Analysis of Spa33 and its Role in the Cytoplasmic Sorting Platform of Shigella

By
Ryan R. Skaar
B.A, Simpson College, 2017

Submitted to the graduate degree program in Pharmaceutical Chemistry and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Science in Pharmaceutical Chemistry

Committee Members

______________
Chair: Dr. William D. Picking

______________
Dr. Russ Middaugh

______________
Dr. Cory J. Berkland

Date Defended: August 8th, 2019
The dissertation committee for Ryan R. Skaar certifies that this is the approved version of the following dissertation

**Analysis of Spa33 and its Role in the Cytoplasmic Sorting Platform of *Shigella***

Chair: William D. Picking

Date Approved: August 8th, 2019
Abstract

Enteric infections, particularly those leading to diarrhea, can profoundly disrupt intestinal function and have a major impact on global mortality and morbidity rates.² Globally, there are 1.7 billion cases of childhood diarrheal disease every year, which results in the death of approximately 525,000 children under the age of 5.² Four pathogens were identified to be the key contributors to childhood diarrheal cases. These pathogens include rotavirus, Cryptosporidium, Shigella, and ST-ETEC. While rotavirus is the most prominent contributor to diarrheal episodes, specifically in infants (0 – 11 months), Shigella’s influence grows and becomes the primary contributor to diarrheal episodes as the child reaches the toddler stage of development (24 – 59 months).³ Ingestion of Shigella bacterium results in the luminal infection referred to as shigellosis. Before symptoms appear, Shigella first crosses the colonic epithelium via M cells where it is engulfed by macrophages that then undergo pyroptosis. After release into the submucosa, Shigella invades intestinal epithelial cells using its type III secretion system (T3SS).

The T3SS or injectisome is essential for Shigella virulence. The injectisome consists of three major components: extracellular needle, basal body, and a cytoplasmic sorting platform. Effector secretion is triggered by host cell contact and controlled by the sorting platform (SP). Shigella’s SP is comprised of five essential virulence proteins (MxiG, MxiK, Spa33, MxiN, and Spa47), which are highly conserved across all Shigella species. Characterization of the cytoplasmic SP via cryo-electron tomography allowed us to generate a 3D model of the SP containing six pod-like structures (comprised most prominently of Spa33). Spa33 is then connected to MxiN spokes that link it to the central ATPase (Spa47) and to MxiK, which links it to the basal body. This model differs dramatically from the contiguous arrangement of C ring proteins seen within the distantly related Gram-negative flagellum. Deletion of spa33 results in complete loss of the SP.
Spa33 consists minimally of a dimer of C-terminal Spa33 domains (Spa33\textsuperscript{C}) and a full-length copy (Spa33\textsuperscript{FL}). Spa33\textsuperscript{FL} is unstable on its own, but is greatly stabilized by association with its alternatively translated Spa33\textsuperscript{C} dimer. Here I provide support using various biophysical techniques \textit{in vivo} and \textit{in vitro} that the association between Spa33\textsuperscript{FL} and Spa33\textsuperscript{C} is essential to T3SS function, suggesting that both Spa33\textsuperscript{FL} and Spa33\textsuperscript{C} are necessary for SP assembly in \textit{Shigella}. 
# Table of Contents

**Background and Significance** ........................................................................................................ 1

ESKAPE and Enterobacteriaceae ........................................................................................................... 1

Diarrheal Diseases ................................................................................................................................. 2

Shigellosis and *Shigella* ....................................................................................................................... 2

*Shigella* Virulence Pathway ................................................................................................................ 4

Stomach Survival .................................................................................................................................. 4

Luminal Epithelium Localization and Apical Mucosa Modification ..................................................... 5

Shigella Macrophage Engagement and Escape ...................................................................................... 5

*Shigella* Intracellular Actin Driven Motility ....................................................................................... 6

Type III Secretion System (T3SS) ......................................................................................................... 7

T3SS Structure ...................................................................................................................................... 10

Cytoplasmic Sorting Platform ............................................................................................................... 15

SctQ/Spa33 ............................................................................................................................................ 26

**Significance** ...................................................................................................................................... 29

**Methods** .......................................................................................................................................... 30

**Results** ........................................................................................................................................... 36

Observation of Spa33\(^{FL/C}\) Co-purification .................................................................................... 36

SP Protein-Protein Interactions Assessed by BACTH Assay .................................................................. 41

Spa33\(^{FL}\) with Spa33\(^{C}\) Permutations ................................................................................................. 44

Spa33\(^{C}\) with Spa33\(^{C}\) Permutations ................................................................................................. 44

Spa33\(^{WT}\) permutations and Spa33\(^{FL}\) with Spa33\(^{FL}\) Permutations .................................................... 46

Spa33\(^{WT}\) with MxiN Permutations .................................................................................................... 49

Spa33\(^{FL}\) with MxiN and Spa33\(^{C}\) with MxiN Permutations ................................................................. 49

Spa33/ MxiK Permutations ................................................................................................................... 53

BACTH Summary .................................................................................................................................. 53

Determining Structural Form of Spa33 ................................................................................................. 54

Circular Dichroism (CD) Spectroscopy .................................................................................................. 58

Functional Assessment *in vivo* of Spa33\(^{WT}\) Components ............................................................... 60

**Discussion** ....................................................................................................................................... 64

**References** ....................................................................................................................................... 68
Background and Significance

ESKAPE and Enterobacteriaceae
Antimicrobial resistance (AMR) is not a new phenomenon. Its presence has been felt since the first antibiotics were used to treat bacterial infections.\(^4,5\) The unfortunate circumstance of AMR is that novel antimicrobial medications developed in response to microbial resistance will ultimately become the catalyst for the next wave of newly resistant bacteria.\(^6\) Hence, we must be able to produce novel antimicrobial agents at a much greater pace and/or agents that have broad spectrum capabilities if we hope to keep up with AMR. This is a daunting task and one we do not seem to be winning. While several new clinically useful antibiotics have been approved since 2013, reversing the downward trend in FDA approval, the urgency to develop new broad spectrum drugs should not be deterred.\(^7,8\) The Infectious Diseases Society of America (IDSA) continues to view with concern the lean pipeline for novel therapeutics to treat drug-resistant infections, especially those caused by gram-negative pathogens.\(^9\) Extensive research has identified a group of AMR pathogens that have been reported to have a high propensity to ‘escape’ the effects of current antimicrobial agents.\(^10,11\) The appropriately named ESKAPE pathogens include: *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumonia*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp. AMR has also been reported in the gram negative pathogens *Escherichia coli*, *Salmonella typhimurium*, and *Shigella flexneri* who are all members of the Enterobacteriaceae family and fall under the ESKAPE pathogens umbrella.\(^12-14\) The Enterobacteriaceae are a family of gram-negative, rod-shaped, non-sporulating bacterial pathogens. Many of its members live in the gut of mammals where some can cause enteric diseases.\(^15\)
Diarrheal Diseases
Enteric infections, particularly those leading to diarrhea, can profoundly disrupt intestinal function and have a major impact on global mortality and morbidity rates.\textsuperscript{1} Globally, there are 1.7 billion cases of childhood diarrheal disease every year, which results in the death of approximately 525,000 children under the age of 5.\textsuperscript{2} The majority of these deaths occur in developing countries. While the mortality rates have been steadily decreasing, morbidity has remained steady. Morbidity as a result of repeated infections by enteric pathogens triggering diarrhea, especially in the first 2 years of child-hood, greatly affect nutrient absorption and has a lasting impact on the growth and development of the child later on in life.\textsuperscript{1} Consequently, the long-term developmental and cognitive deficits seen with early child-hood diarrheal illnesses may be far more costly (in economic, as well as in human health) than previously thought.\textsuperscript{16}

In 2013, a study funded by the Bill and Melinda Gates Foundation was conducted to assess the burden and etiology of diarrheal disease in infants and young children in developing countries.\textsuperscript{3} In this study, four pathogens were identified to be the key contributors to childhood diarrheal cases. These pathogens include rotavirus, Cryptosporidium, Shigella, and ST-ETEC. An interesting dynamic of this group is that while rotavirus is the most prominent contributor to diarrheal episodes, specifically in infants (0 – 11 months), Shigella’s influence grows and becomes the primary contributor to diarrheal episodes as the child reaches the toddler stage of development (24 – 59 months).\textsuperscript{3}

Shigellosis and Shigella
Shigellosis is an acute invasive enteric infection triggered by bacteria belonging to the genus \textit{Shigella}.\textsuperscript{17} The \textit{Shigella} genus is named after Kiyoshi Shiga, who isolated it in 1897. \textit{Shigella} pathogens, members of the Enterobacteriaceae family, are Gram-negative, rod-shaped bacteria
and are phylogenetically separated into four subgroups with various serotypes: *S. boydii* (19 serotypes), *S. flexneri* (6 serotypes), *S. dysenteriae* (15C serotypes), and *S. soneii* (1 serotype). Shigellosis is endemic in most developing countries and was estimated to be responsible for 212,438 deaths globally among all ages in 2016. S. *dysenteriae* and *S. flexneri* are the predominant and most virulent of the four species especially in resource limited areas, whereas *S. soneii* is more common in affluent regions. The number of diarrheal cases and deaths fortunately are decreasing due to massive efforts taken to improve general sanitary practices and health conditions in developing countries.

Ingestion of as few as 10 organisms, usually from a contaminated food, water, or contact with fecal matter, may result in shigellosis. Symptoms begin 1-4 days after ingestion, and range from watery diarrhea to severe inflammatory dysentery characterized by strong abdominal cramps, fever, and stools containing blood and mucus. Isolation of the organism and serotyping is the only definitive method for diagnosis of *Shigella* because is cannot be distinguished from other causes of dysentery symptoms manifestation. Prevention of *Shigella* induced dysentery can be achieved by simple hand-washing, the availability of safe drinking water, proper disposal of human waste, breastfeeding of infants, safe food handling, and control of flies.

Current methods for treating shigellosis and other diseases caused by enteric pathogens often attempt to target universal mechanisms that are vital for pathogen survival. Historically, ampicillin, trimethoprim-sulfamethoxazole, tetracycline, nalidixic acid, and chloramphenicol have been effective antimicrobial treatment options for targeting vital pathogen functions. However, with the rise of AMR, even fluoroquinolones such as ciprofloxacin and other broad ranging antibiotics are additionally proving ineffective towards certain serotypes of *Shigella*. Vaccine effectiveness has also been hindered due to *Shigella’s* ability to employ serotype
switching via O-antigenic variation on its cell surface allowing it to evade host immunity.\textsuperscript{24–26} Therefore, an alternative method that should be considered is to target mechanisms needed for virulence of \textit{Shigella}, but are non-essential for the pathogen’s survival. Further study and treatment expansion focused towards the \textit{Shigella} virulence pathway would ideally demonstrate chronic infection prevention of diarrhea manifestation until the immune system is able to clear the pathogen naturally.

\textit{Shigella Virulence Pathway}

\textbf{Stomach Survival}

Upon \textit{Shigella} ingestion, the bacterium is able to survive the highly acidic environment of the stomach, pass through the small intestine and localize to colonic lumen where it can establish an infection in pH neutral environments.\textsuperscript{23} It has been shown that \textit{S. flexneri} implements at least two acid-resistance stationary-phase pathways (AR1 and AR2) to facilitate its survival through the stomach.\textsuperscript{27} AR1 is an acid-induced, glucose-repressed, oxidative, decarboxylase independent pathway that protects \textit{Shigella} between pH 3-5. Additionally AR2, a glutamate decarboxylase dependent system (GDAR) can protect bacterium below pH 3.\textsuperscript{28,29} The GDAR system, responsible for \textit{Shigella} survival as it passes through the stomach, is characterized by the paralogous pyridoxal 5’-phosphate dependent decarboxylase enzymes GadA and GadB, as well as an inner-membrane antiporter GadC.\textsuperscript{30} Cytoplasmic protons are consumed in a GadA/B facilitated reaction that converts glutamate to gamma-amino butyric acid (GABA) and CO\textsubscript{2}. Generated GABA is then transported out of the cell by GadC in a glutamate exchange process. This process allows the pathogen to survive the stomach’s acidic environment.\textsuperscript{30,31}
**Luminal Epithelium Localization and Apical Mucosa Modification**

Unlike many other enteric pathogens, *Shigella* does not express a flagellum and therefore is unable to propel itself through the thick intestinal mucosa environment. Investigations revealed that *Shigella* accounts for its lack of motility machinery through manipulation of intestinal secretory pathways by promoting the trapping and accumulation of mucins (MUC5AC and Muc2) at the cell surface. A gel-like matrix formed by the recruited mucins along with inhibition of antimicrobial peptides produced by the epithelial surface subsequently promotes a favorable environment for *Shigella* localization to the intestinal lumen. Following *Shigella* intestinal lumen localization, the pathogen must cross the colonic epithelial barrier before pathogenesis is observed. *Shigella* facilitates its transfer across the epithelial barrier via microfold cells (M cells). Studies suggest that M cell engulfment of bacterial agents is triggered by recognition receptors TLR-4, PAF-R and α5β1 integrin. After engulfment by the M cells, and trancytosis across the epithelial layer, the bacterium is released into the intraepithelial pocket where it encounters macrophages.

