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.[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.[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 sub-mucosa, *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 ($Spa33^{C}$) and a full-length copy ($Spa33^{FL}$). $Spa33^{FL}$ is unstable on its own, but is greatly stabilized by association with its alternatively translated $Spa33^{C}$ dimer. Here I provide support using various biophysical techniques *in vivo* and *in vitro* that the association between $Spa33^{FL}$ and $Spa33^{C}$ is essential to T3SS function, suggesting that both $Spa33^{FL}$ and $Spa33^{C}$ are necessary for SP assembly in *Shigella*.

Table of Contents

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: ***E****nterococcus faecium*, ***S****taphylococcus aureus*, ***K****lebsiella pneumonia*, ***A****cinetobacter baumannii*, ***P****seudomonas aeruginosa*, and ***E****nterobacter* 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-sporlating 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.[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.[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.[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.[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.[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).[3]

Shigellosis and *Shigella*

Shigellosis is an acute invasive enteric infection triggered by bacteria belonging to the genus *Shigella.*[17] The *Shigella* genus is named after Kiyoshi Shiga, who isolated it in 1897. *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.[18] *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.[19–21] 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.[22] 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.[23] 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.[17] 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*.[25,26] 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.[24–26] Therefore, an alternative method that should be considered is to target mechanisms needed for virulence of *Shigella*, but are non-essential for the pathogen's survival. Further study and treatment expansion focused towards the *Shigella* virulence pathway would ideally demonstrate chronic infection prevention of diarrhea manifestation until the immune system is able to clear the pathogen naturally.

Shigella Virulence Pathway

Stomach Survival

Upon *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.[23] It has been shown that *S. flexneri* implements at least two acid-resistance stationary-phase pathways (AR1 and AR2) to facilitate its survival through the stomach.[27] AR1 is an acid-induced, glucose-repressed, oxidative, decarboxylase independent pathway that protects *Shigella* between pH 3-5. Additionally AR2, a glutamate decarboxylase dependent system (GDAR) can protect bacterium below pH 3.[28,29] The GDAR system, responsible for *Shigella* survival as it passes through the stomach, is characterized by the paralogous pyriodoxal 5'-phosphate dependent decarboxylase enzymes GadA and GadB, as well as an inner-membrane antiporter GadC.[30] Cytoplasmic protons are consumed in a GadA/B facilitated reaction that converts glutamate to gamma-amino butyric acid (GABA) and CO_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 stomachs acidic environment.[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.[32,33] 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.[32,34–36] 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).[37] Studies suggest that M cell engulfment of bacterial agents is triggered by recognition receptors TLR-4, PAF-R and α5β1 integrin.[38] 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.[39] However, the uptake of *Shigella* following M cell trancytosis and entry into the laminal 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.[40] Inside the macrophages, the bacterium is then able to escape the phagosome and kill the macrophage through apoptotic and pyroptotic pathways.[41] These macrophage cell death pathways are triggered by the secretion of IpaD and IpaB proteins.[42,43] 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

Effector Early/Middle/Late	Activity/ Function	Host cell target	Effect on epithelial cells	Effect on inflammation and immune cells
IpaB (Early)	Forms pore (translocon)	Cholesterol, CD44	Bacterial invasion and vacuolar rupture	Bacterial escape from phagosome; Pyroptosis of MUs; Inhibition of EC-mediated inflammation
IpaC (Early)	Forms pore (translocon) Actin polymerization Docking and effector induction	Actin, Src, Cdc42	Bacterial invasion and vacuolar rupture	Bacterial escape from phagosome
IpaD (Early)	T3SS needle "plug"	TLR2		MU and T cell apoptosis
IpaA (Early)	Actin rearrangement and depolymerization	Vinculin	Bacterial invasion Reduction of cell–matrix adhesion	
IpgB1 (Early)	Actin rearrangement	Rac2 and Cdc42	Bacterial invasion	
IpgB2 (Early)	Actin rearrangement	RhoA	Bacterial invasion	
IpgD (Early)	PI4,5P_2 conversion to PI5P	PI4,5P_2	Bacterial invasion and vacuolar rupture Apoptosis inhibition	Inhibition of T cells migration; DAMPs secretion inhibition in EC
VirA (Middle)	Microtubules degradation Inhibition of NFκB signaling	Tubulin, Calpastatin, and Inactivation of Rab1	Bacterial intra-/ intercellular dissemination Apoptosis inhibition Golgi disruption	Inhibition of EC-mediated inflammation
OspF (Middle)	Inhibition of MAPK signaling	ERKp38		Inhibition of EC-mediated inflammation
OspG (Late)	Inhibition of NFκB signaling	Ubiquitin and Dephosorylates MAPKs	Apoptosis inhibition	Inhibition of EC-mediated inflammation
OspI	Inhibition of NFκB signaling	Deamidates UBC13 E2	Apoptosis inhibition	Inhibition of EC-mediated inflammation
OspZ	Inhibition of NFκB signaling	Blocks p65 nuclear translocation	Apoptosis inhibition	Inhibition of EC-mediated inflammation

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]

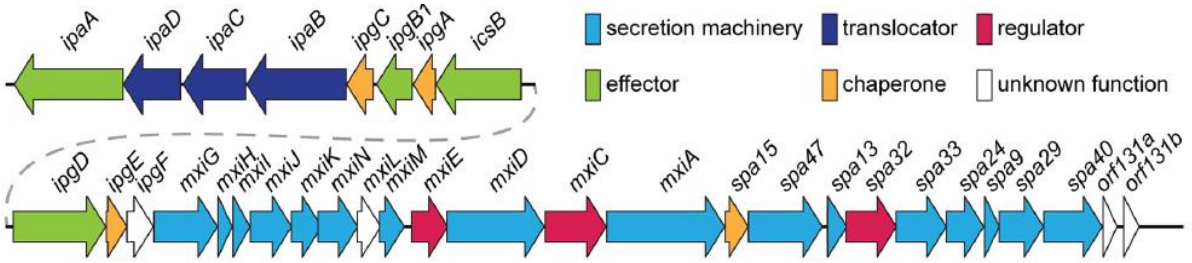


Figure 1 (Above): Genomic organization of the *Shigella* virulence plasmid (VP)- As described in the text, VP genes are clustered in two operons, the *ipa/ipg* and the *mxi/spa* operons. Genes are colored in the legend according to their protein class. Further information about each protein can also 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 comlex, 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

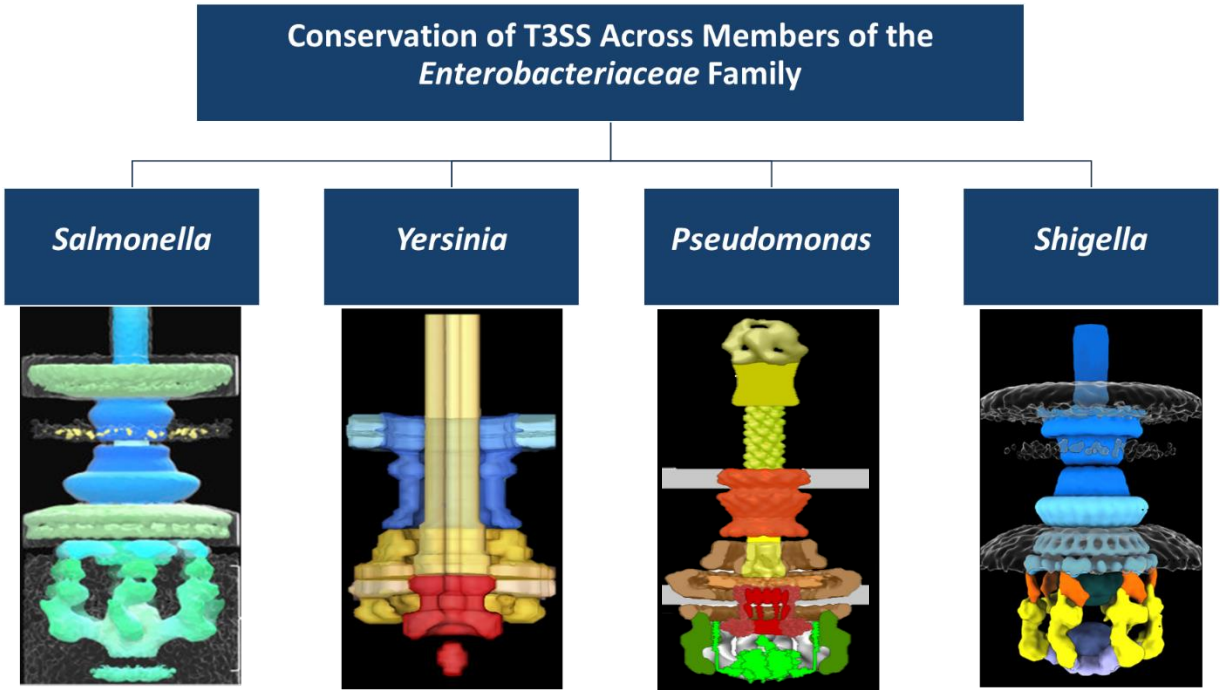


Figure 2 (Above): Multi-Species Overview Comparison of T3SS Injectiomes- As referenced in the text, images above highlight the many similarities found between the injectisomes of *Salmonella enterica*, *Yersinia* species, *Pseudomonas aeruginosa*, and *Shigella flexneri*. It should also be noted that 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.[108,109]

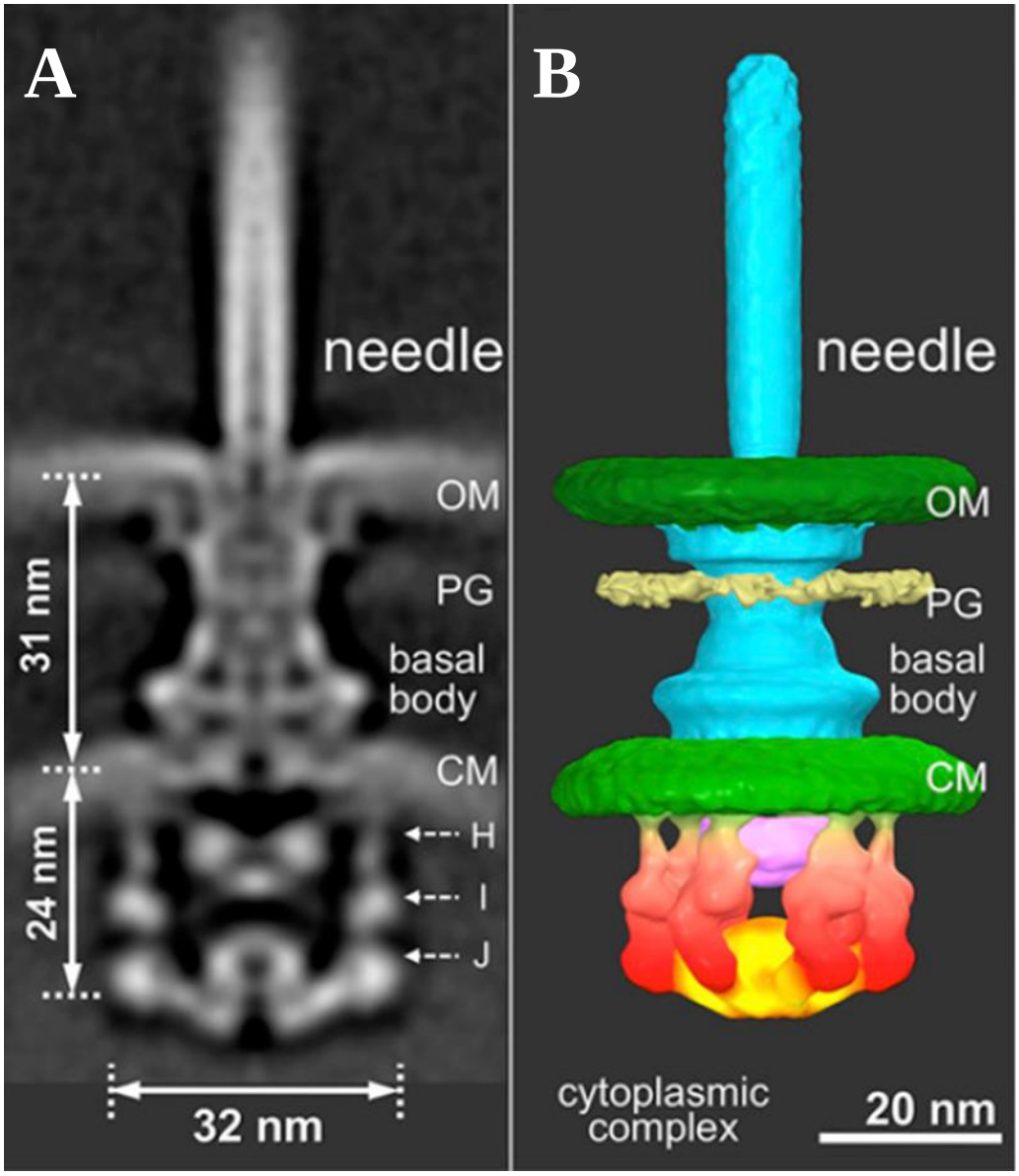


Figure 3: Cryo-ET Model of the Type III Secretion System found in *S. flexineri*- A central subtomogram section average (A) and a 3D surface rendering (B) of the intact injectisome show OM, CM, peptidoglycan (PG), basal body, and needle in detail. Importantly, there is a large cytoplasmic complex that 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).[61]

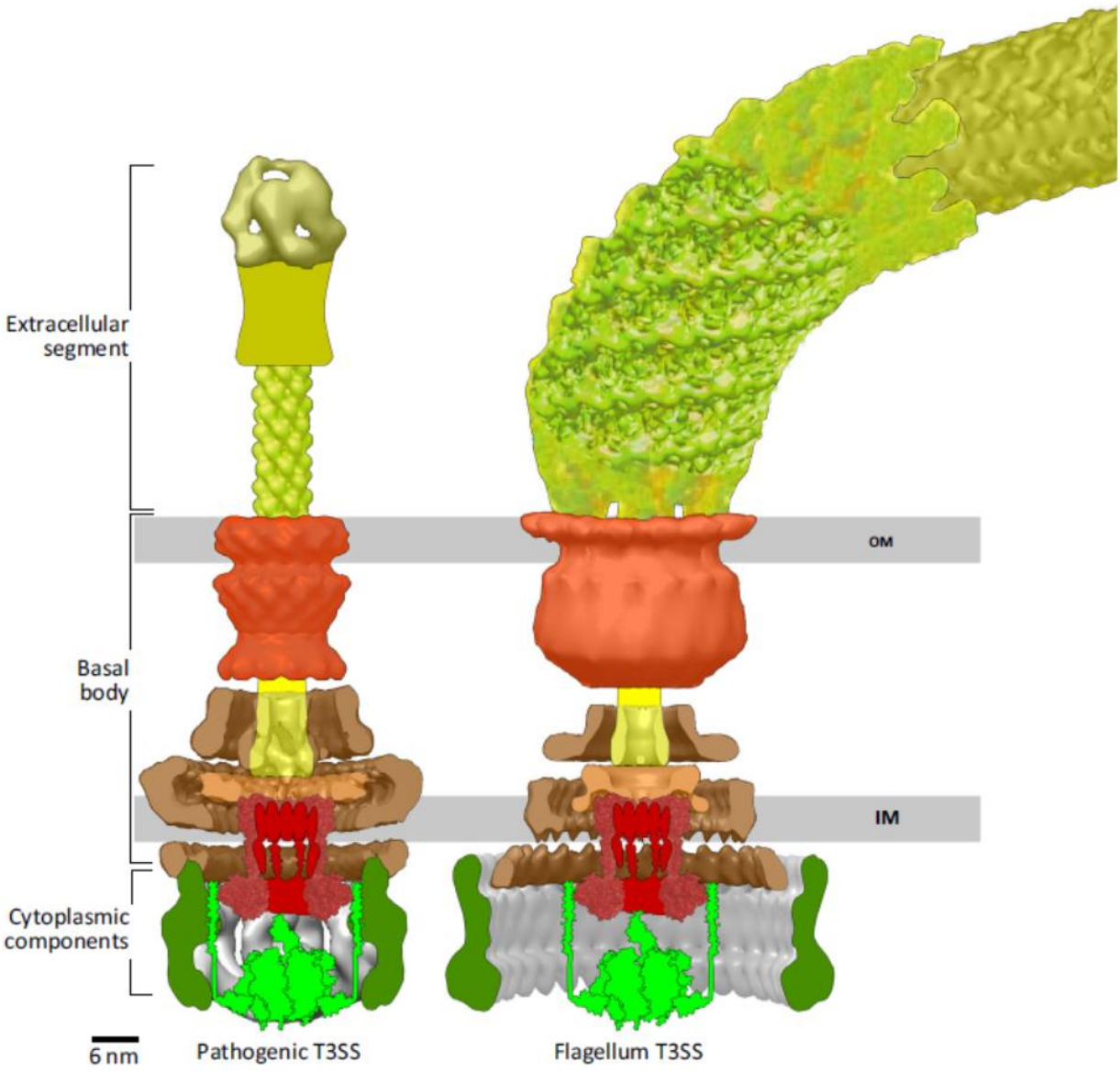


Figure 4: Schematic Comparison of the T3SS Injectisom with Flagellum- As discussed in the text, protein complexes were designed using electron microscopy (EM) structures as a model. With the exception of the extracellular segment, there are many observable structural similarities between the two systems. 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)[110] and EMD accession number 1647 for the hook.[111] 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).[77,110,111]

