the identity of the SsrB recognition element specified by the interaction site sequences. The regulatory architecture governing SsrB input has been elusive despite several SsrB footprints being defined biochemically (33,

100, 109). Our ChIP-on-chip data further suggested that SsrB binding within SPI-2 was specific, with binding peaks overlapping precisely with regions of the DNA footprinted by SsrB (100, 109). By using a genetic screening strategy together with functional and comparative genomics, we were able to define the essential SsrB regulatory element as being an 18-bp palindrome with a conserved 7-4-7 internal organization.

In support of the palindrome as the functional entity we showed the loss of SsrB dependence as a result of deletion of this element for the *ssaG* promoter. Evaluation of the 7-4-7 palindrome in the *ssaG* promoter revealed the minimal architecture and sequence orientation required for transcriptional input. Deletion of the entire palindrome resulted in less than 1% activity in wild type cells, an equivalent level of activity to those lacking *ssrB* entirely. A search of the *S. enterica* genome for this palindromic motif revealed candidates upstream of the previously noted SsrB dependent genes, including two additional SPI-2 sites; one IGR site upstream of *ssaR* that until now had been cryptic, and one intragenic palindrome upstream of *sseA* in the *ssaE* CDS. In both cases these input sites were found to be functional.

Although palindrome architecture was conserved upstream of SsrB-regulated genes, degenerate palindromes in which one half-site was more conserved were also functional. As a result of our mutational analyses we conclude that so long as the orientation of a single heptamer of the palindrome is conserved with respect to the downstream gene, SsrB is tolerant of degeneracy in the adjacent spacer and heptamer sequences. While we were able to identify a number of limited palindrome-like sequences from our bacterial one-hybrid screen, this tolerance in addition to the library size required

to pull out an 18-bp palindrome in large numbers may explain why we isolated functional single heptamer sequences and why degenerate palindromes naturally exist in the genome. A recent report by Carroll et al, postulated that SsrB first interacts with DNA as a monomer, followed by dimerization (116). Our findings also suggest that dimerization is likely required for transcriptional activation however strong recognition by one monomer may stabilize interaction of a second monomer with a less than ideal sequence. The finding that a flexible palindromic sequence can be selective for SsrB input raises many interesting questions around the nature of regulatory evolution. The ability to use a short functional half-site adjacent to an uncharacterized threshold level of tolerated bases would reduce the period of neutral evolution required to generate an inverted repeat sequence twice the length (192), and would limit the loss of intermediate variations to drift while a more desirable palindrome is created by regulatory evolution. For bacteria that make use of horizontal gene transfer, this could increase the tempo with which new DNA is integrated into the regulatory circuitry of the cell.

We showed that the SsrB regulatory palindrome is also present in the orthologous SSR-3 island of the endosymbiont *Sodalis glossinidius* and that the palindrome evolved in *Sodalis* can act as a *cis* regulatory input function in *Salmonella.* Thus, in addition to supporting a pathogenic lifestyle within a host in *Salmonella*, it seems probable that this common promoter architecture may direct the activation of the SSR-3 T3SS of *S. glossinidius* in its endosymbiotic relationship with the tsetse fly host, although we acknowledge this requires experimental validation. The SSR-3 region in *S. glossinidius* is fully conserved in gene synteny and content with that of SPI-2 (185), with the exception

of two effector genes missing in SSR-3 (*sseF* and *sseG)* that are required to localize vacuolar *Salmonella* to the perinuclear Golgi in host cells (137, 193). The SsrB ortholog in *S. glossinidius* is ~30% divergent with SsrB at the protein level, initially leading us to think that they might have different binding site preferences. To the contrary, high local conservation in the promoters evolved in *Salmonella* and *Sodalis* was the crux in defining the functional SsrB input among stochastic noise. This analysis revealed strong palindrome sequence conservation in five promoters identified in SPI-2 and in the orthologous sequences in *Sodalis* SSR-3.

Among palindrome-containing promoters, the *ssrA* promoter is exceptional for two reasons: a lack of conservation between *Salmonella* and *Sodalis*, and the evolution of tandem palindromes in *Salmonella*. One possible interpretation of this divergent regulatory architecture in front of *ssrA* might relate to bacterial lifestyle. *Salmonella* may have retained or evolved SsrB input here to create a positive feedback loop on the regulatory system to rapidly adapt to the host environment during infection, similar to transcriptional surge described for the PhoP response regulator (194). The endosymbiotic relationship of *Sodalis* with the tsetse fly - where long-term vertical transmission has ostensibly been formative in shaping regulatory circuitry at certain promoters - may obviate the need for rapid transcriptional surge, leading to regulatory drift or selection against positive feedback. With the structure of SsrB available (116) and its recognized sequence now identified, future studies will be able to build a picture of how SsrB interacts with both its target DNA, RNA polymerase and potentially other transcription factors including nucleoid associated proteins in order to direct transcription of its regulon.

In summary, this work highlights the evolutionary significance of *cis*-regulatory mutation for the adaptation of *Salmonella* to a host animal. The DNA module that choreographs SsrB-mediated pathogenic behaviour in *Salmonella* appears to have been conserved for mutualism as well, thereby shedding new light on the significance of *cis*-regulatory mutations for bacteria evolving in different ecological settings.

Methods

Ethics statement

All experiments with animals were conducted according to guidelines set by the Canadian Council on Animal Care. The Animal Review Ethics Board at McMaster University approved all protocols developed for this work.

Bacterial strains and growth conditions

The *Salmonella* strain used for microarray and ChIP-on-chip analysis was *Salmonella enterica* serovar Typhimurium strain SL1344. Bacterial strains and plasmids used in this work are described in Table A1.1. Primer sequences used to generate constructs are available upon request. Bacteria were grown in LB medium unless otherwise indicated. Low-phosphate, low magnesium (LPM) medium was used as bacterial growth medium for microarray, ChIP-on-chip, and transcriptional reporter experiments (177). Liquid cultures were routinely grown at 37°C with shaking. Antibiotics were added to media as follows when necessary: ampicillin (Amp, 100 μg/mL), chloramphenicol (CM, 34 μg/mL) kanamycin (Kan, 30 or 50 μg/mL), and streptomycin (SM, 50 μg/mL). NM medium was used in the bacterial one-hybrid experiments as described previously (181).

Transcriptional profiling

Microarray experiments were conducted and analyzed as described previously (174). cDNA was synthesized from RNA harvested from wild type cells and an Δ*ssrB* mutant. cDNA from 2 replicate experiments was hybridized to InGen arrays and analyzed using

ArrayPipe version 1.309 (178). Probe signals underwent a foreground-background correction followed by a printTipLoess normalization by sub-grid. Duplicate spots were merged followed by averaging of the two replicates. Local intensity z scores were calculated for determination of significance.

Genome-wide operon and island analysis

For operon analysis, *S.* Typhimurium SL1344 operons were defined as groups of genes encoded on the same strand with a maximum intergenic distance of 30-bp. Operons selected for further investigation were those possessing at least one significantly down-regulated gene from the cDNA microarray analysis of an *ssrB* mutant. A cDNA microarray score was assigned based on the average fold-change value of all genes within the operon. For ChIP-on-chip analysis, a top ChIP interaction score was defined as that of the highest scoring probe within the first gene or the intergenic region upstream of the first gene of the operon. For the analysis of regions of difference (ROD) between *S. enterica* serovar Typhimurium and *Salmonella bongori*, a reciprocal-best BLAST analysis was performed to identify orthologous genes between *S.* Typhimurium and *S. bongori*. Orthologs were defined as reciprocal best BLAST pairs with E-values less than 0.005. Comparison of gene synteny between regions encoding orthologous genes was performed to identify regions of low conservation including gene deletions and insertions. The location and names of genes flanking the comprehensive list of genomic islands compared to those predicted using IslandViewer (195).

Bacterial one-hybrid screen and analysis

The bacterial one-hybrid (B1H) experiments were conducted as outlined previously using a single-step selection procedure (181). Full-length *phoP* from *E. coli* and the C-terminal domain of *ssrB* (*ssrBc*) from *S*. Typhimurium were cloned into pB1H1 to create a fusion to the αNTD of RNA polymerase. Each bait vector was transformed into *E.coli* Δ*hisB* Δ*pyrF*, purified, and then cells were transformed again with purified prey library that was previously counter-selected for self-activating preys using 5-fluoro-orotic acid. Transformants were recovered for 1 h in SOC medium, washed with NM medium supplemented with 0.1% histidine (NM + his) and allowed to grow for 2 h in this medium. Cells were washed four times with water, once with NM medium lacking histidine (NM – his), then resuspended in NM –his and plated on 150 x 15 mm dishes containing NM – his media supplemented with either 1 mM (for PhoP screen) or 5 mM (SsrBc screen) 3-aminotriazole. Selection was for ~48 hours at 37°C. Individual clones were selected from plates containing <600 colonies, the prey plasmids were isolated and sent for sequencing (Macrogen USA). Sequences were parsed to extract the 18-bp prey sequence, then inputted into MEME (version 4.1.1) for motif generation (183). MEME was run with default parameters and included searching for motifs of length 5-17 bp in either forward or reverse direction and with no limit on the number of occurrences within an input string. Motif logos were generated using Weblogo, version 2.8.2 (196).

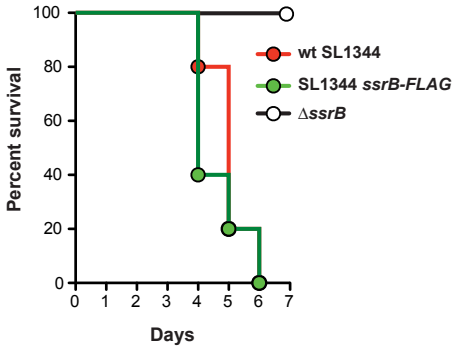
Chromatin immunoprecipitation-on-chip (ChIP-on-chip)

Chromatin immunoprecipitation-on-chip (ChIP-on-chip) was conducted as described previously using an SL1344 strain containing an *ssrB-3xFLAG* allele on the chromosome (33). The primer sequences used to generate the DNA for recombination were: 5'GAG TTA CTT AAC TGT GCC CGA AGA ATG AGG TTA ATA GAG TAT GAC TAC AAA GAC CAT GAC GG3' and 5'ATC AAA ATA TGA CCA ATG CTT AAT ACC ATC GGA CGC CCC TGG CAT ATG AAT ATC CTC CTT AG3'. This strain was generated by an allelic replacement method described previously (197) and causes lethal infection of C57BL/6 mice similar to wild type SL1344 (Figure 3.11). Immunoprecipitated DNA from three experiments under SsrB-inducing conditions (LPM growth medium) and one experiment under non-inducing conditions (exponential growth in LB medium) was hybridized to a single chip printed with four whole genome arrays designed on *S. enterica* serovar Typhimurium strain SL1344 (Oxford Gene Technology, Oxford UK). Signals for each probe within an experiment were normalized to the median channel signal for the respective array. Signal ratios were obtained for both inducing and non-inducing conditions by calculating the ratio of the control probe value and experimental probe value. A final interaction score was obtained by taking the $\log_2$ value of the ratio between the non-inducing and inducing conditions for each probe to remove SsrB interactions that occur under non-inducing conditions. Parsing and data analyses were performed using the Python scripting language. Genome-wide ChIP-on-chip data was plotted using Circos v.0.51 (198).

Figure 3.11. **Mouse virulence data for the *ssrB-FLAG* SL1344 strain.**

Groups of C57BL/6 mice were infected by oral gavage with 10^6 colony forming units of wild type *S.* Typhimurium strain SL1344, SL1344 lacking *ssrB* or SL1344 containing an allelic replacement of *ssrB-FLAG*. Mice were monitored for endpoint and sacrificed when they had lost 20% of their initial body weight. The percent of mice surviving on each day after infection is shown. There is no statistical difference between the groups.

Figure 3.11.


100
80
60
40
20
0
Percent survival
0 1 2 3 4 5 6 7
Days
wt SL1344
SL1344 ssrB-FLAG
ΔssrB

Genome-wide motif analysis

ChIP probes were ordered according to their position on the *S.* Typhimurium SL1344 genome and local maxima for ChIP interaction scores were defined as interaction peaks. Peaks with scores greater than three standard deviations from the mean probe signal were considered significant ChIP interaction peaks and were ranked in order of descending interaction score. The sequence of the top-scoring probe for each peak was exported to a text file and used for analysis by MDscan (184). The background parameter was run with output generated by the included genomebg program from the *S.* Typhimurium SL1344 genome sequence. The initial motif was generated from sequences from the top ten SsrB interaction peaks and refined using the top 25 peak sequences.

To identify instances of the palindrome motif in *S.* Typhimurium, ten 18-bp 7-4-7 palindromic motifs in the promoters of the SPI-2 genes *ssrA*, *ssaB*, *ssaG*, *ssaM*, *ssaR* and their orthologous SSR-3 genes were used as input for MDScan to identify a consensus motif and to determine the position specific scoring matrix (PWM). This PWM was used with MotifLocator to identify instances of this motif in the S. Typhimurium SL1344 genome. A background file specific to SL1344 was generated using the associated script called CreateBackgroundModel. A stringent threshold value of 0.8 was used (186, 187).

Transcriptional reporter experiments

Transcriptional fusions to *lacZ* for the *ssaG* and *sseA* promoter palindrome analysis were generated using chemically synthesized double-stranded DNA (Genscript Corp). Promoter DNA was ligated into pIVET5n, then the plasmid was subsequently conjugated

into SL1344 to generate single-copy transcriptional fusions integrated on the chromosome as described previously (177). Luciferase reporter constructs for the *Sodalis glossinidius SG1292* and *ssaR* promoters were generated by PCR amplification of promoter regions from genomic DNA templates. The luciferase reporter construct for the P*ssaG* scrambled substitution was created by PCR product splicing via overlap extension using the existing *ssaG* promoter cloning primers and two additional internal primers containing the desired mutation sequence (DTM0061R, 5'CGC GAA AGC AAC GAT TAC TCC GGC GCA CG3' and DTM0061.1F, 5'GAG TAA TCG TTG CTT TCG CGA TAC CGG ATG TTC ATT GCT TTC TA3'). This DNA was ligated into pCS26 and transformed into SL1344 to generate plasmid-based reporters. Overnight cultures of *Salmonella* were sub-cultured into LPM medium and grown with shaking for 7 h. Samples were removed hourly to measure β-galactosidase activity via a chemiluminescence-based assay as described previously (177) or luminescence directly from cultures (EnVision, Perkin-Elmer). Output was relative light units (RLU) normalized to OD600. Each experiment was performed in triplicate then averaged. Reporter activity from mutant and rewired promoters was normalized to that from wild type promoters.

Data deposition

All ChIP-on-chip data can be retrieved from the NCBI Gene Expression Omnibus at http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE20192. Data files for viewing in Artemis (http://www.sanger.ac.uk/Software/Artemis/) are available upon request.

Acknowledgments

We thank Michael Lowden for technical assistance with the microarray experiments and Wynne Lock for preliminary bioinformatics analysis. We are grateful to Scot Wolfe and Xiangdong Meng for providing purified prey libraries and technical assistance with the bacterial one-hybrid screens.

Chapter IV – Type VI Secretion System Associated Gene Clusters Contribute to Pathogenesis of *Salmonella enterica* serovar Typhimurium

Chapter IV - Co-authorship statements

Chapter IV consists of the following publication:

Mulder, D.T., Cooper, C.A., Coombes, B.K., (2012) Type VI Secretion System Associated Gene Clusters Contribute to Pathogenesis of *Salmonella enterica* serovar Typhimurium. Infect Immun., 80(6):1996-2007.

The manuscript was written by DTM and BKC.

The following experiments were performed by collaborators other than myself:

(1) Competitive Infection experiments were performed by Cooper, C.A.

Type VI Secretion System Associated Gene Clusters Contribute to Pathogenesis of *Salmonella enterica* serovar Typhimurium

David T. Mulder, Colin A. Cooper, Brian K. Coombes *

Michael G. DeGroote Institute for Infectious Disease Research and the Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario Canada L8N 3Z5

* Correspondence and request for materials should be addressed to BKC

Department of Biochemistry and Biomedical Sciences

McMaster University

HSC-4H17

Hamilton, Ontario

t. +1 905-525-9140 ext. 22159

e. coombes@mcmaster.ca

Running Title: Type VI Associated Gene Clusters in *S.* Typhimurium

Abstract

The enteropathogen *Salmonella enterica* serovar Typhimurium employs a suite of tightly regulated virulence factors within the intracellular compartment of phagocytic host cells resulting in systemic dissemination. A Type VI Secretion System (T6SS) within the *Salmonella* Pathogenicity Island 6 (SPI-6) has been implicated in this process however the regulatory inputs and the roles of non-core genes of this system are not well understood. Here we describe four clusters of non-core T6SS genes in SPI-6 based on a comparative relationship with the T6SS-3 of *Burkholderia mallei* and report that disruption of these genes results in defects in intracellular replication and systemic dissemination in mice. In addition, we show that expression of SPI-6 encoded Hcp and VgrG encoded orthologs is enhanced during late stages of macrophage infection. We identify six regions that are transcriptionally active during cell infections and which have regulatory contributions from the regulators of virulence SsrB, PhoP and SlyA. We show that levels of protein expression are very weak under *in vitro* conditions and that expression is not enhanced upon deletion of *ssrB*, *phoP*, *slyA*, *qseC*, *ompR* or *hfq* suggesting an unknown activating factor. These data suggest that the SPI-6 T6SS has been integrated into the *Salmonella* Typhimurium virulence network and customized for host-pathogen interactions through the action of non-core genes.

Introduction:

Protein secretion is an essential determinant of virulence for pathogenic bacteria. Multiple systems have evolved in order to secrete proteins in to the extracellular environment or across the membranes of target cells (11). In Gram-negative Proteobacteria these systems include the well characterized type III and type IV secretion systems that permit translocation of protein effectors from the bacterial cytoplasm directly in to target cells to modulate the host environment (199, 200). An additional type of protein secretion system involved in protein translocation known as a type VI secretion system (T6SS) has been found to contribute to interactions between bacteria and both bacteria and eukaryotic cells including unicellular and multicellular eukaryotes (201). This system employs an assortment of membrane associated proteins in order to coordinate localization and assembly of the system, however the secretory apparatus itself is composed of proteins that are structurally analogous to that of an inverted contractile bacteriophage tail (153, 154, 202). This coordinated contractile tail is utilized to deliver effector proteins to targets including the actin cytoskeleton in eukaryotic cells and peptidoglycan in bacterial cells (151, 203).

Salmonella enterica subspecies *enterica* serovar Typhimurium (*S.* Typhimurium) employs two well characterized type III protein secretion systems (T3SS) to manipulate host cells through effector translocation in order to invade gut epithelial cells and disseminate systemically within phagocytic cells (11, 34). Most serovars of *Salmonella*

enterica subspecies *enterica* also encode a T6SS within *Salmonella* Pathogenicity Island 6 (SPI-6), while two other classes of T6SS have been described in *Salmonella bongori* and *Salmonella enterica* subspecies *arizonae* at SPI-2 and SPI-6 respectively (26, 28, 204). The *S. enterica* subspecies *enterica* T6SS was first described as part of the *Salmonella enterica* centisome 7 genomic island (SCI) as a contributing factor to eukaryotic pathogenesis (156). Aside from the T6SS, SCI also encodes the adhesin and invasin PagN, the fimbrial gene cluster *safABCD* and the transcriptional regulator *sinR* (205, 206). Deletion of SCI resulted in an approximate 50% reduction in internalization in HEp-2 cells that could not be complemented by PagN expression, however complementation with the *saf*ABCD or *sinR* genes was not tested (156). Disruption of SPI-6 has also been found to result in a defect in systemic dissemination in orally infected BALB/c mice (207). In support, a number of SPI-6 T6SS encoded genes have been implicated in long term persistence in macrophages and mice through transposon mutagenesis, however these transposon mutations have not been verified by precise deletions (208–210). Contrasting with these reports, another study described a hypervirulence phenotype in BALB/c mice upon deletion of the core structural gene *sciS*, and found an increase in numbers of bacteria within macrophages after six hours suggesting a time dependent effect on replication (157). Further supporting a role in interactions with macrophages, genome-wide transcriptional profiling of *S.* Typhimurium during macrophage infections found that many of the SPI-6 T6SS genes were upregulated during macrophage infection although an increase over time between 4 and 12 hours was not observed (211). An earlier gene expression profiling study also observed a trend

towards increased expression of some of these genes however statistical significance was not met (212). The mechanism leading to upregulation in this environment is unknown, however an *in vivo*, host-dependent activation mechanism has been previously described for the T6SS of *Vibrio cholerae* which requires endocytosis by phagocytic cells (213).

Distinct classes of T6SS maintain a core set of thirteen genes required for localization of the system to the bacterial inner membrane and for functional secretion (149). T6SS employ these encoded proteins to coordinate secretion of protein substrates via assembly, contraction and disassembly of a phage-like tail (155, 202). For some T6SS harboring organisms, the core gene encoding the tail-spike protein encodes a fusion protein that allows translocation of an effector domain (155). However, in *S.* Typhimurium and other T6SS harboring organisms, no evolved tail-spike protein is present (28). In addition to the core genes, organisms that possess a T6SS typically encode less well conserved T6SS associated genes that differ extensively in their number and location throughout the genome and some of these genes have been described to act as effector or regulatory proteins (155, 214, 215). Recently, a set of genes encoding T6SS-secreted toxin proteins were identified in *Pseudomonas aeruginosa* that are involved in interbacterial competition (203). While *S.* Typhimurium encodes a number of non-core T6SS associated genes, their role in pathogenesis of the host small intestine or intracellular environment is poorly understood (28).

Non-core T6SS encoded genes are known to play a role in the integration of T6SS in to global regulatory networks. In *Vibrio cholerae*, some strains constitutively express T6SS associated proteins while others do not suggesting niche adaptation (155). Regulatory control has been observed to occur at the transcriptional level by the activator of the alternative sigma factor 54 VasH in *V. cholerae* (216). In addition, post-translational regulation by a T6SS encoded threonine phosphorylation pathway has been described for the T6SS of *Pseudomonas aeruginosa* (217). The non-core T6SS associated gene TagR has been characterized as acting as a post translational repressor of this phosphoylation dependent mechanism (218). Cues from within the host intracellular environment likely play a role in the regulation of the *S.* Typhimurium T6SS as it has been determined that transcription of *sciS* increases approximately sixteen hours post-invasion of macrophages, and transcript levels are negatively correlated with activity of the two component response regulator of the SPI-2 T3SS, SsrB (157). Two component signaling plays a major role in the regulation of virulence in *S.* Typhimurium and these systems may regulate T6SS activity in concert with non-core T6SS encoded genes as is the case for the sensor kinases LadS and RetS in *Pseudomonas syringae* (219). A hallmark of T6SS activity has been the stable expression of the phage-like proteins Hcp and VgrG that form the phage tail-tube and tail-spike respectively (with the ortholog of the latter annotated as VrgS in *S*. Typhimurium). The expression of these proteins in some organisms is absent under experimental conditions, likely due to the absence of inducing stimuli. This can be circumvented by the use of strains with constitutive T6SS activity such as in the case of a *retS* mutant in *P. syringae* or the V52 *V. cholerae* strain (155, 219). Whether or not stable

expression and secretion of the Hcp and VrgS proteins occurs under *in vitro* or *in vivo* conditions in *S.* Typhimurium is currently unknown. However, plasmid based over-expression of the *hcp* ortholog in the SPI-6 T6SS of *S.* Typhi results in secretion *in vitro* suggesting a functional system that can contribute to bacterial pathogenesis (220).

In this study we determined that both the SPI-6 core T6SS and associated genes contribute to systemic dissemination during *in vivo* infections of mice and contribute to intracellular replication in macrophages. We used cell infections to determine that the Hcp and VrgS proteins are most highly expressed during late stages of infection of macrophages. We identified six transcriptionally active regions within the SPI-6 T6SS, three of which receive transcriptional input during growth *in vitro* from SsrB, PhoP, and SlyA. We then show that expression of Hcp and VrgS proteins *in vitro* does not change significantly upon deletion of these and other key virulence regulatory proteins suggesting that induction of this system may rely on an unknown activator specific to the T6SS *in vivo*.

Methods and Materials:

Ethics Statement

All experiments with animals were conducted according to guidelines set by the Canadian Council on Animal Care. The Animal Review Ethics Board at McMaster University approved all protocols developed for this work.

Informatics

Genome sequences were retrieved from NCBI Genbank (*Salmonella enterica* subsp. *enterica* serovar Typhimurium str. LT2 accession number AE006468.1, *Burkholderia_mallei*_ATCC_23344 accession number NC_006349.2, *Burkholderia_pseudomallei*_K96243 accession number NC_006351.1) and Wellcome Trust Sanger Institute (*Salmonella enterica* subsp. *enterica* serovar Typhimurium str. SL1344). BlastP searches were performed online using the NCBI BLAST servers against the non-redundant protein sequence database (nr) and protein sequences were obtained from the *S*. Typhimurium str. LT2 genome file NC_003197.faa. The non-core gene search analysis employed an expect threshold of 1 and target sequences was limited to 1,000. A custom Python script was used to parse the Blast results for organisms with at least two of the proteins queried that did not belong to the Genus *Salmonella*.

Strains and Growth Conditions

All *S. enterica* strains used in this study for experimental work were *Salmonella enterica* subspecies *enterica* serovar Typhimurium strain SL1344 derivatives. Constructs were designed using the *S.* Typhimurium LT2 genome and tested in *S.* Typhimurium SL1344. These genomes differ within the SPI-6 T6SS region by a silent mutation in *sciG*, a M1175T mutation in *sciS*, an in-frame 5 amino acid mutation (A50-L54) in *sciI* and an in-frame 166 amino acid mutation (V620-Q786) in *rhs1*. None of these mutations are expected to have affected the results of experiments. Media employed included LB broth (1%(w/v) tryptone, 1%(w/v) sodium chloride, 0.5%(w/v) yeast extract), acidic low phosphate low magnesium (LPM) minimal media (5mM KCl, 7.5mM $(NH_4)_2SO_4$, 0.5mM K_2SO_4, 80mM MES, 0.3% glycerol, 0.1% casamino acids, 24µM $MgCl_2$, 337µM PO_4^{3-}), M9 minimal media (5mM Na_2HPO_4, 22mM KH_2PO_4, 8.6mM NaCl, 18.6mM NH_4Cl, 11.1mM glucose, 2mM $MgSO_4$, 100uM $CaCl_2$, 0.1% casamino acids, pH 7.4) or Dulbecco's Modified Eagle Medium supplemented with 10% fetal bovine serum (Life Technologies). Bacteria were grown at 37°C with shaking at 225 RPM.

Construct inserts were amplified from *S.* Typhimurium SL1344 genomic DNA using platinum taq DNA polymerase high-fidelity (Life Technologies), purified on PCR purification spin-columns (Qiagen), digested using restriction endonucleases (NEB) and ligated using T4 DNA ligase (NEB). Constructs were propagated through *Escherichia coli* DH5 alpha, and electroporated into *S.* Typhimurium SL1344 competent cells. Integrated chromosomal transcriptional fusion reporters were generated by cloning the intergenic regions of interest into the integrative plasmid pIVET5n and conjugating these into *S.* Typhimurium SL1344 through *E. coli* SM10 lambda *pir* (177). Deletion mutants

were generated using the lambda red mutagenesis method (221). Hemagglutinin tags were introduced on the chromosome using a modified lambda red mutagenesis method (197). Lambda red cassettes were amplified using platinum taq DNA polymerase (Life Technologies) from template plasmids pKD3 (Cm replacement), pKD4 (Kn replacement), pSUB314 (Cm marked HA fusion) or pSUB315 (Kn marked HA fusion), purified on PCR purification spin-columns (Qiagen) and electroporated in to water and glycerol washed competent *S.* Typhimurium SL1344 competent cells containing the helper plasmid pKD46 (197, 221). Plasmid based luciferase transcriptional fusion reporters were generated by cloning the intergenic regions of interest into the pGEN-luxCDABE plasmid (222). Flag epitope expression constructs were generated by cloning the indicated genes in to the pFLAG-CTC expression vector (Sigma), induced with 0.5 mM IPTG and secretion assays were performed in LPM minimal media as previously described (174). Briefly, after six hours growth in LPM, bacteria were pelleted at 10,000 *g*, media was removed and filtered through a 0.2µm acrodisc syringe filter (PALL), trichloroacetic acid was added to 10%, and incubated on ice at 4°C overnight. Samples were pelleted at 18,000 *g* for 30 minutes at 4°C, washed with 1mL acetone, pelleted again at 18,000 *g* for 30 minutes and pellet was allowed to air dry. Pellet was then resuspended in OD600 normalized volume of 2x SDS loading dye. For *sciG* complementation, *sciG* and its upstream region were cloned in to the low copy pWSK29 plasmid (223). Antibiotics were added to media as follows when necessary unless otherwise specified: ampicillin (Amp, 100µg/mL), chloramphenicol (Cm, 34 µg/mL), kanamycin (Kn, 50µg/mL), streptomycin (Sm, 50µg/mL), and gentamycin (Gent, 10 or 100 µg/mL). All strains and plasmids

employed here are listed in Table A1.1. All primers used to generate constructs are listed in Table A1.2.

Mouse infections

Bacteria from LB stationary phase cultures were washed and diluted in inoculation buffer (0.1 M HEPES buffer, 0.9% NaCl) and 10-15 week old female C57BL/6 mice were infected orally with 100 μL of 2 x 10^7 CFU in groups of 3-5. Endpoint analyses and competitive infections were performed as previously described (174). For endpoint experiments, mice were monitored for endpoint and sacrificed when they had lost 20% of their initial body weight. For competitive infections, mixed inoculums of wild type (pseudogene ushA::Cm marked) and mutant were used. At 72 hours post-infection, mice were sacrificed and spleen and liver were harvested, homogenized in PBS and plated to determine bacterial numbers, which were compared to initial inoculum plating numbers as a ratio of CFU (mutant output/wild type output)/(mutant input/wild type input). For *sciG* complementation the wild type reference strain also possessed the empty pWSK29 plasmid. Statistical analyses were performed in Prism using the nonparametric Wilcoxon signed rank test to determine whether mutant CI values differed from 1 ($P < 0.05$ at CI of 95%).