**Shigella Macrophage Engulfment and Escape**

The role of macrophages and their ability to phagocytose and present antigens as a part of the immune response to foreign agents has been widely studied. However, the uptake of *Shigella* following M cell trancytosis and entry into the lamina propria is not completely macrophage driven. While a clear mechanism has not been identified, it is reported that *Shigella* is able to direct its own internalization into macrophages through the secretion of the effector protein IpaC. Inside the macrophages, the bacterium is then able to escape the phagosome and kill the macrophage through apoptotic and pyroptotic pathways. These macrophage cell death pathways are triggered by the secretion of IpaD and IpaB proteins. Once internalized, IpaB causes permeabilization and subsequent disruption of the endolysosome along with the formation
of ion channels to promote K⁺ flux, which alters the membrane potential. Pyroptosis mediated cell death of the macrophage will occur due to the disrupted membrane potential and consequential activation of caspase-1 and release of IL-1β.⁴⁴,⁴⁵ IpaD was alternatively reported to activate caspase-9 (involved in positive-feedback loops and activation of caspases -2, -3, -6, -7, -8, -10), which is part of the mitochondrial apoptotic pathway. ⁴³,⁴⁶ This apoptotic pathway also triggers loss of membrane potential and occurs in parallel with the IpaB pyroptosis pathway. Once the mitochondrial death and inflammatory responses are stimulated, *Shigella* is then able to escape into the submucosa.

**Shigella Intracellular Actin Driven Motility**

Upon entry into the submucosa, *Shigella* will continue to replicate and polymerize actin which directs intracellular movement towards the basolateral side of the intestinal epithelial cell (EC) wall. Actin polymerization needed for intracellular motility of *Shigella* is driven by a VirG dependent process.⁴⁷–⁴⁹ VirG is expressed on the membrane surface and contains 3 domains (N, α, and C). The external presentation of α-domain of VirG is required for F-actin assembly.⁵⁰ Several reports also conclude that the N-terminal two-thirds of the α-domain is responsible for recruitment of F-actin through a direct interaction with vinculin, whereas the α-domain C-terminus aids VirG asymmetric distribution in *Shigella*.⁵¹,⁵² *Shigella* is then able to induce actin-based motility by VirG, N-WASP, and Arp2/3 complex interactions, which mediates filament polymerization and crosslinking needed for propulsion through the sub-mucosa and aid adhesion to the basolateral EC lining.⁴⁷,⁵²,⁵³

In build up to *Shigella* contact with host epithelial cells, I have described a large portion of the *Shigella* virulence pathway through the intestinal tract. Throughout this process *Shigella* has already initiated immune/inflammatory responses by the ongoing secretion of early effector
proteins such as IpaA, B, C, and D, which also aid in prompting macrophage cell death to allow *Shigella* escape into the sub-mucosa. In the following sections, I will discuss the complex protein nanomachine *Shigella* employs to modulate the secretion of the previously mentioned “early” effector invasion plasmid antigen (Ipa) proteins, which are required for infection.\(^{54}\)

Secretion/Injection modulation of effector proteins by this secretion system (see Table 1 for list)\(^{54,55}\) is involved not only in the secretion of early effector proteins, but actually plays a key role in host EC invasion by *Shigella* and secretion of numerous “late” effector proteins that dampen the inflammation response in order to prolong infection.\(^{41,44,56}\) While many Enterobacteriaceae express secretion systems to aid in pathogenicity or other cellular functions.\(^{57}\) *Shigella* are well known for their implementation of the type III secretion system (T3SS), also referred to as the injectisome.

**Type III Secretion System (T3SS)**

T3SS components essential for *Shigella* virulence are encoded by a ~30kb *mxi-spa* operon found in the virulence plasmid of *Shigella*.\(^{58}\) *Shigella* virulence plasmid (VP) (230kb) also maintains the *ipa/ipg* operon, responsible for the encoding of the secreted effector proteins mentioned previously (see Figure 1).\(^{54}\) *Shigella* acquisition of the VP through horizontal gene transfer is generally consider an important part of *Shigella* evolution from *E. coli*.\(^{59}\) While there are undoubtedly many more factors that played a role in *Shigella* evolution from a non-pathogenic species, it is important to note that homologous T3SS injectisomes are prevalent in other pathogens including: *Yersinia* species, *Salmonella enterica*, *Pseudomonas aeruginosa*, *Burkholderia pseudomallei*, and *Chlamydia* species.\(^{60}\) However, *Shigella* is somewhat unique because the operons encoding its virulence components are located on a single VP (Figure 1).
### Table 1

<table>
<thead>
<tr>
<th>Effectors</th>
<th>Activity/ Function</th>
<th>Host cell target</th>
<th>Effect on epithelial cells</th>
<th>Effect on inflammation and immune cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>IpaB</td>
<td>Forms pore (translocon)</td>
<td>Cholesterol, CD44</td>
<td>Bacterial invasion and vacuolar rupture</td>
<td>Bacterial escape from phagosome; Pyroptosis of MUs; Inhibition of EC-mediated inflammation</td>
</tr>
<tr>
<td>IpaC</td>
<td>Forms pore (translocon) Actin polymerization Docking and effector induction</td>
<td>Actin, Src, Cdc42</td>
<td>Bacterial invasion and vacuolar rupture</td>
<td>Bacterial escape from phagosome</td>
</tr>
<tr>
<td>IpaD</td>
<td>T3SS needle “plug”</td>
<td>TLR2</td>
<td></td>
<td>MU and T cell apoptosis</td>
</tr>
<tr>
<td>IpaA</td>
<td>Actin rearrangement and depolymerization</td>
<td>Vinculin</td>
<td>Bacterial invasion</td>
<td>Reduction of cell–matrix adhesion</td>
</tr>
<tr>
<td>lpgB1</td>
<td>Actin rearrangement</td>
<td>Rac2 and Cdc42</td>
<td>Bacterial invasion</td>
<td></td>
</tr>
<tr>
<td>lpgB2</td>
<td>Actin rearrangement</td>
<td>RhoA</td>
<td>Bacterial invasion</td>
<td></td>
</tr>
<tr>
<td>lpgD</td>
<td>PI4,5P2 conversion to PI5P</td>
<td>PI4,5P2</td>
<td>Bacterial invasion and vacuolar rupture Apoptosis inhibition</td>
<td>Inhibition of T cells migration; DAMPs secretion inhibition in EC</td>
</tr>
<tr>
<td>VirA</td>
<td>Microtubules degradation Inhibition of NFkB signaling</td>
<td>Tubulin, Calpastatin, and Inactivation of Rab1</td>
<td>Bacterial intra-/ intercellular dissemination Apoptosis inhibition Golgi disruption</td>
<td>Inhibition of EC-mediated inflammation</td>
</tr>
<tr>
<td>OspF</td>
<td>Inhibition of MAPK signaling</td>
<td>ERKp38</td>
<td></td>
<td>Inhibition of EC-mediated inflammation</td>
</tr>
<tr>
<td>OspG</td>
<td>Inhibition of NFkB signaling Ubiquitin and Dephosorylates MAPKs</td>
<td></td>
<td>Apoptosis inhibition</td>
<td>Inhibition of EC-mediated inflammation</td>
</tr>
<tr>
<td>OspI</td>
<td>Inhibition of NFkB signaling Deamidates UBC13 E2</td>
<td></td>
<td>Apoptosis inhibition</td>
<td>Inhibition of EC-mediated inflammation</td>
</tr>
<tr>
<td>OspZ</td>
<td>Inhibition of NFkB signaling Blocks p65 nuclear translocation</td>
<td></td>
<td>Apoptosis inhibition</td>
<td>Inhibition of EC-mediated inflammation</td>
</tr>
</tbody>
</table>

**Table 1 (Above): Hierarchal secretion of effector proteins and their function** - Secretion of these effector proteins are dependent on T3SS activity. This table as referenced in the text was produced through combination of tables and data from Mattock and Blocker as well as from Belotserkovsky and Sansonetti et al. 54,55
Figure 1 (Above): Genomic organization of the Shigella virulence plasmid (VP)- As described in the text, VP genes are clustered in two operons, the *ipa/ipg* and the *mxi/spa* operons. Genes are colored in the legend according to their protein class. Further information about each protein can also be found in Table 1. **Secretion machinery** (Light Blue) refers to the components that build the T3SS, particularly on the *mxi/spa* operon. **Translocators** (Navy) are components of the translocon, a pore inserted into the host membrane that allows effector translocation into host cell. **Chaperones** (Gold) are protein components that stabilize individual effectors prior to secretion from the bacterium. **Regulators** (Red) modulate T3SS expression and function. This figure was acquired from Mattock and Blocker (2017) who generated it via modifications to the virulence plasmid map produced by Buchrieser et al.(2000).\(^{54,106,107}\)
Alternatively, virulence operons in other pathogens, such as *Salmonella*, are located on chromosomal regions referred to as pathogenicity islands, but both do have the capability for horizontal gene transfer, which provides support for T3SS multi-species conservation. The addressed conservation and virulence mechanism pathway of the T3SS injectisome makes it a compelling candidate for antimicrobial drug targeting because: 1) The T3SS is required for pathogen virulence not survival. Therefore, antimicrobial drugs or vaccines targeting the T3SS may provide less selective pressure and reduced development of AMR risk. 2) Implementation of T3SS antimicrobial therapeutics could have widespread species and serotype effectiveness because the virulence mechanism is so highly conserved. Homologous injectisome component structure and function comparisons will frequently be alluded to throughout the remainder of this work (see Figure 2 for a general comparison).

**T3SS Structure**

The injectisome, is a highly conserved, transmembrane protein nanomachine that consists of three substructures: extracellular needle and tip complex, basal body, and cytoplasmic sorting platform (SP) (Figure 3).\(^{61,62}\) The basal body and SP substructures homologically also share several functional and structural features with flagellar systems.\(^{63}\) The apparent difference can be observed in the expression of a flagellar filament that directs bacterial motility, whereas the T3SS has evolved to express an extracellular needle to mediate intestinal delivery of effectors (Figure 4).\(^{64}\)

Polymerization of MxiH creates a hollow (7nm outer diameter), helical needle filament that protrudes 50 nm extracellularly from the *Shigella* outer membrane (see Figure 5).\(^{65,66}\) Upon contact with the host EC, secretion is stimulated to signal the translocators IpaB-D to be recruited to the MxiH needle tip. It has been reported that *Shigella* mutants lacking IpaB-D and
Figure 2 (Above): Multi-Species Overview Comparison of T3SS Injectomes- As referenced in the text, images above highlight the many similarities found between the injectisomes of *Salmonella enterica*, *Yersinia* species, *Pseudomonas aeruginosa*, and *Shigella flexneri*. It should also be noted that many of the structures are not fully characterized. Some of these gaps will be addressed here. It should be noted that a Cryo/ET *Pseudomonas* model has not been achieved to date. The *Psuedomonas* image above is a conventional model used to represent all T3SS.\textsuperscript{108,109}
Figure 3: Cryo-ET Model of the Type III Secretion System found in *S. flexineri* - A central subtomogram section average (A) and a 3D surface rendering (B) of the intact injectisome show OM, CM, peptidoglycan (PG), basal body, and needle in detail. Importantly, there is a large cytoplasmic complex that is 32 nm in diameter and 24 nm in height. Three cross-sections (indicated in A) of the cytoplasmic complex show sixfold symmetric features. The six pods (colored in red) are linked to the central hub (yellow) by radially arranged (spoke-like) linker densities (yellow). As referenced in text figure and text acquired from Hu et al. (2015).<sup>61</sup>
Figure 4: Schematic Comparison of the T3SS Injectisom with Flagellum- As discussed in the text, protein complexes were designed using electron microscopy (EM) structures as a model. With the exception of the extracellular segment, there are many observable structural similarities between the two systems. The pathogenic injectisome model is generated using: EMD accession numbers 1875, 5720, 5721, and 2669, for the basal body; the model for the major export component (Abrusci et al., 2013) and EMD accession number 1647 for the hook. Homologous components are presented with the same color in the structures. The T3SS comprise the extracellular segment, the basal body, and the cytoplasmic components. Figure and text acquired from Portaliou et al. (2016).
Figure 5: Visualization and Model Schematic of the MxiH Needle Filament Architecture-
As described in the text, images A and B provide low and high resolution reconstructions of the MxiH filaments that are viewed from the top and side. Top view of the structures reveals the helical arrangement of the MxiH subunits. Images provided in A were obtained from Cordes et al. (2003). Image B was acquired from Demers et al. (2014) and provided a higher-resolution structure through the use of solid-state NMR and cryo-electron microscopy. The protein ribbon structure is provided in parallel with the surface electron densities for visualization of the MxiH subunit orientations and arrangement as it polymerizes to form the extracellular needle. 65,66
expressing only MxiH needles, have no secretion activity. More importantly, in the absence of Ipa B-D, *Shigella* invasion into the host cell does not occur. IpaD serves an important role in facilitating IpaB and IpaC host membrane translocon pore formation and control of their secretion as effectors. Further structural information and implication of IpaD in translocon pore formation can be found by referring to Johnson et al. (2006) and Blocker et al. (2008). Following host cell contact and translocon pore formation by IpaB and IpaC, active secretion of host cell effector proteins is triggered. Effector proteins will then be selected, secreted through the export apparatus (basal body), extracellular needle, and then into the host cell where together they can begin to modulate host cell function to trigger EC invasion by *Shigella*. Effector targets and function are listed in Table 1. Once *Shigella* has successfully invaded the host cell, it shifts its focus towards survival and propagation. This is achieved by secretion of VirA and IpaJ, which aid in preventing autophagy. VirG actin-based motility is then used for *Shigella* EC intercellular movement. The pathogen will then protrude into the neighboring EC in order to continually evade immune responses and maintain a favorable propagation niche. Invasion into neighboring EC also stimulates release of interleukin (IL)-8, which recruits polymorphonuclear leukocytes (PMN). The release of proinflammatory cytokines and *Shigella*’s movement through EC lining ultimately results in luminal lining destruction by PMN and symptomatic presentation of shigellosis (abdominal pain, diarrhea, and/or dysentery).