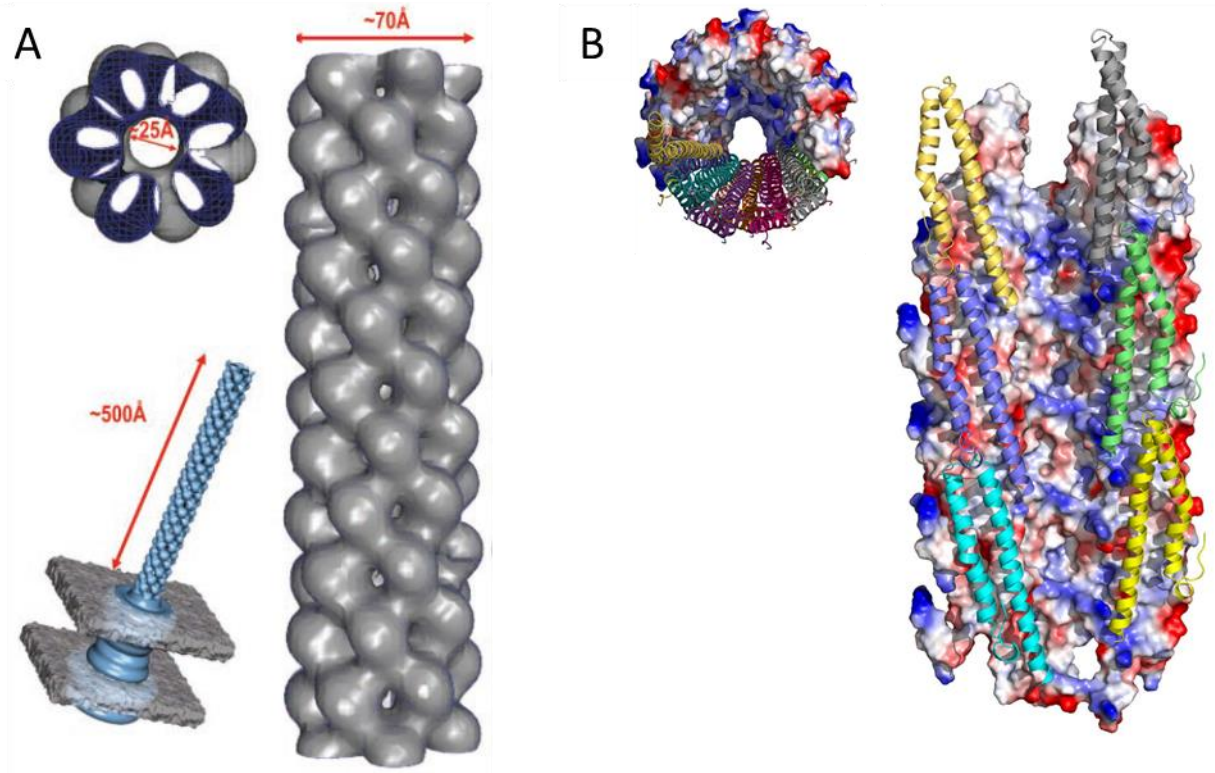


Figure 5: Visualization and Model Schematic of the MxiH Needle Filament Architecture- As described in the text, images A and B provide low and high resolution reconstructions of the MxiH filaments that are viewed from the top and side. Top view of the structures reveals the helical arrangement of the 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.[67] More importantly, in the absence of Ipa B-D, *Shigella* invasion into the host cell does not occur.[68] IpaD serves an important role in facilitating IpaB and IpaC host membrane translocon pore formation and control of their secretion as effectors.[69,70] Further structural information and implication of IpaD in traslocon pore formation can be found by referring to Johnson et al. (2006) and Blocker et al. (2008).[71,72] Following host cell contact and translocon pore formation by IpaB and IpaC, active secretion of host cell effector proteins is triggered.[73] 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.[54,74,75] 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).[44,52–54,56]

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.[76] 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

		Hosts											Flagellar Apparatus
		Humans and Animals								Plants			
Common Name (Sct)	**Function**	**EPEC/EHEC**	***Pseudomonas aeruginosa***	***Yersinia***	***Shigella flexneri***	***Salmonella* sp. SPI-1**	***Salmonella enterica* SPI-2**	***Chlamydia pneumoniae***	***Burkholderia pseudomallei*** [b]	***Pseudomonassyringae***	***Ralstonia solanacearum***	***Xanthomonas* spp.**	
Cytoplasmic Membrane													
SctD	IM ring	EscD	PscD	YscD	MxiG	PrgH	SsaD	CdsD	BsaM	HrpQ	HrpW	HrcD	FliG
Cytoplasmic Components													
SctQ	Cytoplasmic ring	SepQ	PscQ	YscQ	Spa33	SpaO	SsaQ	CdsQ	BsaV	$HrcQ_A$/$HrcQ_B$	HrcQ	$HrcQ_A$; $HrcQ_B$	FliM; FliN
SctL	Stator	EscL	PscL	YscL	MxiN	OrgB	SsaK	CdsL	OrgB	HrpE	HrpF	HrcL	FliH
SctN	ATPase	EscN	PscN	YscN	Spa47	InvC	SsaN	CdsN	BsaS	HrcN	HrcN	HrcN	FliI
SctO	Stalk	EscO	PscO	YscO	Spa13	InvI	SsaO	CdsO	HrpD or BsaT	HrpO	HrpD	HrpB7	FliJ
SctP	Molecular ruler	EscP	PscP	YscP	Spa32	InvJ	SsaP	CdsP	BsaU	HrpP	HpaP	HpaC	FliK
SctW	Gate-keeper	SepL	PopN	YopN/TyeA	MxiC	InvE	SsaL	CopN	BsaP	HrpJ	HpaA	HpaA	/
/	Regulatory compnent	SepD	/	/	/	/	SpiC	/	/	/	/	/	/
SctK	ATPase co-factor	/	PscK	YscK	MxiK	OrgA	/	/	OrgA	HrpD	/	/	/
Chaperones													
/	For early substrates	EscE; EscG	PscE; PscG	YscE; YscG				CdsE; CdsG					
/	For middle substrates	CesAB; CesD; CesD2	PcrG; PcrH	LcrG; SycD; SycB	IpgC; IpgC	SicA	SseA; SsaE	LcrH					FliS; FliT
/	For late substrates	CesT; CesF	SpcU; SpcS	SycE; SycT; SycH; SycN; YscB; YsaK	IpgE; Spa1; IpgA	SicP; SigE; InvB; SigE	SrcA; SscB	SycE	BPSS151	ShcA; ShcM; ShcF; ShcV; ShcO1; ShcS1; ShcS2	HpaB	HpaB	

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).

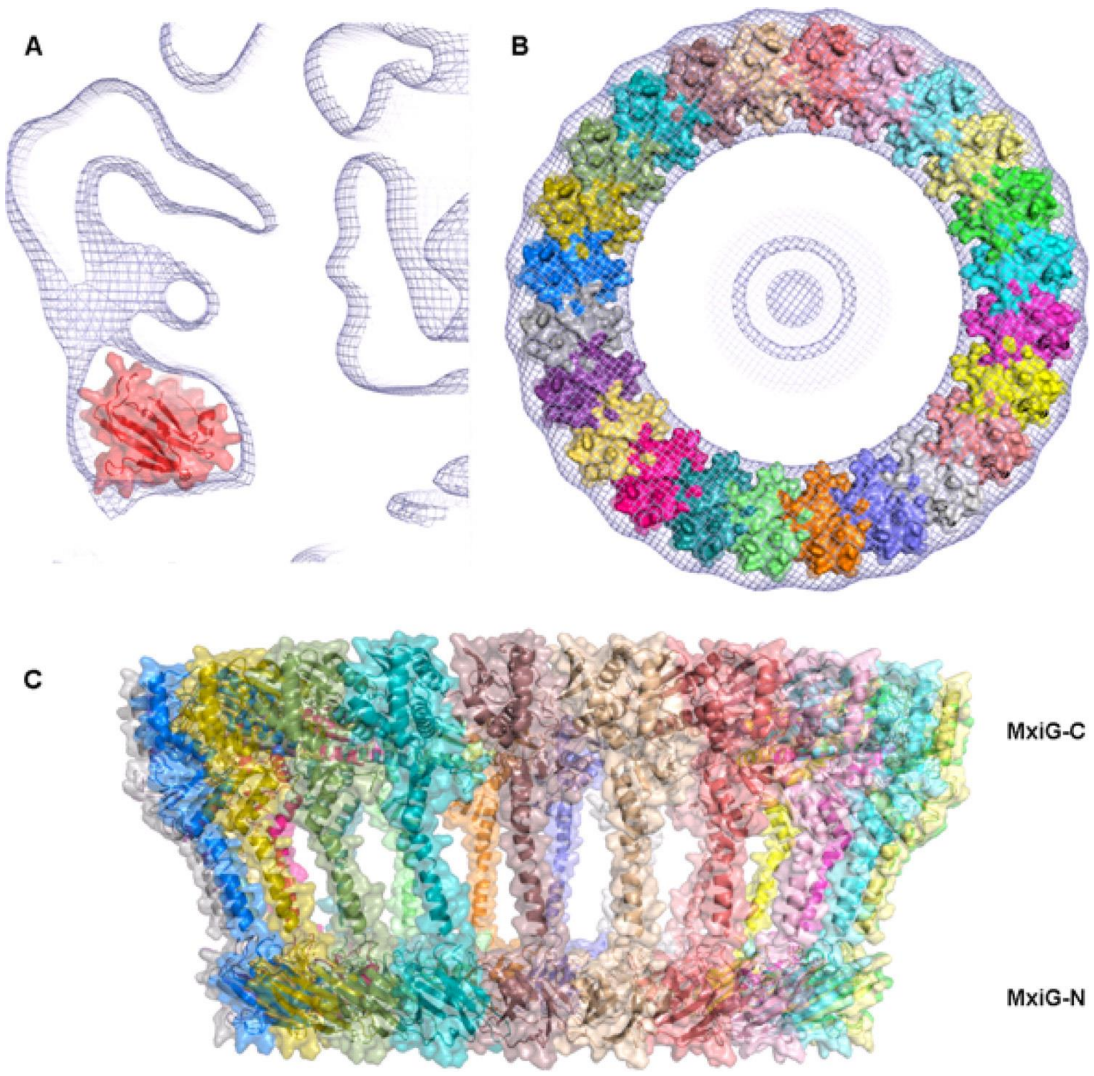


Figure 6: Positioning of MxiG-N(6–112) in the C24 *S. flexneri* NC EM map contoured to 1σ (EMDB 6391). A- A position of MxiG-N(6–112) (*red*) in the density for the leg domains. However, the volume occupied by a single subunit supports the idea that 24 copies can assemble in the ring. **B-** View of 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. [112,113] 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. [83] 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.[84] 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.[61] 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*.[85] This phenomena may lead to the increased exchange rate of YscQ (SctQ) which has been suggested to aid in effector transportation.[86]

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

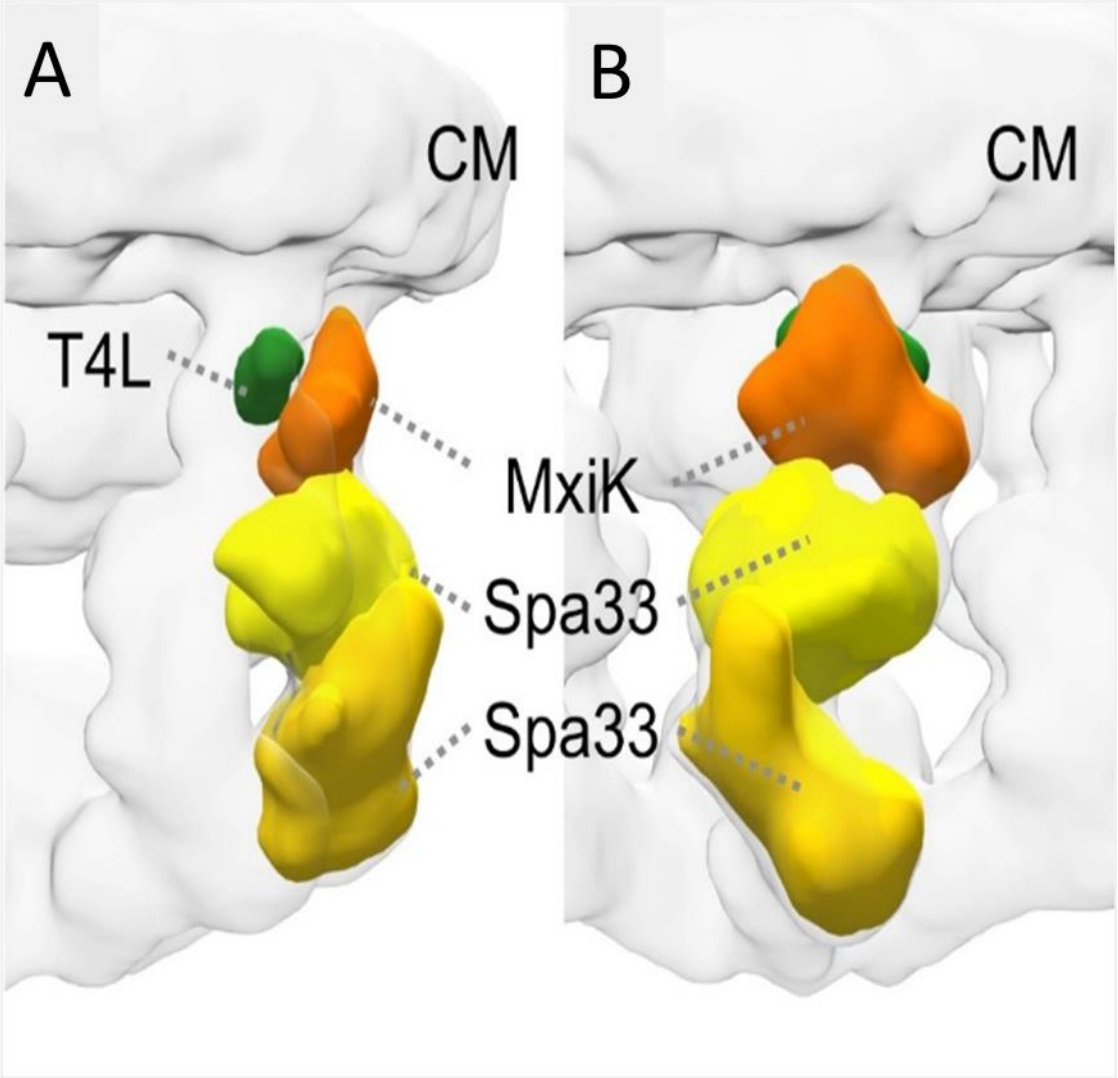


Figure 7: Visualization of the MxiK N terminal-T4L Insertion Mutant- Cryo-ET of MxiK-T4L and Spa33 SP components, expressed in *S. flexineri* mini cells. A sub-tomogram 3D model reveals the localized position of MxiK in between the inner membrane ring and Spa33 below (Tachiyama et al. 2019).