Cell infections

RAW 264.7 macrophages were seeded at 2 x 10^5 CFU per well in 24-well plates and grown at 37°C supplemented with 5% CO_2 in DMEM media supplemented with 10%

fetal bovine serum (Life Technologies). Bacteria from overnight LB cultures were pelleted, resuspended in DMEM/10% FBS and opsonized in human serum and macrophages were infected with an MOI of 50 for 30 minutes followed by two PBS rinses, a 1.5 hour incubation in DMEM/10% FBS + 100µg/mL gentamicin, followed by two PBS rinses, and incubation to the given lysis time point in DMEM/10% FBS + 10µg/mL. (174). At the indicated time points, infected cells were washed with PBS, lysed in cell lysis buffer (1mL PBS, 1% Triton X-100, 0.1% SDS), and plated to establish bacterial counts. For epithelial cell co-culture experiments, HeLa cells and HEp-2 cells were seeded at 1 x 10^5 CFU per well in 24-well plates and co-cultured at an MOI of 10 from pelleted and DMEM/10% FBS washed three hour exponential phase LB cultures. For epithelial cell infections, HeLa or HEp-2 cells were seeded at 1 x 10^5 CFU per well in 24-well plates and infected with an MOI of 50 from pelleted and DMEM/10% FBS washed three hour exponential phase LB cultures for 15 minutes, followed by two PBS washes and allowed to complete invasion in fresh DMEM/10% FBS for 20 minutes. Cells were then washed twice and incubated in DMEM/10% FBS + 100µg/mL gentamicin for 1.5 hours, followed by two PBS washes and lysis in cell lysis buffer. Lysates and infection inoculum were plated and counted to calculate invasion efficiency. For *sciG* complementation the wild type reference strain also possessed the empty pWSK29 plasmid. Statistical analyses were performed in Prism using a one sample T test.

Transcriptional reporter assays

For beta-galactosidase assays, strains were grown overnight in M9 media and then sub-cultured 1:50 into LPM media and grown at 37°C with shaking. At defined time intervals, optical density at 600nm (OD600) was read using a spectrophotometer and 200µL culture samples were pelleted and frozen for analysis. For analysis pellets were resuspended in 200µL PBS, and lysed with the addition of 50µL chloroform. 2µL of the lysate was combined with 100uL of Tropix Galacto-star detection reagent (Applied Biosystems) and Relative Light Units (RLU) were measured using an Envision plate reader (PerkinElmer) (177). For intracellular luminescence assays, macrophages in four sets of wells were identically infected as previously described. At each time point, one set of wells was washed with PBS and luminescence was measured using an Envision plate reader. Cells were then lysed in cell lysis buffer and plated to determine bacterial counts. Statistical analyses were performed in Prism using a one sample T test.

Immunoblotting

Strains were grown overnight in LB media at 37°C. Cultures were pelleted, washed and diluted into DMEM containing 10% FBS. Inoculum was added to 12-well tissue culture plates with or without eukaryotic cells for co-incubation protocol or were infected as previously described. Plates were incubated at 37°C under 5% CO_2. At defined time points, the contents of the well was collected, pelleted and washed in PBS. Pellets were lysed in 200 µL 2x SDS sample buffer (100mM Tris-HCL pH 6.8, 20%% (v/v) glycerol, 4% (w/v) SDS, 0.002% (w/v) bromophenol blue, 4M urea and 0.2M dithiothreitol). Samples were boiled 10 minutes at 100 °C, centrifuged 1 minute at 17,900 *g* and loaded

into 15% SDS-PAGE gels. Protein was transferred to PVDF membrane, incubated overnight with 1:10,000 mouse anti-DnaK (Stressgen), 1:1,000 mouse anti-HA (Covance) or 1:5,000 mouse anti-FLAG antibodies (Sigma). Blots were then washed in Tris-buffered saline tween-20 (TBST), incubated with 1:5,000 goat anti-mouse HRP antibodies (Sigma), washed in TBST, and detected using enhanced chemiluminescence (Western Lightning – PerkinElmer). Low signal Western blots of samples collected from regulator mutant bacteria were detected using SuperSignal West Femto (Thermo Scientific).

Results:

The SPI-6 T6SS non-core genes include T6SS-associated gene pairs

The *S.* Typhimurium SPI-6 T6SS encodes 13 core T6SS genes in addition to other uncharacterized genes (28). Many T6SS have undergone preliminary characterization in other sequenced bacteria and we hoped to gain insight into these uncharacterized genes by identifying a similar characterized system. Using the SPI-6 T6SS core protein sequences, we performed a BlastP search of the NCBI BLAST non-redundant protein sequences (nr) database to identify an organism that encodes a similar T6SS. We observed that the T6SS-3 (Tss3) of *Burkholderia mallei* and *B. pseudomallei* is highly similar (224) (Figure 4.1). These systems have been previously reported to be closely related through phylogenetic analyses (28, 149, 150). The *Burkholderia spp.* systems encode 15 genes in common with *S.* Typhimurium representing 13 core genes and share synteny in core T6SS genes aside from their VgrG ortholog, which is positioned at opposite sides of the island. The two systems differ in the number of additional non-conserved genes: four genes in a single cluster in the *Burkholderia* variants (BMAA Cluster) and ten genes within three clusters in the *S.* Typhimurium T6SS (STM Clusters 1-3). These clusters encode genes *sciJ-sciL* (*STM0274A-STM0278*), *sciQ-sciR* (*STM0283-STM0284*), and *sciT-sciV* (*STM0286-STM0288*) respectively. A fourth cluster of genes (STM Cluster 4): *sciW-sciY* (*STM0290-STM0298*), located between the T6SS and *safABCD* operon, encodes recombination hotspot (Rhs) elements, poorly conserved genes and transposases.

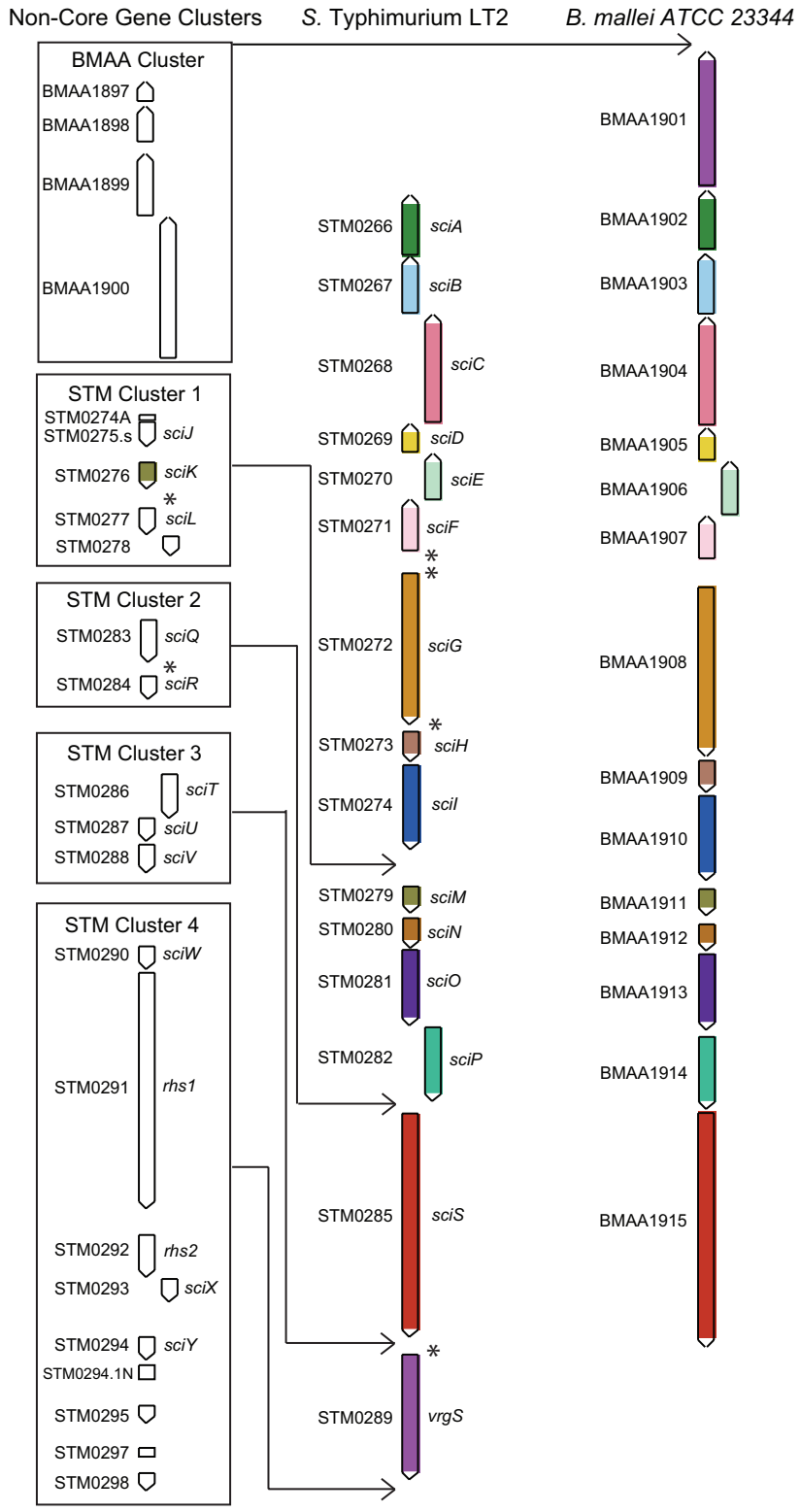
In order to better understand the non-core genes that are not conserved between the *S.* Typhimurium SPI-6 T6SS and the *Burkholderia* Tss-3 we searched for organisms in which orthologs for at least two of these genes exist. We did not include *sciK*, a paralog of *sciM*, in cluster 1 given its extensive conservation as an Hcp encoding gene even though it is absent from the *B. mallei* Tss3. In cluster 1, *STM0274A* and *sciJ* were not well conserved outside of *Salmonella* and were found to exist in only *Escherichia sp.* TW09308. In contrast, orthologs of *sciL* and *STM0278* were found to co-occurr in 14 other genomes, including species of *Escherichia*, *Enterobacter*, *Erwinia*, and *Frateuria*, in close proximity to an *sciS* ortholog (suggestive of a T6SS), and in *Advenella*, *Pantoea*, and *Pseudomonas*, not in close proximity to an *sciS* ortholog. In cluster 2, orthologs of *sciQ* and *sciR* were found to co-occur only in *Agrobacterium tumefaciens* in close proximity to an *sciS* ortholog. In cluster 3, orthologs for *sciU* and *sciV* were found to co-occur in four genomes, including species of *Bordetella* in close proximity to an *sciS* ortholog and in *Delftia*, and *Vibrio*, not in close proximity to an *sciS* ortholog. While not always in close proximity, all the organisms indicated here that possessed these gene pairs did encode at least one *sciS* ortholog. While many genomes were found to encode orthologs of genes within cluster 4, only *Enterobacter hormanechei* ATCC 49162, *Cronobacter sakazakii* ATCC BAA-894, *Cronobacter* turicensis z3032, Escherichia *coli* TA280 and *Pantoea sp.* SL1_M5 encoded orthologs of these genes near orthologs of genes from clusters 1-3.

Our search also recovered the previously identified orphan Hcp encoding gene, *STM3131* which is not linked with the SPI-6 T6SS and likely has a distinct origin from the *S. enterica* subspecies *enterica* and subspecies *arizonae* SPI-6 T6SS (28). Recognizing the possibility that *STM3131* could encode a gene product functionally relevant to the *Salmonella* SPI-6 T6SS, we included this gene in downstream experiments.

Figure 4.1. Genetic diagram of the SPI-6 T6SS.

Gene diagram showing the fifteen T6SS genes encoded in common between the *S.* Typhimurium LT2 SPI-6 T6SS and the *B. mallei* ATCC 23344 Tss-3 T6SS. Orthologs are indicated by shading with the same color. Non-conserved genes in *B. mallei* (BMAA Cluster) and *S.* Typhimurium (STM Clusters) are displayed in boxes on the left side of the figure and their positions relative to the core genes are indicated by black lines. Observed transcriptionally active regions are denoted by an asterisk (*).

Figure 4.1.


Non-Core Gene Clusters
S. Typhimurium LT2
B. mallei ATCC 23344
BMAA Cluster
BMAA1897
BMAA1898
BMAA1899
BMAA1900
STM Cluster 1
STM0274A
STM0275.s
sciJ
STM0276
sciK
STM0277
sciL
STM0278
STM Cluster 2
STM0283
sciQ
STM0284
sciR
STM Cluster 3
STM0286
sciT
STM0287
sciU
STM0288
sciV
STM Cluster 4
STM0290
sciW
STM0291
rhs1
STM0292
rhs2
STM0293
sciX
STM0294
sciY
STM0294.1N
STM0295
STM0297
STM0298
STM0266
sciA
STM0267
sciB
STM0268
sciC
STM0269
sciD
STM0270
sciE
STM0271
sciF
STM0272
sciG
STM0273
sciH
STM0274
sciI
STM0279
sciM
STM0280
sciN
STM0281
sciO
STM0282
sciP
STM0285
sciS
STM0289
vrgS
BMAA1901
BMAA1902
BMAA1903
BMAA1904
BMAA1905
BMAA1906
BMAA1907
BMAA1908
BMAA1909
BMAA1910
BMAA1911
BMAA1912
BMAA1913
BMAA1914
BMAA1915

The SPI-6 T6SS contributes to pathogenesis in a mouse model of typhoid.

We generated mutants for both *sciG* (the *clpV* ortholog), and *sciS* (the *icmF* ortholog), genes which have been established as essential for T6SS assembly and functional secretion in other systems (147, 225). When mice were infected with wild type and mutant strains of *S.* Typhimurium, mice infected with the mutant strains reached endpoint approximately one to two days later than mice infected with wild type *S.* Typhimurium (Figure 4.2A). The major process leading to mortality in this model is systemic dissemination and replication of the bacteria in organs including the spleen and liver. To determine whether the T6SS contributes to this process, we performed competitive infection experiments with the *sciG* and *sciS* mutants against the wild type strain (Figure 4.2B). We obtained combined (liver and spleen) CI values for these mutants of 0.39 and 0.43 respectively that was in agreement with previously reported data for deletion of the entire SPI-6 locus (207). This defect in systemic dissemination was successfully complemented by plasmid-based expression of *sciG* under control of its native promoter.

The SPI-6 non-core gene clusters contribute to pathogenesis in mice.

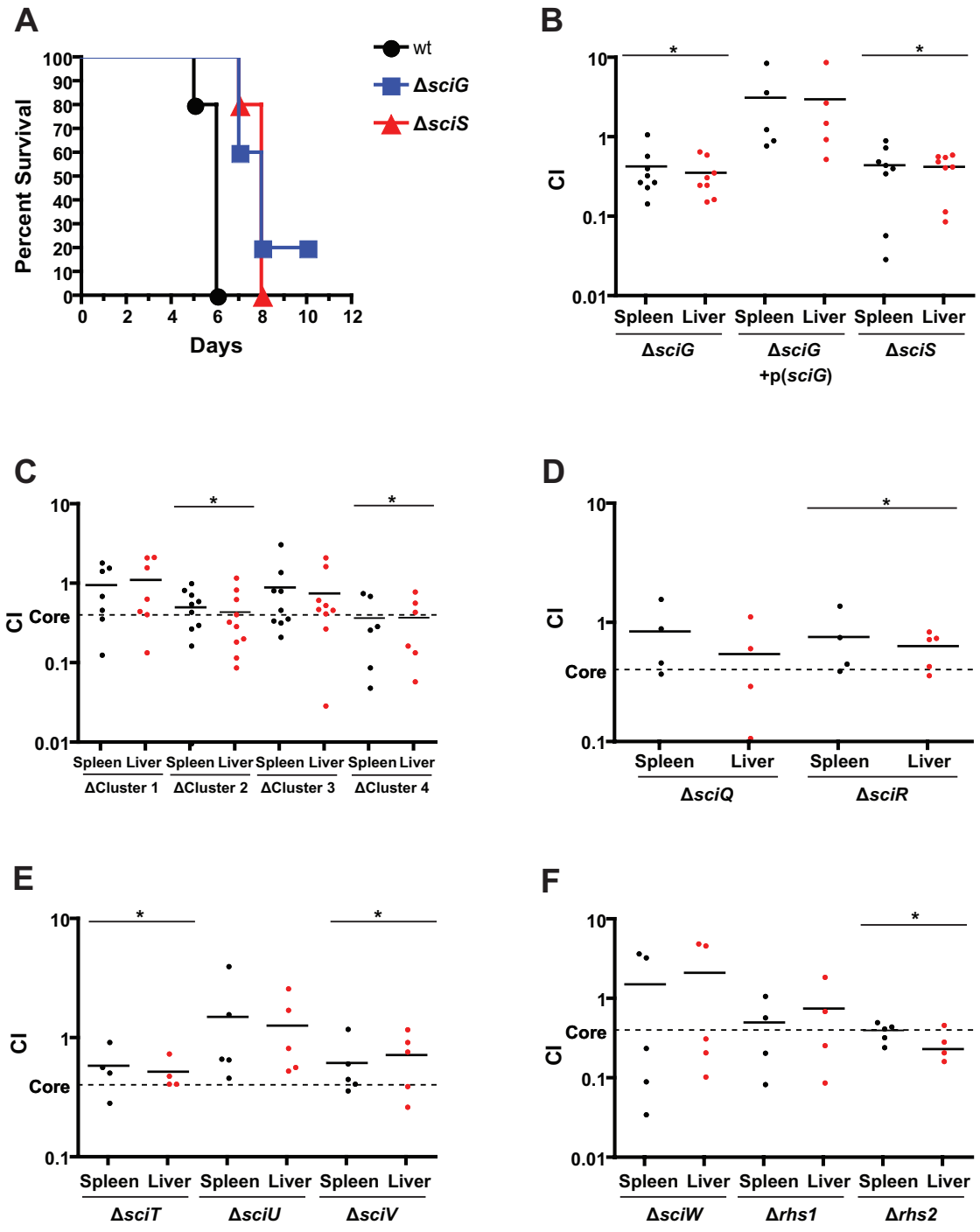
Since we were able to measure a virulence effect for the T6SS core genes by oral infection of mice we employed this model to test the contribution of the SPI-6 non-core gene clusters to systemic dissemination. We generated deletion mutants for clusters 1 through 4. When competed against wild type in oral infections, clusters 2 and 4 yielded combined organ CI values of 0.47 and 0.37, values similar to that obtained by deletion of the core genes *sciG* and *sciS* (Figure 4.2C). Combined CI values for clusters 1 and 3 did

not differ significantly from wild type although we observed clustering of values for the cluster 3 deletion near the CI value of the core mutant deletion (average = 0.82, median = 0.49) and therefore included this cluster for further analysis. To identify specific genes contributing to this fitness defect, we generated additional single gene deletion strains for clusters 2, 3 and 4. When competed against wild type the individual mutants for cluster 2 both had combined CI values of 0.69, and were not as low as the cluster 2 mutant combined CI of 0.47 (Figure 4.2D). CI experiments for individual cluster 3 mutants showed that *sciT* and *sciV* had combined CI values below wild type (0.55 and 0.67) while deletion of *sciU* was found to trend towards increased systemic dissemination with a combined CI value of 1.39, albeit not significant (Figure 4.2E). CI experiments for selected, individual cluster 4 mutants revealed a low combined CI of 0.32 indicating a strong defect for the *rhs2* mutant on systemic dissemination (Figure 4.2F).

Figure 4.2. Contribution of the SPI-6 T6SS to *S.* Typhimurium pathogenesis in mice.

(A) Survival curve of SPI-6 T6SS mutant bacteria in C57BL/6 mice. Mice infected with Δ*sciG* (ATPase*)* and Δ*sciS* (*icmF*) mutant bacteria reached endpoint two days later than those infected with wild type bacteria. (B-F) Competitive infections between wild type and SPI-6 T6SS mutant bacteria in C57BL/6 mice orally infected with 10^6 CFU. (B) Δ*sciG* and Δ*sciS* mutants were less fit in the liver and spleen than wild type bacteria, a defect that could be resolved by plasmid-based complementation of *sciG* (*, $P<0.05$). (C) Deletion of non-core gene clusters 2 and 4 resulted in similar CI values to deletion of essential T6SS genes (core), while deletion of cluster 3 resulted in a CI of intermediate value. (D) Deletion of individual genes within cluster 2 could not recapitulate the defect caused by deletion of the entire cluster. (E) Deletion of individual genes within cluster 3 with a non-significant trend towards increased replication in the *sciU* mutant. (F) Deletion of individual genes within cluster 4 showed a strong defect for Δ*rhs2*.

Figure 4.2.


A
Percent Survival
Days
wt
ΔsciG
ΔsciS
B
CI
Spleen Liver Spleen Liver Spleen Liver
ΔsciG
ΔsciG +p(sciG)
ΔsciS
C
CI
Core
Spleen Liver Spleen Liver Spleen Liver Spleen Liver
ΔCluster 1
ΔCluster 2
ΔCluster 3
ΔCluster 4
D
CI
Core
Spleen Liver Spleen Liver
ΔsciQ
ΔsciR
E
CI
Core
Spleen Liver Spleen Liver Spleen Liver
ΔsciT
ΔsciU
ΔsciV
F
CI
Core
Spleen Liver Spleen Liver Spleen Liver
ΔsciW
Δrhs1
Δrhs2

The SPI-6 T6SS contributes to intracellular replication in macrophages

Systemic dissemination is dependent on survival within phagocytic cells. To determine whether the SPI-6 T6SS contributes to intracellular replication we assessed the intracellular survival/replication rates for T6SS core gene and non-core gene mutants in mixed infections with wild type *Salmonella* 24 hours post infection in the murine macrophage line RAW 264.7 (Figure 4.3A). Compared to the wild type *S.* Typhimurium parent strain, the *sciG* and *sciS* core T6SS gene mutants were found to have a ratio of intracellular replication of approximately 0.5 compared to a ratio of intracellular replication of 0.13 for an *sseC* SPI-2 T3SS translocon mutant that is defective for intracellular persistence. Plasmid based expression of *sciG* was able to complement the *sciG* mutant to near wild type levels of intracellular replication. We observed a significant decrease in replication over the wild type strain for clusters 1 and 3. Individual deletions within cluster 1 revealed a significant contribution for *STM0278*, while no individual gene was found to have a significant contribution for cluster 3 (Figure 4.3B). While not significant, we again observed a trend towards increased replication for the cluster 3 single gene mutant *sciU*.

To determine whether this replication phenotype was specific to macrophages, we performed invasion assays in HeLa and HEp-2 epithelial cells. While we observed a clear defect in invasion for our *hfq* mutant control that is defective for invasion, we saw no significant effect on invasion in both cell types for any of our core and non-core SPI-6 T6SS mutants (Figure 4.4).

Figure 4.3. Assessment of SPI-6 T6SS to replication in macrophages cell culture.

Intracellular replication of SPI-6 T6SS mutants in RAW 264.7 murine macrophage-like cells expressed as a fraction of wild type at 24 hours post-infection compared with a Δ*sseC* SPI-2 T3SS translocon mutant (*, $P<0.05$). Decreased intracellular replication was observed for core gene mutants Δ*sciG* and Δ*sciS*, with near wild type levels for the *sciG* complemented strain. Replication defects were also observed for non-core gene cluster 1 and 3 mutants. Decreased intracellular replication was only observed for *STM0278* when single gene mutants were tested while a non-significant increase was observed for individual mutants of non-core gene cluster 3.

Figure 4.3.

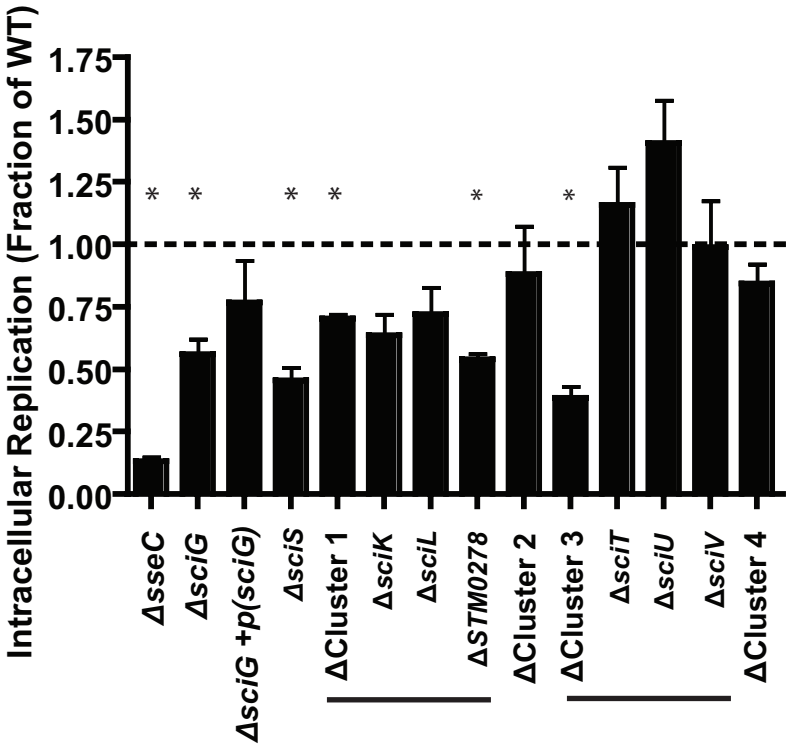
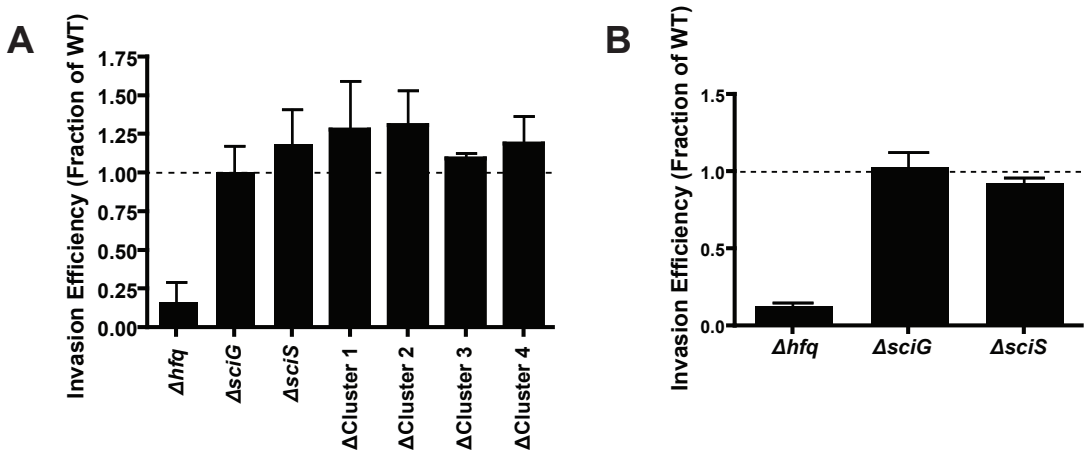

Intracellular Replication (Fraction of WT)
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
ΔsseC
ΔsciG
ΔsciG +p(sciG)
ΔsciS
ΔCluster 1
ΔsciK
ΔsciL
ΔSTM0278
ΔCluster 2
ΔCluster 3
ΔsciT
ΔsciU
ΔsciV
ΔCluster 4

Figure 4.4. Invasion efficiencies of SPI-6 T6SS mutants.

(A) In contrast with an *Δhfq* mutant that is deficient for invasion, core and non-core cluster mutants do not differ significantly from wild type in their ability to invade HeLa epithelial cells. (B) Similar results are observed for HEp-2 epithelial cells.

Figure 4.4.


A
Invasion Efficiency (Fraction of WT)
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Δhfq
ΔsciG
ΔsciS
ΔCluster 1
ΔCluster 2
ΔCluster 3
ΔCluster 4
B
Invasion Efficiency (Fraction of WT)
1.5
1.0
0.5
0.0
Δhfq
ΔsciG
ΔsciS

Hcp and VrgS protein expression is enhanced during infection of macrophages. Secretion of Hcp and VgrG orthologs is the hallmark of a functional T6SS. Hcp-1, Hcp-3 and VrgS were cloned and over-expressed as C-terminal FLAG fusion proteins and assayed for expression and secretion following IPTG induction in wild type and *sciG* mutant strains (Figure 4.5). While all proteins were observed in the cytoplasmic fraction, only Hcp-3 was found to be secreted. In addition it was determined that Hcp-3 secretion to the media was not dependent on the T6SS ATPase SciG. The Hcp proteins were observed at their predicted sizes of approximately 17 kDa while the VrgS protein was observed to migrate at approximately 70 kDa rather than the predicted 80 kDa based on the *S.* Typhimurium LT2 annotation suggesting a possible misannotated start codon. BlastP searches indicate low complexity sequences within the first 100 amino acids and this region falls outside conserved Vgr protein domains TIGR03361, COG3501, TIGR01646 identified by the Blast integrated conserved domain database (226). Interestingly, trace levels of DnaK were observed only in the secreted fraction of VrgS suggesting that over expression of VrgS may be slightly toxic.

In order to characterize the stable expression of SPI-6 T6SS proteins we constructed chromosomal hemagglutinin epitope fusion proteins for the Hcp paralogs SciK (Hcp-1), SciM (Hcp-2), and STM3131 (Hcp-3) and for the VgrG ortholog VrgS such that their expression was controlled by their native chromosomal promoters. During *in vitro* growth conditions, only Hcp-3 was detectable after 24 hours when bacteria were grown in rich (DMEM/10% FBS) or minimal media (LPM) (Figure 4.6A).

Some T6SS have been shown to require target cell contact for activation (213). To determine whether the absence of chromosomal expression of Hcp and VrgS proteins *in vitro* was due to the lack of induction by host cell signals, strains harboring chromosomally tagged genes were used to infect either RAW 264.7 murine-like macrophages or were grown in tissue culture medium in the absence of macrophages (Figure 4.6B). We observed only weak expression for Hcp-1, Hcp-2 and VrgS in the absence of macrophages whereas the abundance of these proteins was increased following macrophage infection. In contrast, Hcp-3 appears to be constitutively expressed under all conditions tested. To determine whether this induction is specific to phagocytic cells or the intracellular environment, HeLa cells were grown in the presence of wild type or invasion deficient *invA* mutant strains. No induction of Hcp-1, Hcp-2 or VrgS expression was observed after 24 hours (Figure 4.7). Hcp-3 was expressed in both backgrounds. It was noted that Hcp-1 appears to have stronger expression than Hcp-2 and VrgS under cell culture conditions both in the absence of macrophages and by wild type cells in the presence of HeLa cells.

A previous report had shown that transcript levels of the *icmF* ortholog, *sciS*, remain low until late time points (>16 hours) in macrophage infections (157). To determine whether Hcp and VrgS expression followed this pattern, we collected the contents of individual wells of infected RAW 264.7 cells at 4, 8, 16 and 24 hours post-infection. Hcp-3

expression was evident at early time points, while VrgS, Hcp-1 and Hcp-2 expression was weak. By 16 h, expression of Hcp-1, Hcp-2 and VrgS was observed (Figure 4.6C).