**Cytoplasmic Sorting Platform**

Contact of the extracellular needle with host ECs activates delivery of virulence effector proteins. Six to twenty of these secreted effectors modulate the function of host regulatory molecules. Secretion of effectors, while dependent on contact with the host cell, is largely controlled by the injectisome’s cytoplasmic sorting platform (SP). The SP is evolutionally well
conserved among the Enterobacteriaceae (e.g. *Yersinia* species, *Salmonella enterica*, *Pseudomonas aeruginosa*, and *Shigella flexneri*) and is similar to flagellar systems. There are five essential protein subcomponents required for the fully assembled SP. The five homologous protein SP components found in Enterobacteriaceae include: SctD, SctK, SctQ, SctL, and SctN. The Sct (secretion and cellular translocation) nomenclature has been applied to unify the homologous T3SS structural components.\(^ {77}\) Table 2 has been provided as an aid to limit confusion as species comparisons between SP protein subunits are made.

Structural studies of the inner membrane ring in *Shigella* have shown that the ring’s protein component is a 371 residue protein called MxiG (SctD homolog). The MxiG ring has 24 fold symmetry (Figure 6).\(^ {78}\) It is also important to note that experiments with *Salmonella* spp SPI-1 and *Yersinia* have revealed that assembly of the inner membrane ring occurs independently of the other four SP components as well as the basal body.\(^ {79-81}\) Therefore, the SctD inner membrane ring forms as the first sequential step in injectisome assembly and it serves as the “architectural” foundation for injectisome SP construction.

After the SctD inner membrane ring has assembled, the remaining SP components (SctK, SctQ, SctL, SctN) are able to sequentially assemble below. While the roles and structural details of the SP components are unclear, efforts have been made to individually assess how the absence of one SP component effects the assembly of the others. Two methods that have been used to assess sorting platform assembly are 2D and 3D 4Pi single-molecule switching nanoscopy (4Pi-SMSN) and cryo-electron tomography (Cryo-ET). 4Pi-SMSN is a newly developed whole cell imaging technique that can reach resolution below 20 nm in 3D.\(^ {82}\) This techniques is able to provide information towards the distribution of SP components within a bacterium. For example, in wild type *Salmonella typhimurium* SpaO (SctQ) mutants fluorescently tagged with mEos3.2, SpaO
<table>
<thead>
<tr>
<th>Hosts</th>
<th>Humans and Animals</th>
<th>Flagellar Apparatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common Name (Sct)</td>
<td>Function</td>
<td>EPEC/EHEC</td>
</tr>
<tr>
<td>SctD</td>
<td>IM ring</td>
<td>EscD</td>
</tr>
</tbody>
</table>

**Cytoplasmic Membrane**

<table>
<thead>
<tr>
<th>SctQ</th>
<th>Cytoplasmic ring</th>
<th>SeptQ</th>
<th>PscQ</th>
<th>YscQ</th>
<th>Spa33</th>
<th>SpaO</th>
<th>SsaQ</th>
<th>CdsQ</th>
<th>BsaV</th>
<th>HrcQ, HrcQe</th>
<th>HrrQ</th>
<th>HrrQe; HrcQe</th>
<th>FilM; FilN</th>
</tr>
</thead>
<tbody>
<tr>
<td>SctL</td>
<td>Stator</td>
<td>EscL</td>
<td>PscL</td>
<td>YscL</td>
<td>MxiN</td>
<td>OrgB</td>
<td>SsaK</td>
<td>CdsL</td>
<td>OrgB</td>
<td>HrpE</td>
<td>HrpF</td>
<td>HrrL</td>
<td>FilH</td>
</tr>
<tr>
<td>SctN</td>
<td>ATPase</td>
<td>EscN</td>
<td>PscN</td>
<td>YscN</td>
<td>Spa47</td>
<td>InvC</td>
<td>SsaN</td>
<td>CdsN</td>
<td>BsaS</td>
<td>HrpN</td>
<td>HrcN</td>
<td>HrrN</td>
<td>FilI</td>
</tr>
<tr>
<td>SctO</td>
<td>Stalk</td>
<td>EscO</td>
<td>PscO</td>
<td>YscO</td>
<td>Spa13</td>
<td>InvI</td>
<td>SsaO</td>
<td>CdsO</td>
<td>HrpD or BsaT</td>
<td>HrpO</td>
<td>HrpD</td>
<td>HrpB7</td>
<td>FilJ</td>
</tr>
<tr>
<td>SctP</td>
<td>Molecular ruler</td>
<td>EscP</td>
<td>PscP</td>
<td>YscP</td>
<td>Spa32</td>
<td>InvJ</td>
<td>SsaP</td>
<td>CdsP</td>
<td>BsaJ</td>
<td>HrpP</td>
<td>HpaP</td>
<td>HpaC</td>
<td>FilK</td>
</tr>
<tr>
<td>SctW</td>
<td>Gatekeeper</td>
<td>SepL</td>
<td>PopN</td>
<td>YopNP</td>
<td>YeeA</td>
<td>MxiC</td>
<td>InvE</td>
<td>SsaL</td>
<td>CopN</td>
<td>BsaP</td>
<td>HrpJ</td>
<td>HpaA</td>
<td>HpaA /</td>
</tr>
<tr>
<td>/</td>
<td>Regulatory component</td>
<td>SepD</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>SpiC</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>SctK</td>
<td>ATPase cofactor</td>
<td>/</td>
<td>PscK</td>
<td>YscK</td>
<td>MxiK</td>
<td>OrgA</td>
<td>/</td>
<td>/</td>
<td>OrgA</td>
<td>HrpD</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
</tbody>
</table>

**Chaperones**

| / | For early substrates | EscE; EscG | PscE; PscG | YscE; YscG | / | / | / | / | / | / | / | / |
| / | For middle substrates | CesAB; CesD; CesD2 | PorG; PcrH | CdsE; CdsG | / | / | / | / | / | / | / | / |
| / | For late substrates | CesT; CesF | SpcU; SpcS | SycE; YscE; YscH; YscN; YscB; YsaK | / | / | / | / | / | / | / | / |
| / | / | / | / | / | / | / | / | / | / | / | / | / |

**Table 2: Unified Cytoplasmic SP Components and Chaperone Nomenclature** - As referenced in text, Table 2 is provided to limit confusion when referencing homologous proteins across various Enterobacteriaceae species. This table was acquired from and developed by Portaliou et al. (2016), who applied common nomenclature rules first suggested by Hueck (1998).
Figure 6: Positioning of MxiG-N(6–112) in the C24 S. flexneri NC EM map contoured to 1σ (EMDB 6391). A- A position of MxiG-N(6–112) (red) in the density for the leg domains. However, the volume occupied by a single subunit supports the idea that 24 copies can assemble in the ring. B- View of the position of MxiG-N(6-112) as described in A symmetrized to give 24 evenly arranged copies within the likely density for the MxiG-N ring. C- Model for a 24-subunit MxiG ring. The transmembrane region is depicted as an arbitrary 20-amino acid helix. Whereas the model shows the likely volume occupied by a 24-member ring, the positioning of the MxiG subunits with respect to each other represents one of many possible conformations. Figure and text acquired from McDowell et al. (2011).
fluorescent clusters were observed near the cell membrane, indicating SpaO localization to the SP. However, in the absence of OrgA and OrgB (SctK/L), SpaO fluorescent signals were evenly dispersed throughout the entire bacterium providing evidence that SP assembly requires the presence of all 5 components. The accessory SctK protein, mentioned above, is the least studied of the five SP components and it has no homolog in flagellar systems, suggesting its role is unique to the T3SS. While SMSN is an extremely powerful technique, it still lacks the atomic resolution needed to determine the position and structural features of the SP components. Fortunately, recent developments in Cryo-ET has enabled 3D reconstruction of SP structural features by aligning 2D subtomogram images. The aligned images are averaged to resolve a 3D model. Cryo-ET imaging by Hu et al. (2015) displayed the presence of pod like electron densities, comprised of Spa33 (SctQ). These densities were connected to the inner membrane ring by either MxiK (SctK) or the alternatively translated C-terminal truncation of Spa33. Unpublished results using MxiK T4-lysosyme insertion mutants, analyzed by Cro-ET in Shigella mini-cells, places MxiK (SctK) as the adapter between the MxiG inner membrane ring and Spa33 (SctQ) (Figure 7) (Tachiyama et al. 2019). The binding affinity between SctK and SctQ SP has been shown to be affected by increased diffusion of Ca$^{2+}$ under secreting conditions in Yersinia. This phenomena may lead to the increased exchange rate of YscQ (SctQ) which has been suggested to aid in effector transportation.

Cryo-ET and crystal structure renderings of SP components show similar structural homology across species even though their residue homology are not similar. Table 3 shows sequence homology comparisons across several Enterobacteriaceae. Amino acid sequence alignment percent similarity scores were acquired using CLUSTALW multiple sequencing alignment and
Figure 7: Visualization of the MxiK N terminal-T4L Insertion Mutant- Cryo-ET of MxiK-T4L and Spa33 SP components, expressed in *S. flexineri* mini cells. A sub-tomogram 3D model reveals the localized position of MxiK in between the inner membrane ring and Spa33 below (Tachiyama et al. 2019).
Table 3: Multiple Sequence Alignment Distance matrix of SP components - SP protein sequences were aligned using the CLUSTALW multiple sequence alignment program. Distance matrices were calculated using UGENE software algorithms and presented as percent sequence similarity as explained in text.