SctD	MxiG *Shigella*	PscD *Pseudomonas*	YscD *Yersinia*	PrgH *Salmonella*	FliG *E.coli*
MxiG *Shigella*	100%	9%	9%	20%	7%
PscD *Pseudomonas*	8%	100%	43%	13%	7%
YscD *Yersinia*	8%	44%	100%	13%	9%
PrgH *Salmonella*	19%	14%	14%	100%	7%
FliG *E.coli*	8%	9%	12%	8%	100%

SctK	MxiK *Shigella*	PscK *Pseudomonas*	YscK *Yersinia*	OrgA *Salmonella*
MxiK *Shigella*	100%	10%	9%	24%
PscK *Pseudomonas*	8%	100%	41%	13%
YscK *Yersinia*	7%	41%	100%	13%
OrgA *Salmonella*	21%	14%	14%	100%

SctQ	Spa33 *Shigella*	PscQ *Pseudomonas*	YscQ *Yersinia*	SpaO *Salmonella*	FliM *E.coli*	FliN *E.coli*
Spa33 *Shigella*	100%	10%	14%	21%	9%	8%
PscQ *Pseudomonas*	10%	100%	40%	18%	6%	9%
YscQ *Yersinia*	14%	40%	100%	20%	8%	10%
SpaO *Salmonella*	20%	19%	20%	100%	10%	9%
FliM *E.coli*	7%	5%	8%	9%	100%	5%
FliN *E.coli*	18%	20%	23%	20%	12%	100%

SctL	MxiN *Shigella*	PscL *Pseudomonas*	YscL *Yersinia*	OrgB *Salmonella*	FliH *E.coli*
MxiN *Shigella*	100%	13%	10%	19%	8%
PscL *Pseudomonas*	14%	100%	54%	15%	19%
YscL *Yersinia*	10%	52%	100%	13%	17%
OrgB *Salmonella*	19%	15%	12%	100%	13%
FliH *E.coli*	8%	18%	16%	13%	100%

Table 3: Multiple Sequence Alignment Distance matrix of SP components- SP protein sequences were alligned 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.[88]

UGENE distance matrix software.[87,88] Homology between SP components (SctD, SctK, SctQ, SctL) were generally below 20 % similarity with exceptions found between *Yersinia entericolitica* 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.[63] 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).[86] In agreement with the *Yersinia* C-ring model, assessments of *Shigella* sorting platform architecture also favored a Spa33 (SctQ) contiguous ring.[76,84] 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*.[89] As mentioned above, T3SS SctQ sequence homology

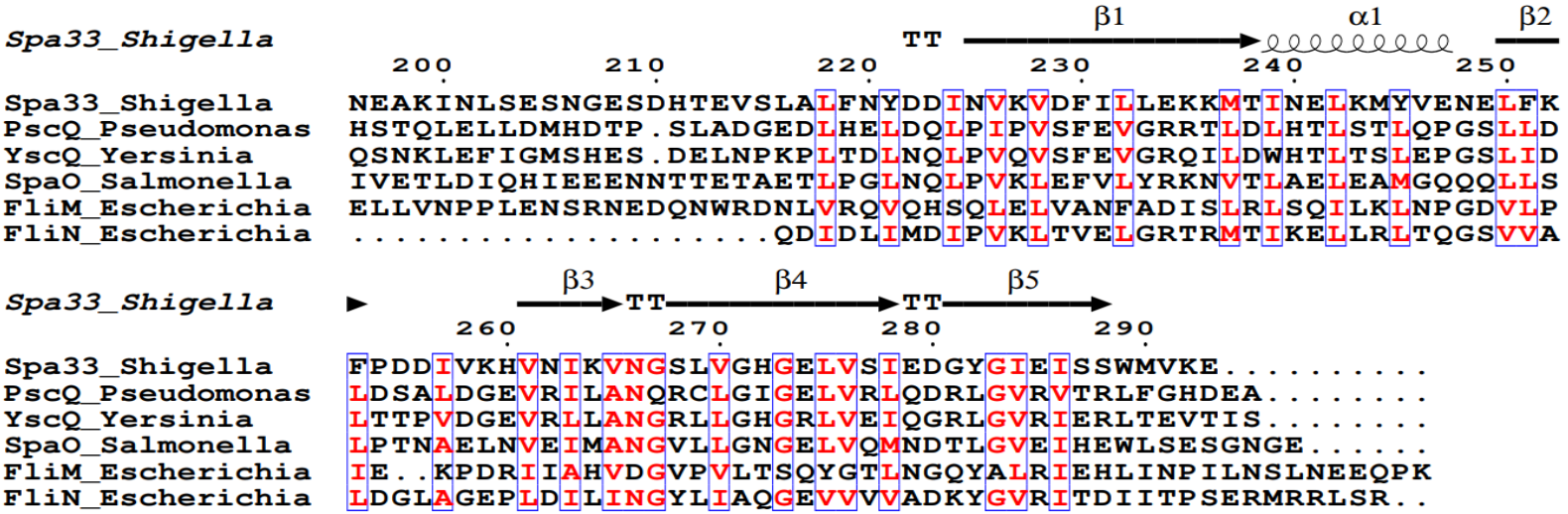


Figure 8: Sequence Alignment of SctQ^{C} Residues- CLUSTALW sequence alignment of SctQ C-terminal residues, rendered by ESPript 3.0, reveals majority of sequence similarity between SctQ SP proteins resides in its C-terminus.[114]

SctQ^{C}	Spa33^{C} *Shigella*	PscQ^{C} *Pseudomonas*	YscQ^{C} *Yersinia*	SpaO^{C} *Salmonella*	FliM *Escherichia*	FliN *E.coli*
Spa33^{C} *Shigella*	100%	19%	22%	34%	6%	25%
PscQ^{C} *Pseudomonas*	19%	100%	55%	32%	11%	27%
YscQ^{C} *Yersinia*	21%	54%	100%	32%	13%	24%
SpaO^{C} *Salmonella*	32%	31%	31%	100%	9%	28%
FliM *E.coli*	5%	10%	12%	9%	100%	12%
FliN *E.coli*	24%	27%	23%	28%	13%	100%

Table 4: Multiple Sequence Alignment Distance matrix of SctQ^{C} components- SP protein sequences were alligned using the CLUSTALW multiple sequence alignment program. Distance matrices were calculated using UGENE software algorithms and presented as percent sequence similarity.[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.

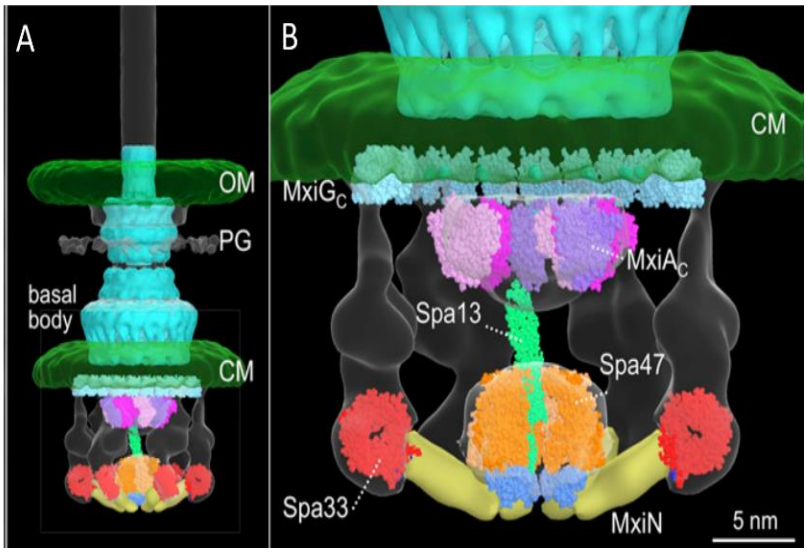


Figure 9 A: Cryo-electron tomography image of the extracellular needle, basal body, and cytoplasmic sorting platform (SP).[1] **B:** A closer look at the sorting platform reveals the 5 essential components of the SP, including Spa33 (red), MxiN (yellow), Spa47 (orange), spa13 (green), and MxiK (grey).[1]

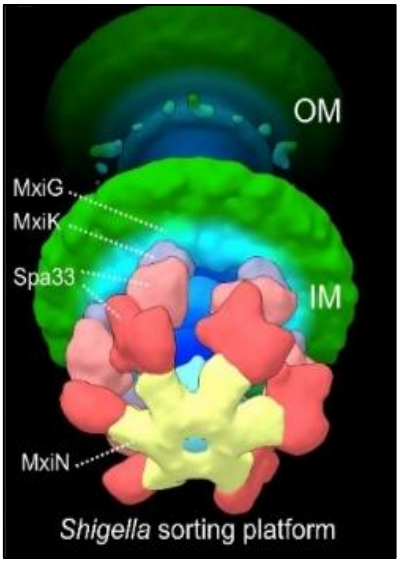


Figure 10: Cryo-ET model of the six radiating spokes of MxiN (Yellow), connected to six Spa33 pods (Red).

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.[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.[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).[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.*[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

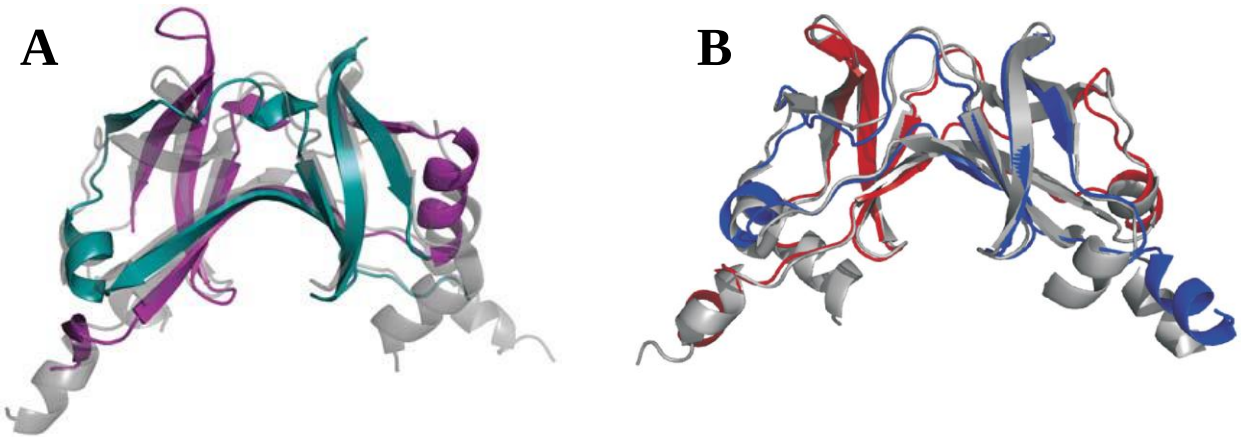


Figure 11: Crystal Structure of SctQ^{C} Overlay with FliN- A. The crystal structure of Spa33^{C} (a.a 208-293) is shown as a ribbon diagram. Chain A and B dimers are colored teal and purple respectively. The Spa33^{C} dimer is superimposed with the *T. maritima* FliN structure (pdb id 1YAB), shown in 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^{C} found in the T3SS SP and the flagellum C-ring.

implying the existence of $SctQ^{C}$ as a dimer.[78,92] Crystal structure alignment of $SctQ^{C}$ 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 $Spa33^{FL}$ and $Spa33^{C}$ ($SctQ^{FL/C}$) subunits was determined to be necessary for assembly of the sorting platform and T3SS function in *Shigella*. Alternatively, *Salmonella* secretion and invasion studies determined that $SpaO^{C}$ ($SctQ^{C}$) is a non-essential component for T3SS function, but is required for the stability of $SpaO^{FL}$ ($SctQ^{FL}$).[94] *In vitro* evidence also suggests that the two protein subunits aggregate to form a heterotrimer through a 1:2 stoichiometric complex of $Spa33^{FL}$-$Spa33^{C}{}_{2}$ ($Spa33^{WT}$).[78,94] Structural observations in *Shigella* are consistent with *in vivo* and *in vitro* results found in *Salmonella* $SpaO^{WT}$ trimeric complex formation.[94] For SpaO, it was determined that the $SpaO^{C}$ component is not essential for SP injectisome function. On the other hand, *Yersinia,* $YscQ^{C}$ ($SctQ^{C}$) was deemed essential for T3SS function. McDowell et al.(2015) suggests that $Spa33^{C}$ is also essential for *Shigella* T3SS, but extensive *in vivo* evidence to support an essential heterotrimeric $Spa33^{WT}$ 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 $Spa33^{WT}$ interactions with MxiN and MxiK.

Characterization of the $Spa33^{FL,C}$ subunit interaction using *in vivo* functional analysis and *in vitro* biophysical methods will provide us with a better understanding of how the $Spa33^{WT}$ 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).[95] 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 His_6-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 flowthorugh 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 *pkT25* expression vector will fuse the T25 subdomain to the proteins N-terminus; alternatively, the *pkNT25* 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 $Spa33^{WT/FL/C}$, 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.[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:

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

θ is the CD signal in mDeg, *mw* is the molecular weight of the protein, C is concentration in mg/mL, ℓ 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 $Spa33^{WT}$ and $Spa33^{C}$. 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.[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 *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 sheering 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 resupended in 400 µL of PBS and 200 µL of 3X SDS-PAGE Buffer. For Western bloting, 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 usedapplied 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.

Contact-mediated hemolysis: The ability of *Shigella* to form and insert functional translocons into target red blood cell (RBC) host membranes was tested by contact-mediated hemolysis. *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 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:

$$resuspension\ volume = \frac{(200\ \mu L\ \text{x OD})}{0.5}$$

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 $Spa33^{WT}$ compliment values.

Results

$Spa33^{FL/C}$ Co-purification

Existence of an alternative translation initiation (ATI) site in the sequence of Spa33 ($Spa33^{WT}$) was first reported by McDowell et al. (2016).[78] 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 $Spa33^{FL}$ and a shorter truncated C-terminal copy of $Spa33^{C}$. 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. $Spa33^{WT}$ was then expressed in *E. coli* Tuner^{TM} 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 $Spa33^{WT}$ yielded two distinct bands at ~11.5 kDa and ~38kDa, which correlates with the molecular weight of $Spa33^{C}$ and N-terminally His-tagged $Spa33^{FL}$. 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 $Spa33^{FL}$ and $Spa33^{C}$ copies have a propensity to interact with each other. To test this hypothesis required isolating $Spa33^{FL}$ and its $Spa33^{C}$ subunits.

Isolation of $Spa33^{C}$ was achieved by selective primer design for nucleotides encoding for residues 208-293 with the same purification protocol as $Spa33^{WT}$. Elution fractions were analyzed by SDS-PAGE (**Figure 15**). $Spa33^{FL}$ expression vector design was achieved by *spa33* inverse PCR with primers containing point mutations at nucleotides 564 and 576, both of which

A. Spa33^{WT} Genomic Sequence

ATGCTAAGAATTAAACATTTTGACGCTAACGAAAAACTACAGATTTTATATGCAA
AGCAACTCTGCGAGCGTTTTTCGATTCAGACATTCAAAAATAAATTTACAGGCAG
TGAAAGTTTAGTCACTCTTACTTCCGTGTGTGGGGATTGGGTAATTCGTATTGATA
CATTATCTTTTTTGAAAAAAAAATACGAGGTATTTTCAGGATTTTCTACACAAGA
ATCTTTACTGCATTTATCAAAATGTGTCTTTATAGAGTCGTCATCTGTATTTTCGAT
TCCAGAACTGTCTGATAAGATTACTTTCCGGATCACGAATGAAATCCAGTATGCA
ACTACTGGGAGTCATTTATGCTGTTTTTCCTCTTCTTTAGGTATTATTTATTTTGAC
AAGATGCCGGTATTACGTAATCAAGTTTCTCTTGACTCATTGCATCATCTTTTAGA
GTTTTGCTTAGGTTCATCTAATGTAAGGCTGGCTACTTTAAAAAGAATTCGCACTG
GTGATATAATCATAGTTCAGAAACTTTATAATTTATTATTGTGTAATCAAGTTATT
ATTGGGGATTATATTGTGAATGATAATAATGAGGCAAAAATTAATCTGTCAGAA
AGTAATGGTGAGTCAGAACACACAGAAGTTTCTTTGGCATTATTCAATTATGATG
ATATCAATGTAAAAGTGGACTTTATTCTTTTAGAAAAAAATATGACAATCAATGA
ACTAAAAATGTATGTAGAAAACGAATTATTCAAGTTTCCCGATGACATAGTTAAA
CATGTAAATATTAAAGTAAATGGTTCTTTGGTTGGGCATGGGGAACTTGTTTCTAT
TGAGGATGGTTATGGTATCGAGATTAGTTCTTGGATGGTAAAGGAGTAA

B. Alternative Translation Initiation (ATI) Site

Spa33^{WT}: [TTATTGGGGATTATATTGTGCACAG**]**

Point Mutations

Spa33^{FL}: [TTATT GGTGATTATATTGTCCACAG**]**

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 (nuceotides 574-776: red). This region permits the translation of two Spa33 products. A full length copy (Spa33^{FL}) and shorter C-terminal copy (Spa33^{C}). **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 Spa33^{C} alternative start codon (ASC-red) respectively. The mutations made in the RBS and ASC prevent activation of the Spa33^{C} ATI site and thus only expression of Spa33^{FL} will occur.