Figure 4.5. Expression of SPI-6 T6SS Hcp and VrgS.

Expression of plasmid-encoded Hcp and VrgS C-terminal FLAG epitope fusion proteins in the minimal media LPM following IPTG induction. (A) Expression of all proteins is observed in the cytoplasmic fraction. (B) Only Hcp-3::FLAG is observed in the secreted fraction. Expression of Hcp-3::FLAG is not dependent on the ATPase sciG.

Figure 4.5.

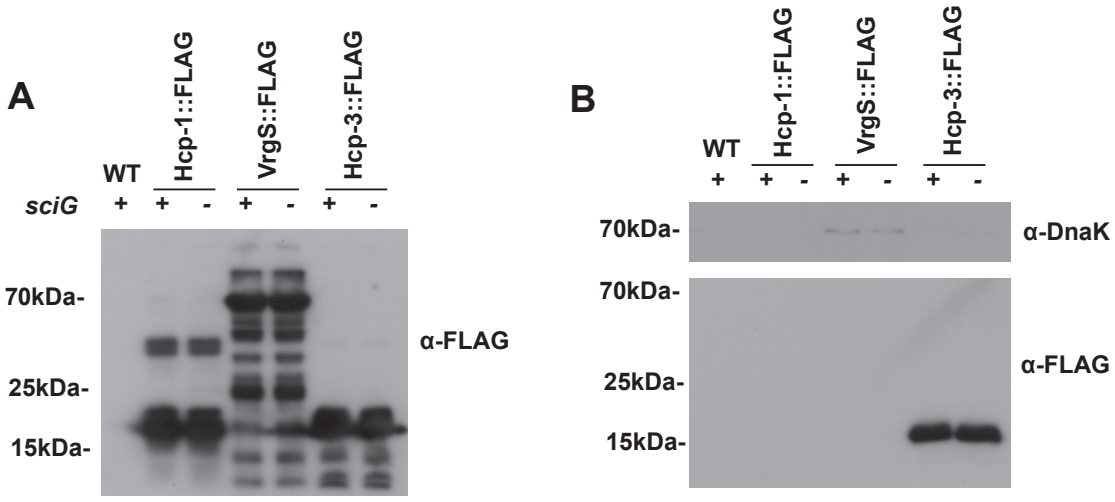

A
WT
Hcp-1::FLAG
VrgS::FLAG
Hcp-3::FLAG
sciG
+
+ -
+ -
+ -
70kDa-
25kDa-
15kDa-
α-FLAG
B
WT
Hcp-1::FLAG
VrgS::FLAG
Hcp-3::FLAG
+
+ -
+ -
+ -
70kDa-
α-DnaK
70kDa-
25kDa-
15kDa-
α-FLAG

Figure 4.6. Expression of Hcp and VrgS C-terminal chromosomal HA epitope fusion proteins.

(A) Following 24 hour growth in DMEM and the minimal media LPM, expression of only Hcp-3::HA is observed. (B) 24 hours post-infection of RAW264.7 macrophages with opsonized bacteria, Hcp-1::HA, Hcp-2::HA and VrgS::HA expression is enhanced compared with bacteria grown under identical cell culture conditions in the absence of eukaryotic cells. Expression of Hcp-3::HA is not enhanced. (C) Expression of Hcp-2::HA and VrgS::HA are detected by 16 hours post-infection of RAW 264.7 macrophages with opsonized bacteria.

Figure 4.6.

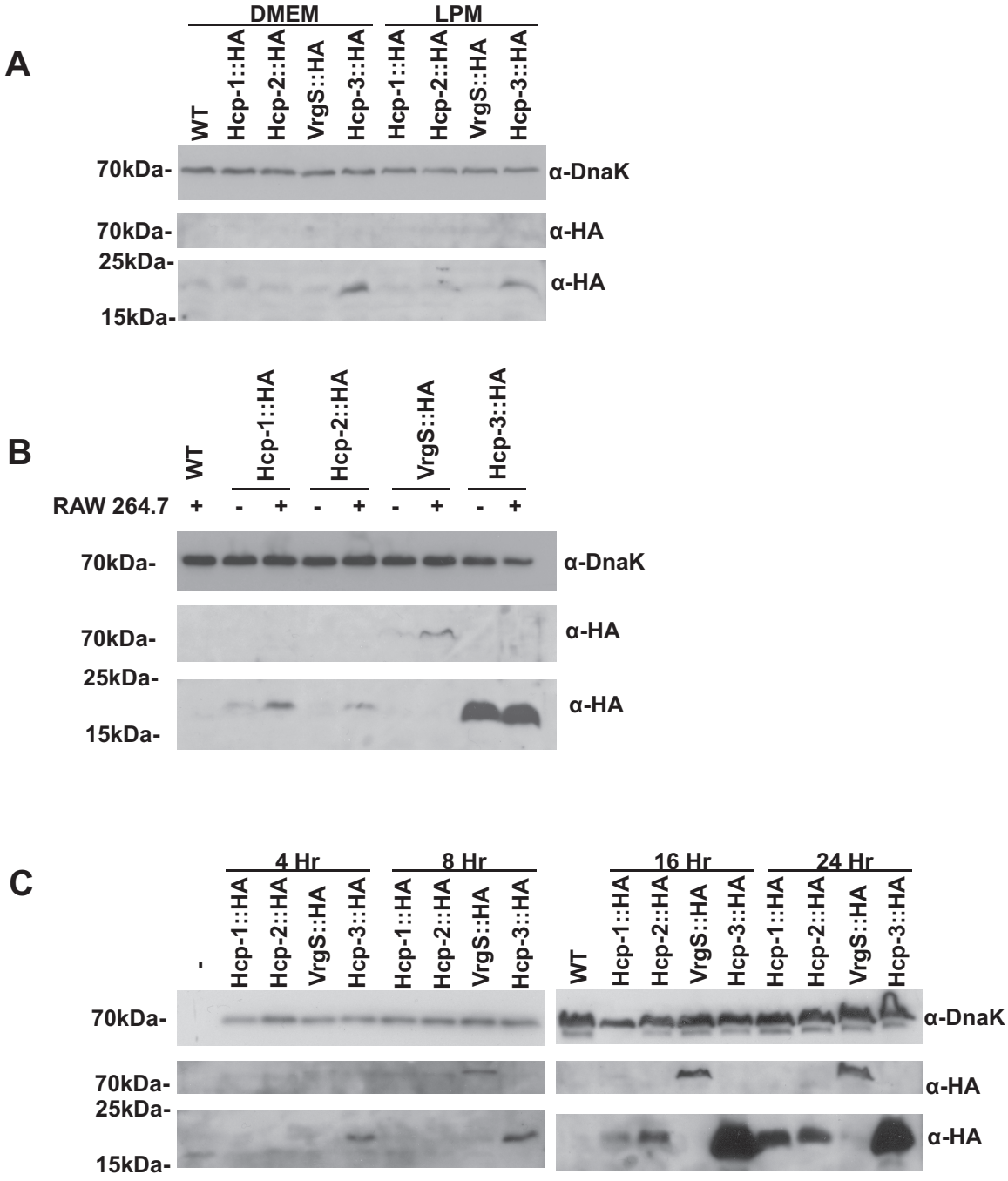
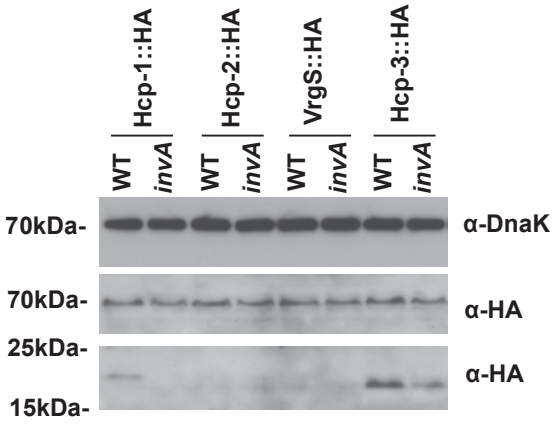

A
DMEM
LPM
WT
Hcp-1::HA
Hcp-2::HA
VrgS::HA
Hcp-3::HA
Hcp-1::HA
Hcp-2::HA
VrgS::HA
Hcp-3::HA
70kDa-
α-DnaK
70kDa-
α-HA
25kDa-
15kDa-
α-HA
B
WT
Hcp-1::HA
Hcp-2::HA
VrgS::HA
Hcp-3::HA
RAW 264.7 + - + - + - + - +
70kDa-
α-DnaK
70kDa-
α-HA
25kDa-
15kDa-
α-HA
C
4 Hr
8 Hr
16 Hr
24 Hr
-
Hcp-1::HA
Hcp-2::HA
VrgS::HA
Hcp-3::HA
Hcp-1::HA
Hcp-2::HA
VrgS::HA
Hcp-3::HA
WT
Hcp-1::HA
Hcp-2::HA
VrgS::HA
Hcp-3::HA
Hcp-1::HA
Hcp-2::HA
VrgS::HA
Hcp-3::HA
70kDa-
α-DnaK
70kDa-
α-HA
25kDa-
15kDa-
α-HA

Figure 4.7. Expression of SPI-6 T6SS Hcp and VrgS in co-culture.

Expression of chromosomal-encoded Hcp and VrgS C-terminal HA epitope fusion proteins after 24 hours co-incubation with HeLa cells. Expression of Hcp-3::HA is observed by both wild type and invasion deficient bacteria. Hcp-2::HA and VrgS::HA expression is not observed by either wild type or invasion deficient bacteria. Expression of Hcp-1::HA is observed only by the wild type cells.

Figure 4.7.


Hcp-1::HA
Hcp-2::HA
VrgS::HA
Hcp-3::HA
WT
invA
WT
invA
WT
invA
WT
invA
70kDa-
α-DnaK
70kDa-
α-HA
25kDa-
15kDa-
α-HA

The SPI-6 T6SS is not induced by deletion of regulators of intracellular virulence.

A previous report suggested that the SPI-6 T6SS is induced in response to host cell signaling as a result of regulatory de-repression, possibly through the transcription factor SsrB (157). In order to further investigate regulatory contributions we identified transcriptionally active regions by cloning intergenic regions greater than 40 base pairs in the SPI-6 T6SS as transcriptional fusions to luciferase. Measurement of luminescence normalized to the number of recovered bacterial cells was performed during RAW 264.7 macrophage cell infections to assess transcriptional activity (Figure 4.8A). At 24 hours post infection, active transcription was observed from regions upstream of the six genes, *sciF*, *sciG*, *sciH*, *sciL*, *sciR*, and *vrgS*. Transcription from regions upstream of *sciF*, *sciL*, and *sciR* were much lower than those of *sciG*, *sciH* and *vrgS* during exponential growth *in vitro* (Figure 4.9).

To determine if these transcriptionally active regions were targets for regulatory input from SsrB, *lacZ* transcriptional reporters were chromosomally integrated downstream of these regions to generate merodiploid reporter strains. The transcriptional activity of these reporters was measured *in vitro* in the defined minimal low phosphate and magnesium medium LPM that provides an environment in which SsrB, and upstream regulators of virulence genes, PhoPQ and SlyA, are active (177). We assessed transcriptional activities of our reporters in backgrounds deficient for these proteins involved in intracellular pathogenesis (175). We observed altered profiles for transcriptionally active regions upstream of *sciF, sciG,* and *vrgS* (Figure 4.8B). In an *ssrAB* mutant background,

transcriptional activity for the regions upstream of *sciF* and vrgS were increased slightly by less than two-fold. Transcriptional profiles in *phoP* and *slyA* mutant backgrounds were identical and revealed an increase in transcription for the region upstream of *sciF* by approximately 2.5 fold and a two-fold decrease for the regions upstream of *sciG* and *vrgS*.

To determine whether expression of the SPI-6 T6SS is repressed by the activity of regulators of intracellular virulence, we assessed levels of the chromosomal fusion protein VrgS::HA expressed under its native promoter in backgrounds deficient for regulators of intracellular pathogenesis (175). These included the two component systems PhoPQ, OmpR/EnvZ, QseBC, PmrAB, SsrAB and the RNA binding protein Hfq, regulators important for intracellular pathogenesis (Figure 4.8C). Expression of VrgS in these mutant backgrounds in LB media after 24 hours remained barely detectable using an enhanced sensitivity chemiluminescence system although expression appeared slightly higher in an *hfq* background. When expression of Hcp-1, Hcp-2, and VrgS were assessed after 24 hours growth in DMEM under cell culture conditions we failed to see an increase in expression for either *ssrB* or *hfq* background strains over wild type (Figure 4.8D). Expression of these proteins was weaker during growth in LPM and also failed to show increased expression over wild type (data not shown).

A PmrA box motif was previously identified upstream of the orphan Hcp-3 gene, *STM3131*, although the regulatory input was not confirmed (227). In order to determine whether a regulatory interaction for PmrA occurs at this site we generated a *lacZ*

transcriptional fusion on the chromosome downstream from *STM3131*. We also generated *lacZ* transcriptional fusions on the chromosome for the PmrA activated promoter of *pmrC* and the PmrA repressed promoter of *pmrD*. Assessment of transcriptional activity indicated that transcription increased two-fold in a *pmrA* mutant background, similar to that of the PmrA repressed promoter of *pmrD*, suggesting that PmrA acts as a transcriptional repressor of this gene (Figure 4.9).

Figure 4.8. Effect of regulator mutations on SPI-6 T6SS gene expression.

(A) Activities of nine SPI-6 T6SS intergenic regions directly upstream of the indicated genes that were cloned as transcriptional fusions with the *luxCDABE* genes. Luminescence data was collected as relative light units (RLU) and normalized to colony forming units (CFU). At 24 hours post infection of RAW264.7 macrophages, six of the nine intergenic regions were transcriptionally active. The constitutive synthetic promoter *em7*, and the SsrB dependent promoter of *ssaG* are also shown as controls. (B) Activities of the six transcriptionally active intergenic regions when cloned as *lacZ* chromosomal transcriptional fusion reporters in regulator mutant backgrounds and measured during log phase growth in LPM (*, $P<0.05$). Luminescence data was collected as RLU and normalized to OD_{600}. Transcriptional activity of the region upstream of *sciF* is increased in *phoP*/*slyA/ssrAB* mutant backgrounds and that of *vrgS* is increased only in an *ssrAB* mutant background. Those of *sciG* and VrgS are decreased in *phoP/slyA* mutant backgrounds. (C) Expression of VrgS::HA protein assessed by Western blot. Trace levels of VrgS::HA fusion protein expressed under native promoter control on the chromosome can be detected after 24 hours growth in LB media and expression is not strongly upregulated in *phoP*, *ssrAB*, *qseC*, *ompR*, *hfq* mutant backgrounds. (D) Expression of native promoter expressed Hcp-1::HA, Hcp-2::HA, and VrgS::HA chromosomal fusion proteins assessed by Western blot in *ssrAB* and *hfq* mutant backgrounds after 24 hours growth in DMEM/10% FBS under cell culture conditions. No significant difference in expression is observed.

Figure 4.8.

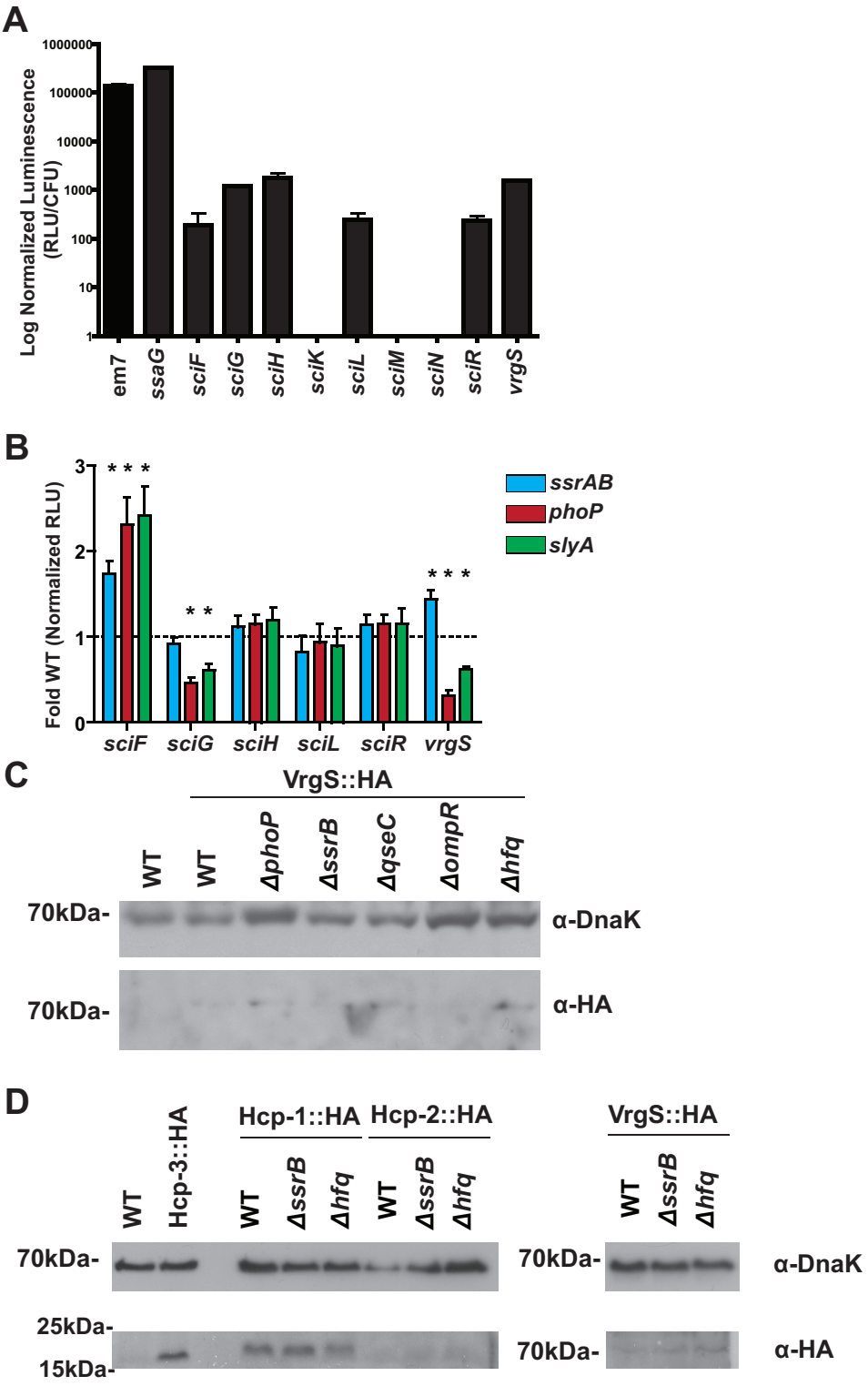
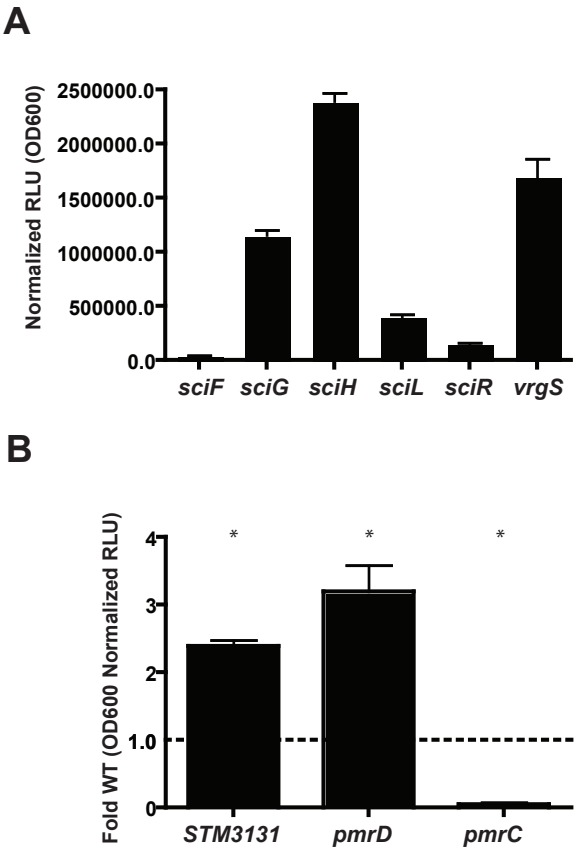

A
Log Normalized Luminescence (RLU/CFU)
1000000
100000
10000
1000
100
10
1
em7
ssaG
sciF
sciG
sciH
sciK
sciL
sciM
sciN
sciR
vrgS
B
Fold WT (Normalized RLU)
ssrAB
phoP
slyA
sciF
sciG
sciH
sciL
sciR
vrgS
C
VrgS::HA
WT
WT
ΔphoP
ΔssrB
ΔqseC
ΔompR
Δhfq
70kDa-
α-DnaK
70kDa-
α-HA
D
Hcp-3::HA
Hcp-1::HA
Hcp-2::HA
VrgS::HA
WT
ΔssrB
Δhfq
70kDa-
25kDa-
15kDa-
α-DnaK
α-HA

Figure 4.9. Transcriptional activity of SPI-6 T6SS promoters.

Transcriptional activities of the intergenic regions upstream of the indicated genes as transcriptional fusions to the *lacZ* reporter gene in merodiploid strains or to the *luxCDABE* operon on the reporter plasmid pGEN. (A) Transcriptional activity of luciferase reporters during exponential growth in the minimal media LPM at 37C with shaking at 225RPM. (B) Transcriptional activity from the region upstream of *STM3131* as a *lacZ* reporter fusion is increased in a *ΔpmrA* mutant in a similar manner to that of the PmrA activated gene *pmrD* unlike the PmrA repressed gene *pmrC* (*, $P<0.05$).

Figure 4.9.


A
Normalized RLU (OD600)
2500000.0
2000000.0
1500000.0
1000000.0
500000.0
0.0
sciF
sciG
sciH
sciL
sciR
vrgS
B
Fold WT (OD600 Normalized RLU)
4
3
2
1.0
0
*
*
*
STM3131
pmrD
pmrC

Discussion:

T6SS variants have been acquired at least three times within the *Salmonella* lineage with little being known about their function and mechanism, although most reports have implicated roles for the *S.* Typhimurium SPI-6 T6SS in pathogenesis of mice and infection of macrophages (19, 21, 26, 28, 157, 207, 208, 228). We found that disruption of the non-core T6SS clusters 2 and 4 caused significant defects in systemic dissemination in mice and that disruption of non-core gene clusters 1 and 3 resulted in a significant intracellular replication defect in macrophages. Further supporting this role in intracellular pathogenesis we showed an increase in Hcp and VrgS protein expression in association with macrophages. Finally, we showed that deletion of a previously identified negative regulator of this system enhances levels of transcription of this system from defined transcriptionally active regions but does not lead to increased levels of Hcp and VrgS protein expression upon deletion.

Distinct T6SS are encoded by *S. bongori, S. enterica* subspecies *arizonae*, and serovars of *S. enterica* subspecies *enterica* as the most proximal element to the tRNA genes within SPI-6 or SPI-2 suggesting relatively recent acquisition events (26, 28). In spite of the number of T6SS in *Salmonella*, an evolved VgrG effector gene has only been identified in *S. enterica* subspecies *arizonae* suggesting that if the *S.* Typhimurium SPI-6 T6SS functions to deliver effector proteins, they are not encoded as evolved VgrG genes (28). The SPI-6 T6SS closely resembles the *Burkholderia* Tss3 (224), present in *B. mallei* and

B. pseudomallei, but lacks a reported role in pathogenesis and is absent in *B. thailandensis* (229). In order to understand the role of the SPI-6 T6SS we focused on non-core T6SS genes. Non-core T6SS gene transposon mutants of *S.* Typhimurium have been found to be defective in long term persistence in mice and macrophages; these include *sciR* in cluster 2 (208); *sciU* in cluster 3 (208); *sciW*, *rhs1*, and *STM0298* in cluster 4 suggesting that these genes may encode proteins important for T6SS activity (208–210). Using a blast-based analysis we found that some of these genes are restricted to *Salmonella* while others exist as gene pairs in other T6SS-encoding organisms. The gene pair *sciL-STM0278* in cluster 1 is extensively conserved and is a particularly interesting candidate for further investigation. Cluster 2 was found to be poorly conserved and exists as a conserved gene pair in only one other non-*Salmonella* organism. The genes *sciU-sciV* also exist as a conserved pair in a limited number of organisms although the association with T6SS is not as strong as that of *sciL-STM0278*. Many orthologs were identified for genes within cluster 4, including for the gene pair *sciW-rhs1,* however the extensive conservation of *rhs*-associated genes and their close similarity made drawing conclusions about these genes difficult. Evidence of conservation of gene pairs within each of these clusters is particularly interesting given the presence of toxin-antitoxin gene pairs encoded in other T6SS encoding organisms (152).

In order to quantify the contribution of the SPI-6 T6SS in host pathogenesis we assessed systemic dissemination of core and non-core mutants in a murine mouse model of typhoid and found significant contributions for non-core gene clusters 2 and 4. Deletion of

individual genes within non-core gene cluster 2 did not decrease systemic dissemination to the same extent as deletion of the entire cluster suggesting that the transcriptionally active region upstream of *sciR* may be a contributing element to pathogenesis. Interestingly, the orthologous gene pair of *sciQ-sciR* in *A. tumefaciens* is in a different orientation and lacks this intergenic region. In the deletions within cluster 3, individual mutants in *sciT* and *sciV* revealed decreased fitness and a trend towards increased fitness for the *sciU* mutant. Deletion of all three genes may result in either the increased or decreased fitness *in vivo* depending on uncontrolled host factors. The *sciU* and *sciV* genes may function as a gene pair given that these two genes exist as an orthologous gene pair in the absence of *sciT* in four other Proteobacteria member genomes. Deletion of cluster 4 had the greatest effect on systemic dissemination of all clusters tested indicating that it may encode an element essential for T6SS-associated pathogenesis *in vivo*. This cluster encodes recombination hotspot (Rhs) family genes and Rhs-associated genes that are commonly linked with T6SS and which are extensively distributed in Gram-negative Proteobacteria (230). We disrupted *sciW, rhs1* and *rhs2* and found that *rhs2* had the same defect on systemic dissemination in mice as disruption of core T6SS genes and the entire SPI-6 locus (207). This Rhs element is encoded with *sciX* in a bicistronic operon and may also function as a toxin/immunity gene pair participating in contact dependent growth inhibition (231).

Survival and replication within host cells permits systemic dissemination *in vivo* and disruption of the SPI-6 T6SS resulted in reduced intracellular replication in macrophages.

In contrast with systemic dissemination, we found that only non-core gene cluster mutants 1 and 3 had significant effects on intracellular replication in cell culture. Deletion of individual genes within cluster 1 revealed a significant replication defect for *STM0278*. Like *rhs2-sciX*, this gene is also part of a bicistronic gene pair with *sciL* and may also function as a toxin/immunity gene pair. We found that orthologs of these genes are present in close association with T6SS loci in a number of other members of the family Enterobactericeae. Given that this gene pair is the most strongly conserved of all non-core T6SS associated genes in clusters 1-3 we are further pursuing their characterization. Deletion of individual genes within cluster 3 were not significantly attenuated, and in fact showed a non-significant increase in intracellular replication. Interestingly, the gene mutant *sciU* within cluster 3 showed a trend towards increased intracellular replication and systemic dissemination upon deletion. These genes may have redundant functions, or attenuation in the non-core gene cluster 3 mutant may be due to polar effects of the disruption. The difference in mutant fitness between *in vivo* systemic dissemination and cell culture intracellular replication experiments may be explained by differences in the length of experiment and cell types encountered as assessment of intracellular replication in cell culture is a simplified model of one aspect of systemic dissemination. Oral infection of mice with *S.* Typhimurium involves passage through the gut environment before systemic dissemination and it is possible that genes in clusters 2 and 4 may be involved in interbacterial competition in this milieu. These multiple gene sets may act to repurpose the general type 6 secretory apparatus for delivery of effector proteins to different cell targets. While other T6SS such as that of *Pseudomonas aeruginosa* and *V.*

cholerae have been found to have antibacterial properties we were unable to observe this phenomenon for the *S. enterica* T6SS using solid LB agar-based competition assays against *E. coli* (data not shown) (152, 232). In addition, a previous report found that deletion of the entire SCI genomic island resulted in a 50% reduction of HEp-2 epithelial cells but we were unable to observe an effect on invasion of both HeLa and HEp-2 epithelial cells for our *sciG* and *sciS* mutants suggesting that gut epithelial cells may not be the target for this system and that this reduction in invasion may be due to a different SCI-encoded factor (156). Interestingly, *sciL* (cluster 1) and *rhs2 and sciX* (cluster 4) were found to be upregulated in macrophages and not HeLa epithelial cells (211). *SciL* is encoded in a bicistronic operon with *STM0278* in cluster 1 which had the most notable contribution to intracellular replication in macrophages. *rhs2* and *sciX* are encoded as a bicistronic operon in cluster 4, and *rhs2* had the most notable contribution to systemic dissemination in mice. Additionally, the *sciL*, *STM0278*, *rhs1* and *rhs2* genes have the largest number of orthologs of all genes in clusters 1-4.

Given the number of environments encountered by *S.* Typhimurium, virulence systems require regulatory control to achieve appropriate situational expression. Arguing against a general role for the SPI-6 T6SS, we observed only weak chromosomal expression and we did not observe plasmid-based secretion of Hcp-1, Hcp-2 or VrgS following growth under *in vitro* conditions after 24 hours. In contrast, Hcp-3 appears to be constitutively expressed, negatively regulated by PmrA and secreted under these conditions however the secretion of Hcp-3 is curiously not dependent on *sciG.* The ClpV ATPase is a core T6SS

gene and its abrogation results in the inability of the T6SS to assemble and secrete Hcp in *V. cholerae* (225). Hcp-3 differs significantly in coding sequence from Hcp-1 and Hcp-2, more closely resembling the Hcp proteins of *S. bongori* and *Pseudomonas* and may have evolved to perform alternative functions as is the case for the fourth *S. enterica* Hcp encoding gene paralog *hilE* (156). Hcp-3 may therefore have an unrelated function and may be accessing the extracellular environment through an alternate secretory pathway. *SciG* may be dispensable for functional secretion in *S.* Typhimurium however, the equivalent defect observed for *sciG* and *sciS* in both systemic dissemination in mice and intracellular replication in macrophages makes this unlikely. The expression of Hcp-1 and Hcp-2 are not identical as we observed expression of Hcp-1 in the absence of macrophages in DMEM media, in contrast with Hcp-2 and VrgS, the expression of which is barely detectable. Hcp-1 is encoded by *sciK* within non-core gene cluster 1 and its expression is perhaps a reflection of promoter changes during its acquisition or duplication from *sciM*.