<table>
<thead>
<tr>
<th>SctD</th>
<th>MxiG Shigella</th>
<th>PscD Pseudomonas</th>
<th>YscD Yersinia</th>
<th>PrgH Salmonella</th>
<th>FliG E.coli</th>
</tr>
</thead>
<tbody>
<tr>
<td>MxiG Shigella</td>
<td>100%</td>
<td>9%</td>
<td>9%</td>
<td>20%</td>
<td>7%</td>
</tr>
<tr>
<td>PscD Pseudomonas</td>
<td>8%</td>
<td>100%</td>
<td>43%</td>
<td>13%</td>
<td>7%</td>
</tr>
<tr>
<td>YscD Yersinia</td>
<td>8%</td>
<td>44%</td>
<td>100%</td>
<td>13%</td>
<td>9%</td>
</tr>
<tr>
<td>PrgH Salmonella</td>
<td>19%</td>
<td>14%</td>
<td>14%</td>
<td>100%</td>
<td>7%</td>
</tr>
<tr>
<td>FliG E.coli</td>
<td>8%</td>
<td>9%</td>
<td>12%</td>
<td>8%</td>
<td>100%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SctK</th>
<th>MxiK Shigella</th>
<th>PscK Pseudomonas</th>
<th>YscK Yersinia</th>
<th>OrgA Salmonella</th>
</tr>
</thead>
<tbody>
<tr>
<td>MxiK Shigella</td>
<td>100%</td>
<td>10%</td>
<td>9%</td>
<td>24%</td>
</tr>
<tr>
<td>PscK Pseudomonas</td>
<td>8%</td>
<td>100%</td>
<td>41%</td>
<td>13%</td>
</tr>
<tr>
<td>YscK Yersinia</td>
<td>7%</td>
<td>41%</td>
<td>100%</td>
<td>13%</td>
</tr>
<tr>
<td>OrgA Salmonella</td>
<td>21%</td>
<td>14%</td>
<td>14%</td>
<td>100%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SctQ</th>
<th>Spa33 Shigella</th>
<th>PscQ Pseudomonas</th>
<th>YscQ Yersinia</th>
<th>SpaO Salmonella</th>
<th>FliM E.coli</th>
<th>FliN E.coli</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spa33 Shigella</td>
<td>100%</td>
<td>10%</td>
<td>14%</td>
<td>21%</td>
<td>9%</td>
<td>8%</td>
</tr>
<tr>
<td>PscQ Pseudomonas</td>
<td>10%</td>
<td>100%</td>
<td>40%</td>
<td>18%</td>
<td>6%</td>
<td>9%</td>
</tr>
<tr>
<td>YscQ Yersinia</td>
<td>14%</td>
<td>40%</td>
<td>100%</td>
<td>20%</td>
<td>8%</td>
<td>10%</td>
</tr>
<tr>
<td>SpaO Salmonella</td>
<td>20%</td>
<td>19%</td>
<td>20%</td>
<td>100%</td>
<td>10%</td>
<td>9%</td>
</tr>
<tr>
<td>FliM E.coli</td>
<td>7%</td>
<td>5%</td>
<td>8%</td>
<td>9%</td>
<td>100%</td>
<td>5%</td>
</tr>
<tr>
<td>FliN E.coli</td>
<td>18%</td>
<td>20%</td>
<td>23%</td>
<td>20%</td>
<td>12%</td>
<td>100%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SctL</th>
<th>MxiN Shigella</th>
<th>PscL Pseudomonas</th>
<th>YscL Yersinia</th>
<th>OrgB Salmonella</th>
<th>FliH E.coli</th>
</tr>
</thead>
<tbody>
<tr>
<td>MxiN Shigella</td>
<td>100%</td>
<td>13%</td>
<td>10%</td>
<td>19%</td>
<td>8%</td>
</tr>
<tr>
<td>PscL Pseudomonas</td>
<td>14%</td>
<td>100%</td>
<td>54%</td>
<td>15%</td>
<td>19%</td>
</tr>
<tr>
<td>YscL Yersinia</td>
<td>10%</td>
<td>52%</td>
<td>100%</td>
<td>13%</td>
<td>17%</td>
</tr>
<tr>
<td>OrgB Salmonella</td>
<td>19%</td>
<td>15%</td>
<td>12%</td>
<td>100%</td>
<td>13%</td>
</tr>
<tr>
<td>FliH E.coli</td>
<td>8%</td>
<td>18%</td>
<td>16%</td>
<td>13%</td>
<td>100%</td>
</tr>
</tbody>
</table>
Homology between SP components (SctD, SctK, SctQ, SctL) were generally below 20% similarity with exceptions found between *Yersinia enterocolitica* and *Psuedomonas aeruginosa* SP components, whose alignment scores ranged between 40 and 50%. It should also be noted that while SctQ similarity scores were under 20%, the aligned residues contributing to this score were contained in the C-terminus (*Figure 8*). This is supported by an approximate 10% increase in similarity when comparing alignment scores of just the SctQ C-terminus residues (*Table 4*). The substantial difference in overall sequence alignment scores makes the identification of essential interaction domains difficult. The poor sequence alignment scores may also explain differences between observed SctQ arrangement in T3SS and flagellar systems. Flagellar contiguous arrangement of SctQ (FliM/N) in a “C-ring” is well documented, and has similar structural homology to the T3SS. Early *Yersina* studies suggested that YscQ (SctQ) also assembled in a continuous C-ring with symmetry to SctD inner membrane ring proteins. The C-ring is connected by SctL (YscL) spokes, which radiate centrally from SctN (YscN) to SctQ (YscQ). In agreement with the *Yersinia* C-ring model, assessments of *Shigella* sorting platform architecture also favored a Spa33 (SctQ) contiguous ring. However, a separate study performed by Hu et al., in 2015, provided SP visualization, which provided an alternative sorting platform architecture shown in *Figure 9*. The cytoplasmic SP model, proposed by Hu et al. (2015), was obtained via cryo-ET, and it suggests that in *Shigella*, SctQ (Spa33) is arranged into six pod-like structures, connected to six SctL (MxiN) spokes (*Figure 10*). This model is distinctly different from the earlier proposed contiguous T3SS C-ring arrangement in *Yersinia*. Visualization of the SctQ pod arrangement was also observed in the T3SS of *Salmonella*. As mentioned above, T3SS SctQ sequence homology
Figure 8: Sequence Alignment of SctQC Residues- CLUSTALW sequence alignment of SctQC C-terminal residues, rendered by ESPript 3.0, reveals majority of sequence similarity between SctQ SP proteins resides in its C-terminus.\textsuperscript{114}

Table 4: Multiple Sequence Alignment Distance matrix of SctQC components- SP protein sequences were aligned using the CLUSTALW multiple sequence alignment program. Distance matrices were calculated using UGENE software algorithms and presented as percent sequence similarity.\textsuperscript{88} Scores increased by approximately 10% compared to results of Table 3 supporting the findings of the majority of homology is contained in the C-terminus of the SctQ proteins.
Figure 9  A: Cryo-electron tomography image of the extracellular needle, basal body, and cytoplasmic sorting platform (SP).\textsuperscript{1}  B: A closer look at the sorting platform reveals the 5 essential components of the SP, including Spa33 (red), MxiN (yellow), Spa47 (orange), spa13 (green), and MxiK (grey).\textsuperscript{1}

Figure 10: Cryo-ET model of the six radiating spokes of MxiN (Yellow), connected to six Spa33 pods (Red).
resides in the C-terminus. The C-terminal homology is also conserved in comparison with the flagellar FliN protein (25-28%). As mentioned, T3SS pathogens are proposed to evolve from flagellated bacteria, and both systems utilize an ATPase protein that is supported by the radiating SctL spokes. A well conserved interface between SctL and SctQ connecting the ATPase “motor” protein to the rest of the system may reflect the homology found between SctQ C-terminal residues and FliN. Alternatively, the poorer sequence homology observed between SctQ N-terminus and the flagellar FliM (5%) may explain how T3SS pathogens evolved to use their SP to drive secretion and not flagellar motion. This selection for an alternative SctQ arrangement from a contiguous FliM to non-contiguous radiating pods (Figure 4) would be favorable for secretion system evolution because a non-contiguous arrangement of pods increases the space available for the SP to access and select effector proteins for secretion. SP spacing is further enhanced because SctQ interactions with the rest of the SP components are considered to be transient and involve weak protein-protein interactions. This also implies that Spa33’s role within the injectisome is dynamic as secretion of effector proteins occurs.86

It is common for the stability of complex protein assemblies to be affected by the absence of one or more of its components. Even though SctQ interactions may be dynamic, its presence and interaction with SctL is essential to Shigella T3SS function. Deletion of sctQ (spa33) gene prevents the localization of SctL and SctN to the SP, which results in complete loss of the injectisome’s ability to construct the needle complex (NC) and subsequently direct effector protein secretion. Upon further analysis, the loss of secretion can be attributed to the sorting platforms inability to assemble in the absence of SctQ.61,90 Deletion of the sctL (mxiN) gene, also prevented the localization of SctN and formation of the needle complex, however, the SctQ pod densities were still visible. These data would suggest that SctQ plays an important role in the
sequential assembly of SctD-SctK-SctQ-SctL-SctN and that the presence of all 5 components are essential for modulation of the T3SS activity.\textsuperscript{83}

**SctQ/Spa33**

SctQ homologs are approximately 300 aa in length, but as discussed above, they have low sequence identity (Table 3). However, even with poor sequence homology, a unique characteristic in the translation of the $sctQ$ gene has been identified in *Yersinia, Salmonella,* and *Shigella.* While it was originally believed that prokaryotic translation is only initiated by start codons (AUG), recent discoveries of non-AUG alternative translation initiation (ATI) sites reveal that alternative forms of these proteins could be generated for specific cellular functions.\textsuperscript{91}

The ATI site for T3SS of Enterobacteriaceae was first identified in the SsaQ/SctQ component of *Salmonella* spp. SPI-2 T3SS by Yu et al. (2011).\textsuperscript{92} SsaQ (SctQ) translation from the ATI site was shown to occur in parallel with the conventional AUG start codon to produce a shorter SsaQ C-terminal product (106 residues) along with the expected SsaQ protein (322 residues). Evidence for parallel ATI production of a truncated C-terminal and full length SctQ was also shown to occur in *Salmonella* spp SPI-1, *Yersinia* and *Shigella.*\textsuperscript{78,93} Various, nomenclatures have been used to denote the alternatively translated products. For consistency, I have chosen to denote the alternatively translated shorter C-terminal SctQ protein as SctQ$^C$ (homologs: SsaQ$^C$ SpaO$^C$, YscQ$^C$, Spa33$^C$) and the conventionally expressed SctQ version as SctQ$^{FL}$ (homologs: SsaQ$^{FL}$, SpaO$^{FL}$, YscQ$^{FL}$, Spa33$^{FL}$). The existence of an alternative translation initiation site in the $sctQ$ gene suggests its assembly in the SP is more complicated than originally thought. Recombinant expression and purification of the SctQ components resulted in a single peak, which suggests that the co-purification of SctQ$^C$ and SctQ$^{FL}$ components may form a complex. Crystal structures (Figure 11) of *Shigella* and *Yersinia* SctQ$^C$ revealed two overlapping SctQ$^C$ chains
Figure 11: Crystal Structure of SctQ\textsuperscript{C} Overlay with FliN- A. The crystal structure of Spa33\textsuperscript{C} (a.a 208-293) is shown as a ribbon diagram. Chain A and B dimers are colored teal and purple respectively. The Spa33\textsuperscript{C} dimer is superimposed with the \textit{T. maritima} FliN structure (pdb id 1YAB), shown in grey. B. Three-dimensional structure of SsaQS and FliN. Superposition of the FliN (gray) and SsaQS (blue/red) dimer structures. Figure and text was combined from Yu et al. (2011) and McDowell et al. (2015) to make comparisons and show the homological similarities of the SctQ\textsuperscript{C} found in the T3SS SP and the flagellum C-ring.
implying the existence of SctQC as a dimer.\textsuperscript{78,92} Crystal structure alignment of SctQC with its flagellum homolog FliN also illustrated remarkable structural homology. The behavior of SctQ in the SP is suggested to be dynamic in nature, however, very little is understood about its regulation. The role of SctQ in the SP is further complicated by the above-mentioned discovery of the ATI site and hypothesized interactions between the FL and truncated C-terminal components. Production of both Spa33\textsuperscript{FL} and Spa33\textsuperscript{C} (SctQ\textsuperscript{FL/C}) subunits was determined to be necessary for assembly of the sorting platform and T3SS function in \textit{Shigella}. Alternatively, \textit{Salmonella} secretion and invasion studies determined that SpaO\textsuperscript{C} (SctQ\textsuperscript{C}) is a non-essential component for T3SS function, but is required for the stability of SpaO\textsuperscript{FL} (SctQ\textsuperscript{FL}).\textsuperscript{94} \textit{In vitro} evidence also suggests that the two protein subunits aggregate to form a heterotrimer through a 1:2 stoichiometric complex of Spa33\textsuperscript{FL}-Spa33\textsuperscript{C\textsubscript{2}} (Spa33\textsuperscript{WT}).\textsuperscript{78,94} Structural observations in \textit{Shigella} are consistent with \textit{in vivo} and \textit{in vitro} results found in \textit{Salmonella} SpaO\textsuperscript{WT} trimeric complex formation.\textsuperscript{94} For SpaO, it was determined that the SpaO\textsuperscript{C} component is not essential for SP injectisome function. On the other hand, \textit{Yersinia}, YscQC (SctQC) was deemed essential for T3SS function. McDowell et al.(2015) suggests that Spa33\textsuperscript{C} is also essential for \textit{Shigella} T3SS, but extensive \textit{in vivo} evidence to support an essential heterotrimeric Spa33\textsuperscript{WT} complex for T3SS activity was not provided. Unfortunately, a literature review has also been unable to provide suggestions of amino acid composition of the domains necessary for Spa33\textsuperscript{WT} interactions with MxiN and MxiK.

Characterization of the Spa33\textsuperscript{FL/C} subunit interaction using \textit{in vivo} functional analysis and \textit{in vitro} biophysical methods will provide us with a better understanding of how the Spa33\textsuperscript{WT} complex affects sorting platform assembly and modulation of secretion. The information gained
by characterizing the functional domains that Spa33$^{WT}$ uses to interact with surrounding sorting platform architecture could then be used for the development of novel T3SS inhibitors.

**Significance**

Antimicrobial resistance (AMR) emergence will have major effects on world health and economic infrastructure if not addressed. The unfortunate circumstance of AMR is that novel antimicrobial medication developed in response to microbial resistance will ultimately become the catalyst for the next generation of newly resistant bacteria.$^6$ Therefore, we must be able to produce novel antimicrobial agents with broad spectrum capabilities if we hope to keep up with AMR. Several species from the Enterobacteriaceae have been noted to have a propensity to ‘escape’ the effects of current antimicrobial agents as well as evade the host immune system.$^{10-14}$ Targeting mechanisms are needed for virulence, but are nonessential for a pathogen’s survival. They could serve as a method for preventing the onset of infection until the immune system is able to clear the pathogen naturally. A defining evolutionarily conserved virulence feature for many of these pathogens is their expression and implementation of a nanomachine referred to as the type III secretion system (T3SS).$^{57}$ Notably, the enteric pathogen *Shigella* employs its injectisome to modulate secretion/injection of effector proteins to aid with contact and invasion of host EC.$^{41,44,54-56}$ *Shigella* movement across the EC lining ultimately results in luminal wall destruction and symptomatic presentation of bacillary dysentery or shigellosis. Symptoms include: abdominal pain, diarrhea, and/or dysentery.$^{44,52-54,56}$ The T3SS injectisome is therefore a compelling candidate for antimicrobial drug targeting because: 1) The T3SS is required for pathogen virulence, not survival. Therefore antimicrobial drugs or vaccines targeting the T3SS may provide less selective pressure and reduced AMR risk. 2) Implantation of injectisome
antimicrobial therapeutics could have widespread species or serotype effectiveness because the virulence mechanism is so highly conserved.