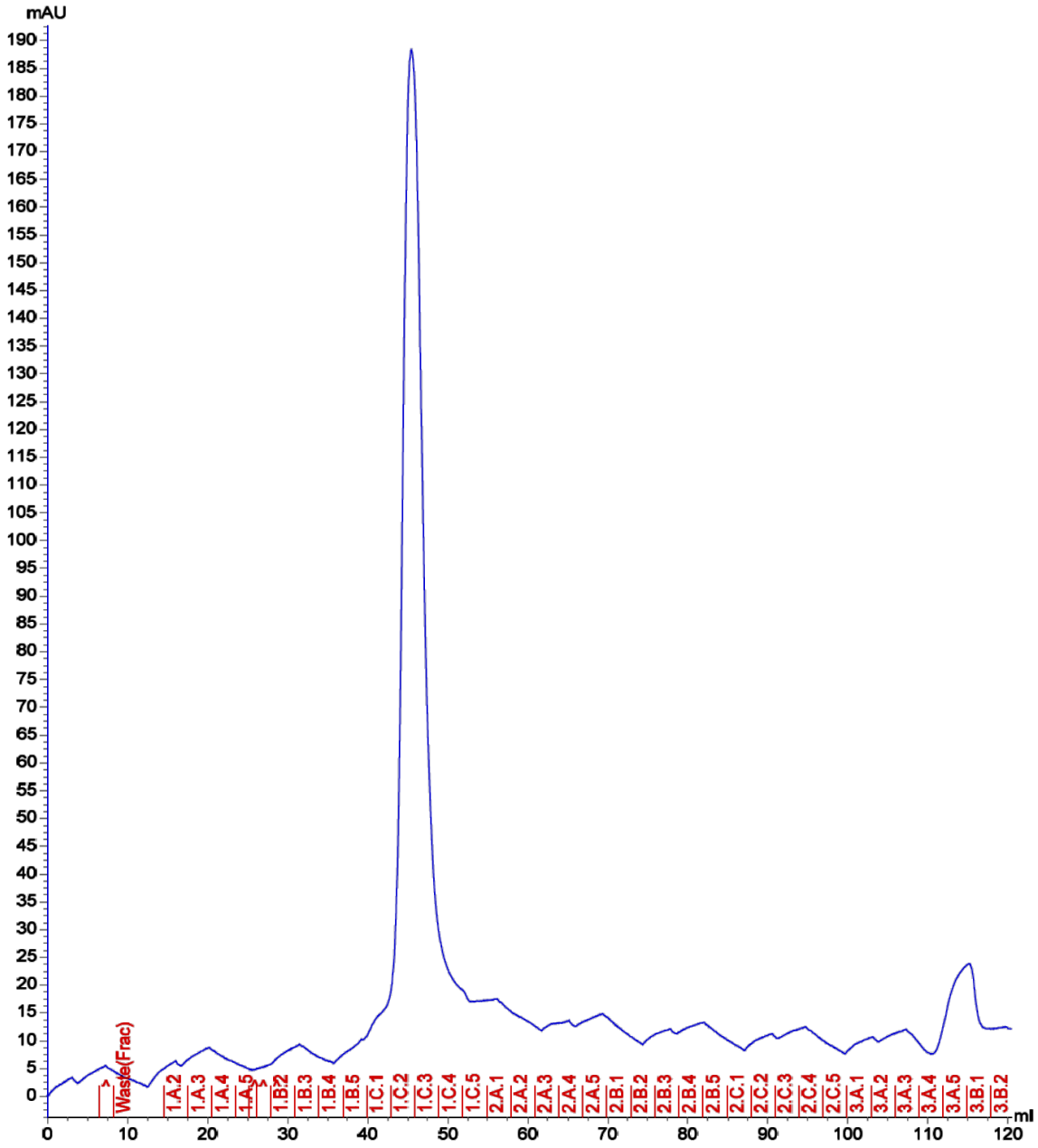
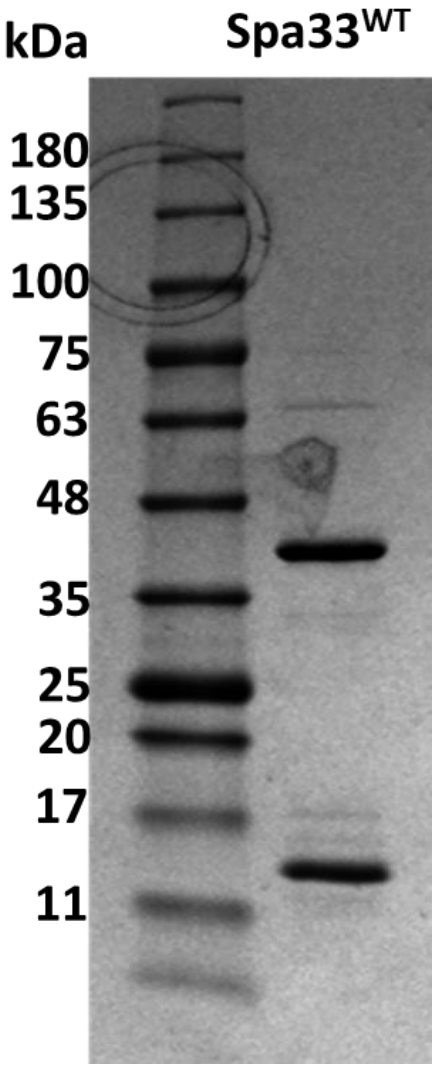


Figure 13: Size Exclusion Gel-Chromatogram of Spa33^{WT}- Ni-affinity purification followed by SEC purification of expressed Spa33 yields a singular peak following from 120ml Superdex gel matrix column. This is suggestive that the Spa33^{FL} copy and the alternatively translated Spa33^{C} copy are able to form an interacting complex.



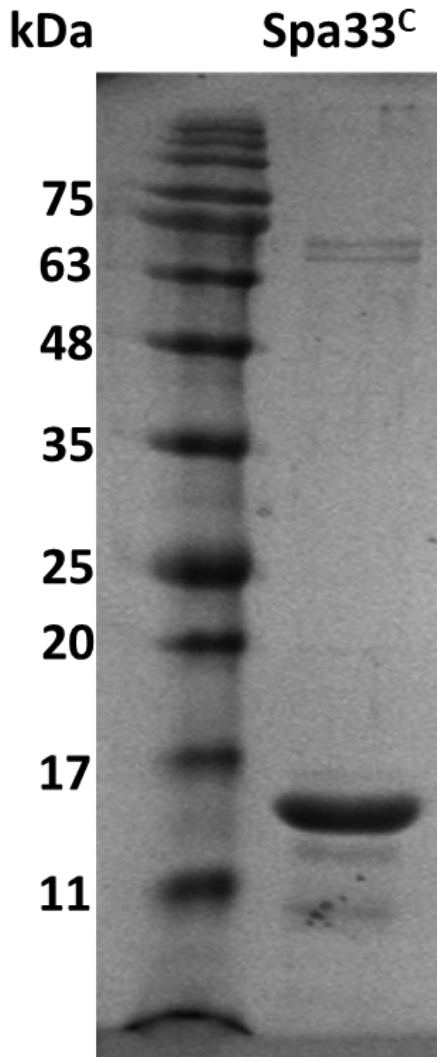


Figure 14 (Left): SDS-PAGE Purified Spa33^{WT}- Post recombinant N-terminal His-tagged Spa33^{WT} expression and successive Ni-Affinity, SEC purification, and SDS-PAGE analysis of the singular elution peak from the SEC chromatogram in **Figure 13** yielded two distinct bands, which run equivalently to the molecular weights of N-terminal His-Spa33^{FL} (~38kDa) and Spa33^{C} (~11.5 kDa). Band identity was confirmed by mass spectrometry (data not included).

Figure 15 (Right): SDS-PAGE of Purified Spa33^{C}- Post recombinant N-terminal His-tagged Spa33^{C} 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 Spa33^{C+}(~11.5 kDa). Based on multi-angle dynamic light scattering and size exclusion chromatography, Spa33^{C} may exist as a dimer in solution (hydrodynamic dimeter = 8 nm) that is able to interact with a single copy of Spa33^{FL} to give a heterotrimer of Spa33^{FL}(Spa33^{C})$_2$ (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 $Spa33^{C}$ ATI site recognition resulting in the expression of only $Spa33^{FL}$. We were unable to purify $Spa33^{FL}$ after recombinant expression because of its insoluble nature in the absence of the $Spa33^{C}$ component. $Spa33^{FL}$ was poorly soluble in the absence of $Spa33^{C}$, 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 $Spa33^{FL}$ is N-terminally His-tagged, whereas the ATI site for $Spa33^{C}$ 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 $Spa33^{FL}$ and $Spa33^{C}$ interact minimally in 1:2 fashion ($Spa33^{FL}$-$Spa33^{C}_{2}$ equivalent to $Spa33^{WT}$) 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 $Spa33^{WT}$ and the Ishihara et al. (2005) report of $Spa33^{WT}$ interaction with other SP components using GST-pull down assays are all *in vitro* techniques.[90] 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 $Spa33^{WT}$ 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.

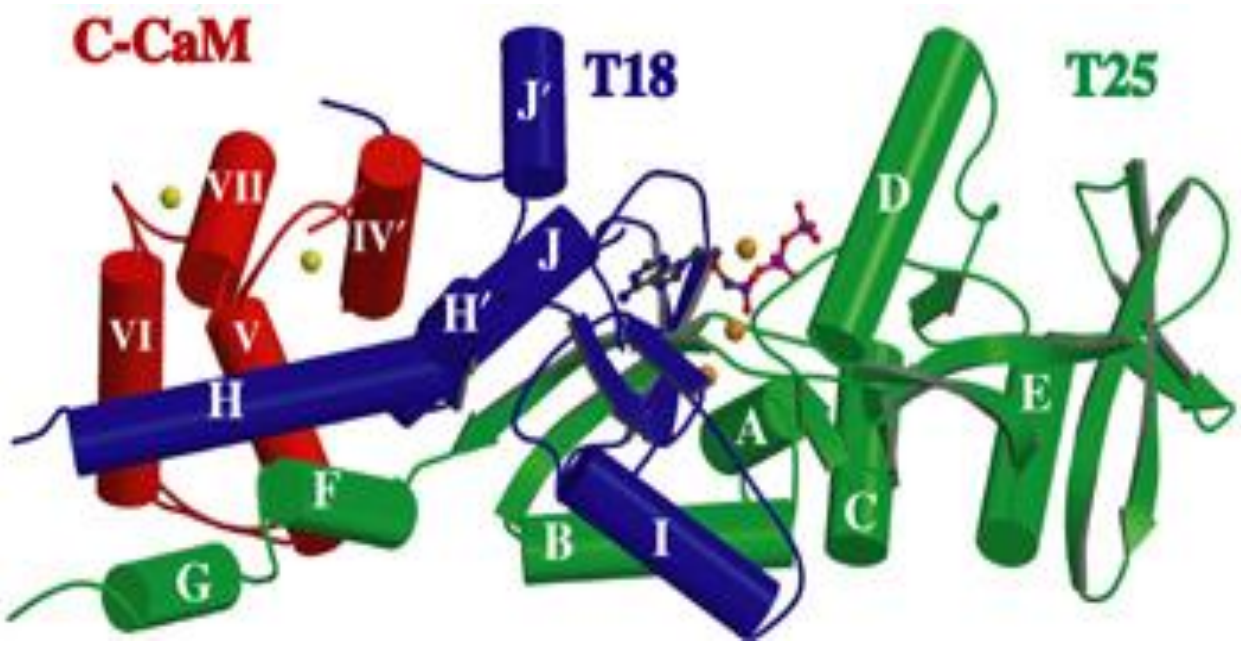


Figure 16: **Secondary structure of CaM-Bound to CyaA Catalytic Domain-** T25 and T18 domains are colored in green and purple. Calmodulin (red) is bound to the T18/T15 subdomains which will inherently restore cAMP synthesis. Figure acquired from Guo et al. (2005). [101]

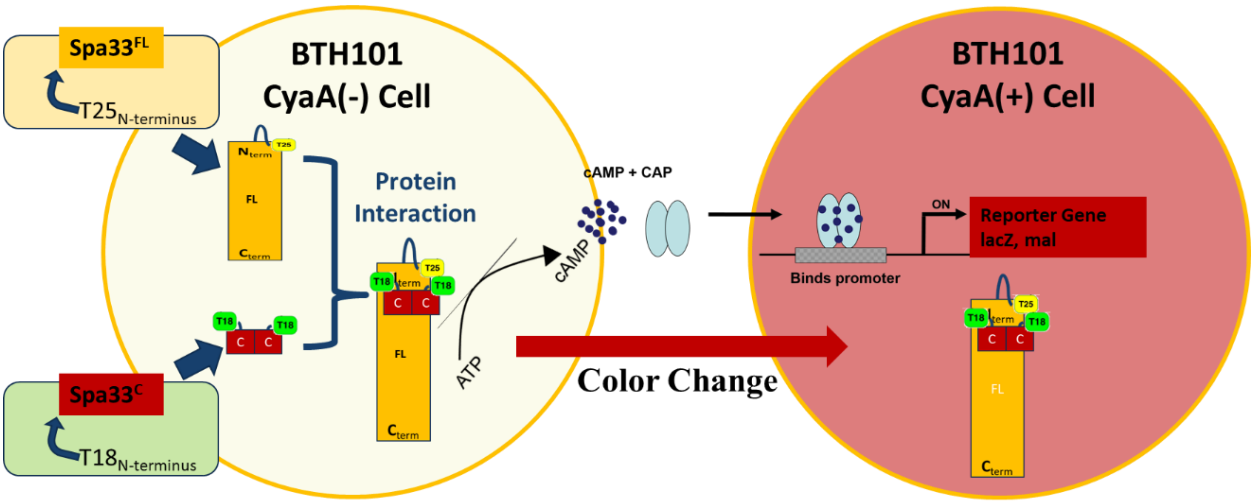


Figure 17: BACTH Analysis Model- A BACTH assay was used to test for interactions between different Spa33 forms. The proteins of interests are fused with the inactive adenylate cyclase catalytic domains T25 and T18. A positive interaction between two hybrid proteins brings T18 and T25 together and restores CyaA activity in the *E. coli* BTH101 cells. Restored CyaA activity results in a BTH101 phenotypic color change.

$Spa33^{FL}$ with $Spa33^{C}$ Permutations

While previous studies have provided *in vitro* support for the $Spa33^{FL/C}$ 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 $Spa33^{FL}$ and $Spa33^{C}$ on LB-XGal and MacConkey indicator media. These combinations include: $Spa33^{FL}_{Nterm\text{-}T25}$ with $Spa33^{C}_{Cterm\text{-}T18}$, $Spa33^{FL}_{Nterm\text{-}T25}$ with $Spa33^{C}_{Nterm\text{-}T18}$, and $Spa33^{FL}_{Nterm\text{-}T18}$ with $Spa33^{C}_{Cterm\text{-}T25}$. When comparing the positive combinations in **Figure 18**, we observe that in each of the three cases, only the $Spa33^{C}$ domain permutations are shown to interact with a $Spa33^{FL}$ permutation that has T25/T18 fused to its N-terminus. Assuming that the N and C-termini of $Spa33^{FL}$ are located on polar opposite ends of each other, we could hypothesize that the N-terminal region of $Spa33^{FL}$ contains the domain for the binding of $Spa33^{C}$. In parallel, it appears that both $Spa33^{C}$ N and C-terminal-ends are capable of interacting with $Spa33^{FL}$.

$Spa33^{C}$ with $Spa33^{C}$ Permutations

Even with confirmation of the direct intracellular protein-protein interaction of $Spa33^{FL}$ with $Spa33^{C}$ in **Figure 18**, this interaction required further investigation. I hypothesize that there are two possible scenarios for the interaction of $Spa33^{C}$ with the N-terminus of $Spa33^{FL}$: **1)** Two $Spa33^{C}$ domains dimerize prior to interacting with the FL protein or **2)** Two $Spa33^{C}$ domains interact separately with the FL protein. Previous research concerning *Shigella* and *Yersinia*, has provided crystal structures (**Figure 11**) of the $Spa33^{C}$ and $YscQ^{C}$ domains, with a crystal unit cell containing two $Spa33^{C}$ chains interacting to form a dimer. This provides support for scenario **1**

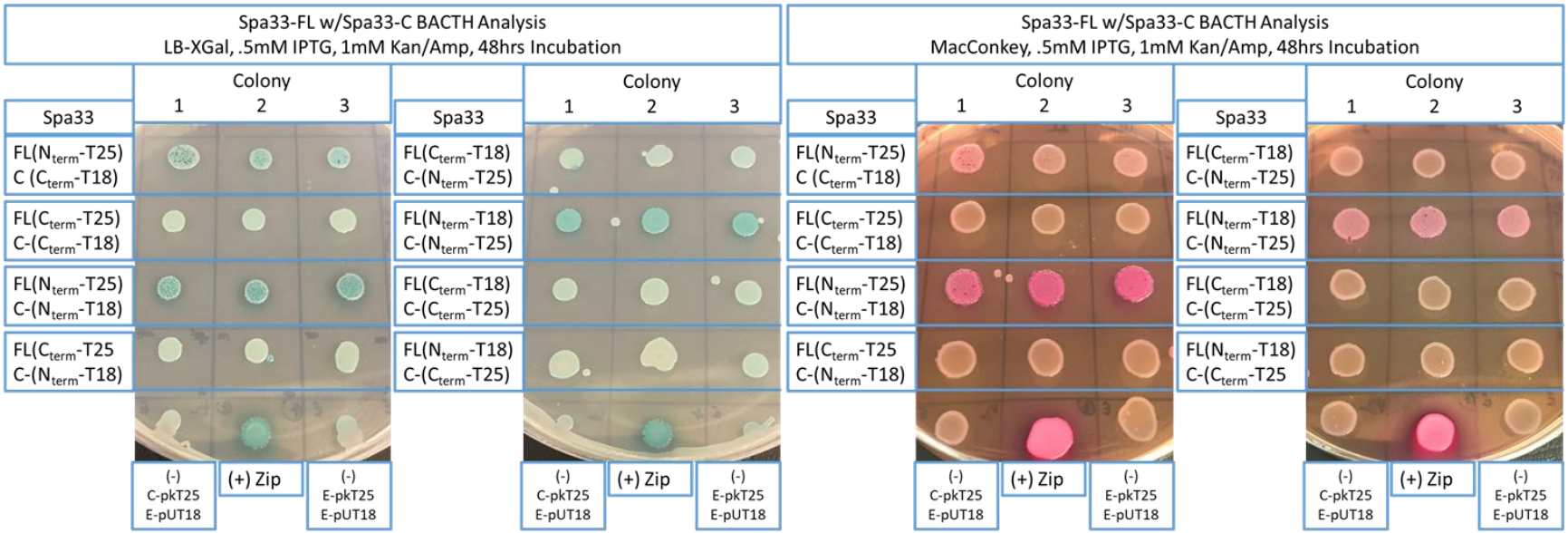


Figure 18: **Spa33^{FL} /Spa33^{C} Permutation BACTH Screening Plates**- BACTH analysis for the screening of Spa33^{FL} and Spa33^{C} protein-protein interactions. A positive protein-protein interaction is dictated by the phenotypic expression and observation of blue colonies on LB-XGal and red colonies on MacConkey indicator plates. A positive control is indicated by the (+) Zip protein, and the negative controls (-) are indicated accordingly.