The absence of expression and secretion of Hcp-2 and VrgS proteins suggests a system under tight regulatory control. Expression of *S.* Typhimurium SPI-2 T3SS virulence genes important for replication and survival in macrophages are upregulated upon exposure to the intracellular compartment (61). In a similar manner, we found that the Hcp and VrgS genes are differentially expressed in the presence or absence of macrophages and between *in vitro* and tissue culture conditions. Hcp-1 is encoded within non-core gene cluster 1, and like Hcp-2 and VrgS, expression was clearly enhanced

during late stages of macrophage infections. The difference in expression between *in vitro* and cell culture conditions for these proteins suggests an activating signal specific to growth in the presence of macrophages but may also be in response to general cell stress. Expression of the SPI-6 T6SS may rely on regulatory inputs from the host environment in a similar manner to the T6SS in *Vibrio cholerae* that requires internalization of bacteria (213). With regard to *S.* Typhimurium, expression of the T6SS does not appear to occur until after 8 hours of infection and has been previously hypothesized to be a result of the loss of repression through an SsrB dependent regulatory pathway. It was hypothesized that induction may be relevant to host cell death and bacterial escape (157). Another group reported upregulation of some SPI-6 T6SS genes during macrophage infection but did not observe an increase over time (211). It is possible that this induction relies on the onset of host cell death which can be affected by differing multiplicities of infection and experimental conditions. Indeed, when MOI values were much reduced we observed that Hcp and VrgS expression was lower (data not shown). An important consideration is that expression levels of this system, even in the context of macrophage infection, remain low and we have not yet been able to observe translocation of tagged proteins to the host cells and therefore further work will be necessary to determine whether this system is activated within the intracellular compartment or activated simply in the presence of macrophages.

Induction of the SPI-6 T6SS could be similar to that of the HSI-I T6SS in *Pseudomonas spp.* where deletion of the sensor kinase RetS leads to enhanced secretory activity of effector proteins (152). We pursued this hypothesis by identifying transcriptionally active

regions within the SPI-6 T6SS. We found six transcriptionally active regions including one within the non-core gene cluster 2, active *in vitro* and in association with host macrophages. A previous report observed increased numbers of transcripts of *sciS* in *ssrB* null mutants during late stages of host cell infection (157). We also found increased levels of transcription for the T6SS in an *ssrB* mutant background, with increased transcriptional activity from regions upstream of *sciF* and *vrgS*. The regulatory proteins PhoP and SlyA are involved in activation of SsrB (103, 106) and we found that transcriptional activity upstream of *sciF* also increases in this mutant background, however activity for *sciG* and *vrgS* decrease suggesting additional regulatory inputs to this system between PhoP and SsrB. Transcriptional profiles for all six reporters in both *phoP* and *slyA* mutant backgrounds were identical suggesting that they act through a common regulatory input. We previously reported the genome-wide interaction map of SsrB using chromatin immunoprecipitation and found only two interaction sites within the SPI-6 T6SS, upstream of *sciF* and *sciR* (233). Given our observation of increased transcriptional activity upstream of *sciF* and the previously reported increase in *sciS* transcripts in a *ssrB* mutant background, we asked whether the loss of these regulatory inputs have a significant effect on VrgS protein expression through regulatory deregulation. Other regulatory proteins aside from PhoP-PhoQ, SsrA-SsrB and SlyA are necessary for intracellular survival during host infection (175). VrgS levels were not affected in *phoP*, *ssrB*, qseC, *ompR,* or *hfq* mutant backgrounds although disruption of these proteins results in extensive changes to *S.* Typhimurium gene expression and leads to abrogation of pathogenesis (175). While not extensive, we believe that these results argue for the

presence of an activating factor that controls T6SS upregulation in response to inducing signals. An unbiased transposon screening approach will likely be necessary to identify such a regulator.

In summary, these findings further support a role for the SPI-6 T6SS during host infection, likely mediated through interactions with macrophages and enabled by the SPI-6 T6SS encoded non-core genes *STM0278* and *rhs2*. Two transcriptionally active regions of this system receive negative regulatory input through the SsrB regulatory network however expression of this system appears to be controlled by a yet unknown regulatory factor.

Acknowledgements:

This work was funded by an operating grant from the Canadian Institutes of Health Research (MOP 82704), an infrastructure grant from the Canada Foundation for Innovation and the Canada Research Chairs Program from the Government of Canada. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. DTM is supported by a Canada Graduate Scholarship from the Canadian Institutes of Health Research. CAC was the recipient of a Canada Graduate Scholarship from the Natural Sciences and Engineering Research Council of Canada. BKC is a Canadian Institutes of Health Research New Investigator and the Canada Research Chair in Infectious Disease Pathogenesis.

Chapter V - Discussion

Chapter V – Discussion

The preceding chapters described advances in understanding the structure and function of the two-component regulatory system SsrA-SsrB as well as the genes regulated by SsrB that contribute to systemic dissemination *in vivo* in mice and *in vitro* in cultured cells. The following discussion places this work in the context of the field and proposes future avenues of investigation for these areas of research.

Characterization of the Sensor Kinase SsrA

Chapter 2 discussed the two-component regulatory system sensor kinase SsrA. It has been established that acidification of the intracellular vacuolar environment is essential for correct maturation of the phagolysosome as well as full induction of virulence in *S. enterica* (49). This intracellular host cue is integrated by multiple sensors, including PhoQ and PmrB (76, 234). Preliminary work in the field suggested that SsrA also integrates a pH cue during induction *in vitro* and *in vivo,* however the mechanism of this process was unknown (61, 99, 115). In this work we showed that pH sensing is dependent on histidine residues within loop regions of the membrane-proximal PDC domain.

A mechanism of integrating input through compounding conformational change could be at work in sensing pH by SsrA. Six histidine residues distributed throughout flexible loop regions drive two conformational changes in the diphtheria toxin T domain of *Corynebacterium diptheriae* in response to acidification of the vacuole (235). In

addition, histidine residues have been implicated in pH sensing by the sensor kinase ArsS of *Helicobacter pylori* (236). Histidine residues are also used as biological switches in the sensor kinases PmrB and PhoQ (76, 234). Only in the case of the diptheria toxin T domain have structural methods been used to understand the effects of histidine protonation on overall structure, where it was found that at least three such events were necessary to generate sufficient conformational strain to drive large scale structural transition to an alternate form (235). Further structural investigation of SsrA under various pH conditions could determine whether strain is generated by protonation of the imidazole rings and whether this strain is resolved through adoption of an alternate structural conformation. Preliminary investigation using alternate histidine mutants suggests that each histidine protonation state is unique in the role it plays in stabilizing the two states (Figure A1.5). We used a structural model to interpret our data; however, there are regions that are of low confidence, such as the histidine containing loop regions and the membrane-distal PDC domain. Purification of SsrA and elucidication of the SsrA structure would therefore provide great insight for further investigation. Preliminary work with various truncated forms of SsrA showed it is not sufficiently soluble for purification due to a tendency to aggregate in inclusion bodies.

Recent work on PhoQ showed that sensor kinases can be tuned to environmental signatures that consist of multiple cues rather than a single ligand (75–77). While a pH cue is clearly integrated by SsrA, it also modulates a basal level of signal, the source of which is unknown. We showed that a mutant defective in pH sensing had similar levels of activity under acidic and neutral pH conditions however this mutant was highly activated

under otherwise inducing conditions compared to the null *ssrA* mutant strain. We hypothesize that this pH-independent activation is a result of activation of the membrane distal domain. This region encodes a ligand-binding pocket and specificity-determining major and minor binding loops which are highly conserved amongst related orthologs (74). We showed that key residues within the ligand binding pocket are conserved and that those residues within the major and minor ligand binding loops that confer ligand specificity in other orthologs are conserved within *Salmonella.* However, they are distinct from those in orther orthologs, suggesting that SsrA recognizes a specific ligand. In contrast with SsrA, the pH-sensing histidine residue of PhoQ is positioned within the membrane-distal domain while primary antimicrobial peptide signaling occurs by way of the membrane-proximal domain (237). Therefore the double PDC domain scaffold appears to be a useful starting point for environmental sensing on which selection for various cues can act.

If such a putative SsrA ligand exists, then it would be expected to be present within the intracellular environment of the SCV. The overall trend for non-critical residues mutagenized suggests that destabilizing the SsrA structure drives pH-independent activation. In addition, lowering concentrations *in vitro* of divalent cations leads to increased SsrA-dependent signaling, suggesting that a specific signal for SsrA *in vivo* may function to drive signaling by way of promoting an alternate conformational state (Figure A1.2). Limiting concentrations of divalent cations *in vitro* and antimicrobial peptides *in vivo* both destabilize the sensor platform and drive activation of the sensor kinase PhoQ (77). Alternatively, such a ligand may not drive transition to an alternate

conformational state, but instead may simply stabilize this alternate conformation. An answer to this question requires identification of the ligand, which may be possible using a high throughput screen where a pH-insensitive SsrA mutant strain is exposed to a panel of small molecules.

An additional question that remains unresolved is whether dimerization of SsrA has a role in function. Characterized sensor kinases require dimerization in order to undergo autotransphosphorylation (79). Similar sensor kinases have been crystalized in a dimeric form where the monomers interact along the length of the N-terminal alpha helical stalk, burying this stalk in the center, with hydrophobic contacts along the periphery of the PDC domains (73). This configuration both buries exposed hydrophobic regions and brings the four transmembrane helices into position to form a trans-membrane coiled-coil helical bundle. We found evidence to support such a structure – for SsrA – specifically, conserved hydrophobic residues in otherwise solvent-exposed regions of the alpha-helical stalk. The thermodynamic implications of such residues are that SsrA would be predicted to form a dimer under all conditions. This idea could be assayed *in vitro* through the use of FRET, where two SsrA monomers with translationally fused FRET compatible fluorophores are co-expressed (238). Co-expression of these constructs should result in FRET under all conditions, or only under SsrA inducing conditions if dimerization only occurs under such conditions. If dimerization is required for activity and oligomerization only occurs in the activated state then this may present an additional level of regulation for SsrA.

A topic not addressed in the literature is whether SsrA has SsrB-specific phosphatase activity and whether this activity is constitutive or regulated by the N-terminal sensor platform. Most work on sensor kinase phosphatase activity has occurred *in vitro* and debate continues about the importance of this activity *in vivo* (239). However, a series of mutants could be used to indirectly examine phosphatase activity *in vivo*, particularly a phosphorylation defective SsrA mutant D739E. SsrB activity would be predicted to be lower in the presence of this D739E mutant than in the absence of SsrA entirely. SsrA phosphatase activity might explain the constitutive activity of SsrB when overexpressed. Small molecule phosphodonors, and a long half-life of phosphorylated SsrB, or weak autodephosphorylation activity may necessitate phosphatase activity by SsrA to minimize SsrB activity (81). This might also explain the added level of regulation of SsrB made possible by the *ssrB* promoter (105). In any case, while multiple groups have used genetic tools to show the connection between SsrA and SsrB, specific interaction remains to be demonstrated *in vitro* using purified protein to demonstrate either phosphorylation or phosphatase activity.

In summary, this work provides an additional system where a simple environmental trigger can be used to study the relationship between protein conformation and function, and how alternative protein states are stabilized and maintained *in vivo*.

Characterization of the Response Regulator SsrB

Chapter 3 described identification of the regulon and DNA-binding sequence of the two-component regulatory system response regulator SsrB. Previous work by

numerous groups identified individual members of the SsrB regulon through usage of a *ssrA* DNA microarray (113), DNase I footprinting (100, 105, 109), and transcriptional reporters (118). Since the complete data set for the former microarray was not publically available, and due to the possibility that SsrA activated other downstream proteins, we used a DNA microarray to determine the complete regulon for SsrB. We validated this set of genes with those members of the regulon for which there is transcriptional evidence as well as generated a list of genes for future investigation for which traditional roles in virulence are not immediately evident. Traditionally, genes that appear to play a role in metabolism have been ignored; however, recent studies showed that *Salmonella* can use the inflammatory environment to generate nutrient sources during infection (240). One SsrB regulated protein has been recently found to contribute to intracellular virulence by directing transport of D-alanine within the SCV(241). Therefore, identification of these genes unexpectedly regulated by SsrB may present an opportunity to discover other novel strategies of host exploitation by a pathogen.

The DNA-binding motif was configured as a heptameric palindrome, suggesting that the functional state of SsrB is that of a dimer, despite the DNA-binding C-terminal truncated monomeric form being constitutively active (116). We found that the *sseA* promoter has only one well defined heptamer although other reports have described an approximate 17bp DNAaseI footprint (100). This finding suggests that in the presence of one well-defined heptamer, the second monomer has high tolerance for sequence degeneracy. Transcriptional reporter experiments where portions of the SsrB motif were scrambled in an attempt to address this question supported this tolerance of degeneracy;

however, strategies that more directly assess the interaction between DNA and protein such as electromobility shift assays (EMSA) would be more useful in addressing this question. Another question relating to the oligomeric state of SsrB is whether dimerization follows recruitment to the promoter, phosphorylation or whether it exists as a dimer in both active and inactive conformations (80, 81). The oligomeric state of SsrB could be assayed in a similar manner to that of SsrA by employing FRET and fluorophore tagged constructs (242). A D56A mutant unable to be phosphorylated and a constitutively active D56E mutant could be used as controls to explore this question. They have been generated and verified for this purpose.

Previously, a cryptic promoter was identified upstream of *ssrB* demonstrating that transcription of *ssrA* and *ssrB* was uncoupled (105). We discovered and characterized an additional promoter upstream of *ssaR*. In addition, we observed a strong interaction site upstream of *ssaE* which was found to contribute to transcriptional activation of downstream genes. These additional promoters provide added sites at which spatiotemporal regulation might occur in order to properly assemble the T3SS apparatus (123). Additionally, many interaction sites for SsrB were found throughout the genome, both within promoters and within the coding regions of genes. The latter set presents an interesting question as to whether these sites confer positive or negative regulatory effects. An example of such a negative regulatory interaction is observed upstream of a SPI-6 T6SS gene, the expression of which is increased in the absence of SsrB (157).

During our investigation we used sequence alignments between promoter regions of SPI-2 SsrB-dependent genes with those of SSR-3, a genomic island of *Sodalis*

glossinidius encoding orthologs of the former island with identical gene synteny but lacking *sseF* and *sseG* (185). *S. enterica* and *S. glossinidius* are of sufficient evolutionary distance that sequence drift has occurred at those positions under neutral selection, revealing positions conserved under positive selection by sequence alignments. Sequences other than the SsrB DNA-binding motif in the promoters of SsrB-dependent genes were also conserved, notably a stretch within the promoter of *ssaG* suggesting that this strategy may be useful to identify additional regulatory elements, aside from SsrB, for SPI-2 promoters not yet identified (64, 96). Experiments focusing on these elements would not only provide insight in to the regulation of SPI-2 genes in *S. enterica* but also SSR-3 genes in *S. glossinidius*, an organism which encodes two additional T3SS but which appears to be a transitional obligate endosymbiont of an insect fly host (185, 243).

An interesting observation made from this work was that the SsrB motif identified by one-hybrid analysis, the one within SPI-2 and the one outside of SPI-2 were not identical in their degeneracies. This variation may reflect different conditions under which these sequences evolved. The first was selected *in vitro* for high expression to rescue an auxotrophic strain. The second was evolved in an unknown organism and acquired as part of SPI-2, and the third was evolved in *Salmonella*. This provides an opportunity to investigate how constraints on a system feed into regulatory evolution.

One of the SsrB recognition sites identified in this work was for a SsrB-regulated promoter upstream of a gene ancestral to both *S. enterica* and *S. bongori*, which was found to have evolved to be regulated by SsrB following acquisition of SPI-2 (33). The modular nature of SsrA-SsrB and its T3SS therefore provides an exciting opportunity to

study the global effects on a genome as a result of acquisition of a regulatory protein. The genomes of *S. enterica* and *S. bongori* are highly similar aside from differences in horizontal acquisitions, notably at SPI-2 where a T3SS is present in the former and a T6SS is present in the latter (26). An evolved genome can therefore be compared to a naïve genome with respect to the presence of SsrB, which allows for identification of ancestral gene promoters that have undergone regulatory evolution as a result of SsrA-SsrB acquisition. The extensive number of SsrB interaction sites throughout the genome in *S. enterica* suggests that SsrB could have recruited RNA polymerase sufficiently close to many transcriptional initiation sites following acquisition and expression. Whole transcriptome sequencing (RNAseq) of *S. enterica* (*ssrAB* mutant) and *S. bongori* strains expressing SsrB or an empty plasmid would allow characterization of all these SsrB-dependent transcripts. A first internal comparison (SsrB + : SsrB -) would identify genes under the control of SsrB and a second comparison (SsrB evolved : SsrB naïve) would identify ancestral genes that have either gained or lost SsrB regulation by selection. While one ancestral gene's promoter that has undergone regulatory evolution has been identified in *Salmonella*, gene promoters that have undergone negative selection following acquisition of a novel regulatory protein have not been identified or characterized in *S. enterica* nor has it been a topic in the literature in general.

In summary, with the knowledge of SsrB and genetic tools at our disposal, this unique system in *S. enterica* provides an opportunity to learn about regulatory evolution which could be applied to other organisms and particularly in the context of how

horizontally acquired genes such as antibiotic resistance genes are integrated in to the global regulon of pathogens.

Characterization of the SPI-6 T6SS

Chapter 4 discussed the contributions of SPI-6 T6SS genes to *S. enterica* virulence and the regulatory inputs to this system which include SsrB. Previous work by other groups established that selected genes of this system contribute to pathogenesis (157, 207) and we were able to determine that both core genes, and three of four accessory islands identified by comparative genomics, are involved in systemic dissemination in mice.

It remains unclear how the T6SS contributes to systemic dissemination during murine infections. Many T6SS in proteobacteria have interbacterial competition functions (244); however, we showed that the contribution to systemic dissemination was identical when the bacteria were delivered by intraperitoneal injection thereby demonstrating that the effect on systemic dissemination was not a result of interactions with the host microbiota encountered during oral delivery. This finding suggests that the contribution to intracellular dissemination of the T6SS is one between host cells and bacterium. It is possible that the T6SS has an interbacterial function, as expression of this system was not observed under most conditions *in vitro*. Indeed, the presence of a number of paired genes suggests possible toxin/immunity pairs involved in interbacterial competition. Despite these observations, previous work had suggested that the T6SS may be involved in late stages of intracellular infection (157). This idea is supported by our observation that

expression of T6SS proteins peaked during late stages of intracellular growth in cell culture. The exact spatiotemporal role of the T6SS therefore remains to be elucidated. Generating transcriptional and translational chromosomal fusions followed by animal infections, recovery of infected cells by fluorescence activated cell sorting (FACS) and examination by microscopy will be required. This investigation should be extended to include promoters and genes of the SPI-1 and SPI-2 T3SS not only to understand the spatiotemporal role of the T6SS, but also how it relates to these other systems. Recent work has begun to show that there is T3SS regulatory cross-talk and that expression of apparatus and specificity of effectors overlap (93, 245). It seems likely therefore that the T6SS has also been integrated at a higher level of the *S. enterica* virulence program (246). While it is possible that the SPI-6 T6SS contributes to vacuolar escape or modulation of pyroptosis during late stages of host cell infection, function in the host will remain unknown in the absence of identification of a secreted effector protein.

In support for regulatory cross-talk between secretion systems in *S. enterica*, our work and that of others has identified a role for SsrB in T6SS repression (157). SsrB has no characterized repression activity and the low levels observed suggest that repression is mediated by an intermediate regulatory factor. However, we did identify an SsrB DNA binding interaction upstream of *sciS*, the downregulated SPI-6 T6SS gene. If this regulation is mediated by another factor its identity needs to be elucidated, possibly through promoter mapping of SsrB-regulated T6SS promoters as has been done for the SPI-2 T3SS (100, 105, 109).

A major limitation encountered during investigation was the low level of expression of the T6SS *in vitro*. While a phenotype could be clearly identified *in vivo*, expression *in vitro* was low under conditions tested, including low salt conditions known to activate T6SS in other bacteria (203). While this limitation has been overcome in other organisms by identification of a mutant with constitutive T6SS expression (147, 247) or knowledge of regulatory repressing proteins (219), we were unable to identify such a mutant in *S. enterica* using transposon mutagenesis. Identification of conditions *in vitro* that stimulate high levels of activation of this system or a mutant with increased T6SS activity would be useful in identifying a secreted effector protein that would likely provide insight as to SPI-6 T6SS function.

A direct question resulting from our investigation pertains to the function of the accessory island genes. Three of four islands have a role in systemic dissemination in mice however it is not clear whether these genes represent apparatus, effector, chaperone or other accessory genes. Further complicating understanding is that such T6SS accessory genes are poorly conserved and present in few genomes, most of which have T6SS, suggesting that these genes have evolved to provide very specific functions, possibly acting to specialize the T6SS for specific host targets (150). Interestingly, *S. enterica* serovar Typhimurium is primarily a pathogen of humans, yet the SPI-6 T6SS was found to have a role in a murine model of typhoid-like disease suggesting it may target processes conserved amongst mammals rather than having a role in specific host tropism like those of *Burkholderia* (229).

In summary, it appears that the *S. enterica* SPI-6 T6SS has been applied to a host-pathogen interaction rather than interbacterial competition. This system provides an additional delivery mechanism for effectors during the spatiotemporal interaction between *Salmonella* and host cells during infection. Better understanding of the function and regulation of this system might provide a target for antivirulence strategies or provide understanding of late stages of infection of infection of host cells by bacterial pathogens, particularly in the context of pyroptosis.

References

1. **Hohmann EL**. 2001. Nontyphoidal salmonellosis. Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am. **32**:263–269.
2. **Wangdi T, Winter SE, Bäumler AJ**. 2012. Typhoid fever: "you can't hit what you can't see." Gut Microbes **3**:88–92.
3. **Dougan G, Baker S**. 2014. Salmonella enterica Serovar Typhi and the Pathogenesis of Typhoid Fever. Annu. Rev. Microbiol. **68**:317–336.
4. **Monack DM**. 2012. Salmonella persistence and transmission strategies. Curr. Opin. Microbiol. **15**:100–107.
5. **Hall RM**. 2010. Salmonella genomic islands and antibiotic resistance in Salmonella enterica. Future Microbiol. **5**:1525–1538.
6. **Feasey NA, Dougan G, Kingsley RA, Heyderman RS, Gordon MA**. 2012. Invasive non-typhoidal salmonella disease: an emerging and neglected tropical disease in Africa. Lancet **379**:2489–2499.
7. **Rabsch W, Andrews HL, Kingsley RA, Prager R, Tschäpe H, Adams LG, Bäumler AJ**. 2002. Salmonella enterica serotype Typhimurium and its host-adapted variants. Infect. Immun. **70**:2249–2255.
8. **Hernández-Reyes C, Schikora A**. 2013. Salmonella, a cross-kingdom pathogen infecting humans and plants. FEMS Microbiol. Lett. **343**:1–7.
9. **Tsolis RM, Xavier MN, Santos RL, Bäumler AJ**. 2011. How to become a top model: impact of animal experimentation on human Salmonella disease research. Infect. Immun. **79**:1806–1814.
10. **Reddick LE, Alto NM**. 2014. Bacteria fighting back: how pathogens target and subvert the host innate immune system. Mol. Cell **54**:321–328.
11. **Gerlach RG, Hensel M**. 2007. Protein secretion systems and adhesins: the molecular armory of Gram-negative pathogens. Int. J. Med. Microbiol. IJMM **297**:401–415.
12. **Darmon E, Leach DRF**. 2014. Bacterial genome instability. Microbiol. Mol. Biol. Rev. MMBR **78**:1–39.
13. **Alvarez-Ordóñez A, Begley M, Prieto M, Messens W, López M, Bernardo A, Hill C**. 2011. Salmonella spp. survival strategies within the host gastrointestinal tract. Microbiol. Read. Engl. **157**:3268–3281.
14. **Bowe F, Lipps CJ, Tsolis RM, Groisman E, Heffron F, Kusters JG**. 1998. At Least Four Percent of the Salmonella typhimurium Genome Is Required for Fatal Infection of Mice. Infect. Immun. **66**:3372–3377.
15. **McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F, Hou S, Layman D, Leonard S, Nguyen C, Scott K, Holmes A, Grewal N, Mulvaney E, Ryan E, Sun H, Florea L, Miller W, Stoneking T, Nhan M, Waterston R, Wilson RK**. 2001. Complete genome

sequence of Salmonella enterica serovar Typhimurium LT2. Nature **413**:852–856.

16. **Koonin EV**. 2003. Comparative genomics, minimal gene-sets and the last universal common ancestor. Nat. Rev. Microbiol. **1**:127–136.
17. **Langille MGI, Hsiao WWL, Brinkman FSL**. 2010. Detecting genomic islands using bioinformatics approaches. Nat. Rev. Microbiol. **8**:373–382.
18. **Ho Sui SJ, Fedynak A, Hsiao WWL, Langille MGI, Brinkman FSL**. 2009. The association of virulence factors with genomic islands. PloS One **4**:e8094.
19. **Lawley TD, Chan K, Thompson LJ, Kim CC, Govoni GR, Monack DM**. 2006. Genome-wide screen for Salmonella genes required for long-term systemic infection of the mouse. PLoS Pathog. **2**:e11.
20. **Chan K, Kim CC, Falkow S**. 2005. Microarray-based detection of Salmonella enterica serovar Typhimurium transposon mutants that cannot survive in macrophages and mice. Infect. Immun. **73**:5438–5449.
21. **Klumpp J, Fuchs TM**. 2007. Identification of novel genes in genomic islands that contribute to Salmonella typhimurium replication in macrophages. Microbiol. Read. Engl. **153**:1207–1220.
22. **Hensel M**. 2004. Evolution of pathogenicity islands of Salmonella enterica. Int. J. Med. Microbiol. IJMM **294**:95–102.
23. **Porwollik S, McClelland M**. 2003. Lateral gene transfer in Salmonella. Microbes Infect. Inst. Pasteur **5**:977–989.
24. **Jacobsen A, Hendriksen RS, Aaresturp FM, Ussery DW, Friis C**. 2011. The Salmonella enterica pan-genome. Microb. Ecol. **62**:487–504.
25. **Que F, Wu S, Huang R**. 2013. Salmonella pathogenicity island 1(SPI-1) at work. Curr. Microbiol. **66**:582–587.
26. **Fookes M, Schroeder GN, Langridge GC, Blondel CJ, Mammina C, Connor TR, Seth-Smith H, Vernikos GS, Robinson KS, Sanders M, Petty NK, Kingsley RA, Bäumler AJ, Nuccio S-P, Contreras I, Santiviago CA, Maskell D, Barrow P, Humphrey T, Nastasi A, Roberts M, Frankel G, Parkhill J, Dougan G, Thomson NR**. 2011. Salmonella bongori provides insights into the evolution of the Salmonellae. PLoS Pathog. **7**:e1002191.
27. **Desai PT, Porwollik S, Long F, Cheng P, Wollam A, Bhonagiri-Palsikar V, Hallsworth-Pepin K, Clifton SW, Weinstock GM, McClelland M**. 2013. Evolutionary Genomics of Salmonella enterica Subspecies. mBio **4**.
28. **Blondel CJ, Jiménez JC, Contreras I, Santiviago CA**. 2009. Comparative genomic analysis uncovers 3 novel loci encoding type six secretion systems differentially distributed in Salmonella serotypes. BMC Genomics **10**:354.
29. **Ono S, Goldberg MD, Olsson T, Esposito D, Hinton JCD, Ladbury JE**. 2005. H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. Biochem. J. **391**:203–213.
30. **Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, Fang FC**. 2006. Selective silencing of foreign DNA with low GC content by the H-NS protein in Salmonella. Science **313**:236–238.

31. **Lucchini S, Rowley G, Goldberg MD, Hurd D, Harrison M, Hinton JCD**. 2006. H-NS mediates the silencing of laterally acquired genes in bacteria. PLoS Pathog. **2**:e81.
32. **Vernikos GS, Thomson NR, Parkhill J**. 2007. Genetic flux over time in the Salmonella lineage. Genome Biol. **8**:R100.
33. **Osborne SE, Walthers D, Tomljenovic AM, Mulder DT, Silphaduang U, Duong N, Lowden MJ, Wickham ME, Waller RF, Kenney LJ, Coombes BK**. 2009. Pathogenic adaptation of intracellular bacteria by rewiring a cis-regulatory input function. Proc. Natl. Acad. Sci. U. S. A. **106**:3982–3987.
34. **Haraga A, Ohlson MB, Miller SI**. 2008. Salmonellae interplay with host cells. Nat. Rev. Microbiol. **6**:53–66.
35. **Ouellette AJ**. 2011. Paneth cell α-defensins in enteric innate immunity. Cell. Mol. Life Sci. CMLS **68**:2215–2229.
36. **Martinez Rodriguez NR, Eloi MD, Huynh A, Dominguez T, Lam AHC, Carcamo-Molina D, Naser Z, Desharnais R, Salzman NH, Porter E**. 2012. Expansion of Paneth cell population in response to enteric Salmonella enterica serovar Typhimurium infection. Infect. Immun. **80**:266–275.
37. **Steele-Mortimer O**. 2008. The Salmonella-containing vacuole: moving with the times. Curr. Opin. Microbiol. **11**:38–45.
38. **Patel JC, Galán JE**. 2005. Manipulation of the host actin cytoskeleton by Salmonella--all in the name of entry. Curr. Opin. Microbiol. **8**:10–15.
39. **Patel S, McCormick BA**. 2014. Mucosal Inflammatory Response to Salmonella typhimurium Infection. Front. Immunol. **5**:311.
40. **Jepson MA, Clark MA**. 2001. The role of M cells in Salmonella infection. Microbes Infect. Inst. Pasteur **3**:1183–1190.
41. **Alpuche-Aranda CM, Racoosin EL, Swanson JA, Miller SI**. 1994. Salmonella stimulate macrophage macropinocytosis and persist within spacious phagosomes. J. Exp. Med. **179**:601–608.
42. **Santos RL, Bäumler AJ**. 2004. Cell tropism of Salmonella enterica. Int. J. Med. Microbiol. IJMM **294**:225–233.
43. **Monack DM, Bouley DM, Falkow S**. 2004. Salmonella typhimurium Persists within Macrophages in the Mesenteric Lymph Nodes of Chronically Infected Nramp1+/+ Mice and Can Be Reactivated by IFNγ Neutralization. J. Exp. Med. **199**:231–241.
44. **Worley MJ, Nieman GS, Geddes K, Heffron F**. 2006. Salmonella typhimurium disseminates within its host by manipulating the motility of infected cells. Proc. Natl. Acad. Sci. **103**:17915–17920.
45. **Fink SL, Cookson BT**. 2007. Pyroptosis and host cell death responses during Salmonella infection. Cell. Microbiol. **9**:2562–2570.
46. **Sarantis H, Grinstein S**. 2012. Subversion of phagocytosis for pathogen survival. Cell Host Microbe **12**:419–431.
47. **Flannagan RS, Cosío G, Grinstein S**. 2009. Antimicrobial mechanisms of phagocytes and bacterial evasion strategies. Nat. Rev. Microbiol. **7**:355–366.