Previous research in our laboratory has focused mainly on the general modulation, location, and structural characterization of Shigella proteins within the T3SS cytoplasmic sorting platform and tip complex. Thus, in response to the limited biophysical analyses available towards characterizing the interactions between sorting platform components, particularly Spa33, we hope to provide more clarity in this area. Spa33 interactions with surrounding proteins is suggested to be dynamic and essential for platform assembly and function. An improved understanding of the Spa33 functional interactions that are used to assemble its trimeric complex within the SP would be useful to learn how the Shigella T3SS is modulated. Ultimately, this research may be translated to develop T3SS inhibitor drugs that target the Spa33 functional domains. Inhibition of the Spa33 functional domains would prevent sorting platform assembly and thereby inhibit the injectisome functions necessary for Shigella host infection.

**Methods**

**Expression Plasmid and Protein Expression:** The coding sequences of Spa33 of *S. flexneri* was cloned into the pT7HMT expression plasmid. Gene specific primers were generated using the freely available online tool CloningTroll (http://biotroll.com). The corresponding genes were amplified by PCR from their respective genomic DNA. Amplified PCR products were restricted and In-Fusion Enzyme Premix (Takara Bio, Inc.) was used to ligate the 5’ and 3’ ends of the plasmid. Cloned plasmids were engineered to possess a N-terminal His6-tag followed by a TEV protease cleavage site. The presence of the gene insert and reading frame in the recombinant plasmids were verified using DNA sequencing services provided by Genescript™. Recombinant plasmids were transformed into *Escherichia coli* Tuner™ cells and inoculated in Luria broth.
(LB) containing 1 mM kanamycin antibiotic. For large-scale expression transformed cells were taken from glycerol stocks and inoculated into 15ml of LB and grown overnight. The starter culture was diluted and fresh cells were grown at 37 °C until $A_{600}$ 0.6 was achieved (mid log phase). Cultures were cooled to 21 °C, and protein expression was induced with 2ml of 0.5 M isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated overnight at 16 °C under constant shaking conditions (200rpm).

Protein Purification: The expression cells were harvested by centrifugation at 3470 x g for 10 min at 4 °C. The cellular pellet was suspended in 1x IMAC binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris) pH 7.5. The lysate was sonicated on ice at 70 % amplitude, 15 s on : 30 s off for 12min and centrifuged at 21689 x g for 30min at 4 °C. The supernatant was applied to an immobilized metal affinity chromatography column (Ni-IMAC) with Ni-NTA resin and purified by a GE AKTA™ purification system. The GE HisTrap™ FF crude Ni-NTA resin was equilibrated with binding buffer and the column was washed with increasing concentrations of elution buffer (0.5 M imidazole, 0.5 M NaCl, and 20mM Tris) pH 7.5 applied in a stepwise gradient until the bound protein selectively eluted from the Ni-NTA column. His-tag removal was achieved by the addition of TEV protease, 1mM DTT, dialyzed overnight in Binding Buffer. The protein was then applied to a second Ni-IMAC column. The flowthrough was collected, concentrated, and buffer exchanged with 20mM Tris, 200mM NaCl, pH 7.5 and subsequently loaded onto an equilibrated Superdex 200 pg (GE) size exclusion gel filtration column (SEC). All fractions throughout purification were analyzed by SDS-PAGE.

Bacterial Adenylate Two-Hybrid (BACTH) assay: The BACTH assay is useful for the detection of an interaction occurring between two proteins in an in vivo environment. Two proteins of interest are fused with a T25 or T18 adenylate cyclase toxin (CyaA) subdomain and co-expressed
in BTH101 *E. coli* cells that lack endogenous adenylate cyclase activity (CyaA). When two proteins of interest interact, the T25 and T18 subdomains will be brought together and subsequently restore CyaA activity. Restored CyaA activity in BTH101 cells results in a media dependent phenotypic color change of the plated cells (Blue: LB-XGal, Red: MacConkey, and Blue: M63). The observed color change indicates a protein-protein interaction. **Preparation of Sorting Platform T25/T18 Fusion Proteins:** Subdomain fusion proteins were made by cloning the proteins of interest genomic sequence into the T25 or T18 compatible vectors. The subdomains will be fused to either the N-terminal or C-terminal end of our protein of interest, which is dependent on the vector used. The *pk*T25 expression vector will fuse the T25 subdomain to the proteins N-terminus; alternatively, the *pk*NT25 vector will fuse the T25 subdomain to the C terminal end of our protein. The same can be achieved for fusion of the T18 subunit using the *pUT18* or *pUT18C* vectors, which fuses the T18 unit on the N and C-terminus, respectively. When probing for the interaction between two unique proteins it is important to screen using all of the T25/T18 fusion permutation combinations possible. This is because the binding of one protein to another in BACTH analysis requires that the T25 and T18 domains be oriented such that they form an active adenylate cyclase. Some orientations may not allow this to occur, which could give rise to a false negative result. Thus, examining all possible permutations for the BACTH fusion we decrease the likelihood of a false negative result.

Accuracy of the coding Spa33WT/FL/C, MxiK, and MxiN sequences was verified by Genscript™ sequencing using T25 5’ primer (gcgcagttcggtgaccagcg) and T18 (cggataacaatttcacacag) primer. Two compatible hybrid plasmids covering all the combinations were co-transformed into *E. coli* BTH101 host strain lacking functional CyaA. Colonies expressing the fusion plasmids were grown in LB broth with 0.5mM IPTG overnight. The colonies were then washed with M63
minimal medial to remove LB. Colonies were spotted on indicator plates, LB agar with 40 µg/mL X-Gal and 0.5 mM IPTG, MacConkey agar with 1 % (w/v) maltose. All the growth media were supplemented with selection antibiotics (T18: ampicillin and T25: kanamycin). Plates were incubated at 30 °C for 2-4 days. More details are available are referenced in the Euromedex BACTH System Kit.\textsuperscript{96,97}

**Circular Dichroism:** CD measurements were performed with Jasco J-1500 spectrophotometer (Jasco, Easton, MD, USA) and the results were expressed as the mean molar ellipticity ([θ]). Molar ellipticity is calculated by the following formula:

\[
\text{Molar ellipticity} \ [\theta] = \frac{\theta \times 100 \times mw}{C \times \ell \times n \times 1000}.
\]

θ is the CD signal in mDeg, \(mw\) is the molecular weight of the protein, C is concentration in mg/mL, \(\ell\) is the path length in cm, n is number of amino acid residues in the protein. 100 is a factor to convert path length from cm into meters and 1000 is a factor to convert the raw CD signal from mDeg to Degrees. All spectroscopic measurements were carried out at 10 °C. CD-measurements were done to analyze the secondary structure of Spa33\textsuperscript{WT} and Spa33\textsuperscript{C}. Purified protein was buffer exchanged into PBS (pH 7.4) to working concentration of 0.5 to 1 mg/mL. CD-spectra were obtained over a range of 195-260 nm, and temperature was increased during the variable temperature analysis from 10 °C - 90 °C in 2° increments. CD values were converted to molar ellipticity and plotted against wavelength and temperature.\textsuperscript{98,99}

**Overnight Steady State Secretion Assay:** T3SS secretion activity was tested by detection of IpaB, IpaC, and IpaD secretory protein in overnight bacterial culture supernatants. Colonies of *S. flexneri* and complimented *S. flexneri* mutants were plated on tryptic soy agar (TSA) plates with ampicillin (Amp) and/or kanamycin (Kan) depending on the strain plated. These recombinant
plasmids were transduced into \textit{spa33 S. flexneri null} competent cells by electroporation. Several colonies were selected and inoculated into tryptic soy broth (TSB) with strain dependent antibiotics (Kan/Amp). The bacterial cultures were grown at 37 °C overnight with slow shaking (100rpm) to prevent cell shearing and lysis and an 11 mL alliquot of the culture was collected and centrifuged at 4000 rpm for 15min at 4 °C. Fractions (9 mL) of centrifuged supernatant were then collected with care taken not to disturb the cell pellet. Proteins were precipitated from the fraction by adding 1 mL of a 100% trichloracetic acid (TCA) and incubated on ice for 30 min. Fractions were then centrifuged at 10,000 rpm for 15 min at 4 °C to collect protein precipitants. The supernatant was discarded and the protein pellet was washed with 5% TCA, again centrifuged, and followed by 2 sequential ice cold acetone wash steps. Acetone was evaporated and the precipitated protein was resuspended in 400 µL of PBS and 200 µL of 3X SDS-PAGE Buffer. For Western blotting, a 10% polyacrylamide gel was used for separation and transfer of proteins onto a nitrocellulose membranes. Rabbit anti-IpaB, IpaC, and IpaD and donkey anti-rabbit IgG with infrared tag (Li-Cor, Lincoln, NE) were used to detect IpaB, C, and D. Blots were imaged using an Odyssey Li-Cor CLx system and near-infrared fluorescence was used to detect the fluorescently tagged donkey anti-rabbit IgG. Li-cor Image Studio Lite software provided blot densitometry quantification, which allowed us to compare relative amounts of secreted IpaB, C, and D.

\textbf{Contact-mediated hemolysis:} The ability of \textit{Shigella} to form and insert functional translocons into target red blood cell (RBC) host membranes was tested by contact-mediated hemolysis. \textit{S. flexneri} strains were inoculated from glycerol stocks and grown on TSA with Amp and Kan plates overnight. Colonies from each plate were selected and inoculated into 15ml fresh TSB and grown at 37 °C until the mid-log growth phase, between 0.5 and 1.0 \textit{A}_{600}, was reached. Ten mL
of culture was centrifuged at 30 °C for 10 min. The supernatant was discarded and bacterial pellets were resuspended in PBS to equal bacterial densities by:

\[
\text{resuspension volume} = \frac{(200 \, \mu L \times \text{OD})}{0.5}
\]

Defibrinated sheep red blood cells (RBC) from Colorado Serum Co. Denver, CO) were aseptically withdrawn (3.5 mL) and washed twice in 40 mL PBS (10 mM phosphate, pH 7.4/150mM NaCl, 2.5mM KCl, pH 7.4) and centrifuged at 2876 x g at 30 °C for 10 min. The supernatant was discarded and the pellet was resuspended in 2.5 mL PBS. 10 µL of RBC was diluted 10,000 fold in PBS and RBC counts were measured with a Millipore Handheld Cell Counter Scepter. A measured RBC count on the order of 4x10^9 RBC/mL is desirable.

Resuspended bacteria and washed RBCs were mixed 1:1 (v/v -100 µL total) on a 96 well plate. The plate was centrifuged at 2876 x g at 30 °C to promote contact, and then incubated at 37 °C for 15 min. After incubation, 100 µL of ice cold PBS was subsequently added to resuspend the cell mixture. The plate was then centrifuged at 3500 rpm for 15 min at 10 °C. Supernatants were obtained from the mixtures and transferred into new 96 well plates. Water and PBS were used for non-bacteria positive and negative controls respectively. Hemoglobin content was measured by absorbance of 545 nm in a plate reader. PBS blank background absorbance was subtracted and absorbance measurements were then normalized to Spa33WT compliment values.
Results

Spa33<sup>FL/C</sup> Co-purification

Existence of an alternative translation initiation (ATI) site in the sequence of Spa33 (Spa33<sup>WT</sup>) was first reported by McDowell et al. (2016). The spa33 ATI site begins with the ribosomal binding site (RBS) at nucleotide 562 and ends at the alternative start codon at nucleotide 577 (Figure 12). Translation of the spa33 gene was shown to generate a full length copy Spa33<sup>FL</sup> and a shorter truncated C-terminal copy of Spa33<sup>C</sup>. We were able replicate findings observed by McDowell et al. (2016), using pT7Hmt expression vectors designed to N-terminally tag Spa33 with six histidine residues. Spa33<sup>WT</sup> was then expressed in E. coli Tuner<sup>TM</sup> cells and purified by Ni-affinity and size exclusion chromatography (SEC). Both protein copies were shown to co-purify in a single SEC elution peak (Figure 13) when the protein composition was analyzed by SDS-PAGE (Figure 14). Analysis of purified Spa33<sup>WT</sup> yielded two distinct bands at ~11.5 kDa and ~38kDa, which correlates with the molecular weight of Spa33<sup>C</sup> and N-terminally His-tagged Spa33<sup>FL</sup>. Band identities were further confirmed by mass spectrometry (data not included). SDS-PAGE visualization of the co-eluting bands support the observations made by McDowell et al. (2016) of an ATI site existing in the Spa33 gene. The co-elution following Ni-affinity and SEC, suggests that Spa33<sup>FL</sup> and Spa33<sup>C</sup> copies have a propensity to interact with each other. To test this hypothesis required isolating Spa33<sup>FL</sup> and its Spa33<sup>C</sup> subunits.