Where the $Spa33^{C}$ domain forms a dimer prior to its interaction with the FL protein. If this dimerization occurs *in vivo*, we should be able to easily probe this interaction using BACTH analysis. **Figure 19** BATCH screening indeed supports our hypothesis that $Spa33^{C}$ can form a dimer in the absence of the FL protein. The observed positive interactions in **Figure 19** include $Spa33^{C}_{Nterm\text{-}T25}$ with $Spa33^{C}_{Cterm\text{-}T18}$ and $Spa33^{C}_{Nterm\text{-}T25}$ with $Spa33^{C}_{Nterm\text{-}T18}$. When we examine the crystal structures in **Figure 11,** the observed positive interaction between two permutation combinations is not unexpected because the N and C-terminal-ends of a single $Spa33^{C}$ 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 $Spa33^{C}$ interacting with the N-terminus of $Spa33^{FL}$ in **Figure 19**.

$Spa33^{WT}$ permutations and $Spa33^{FL}$ with $Spa33^{FL}$ Permutations

At this point in our Spa33 BATCH screens, we have provided data to support a protein-protein interaction that involves $Spa33^{C}$ dimer with the N-terminus of $Spa33^{FL}$ to form a heterotrimer complex ($Spa33^{WT}$). In **Figure 20,** T25/T18 permutations of $Spa33^{WT}$ 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 $Spa33^{WT}$ plasmids are expressing $Spa33^{FL}$ and $Spa33^{C}$ 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 $Spa33^{WT}$ oligomerization did occur it would probably involve interactions between $Spa33^{FL}$ components.

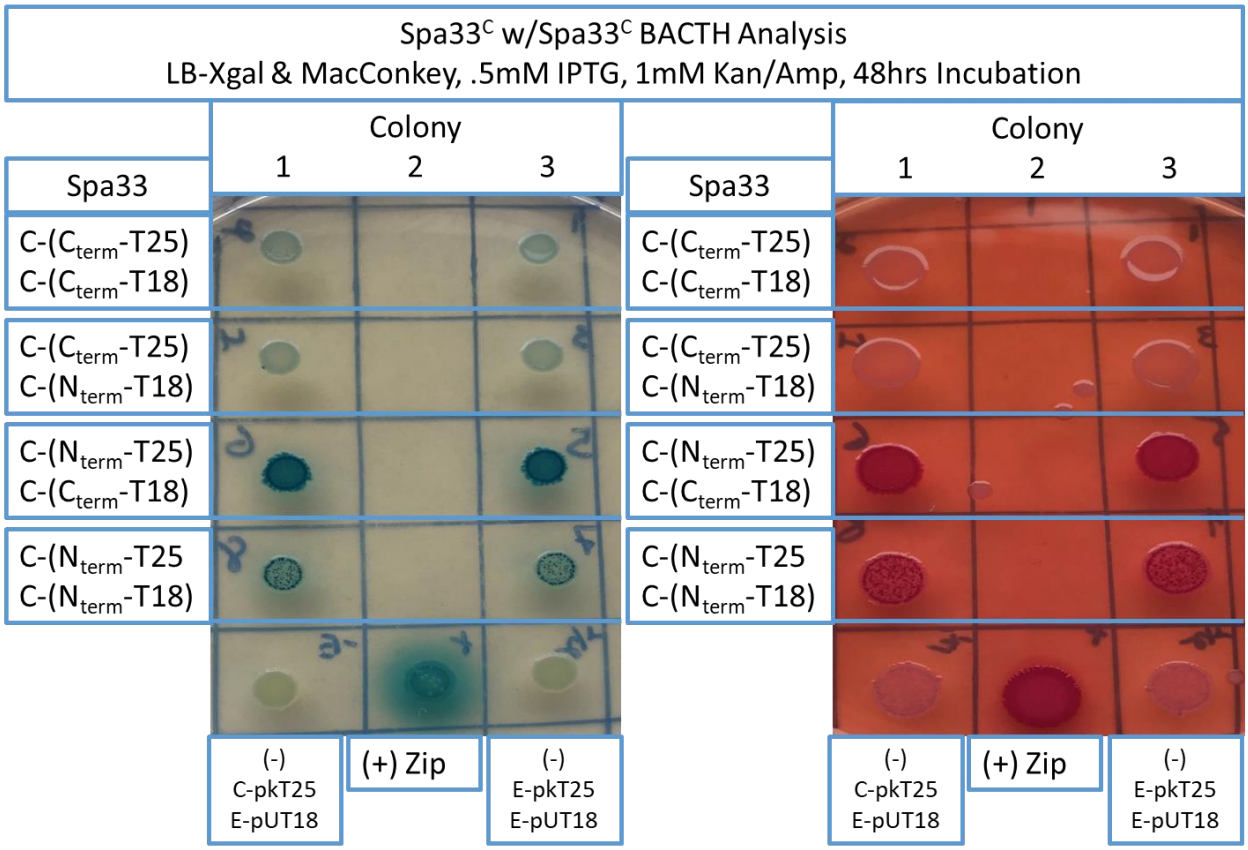


Figure 19: Spa33^{C} /Spa33^{C} Permutation BACTH Screening Plates- BACTH analysis for the screening of Spa33^{C} and Spa33^{C} protein-protein interaction. A positive protein-protein interaction is dictated by the phenotypic expression and observation of blue colonies on LB-XGal and red colonies on MacConkey indicator plates. A positive control is indicated by the (+) Zip protein, and the negative controls (-) are indicated accordingly.

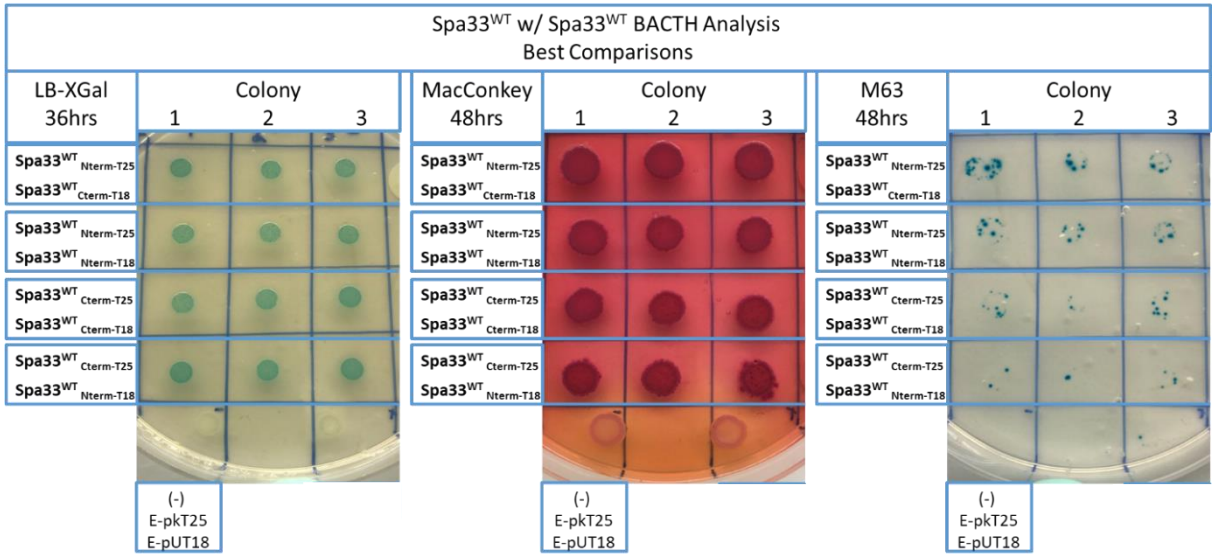


Figure 20: Spa33^{WT} /Spa33^{WT} Permutation BACTH Screening Plates- BACTH analysis for the screening of Spa33^{WT} and Spa33^{WT} protein-protein interaction. A positive protein-protein interaction is dictated by the phenotypic expression and observation of red colonies on MacConkey and blue colonies 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^{FL} /Spa33^{FL} Permutation BACTH Screening Plates- BACTH analysis for the screening of Spa33^{FL} and Spa33^{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.

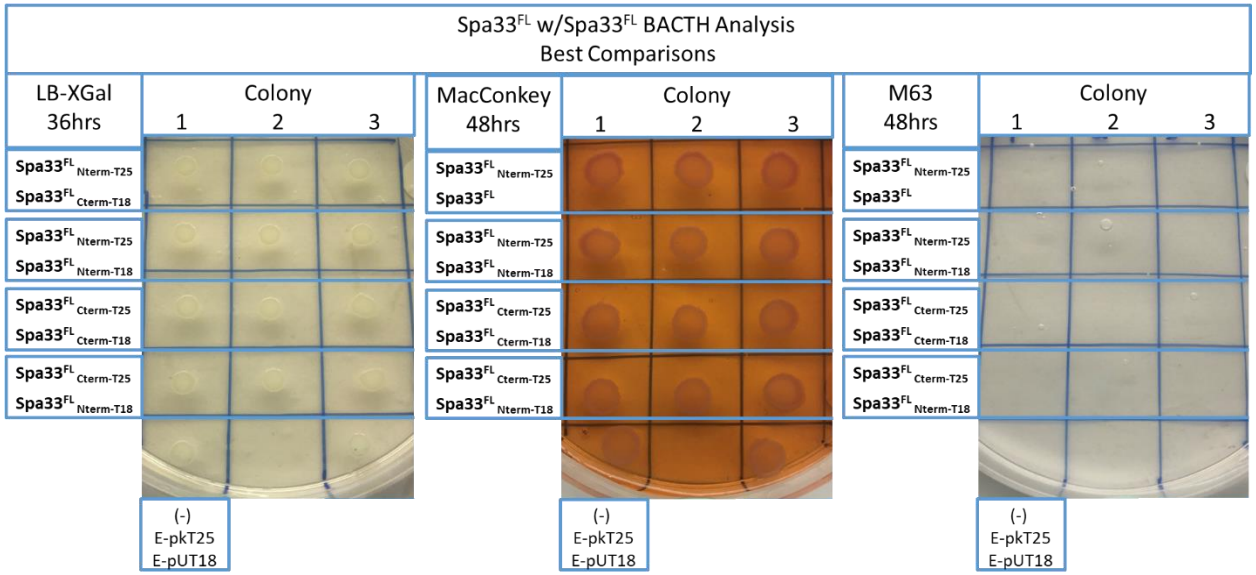


Figure 21 depicts screening for the protein-protein interaction of $Spa33^{FL}$ and $Spa33^{FL}$. As we can see in **Figure 21**, no positive interactions were observed, supporting the idea that $Spa33^{WT}$ heterohexamer formation is unlikely.

$Spa33^{WT}$ with MxiN Permutations

We next explored how $Spa33^{WT}$ 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 $Spa33^{WT}$ complex and MxiN was investigated and reveals two combinations expressing a positive interaction. The positive interactions included $MxiN_{Nterm\text{-}T25}$ with $Spa33^{WT}{}_{Cterm\text{-}T18}$, and $MxiN_{Nterm\text{-}T25}$ with $Spa33^{WT}{}_{Nterm\text{-}T18}$. These results could indicate that the N or C-terminus of the $Spa33^{WT}$ complex is able to interact only with the N-terminus of MxiN. To investigate this interaction further, $Spa33^{FL}$ and $Spa33^{C}$ interactions with MxiN were individually assessed.

$Spa33^{FL}$ with MxiN and $Spa33^{C}$ with MxiN Permutations

BACTH analysis in **Figure 23** shows a positive interaction with $MxiN_{Nterm\text{-}T25}$ and $Spa33^{FL}{}_{C\text{-}term\text{-}T18}$. These data suggest that the C-terminus of $Spa33^{FL}$ 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 $Spa33^{WT}$ 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 $Spa33^{C}$ doesn't interact with MxiN. Therefore, the observed interaction (**Figure 23, Row 2)** between the $Spa33^{WT}$ N-terminus with MxiN probably doesn't involve the presence of the $Spa33^{C}$ dimer.

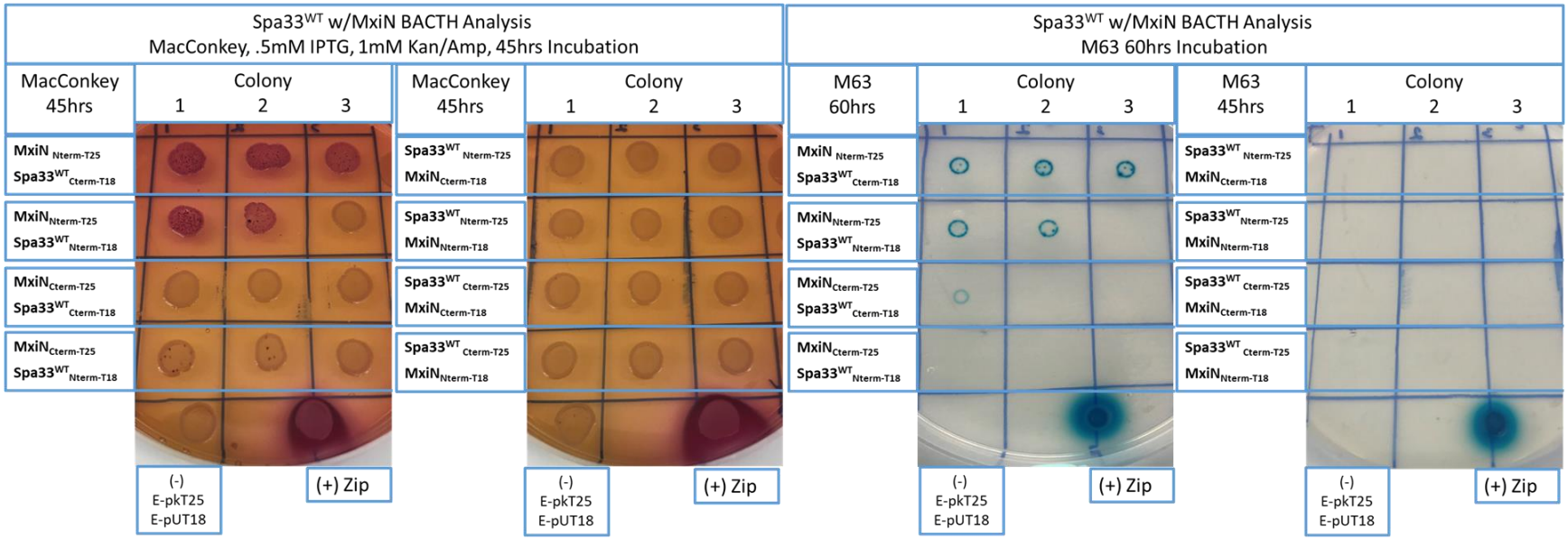


Figure 22: Spa33^{WT} /MxiN Permutation BACTH Screening Plates- BACTH analysis for the screening of MxiN and Spa33^{WT} protein-protein interaction. A positive protein-protein interaction is reflected by the phenotypic expression and observation of red colonies on MacConkey and blue colonies on M63 indicator plates. A positive control is indicated by the (+) Zip protein, and the negative controls (-) are indicated accordingly.

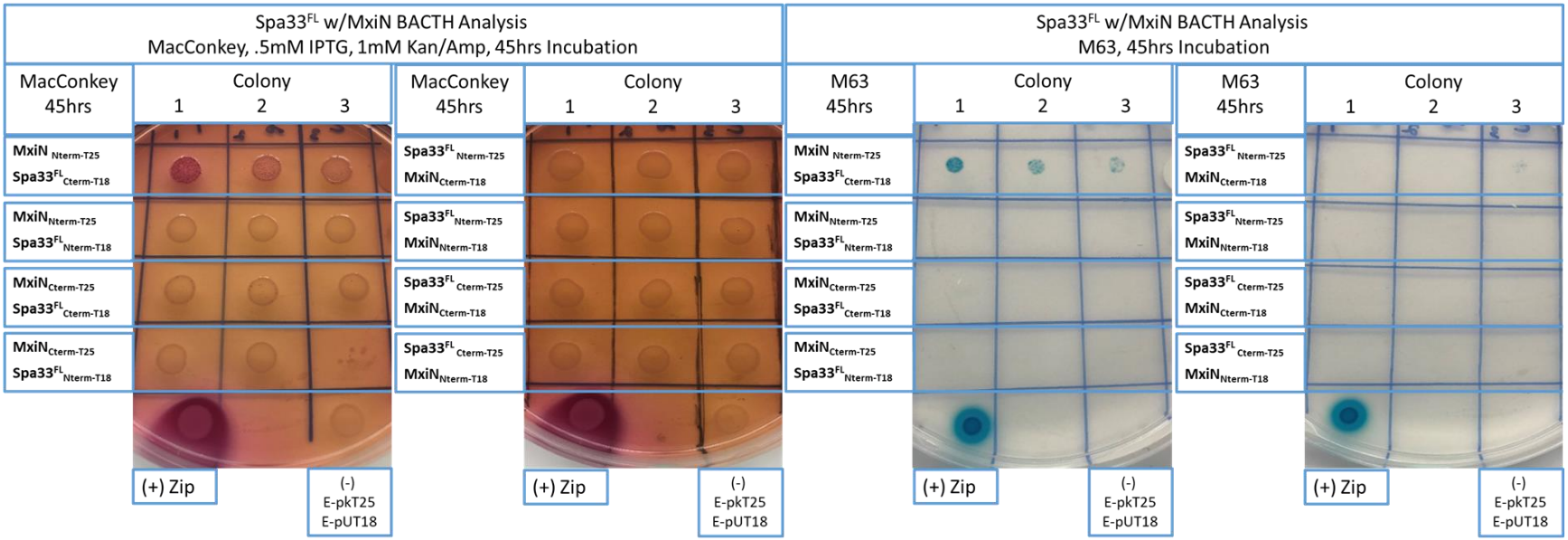


Figure 23: Spa33^{FL} /MxiN Permutation BACTH Screening Plates- BACTH analysis for the screening of MxiN and Spa33^{FL} protein-protein interactions. A positive protein-protein interaction is reflected in the phenotypic expression and observation of red colonies on MacConkey and blue colonies on M63 indicator plates. A positive control is indicated by the (+) Zip protein, and the negative controls (-) are indicated accordingly.