48. **Huynh KK, Grinstein S**. 2007. Regulation of vacuolar pH and its modulation by some microbial species. Microbiol. Mol. Biol. Rev. MMBR **71**:452–462.
49. **Arpaia N, Godec J, Lau L, Sivick KE, McLaughlin LM, Jones MB, Dracheva T, Peterson SN, Monack DM, Barton GM**. 2011. TLR signaling is required for Salmonella typhimurium virulence. Cell **144**:675–688.
50. **Zhao B, Houry WA**. 2010. Acid stress response in enteropathogenic gammaproteobacteria: an aptitude for survival. Biochem. Cell Biol. Biochim. Biol. Cell. **88**:301–314.
51. **Foster JW**. 1999. When protons attack: microbial strategies of acid adaptation. Curr. Opin. Microbiol. **2**:170–174.
52. **Gahan CG, Hill C**. 1999. The relationship between acid stress responses and virulence in Salmonella typhimurium and Listeria monocytogenes. Int. J. Food Microbiol. **50**:93–100.
53. **Rabsch W, Voigt W, Reissbrodt R, Tsolis RM, Baumler AJ**. 1999. Salmonella typhimurium IroN and FepA Proteins Mediate Uptake of Enterobactin but Differ in Their Specificity for Other Siderophores. J. Bacteriol. **181**:3610–3612.
54. **Nairz M, Haschka D, Demetz E, Weiss G**. 2014. Iron at the interface of immunity and infection. Front. Pharmacol. **5**:152.
55. **Ganz T**. 2003. Defensins: antimicrobial peptides of innate immunity. Nat. Rev. Immunol. **3**:710–720.
56. **Peschel A**. 2002. How do bacteria resist human antimicrobial peptides? Trends Microbiol. **10**:179–186.
57. **Chen HD, Groisman EA**. 2013. The biology of the PmrA/PmrB two-component system: the major regulator of lipopolysaccharide modifications. Annu. Rev. Microbiol. **67**:83–112.
58. **Fang FC**. 2011. Antimicrobial actions of reactive oxygen species. mBio **2**.
59. **Cabiscol E, Tamarit J, Ros J**. 2000. Oxidative stress in bacteria and protein damage by reactive oxygen species. Int. Microbiol. Off. J. Span. Soc. Microbiol. **3**:3–8.
60. **Husain M, Jones-Carson J, Song M, McCollister BD, Bourret TJ, Vázquez-Torres A**. 2010. Redox sensor SsrB Cys203 enhances Salmonella fitness against nitric oxide generated in the host immune response to oral infection. Proc. Natl. Acad. Sci. U. S. A. **107**:14396–14401.
61. **Löber S, Jäckel D, Kaiser N, Hensel M**. 2006. Regulation of Salmonella pathogenicity island 2 genes by independent environmental signals. Int. J. Med. Microbiol. IJMM **296**:435–447.
62. **Ellermeier JR, Slauch JM**. 2007. Adaptation to the host environment: regulation of the SPI1 type III secretion system in Salmonella enterica serovar Typhimurium. Curr. Opin. Microbiol. **10**:24–29.
63. **Yoon H, McDermott JE, Porwollik S, McClelland M, Heffron F**. 2009. Coordinated regulation of virulence during systemic infection of Salmonella enterica serovar Typhimurium. PLoS Pathog. **5**:e1000306.

64. **Xu X, Hensel M**. 2010. Systematic analysis of the SsrAB virulon of Salmonella enterica. Infect. Immun. **78**:49–58.
65. **Dorman CJ**. 2009. Global regulators and environmental adaptation in Gram-negative pathogens. Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis. **15 Suppl 1**:47–50.
66. **Ninfa AJ, Magasanik B**. 1986. Covalent modification of the glnG product, NRI, by the glnL product, NRII, regulates the transcription of the glnALG operon in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. **83**:5909–5913.
67. **Nixon BT, Ronson CW, Ausubel FM**. 1986. Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes ntrB and ntrC. Proc. Natl. Acad. Sci. U. S. A. **83**:7850–7854.
68. **Koretke KK, Lupas AN, Warren PV, Rosenberg M, Brown JR**. 2000. Evolution of two-component signal transduction. Mol. Biol. Evol. **17**:1956–1970.
69. **Mascher T, Helmann JD, Unden G**. 2006. Stimulus perception in bacterial signal-transducing histidine kinases. Microbiol. Mol. Biol. Rev. MMBR **70**:910–938.
70. **Rowland MA, Deeds EJ**. 2014. Crosstalk and the evolution of specificity in two-component signaling. Proc. Natl. Acad. Sci. U. S. A. **111**:5550–5555.
71. **Podgornaia AI, Laub MT**. 2013. Determinants of specificity in two-component signal transduction. Curr. Opin. Microbiol. **16**:156–162.
72. **Galperin MY**. 2010. Diversity of structure and function of response regulator output domains. Curr. Opin. Microbiol. **13**:150–159.
73. **Cheung J, Hendrickson WA**. 2010. Sensor domains of two-component regulatory systems. Curr. Opin. Microbiol. **13**:116–123.
74. **Zhang Z, Hendrickson WA**. 2010. Structural characterization of the predominant family of histidine kinase sensor domains. J. Mol. Biol. **400**:335–353.
75. **Viarengo G, Sciara MI, Salazar MO, Kieffer PM, Furlán RLE, García Véscovi E**. 2013. Unsaturated long chain free fatty acids are input signals of the Salmonella enterica PhoP/PhoQ regulatory system. J. Biol. Chem. **288**:22346–22358.
76. **Prost LR, Daley ME, Le Sage V, Bader MW, Le Moual H, Klevit RE, Miller SI**. 2007. Activation of the bacterial sensor kinase PhoQ by acidic pH. Mol. Cell **26**:165–174.
77. **Cho US, Bader MW, Amaya MF, Daley ME, Klevit RE, Miller SI, Xu W**. 2006. Metal bridges between the PhoQ sensor domain and the membrane regulate transmembrane signaling. J. Mol. Biol. **356**:1193–1206.
78. **Dago AE, Schug A, Procaccini A, Hoch JA, Weigt M, Szurmant H**. 2012. Structural basis of histidine kinase autophosphorylation deduced by integrating genomics, molecular dynamics, and mutagenesis. Proc. Natl. Acad. Sci. U. S. A. **109**:E1733–1742.

79. **Stock AM, Robinson VL, Goudreau PN**. 2000. Two-component signal transduction. Annu. Rev. Biochem. **69**:183–215.
80. **Bourret RB**. 2010. Receiver domain structure and function in response regulator proteins. Curr. Opin. Microbiol. **13**:142–149.
81. **Gao R, Stock AM**. 2010. Molecular Strategies for Phosphorylation-Mediated Regulation of Response Regulator Activity. Curr. Opin. Microbiol. **13**:160–167.
82. **Volz K**. 1993. Structural conservation in the CheY superfamily. Biochemistry (Mosc.) **32**:11741–11753.
83. **Lee SY, Cho HS, Pelton JG, Yan D, Berry EA, Wemmer DE**. 2001. Crystal structure of activated CheY. Comparison with other activated receiver domains. J. Biol. Chem. **276**:16425–16431.
84. **Silversmith RE, Appleby JL, Bourret RB**. 1997. Catalytic mechanism of phosphorylation and dephosphorylation of CheY: kinetic characterization of imidazole phosphates as phosphodonors and the role of acid catalysis. Biochemistry (Mosc.) **36**:14965–14974.
85. **Wolfe AJ**. 2010. Physiologically relevant small phosphodonors link metabolism to signal transduction. Curr. Opin. Microbiol. **13**:204–209.
86. **Lukat GS, McCleary WR, Stock AM, Stock JB**. 1992. Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors. Proc. Natl. Acad. Sci. U. S. A. **89**:718–722.
87. **Zhao R, Collins EJ, Bourret RB, Silversmith RE**. 2002. Structure and catalytic mechanism of the E. coli chemotaxis phosphatase CheZ. Nat. Struct. Biol. **9**:570–575.
88. **Pioszak AA, Ninfa AJ**. 2004. Mutations altering the N-terminal receiver domain of NRI (NtrC) That prevent dephosphorylation by the NRII-PII complex in Escherichia coli. J. Bacteriol. **186**:5730–5740.
89. **Pazy Y, Motaleb MA, Guarnieri MT, Charon NW, Zhao R, Silversmith RE**. 2010. Identical phosphatase mechanisms achieved through distinct modes of binding phosphoprotein substrate. Proc. Natl. Acad. Sci. U. S. A. **107**:1924–1929.
90. **Aravind L, Anantharaman V, Balaji S, Babu MM, Iyer LM**. 2005. The many faces of the helix-turn-helix domain: transcription regulation and beyond. FEMS Microbiol. Rev. **29**:231–262.
91. **Browning DF, Busby SJ**. 2004. The regulation of bacterial transcription initiation. Nat. Rev. Microbiol. **2**:57–65.
92. **D'haeseleer P**. 2006. How does DNA sequence motif discovery work? Nat. Biotechnol. **24**:959–961.
93. **Fass E, Groisman EA**. 2009. Control of Salmonella pathogenicity island-2 gene expression. Curr. Opin. Microbiol. **12**:199–204.
94. **Beier D, Gross R**. 2006. Regulation of bacterial virulence by two-component systems. Curr. Opin. Microbiol. **9**:143–152.

95. **Brown NF, Vallance BA, Coombes BK, Valdez Y, Coburn BA, Finlay BB**. 2005. Salmonella pathogenicity island 2 is expressed prior to penetrating the intestine. PLoS Pathog. **1**:e32.
96. **Osborne SE, Coombes BK**. 2011. Transcriptional priming of Salmonella Pathogenicity Island-2 precedes cellular invasion. PloS One **6**:e21648.
97. **Silphaduang U, Mascarenhas M, Karmali M, Coombes BK**. 2007. Repression of intracellular virulence factors in Salmonella by the Hha and YdgT nucleoid-associated proteins. J. Bacteriol. **189**:3669–3673.
98. **Vivero A, Baños RC, Mariscotti JF, Oliveros JC, García-del Portillo F, Juárez A, Madrid C**. 2008. Modulation of horizontally acquired genes by the Hha-YdgT proteins in Salmonella enterica serovar Typhimurium. J. Bacteriol. **190**:1152–1156.
99. **Walthers D, Li Y, Liu Y, Anand G, Yan J, Kenney LJ**. 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.
100. **Walthers D, Carroll RK, Navarre WW, Libby SJ, Fang FC, Kenney LJ**. 2007. The response regulator SsrB activates expression of diverse Salmonella pathogenicity island 2 promoters and counters silencing by the nucleoid-associated protein H-NS. Mol. Microbiol. **65**:477–493.
101. **Cameron ADS, Dorman CJ**. 2012. A Fundamental Regulatory Mechanism Operating through OmpR and DNA Topology Controls Expression of Salmonella Pathogenicity Islands SPI-1 and SPI-2. PLoS Genet **8**:e1002615.
102. **Duong N, Osborne S, Bustamante VH, Tomljenovic AM, Puente JL, Coombes BK**. 2007. Thermosensing coordinates a cis-regulatory module for transcriptional activation of the intracellular virulence system in Salmonella enterica serovar Typhimurium. J. Biol. Chem. **282**:34077–34084.
103. **Okada N, Oi Y, Takeda-Shitaka M, Kanou K, Umeyama H, Haneda T, Miki T, Hosoya S, Danbara H**. 2007. Identification of amino acid residues of Salmonella SlyA that are critical for transcriptional regulation. Microbiol. Read. Engl. **153**:548–560.
104. **Lee AK, Detweiler CS, Falkow S**. 2000. OmpR regulates the two-component system SsrA-ssrB in Salmonella pathogenicity island 2. J. Bacteriol. **182**:771–781.
105. **Feng X, Oropeza R, Kenney LJ**. 2003. Dual regulation by phospho-OmpR of ssrA/B gene expression in Salmonella pathogenicity island 2. Mol. Microbiol. **48**:1131–1143.
106. **Bijlsma JJE, Groisman EA**. 2005. The PhoP/PhoQ system controls the intramacrophage type three secretion system of Salmonella enterica. Mol. Microbiol. **57**:85–96.
107. **Choi E, Kim H, Lee H, Nam D, Choi J, Shin D**. 2014. The Iron-Sensing Fur Regulator Controls Expression Timing and Levels of Salmonella Pathogenicity

Island 2 Genes in the Course of Environmental Acidification. Infect. Immun. **82**:2203–2210.

108. **Choi J, Groisman EA**. 2013. The lipopolysaccharide modification regulator PmrA limits Salmonella virulence by repressing the type three-secretion system Spi/Ssa. Proc. Natl. Acad. Sci. **110**:9499–9504.
109. **Feng X, Walthers D, Oropeza R, Kenney LJ**. 2004. The response regulator SsrB activates transcription and binds to a region overlapping OmpR binding sites at Salmonella pathogenicity island 2. Mol. Microbiol. **54**:823–835.
110. **Ochman H, Soncini FC, Solomon F, Groisman EA**. 1996. Identification of a pathogenicity island required for Salmonella survival in host cells. Proc. Natl. Acad. Sci. U. S. A. **93**:7800–7804.
111. **Cirillo DM, Valdivia RH, Monack DM, Falkow S**. 1998. Macrophage-dependent induction of the Salmonella pathogenicity island 2 type III secretion system and its role in intracellular survival. Mol. Microbiol. **30**:175–188.
112. **Hensel M, Shea JE, Waterman SR, Mundy R, Nikolaus T, Banks G, Vazquez-Torres A, Gleeson C, Fang FC, Holden DW**. 1998. Genes encoding putative effector proteins of the type III secretion system of Salmonella pathogenicity island 2 are required for bacterial virulence and proliferation in macrophages. Mol. Microbiol. **30**:163–174.
113. **Rytkönen A, Poh J, Garmendia J, Boyle C, Thompson A, Liu M, Freemont P, Hinton JCD, Holden DW**. 2007. SseL, a Salmonella deubiquitinase required for macrophage killing and virulence. Proc. Natl. Acad. Sci. U. S. A. **104**:3502–3507.
114. **Deiwick J, Nikolaus T, Erdogan S, Hensel M**. 1999. Environmental regulation of Salmonella pathogenicity island 2 gene expression. Mol. Microbiol. **31**:1759–1773.
115. **Miao EA, Freeman JA, Miller SI**. 2002. Transcription of the SsrAB regulon is repressed by alkaline pH and is independent of PhoPQ and magnesium concentration. J. Bacteriol. **184**:1493–1497.
116. **Carroll RK, Liao X, Morgan LK, Cicirelli EM, Li Y, Sheng W, Feng X, Kenney LJ**. 2009. Structural and functional analysis of the C-terminal DNA binding domain of the Salmonella typhimurium SPI-2 response regulator SsrB. J. Biol. Chem. **284**:12008–12019.
117. **Miao EA, Scherer CA, Tsolis RM, Kingsley RA, Adams LG, Bäumler AJ, Miller SI**. 1999. Salmonella typhimurium leucine-rich repeat proteins are targeted to the SPI1 and SPI2 type III secretion systems. Mol. Microbiol. **34**:850–864.
118. **Worley MJ, Ching KH, Heffron F**. 2000. Salmonella SsrB activates a global regulon of horizontally acquired genes. Mol. Microbiol. **36**:749–761.
119. **Bhavsar AP, Guttman JA, Finlay BB**. 2007. Manipulation of host-cell pathways by bacterial pathogens. Nature **449**:827–834.

120. **Abby SS, Rocha EPC**. 2012. The Non-Flagellar Type III Secretion System Evolved from the Bacterial Flagellum and Diversified into Host-Cell Adapted Systems. PLoS Genet **8**:e1002983.
121. **Erhardt M, Namba K, Hughes KT**. 2010. Bacterial Nanomachines: The Flagellum and Type III Injectisome. Cold Spring Harb. Perspect. Biol. **2**:a000299.
122. **Chatterjee S, Chaudhury S, McShan AC, Kaur K, De Guzman RN**. 2013. Structure and Biophysics of Type III Secretion in Bacteria. Biochemistry (Mosc.) **52**:2508–2517.
123. **Osborne SE, Coombes BK**. 2011. Expression and secretion hierarchy in the nonflagellar type III secretion system. Future Microbiol. **6**:193–202.
124. **Yu X-J, McGourty K, Liu M, Unsworth KE, Holden DW**. 2010. pH sensing by intracellular Salmonella induces effector translocation. Science **328**:1040–1043.
125. **Pizarro-Cerdá J, Cossart P**. 2006. Bacterial adhesion and entry into host cells. Cell **124**:715–727.
126. **Bruno VM, Hannemann S, Lara-Tejero M, Flavell RA, Kleinstein SH, Galán JE**. 2009. Salmonella Typhimurium type III secretion effectors stimulate innate immune responses in cultured epithelial cells. PLoS Pathog. **5**:e1000538.
127. **Galyov EE, Wood MW, Rosqvist R, Mullan PB, Watson PR, Hedges S, Wallis TS**. 1997. A secreted effector protein of Salmonella dublin is translocated into eukaryotic cells and mediates inflammation and fluid secretion in infected ileal mucosa. Mol. Microbiol. **25**:903–912.
128. **Lee CA, Silva M, Siber AM, Kelly AJ, Galyov E, McCormick BA**. 2000. A secreted Salmonella protein induces a proinflammatory response in epithelial cells, which promotes neutrophil migration. Proc. Natl. Acad. Sci. U. S. A. **97**:12283–12288.
129. **Wu H, Jones RM, Neish AS**. 2012. The Salmonella effector AvrA mediates bacterial intracellular survival during infection in vivo. Cell. Microbiol. **14**:28–39.
130. **Hensel M, Shea JE, Gleeson C, Jones MD, Dalton E, Holden DW**. 1995. Simultaneous identification of bacterial virulence genes by negative selection. Science **269**:400–403.
131. **Figueira R, Holden DW**. 2012. Functions of the Salmonella pathogenicity island 2 (SPI-2) type III secretion system effectors. Microbiol. Read. Engl. **158**:1147–1161.
132. **Dumont A, Boucrot E, Drevensek S, Daire V, Gorvel J-P, Poüs C, Holden DW, Méresse S**. 2010. SKIP, the host target of the Salmonella virulence factor SifA, promotes kinesin-1-dependent vacuolar membrane exchanges. Traffic Cph. Den. **11**:899–911.
133. **Beuzón CR, Méresse S, Unsworth KE, Ruíz-Albert J, Garvis S, Waterman SR, Ryder TA, Boucrot E, Holden DW**. 2000. Salmonella maintains the integrity of its intracellular vacuole through the action of SifA. EMBO J. **19**:3235–3249.

134. **Beuzón CR, Salcedo SP, Holden DW**. 2002. Growth and killing of a Salmonella enterica serovar Typhimurium sifA mutant strain in the cytosol of different host cell lines. Microbiol. Read. Engl. **148**:2705–2715.
135. **Garcia-del Portillo F, Zwick MB, Leung KY, Finlay BB**. 1993. Salmonella induces the formation of filamentous structures containing lysosomal membrane glycoproteins in epithelial cells. Proc. Natl. Acad. Sci. U. S. A. **90**:10544–10548.
136. **Stein MA, Leung KY, Zwick M, Garcia-del Portillo F, Finlay BB**. 1996. Identification of a Salmonella virulence gene required for formation of filamentous structures containing lysosomal membrane glycoproteins within epithelial cells. Mol. Microbiol. **20**:151–164.
137. **Deiwick J, Salcedo SP, Boucrot E, Gilliland SM, Henry T, Petermann N, Waterman SR, Gorvel J-P, Holden DW, Méresse S**. 2006. The translocated Salmonella effector proteins SseF and SseG interact and are required to establish an intracellular replication niche. Infect. Immun. **74**:6965–6972.
138. **Kuhle V, Hensel M**. 2002. SseF and SseG are translocated effectors of the type III secretion system of Salmonella pathogenicity island 2 that modulate aggregation of endosomal compartments. Cell. Microbiol. **4**:813–824.
139. **Haraga A, Miller SI**. 2003. A Salmonella enterica serovar typhimurium translocated leucine-rich repeat effector protein inhibits NF-kappa B-dependent gene expression. Infect. Immun. **71**:4052–4058.
140. **Mazurkiewicz P, Thomas J, Thompson JA, Liu M, Arbibe L, Sansonetti P, Holden DW**. 2008. SpvC is a Salmonella effector with phosphothreonine lyase activity on host mitogen-activated protein kinases. Mol. Microbiol. **67**:1371–1383.
141. **Haneda T, Ishii Y, Shimizu H, Ohshima K, Iida N, Danbara H, Okada N**. 2012. Salmonella type III effector SpvC, a phosphothreonine lyase, contributes to reduction in inflammatory response during intestinal phase of infection. Cell. Microbiol. **14**:485–499.
142. **Méresse S, Unsworth KE, Habermann A, Griffiths G, Fang F, Martínez-Lorenzo MJ, Waterman SR, Gorvel JP, Holden DW**. 2001. Remodelling of the actin cytoskeleton is essential for replication of intravacuolar Salmonella. Cell. Microbiol. **3**:567–577.
143. **Poh J, Odendall C, Spanos A, Boyle C, Liu M, Freemont P, Holden DW**. 2008. SteC is a Salmonella kinase required for SPI-2-dependent F-actin remodelling. Cell. Microbiol. **10**:20–30.
144. **Miao EA, Brittnacher M, Haraga A, Jeng RL, Welch MD, Miller SI**. 2003. Salmonella effectors translocated across the vacuolar membrane interact with the actin cytoskeleton. Mol. Microbiol. **48**:401–415.
145. **Bladergroen MR, Badelt K, Spaink HP**. 2003. Infection-blocking genes of a symbiotic Rhizobium leguminosarum strain that are involved in temperature-dependent protein secretion. Mol. Plant-Microbe Interact. MPMI **16**:53–64.

146. **Rao PSS, Yamada Y, Tan YP, Leung KY**. 2004. Use of proteomics to identify novel virulence determinants that are required for Edwardsiella tarda pathogenesis. Mol. Microbiol. **53**:573–586.
147. **Pukatzki S, Ma AT, Sturtevant D, Krastins B, Sarracino D, Nelson WC, Heidelberg JF, Mekalanos JJ**. 2006. Identification of a conserved bacterial protein secretion system in Vibrio cholerae using the Dictyostelium host model system. Proc. Natl. Acad. Sci. U. S. A. **103**:1528–1533.
148. **Shrivastava S, Mande SS**. 2008. Identification and functional characterization of gene components of Type VI Secretion system in bacterial genomes. PloS One **3**:e2955.
149. **Boyer F, Fichant G, Berthod J, Vandenbrouck Y, Attree I**. 2009. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? BMC Genomics **10**:104.
150. **Bingle LE, Bailey CM, Pallen MJ**. 2008. Type VI secretion: a beginner's guide. Curr. Opin. Microbiol. **11**:3–8.
151. **Pukatzki S, Ma AT, Revel AT, Sturtevant D, Mekalanos JJ**. 2007. Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. Proc. Natl. Acad. Sci. U. S. A. **104**:15508–15513.
152. **Hood RD, Singh P, Hsu F, Güvener T, Carl MA, Trinidad RRS, Silverman JM, Ohlson BB, Hicks KG, Plemel RL, Li M, Schwarz S, Wang WY, Merz AJ, Goodlett DR, Mougous JD**. 2010. A type VI secretion system of Pseudomonas aeruginosa targets a toxin to bacteria. Cell Host Microbe **7**:25–37.
153. **Leiman PG, Basler M, Ramagopal UA, Bonanno JB, Sauder JM, Pukatzki S, Burley SK, Almo SC, Mekalanos JJ**. 2009. Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. Proc. Natl. Acad. Sci. U. S. A. **106**:4154–4159.
154. **Pell LG, Kanelis V, Donaldson LW, Howell PL, Davidson AR**. 2009. The phage lambda major tail protein structure reveals a common evolution for long-tailed phages and the type VI bacterial secretion system. Proc. Natl. Acad. Sci. U. S. A. **106**:4160–4165.
155. **Pukatzki S, McAuley SB, Miyata ST**. 2009. The type VI secretion system: translocation of effectors and effector-domains. Curr. Opin. Microbiol. **12**:11–17.
156. **Folkesson A, Löfdahl S, Normark S**. 2002. The Salmonella enterica subspecies I specific centisome 7 genomic island encodes novel protein families present in bacteria living in close contact with eukaryotic cells. Res. Microbiol. **153**:537–545.
157. **Parsons DA, Heffron F**. 2005. sciS, an icmF homolog in Salmonella enterica serovar Typhimurium, limits intracellular replication and decreases virulence. Infect. Immun. **73**:4338–4345.
158. **Mougous JD, Cuff ME, Raunser S, Shen A, Zhou M, Gifford CA, Goodman AL, Joachimiak G, Ordoñez CL, Lory S, Walz T, Joachimiak A, Mekalanos JJ**.

2006. A virulence locus of Pseudomonas aeruginosa encodes a protein secretion apparatus. Science **312**:1526–1530.
159. **Steele-Mortimer O**. 2008. The Salmonella-containing Vacuole – Moving with the Times. Curr. Opin. Microbiol. **11**:38–45.
160. **Lee E-J, Groisman EA**. 2012. Control of a Salmonella virulence locus by an ATP-sensing leader messenger RNA. Nature **486**:271–275.
161. **Bang IS, Audia JP, Park YK, Foster JW**. 2002. Autoinduction of the ompR response regulator by acid shock and control of the Salmonella enterica acid tolerance response. Mol. Microbiol. **44**:1235–1250.
162. **Bearson BL, Wilson L, Foster JW**. 1998. A low pH-inducible, PhoPQ-dependent acid tolerance response protects Salmonella typhimurium against inorganic acid stress. J. Bacteriol. **180**:2409–2417.
163. **Riesenberg-Wilmes MR, Bearson B, Foster JW, Curtis R 3rd**. 1996. Role of the acid tolerance response in virulence of Salmonella typhimurium. Infect. Immun. **64**:1085–1092.
164. **Bustamante VH, Martínez LC, Santana FJ, Knodler LA, Steele-Mortimer O, Puente JL**. 2008. HilD-mediated transcriptional cross-talk between SPI-1 and SPI-2. Proc. Natl. Acad. Sci. U. S. A. **105**:14591–14596.
165. **Ochman H, Lawrence JG, Groisman EA**. 2000. Lateral gene transfer and the nature of bacterial innovation. Nature **405**:299–304.
166. **Prud'homme B, Gompel N, Carroll SB**. 2007. Emerging principles of regulatory evolution. Proc. Natl. Acad. Sci. U. S. A. **104 Suppl 1**:8605–8612.
167. **Stone JR, Wray GA**. 2001. Rapid evolution of cis-regulatory sequences via local point mutations. Mol. Biol. Evol. **18**:1764–1770.
168. **Isalan M, Lemerle C, Michalodimitrakis K, Horn C, Beltrao P, Raineri E, Garriga-Canut M, Serrano L**. 2008. Evolvability and hierarchy in rewired bacterial gene networks. Nature **452**:840–845.
169. **Mayo AE, Setty Y, Shavit S, Zaslaver A, Alon U**. 2006. Plasticity of the cis-regulatory input function of a gene. PLoS Biol. **4**:e45.
170. **Perez JC, Groisman EA**. 2009. Evolution of transcriptional regulatory circuits in bacteria. Cell **138**:233–244.
171. **Shea JE, Hensel M, Gleeson C, Holden DW**. 1996. Identification of a virulence locus encoding a second type III secretion system in Salmonella typhimurium. Proc. Natl. Acad. Sci. U. S. A. **93**:2593–2597.
172. **Cornelis GR**. 2006. The type III secretion injectisome. Nat. Rev. Microbiol. **4**:811–825.
173. **Galán JE, Wolf-Watz H**. 2006. Protein delivery into eukaryotic cells by type III secretion machines. Nature **444**:567–573.
174. **Coombes BK, Wickham ME, Lowden MJ, Brown NF, Finlay BB**. 2005. Negative regulation of Salmonella pathogenicity island 2 is required for contextual control of virulence during typhoid. Proc. Natl. Acad. Sci. U. S. A. **102**:17460–17465.

175. **Yoon H, McDermott JE, Porwollik S, McClelland M, Heffron F**. 2009. Coordinated regulation of virulence during systemic infection of Salmonella enterica serovar Typhimurium. PLoS Pathog. **5**:e1000306.
176. **Ochman H, Groisman EA**. 1996. Distribution of pathogenicity islands in Salmonella spp. Infect. Immun. **64**:5410–5412.
177. **Coombes BK, Brown NF, Valdez Y, Brumell JH, Finlay BB**. 2004. Expression and secretion of Salmonella pathogenicity island-2 virulence genes in response to acidification exhibit differential requirements of a functional type III secretion apparatus and SsaL. J. Biol. Chem. **279**:49804–49815.
178. **Hokamp K, Roche FM, Acab M, Rousseau M-E, Kuo B, Goode D, Aeschliman D, Bryan J, Babiuk LA, Hancock REW, Brinkman FSL**. 2004. ArrayPipe: a flexible processing pipeline for microarray data. Nucleic Acids Res. **32**:W457–459.
179. **Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND, Koonin EV**. 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. **29**:22–28.
180. **Dobrindt U, Hochhut B, Hentschel U, Hacker J**. 2004. Genomic islands in pathogenic and environmental microorganisms. Nat. Rev. Microbiol. **2**:414–424.
181. **Meng X, Brodsky MH, Wolfe SA**. 2005. A bacterial one-hybrid system for determining the DNA-binding specificity of transcription factors. Nat. Biotechnol. **23**:988–994.
182. **Zwir I, Shin D, Kato A, Nishino K, Latifi T, Solomon F, Hare JM, Huang H, Groisman EA**. 2005. Dissecting the PhoP regulatory network of Escherichia coli and Salmonella enterica. Proc. Natl. Acad. Sci. U. S. A. **102**:2862–2867.
183. **Bailey TL, Williams N, Misleh C, Li WW**. 2006. MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Res. **34**:W369–373.
184. **Liu XS, Brutlag DL, Liu JS**. 2002. An algorithm for finding protein-DNA binding sites with applications to chromatin-immunoprecipitation microarray experiments. Nat. Biotechnol. **20**:835–839.
185. **Toh H, Weiss BL, Perkin SAH, Yamashita A, Oshima K, Hattori M, Aksoy S**. 2006. Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of Sodalis glossinidius in the tsetse host. Genome Res. **16**:149–156.
186. **Thijs G, Lescot M, Marchal K, Rombauts S, De Moor B, Rouzé P, Moreau Y**. 2001. A higher-order background model improves the detection of promoter regulatory elements by Gibbs sampling. Bioinforma. Oxf. Engl. **17**:1113–1122.
187. **Thijs G, Marchal K, Lescot M, Rombauts S, De Moor B, Rouzé P, Moreau Y**. 2002. A Gibbs sampling method to detect overrepresented motifs in the upstream regions of coexpressed genes. J. Comput. Biol. J. Comput. Mol. Cell Biol. **9**:447–464.