Isolation of Spa33<sup>C</sup> was achieved by selective primer design for nucleotides encoding for residues 208-293 with the same purification protocol as Spa33<sup>WT</sup>. Elution fractions were analyzed by SDS-PAGE (Figure 15). Spa33<sup>FL</sup> expression vector design was achieved by spa33 inverse PCR with primers containing point mutations at nucleotides 564 and 576, both of which
A. Spa33WT Genomic Sequence
ATGCTAAGAATTTAACATTTTTGACGCTACGACAGAAAACATACAGATTTTATATGCAA
AGCAACTCTGCGAGGCTTTTTCGATTTGACATTCAACTTCTCTCGAGTTTTTCGATTTCTTATTCTGGAATATTGATTTGGATGATTTTTCCGAT
TCCAGAAGTCTGATAAGATTACTTTCGGGTACAGAATGGAATCTAGATAGCA
ATCTTTACTGCATATTATCAAAATGTTGCTTTTATAGAGATGATCTCATTCTTTGAC
AAGATGCGGTTACGTAATCAAGTTTCTCTTGAATTCGACTCTATGCAATCTCTTATA
GTTTTCGCTTTGATCTTAAATGAAAGGGGTGTGCTTATTTAAAAAGAATTAGCAGT
GCACTTAAATCATAGTCCGAAACTTTATATATTATATTATGTGTAATCAAGTATT
ATTGGGTTATATATGAGAATGTTAATAATATGAAGCCAATTAATCTGTGCAAA
AGTAAATGGTGAGTCAGAACACACAGAAGTTTCTTGGCATTATTCATATATGTA
ATATCAATGTAAGTTTTAGATCTTTTATAGAATTAATGACAATTCAATCTTGA
ACTAAAAATGATAGTTAGAAAACGAATTATTCATTTCCGATGACATAGTTAAA
CATGTAATATCAAAGTAAATGATGTTTGCTTTGATTGCGATGGAAACTTTGTCTAT
TGAGGATGTTATGTTATCTGAGATTATTTCTGGATGTAAGAGTAA

B. Alternative Translation Initiation (ATI) Site

Spa33WT: [TTATTGGGGAATTATTTGACGACAG]

Spa33FL: [TTATTGTTGACATTATTTGACGACAG]

Point Mutations

Figure 12: A. spa33 Genomic sequence- As described in text, the alternative translation initiation site, highlighted in grey, begins with the RBS (nucleotides 562-568: blue) and ends with the alternative start codon (nucleotides 574-776: red). This region permits the translation of two Spa33 products. A full length copy (Spa33FL) and shorter C-terminal copy (Spa33C). B. Alternative Translation Initiation (ATI) site (grey)- spa33 inverse PCR primers were designed to insert guanine to cytosine point mutations in the ATI site for nucleotides 564 and 576. Nucleotides 564 and 576 are located in the ribosomal binding site (RBS-blue) and the Spa33C alternative start codon (ASC-red) respectively. The mutations made in the RBS and ASC prevent activation of the Spa33C ATI site and thus only expression of Spa33FL will occur.
Figure 13: Size Exclusion Gel-Chromatogram of Spa33^WT-. Ni-affinity purification followed by SEC purification of expressed Spa33 yields a singular peak following from 120ml Superdex gel matrix column. This is suggestive that the Spa33^FL copy and the alternatively translated Spa33^C copy are able to form an interacting complex.
Figure 14 (Left): SDS-PAGE Purified Spa33WT - Post recombinant N-terminal His-tagged Spa33WT expression and successive Ni-Affinity, SEC purification, and SDS-PAGE analysis of the singular elution peak from the SEC chromatogram in Figure 13 yielded two distinct bands, which run equivalently to the molecular weights of N-terminal His-Spa33FL (~38kDa) and Spa33C (~11.5 kDa). Band identity was confirmed by mass spectrometry (data not included).

Figure 15 (Right): SDS-PAGE of Purified Spa33C - Post recombinant N-terminal His-tagged Spa33C expression and successive Ni-Affinity, SEC, TEV protease His-tag cleavage, 2° Ni-Affinity purification, and SDS-PAGE analysis of 2° flow through fraction yielded a single band corresponding to Spa33C+ (~11.5 kDa). Based on multi-angle dynamic light scattering and size exclusion chromatography, Spa33C may exist as a dimer in solution (hydrodynamic diameter = 8 nm) that is able to interact with a single copy of Spa33FL to give a heterotrimer of Spa33FL(Spa33C)2 (hydrodynamic diameter = 21 nm).
replace a guanine for a cytosine (Figure 12). The mutations made to the ribosomal binding site (564) and alternative start codon (576) prevent Spa33C ATI site recognition resulting in the expression of only Spa33FL. We were unable to purify Spa33FL after recombinant expression because of its insoluble nature in the absence of the Spa33C component. Spa33FL was poorly soluble in the absence of Spa33C, indicating the importance of both components being present for proper SP assembly. If the copies were not in a complex with each other, they would not have co-purified during the Ni-affinity step because Spa33FL is N-terminally His-tagged, whereas the ATI site for Spa33C is a C-terminal domain, consequently lacking a His tag and would theoretically lack affinity for the Ni-column.

McDowell et al. provided further evidence using mass spectrometry that Spa33FL and Spa33C interact minimally in 1:2 fashion (Spa33FL-Spa33C2 equivalent to Spa33WT) with the ability to oligomerize further. McDowell proposed that Spa33 has a propensity to oligomerize, shown by SEC-MALS analysis, which was interpreted as the proteins assembling into a C-ring formation similar to that in flagellar systems. It should be noted that the above mentioned experiments suggesting oligomerization of Spa33WT and the Ishihara et al. (2005) report of Spa33WT interaction with other SP components using GST-pull down assays are all in vitro techniques. While the insight provided by these in vitro studies is useful, additional studies will need to be done to fully address the assembly and function of the Spa33 FL and C components.

Also included in the Ishihara et al. (2005) report, was their attempt to investigate the capacity of Spa33 to interact with SP associated proteins using a yeast two-hybrid system. However, screening with Spa33 and the other 5 essential SP constructs yielded no positive clones, suggesting these proteins were not interactive in the yeast two-hybrid system.
**SP Protein-Protein Interactions Assessed by the BACTH Assay**

To detect Spa33WT interactions with the other proposed components of the type III secretion system (T3SS) cytoplasmic sorting platform, we utilized a BACTH assay. The BATCH system takes advantage of the adenylate cyclase toxin (CyaA) from *Bordetella pertussis*. CyaA is activated by eukaryotic calmodulin (CaM). Upon activation by calmodulin, CyaA catalyzes the synthesis of cyclic adenosine monophosphate (cAMP).¹⁰⁰ CyaA catalytic domain (AC) contains two subdomains, T25 and T18 (Figure 16).¹⁰¹ The T18 subdomain carries the binding site for CaM, whereas T25 contains the catalytic domain.¹⁰⁰ Co-expression of the T25/T18 subdomains restores the synthesis of cAMP in the presence of calmodulin.¹⁰² Complementation of the T25 and T18 subdomains can also be achieved by fusing each subdomain to proteins of interest. If the two proteins of interest interact they will bring together the T25 and T18 subdomains and subsequently restore CyaA activity when co-expressed in Cya deficient cells. The restored activity of CyaA will initiate the synthesis of cAMP, which then forms a complex with the catabolite activator protein (CAP). The CAP complex then activates transcription of the lactose and maltose catabolic operons which can be easily detected by several different assays (Figure 17).⁹⁶ A qualitative measure of protein-protein interaction is achieved by plating BTH101 cya E. coli transformed with our proteins of interest on indicator media (LB-XGal: Blue, MacConkey: Red, and M63: Blue). A phenotypic color change of the plated bacteria indicates a positive protein-protein interaction. A quantitative measure for detection of protein-protein interactions can be done by assessing β-galactosidase enzymatic activity, which is associated with cAMP production.
Figure 16: Secondary structure of CaM-Bound to CyaA Catalytic Domain- T25 and T18 domains are colored in green and purple. Calmodulin (red) is bound to the T18/T15 subdomains which will inherently restore cAMP synthesis. Figure acquired from Guo et al. (2005).
Figure 17: BACTH Analysis Model- A BACTH assay was used to test for interactions between different Spa33 forms. The proteins of interests are fused with the inactive adenylate cyclase catalytic domains T25 and T18. A positive interaction between two hybrid proteins brings T18 and T25 together and restores CyaA activity in the E. coli BTH101 cells. Restored CyaA activity results in a BTH101 phenotypic color change.
Spa33FL with Spa33C Permutations
While previous studies have provided in vitro support for the Spa33FL/C domains interacting in a 1:2 fashion, confirmation of the SP protein-protein interactions in vivo has yet to be accomplished. In order to probe and confirm the protein-protein interactions between the Spa33 FL and C protein domains in an in vivo setting, BACTH analysis was done. In Figure 18, of the 8 permutation combinations, there are three combinations that detect a protein-protein interaction between Spa33FL and Spa33C on LB-XGal and MacConkey indicator media. These combinations include: Spa33FL_Nterm-T25 with Spa33C_Cterm-T18, Spa33FL_Nterm-T25 with Spa33C_Nterm-T18, and Spa33FL_Nterm-T18 with Spa33C_Cterm-T25. When comparing the positive combinations in Figure 18, we observe that in each of the three cases, only the Spa33C domain permutations are shown to interact with a Spa33FL permutation that has T25/T18 fused to its N-terminus. Assuming that the N and C-termini of Spa33FL are located on polar opposite ends of each other, we could hypothesize that the N-terminal region of Spa33FL contains the domain for the binding of Spa33C. In parallel, it appears that both Spa33C N and C-terminal-ends are capable of interacting with Spa33FL.

Spa33C with Spa33C Permutations
Even with confirmation of the direct intracellular protein-protein interaction of Spa33FL with Spa33C in Figure 18, this interaction required further investigation. I hypothesize that there are two possible scenarios for the interaction of Spa33C with the N-terminus of Spa33FL: 1) Two Spa33C domains dimerize prior to interacting with the FL protein or 2) Two Spa33C domains interact separately with the FL protein. Previous research concerning Shigella and Yersinia, has provided crystal structures (Figure 11) of the Spa33C and YscQC domains, with a crystal unit cell containing two Spa33C chains interacting to form a dimer. This provides support for scenario 1.
Figure 18: Spa33\textsuperscript{FL}/Spa33\textsuperscript{C} Permutation BACTH Screening Plates - BACTH analysis for the screening of Spa33\textsuperscript{FL} and Spa33\textsuperscript{C} protein-protein interactions. A positive protein-protein interaction is dictated by the phenotypic expression and observation of blue colonies on LB-XGal and red colonies on MacConkey indicator plates. A positive control is indicated by the (+) Zip protein, and the negative controls (-) are indicated accordingly.
Where the Spa33C domain forms a dimer prior to its interaction with the FL protein. If this dimerization occurs \textit{in vivo}, we should be able to easily probe this interaction using BACTH analysis. \textbf{Figure 19} BATCH screening indeed supports our hypothesis that Spa33C can form a dimer in the absence of the FL protein. The observed positive interactions in \textbf{Figure 19} include Spa33C\textsubscript{Nterm-T25} with Spa33C\textsubscript{Cterm-T18} and Spa33C\textsubscript{Nterm-T25} with Spa33C\textsubscript{Nterm-T18}. When we examine the crystal structures in \textbf{Figure 11}, the observed positive interaction between two permutation combinations is not unexpected because the N and C-terminal-ends of a single Spa33C chain are within close proximity to the other chains N and C-terminal-ends. The close proximity of both chains’ terminal ends also indicates why we were able to observe multiple permutations of Spa33C interacting with the N-terminus of Spa33FL in \textbf{Figure 19}.

\textbf{Spa33WT permutations and Spa33FL with Spa33FL Permutations}

At this point in our Spa33 BATCH screens, we have provided data to support a protein-protein interaction that involves Spa33C dimer with the N-terminus of Spa33FL to form a heterotrimer complex (Spa33\textsuperscript{WT}). In \textbf{Figure 20}, T25/T18 permutations of Spa33\textsuperscript{WT} were screened. Binding is observed for all possible permutation combinations. This result suggests that Spa33 is able to form a heterohexamer. However, based on our previous results and the radiating pods shown by cryo-ET, a heterohexamer formation seems unlikely. This particular BACTH screen is further complicated because both T18 and T25 Spa33\textsuperscript{WT} plasmids are expressing Spa33FL and Spa33C components, which are free to interact in various combinations with the other expressed components to assemble a heterotrimer as opposed to a heterohexamer. Thus, interpretation of these results should be considered with care given the multitude of scenarios that could lead to a positive interaction for this permutation screen. It also should be noted that if Spa33\textsuperscript{WT} oligomerization did occur it would probably involve interactions between Spa33FL components.
Figure 19: Spa33C/Spa33C Permutation BACTH Screening Plates - BACTH analysis for the screening of Spa33C and Spa33C protein-protein interaction. A positive protein-protein interaction is dictated by the phenotypic expression and observation of blue colonies on LB-XGal and red colonies on MacConkey indicator plates. A positive control is indicated by the (+) Zip protein, and the negative controls (-) are indicated accordingly.
Figure 20: Spa33\textsuperscript{WT} /Spa33\textsuperscript{WT} Permutation BACTH Screening Plates- BACTH analysis for the screening of Spa33\textsuperscript{WT} and Spa33\textsuperscript{WT} protein-protein interaction. A positive protein-protein interaction is dictated by the phenotypic expression and observation of red colonies on MacConkey and blue colonies on LB and M63 indicator plates. A positive control is indicated by the (+) Zip protein, and the negative controls (-) are indicated accordingly.