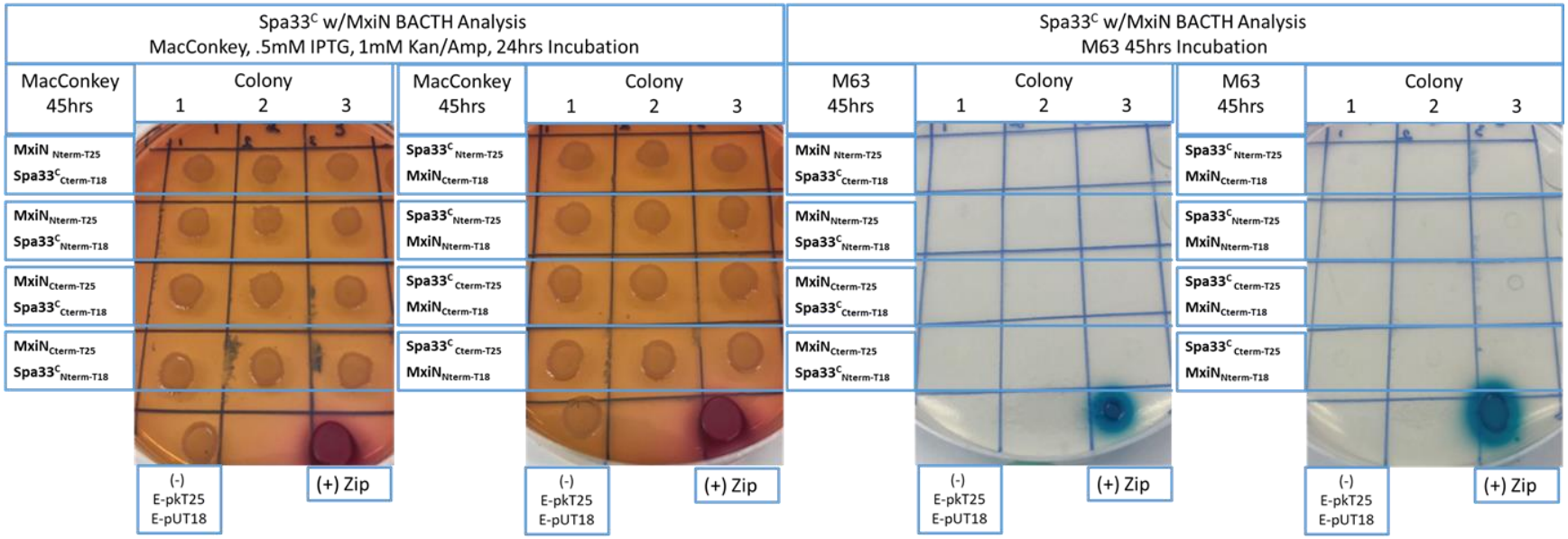


Figure 24: Spa33^{C} /MxiN Permutation BACTH Screening Plates- BACTH analysis for the screening of MxiN and Spa33^{C} protein-protein interactions. A positive protein-protein interaction is reflected in the phenotypic expression and observation of red colonies on MacConkey and blue colonies on M63 indicator 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 $Spa33^{WT}$ 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 $Spa33^{WT}{}_{N\text{-terminus}}$ (**Figure 25: K).** We were able to determine that this interaction is facilitated by the $Spa33^{FL}{}_{N\text{-terminus}}$ domain, and that $Spa33^{C}$ doesn't appear to be essential for $Spa33^{FL}$ 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 $Spa33^{FL}$ (T25 or T18 fused to its N-terminus) and $Spa33^{C}$ (T18 or T25 fused to its N-terminus). Comparisons reveal that the $Spa33^{C}$ domain only interacts with $Spa33^{FL}$ that has T25/T18 fused to its N-terminus. Therefore, I hypothesized that the N-terminal region of $Spa33^{FL}$ contains the binding domain for $Spa33^{C}$. $Spa33^{C}$ 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 $Spa33^{FL}$-$Spa33^{C}{}_{2}$ trimeric complex binding domains with the rest of the SP platform components (**Figure 26).**

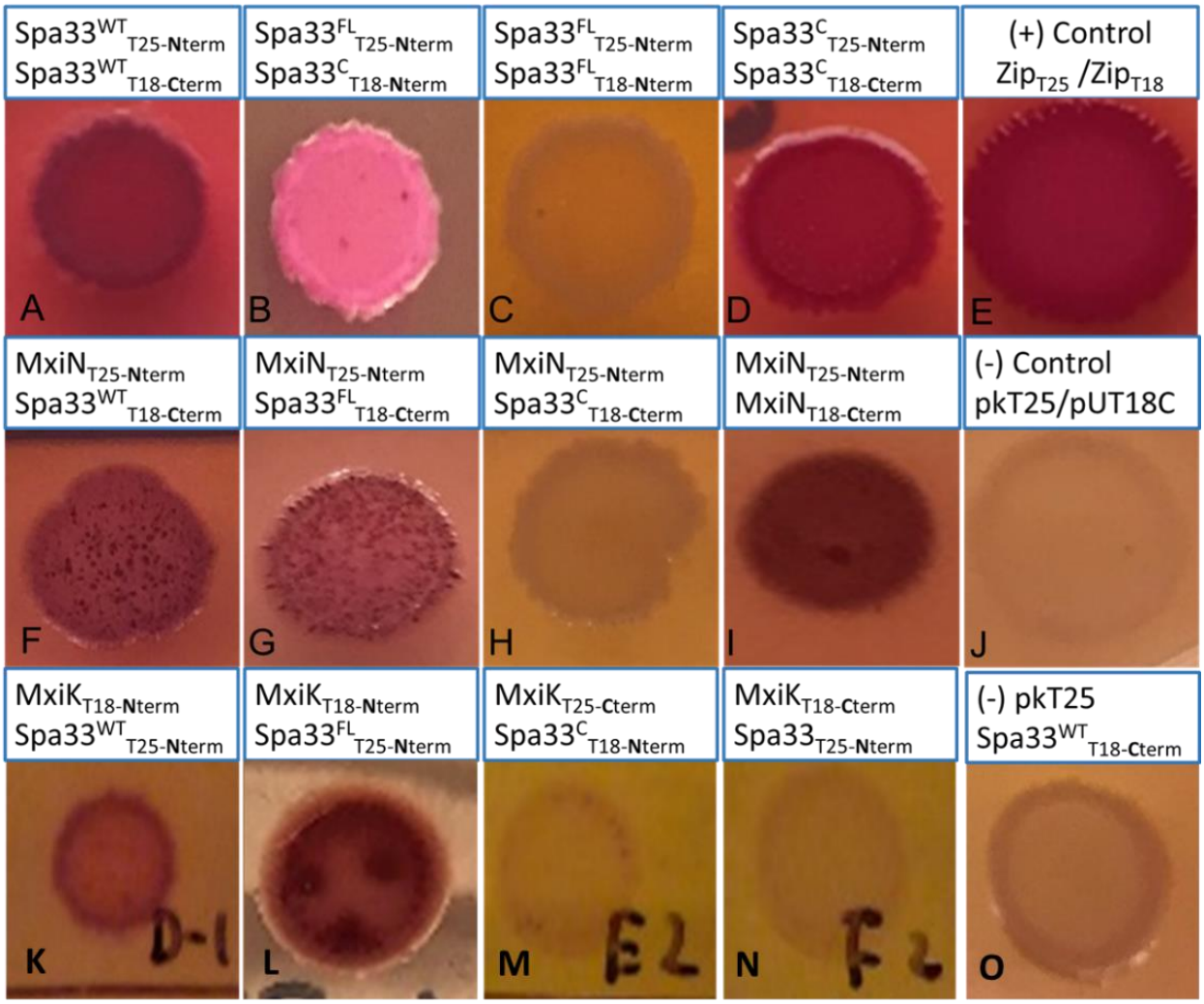


Figure 25: Summary of Key BACTH Assay Results- BACTH analysis summary for the screening of Spa33 protein-protein interactions with the other SP components. A positive protein-protein interaction is dictated by the phenotypic expression and observation of red colonies on MacConkey. A positive control is indicated by the (+) Zip protein, and the negative controls (-) are indicated accordingly.

	Protein 1	Protein 2	Interaction
A	$Spa33^{WT}$	$Spa33^{WT}$	Yes
B	$Spa33^{FL}$ N-terminus	$Spa33^{C}$ (N or C)-terminus	Yes
C	$Spa33^{FL}$	$Spa33^{FL}$	No
D	$Spa33^{C}$	$Spa33^{C}$	Yes
F	$Spa33^{WT}$	MxiN N-terminus	Yes
G	$Spa33^{FL}$ C-terminus	MxiN N-terminus	Yes
H	$Spa33^{C}$	MxiN	No
I	MxiN	MxiN	Yes
K	MxiK N-terminus	$Spa33^{WT}$ N-terminus	Yes
L	MxiK N-terminus	$Spa33^{FL}$ N-terminus	Yes
M	MxiK	$Spa33^{C}$	No

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)**.

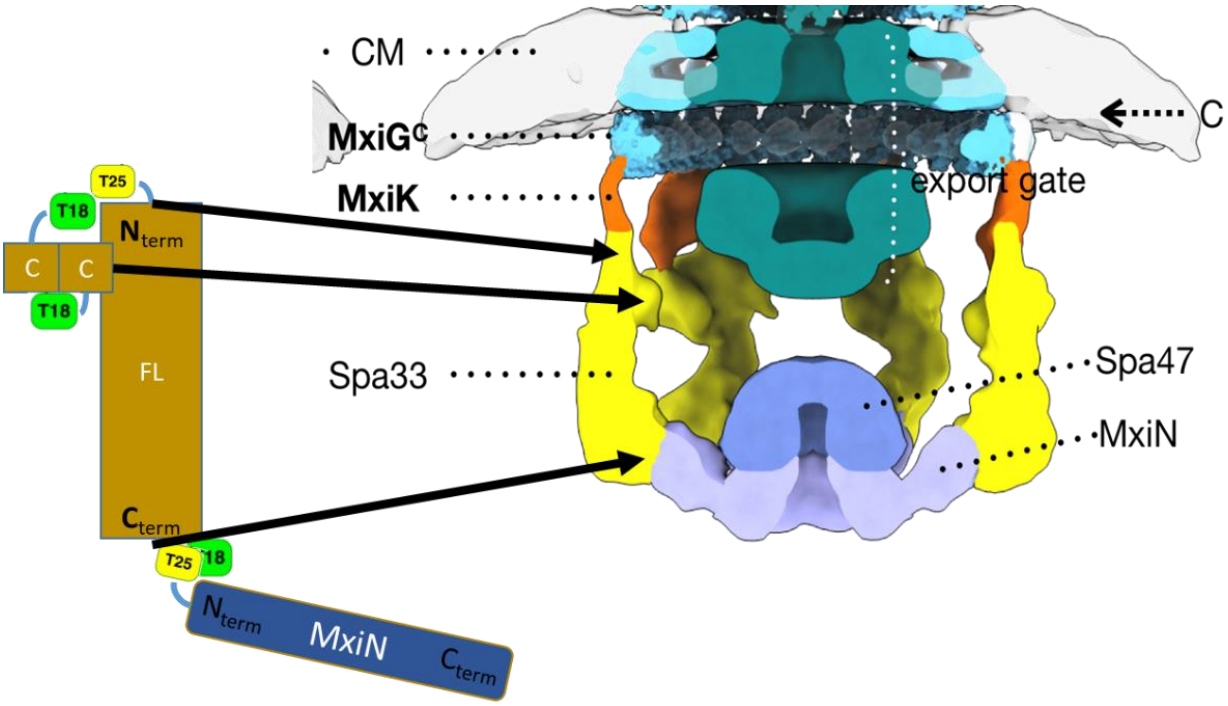


Figure 26: Estimated Localization of Each of the SP Components and their Termini (Left): The graphic depicts the hypothesized localization and orientation of the Spa33^{WT} components (Spa33^{FL} and Spa33^{C}) within the sorting platform derived from the BACTH results. **(Right):** Cryo-electron tomography 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).[61] 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 lysosyme insertion mutants would also aid in orienting the protien 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.[103] Ultimately, the crystal structure must be determined to provide complete characterization of the proteins. We have already referred to the crystal structure of the $Spa33^{C}$ (McDowell et al. 2015) dimer, but to date, attempts to solve crystal structures for the $Spa33^{WT}$ trimeric complex and $Spa33^{FL}$ 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^{WT} and Spa33^{C} are required. Spa33^{FL} is insoluble in the absence of Spa33^{C}, which has limited our ability to characterize the Spa33^{FL} subunit. Nonetheless, the secondary structure of purified Spa33^{WT} and Spa33^{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.[98] This method is thus able to provide the secondary structure of purified Spa33^{FL} and Spa33^{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^{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^{C} after thermal unfolding and then cooling suggests that Spa33^{C} is able to refold after unfolding at 90° C and subsequent cooling back to 10° C. The Spa33^{C} CD spectrum is supportive of the already solved crystal structure in which β-sheet represent the majority of the structure (**Figure 11)**.[78] Alternatively, when analyzing the CD spectrum of Spa33^{WT}, a double minimum is seen (208 and 222nm) suggesting Spa33^{WT} is more α-helical in nature. Interestingly, the Spa33^{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^{WT} α-helical

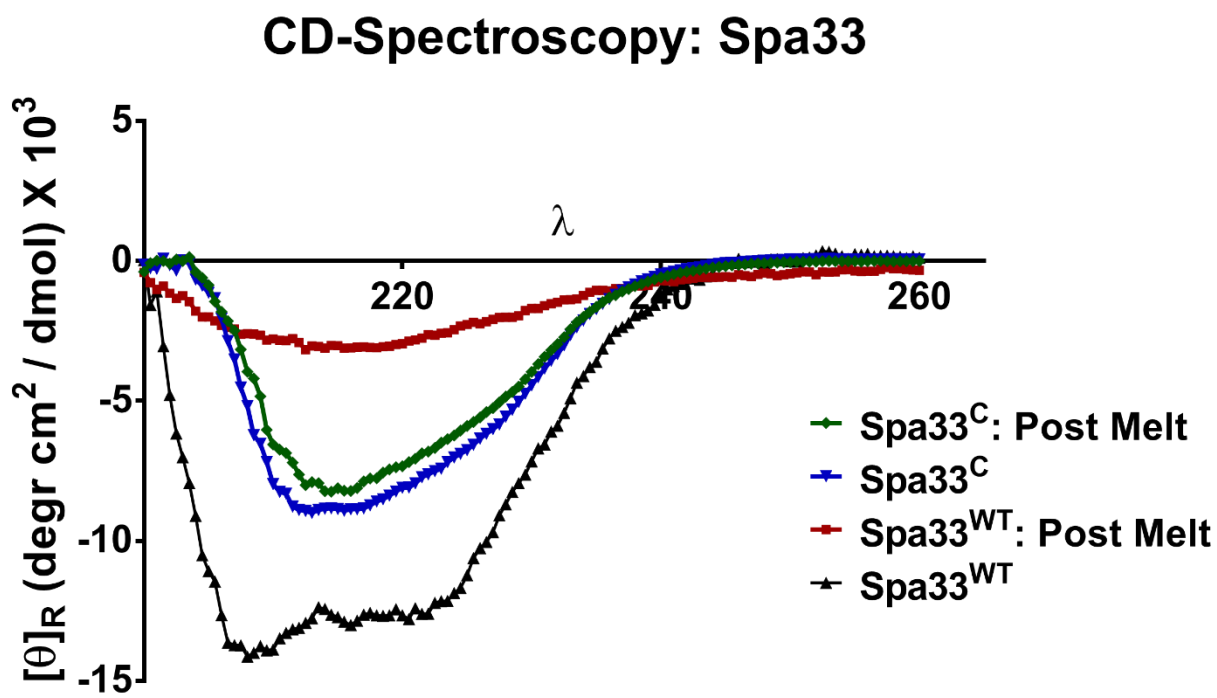


Figure 27: Circular Dichroism (CD) Spectroscopy of purified Spa33^{WT} and Spa33^{C}- The Spa33^{C} (blue) spectrum, with a minimum near 215nm, suggests β-sheet 2° structure is present and this is consistent with its crystal structure. Spa33^{C} post melt spectrum (green) shows that the Spa33^{C} is very stable and is probably able to refold after melting and cooling back to 10 °C. The spectrum of Spa33^{WT} (black), contains minima at 208nm and 220nm, which indicates a strong α-helical 2° structure presence. Spa33^{WT} post melt spectrum (red), reveals that a β-sheet 2° structure, which could indicated that after reaching melting conditions Spa33^{FL} becomes denatured or aggregated while the Spa33^{C} dimer is able to reobtain its 2° structure qualities. We thus hypothesize that the uncharacterized Spa33^{FL} subunit has a mixed α/ β structure. Crystal Structure: The structure of Spa33^{C} was solved by McDowell et al.

properties, measured before melting, could be entirely accounted by the $Spa33^{FL}$ 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 $Spa33^{C}$ to the $Spa33^{WT}$ complex ellipticity would be overshadowed if $Spa33^{FL}$ has an N-terminus rich in α-helical structure. Consequently, after $Spa33^{WT}$ is exposed to unfolding temperatures, $Spa33^{FL}$ will disassociate from the $Spa33^{C}$ dimer and destabilize its secondary structure. However, because of the insoluble nature of $Spa33^{FL}$ and its inability to refold its secondary structure during subsequent cooling to 10° C, the spectrum of the disassociated $Spa33^{C}$ dimer is now detectable because $Spa33^{FL}$ is no longer able to contribute to the observed ellipticity. I thus hypothesize that the uncharacterized $Spa33^{FL}$ subunit, particularly its N-terminus, contains an elevated α-helical secondary structure compared to $Spa33^{C}$. A CD-spectrum overlay for all of these conditions is provided in **Figure 27**.