188. **Coombes BK**. 2009. Type III secretion systems in symbiotic adaptation of pathogenic and non-pathogenic bacteria. Trends Microbiol. **17**:89–94.
189. **Dale C, Plague GR, Wang B, Ochman H, Moran NA**. 2002. Type III secretion systems and the evolution of mutualistic endosymbiosis. Proc. Natl. Acad. Sci. U. S. A. **99**:12397–12402.
190. **Cho B-K, Knight EM, Barrett CL, Palsson BØ**. 2008. Genome-wide analysis of Fis binding in Escherichia coli indicates a causative role for A-/AT-tracts. Genome Res. **18**:900–910.
191. **Shimada T, Ishihama A, Busby SJW, Grainger DC**. 2008. The Escherichia coli RutR transcription factor binds at targets within genes as well as intergenic regions. Nucleic Acids Res. **36**:3950–3955.
192. **MacArthur S, Brookfield JFY**. 2004. Expected rates and modes of evolution of enhancer sequences. Mol. Biol. Evol. **21**:1064–1073.
193. **Salcedo SP, Holden DW**. 2003. SseG, a virulence protein that targets Salmonella to the Golgi network. EMBO J. **22**:5003–5014.
194. **Shin D, Lee E-J, Huang H, Groisman EA**. 2006. A positive feedback loop promotes transcription surge that jump-starts Salmonella virulence circuit. Science **314**:1607–1609.
195. **Langille MGI, Brinkman FSL**. 2009. IslandViewer: an integrated interface for computational identification and visualization of genomic islands. Bioinforma. Oxf. Engl. **25**:664–665.
196. **Crooks GE, Hon G, Chandonia J-M, Brenner SE**. 2004. WebLogo: a sequence logo generator. Genome Res. **14**:1188–1190.
197. **Uzzau S, Figueroa-Bossi N, Rubino S, Bossi L**. 2001. Epitope tagging of chromosomal genes in Salmonella. Proc. Natl. Acad. Sci. U. S. A. **98**:15264–15269.
198. **Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA**. 2009. Circos: an information aesthetic for comparative genomics. Genome Res. **19**:1639–1645.
199. **Cascales E, Christie PJ**. 2003. The versatile bacterial type IV secretion systems. Nat. Rev. Microbiol. **1**:137–149.
200. **Kuhle V, Hensel M**. 2004. Cellular microbiology of intracellular Salmonella enterica: functions of the type III secretion system encoded by Salmonella pathogenicity island 2. Cell. Mol. Life Sci. CMLS **61**:2812–2826.
201. **Schwarz S, Hood RD, Mougous JD**. 2010. What is type VI secretion doing in all those bugs? Trends Microbiol. **18**:531–537.
202. **Basler M, Pilhofer M, Henderson GP, Jensen GJ, Mekalanos JJ**. 2012. Type VI secretion requires a dynamic contractile phage tail-like structure. Nature **483**:182–186.
203. **Russell AB, Hood RD, Bui NK, LeRoux M, Vollmer W, Mougous JD**. 2011. Type VI secretion delivers bacteriolytic effectors to target cells. Nature **475**:343–347.

204. **Sabbagh SC, Forest CG, Lepage C, Leclerc J-M, Daigle F**. 2010. So similar, yet so different: uncovering distinctive features in the genomes of Salmonella enterica serovars Typhimurium and Typhi. FEMS Microbiol. Lett. **305**:1–13.
205. **Folkesson A, Advani A, Sukupolvi S, Pfeifer JD, Normark S, Löfdahl S**. 1999. Multiple insertions of fimbrial operons correlate with the evolution of Salmonella serovars responsible for human disease. Mol. Microbiol. **33**:612–622.
206. **Lambert MA, Smith SGJ**. 2008. The PagN protein of Salmonella enterica serovar Typhimurium is an adhesin and invasin. BMC Microbiol. **8**:142.
207. **Haneda T, Ishii Y, Danbara H, Okada N**. 2009. Genome-wide identification of novel genomic islands that contribute to Salmonella virulence in mouse systemic infection. FEMS Microbiol. Lett. **297**:241–249.
208. **Chan K, Kim CC, Falkow S**. 2005. Microarray-based detection of Salmonella enterica serovar Typhimurium transposon mutants that cannot survive in macrophages and mice. Infect. Immun. **73**:5438–5449.
209. **Klumpp J, Fuchs TM**. 2007. Identification of novel genes in genomic islands that contribute to Salmonella typhimurium replication in macrophages. Microbiol. Read. Engl. **153**:1207–1220.
210. **Lawley TD, Chan K, Thompson LJ, Kim CC, Govoni GR, Monack DM**. 2006. Genome-wide screen for Salmonella genes required for long-term systemic infection of the mouse. PLoS Pathog. **2**:e11.
211. **Hautefort I, Thompson A, Eriksson-Ygberg S, Parker ML, Lucchini S, Danino V, Bongaerts RJM, Ahmad N, Rhen M, Hinton JCD**. 2008. During infection of epithelial cells Salmonella enterica serovar Typhimurium undergoes a time-dependent transcriptional adaptation that results in simultaneous expression of three type 3 secretion systems. Cell. Microbiol. **10**:958–984.
212. **Eriksson S, Lucchini S, Thompson A, Rhen M, Hinton JCD**. 2003. Unravelling the biology of macrophage infection by gene expression profiling of intracellular Salmonella enterica. Mol. Microbiol. **47**:103–118.
213. **Ma AT, McAuley S, Pukatzki S, Mekalanos JJ**. 2009. Translocation of a Vibrio cholerae type VI secretion effector requires bacterial endocytosis by host cells. Cell Host Microbe **5**:234–243.
214. **Bernard CS, Brunet YR, Gueguen E, Cascales E**. 2010. Nooks and crannies in type VI secretion regulation. J. Bacteriol. **192**:3850–3860.
215. **Filloux A, Hachani A, Bleves S**. 2008. The bacterial type VI secretion machine: yet another player for protein transport across membranes. Microbiol. Read. Engl. **154**:1570–1583.
216. **Kitaoka M, Miyata ST, Brooks TM, Unterweger D, Pukatzki S**. 2011. VasH is a transcriptional regulator of the type VI secretion system functional in endemic and pandemic Vibrio cholerae. J. Bacteriol. **193**:6471–6482.

217. **Mougous JD, Gifford CA, Ramsdell TL, Mekalanos JJ**. 2007. Threonine phosphorylation post-translationally regulates protein secretion in Pseudomonas aeruginosa. Nat. Cell Biol. **9**:797–803.
218. **Silverman JM, Austin LS, Hsu F, Hicks KG, Hood RD, Mougous JD**. 2011. Separate inputs modulate phosphorylation-dependent and -independent type VI secretion activation. Mol. Microbiol. **82**:1277–1290.
219. **Records AR, Gross DC**. 2010. Sensor kinases RetS and LadS regulate Pseudomonas syringae type VI secretion and virulence factors. J. Bacteriol. **192**:3584–3596.
220. **Wang M, Luo Z, Du H, Xu S, Ni B, Zhang H, Sheng X, Xu H, Huang X**. 2011. Molecular Characterization of a Functional Type VI Secretion System in Salmonella enterica serovar Typhi. Curr. Microbiol.
221. **Datsenko KA, Wanner BL**. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. U. S. A. **97**:6640–6645.
222. **Lane MC, Alteri CJ, Smith SN, Mobley HLT**. 2007. Expression of flagella is coincident with uropathogenic Escherichia coli ascension to the upper urinary tract. Proc. Natl. Acad. Sci. U. S. A. **104**:16669–16674.
223. **Wang RF, Kushner SR**. 1991. Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in Escherichia coli. Gene **100**:195–199.
224. **Shalom G, Shaw JG, Thomas MS**. 2007. In vivo expression technology identifies a type VI secretion system locus in Burkholderia pseudomallei that is induced upon invasion of macrophages. Microbiol. Read. Engl. **153**:2689–2699.
225. **Bonemann G, Pietrosiuk A, Diemand A, Zentgraf H, Mogk A**. 2009. Remodelling of VipA/VipB tubules by ClpV-mediated threading is crucial for type VI protein secretion. EMBO J **28**:315–325.
226. **Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH**. 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. **39**:D225–229.
227. **Marchal K, De Keersmaecker S, Monsieurs P, van Boxel N, Lemmens K, Thijs G, Vanderleyden J, De Moor B**. 2004. In silico identification and experimental validation of PmrAB targets in Salmonella typhimurium by regulatory motif detection. Genome Biol. **5**:R9.
228. **Libby SJ, Brehm MA, Greiner DL, Shultz LD, McClelland M, Smith KD, Cookson BT, Karlinsey JE, Kinkel TL, Porwollik S, Canals R, Cummings LA, Fang FC**. 2010. Humanized nonobese diabetic-scid IL2rgammanull mice are

susceptible to lethal Salmonella Typhi infection. Proc. Natl. Acad. Sci. U. S. A. **107**:15589–15594.

229. **Schwarz S, West TE, Boyer F, Chiang W-C, Carl MA, Hood RD, Rohmer L, Tolker-Nielsen T, Skerrett SJ, Mougous JD**. 2010. Burkholderia type VI secretion systems have distinct roles in eukaryotic and bacterial cell interactions. PLoS Pathog. **6**.
230. **Wang YD, Zhao S, Hill CW**. 1998. Rhs elements comprise three subfamilies which diverged prior to acquisition by Escherichia coli. J. Bacteriol. **180**:4102–4110.
231. **Poole SJ, Diner EJ, Aoki SK, Braaten BA, t' Kint de Roodenbeke C, Low DA, Hayes CS**. 2011. Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (CDI) and rearrangement hotspot (Rhs) systems. PLoS Genet. **7**:e1002217.
232. **MacIntyre DL, Miyata ST, Kitaoka M, Pukatzki S**. 2010. The Vibrio cholerae type VI secretion system displays antimicrobial properties. Proc. Natl. Acad. Sci. U. S. A. **107**:19520–19524.
233. **Tomljenovic-Berube AM, Mulder DT, Whiteside MD, Brinkman FSL, Coombes BK**. 2010. Identification of the regulatory logic controlling Salmonella pathoadaptation by the SsrA-SsrB two-component system. PLoS Genet. **6**:e1000875.
234. **Perez JC, Groisman EA**. 2007. Acid pH activation of the PmrA/PmrB two-component regulatory system of Salmonella enterica. Mol. Microbiol. **63**:283–293.
235. **Perier A, Chassaing A, Raffestin S, Pichard S, Masella M, Ménez A, Forge V, Chenal A, Gillet D**. 2007. Concerted protonation of key histidines triggers membrane interaction of the diphtheria toxin T domain. J. Biol. Chem. **282**:24239–24245.
236. **Müller S, Götz M, Beier D**. 2009. Histidine Residue 94 Is Involved in pH Sensing by Histidine Kinase ArsS of Helicobacter pylori. PLoS ONE **4**.
237. **Prost LR, Miller SI**. 2008. The Salmonellae PhoQ sensor: mechanisms of detection of phagosome signals. Cell. Microbiol. **10**:576–582.
238. **Scheu PD, Liao Y-F, Bauer J, Kneuper H, Basché T, Unden G, Erker W**. 2010. Oligomeric sensor kinase DcuS in the membrane of Escherichia coli and in proteoliposomes: chemical cross-linking and FRET spectroscopy. J. Bacteriol. **192**:3474–3483.
239. **Kenney LJ**. 2010. How important is the phosphatase activity of sensor kinases? Curr. Opin. Microbiol. **13**:168–176.
240. **Winter SE, Thiennimitr P, Winter MG, Butler BP, Huseby DL, Crawford RW, Russell JM, Bevins CL, Adams LG, Tsolis RM, Roth JR, Bäumler AJ**. 2010. Gut inflammation provides a respiratory electron acceptor for Salmonella. Nature **467**:426–429.
241. **Osborne SE, Tuinema BR, Mok MCY, Lau PS, Bui NK, Tomljenovic-Berube AM, Vollmer W, Zhang K, Junop M, Coombes BK**. 2012. Characterization of

DalS, an ATP-binding cassette transporter for D-alanine, and its role in pathogenesis in Salmonella enterica. J. Biol. Chem. **287**:15242–15250.

242. **Gao R, Tao Y, Stock AM**. 2008. System-level mapping of Escherichia coli response regulator dimerization with FRET hybrids. Mol. Microbiol. **69**:1358–1372.
243. **Dale C, Jones T, Pontes M**. 2005. Degenerative evolution and functional diversification of type-III secretion systems in the insect endosymbiont Sodalis glossinidius. Mol. Biol. Evol. **22**:758–766.
244. **Kapitein N, Mogk A**. 2013. Deadly syringes: type VI secretion system activities in pathogenicity and interbacterial competition. Curr. Opin. Microbiol. **16**:52–58.
245. **Golubeva YA, Sadik AY, Ellermeier JR, Slauch JM**. 2012. Integrating global regulatory input into the Salmonella pathogenicity island 1 type III secretion system. Genetics **190**:79–90.
246. **Leung KY, Siame BA, Snowball H, Mok Y-K**. 2011. Type VI secretion regulation: crosstalk and intracellular communication. Curr. Opin. Microbiol. **14**:9–15.
247. **Unterweger D, Kitaoka M, Miyata ST, Bachmann V, Brooks TM, Moloney J, Sosa O, Silva D, Duran-Gonzalez J, Provenzano D, Pukatzki S**. 2012. Constitutive type VI secretion system expression gives Vibrio cholerae intra- and interspecific competitive advantages. PloS One **7**:e48320.

Appendix 1

This appendix contains material relevant to the previous chapters.

The following tables are referenced by the previous chapters but have been placed in the appendix for brevity.

Chapter 2:
- Table A1.1. Primers and strains used in this study
- Table A1.2. Strains used for generation of data in this study

Chapter 3:
- Table A1.3. Strains and plasmids used in this study

Chapter 4:
- Table A1.4: Strains used in this study.
- Table A1.5: Primers used in this study.

The following figures are relevant to the previous chapters but were not included in the original manuscripts.

Chapter 2:
- Figure A1.1. Diagram of the dual plasmid reporter and complementation system.
- Figure A1.2. SsrA-SsrB-dependent reporter activity in various media.
- Figure A1.3. Conservation of SsrA platform and PDC-2-domains.
- Figure A1.4. SsrA H48 and H286 mutagenesis data.
- Figure A1.5. Additional SsrA histidine mutagenesis data.

Chapter 3:
- Figure A1.6. Conservation of *S. enterica* SPI-2 T3SS and *S. glossinidius* SSR-3.

Chapter 4:
- Figure A1.7. Complementation of *S. enterica ssrB* with *S. glossinidius SG1279*
- Figure A1.8. Alignment of *S. bongori* SPI-2 T6SS with those of *Pseudomonas spp.*
- Figure A1.9. Ancestral T6SS associated genes in *S.* Typhimurium and *S. bongori.*
- Figure A1.10. Transcriptional activity of T6SS promoters in regulatory mutants.

Table A1.1. Primers used in this study

Name	Sequence	Description
DTM_0177_F	ACGCGTCGACacataaatgggagtttctatcaaattc	PssrA-SsrA forward cloning primer (SalI)
DTM_0177_R	GCTCTAGAttaagtaatggtgtagtttttgaagatca	PssrA-SsrA reverse cloning primer (xbaI)
DTM_0208_R	GCTCTAGAtta**cttgtcgtcatcgtccttgtagtc**agtaatggtgtagtttttgaagatcat	ssrA flag tag, primer, use with 17F or 177F
DTM_285_F	**caacaataattatttggctgctatctgtgcttaccgcagcttatatatca**gtgtaggctggagctgcttcg	amplification of pKD4 cassette for periplasmic ssrA domain disruption
DTM_285_R	**aaccagcaaaaagccgacgtcatcaacaccaatgctgttaatgtaaaggg**catatgaatatcctcctta	
DTM_301_F	atgaataaatttgctcgccattttctgccgctgtcgctgcgggttcggtgtaggctggagctgcttcg	phoQ lambda red primers to generate phoQ mutant
DTM_301_R	ttattcctctttctgtgtgggatgctgtcggccaaaaacgacctccataccatatgaatatcctcctta	
DTM_302_F	ctgcgggaaagtcataccat	phoQ deletion screening primers
DTM_302_R	caagaaagtcgggccagtta	
DTM_0179_F	atctcagtcgctaatga**ggc**ttgatacattaaatttttag	SsrA C82A mutant (2)
DTM_0179_R	aaaatttaatgtatcaa**gcc**tcattagcgactgagattca	
DTM_0180_F	ccgtttagcgtcggcgt**ggc**atttgaaggaccgacagata	SsrA C108A mutant (3)
DTM_0180_R	ctgtcggtccttcaaat**gcc**acgccgacgctaaacggaga	
DTM_0181_F	tcatcgatatcagagga**ggc**cagaaagagacggtgcttct	SsrA Q122A mutant (4)
DTM_0181_R	agcaccgtctctttctg**gcc**tcctctgatatcgatgaaaa	
DTM_0182_F	tgattaagaataaaact**ggc**gcgacgaaagctattttcat	SsrA D134A mutant (5)
DTM_0182_R	aaaatagctttcgtcgc**gcc**agttttattcttaatcataa	
DTM_0183_F	aaaggctgtagagttga**ggc**atctgaagggttatcagtag	SsrA Y155A mutant (6)
DTM_0183_R	ctgataacccttcagat**gcc**tcaactctacagcctttaac	
DTM_0184_F	ctccagtaaaacccggc**ggc**ggttgggtataaaggaaagc	SsrA H172A mutant (7)
DTM_0184_R	ttcctttatacccaacc**gcc**gccgggttttactggagtga	
DTM_0185_F	**ggc**tcctttgccgtttatgtattctggttcact**ggc**gtaaaacccggcatgggttg	SsrA W177A,W188A mutant (8)
DTM_0185_R	**gcc**agtgaaccagaatacataaacggcaaagga**gcc**cacgcttccgttgcggttgc	
DTM_0186_F	aataagtggttgttttg**ggc**cagccatactcgaatactat	SsrA D228A mutant (9)
DTM_0186_R	gtattcgagtatggctg**gcc**caaaacaaccacttattgcc	
DTM_0187_F	**ggc**tccggggccatgcaaggttgtgcgtaatatcagaaatccgggaatttgctg**ggc**tccatcatgcagcgttacat	SsrA W256A,W274A mutant (10)
DTM_0187_R	**gcc**cagcaaattcccggatttctgatattacgcacaaccttgcatggccccgga**gcc**agtctggttacgctgtaccc	
DTM_0188_F	gagcgatgtagtaacca**ggc**aaaagccgacgtcatcaaca	SsrA C312A mutant (11)
DTM_0188_R	tgatgacgtcggcttttg**gcc**tggttactacatcgctcact	
DTM_219_F	tagatgatagtattcgagta**gcc**ctggatcaaaacaaccacttattg	SsrA W226A (15)
DTM_219_R	caataagtggttgttttgatccag**ggc**tactcgaatactatcatcta	
DTM_220_F	ttcccggatttctgatatta**gcc**acaaccttgcatggccccggatg	SsrA R266A (16)
DTM_220_R	catccggggccatgcaaggttgt**ggc**taatatcagaaatccgggaa	
DTM_226_F	ggtttgaagaagctgaacgt**gcc**gctaaaaatttaatgtatca	SsrA D74A (22)
DTM_226_R	tgatacattaaatttttagc**ggc**acgttcagcttcttcaaacc	
DTM_227_F	atcaacggtttgaagaagct**gcc**cgtgacgctaaaaatttaatgta	SsrA E72A (23)
DTM_227_R	tacattaaatttttagcgtcacg**ggc**agcttcttcaaaccgttgat	
DTM_230_F	acataaacggcaaaggatgg**gcc**gcttccgttgcggttgccga	SsrA H189A

DTM_230_R	tcggcaaccgcaacggaagc**ggc**ccatcctttgccgtttatgt	
DTM_231_F	ccgatctcattactaagagc**gcc**ctgccattagatgatagtat	SsrA H216A
DTM_231_R	atactatcatctaatggcag**ggc**gctcttagtaatgagatcgg	
DTM_232_F	tatggctggatcaaaacaac**gcc**ttattgccgttttcatacat	SsrA H232A
DTM_232_R	atgtatgaaaacggcaataa**ggc**gttgttttgatccagccata	
DTM_233_F	agttagaaaatgtaacgctg**gcc**gatggatggcagcaaattcc	SsrA H253A
DTM_233_R	ggaatttgctgccatccatc**ggc**cagcgttacattttctaact	
DTM_234_F	tgatattacgcacaaccttg**gcc**ggccccggatggagtctggt	SsrA H270A
DTM_234_R	accagactccatccggggcc**ggc**caaggttgtgcgtaatatca	
DTM_241_F	tgcatgatggatg**ggcc**caaattcccgggtttctgatattacgcacaac	SsrA Q257A (XmaI)
DTM_241_R	cgtaatatcagaaacccgggaatttg**ggc**ccatccatcatgcagcgtta	
DTM_270_F	caatggttcagaaacggcag**gcg**ataatagaggatttatccgttc	SsrA H48A
DTM_270_R	gaacggataaatcctctattat**cgc**ctgccgtttctgaaccattg	
DTM_271_F	caatgctcattagcgactgagatt**gcggcg**aacgatattttccctgaggtgagc	SsrA H89A H90A
DTM_271_R	gctcacctcagggaaaatatcgtt**cgccgc**aatctcagtcgctaatgagcattg	
DTM_272_F	cgatattttccctgaggtgagccgg**gcg**ctatctgtcggtccttcaaattgc	SsrA H100A
DTM_272_R	gcaatttgaaggaccgacagatag**cgc**ccggctcacctcagggaaaatatcg	
DTM_273_F	cgacgctaaacggagagaag**gcg**cgtctctttctgcagtcctc	SsrA H177A
DTM_273_R	gaggactgcagaaagagacg**cgc**cttctctccgtttagcgtcg	
DTM_274_F	gctttcgtcgcgatagttttattcttaat**gcg**aaaaatgagatttcgttattatctac	SsrA H140A
DTM_274_R	gtagataataacgaaatctcatttttt**cgc**attaagaataaaactatcgcgacgaaagc	
DTM_275_F	cgctgtacccatacggtaatcta**gcg**aatcgcatcttaaaaattatccttc	SsrA H286A
DTM_275_R	gaaggataatttttaagatgcgatt**cgc**tagattaccgtatgggtacagcg	
DTM_313_F	caacccatgccgggttttac**gcg**agtgaaccagaatacataaacgg	SsrA W177A
DTM_313_R	ccgtttatgtattctggttcact**cgc**gtaaaacccggcatgggttg	
DTM_314_F	gaatacataaacggcaaagga**gcg**cacgcttccgttgcggttgccg	SsrA W188A
DTM_314_R	cggcaaccgcaacggaagcgtg**cgc**tcctttgccgtttatgtattc	
DTM_315_F	gtaacgctgcatgatgga**gcg**cagcaaattcccggatttctg	SsrA W256A
DTM_315_R	cagaaatccgggaatttgctg**cgc**tccatcatgcagcgttac	
DTM_316_F	caaccttgcatggccccgga**gcg**agtctggttacgctgtacccatac	SsrA W274A
DTM_316_R	gtatgggtacagcgtaaccagact**cgc**tccggggccatgcaaggttg	
DTM_325_F	gaatattgtactaagcaatcaa**gcg**tttgaagaagctgaacgtgac	SsrA R67A
DTM_325_R	gtcacgttcagcttcttcaaa**cgc**ttgattgcttagtacaatattc	
DTM_326_F	gtactaagcaatcaacgg**gcg**gaagaagctgaacgtgacgc	SsrA F68A
DTM_326_R	gcgtcacgttcagcttcttc**cgc**ccgttgattgcttagtac	
DTM_327_F	cgatattttccctgaggtg**gcg**cggcatctatctgtcggtcc	SsrA S98A
DTM_327_R	ggaccgacagatagatgccg**cgc**cacctcagggaaaatatcg	
DTM_328_F	cttaatcataaaaatgagatt**gcg**ttattatctactgataaccc	SsrA S145A
DTM_328_R	gggttatcagtagataataa**cgc**aatctcatttttatgattaag	
DTM_329_F	ggttttactggagtgaacca**gcg**tacataaacggcaaaggatg	SsrA E181A
DTM_329_R	catcctttgccgtttatgta**cgc**tggttcactccagtaaaacc	

Table A1.2. Strains used for generation of data in this study

Genotype	Shorthand	Plasmid 1	Plasmid 2
wild type	wt		
del ssrA			
ssrA:ssrA10	ssrA10		
ushA:cm10			
wild type			pGEN:PsseA-lux
del ssrA			pGEN:PsseA-lux
ssrA:ssrA10			pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA	pGEN:PsseA-lux
del ssrA		pWSK129	pGEN:PsseA-lux
del ssrAB		pWSK129	pGEN:PsseA-lux
del ssrA del phoQ		pWSK129:PssrA-ssrA	pGEN:PsseA-lux
del ssrA del phoQ		pWSK129	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA E72A,D74A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA C82A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA C108A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA Q122A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA D134A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA Y155A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA H172A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA D228A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA C312A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA W226A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA R266A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA D74A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA E72A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA H189A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA H216A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA H232A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA H253A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA H270A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA Q257A (XmaI)	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA H48A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA H89A H90A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA H100A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA H177A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA H140A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA H286A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA W177A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA W188A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA W256A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA W274A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA R67A	pGEN:PsseA-lux

del ssrA		pWSK129:PssrA-ssrA F68A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA S98A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA S145A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA E181A	pGEN:PsseA-lux
del ssrA	N5	pWSK129:PssrA-ssrA H89A H90A H100A H117A H140A	pGEN:PsseA-lux
del ssrA	C2	pWSK129:PssrA-ssrA H172A H253A	pGEN:PsseA-lux
del ssrA		pWSK129:PssrA-ssrA H172A H253A H270A	pGEN:PsseA-lux
del ssrA	C4	pWSK129:PssrA-ssrA H172A H232A H253A H270A	pGEN:PsseA-lux
del ssrA	C5	pWSK129:PssrA-ssrA H172A H216A H232A H253A H270A	pGEN:PsseA-lux
del ssrA	10	pWSK129:PssrA-ssrA H89A H90A H100A H117A H140A H172A H216A H232A H253A H270A	pGEN:PsseA-lux
del ssrA	11	pWSK129:PssrA-ssrA H89A H90A H100A H117A H140A H172A H216A H232A H253A H270A H286A	pGEN:PsseA-lux
del ssrA	C3	pWSK129:PssrA-ssrA H172A H216A H253A	pGEN:PsseA-lux

Table A1.3. Strains and plasmids used in this study

Strain or plasmid	Genotype or description	Reference
Strains		
DH5α	supE44 Dlacu169 (f80 lacZDM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1	Lab strain
DH5α λpir	supE44 Dlacu169 (f80 lacZDM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1	λpir Lab strain
E. coli SM10 λpir	thi recA thr leu tonA lacY supE RP4–2–Tc::Mu λpir	Lab strain
E. coli B1H	uso F' pyrF hisB; bacterial one-hybrid host strain	[1]
SL1344	Wild-type S. enterica sv. Typhimurium, SmR	[2]
SL1344 ssrB::aphT	Marked deletion of ssrB, KanR	[3]
SL1344 ΔssrB	Unmarked, in-frame deletion of ssrB	This work
SL1344 ΔssaE	Unmarked, in-frame deletion of ssaE	This work
SL1344 ssrB::3FLAG-kan	Encodes chromosomally FLAG-tagged ssrB, KanR	[4]
SL1344 PssaG::pIVET-PssaG	Merodiploid containing integrated PssaG-tnpR-lacZ reporter	This work
SL1344 PssaG::pIVET-PssaG-X47	Merodiploid containing integrated PssaG-X47-tnpR-lacZ reporter	This work
SL1344 PssaG::pIVET-PssaG-7X7	Merodiploid containing integrated PssaG-7X7-tnpR-lacZ reporter	This work
SL1344 PssaG::pIVET-PssaG-74X	Merodiploid containing integrated PssaG-74X-tnpR-lacZ reporter	This work
SL1344 PssaG::pIVET-PssaG-XXX	Merodiploid containing integrated PssaG-XXX-tnpR-lacZ reporter	This work
SL1344 PssaG::pIVET-PssaG-H2H	Merodiploid containing integrated PssaG-H2H-tnpR-lacZ reporter	This work
SL1344 PssaG::pIVET-PssaG-Rev	Merodiploid containing integrated PssaG-Rev-tnpR-lacZ reporter	This work
SL1344 PssaG::pIVET-PssaG-7"-4-7'	Merodiploid containing integrated PssaG-7"-4-7'-tnpR-lacZ reporter	This work
SL1344 PsseA::pIVET-PsseA	Merodiploid containing integrated PsseA-tnpR-lacZ reporter	This work
SL1344 PsseA::pIVET-PsseAdel	Merodiploid containing integrated PsseAdel-tnpR-lacZ reporter	This work
SL1344 ΔssrB PsseA::pIVET-PsseA	ssrB mutant, merodiploid containing integrated PsseA-tnpR-lacZ reporter	This work
SL1344 ΔssrB PsseA::pIVET-PsseAdel	ssrB mutant, merodiploid containing integrated PsseAdel-tnpR-lacZ reporter	This work
SL1344 ΔssaE PsseA::pIVET-PsseA	ssaE mutant, merodiploid containing integrated PsseA-tnpR-lacZ reporter	This work
SL1344 ΔssaE PsseA::pIVET-PsseAdel	ssaE mutant, merodiploid containing integrated PsseA-tnpR-lacZdel reporter	This work
SL1344 PssaR::pIVET-PssaR	Merodiploid containing integrated PssaR-tnpR-lacZ reporter	This work
Plasmids		
pB1H1	Bacterial one-hybrid bait plasmid, CmR	[1]
pB1H1-ssrBc	Bait plasmid carrying the C-terminal domain of ssrB (ssrBc), CmR	This work
pB1H1-phoP	Bait plasmid carrying E. coli phoP, CmR	This work
pH3U3	Bacterial one-hybrid prey plasmid, KanR	[1]
pIVET5n	tnpR-lacZ, sacB, R6K ori, bla, AmpR	[5]
pCS26	pSC101 ori, luxCDABE, KanR	[6]
pIVET-PssaG	ssaG promoter fused to tnpR-lacZ in pIVET5n, AmpR	This work
pIVET-PssaG-X47	ssaG promoter with "X-4-7" mutation fused to tnpR-lacZ in pIVET5n, AmpR	This work
pIVET-PssaG-7X7	ssaG promoter with "7-X-7" mutation fused to tnpR-lacZ in pIVET5n, AmpR	This work
pIVET-PssaG-74X	ssaG promoter with "7-4-X" mutation fused to tnpR-lacZ in pIVET5n, AmpR	This work
pIVET-PssaG-XXX	ssaG promoter with "X-X-X" mutation fused to tnpR-lacZ in pIVET5n, AmpR	This work

pIVET-PssaG-H2H	ssaG promoter with "H2H" mutation fused to tnpR-lacZ in pIVET5n, AmpR	This work
pIVET-PssaG-Rev	ssaG promoter with 'Rev" mutation fused to tnpR-lacZ in pIVET5n, AmpR	This work
pIVET-PssaG-7"-4-7'	ssaG promoter with "7"-4-7"' mutation fused to tnpR-lacZ in pIVET5n, AmpR	This work
pIVET-PsseA	sseA promoter fused to tnpR-lacZ in pIVET5n, AmpR	This work
pIVET-PsseAdel	sseA promoter with motif deletion mutation fused to tnpR-lacZ in pIVET5n, AmpR	This work
pIVET-PssaR	ssaR promoter fused to tnpR-lacZ in pIVET5n, AmpR	This work
pCS26-PssaG	ssaG promoter fused to luxCDABE in pCS26, KanR	This work
pCS26-PssaG-XXX	ssaG promoter with “X-X-X” deletion fused to luxCDABE in pCS26, KanR	This work
pCS26-PSG1292	Sodalis glossinidius SG1292 promoter fused to luxCDABE in pCS26, KanR	This work
pCS26-PSG1292-XXX	Sodalis glossinidius SG1292 promoter with “X-X-X” deletion fused to luxCDABE in pCS26, KanR	This work
pCS26-PssaR	ssaR promoter fused to luxCDABE in pCS26, KanR	This work
pCS26-PssaG mut(1-9)	Scrambled ssaG promoter fused to luxCDABE in pCS26, KanR	This work

Reference list for Table

1. Meng X, Brodsky MH, Wolfe SA (2005) A bacterial one-hybrid system for determining the DNA-binding specificity of transcription factors. Nat Biotechnol 23: 988-994.
2. Wray C, Sojka WJ (1978) Experimental Salmonella typhimurium infection in calves. Res Vet Sci 25: 139-143.
3. Knodler LA, Celli J, Hardt WD, Vallance BA, Yip C, et al. (2002) Salmonella effectors within a single pathogenicity island are differentially expressed and translocated by separate type III secretion systems. Mol Microbiol 43: 1089-1103.
4. Duong N, Osborne S, Bustamante VH, Tomljenovic AM, Puente JL, et al. (2007) Thermosensing coordinates a cis-regulatory module for transcriptional activation of the intracellular virulence system in Salmonella enterica serovar Typhimurium. J Biol Chem 282: 34077-34084.
5. Brown NF, Vallance BA, Coombes BK, Valdez Y, Coburn BA, et al. (2005) Salmonella Pathogenicity Island 2 Is Expressed Prior to Penetrating the Intestine. PLoS Pathog 1:e32.
6. Beeston AL, Surette MG (2002) pfs-dependent regulation of autoinducer 2 production in Salmonella enterica serovar Typhimurium. J Bacteriol 184: 3450-3456.