Figure 21: Spa33\textsuperscript{FL} /Spa33\textsuperscript{FL} Permutation BACTH Screening Plates- BACTH analysis for the screening of Spa33\textsuperscript{FL} and Spa33\textsuperscript{FL} protein-protein interaction. A positive protein-protein interaction is dictated by the phenotypic expression and observation of red colonies on MacConkey and blue colonies on LB and M63 indicator plates. A positive control is indicated by the (+) Zip protein, and the negative controls (-) are indicated accordingly.
Figure 21 depicts screening for the protein-protein interaction of Spa33FL and Spa33FL. As we can see in Figure 21, no positive interactions were observed, supporting the idea that Spa33WT heterohexamer formation is unlikely.

Spa33WT with MxiN Permutations
We next explored how Spa33WT and its individual components (FL and C domains) may interact with the surrounding sorting platform architecture using BACTH analysis. In Figure 22, the interaction between the Spa33WT complex and MxiN was investigated and reveals two combinations expressing a positive interaction. The positive interactions included MxiN_{Nterm-T25} with Spa33^{WT}_{Cterm-T18}, and MxiN_{Nterm-T25} with Spa33^{WT}_{Nterm-T18}. These results could indicate that the N or C-terminus of the Spa33WT complex is able to interact only with the N-terminus of MxiN. To investigate this interaction further, Spa33FL and Spa33C interactions with MxiN were individually assessed.

Spa33FL with MxiN and Spa33C with MxiN Permutations
BACTH analysis in Figure 23 shows a positive interaction with MxiN_{Nterm-T25} and Spa33^{FL}_{C-term-T18}. These data suggest that the C-terminus of Spa33FL interacts with the N-terminus of MxiN. Comparison between the results shown in Figures 23 and 24 allows us to suggest that the interaction observed between the C-terminus of the Spa33WT and MxiN (Figure 22) can be accounted for by the C-terminus of the complex’s FL domain interacting with MxiN. This is supported by the absence of an interaction as shown in Figure 23 (Row 2). BACTH analysis illustrated in Figure 24 revealed that Spa33C doesn’t interact with MxiN. Therefore, the observed interaction (Figure 23, Row 2) between the Spa33WT N-terminus with MxiN probably doesn’t involve the presence of the Spa33C dimer.
Figure 22: Spa33\textsuperscript{WT}/MxiN Permutation BACTH Screening Plates - BACTH analysis for the screening of MxiN and Spa33\textsuperscript{WT} protein-protein interaction. A positive protein-protein interaction is reflected by the phenotypic expression and observation of red colonies on MacConkey and blue colonies on M63 indicator plates. A positive control is indicated by the (+) Zip protein, and the negative controls (−) are indicated accordingly.
**Figure 23: Spa33<sup>FL</sup> /MxiN Permutation BACTH Screening Plates** - BACTH analysis for the screening of MxiN and Spa33<sup>FL</sup> protein-protein interactions. A positive protein-protein interaction is reflected in the phenotypic expression and observation of red colonies on MacConkey and blue colonies on M63 indicator plates. A positive control is indicated by the (+) Zip protein, and the negative controls (-) are indicated accordingly.
**Figure 24: Spa33\(^{C}\) /MxiN Permutation BACTH Screening Plates** - BACTH analysis for the screening of MxiN and Spa33\(^{C}\) protein-protein interactions. A positive protein-protein interaction is reflected in the phenotypic expression and observation of red colonies on MacConkey and blue colonies on M63 indicator plates. A positive control is indicated by the (+) Zip protein, and the negative controls (-) are indicated accordingly.
Spa33/ MxiK Permutations

Employing the BACTH system we have been able to provide in vivo evidence for the interaction and assembly of many of the SP proteins. In particular proteins interacting with the Spa33\textsuperscript{WT} heterotrimer’s FL and C-terminal components. A summary of the key interactions discussed above is provided in Figure 25 and Table 5. Also included in Figure 25 is the BACTH assessment of MxiK with the Spa33 components. The results indicate that MxiK is able to interact with Spa33\textsuperscript{WT}\textsubscript{N-terminus} (Figure 25: K). We were able to determine that this interaction is facilitated by the Spa33\textsuperscript{FL}\textsubscript{N-terminus} domain, and that Spa33\textsuperscript{C} doesn’t appear to be essential for Spa33\textsuperscript{FL} binding to the MxiK interface (Figure 25: L & M).

BACTH Summary

In response to the inability of the yeast two-hybrid system to screen for functional protein-protein interactions of Shigella SP components, I was able to provide evidence for in vivo protein-protein interactions using a bacterial adenylate cyclase two-hybrid (BACTH) assay. After screening multiple permutations, positive protein-protein interactions were observed between Spa33\textsuperscript{FL} (T25 or T18 fused to its N-terminus) and Spa33\textsuperscript{C} (T18 or T25 fused to its N-terminus). Comparisons reveal that the Spa33\textsuperscript{C} domain only interacts with Spa33\textsuperscript{FL} that has T25/T18 fused to its N-terminus. Therefore, I hypothesized that the N-terminal region of Spa33\textsuperscript{FL} contains the binding domain for Spa33\textsuperscript{C}. Spa33\textsuperscript{C} dimerization was also detected. By including the compiled BACTH results in Figure 25 and Table 5, I have provided a hypothetical assembly model for the localization of the Spa33\textsuperscript{FL}-Spa33\textsuperscript{C}\textsubscript{2} trimeric complex binding domains with the rest of the SP platform components (Figure 26).
Figure 25: Summary of Key BACTH Assay Results- BACTH analysis summary for the screening of Spa33 protein-protein interactions with the other SP components. A positive protein-protein interaction is dictated by the phenotypic expression and observation of red colonies on MacConkey. A positive control is indicated by the (+) Zip protein, and the negative controls (-) are indicated accordingly.
Table 5: Summary of Key BATCH Assay Results Including the Terminus of Interaction

Table addressing the summarized Figure 25 (above) protein-protein interaction chart (If no terminus is mentioned then it should be assumed the interaction was shown to be located at either the N or C-terminus. With this summary we can take this information and develop a comprehensive model, for how Spa33 and its subdomains may be oriented in the cytoplasmic sorting platform (Figure 26).

<table>
<thead>
<tr>
<th>Protein 1</th>
<th>Protein 2</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Spa33&lt;sup&gt;WT&lt;/sup&gt;</td>
<td>Spa33&lt;sup&gt;WT&lt;/sup&gt;</td>
</tr>
<tr>
<td>B</td>
<td>Spa33&lt;sup&gt;FL&lt;/sup&gt; N-terminus</td>
<td>Spa33&lt;sup&gt;C&lt;/sup&gt; (N or C)-terminus</td>
</tr>
<tr>
<td>C</td>
<td>Spa33&lt;sup&gt;FL&lt;/sup&gt;</td>
<td>Spa33&lt;sup&gt;FL&lt;/sup&gt;</td>
</tr>
<tr>
<td>D</td>
<td>Spa33&lt;sup&gt;C&lt;/sup&gt;</td>
<td>Spa33&lt;sup&gt;C&lt;/sup&gt;</td>
</tr>
<tr>
<td>F</td>
<td>Spa33&lt;sup&gt;WT&lt;/sup&gt;</td>
<td>MxiN N-terminus</td>
</tr>
<tr>
<td>G</td>
<td>Spa33&lt;sup&gt;FL&lt;/sup&gt; C-terminus</td>
<td>MxiN N-terminus</td>
</tr>
<tr>
<td>H</td>
<td>Spa33&lt;sup&gt;C&lt;/sup&gt;</td>
<td>MxiN</td>
</tr>
<tr>
<td>I</td>
<td>MxiN</td>
<td>MxiN</td>
</tr>
<tr>
<td>K</td>
<td>MxiK N-terminus</td>
<td>Spa33&lt;sup&gt;WT&lt;/sup&gt; N-terminus</td>
</tr>
<tr>
<td>L</td>
<td>MxiK N-terminus</td>
<td>Spa33&lt;sup&gt;FL&lt;/sup&gt; N-terminus</td>
</tr>
<tr>
<td>M</td>
<td>MxiK</td>
<td>Spa33&lt;sup&gt;C&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
Figure 26: Estimated Localization of Each of the SP Components and their Termini (Left): The graphic depicts the hypothesized localization and orientation of the Spa33<sup>WT</sup> components (Spa33<sup>FL</sup> and Spa33<sup>C</sup>) within the sorting platform derived from the BACTH results. (Right): Cryo-electron tomography model of the cytoplasmic sorting platform. Protein components are labeled as follows: MxiK (Orange), Spa33 (Yellow), and MxiN (Blue).
Determining Structural Form of Spa33

Characterization of the cytoplasmic sorting complex, by cryoelectron tomography, has allowed us to generate a model of the sorting platform containing six pod-like structures (comprised most prominently of Spa33) connected to six spokes (MxiN) to a central ATPase (Spa47). This model is distinctly different from the contiguous arrangement of Spa33 seen with the distantly related flagellar C-ring. Using cryo-ET data, the location of the Spa33 FL and C domains cannot be pinpointed within the structure. The aforementioned BACTH results in combination with cryo-ET 3D subtomogram models provide in situ evidence that allows us to hypothesize the localization of the Spa33 proteins and other SP binding domains in a functional SP. Cryo-ET analysis of Spa33 containing higher electron density T4 lysosome insertion mutants would also aid in orienting the protein within the SP. However, while major improvements have been made in cryo-ET methodology, reaching atomic resolution visualization in some cases, cryo-ET is limited by specimen thickness and total tolerable electron dose that can be deposited on one sample. Ultimately, the crystal structure must be determined to provide complete characterization of the proteins. We have already referred to the crystal structure of the Spa33C (McDowell et al. 2015) dimer, but to date, attempts to solve crystal structures for the Spa33WT trimeric complex and Spa33FL have been unsuccessful. Fortunately, there are other methods available to aid in general biophysical characterization of Spa33. These techniques include circular dichroism and dynamic light scattering.

Characterization of the Spa33 subunits by these methods will provide not only low-resolution structural information, but insight into the extent to which the subunits interact with themselves and other proteins in the sorting platform. Further experiments utilizing these methods, combined with permutations to the characterized native subunits may permit identification of key amino acids necessary for Spa33 component interactions with the rest of the platform.
To implement these biophysical techniques purified Spa33\textsuperscript{WT} and Spa33\textsuperscript{C} are required. Spa33\textsuperscript{FL} is insoluble in the absence of Spa33\textsuperscript{C}, which has limited our ability to characterize the Spa33\textsuperscript{FL} subunit. Nonetheless, the secondary structure of purified Spa33\textsuperscript{WT} and Spa33\textsuperscript{C} was characterized using circular dichroism spectroscopy (CD).

**Circular Dichroism (CD) Spectroscopy**

CD spectroscopy is a useful method that provides information about the secondary structure of a protein of interest.\textsuperscript{98} This method is thus able to provide the secondary structure of purified Spa33\textsuperscript{FL} and Spa33\textsuperscript{C} proteins. For example, if Spa33 proteins are predominately α-helical, their CD spectra will have wavelength minima near 208 and 222nm. Alternatively if the subunits secondary structure is more β-sheet or random coil like, their spectra will manifest a minima near 217-218nm (β-sheet) or 195-200nm (random coil), respectively.

Our measured Spa33\textsuperscript{C} molar ellipticity at 10° C yielded a single minimum between 215-218nm, suggesting a protein rich in β-sheet 2° structure (**Figure 27**). The spectrum of Spa33\textsuperscript{C} after thermal unfolding and then cooling suggests that Spa33\textsuperscript{C} is able to refold after unfolding at 90° C and subsequent cooling back to 10° C. The Spa33\textsuperscript{C} CD spectrum is supportive of the already solved crystal structure in which β-sheet represent the majority of the structure (**Figure 11**).\textsuperscript{78} Alternatively, when analyzing the CD spectrum of Spa33\textsuperscript{WT}, a double minimum is seen (208 and 222nm) suggesting Spa33\textsuperscript{WT} is more α-helical in nature. Interestingly, the Spa33\textsuperscript{WT} post melt spectrum (red) reveals a reduction in molar ellipticity and a change from α-helical to β-sheet secondary structure. This observation suggests that the Spa33\textsuperscript{WT} α-helical
Figure 27: Circular Dichroism (CD) Spectroscopy of purified Spa33\textsuperscript{WT} and Spa33\textsuperscript{C} - The Spa33\textsuperscript{C} (blue) spectrum, with a minimum near 215nm, suggests β-sheet 2\textdegree structure is present and this is consistent with its crystal structure. Spa33\textsuperscript{C} post melt spectrum (green) shows that the Spa33\textsuperscript{C} is very stable and is probably able to refold after melting and cooling back to 10 °C. The spectrum of Spa33\textsuperscript{WT} (black), contains minima at 208nm and 220nm, which indicates a strong α-helical 2\textdegree structure presence. Spa33\textsuperscript{WT} post melt spectrum (red), reveals that a β-sheet 2\textdegree structure, which could indicated that after reaching melting conditions Spa33\textsuperscript{FL} becomes denatured or aggregated while the Spa33\textsuperscript{C} dimer is able to reobtain its 2\textdegree structure qualities. We thus hypothesize that the uncharacterized Spa33\textsuperscript{FL} subunit has a mixed α/β structure.