Functional Assessment *in vivo* of $Spa33^{WT}$ Components

The behavior of $Spa33^{WT}$ (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 $Spa33^{C}$ plays in maintaining $Spa33^{FL}$ 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 $Spa33^{FL}$ and $Spa33^{C}$ ($SctQ^{FL/C}$) subunits was proposed to be essential for SP assembly and T3SS function in *Shigella*. Alternatively, *Salmonella* secretion and invasion studies determined that $SpaO^{S}$ ($SctQ^{C}$) is a non-essential component for T3SS function, but is required for the stability of $SpaO^{FL}$ ($SctQ^{FL}$).[94] *In*

vitro evidence also suggests that these two protein subunits also associate to form a heterotrimer with a 1:2 stoichiometric complex of $SctQ^{FL}$-$ScctQ^{C}{}_{2}$ ($SctQ^{WT}$).[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 $\Delta Spa33^{FL}$ or $\Delta Spa33^{C}$ *S. flexneri* mutants supporting their hypothesis that both components are essential to T3SS activity. Contrary to this, Terjo et al. found that $\Delta Spao^{C}$ *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

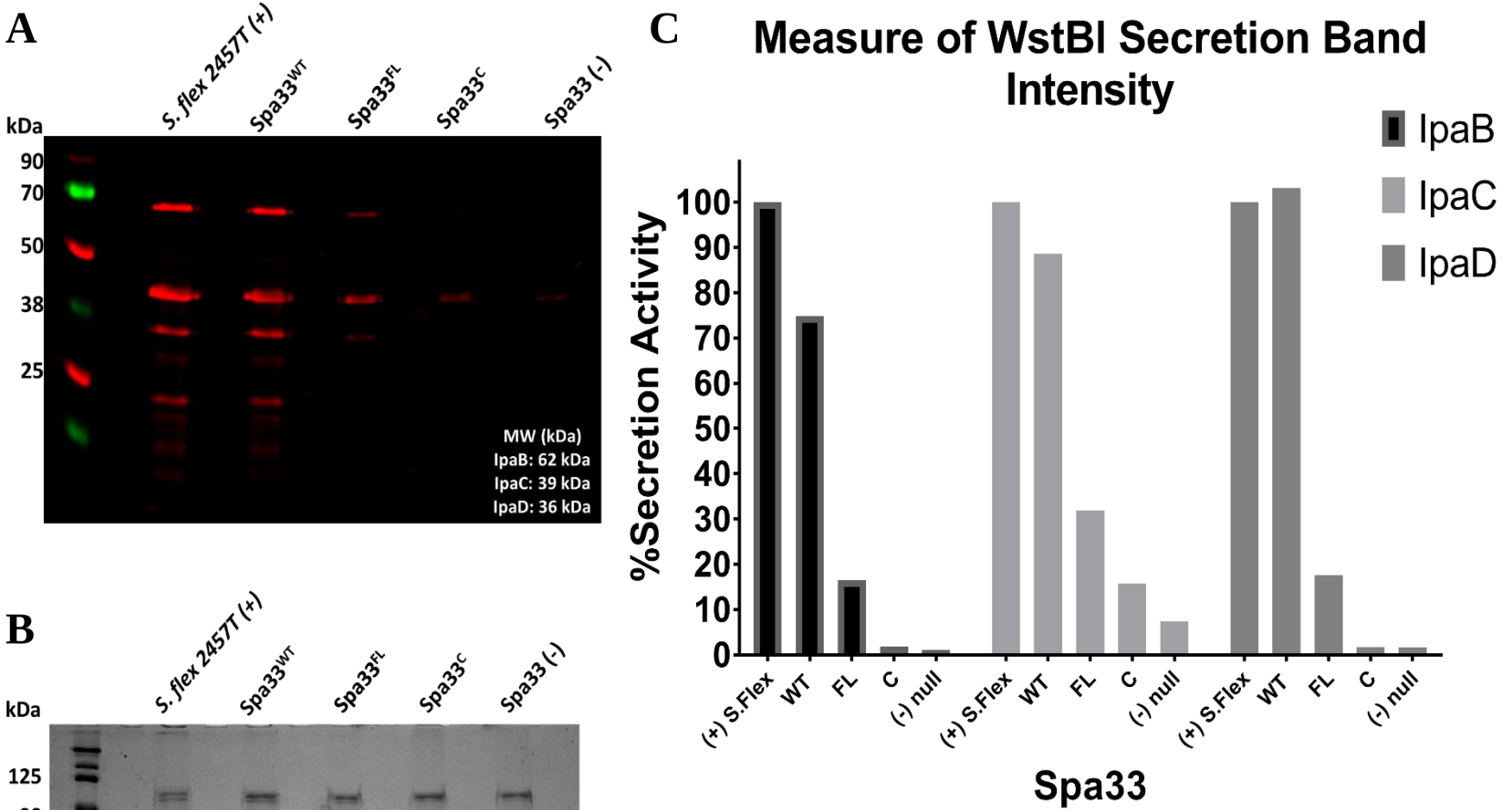


Figure 28: Overnight Secretion of *S. flexneri* Complemented Spa33 Mutants- A. Western Blot analysis of IpaB, IpaC, and IpaD *S. flexneri* overnight secretion in the presence or absence of $Spa33^{FL}$ or $Spa33^{C}$. Lanes (Left-Right) *S. flex.* 2457T (+), *S.flex*. null bacteria complimented with $Spa33^{WT}$, $Spa33^{FL}$, $Spa33^{C}$, 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 $Spa33^{C}$ and Spa33(-) null bacteria. In other words, the absence of $Spa33^{C}$ may not completely prevent T3SS activity in *Shigella* as McDowell et al. had proposed.While the absence of $Spa33^{C}$ may not eliminate all T3SS activity, its reduction was profound compared to wild type *S. flexneri* 2457T and the $Spa33^{WT}$ 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 $Spa33^{WT}$, $Spa33^{FL}$, and $Spa33^{C}$ 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 uses 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

in vivo measure of T3SS activity compared to secretion assays because RBC hemolysis is dependent on the formation and insertion of a functional translocon pore.[104,105] Measurement 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

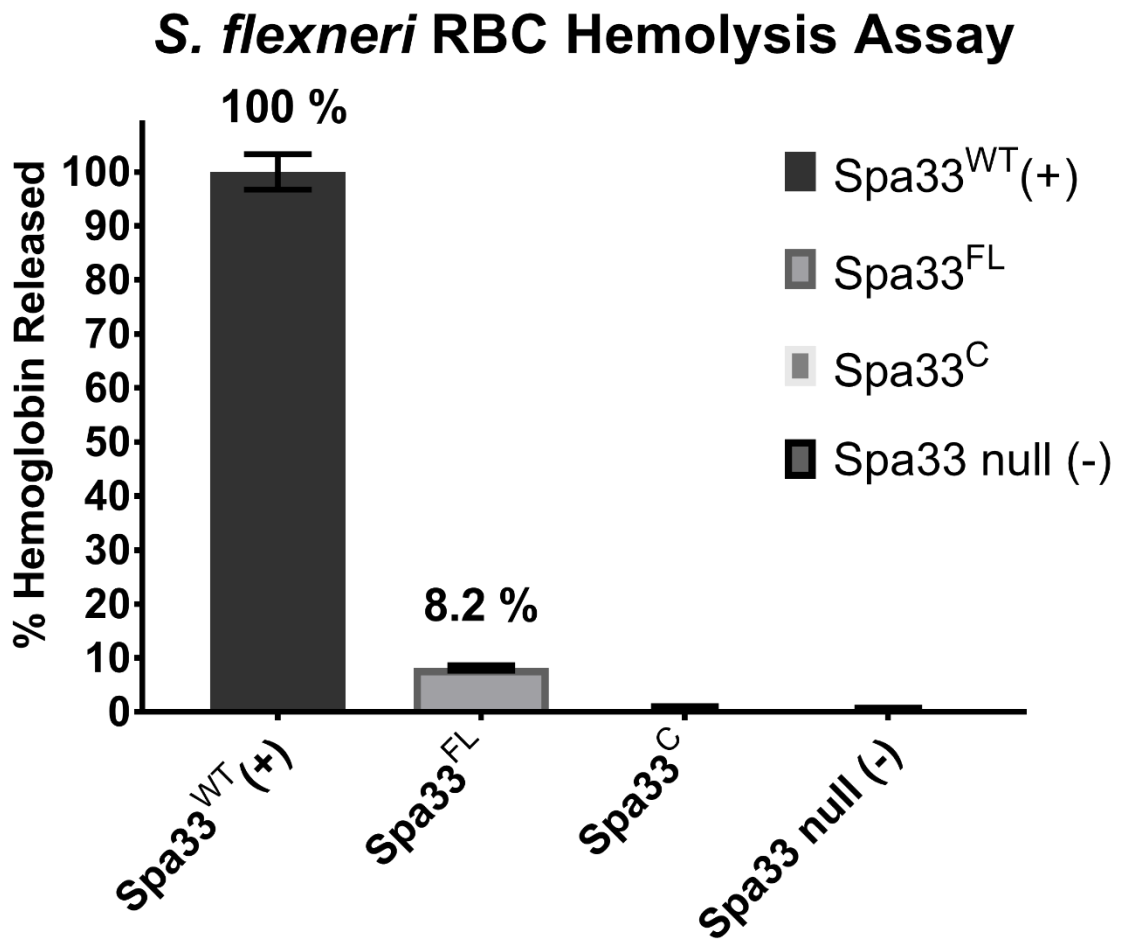


Figure 29: *S. flexneri* RBC Hemolysis Assay- Hemolytic assay measuring the release of hemoglobin from RBCs (Abs 545nm), post centrifugation and incubation (37°C) with *S. flexneri spa33* null bacterium complimented with Spa33 WT, FL and C plasmids. The ability of the complimented plasmids to restore 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 $Spa33^{WT}$.

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 $Spa33^{FL}$-$Spa33^{C}{}_{2}$ 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 these 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

1. Petri, W. A. *et al.* Enteric infections, diarrhea, and their impact on function and development. *Journal of Clinical Investigation* **118**, 1277–1290 (2008).
2. Diarrhoeal disease. Available at: https://www.who.int/news-room/fact-sheets/detail/diarrhoeal-disease. (Accessed: 29th June 2019)
3. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study | Elsevier Enhanced Reader. doi:10.1016/S0140-6736(13)60844-2
4. *Hugo and Russell's pharmaceutical microbiology*. (Blackwell Science, 2004).
5. Pendleton, J. N., Gorman, S. P. & Gilmore, B. F. Clinical relevance of the ESKAPE pathogens. *Expert Review of Anti-Infective Therapy; London* **11**, 297–308 (2013).
6. Rosenblatt-Farrell, N. The Landscape of Antibiotic Resistance. *Environ Health Perspect* **117**, A244–A250 (2009).
7. Talbot, G. H. *et al.* The Infectious Diseases Society of America's 10 × '20 Initiative (10 New Systemic Antibacterial Agents US Food and Drug Administration Approved by 2020): Is 20 × '20 a Possibility? *Clinical Infectious Diseases* **69**, 1–11 (2019).
8. Nüesch-Inderbinen, M. *et al.* Shigella Antimicrobial Drug Resistance Mechanisms, 2004-2014. *Emerging Infect. Dis.* **22**, 1083–1085 (2016).
9. Boucher, H. W. *et al.* Bad Bugs, No Drugs: No ESKAPE! An Update from the Infectious Diseases Society of America. *Clinical Infectious Diseases* **48**, 1–12 (2009).
10. Karlowsky, J. A., Hoban, D. J., Hackel, M. A., Lob, S. H. & Sahm, D. F. Antimicrobial susceptibility of Gram-negative ESKAPE pathogens isolated from hospitalized patients with intra-abdominal and urinary tract infections in Asia–Pacific countries: SMART 2013–2015. *Journal of Medical Microbiology* **66**, 61–69 (2017).

11. Rice, L. B. Federal Funding for the Study of Antimicrobial Resistance in Nosocomial Pathogens: No ESKAPE. *The Journal of Infectious Diseases* **197**, 1079–1081 (2008).
12. Agaisse, H. Shigella flexneri serotype 3a: the rise of a superbug. *Lancet Infect Dis* **15**, 867–868 (2015).
13. Poirel, L. *et al.* Antimicrobial Resistance in Escherichia coli. *Microbiology Spectrum* **6**, (2018).
14. Karkey, A., Thwaites, G. & Baker, S. The evolution of antimicrobial resistance in Salmonella Typhi. *Current Opinion in Gastroenterology* **34**, 25–30 (2018).
15. Morales-López, S., Yepes, J. A., Prada-Herrera, J. C. & Torres-Jiménez, A. Enterobacteria in the 21st century: a review focused on taxonomic changes. *1* **13**, 265–273 (2019).
16. Guerrant, R. L., Kosek, M., Lima, A. A. M., Lorntz, B. & Guyatt, H. L. Updating the DALYs for diarrhoeal disease. *Trends in Parasitology* **18**, 191–193 (2002).
17. WHO | Guidelines for the control of shigellosis, including epidemics due to Shigella dysenteriae type 1. *WHO* Available at: http://www.who.int/cholera/publications/shigellosis/en/. (Accessed: 25th June 2018)
18. Khalil, I. A. *et al.* Morbidity and mortality due to shigella and enterotoxigenic Escherichia coli diarrhoea: the Global Burden of Disease Study 1990–2016. *The Lancet Infectious Diseases* **18**, 1229–1240 (2018).
19. Hale, T. L. & Keusch, G. T. Shigella. in *Medical Microbiology* (ed. Baron, S.) (University of Texas Medical Branch at Galveston, 1996).
20. Admoni, O. *et al.* Epidemiological, clinical and microbiological features of shigellosis among hospitalized children in northern Israel. *Scand. J. Infect. Dis.* **27**, 139–144 (1995).

21. Sivapalasingam, S. *et al.* High Prevalence of Antimicrobial Resistance among Shigella Isolates in the United States Tested by the National Antimicrobial Resistance Monitoring System from 1999 to 2002. *Antimicrob Agents Chemother* **50**, 49–54 (2006).

22. Opintan, J. & Newman, M. J. Distribution of Serogroups and Serotypes of Multiple Drug Resistant Shigella Isolates. *Ghana Med J* **41**, 8–29 (2007).

23. Schroeder, G. N. & Hilbi, H. Molecular Pathogenesis of Shigella spp.: Controlling Host Cell Signaling, Invasion, and Death by Type III Secretion. *Clinical Microbiology Reviews* **21**, 134–156 (2008).

24. Kozak, G. K., MacDonald, D., Landry, L. & Farber, J. M. Foodborne outbreaks in Canada linked to produce: 2001 through 2009. *J. Food Prot.* **76**, 173–183 (2013).

25. Lerouge, I. & Vanderleyden, J. O-antigen structural variation: mechanisms and possible roles in animal/plant–microbe interactions. *FEMS Microbiology Reviews* **26**, 17–47 (2002).

26. Cui, X. *et al.* Antimicrobial Resistance of Shigella flexneri Serotype 1b Isolates in China. *PLOS ONE* **10**, e0129009 (2015).

27. Jennison, A. V. & Verma, N. K. The acid-resistance pathways of Shigella flexneri 2457T. *Microbiology* **153**, 2593–2602 (2007).

28. Lin, J., Lee, I. S., Frey, J., Slonczewski, J. L. & Foster, J. W. Comparative analysis of extreme acid survival in Salmonella typhimurium, Shigella flexneri, and Escherichia coli. *J Bacteriol* **177**, 4097–4104 (1995).

29. Bhagwat, A. A. & Bhagwat, M. Comparative analysis of transcriptional regulatory elements of glutamate-dependent acid-resistance systems of *Shigella flexneri* and *Escherichia coli* O157:H7. *FEMS Microbiology Letters* **234**, 139–147 (2004).

30. Zhao, B. & Houry, W. A. Acid stress response in enteropathogenic gammaproteobacteria: an aptitude for survivalThis paper is one of a selection of papers published in this special issue entitled “Canadian Society of Biochemistry, Molecular & Cellular Biology 52nd Annual Meeting — Protein Folding: Principles and Diseases” and has undergone the Journal’s usual peer review process. *Biochemistry and Cell Biology* **88**, 301–314 (2010).
31. Bearson, S., Bearson, B. & Foster, J. W. Acid stress responses in enterobacteria. *FEMS Microbiology Letters* **147**, 173–180 (2006).
32. Sperandio, B. *et al.* Virulent Shigella flexneri Affects Secretion, Expression, and Glycosylation of Gel-Forming Mucins in Mucus-Producing Cells. *Infection and Immunity* **81**, 3632–3643 (2013).
33. Ranganathan, S. *et al.* Evaluating *Shigella flexneri* Pathogenesis in the Human Enteroid Model. *Infection and Immunity* **87**, (2019).
34. In, J. *et al.* Enterohemorrhagic Escherichia coli reduce mucus and intermicrovillar bridges in human stem cell-derived colonoids. *Cell Mol Gastroenterol Hepatol* **2**, 48-62.e3 (2016).
35. Wilson, S. S. *et al.* Alpha-defensin-dependent enhancement of enteric viral infection. *PLOS Pathogens* **13**, e1006446 (2017).
36. Gudmundsson, G. H., Bergman, P., Andersson, J., Raqib, R. & Agerberth, B. Battle and balance at mucosal surfaces – The story of Shigella and antimicrobial peptides. *Biochemical and Biophysical Research Communications* **396**, 116–119 (2010).
37. WASSEF, J. S., KEREN, D. F. & MAILLOUX, J. L. Role of M Cells in Initial Antigen Uptake and in Ulcer Formation in the Rabbit Intestinal Loop Model of Shigellosis. *INFECT. IMMUN.* **57**, 6 (1989).

38. Tyrer, P. C., Ruth Foxwell, A., Kyd, J. M., Otczyk, D. C. & Cripps, A. W. Receptor mediated targeting of M-cells. *Vaccine* **25**, 3204–3209 (2007).

39. Elhelu, M. A. The Role of Macrophages in Immunology. *J Natl Med Assoc* **75**, 314–317 (1983).

40. Kuwae, A. *et al.* Shigella Invasion of Macrophage Requires the Insertion of IpaC into the Host Plasma Membrane FUNCTIONAL ANALYSIS OF IpaC. *J. Biol. Chem.* **276**, 32230–32239 (2001).

41. Ashida, H., Mimuro, H. & Sasakawa, C. Shigella Manipulates Host Immune Responses by Delivering Effector Proteins with Specific Roles. *Front Immunol* **6**, (2015).

42. Yang, S.-C., Hung, C.-F., Aljuffali, I. A. & Fang, J.-Y. The roles of the virulence factor IpaB in Shigella spp. in the escape from immune cells and invasion of epithelial cells. *Microbiological Research* **181**, 43–51 (2015).

43. Arizmendi, O., Picking, W. D. & Picking, W. L. Macrophage Apoptosis Triggered by IpaD from Shigella flexneri. *Infection and Immunity* **84**, 1857–1865 (2016).

44. Ashida, H. *et al.* Shigella deploy multiple countermeasures against host innate immune responses. *Current Opinion in Microbiology* **14**, 16–23 (2011).

45. Senerovic, L. *et al.* Spontaneous formation of IpaB ion channels in host cell membranes reveals how Shigella induces pyroptosis in macrophages. *Cell Death Dis* **3**, e384 (2012).

46. Slee, E. A. *et al.* Ordering the cytochrome c-initiated caspase cascade: hierarchical activation of caspases-2, -3, -6, -7, -8, and -10 in a caspase-9-dependent manner. *J. Cell Biol.* **144**, 281–292 (1999).

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

intercellular spread through interaction with F-actin. *Proc Natl Acad Sci U S A* **86**, 3867–3871 (1989).

48. Lett, M. C. *et al.* virG, a plasmid-coded virulence gene of Shigella flexneri: identification of the virG protein and determination of the complete coding sequence. *J Bacteriol* **171**, 353–359 (1989).
49. Goldberg, M. B. Actin-Based Motility of Intracellular Microbial Pathogens. *Microbiol Mol Biol Rev* **65**, 595–626 (2001).
50. Suzuki, T., Lett, M.-C. & Sasakawa, C. Extracellular Transport of VirG Protein in Shigella. *J. Biol. Chem.* **270**, 30874–30880 (1995).
51. Suzuki, T., Saga, S. & Sasakawa, C. Functional Analysis of Shigella VirG Domains Essential for Interaction with Vinculin and Actin-based Motility. *J. Biol. Chem.* **271**, 21878–21885 (1996).
52. Suzuki, T. & Sasakawa, C. Molecular Basis of the Intracellular Spreading of Shigella. *Infect Immun* **69**, 5959–5966 (2001).
53. Torres, A. G. Current aspects of Shigella pathogenesis. *Rev Latinoam Microbiol* 10
54. Mattock, E. & Blocker, A. J. How Do the Virulence Factors of Shigella Work Together to Cause Disease? *Front Cell Infect Microbiol* **7**, (2017).
55. Belotserkovsky, I. & Sansonetti, P. J. Shigella and Enteroinvasive Escherichia Coli. in *Escherichia coli, a Versatile Pathogen* (eds. Frankel, G. & Ron, E. Z.) 1–26 (Springer International Publishing, 2018). doi:10.1007/82_2018_104
56. Killackey, S. A., Sorbara, M. T. & Girardin, S. E. Cellular Aspects of Shigella Pathogenesis: Focus on the Manipulation of Host Cell Processes. *Front Cell Infect Microbiol* **6**, (2016).

57. Green, E. R. & Mecsas, J. Bacterial Secretion Systems – An overview. *Microbiol Spectr* **4**, (2016).

58. Yang, F. Genome dynamics and diversity of Shigella species, the etiologic agents of bacillary dysentery. *Nucleic Acids Research* **33**, 6445–6458 (2005).

59. Zhu, L. *et al.* Global analysis of a plasmid-cured Shigella flexneri strain: new insights into the interaction between the chromosome and a virulence plasmid. *J. Proteome Res.* **9**, 843–854 (2010).

60. Coburn, B., Sekirov, I. & Finlay, B. B. Type III Secretion Systems and Disease. *Clin Microbiol Rev* **20**, 535–549 (2007).

61. Hu, B. *et al.* Visualization of the type III secretion sorting platform of *Shigella flexneri*. *Proceedings of the National Academy of Sciences* **112**, 1047–1052 (2015).

62. Kubori, T. Supramolecular Structure of the Salmonella typhimurium Type III Protein Secretion System. *Science* **280**, 602–605 (1998).

63. Diepold, A. & Armitage, J. P. Type III secretion systems: the bacterial flagellum and the injectisome. *Philos Trans R Soc Lond B Biol Sci* **370**, (2015).

64. Portaliou, A. G. *et al.* Hierarchical protein targeting and secretion is controlled by an affinity switch in the type III secretion system of enteropathogenic Escherichia coli. *EMBO J* **36**, 3517–3531 (2017).

65. Cordes, F. S. *et al.* Helical Structure of the Needle of the Type III Secretion System of Shigella flexneri. *J. Biol. Chem.* **278**, 17103–17107 (2003).

66. Demers, J.-P. *et al.* High-resolution structure of the Shigella type-III secretion needle by solid-state NMR and cryo-electron microscopy. *Nature Communications* **5**, (2014).

67. Veenendaal, A. K. J. *et al.* The type III secretion system needle tip complex mediates host cell sensing and translocon insertion. *Molecular Microbiology* **63**, 1719–1730 (2007).

68. Ménard, R., Sansonetti, P. J. & Parsot, C. Nonpolar mutagenesis of the ipa genes defines IpaB, IpaC, and IpaD as effectors of Shigella flexneri entry into epithelial cells. *J. Bacteriol.* **175**, 5899–5906 (1993).

69. Picking, W. L. *et al.* IpaD of Shigella flexneri Is Independently Required for Regulation of Ipa Protein Secretion and Efficient Insertion of IpaB and IpaC into Host Membranes. *Infect Immun* **73**, 1432–1440 (2005).

70. Espina, M. *et al.* IpaD Localizes to the Tip of the Type III Secretion System Needle of Shigella flexneri. *Infection and Immunity* **74**, 4391–4400 (2006).

71. Johnson, S. *et al.* Expression, limited proteolysis and preliminary crystallographic analysis of IpaD, a component of the *Shigella flexneri* type III secretion system. *Acta Crystallographica Section F Structural Biology and Crystallization Communications* **62**, 865–868 (2006).

72. Blocker, A. J. *et al.* What's the point of the type III secretion system needle? *PNAS* **105**, 6507–6513 (2008).

73. Murillo, I., Martinez-Argudo, I. & Blocker, A. J. Genetic Dissection of the Signaling Cascade that Controls Activation of the Shigella Type III Secretion System from the Needle Tip. *Sci Rep* **6**, (2016).

74. Dong, N. *et al.* Structurally distinct bacterial TBC-like GAPs link Arf GTPase to Rab1 inactivation to counteract host defenses. *Cell* **150**, 1029–1041 (2012).

75. Burnaevskiy, N., Peng, T., Reddick, L. E., Hang, H. C. & Alto, N. M. Myristoylome profiling reveals a concerted mechanism of ARF GTPase deacylation by the bacterial protease IpaJ. *Mol. Cell* **58**, 110–122 (2015).

76. Cornelis, G. R. The type III secretion injectisome. *Nature Reviews Microbiology* **4**, 811–825 (2006).

77. Portaliou, A. G., Tsolis, K. C., Loos, M. S., Zorzini, V. & Economou, A. Type III Secretion: Building and Operating a Remarkable Nanomachine. *Trends in Biochemical Sciences* **41**, 175–189 (2016).

78. McDowell, M. A. *et al.* Characterisation of *Shigella* Spa33 and *Thermotoga* FliM/N reveals a new model for C-ring assembly in T3SS: Uniform C-ring assembly by NF- and flagellar-T3SS. *Molecular Microbiology* **99**, 749–766 (2016).

79. Diepold, A. *et al.* Deciphering the assembly of the Yersinia type III secretion injectisome. *EMBO J* **29**, 1928–1940 (2010).

80. Kimbrough, T. G. & Miller, S. I. Contribution of Salmonella typhimurium type III secretion components to needle complex formation. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11008–11013 (2000).

81. Diepold, A., Wiesand, U., Amstutz, M. & Cornelis, G. R. Assembly of the Yersinia injectisome: the missing pieces. *Molecular Microbiology* **85**, 878–892 (2012).

82. Huang, F. *et al.* Ultra-High Resolution 3D Imaging of Whole Cells. *Cell* **166**, 1028–1040 (2016).

83. Zhang, Y., Lara-Tejero, M., Bewersdorf, J. & Galán, J. E. Visualization and characterization of individual type III protein secretion machines in live bacteria. *PNAS* **114**, 6098–6103 (2017).

84. Cryo-Electron Tomography - an overview | ScienceDirect Topics. Available at: https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/cryo-electron-tomography. (Accessed: 16th July 2019)

85. Diepold, A. *et al.* A dynamic and adaptive network of cytosolic interactions governs protein export by the T3SS injectisome. *Nat Commun* **8**, (2017).

86. Diepold, A., Kudryashev, M., Delalez, N. J., Berry, R. M. & Armitage, J. P. Composition, Formation, and Regulation of the Cytosolic C-ring, a Dynamic Component of the Type III Secretion Injectisome. *PLOS Biology* **13**, e1002039 (2015).

87. Multiple Sequence Alignment - CLUSTALW. Available at: https://www.genome.jp/tools-bin/clustalw. (Accessed: 6th July 2019)

88. Okonechnikov, K., Golosova, O. & Fursov, M. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* **28**, 1166–1167 (2012).

89. Park, D. *et al.* Visualization of the type III secretion mediated Salmonella–host cell interface using cryo-electron tomography. *eLife* **7**, e39514 (2018).

90. Morita-Ishihara, T. *et al. Shigella* Spa33 Is an Essential C-ring Component of Type III Secretion Machinery. *Journal of Biological Chemistry* **281**, 599–607 (2006).

91. Kearse, M. G. & Wilusz, J. E. Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes & Development* **31**, 1717–1731 (2017).

92. Yu, X.-J., Liu, M. R., Matthews, S. J. & Holden, D. W. Tandem Translation Generates a Chaperone for the Salmonella Type III Secretion System Protein SsaQ*. in *The Journal of biological chemistry* (2011). doi:10.1074/jbc.M111.278663

93. Bzymek, K. P., Hamaoka, B. Y. & Ghosh, P. Two translation products of Yersinia yscQ assemble to form a complex essential to type III secretion. *Biochemistry* **51**, 1669–1677 (2012).

94. Lara-Tejero, M. *et al.* Role of SpaO in the assembly of the sorting platform of a Salmonella type III secretion system. *PLOS Pathogens* **15**, e1007565 (2019).

95. BioTroll Tools: Cloning Troll a Hassle-Free PCR Primer design tool. SeqTroll: Fast and easy to use sequencing data analysis tool. OptiTroll: Analyses and optimizes the codon usage improving the translation efficiency. Available at: http://biotroll.com/. (Accessed: 19th July 2019)

96. Battesti, A. & Bouveret, E. The bacterial two-hybrid system based on adenylate cyclase reconstitution in Escherichia coli. *Methods* **58**, 325–334 (2012).

97. Euromedex, life science products. Available at: https://www.euromedex.com/reagentsShop.xhtml. (Accessed: 11th July 2019)

98. Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1751**, 119–139 (2005).

99. CAPITO - CD Analysing and Plotting Tool. Available at: https://capito.nmr.leibniz-fli.de//index.php. (Accessed: 11th July 2019)

100. Ladant, D. & Ullmann, A. Bordetella pertussis adenylate cyclase: a toxin with multiple talents. **7**, 5 (1999).

101. Guo, Q. *et al.* Structural basis for the interaction of Bordetella pertussis adenylyl cyclase toxin with calmodulin. *The EMBO Journal* **24**, 3190–3201 (2005).

102. LadantS, D., Michelsong, S., PredeleanuS, R. & BarzuS, O. Characterization of the Calmodulin-binding andof the Catalytic Domains of Bordetella pertussisAdenylate Cyclase. 6

103. Koning, R. I. Chapter 24 - Cryo-Electron Tomography of Cellular Microtubules. in *Methods in Cell Biology* (eds. Cassimeris, L. & Tran, P.) **97**, 455–473 (Academic Press, 2010).

104. Bahrani, F. K., Sansonetti, P. J. & Parsot, C. Secretion of Ipa proteins by Shigella flexneri: inducer molecules and kinetics of activation. *Infect Immun* **65**, 4005–4010 (1997).

105. Pope, L. M., Reed, K. E. & Payne, S. M. Increased protein secretion and adherence to HeLa cells by Shigella spp. following growth in the presence of bile salts. *Infect. Immun.* **63**, 3642–3648 (1995).

106. Roerich-Doenitz, A. D. Regulation of Type III Secretion Hierachy in Shigella flexneri. (Ph. D. thesis. University of Bristol, 2013).

107. Buchrieser, C. *et al.* The virulence plasmid pWR100 and the repertoire of proteins secreted by the type III secretion apparatus of Shigella flexneri. *Mol. Microbiol.* **38**, 760–771 (2000).

108. Kudryashev, M. *et al.* In situ structural analysis of the Yersinia enterocolitica injectisome. *eLife* **2**, (2013).

109. Lara-Tejero, M. & Galán, J. E. The Injectisome, a Complex Nanomachine for Protein Injection into Mammalian Cells. *EcoSal Plus* **8**, (2019).

110. Abrusci, P. *et al.* Architecture of the major component of the type III secretion system export apparatus. *Nat Struct Mol Biol* **20**, 99–104 (2013).

111. Erhardt, M., Namba, K. & Hughes, K. T. Bacterial Nanomachines: The Flagellum and Type III Injectisome. *Cold Spring Harb Perspect Biol* **2**, (2010).

112. Marlovits, T. C. *et al.* Assembly of the inner rod determines needle length in the type III secretion injectisome. *Nature* **441**, 637–640 (2006).

113. Claret, L., Calder, S. R., Higgins, M. & Hughes, C. Oligomerization and activation of the FliI ATPase central to bacterial flagellum assembly. *Mol Microbiol* **48**, 1349–1355 (2003).

114. ESPript 3. *ESPript 3* Available at: http://espript.ibcp.fr/ESPript/ESPript/. (Accessed: 18th July 2019)