Table A1.4: Strains used in this study.

Strains	Relevant Characteristics	Source
T6SS mutant strains	**Locus ID**	
ΔsciG	ΔSTM0272	This study
ΔsciG complementation	ΔSTM0272, pWSK29:PSTM0272-STM0272	This study
ΔsciS	ΔSTM0285	This study
Δcluster 1	ΔSTM0274A-STM0278	This study
Δcluster 2	ΔSTM0283-STM0284	This study
Δcluster 3	ΔSTM0286-STM0288	This study
Δcluster 4	ΔSTM0290-STM0298	This study
ΔsciK	STM0276::kn	This study
ΔsciL	STM0277::kn	This study
ΔSTM0278	STM0278::kn	This study
ΔsciQ	ΔSTM0283	This study
ΔsciR	ΔSTM0284	This study
ΔsciT	ΔSTM0286	This study
ΔsciU	ΔSTM0287	This study
ΔsciV	ΔSTM0288	This study
ΔsciW	ΔSTM0290	This study
Δrhs1	ΔSTM0291	This study
Δrhs2	ΔSTM0292	This study
Additional mutant strains	**Locus ID**	
Δhfq	STM4361::kn	This study
ΔsseC	STM1400::kn	This study
ΔpmrA	ΔSTM4292	This study
ΔssrAB	ΔSTM1392-STM1391	This study
ΔphoP	STM1231::Cm	This study
ΔqseC	STM3178::Cm	This study
ΔompR	ΔSTM3502	This study
Δhfq	STM4361::Cm	This study
ΔinvA	STM2896::Kn	This study
HA chromosomal epitope fusion expression strains	**Multiple strain backgrounds**	
STM0276::HA-Kn	wt, phoP::Cm, ΔssrAB, qseC::Cm, ΔompR, hfq::Cm	This study
STM0279::HA-Kn	wt, phoP::Cm, ΔssrAB, qseC::Cm, ΔompR, hfq::Cm	This study
STM0289::HA-Kn	wt, phoP::Cm, ΔssrAB, qseC::Cm, ΔompR, hfq::Cm	This study
STM3131::HA-Kn	wt	This study
STM0276::Cm-Kn	invA::kn	This study
STM0279::Cm-Kn	invA::kn	This study
STM0289::Cm-Kn	invA::kn	This study
STM3131::Cm-Kn	invA::kn	This study

Flag ectopic epitope fusion expression strains	**Multiple strain backgrounds**	
pCTC-Flag:STM0279	wt, ΔSTM0272, pCTC-Flag:STM0279	This study
pCTC-Flag:STM0289	wt, ΔSTM0272, pCTC-Flag:STM0289	This study
pCTC-Flag:STM3131	wt, ΔSTM0272, pCTC-Flag:STM3131	This study

Ectopic luciferase transcriptional reporter strains	**Multiple strain backgrounds**	
pGEN-LuxCDABE:Pem7	wt, pGEN-LuxCDABE:Pem7	This study
pGEN-LuxCDABE:PssaG	wt, pGEN-LuxCDABE:PssaG	This study
pGEN-LuxCDABE:Pstm0271	wt, pGEN-LuxCDABE:Pstm0271	This study
pGEN-LuxCDABE:Pstm0272	wt, pGEN-LuxCDABE:Pstm0272	This study
pGEN-LuxCDABE:Pstm0273	wt, pGEN-LuxCDABE:Pstm0273	This study
pGEN-LuxCDABE:Pstm0276	wt, pGEN-LuxCDABE:Pstm0276	This study
pGEN-LuxCDABE:Pstm0277	wt, pGEN-LuxCDABE:Pstm0277	This study
pGEN-LuxCDABE:Pstm0279	wt, pGEN-LuxCDABE:Pstm0279	This study
pGEN-LuxCDABE:Pstm0280	wt, pGEN-LuxCDABE:Pstm0280	This study
pGEN-LuxCDABE:Pstm0284	wt, pGEN-LuxCDABE:Pstm0284	This study
pGEN-LuxCDABE:Pstm0289	wt, pGEN-LuxCDABE:Pstm0289	This study

Merodiploid lacZ transcriptional reporter strains	**Multiple strain backgrounds**	
PSTM0271:PSTM0271-tnpR-lacZ	wt, ΔssrAB, phoP::cm, slyA::kn, ΔompR, hfq::kn, ΔqseC	This study
PSTM0272:PSTM0272-tnpR-lacZ	wt, ΔssrAB, phoP::cm, slyA::kn, ΔompR, hfq::kn, ΔqseC	This study
PSTM0273:PSTM0273-tnpR-lacZ	wt, ΔssrAB, phoP::cm, slyA::kn, ΔompR, hfq::kn, ΔqseC	This study
PSTM0277:PSTM0277-tnpR-lacZ	wt, ΔssrAB, phoP::cm, slyA::kn, ΔompR, hfq::kn, ΔqseC	This study
PSTM0284:PSTM0284-tnpR-lacZ	wt, ΔssrAB, phoP::cm, slyA::kn, ΔompR, hfq::kn, ΔqseC	This study
PSTM0289:PSTM0289-tnpR-lacZ	wt, ΔssrAB, phoP::cm, slyA::kn, ΔompR, hfq::kn, ΔqseC	This study
PSTM3131:PSTM3131-tnpR-lacZ	wt, ΔpmrA	This study
PpmrD:PpmrD-tnpR-lacZ	wt	This study
PpmrC:PpmrC-tnpR-lacZ	wt	This study

Plasmids	**Description**	
pKD46	lambda red helper plasmid	Datsenko (Wanner) (12)
pKD3	lambda red template plasmid	Datsenko (Wanner) (12)
pKD4	lambda red template plasmid	Datsenko (Wanner) (12)
pCP20	lambda red curing plasmid	Datsenko (Wanner) (12)
pSUB315	lambda red template plasmid	Uzzau (Bossi) (54)
pGEN-LuxCDABE	luciferase reporter plasmid	Lane (Mobley) (28)
pWSK29	expression plasmid	Coombes (Finlay) (56)
pIVET5n	integrative lacZ reporter plasmid	Coombes (Finlay) (10)

Table A1.5: Primers used in this study.

Primer	Description	Sequence
Lambda red mutagenesis primers		
DTM_0011_F	STM0272 F	cctgtttcacgcagtgcgttgtatggaaaactggccggcccactattccgggtgtaggctggagctgcttcg
DTM_0011_R	STM0272 R	gatctcaaaaacaatctgctcatcaccattaacgccgatatcaatctgcctcatatgaatatcctcctta
DTM_0012_F	STM0285 F	tttcttagtctgcttttttcccggcgcgcgctggcagttgtgggcgttctggtgtaggctggagctgcttcg
DTM_0012_R	STM0285 R	cctcgccgggcaactgaagttctgcaacagctcccggctgaccgggttaaacatatgaatatcctcctta
DTM_0005_F	pmrA F	ctgattgttgaagacgacacgctattattacaggggttaatactcgccgtgtaggctggagctgcttcg
DTM_0005_R	pmrA R	ctcagtggcaaccagcatgtagccaaacccgcgaaccgtgcgaatgcgcatatgaatatcctcctta
DTM_0057_F	STM0291 F	tatgaagcagcccgtgtggatgatcctatctaccacaccagcgcgctcgcgtgtaggctggagctgcttcg
DTM_0057_R	STM0291 R	tgcttcaacttccccattaattttctcgccccttacagttatgttttctgcatatgaatatcctcctta
DTM_0081_F	CDS STM0274A-STM0278 F	acaaaactgccgtcagtgaagcagggcaatgcctgatatatattttgtgaGTGTAGGCTGGAGCTGCTTCG
DTM_0081_R	CDS STM0274A-STM0278 R	ccataaattcatcagagcaagagaacgtatgcttttaatatctttgaattCATATGAATATCCTCCTTA
DTM_0082_F	CDS STM0283-STM0284 F	gggcgaaaaaccgtcgggtggaaattctggtagtggagtgaatgaGTGTAGGCTGGAGCTGCTTCG
DTM_0082_R	CDS STM0283-STM0284 R	gcagactaagaaatttttgcattatatttcgactcgccacccaataatcaatattCATATGAATATCCTCCTTA
DTM_0083.1_F	CDS STM0286-STM0288 F	tcagccgggagctgttgcagaacttcagttgcccggcgagggcgctgtaaGTGTAGGCTGGAGCTGCTTCG
DTM_0083_R	CDS STM0286-STM0288 R	gtcaaccacagcacgtccagacacgataatttcatccggttttaacggtaCATATGAATATCCTCCTTA
DTM_0084_F	CDS STM0290-STM0298 F	attgccccccttctcacctctgagtgatctgaattaaacctggagttctcGTGTAGGCTGGAGCTGCTTCG
DTM_0084_R	CDS STM0290-STM0298 R	aatagacccgttttagttccatgcatttttaagattccggccagataacgCATATGAATATCCTCCTTA
DTM_0096_F	STM0283 F	aaatcaacctatgatgtgtctcatcattcggcagtatgtggcgtgacggtgtaggctggagctgcttcg
DTM_0096_R	STM0283 R	gggatactctattaatttatttatgataagttcagccgcgcgagaattcatatgaatatcctcctta
DTM_0097_F	STM0284 F	ttggtgggtgttcctgccatttcttatgcgcacgattatgggtgtgctgtgtaggctggagctgcttcg
DTM_0097_R	STM0284 R	ggataaagtaatataaccattaaatcttacagaacactcatcagcattcatatgaatatcctcctta
DTM_0098_F	STM0286 F	gcagtggtatcagccccggcgatgttcggcaggctgccggaccagcgcgtgtaggctggagctgcttcg
DTM_0098_R	STM0286 R	ctgacctctatttttttcatccaccaaccgcagcgcatccacataccgcatatgaatatcctcctta
DTM_0099_F	STM0287 F	gctgtagaggttctgtttatgtctgcctgtgcaggaatgttcctgcccgtgtaggctggagctgcttcg
DTM_0099_R	STM0287 R	tgtacaacgtatcattagcgccagcttctgatcccgtacagactggttcatatgaatatcctcctta
DTM_0100_F	STM0288 F	agactgacttttgacgggcaaaagctgacatggcctggtatcgggatagtgtaggctggagctgcttcg
DTM_0100_R	STM0288 R	accattcgtttgctgaccagaaacatacttaacattaaccgtaaatgtcatatgaatatcctcctta
DTM_0107_F	SL4295 (hfq) forward	caatctttacaagatccgttcctgaacgcattgcgtcgggaacgtgttgtgtaggctggagctgcttcg
DTM_0107_R	SL4295 (hfq) reverse	catcacggtagcaacgcgcaggggtctactgcgcaacaggacagcgaacatatgaatatcctcctta
DTM_0159_F	STM0276 forward	atggcttatgacatttttttgaaaattgacggcattgatggcgagtcaatggtgtaggctggagctgcttcg
DTM_0159_R	STM0276 reverse	aaaacaataaaaccagccagacataacatctggccggaaaaacagccgttacatatgaatatcctcctta
DTM_0160_F	STM0277 forward	atgaacagaccttcattcaatgaagcgtggttagcttttaggaaggtgaatgtgtaggctggagctgcttcg
DTM_0160_R	STM0277 reverse	tcacggcagtatccacagtgtcccaacttcaggaacaaatggtccattatccatatgaatatcctcctta
DTM_0161_F	STM0278 forward	gtggatactgccgtgaaatggttaatattagttactttcagcataagtggagtgtaggctggagctgcttcg
DTM_0161_R	STM0278 reverse	ttacttctcgacacgccttttaattaattcatttaattccctgtcatggatcatatgaatatcctcctta
DTM_0162_F	STM0290 forward	atggatcgaccataccgcatacaggaagggtgttttgtcctgcctgaaacagtgtaggctggagctgcttcg
DTM_0162_R	STM0290 reverse	ttattcgtttttatccggctgaaagcctgccagccaggtgttccagagtagcatatgaatatcctcctta
DTM_0164_F	STM0292 forward	cccatcaatccggagacgcgctaccggtacgatgcgctgggcaggcgggtggtgtaggctggagctgcttcg
DTM_0164_R	STM0292 reverse	ttaacatttatcatctcccttagagtcagagcgatgcggatgttgtgggcccatatgaatatcctcctta

DTM_0209_F	STM0293 forward	atgttaaattcaaatatgtctgaacttagaatcgaactggagaatgcggtgtaggctggagctgcttcg
DTM_0209_R	STM0293 reverse	ttacttaatggttgagcatacaatcaaatcaccatgatcattttcacacatatgaatatcctcctta

Complementation cloning primers

DTM_0158_F	STM0272 xbaI F	gc**tctaga**ggcggcagtaaatacgatgt
DTM_0158_R	STM0272 salI R	acgc**gtcgac**ttagatgaatttctgcgcgctat

Lambda red epitope tagging primers

DTM_0036_F	STM0276 forward	gcggcggcaccatcaccgcaggctacgacttcaaggccaacaaagaaattTATCCGTATGATGTTCC TGA
DTM_0036_R	STM0276 reverse	tcaaaacaataaaaccagccagacataacatctggccggaaaaacagccgATATGAATATCCTCCTT AG
DTM_0037_F	STM0279 forward	gcggcggcaccatcaccgcaggctacgacttcaaggccaacaaagaaattTATCCGTATGATGTTCC TGA
DTM_0037_R	STM0279 reverse	atcaaaataataaatccggccagatataaatctggccggaaaaacagccgATATGAATATCCTCCTT AG
DTM_0038_F	STM0289 forward	gcggcacagccgtaaccgccgatgatattgccccccttctcacctctgagTATCCGTATGATGTTCCT GA
DTM_0038_R	STM0289 reverse	ttcctgtatgcggtatggtcgatccatgagaactccaggtttaattcagaATATGAATATCCTCCTTAG
DTM_0039_F	STM3131 forward	aggagtccagaacgtcagtatgggatctggcgatgaaccagatagggaaaTATCCGTATGATGTTCC TGA
DTM_0039_R	STM3131 reverse	ctcaaccaagagcctgtcgaatagcgcgaccggaatggttggtgtcatcaATATGAATATCCTCCTTA G

Flag epitope cloning primers

DTM_0194_F	STM0276 NdeI F	GGAATTC**CATATG**ATGGCTTATGACATTTTTTTGAAAATTG
DTM_0194_R	STM0276 KpnI R	GG**GGTACC**AATTTCTTTGTTGGCCTTGAAG
DTM_0195_F	STM0279 NdeI F	GGAATTC**CATATG**ATGTCTTATGACATTTTTCTGAAAATTG
DTM_0196_F	STM3131 NdeI F	GGAATTC**CATATG**GATGCGATTTTTTTAAAGCTCG
DTM_0196_R	STM3131 KpnI R	GG**GGTACC**TTTCCCTATCTGGTTCATCGC

pIVET cloning primers

DTM_0006_F	P(STM3131) (XhoI) F	ccg**ctcgag**cggcgtatggtttgcaacctgtg
DTM_0006.1_R	P(STM3131) (EcoRI) R	ccg**gaattc**cggttacgcatccatgatttatctcc
DTM_0007_F	P(STM0271) (XhoI) F	ccg**ctcgag**cggggatctggtaatgccgaaga
DTM_0007.1_R	P(STM0271) (EcoRI) R	ccg**gaattc**cggttaggcggcagtaaatacgatgt
DTM_0008_F	P(STM0272) (XhoI) F	ccg**ctcgag**cggggcggcagtaaatacgatgt
DTM_0008.1_R	P(STM0272) (EcoRI) R	ccg**gaattc**cggttagatctggtaatgccgaagaa
DTM_0009_F	P(pmrC) (XhoI) F	ccg**ctcgag**cggcaatcttcgcacgggttaat
DTM_0009.1_R	P(pmrC) (EcoRI) R	ccg**gaattc**cggttacagcgagtttaacggtaggt
DTM_0021_F	P(pmrD) (XhoI) F	ccg**ctcgag**agacgtgaacctcgctgaat
DTM_0022_R	P(pmrD) (EcoRI) R	ccg**gaattc**ttacttaaccaaccattccatag
DTM_0052_F	P(STM0284) (XhoI) F	ccg**ctcgag**ccgattaacgcaggattcat
DTM_0052.1_R	P(STM0284) (MfeI) R	ccg**caattg**ttacaacagacagaccattacct
DTM_0171_F	P(STM0273) (XhoI) F	ccg**ctcgag**tgatatcggcgttaatggtg
DTM_0171_R	P(STM0273) (MfeI) R	ccg**caattg**ttagaatttctgcgcgctattgtt
DTM_0172_F	P(STM0276) (XhoI) F	ccg**ctcgag**acgcaaccaatacctgcttc
DTM_0172_R	P(STM0276) (MfeI) R	ccg**caattg**ttaatcaatgccgtcaattttcaa
DTM_0173_F	P(STM0277) (XhoI) F	ccg**ctcgag**gctacgacttcaaggccaac

DTM_0173_R	P(STM0277) (MfeI) R	ccg**caattg**ttagacggaatgattcaccttc
DTM_0174_F	P(STM0279) (XhoI) F	ccg**ctcgag**gggttgaaatataacggttcg
DTM_0174_R	P(STM0279) (MfeI) R	ccg**caattg**ttagtgcatggtggattcctgat
DTM_0175_F	P(STM0280) (XhoI) F	ccg**ctcgag**gctacgacttcaaggccaac
DTM_0175_R	P(STM0280) (MfeI) R	ccg**caattg**ttactgtgattcgaggtcaggt
DTM_0176_F	P(STM0289) (XhoI) F	ccg**ctcgag**ttctggtcagcaaacgaatg
DTM_0176_R	P(STM0289) (MfeI) R	ccg**caattg**ttaccctcccataccggatttat

pGEN-LuxCDABE cloning primers

DTM_0072_F	P(STM0271) (-) (BamHI)	CG**GGATCC**ggatctggtaatgccgaaga
DTM_0072_R	P(STM0271) (-) (SnaBI)	gct**TACGTA**ttaggcggcagtaaatacgatgt
DTM_0073_F	P(STM0272) (+) (BamHI)	CG**GGATCC**ggcggcagtaaatacgatgt
DTM_0073_R	P(STM0272) (+) (SnaBI)	gct**TACGTA**ttagatctggtaatgccgaagaa
DTM_0074_F	P(STM0284) (-) (BamHI)	CG**GGATCC**ccgattaacgcaggattcat
DTM_0074_R	P(STM0284) (-) (SnaBI)	gct**TACGTA**ttacaacagacagaccattacct
DTM_0075_F	P(STM0289) (-) (BamHI)	CG**GGATCC**ttctggtcagcaaacgaatg
DTM_0075_R	P(STM0289) (-) (SnaBI)	gct**TACGTA**ttaccctcccataccggatttat
DTM_0076_F	P(STM0273) (-) (BamHI)	CG**GGATCC**tgatatcggcgttaatggtg
DTM_0076_R	P(STM0273) (-) (SnaBI)	gct**TACGTA**ttagaatttctgcgcgctattgtt
DTM_0077_F	P(STM0276) (-) (BamHI)	CG**GGATCC**acgcaaccaatacctgcttc
DTM_0077_R	P(STM0276) (-) (SnaBI)	gct**TACGTA**ttaatcaatgccgtcaattttcaa
DTM_0078_F	P(STM0277) (-) (BamHI)	CG**GGATCC**gctacgacttcaaggccaac
DTM_0078_R	P(STM0277) (-) (SnaBI)	gct**TACGTA**ttagacggaatgattcaccttc
DTM_0079_F	P(STM0279) (-) (BamHI)	CG**GGATCC**gggttgaaatataacggttcg
DTM_0079_R	P(STM0279) (-) (SnaBI)	gct**TACGTA**ttagtgcatggtggattcctgat
DTM_0080_F	P(STM0280) (-) (BamHI)	CG**GGATCC**gctacgacttcaaggccaac
DTM_0080_R	P(STM0280) (-) (SnaBI)	gct**TACGTA**ttactgtgattcgaggtcaggt

Figure A1.1. Diagram of the dual plasmid reporter and complementation system.

SsrA is provided *in trans* from a low copy plasmid, SsrB is expressed from the chromosome, which is activated by SsrA in response to extracellular signals and goes on to activate downstream promoters including P*sseA* which is cloned on to a second plasmid as a transcriptional fusion with the luciferase operon *luxCDABE*.

Figure A1.1.

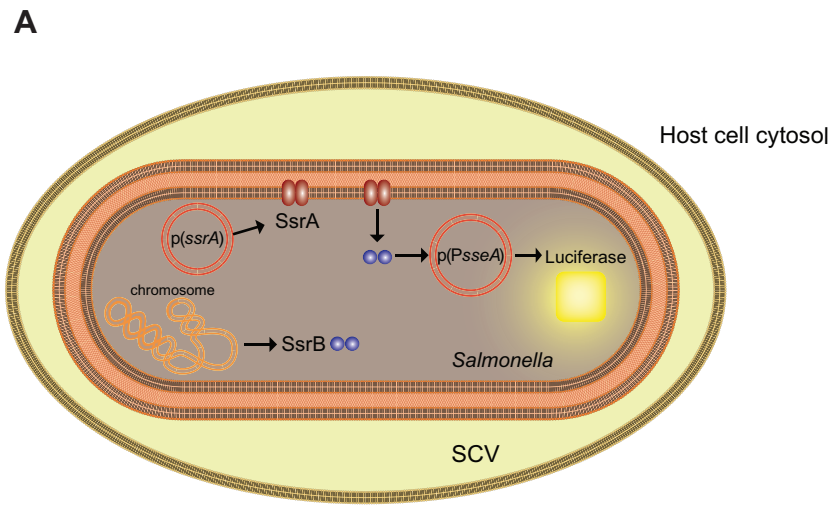

A
Host cell cytosol
p(ssrA)
SsrA
p(PsseA)
Luciferase
chromosome
SsrB
Salmonella
SCV

Figure A1.2. SsrA-SsrB-dependent reporter activity in various media.

(A) SsrA-SsrB-dependent reporter activity in the defined rich media M9 and defined minimal media LPM. Data are from cultures normalized to optical density and expressed as a fold-change above reporter activity from an Δ*ssrA* mutant. Data are the means with standard errors from three independent experiments. (B-C) Reporter activity at 4 hours is decreased by approximately 50% by media supplementation at pH 7 or pH 4.25 but pH response is unaffected. (D-E) Reporter activity of SsrA mutants are similarly affected at pH 7 and pH 4.25 at times indicated. G = Glucose, M = Magnesium, P = Phosphate, N = Sodium, C = Calcium. Data are the means with standard errors from three separate experiments.

Figure A1.2

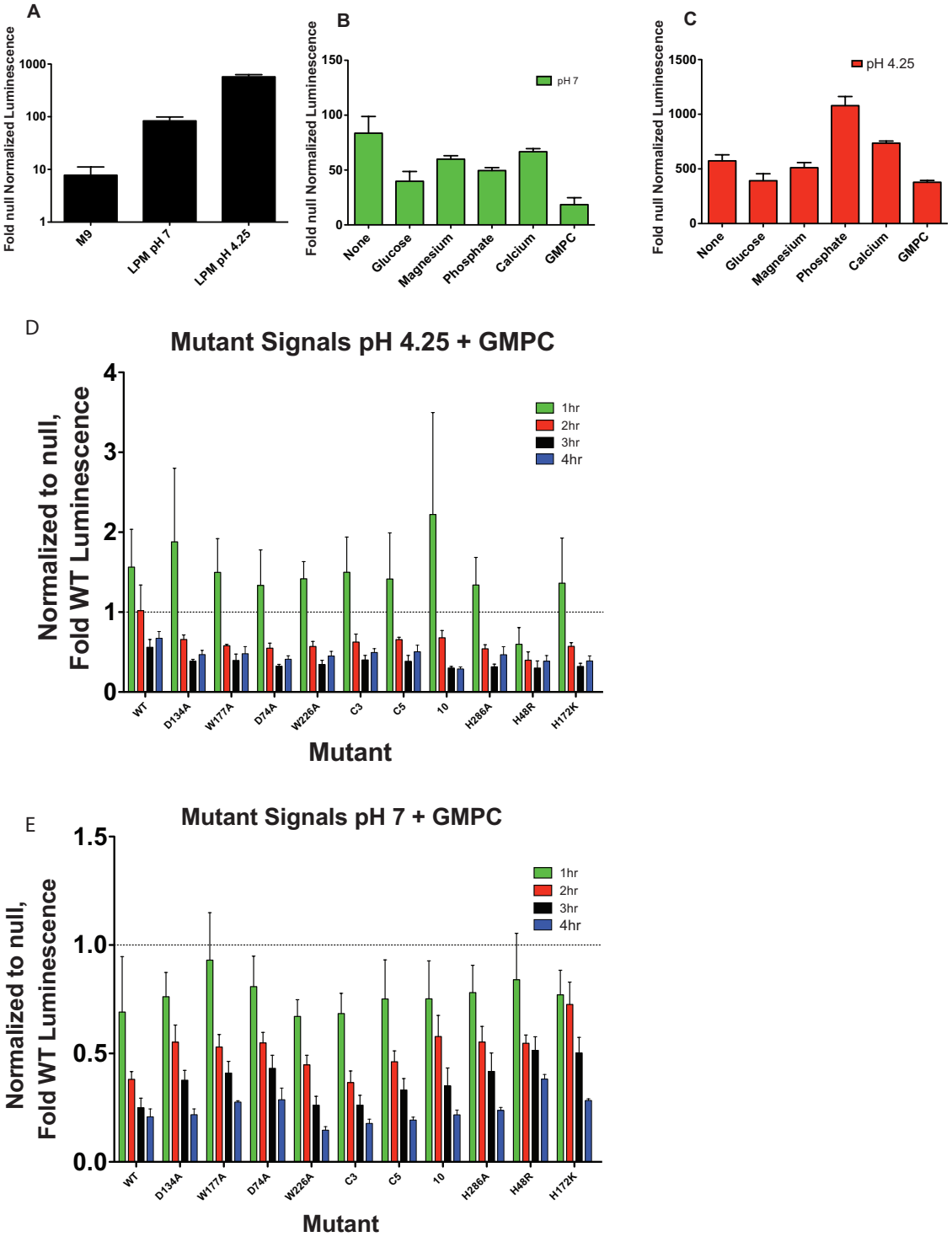

A
Fold null Normalized Luminescence
1000
100
10
1
M9
LPM pH 7
LPM pH 4.25
B
Fold null Normalized Luminescence
150
100
50
0
pH 7
None
Glucose
Magnesium
Phosphate
Calcium
GMPC
C
Fold null Normalized Luminescence
1500
1000
500
0
pH 4.25
None
Glucose
Magnesium
Phosphate
Calcium
GMPC
D
Mutant Signals pH 4.25 + GMPC
Normalized to null, Fold WT Luminescence
4
3
2
1
0
1hr
2hr
3hr
4hr
WT
D134A
W177A
D74A
W226A
C3
C5
10
H286A
H48R
H172K
Mutant
E
Mutant Signals pH 7 + GMPC
Normalized to null, Fold WT Luminescence
1.5
1.0
0.5
0.0
1hr
2hr
3hr
4hr
WT
D134A
W177A
D74A
W226A
C3
C5
10
H286A
H48R
H172K
Mutant

Figure A1.3. Conservation of SsrA platform and PDC-2-domains.

An overlay of sequence alignment information between SsrA and *Salmonella* orthologs (blue) and non-*Salmonella* orthologs (orange) for the platform domain. Structural information is indicated above (α-helix:green, β-sheet:yellow, ligand binding loops:grey). Regions of high conservation only in *Salmonella* are indicated with an asterisk. (C) An overlay of sequence alignment information between SsrA and *Salmonella* orthologs (blue) and non-*Salmonella* orthologs (orange) for the membrane-proximal PDC 2 domain. Structural information is indicated above (α-helix:green, β-sheet:yellow, ligand binding loops:grey). Regions of high conservation only in *Salmonella* are indicated with an asterisk.

Figure A1.3.

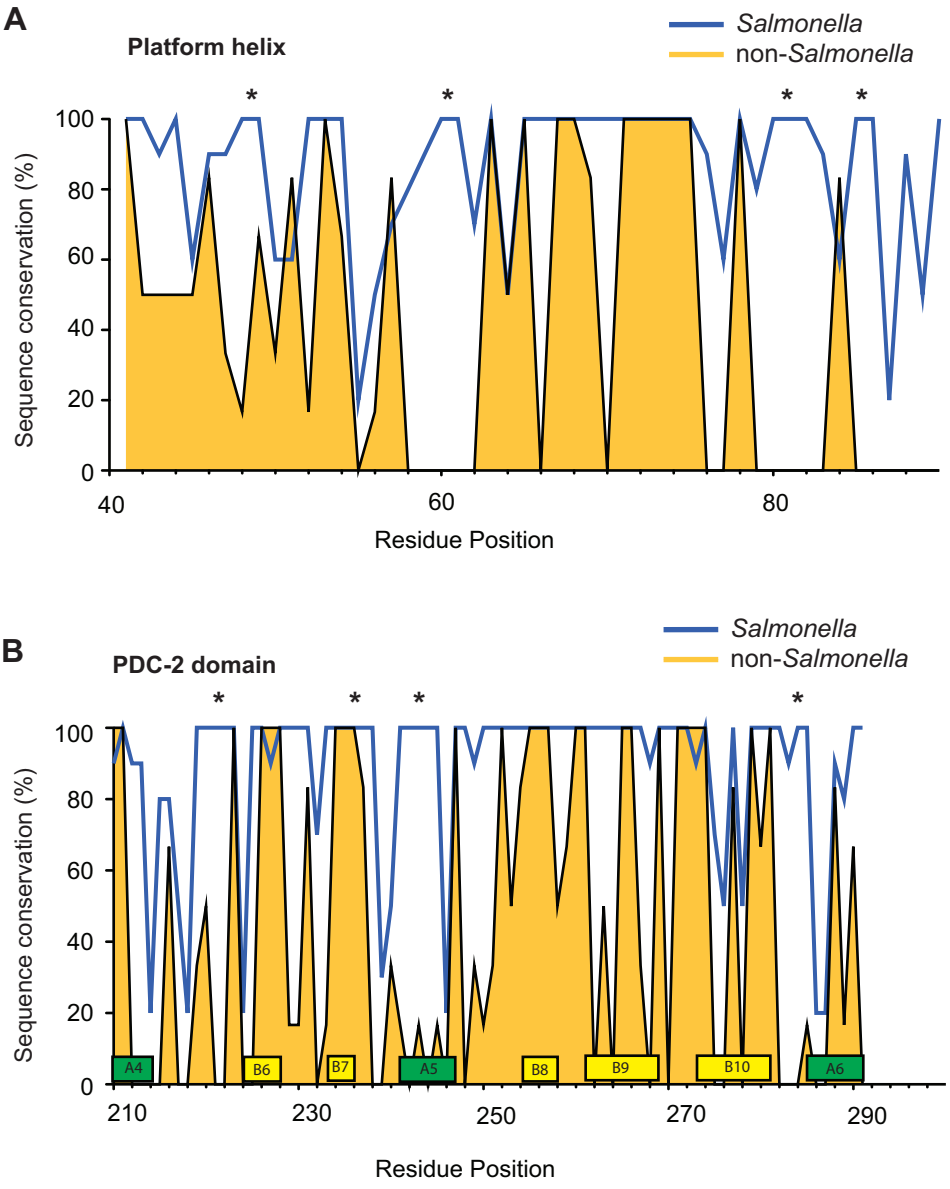

A
Platform helix
Salmonella
non-Salmonella
Sequence conservation (%)
Residue Position
B
PDC-2 domain
Salmonella
non-Salmonella
Sequence conservation (%)
A4
B6
B7
A5
B8
B9
B10
A6
Residue Position

Figure A1.4. SsrA H48 and H286 mutagenesis data.

Mutant reporter activity in LPM media for indicated double plasmid complementation/reporter mutants. (A) H48 mutants at pH 4.25 and pH 7.00. (B) H286 mutants at pH 4.25 and pH 7.00. Reporter activity profiles over the range of pH 4.25 to pH 7.00 with each mutant reading normalized to the pH 7.00 value. Data are from cultures normalized to optical density. Data are the means with standard errors from three separate experiments.

Figure A1.4.

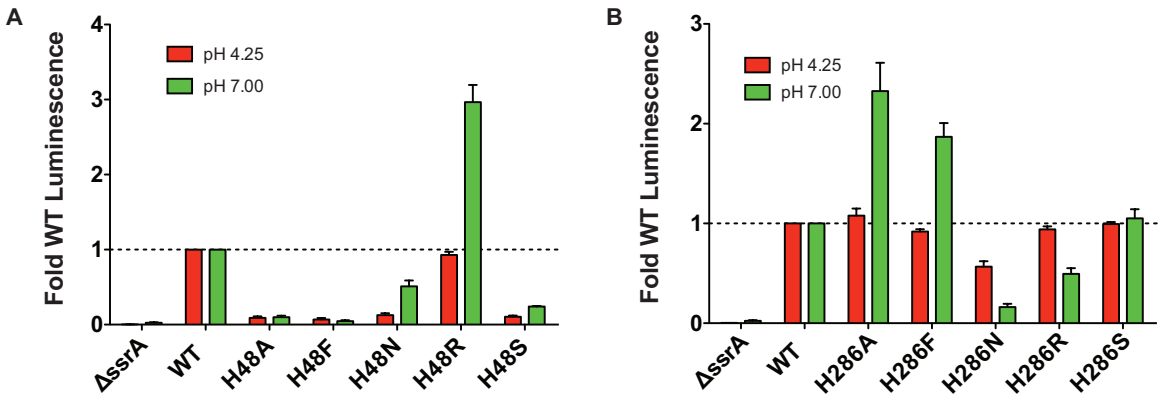

A
Fold WT Luminescence
pH 4.25
pH 7.00
0
1
2
3
4
ΔssrA
WT
H48A
H48F
H48N
H48R
H48S
B
Fold WT Luminescence
pH 4.25
pH 7.00
0
1
2
3
ΔssrA
WT
H286A
H286F
H286N
H286R
H286S

Figure A1.5. Additional SsrA histidine mutagenesis data.

Mutant reporter activity in LPM media for indicated double plasmid complementation/reporter mutants. (A) pH sensing histidine mutagenesis to arginine at pH 4.25 and pH 7.00. (B) Histidine to phenylalanine mutants at pH 4.25 and pH 7.00. (C) pH induction profiles for histidine to phenylalanine mutagenesis. Reporter activity profiles over the range of pH 4.25 to pH 7.00 with each mutant reading normalized to the pH 7.00 value. Data are from cultures normalized to optical density. Data are the means with standard errors from three separate experiments. (D) H172 mutagenesis at pH 4.25 and pH 7.00. (E) pH induction profiles for H172 mutagenesis. Reporter activity profiles over the range of pH 4.25 to pH 7.00 with each mutant reading normalized to the pH 7.00 value. Data are from cultures normalized to optical density. Data are the means with standard errors from three separate experiments.

Figure A1.5.

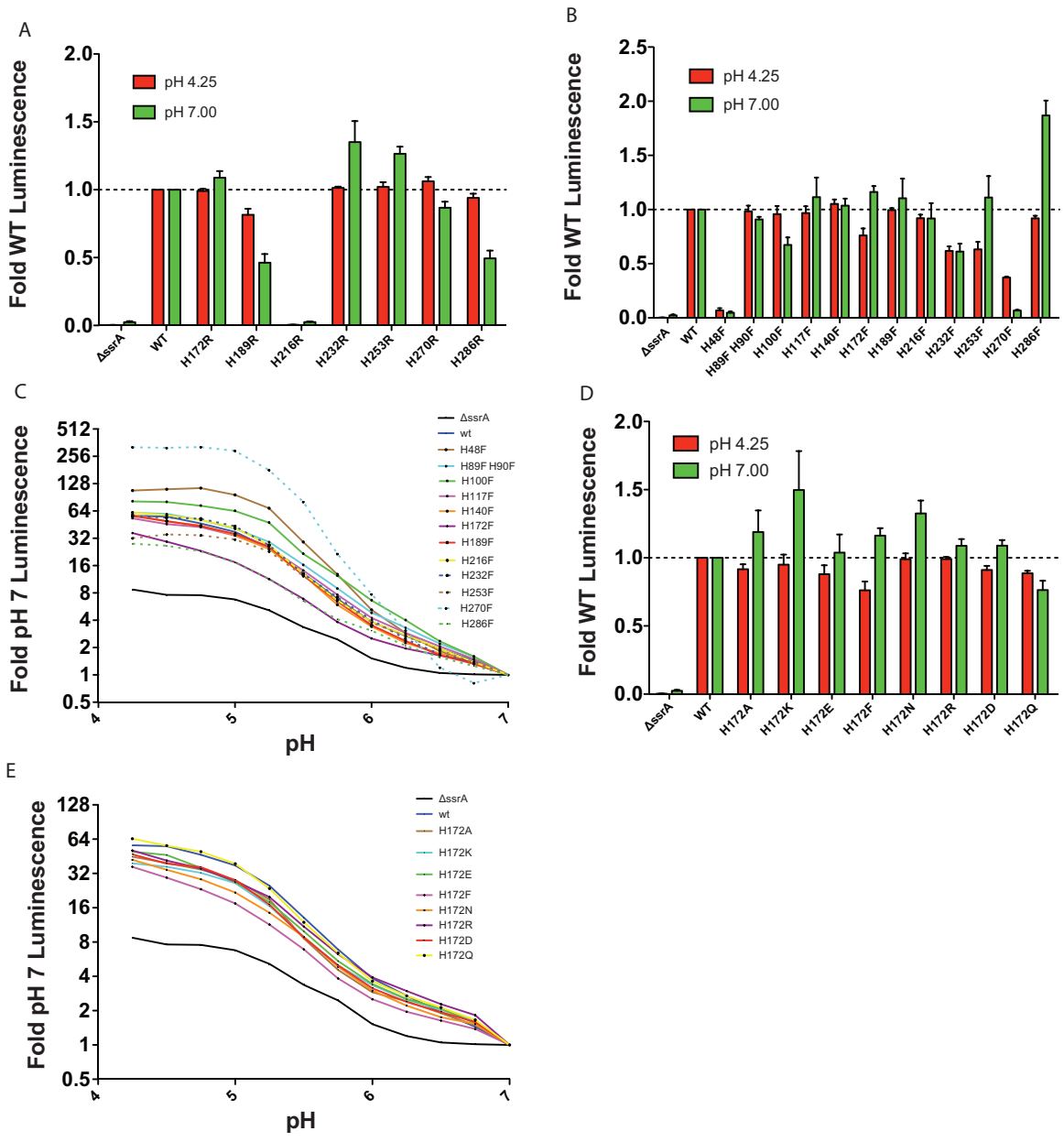

A
Fold WT Luminescence
pH 4.25
pH 7.00
ΔssrA
WT
H172R
H189R
H216R
H232R
H253R
H270R
H286R
B
Fold WT Luminescence
pH 4.25
pH 7.00
ΔssrA
WT
H48F
H89F H90F
H100F
H117F
H140F
H172F
H189F
H216F
H232F
H253F
H270F
H286F
C
Fold pH 7 Luminescence
pH
ΔssrA
wt
H48F
H89F H90F
H100F
H117F
H140F
H172F
H189F
H216F
H232F
H253F
H270F
H286F
D
Fold WT Luminescence
pH 4.25
pH 7.00
ΔssrA
WT
H172A
H172K
H172E
H172F
H172N
H172R
H172D
H172Q
E
Fold pH 7 Luminescence
pH
ΔssrA
wt
H172A
H172K
H172E
H172F
H172N
H172R
H172D
H172Q

Figure A1.6. Conservation of *S. enterica* SPI-2 T3SS and *S. glossinidius* SSR-3.

BlastP conservation scores are presented for *S. enterica* SPI-2 T3SS proteins compared to *S. glossinidius* SSR-3 orthologs.

Figure A1.6.

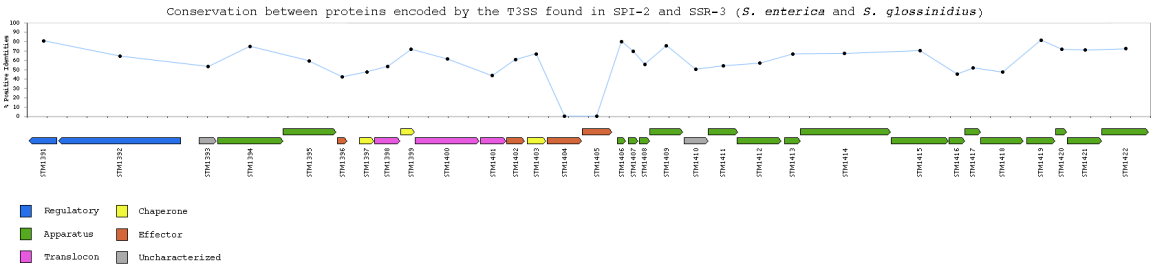

Conservation between proteins encoded by the T3SS found in SPI-2 and SSR-3 (S. enterica and S. glossinidius)
% Positive Identities
100
90
80
70
60
50
40
30
20
10
0
STM1391
STM1392
STM1393
STM1394
STM1395
STM1396
STM1397
STM1398
STM1399
STM1400
STM1401
STM1402
STM1403
STM1404
STM1405
STM1406
STM1407
STM1408
STM1409
STM1410
STM1411
STM1412
STM1413
STM1414
STM1415
STM1416
STM1417
STM1418
STM1419
STM1420
STM1421
STM1422
Regulatory
Apparatus
Translocon
Chaperone
Effector
Uncharacterized

Figure A1.7. Complementation of *S. enterica ssrB* with *S. glossinidius SG1279.*

Sodalis glossinidius SsrB (SG1279) can complement for SsrB. Cloning of *SG1280* (ortholog of *ssrA*) was not successful after many attempts and expression is likely toxic to *E. coli*, requiring cloning under an inducible, silenced promoter.

Figure A1.7.

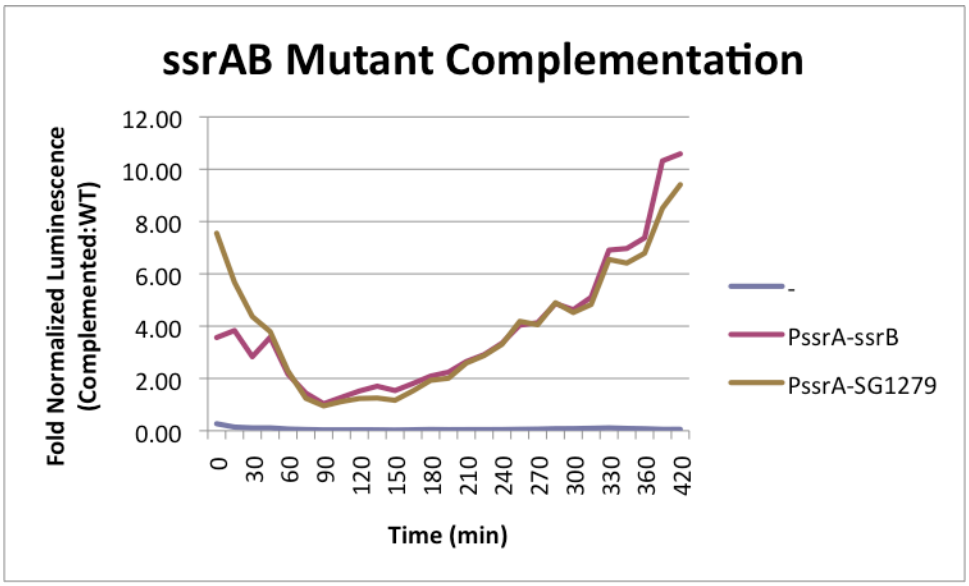

ssrAB Mutant Complementation
Fold Normalized Luminescence (Complemented:WT)
12.00
10.00
8.00
6.00
4.00
2.00
0.00
0
30
60
90
120
150
180
210
240
270
300
330
360
420
Time (min)
-
PssrA-ssrB
PssrA-SG1279

Figure A1.8. Alignment of *S. bongori* SPI-2 T6SS with those of *Pseudomonas spp.*

The *S. bongori* T6SS at SPI-2 strongly resembles a T6SS encoded within *Pseudomonas*.

Figure A1.8.

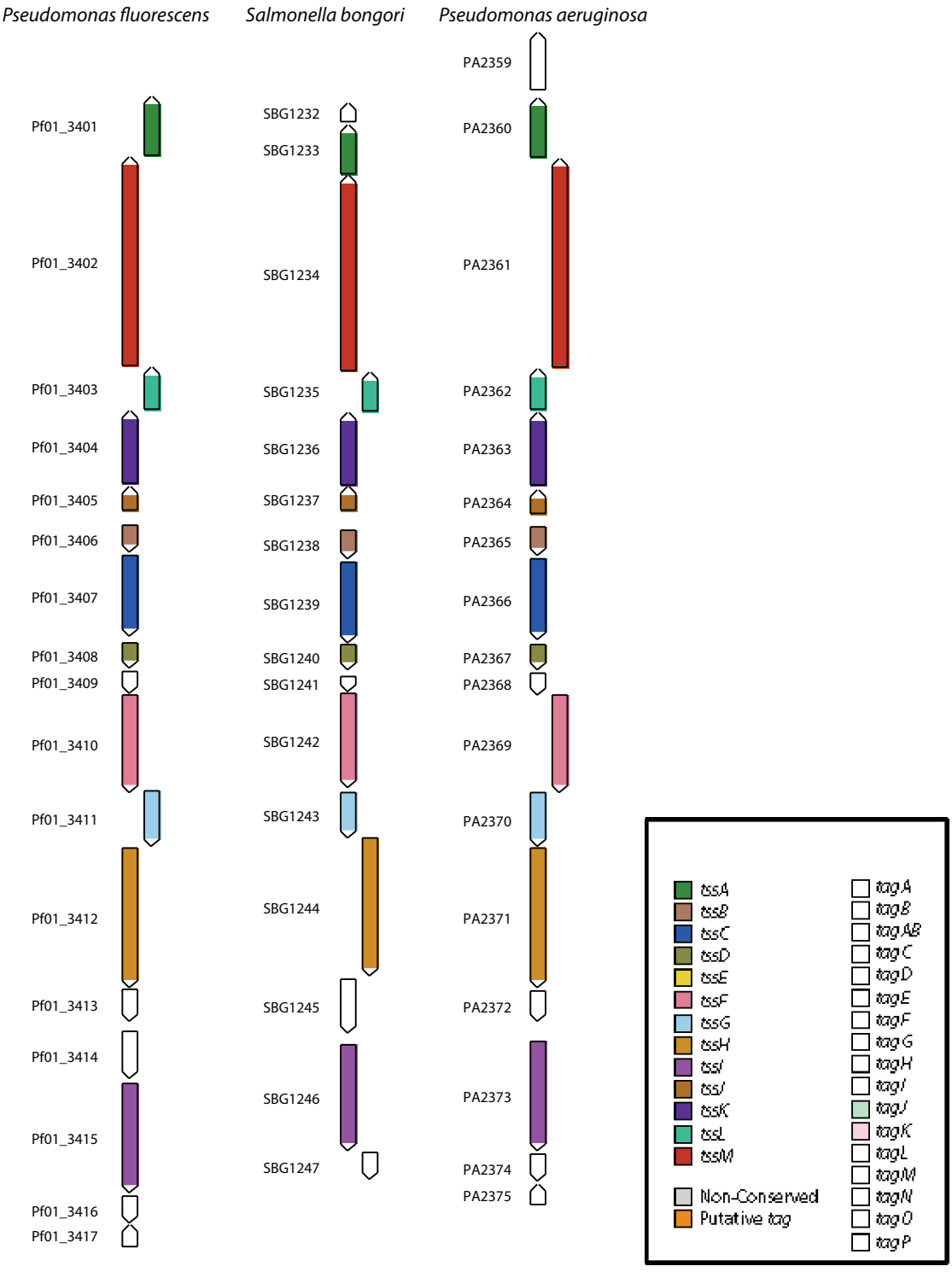

Pseudomonas fluorescens
Salmonella bongori
Pseudomonas aeruginosa
Pf01_3401
Pf01_3402
Pf01_3403
Pf01_3404
Pf01_3405
Pf01_3406
Pf01_3407
Pf01_3408
Pf01_3409
Pf01_3410
Pf01_3411
Pf01_3412
Pf01_3413
Pf01_3414
Pf01_3415
Pf01_3416
Pf01_3417
SBG1232
SBG1233
SBG1234
SBG1235
SBG1236
SBG1237
SBG1238
SBG1239
SBG1240
SBG1241
SBG1242
SBG1243
SBG1244
SBG1245
SBG1246
SBG1247
PA2359
PA2360
PA2361
PA2362
PA2363
PA2364
PA2365
PA2366
PA2367
PA2368
PA2369
PA2370
PA2371
PA2372
PA2373
PA2374
PA2375
tssA
tssB
tssC
tssD
tssE
tssF
tssG
tssH
tssI
tssJ
tssK
tssL
tssM
Non-Conserved
Putative tag
tagA
tagB
tagAB
tagC
tagD
tagE
tagF
tagG
tagH
tagI
tagJ
tagK
tagL
tagM
tagN
tagO
tagP

Figure A1.9. Ancestral T6SS associated genes in *S.* Typhimurium and *S. bongori.*

It was found that one of the T6SS-associated genes encoding an *hcp* ortholog is encoded within both *S.* Typhimurium and *S. bongori* even though these organisms do not share a common T6SS.

Figure A1.9.

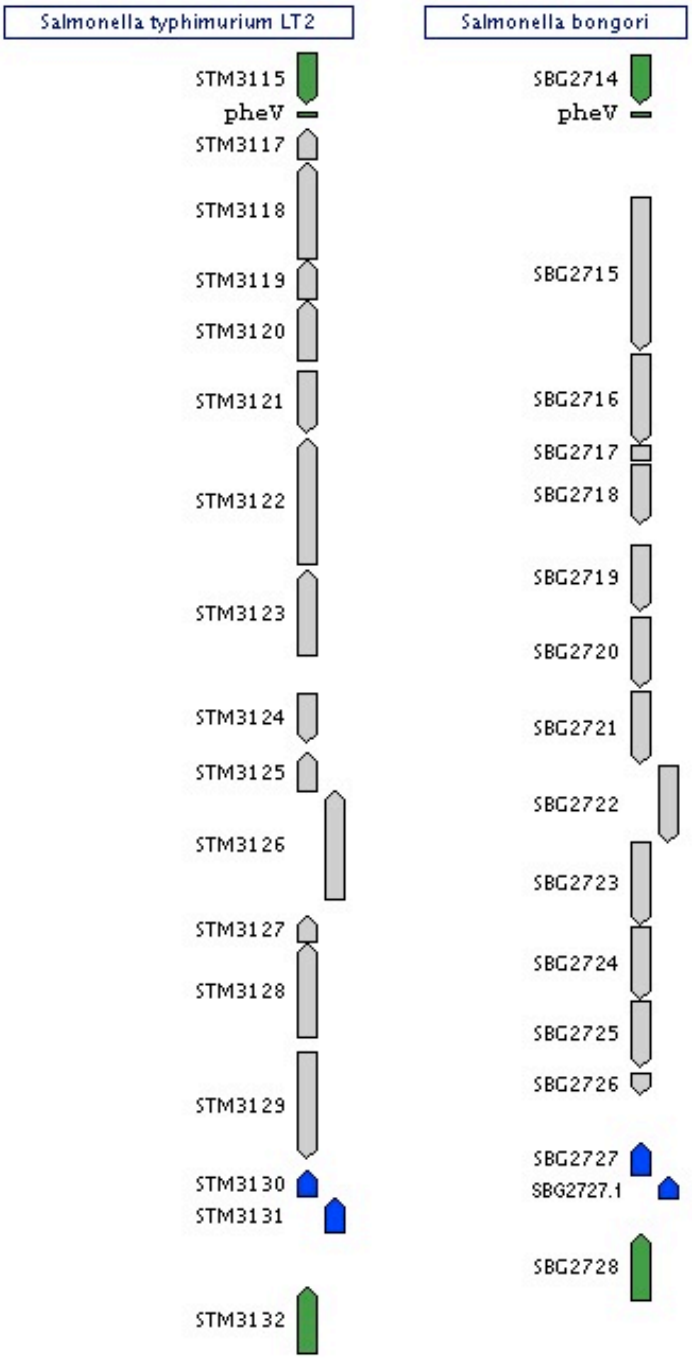
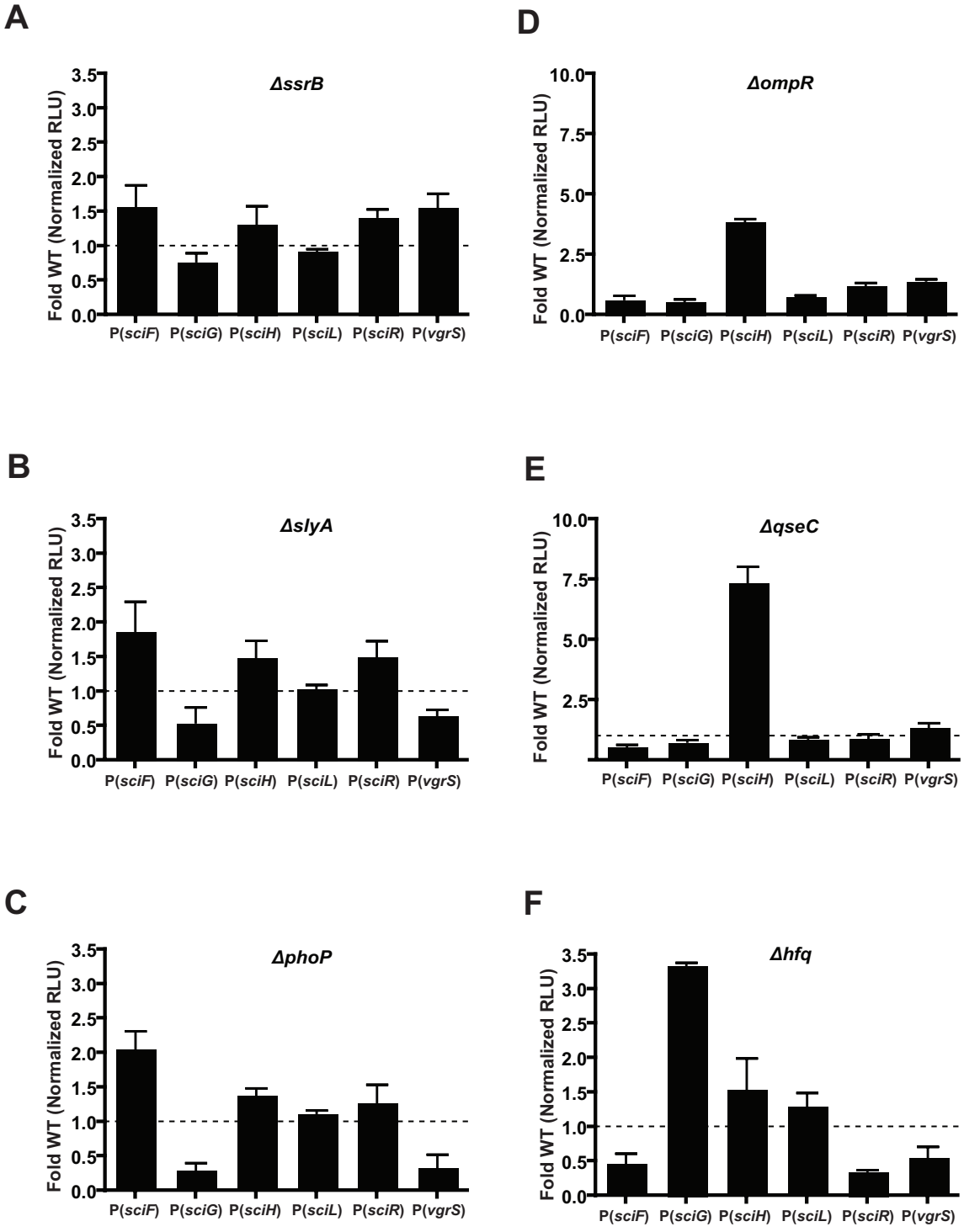

Salmonella typhimurium LT2
STM3115
pheV
STM3117
STM3118
STM3119
STM3120
STM3121
STM3122
STM3123
STM3124
STM3125
STM3126
STM3127
STM3128
STM3129
STM3130
STM3131
STM3132
Salmonella bongori
SBG2714
pheV
SBG2715
SBG2716
SBG2717
SBG2718
SBG2719
SBG2720
SBG2721
SBG2722
SBG2723
SBG2724
SBG2725
SBG2726
SBG2727
SBG2727.1
SBG2728

Figure A1.10. Transcriptional activity of T6SS promoters in regulatory mutants.

Transcriptional activity of promoter fusions with a *lacZ* reporter. From experiments done in triplicate. *S. enterica* SL1344 genotype as indicated.

Figure A1.10.


A
ΔssrB
Fold WT (Normalized RLU)
P(sciF) P(sciG) P(sciH) P(sciL) P(sciR) P(vgrS)
B
ΔslyA
C
ΔphoP
D
ΔompR
E
ΔqseC
F
Δhfq