Crystal Structure: The structure of Spa33\textsuperscript{C} was solved by McDowell et al.
properties, measured before melting, could be entirely accounted by the Spa33FL component, which is currently uncharacterized due to its insoluble nature during protein expression/purification. Alternatively, this could suggest the formation of intermolecular β-structure aggregation. CD α-helical signals tend to be stronger than those of β-sheet signals. Therefore, any β-sheet qualities contributed by Spa33C to the Spa33WT complex ellipticity would be overshadowed if Spa33FL has an N-terminus rich in α-helical structure. Consequently, after Spa33WT is exposed to unfolding temperatures, Spa33FL will disassociate from the Spa33C dimer and destabilize its secondary structure. However, because of the insoluble nature of Spa33FL and its inability to refold its secondary structure during subsequent cooling to 10° C, the spectrum of the disassociated Spa33C dimer is now detectable because Spa33FL is no longer able to contribute to the observed ellipticity. I thus hypothesize that the uncharacterized Spa33FL subunit, particularly its N-terminus, contains an elevated α-helical secondary structure compared to Spa33C. A CD-spectrum overlay for all of these conditions is provided in Figure 27.

Functional Assessment in vivo of Spa33WT Components
The behavior of Spa33WT (SctQ) in the SP has been suggested to be a dynamic process due to its various interactions in vivo as shown by the BACTH assay and the role Spa33C plays in maintaining Spa33FL solubility as shown in vitro by CD-spectroscopy. However, very little is understood about its role in regulating type III secretion. SctQ’s role in the SP is further complicated by the above-mentioned discovery of the ATI site and hypothesized interactions between the FL and truncated C-terminal components. Production of both Spa33FL- and Spa33C (SctQFL/C) subunits was proposed to be essential for SP assembly and T3SS function in Shigella. Alternatively, Salmonella secretion and invasion studies determined that SpaO\textsuperscript{S} (SctQ\textsuperscript{C}) is a non-essential component for T3SS function, but is required for the stability of SpaO\textsuperscript{FL} (SctQ\textsuperscript{FL}).

In
*vitro* evidence also suggests that these two protein subunits also associate to form a heterotrimer with a 1:2 stoichiometric complex of SctQ_{FL}-ScctQ_{C2} (SctQ_{WT}).^{78,94} Observations in *Shigella* are consistent with what is seen for *Salmonella* SpaO_{WT} trimeric complex formation.^{94} Although it was determined that SpaO_{C} is a non-essential SP component needed for *Salmonella* activity there are reports that Spa33_{C} is essential for *Shigella* T3SS activity.^{94}

To address differences brought forth by McDowell et al. (2015) and Tejero et al. (2019) regarding the involvement of Spa33_{FL/C}/SpaO_{FL/C} towards SP function in *Shigella* and *Salmonella*, I investigated *S. flexneri in vivo* T3SS activity. McDowell et al. investigated T3SS activity by examining the secretion activity of Ipa(A,B,C,D), IpgD, and VirA through Congo red induction of type III secretion. Their results, analyzed by SDS-PAGE and silver-staining did not detect any secretion activity by ΔSpa33_{FL} or ΔSpa33_{C} *S. flexneri* mutants supporting their hypothesis that both components are essential to T3SS activity. Contrary to this, Terjo et al. found that ΔSpao_{C} *Salmonella* SPI-1 displayed only a slight reduction in secretion of SipB, SipC and InvJ (homologs to *Shigella* secreted proteins) compared to the wild type strain. This suggests, as previously discussed, that SpaO_{C} is not essential to T3SS activity in *Salmonella* SPI-1.

Our investigation into SP *in vivo* activity contributions by Spa33_{FL} and Spa33_{C} required us to construct *spa33* *S. flexneri* null strains complemented with Spa33_{WT}, Spa33_{FL}, and Spa33_{C} expressing plasmids. After overnight secretion at 37 °C in TSB and western detection of IpaB, C, and D in culture supernatants revealed that in the absence of Spa33_{FL} no secretion activity was observed for IpaB and IpaD with minimal band intensity for IpaC (*Figure 28A*). However, residual secretion by the Spa33_{FL} complimented strain was observed in the absence of Spa33_{C} (*Figure 28A*). Odyssey CLX band intensity measurements (*Figure 28C*) normalized to the
Figure 28: Overnight Secretion of *S. flexneri* Complemented Spa33 Mutants-
A. Western Blot analysis of IpaB, IpaC, and IpaD *S. flexneri* overnight secretion in the presence or absence of Spa33FL or Spa33C. Lanes (Left-Right) *S. flex.* 2457T (+), *S. flex.* null bacteria complimented with Spa33WT, Spa33FL, Spa33C, and null (-) plasmids. B. SDS-PAGE of overnight secretion supernatant to detect Type II secretion of SepA. C. Quantitative measure of the western blot (A) band intensities. Results were normalized to wild type *S. flexneri* 2457T (+).
S. flexneri band intensities for secreted Ipa proteins would suggest that residual secretion activity of Spa33\(^{FL}\) compliment is detected in comparison to the Spa33\(^{\text{C}}\) and Spa33\(^{(-)}\) null bacteria. In other words, the absence of Spa33\(^{\text{C}}\) may not completely prevent T3SS activity in Shigella as McDowell et al. had proposed. While the absence of Spa33\(^{\text{C}}\) may not eliminate all T3SS activity, its reduction was profound compared to wild type S. flexneri 2457T and the Spa33\(^{\text{WT}}\) complimented strain.

While secretory detection of effector proteins by western blot is an accepted method for assessing T3SS activity in Shigella, this method is not the most sensitive for measuring quantitative differences. Thus, it is important to determine if the Spa33\(^{FL}\) residual secretion activity, observed in Figure 28, is a significant observation or an artifact of cell lysis. To test this, we measured the ability for S. flexneri expressing Spa33\(^{\text{WT}}\), Spa33\(^{FL}\), and Spa33\(^{\text{C}}\) to cause lysis of red blood cells (RBC). For successful invasion and the lysis of RBCs, Shigella must be able to form a translocon pore in the host cell membrane. Shigella will use this pore to allow effector proteins to pass into the host cell. The hemolysis method is performed by growing Shigella cultures to 0.5 OD\(^{600}\) and forcing contact by Shigella centrifugation with the RBC. Once the injectisome has made contact with the host RBC, insertion of the IpaB/C translocon pore and secretion of effector proteins will occur causing the RBC to lyse. Hemoglobin will consequently only be released from the RBC if the translocon pore is assembled. If the injectisome is able to form the translocon pore and cause RBC lysis, higher concentrations of hemoglobin will be released, correlating with elevated T3SS activity. Alternatively, if the translocon pore is unable to form because of the SP inability to assemble, then the hemoglobin concentration will be low. Poor hemoglobin release correlates with reduced T3SS activity. This analysis is a more sensitive
Measurements of RBC hemoglobin release by *S. flexneri* expressing Spa33\(^{WT}\), Spa33\(^{FL}\), and Spa33\(^{C}\) is described in **Figure 29**. Hemolytic activity was normalized to the Spa33\(^{WT}\) compliment, which expresses both the Spa33\(^{FL}\) and Spa33\(^{C}\) components. In the absence of Spa33\(^{C}\), the Spa33\(^{FL}\) complement bacteria hemolytic activity was reduced to about 8%. In the absence of Spa33\(^{FL}\) and the Spa33\(^{C}\) compliment no hemolytic activity was seen, which is the same as seen with the Spa33 null strain. These results suggest that the secretion activity of Spa33\(^{FL}\) compliment (~16% IpaB secretion) observed in **Figure 28** is real although at a greatly reduced level. While our results do not entirely agree with McDowell et al. (2015) who proposed that the presence of both Spa33\(^{FL/C}\) components is completely essential for *Shigella* injectisome activity, our hemolysis and secretion data certainly do suggest that Spa33\(^{FL}\) ability to complex with Spa33\(^{C}\) seems to have an elevated importance to *Shigella* injectisome activity relative to its SpaO homologs in *Salmonella* where the absence of SpaO\(^{C}\) little effect on injectisome secretion activity (Tejero et al. 2019).

**Discussion**

Enteric infections, particularly those leading to diarrhea, can profoundly disrupt intestinal function and have a major impact on global mortality and morbidity rates.\(^1\) Rotavirus is the most prominent contributor to diarrheal episodes, specifically in infants (0 – 11 months). *Shigella*’s influence is shown to grow and becomes the primary contributor to diarrheal episodes as the child reaches the toddler stage of development (24 – 59 months).\(^3\) *Shigella* causes enteric
**S. flexneri RBC Hemolysis Assay**

![Graph showing hemoglobin release percentages for different conditions.](image)

**Figure 29: S. flexneri RBC Hemolysis Assay** - Hemolytic assay measuring the release of hemoglobin from RBCs (Abs 545nm), post centrifugation and incubation (37°C) with *S. flexneri* spa33 null bacterium complimented with Spa33 WT, FL and C plasmids. The ability of the complimented plasmids to restore hemolytic activity was then assessed. Error bars were calculated by finding the standard deviation between triplicate samples. Results were normalized to the activity of *Shigella* complemented with Spa33WT.
infection through invasion of the intestinal epithelial cells using its type III secretion system (T3SS). The T3SS or injectisome is essential for *Shigella* virulence. The injectisome consists of three major components: extracellular needle, basal body, and a cytoplasmic sorting platform. Effector secretion is triggered by host cell contact and controlled by the sorting platform (SP).

Previous research in our laboratory has focused mainly on the general modulation, location, and structural characterization of *Shigella* proteins within the T3SS cytoplasmic sorting platform and tip complex. Thus, in response to the limited biophysical information available towards characterizing the interactions between SP components, we have tried to provide more clarity in this area. Experimental results using various *in vivo* and *in vitro* techniques highlights the essential role Spa33 (SctQ) plays in modulating *Shigella* injectisome activity and SP assembly. Evidence for *in vivo* protein-protein interactions using the BACTH assay compiled in Figure 25 and Table 5, supports my hypothesized assembly model for the binding domain locations of the Spa33^{FL}-Spa33^{C2} trimeric complex with respect to the other SP components (Figure 26). This model suggests that 1) MxiK (SctK) serves as the adapter protein between MxiG (SctD) and Spa33^{WT} (SctQ^{WT}), 2) MxiK interaction with Spa33^{WT} occurs between MxiK and Spa33^{FL} N-terminal domains, 3) Spa33^{C} dimer was not observed to interact with MxiK, but does appear to localize to the N-terminus of Spa33^{FL}, and 4) The C-terminus of Spa33^{FL} interacts with the N-terminus of MxiN. Further biophysical characterization of Spa33^{C} binding with Spa33^{FL} determined that Spa33^{FL} stability and type III secretion function is greatly affected by the presence or absence of the Spa33^{C} dimer (Figures 27-29).

Spa33’s involvement in the SP is complex due to its ability to interact with a multitude of cytosolic components. This complexity is further enhanced due to the presence of an ATI site, which enables the *spa33* gene to encode for FL and C-terminal copies of itself. This ATI site is
conserved in other T3SS, such as those from *Yersinia* and *Salmonella*, who also use it to express interacting SctQ^{FL} and SctQ^{C} proteins. The C-terminus of SctQ also shares some homology with the flagellar C-ring protein FliN. Because SctQ^{WT}'s role in injectisome activity involves interactions with the other SP components, development of inhibitor drugs targeting Spa33 could also be effective against SctQ proteins in other organisms. If so, Spa33 inhibitory drugs could have broad spectrum microbial effectiveness and be useful for treatment of various T3SS modulated diseases.

The role of SctQ appears to impact T3SS activities differently depending on the organism encoding the injectisome. For example, SctQ has been proposed to assemble in a contiguous C-ring arrangement in *Yersinia* and the flagellum of Enterobacteriaceae. Alternatively, the arrangement of SctQ in *Shigella* and *Salmonella Typhimurium* is observed by Cryo-ET to assemble as 6 individual non-contiguous pods. It was also found that the absence of Spa33^{C}, compared to its SpaO^{C} homolog, seems to be more detrimental to the *Shigella* injectisome activity relative to *Salmonella* activity ([Figures 28-29](#)). Even with these differences, it is still apparent the interaction of SctQ^{C} and Spa33^{FL} is important to T3SS activity. Further research towards their individual roles should enhance our ability to develop small molecule/vaccine therapeutics against T3SS mediated pathogenesis.
References


47. Bernardini, M. L., Mounier, J., d’Hauteville, H., Coquis-Rondon, M. & Sansonetti, P. J. Identification of icsA, a plasmid locus of Shigella flexneri that governs bacterial intra- and


97. Euromedex, life science products. Available at:


