2016

Structural and Functional Characterization of SPOA Domains in Salmonella Typhimurium Type III Secretion Systems

Ryan Q. Notti

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF SPOA DOMAINS IN *SALMONELLA TYPHIMURIUM* TYPE III SECRETION SYSTEMS

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Ryan Q. Notti

June 2016

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF SPOA DOMAINS IN *SALMONELLA TYPHIMURIUM* TYPE III SECRETION SYSTEMS

Ryan Q. Notti, Ph.D.
The Rockefeller University 2016

Once optimistically believed to be a relic of the pre-antibiotic era, bacterial pathogens remain a substantial threat to human health, and the growing epidemic of antibiotic resistance has raised concerns for the long term prospects of antimicrobial therapy. By understanding the mechanisms used by bacteria to manipulate their host and cause disease, it is hypothesized that we might more rationally approach anti-infective therapeutic design.

Type III secretion systems (T3SS) are employed by some gram-negative human pathogens to manipulate the host environment. One T3SS subtype, known as the “injectisome,” delivers virulence factors directly into host cells. The other T3SS subtype secretes the the polymeric flagellar filament used for motility. While both systems share related elements of a cytoplasmic “sorting platform” that facilitates the hierarchical secretion of protein substrates, the structural mechanism of its assembly remains unclear.

The work described in this thesis makes strides towards the mechanistic understanding of T3SS sorting platform assembly by applying structural, biochemical, and genetic techniques to the characterization of the SctQ/FliM/FliN protein family and their interactions with other sorting platform components. These proteins uniquely possess **S**urface **P**resentation **O**f **A**ntigens (SPOA) folds, and I will present the molecular structures of distinct homotypic and

heterotypic SPOA-SPOA interactions in the *Salmonella typhimurium* SPI-1 sorting platform protein SpaO (Chapter 2). In Chapter 3, I structurally characterize the interaction of the heterotypic SPOA complex with a regulator of the SPI-1 ATPase and demonstrate the necessity of the interaction for T3SS secretory function. Then, I will present the homologous structures from the *S. typhimurium* flagellar apparatus and compare and contrast them with their SPI-1 homologues, providing an explanation for the observed subtype specificity in sorting platform assembly (Chapter 4). In Chapter 5, biochemical evidence for an interaction of the SpaO amino-terminal domain(s) with the homotypic SPOA complex is presented. These results provide a model for the subtype-specific assembly of T3SS sorting platforms and will support further mechanistic analysis and anti-virulence drug design (Chapter 6).

This document is dedicated to my family, without whose unconditional love and support this work would not have been possible.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. C. Erec Stebbins, for his advice, guidance, and support during the duration of my doctoral training. Additionally, I would like to thank Drs. Howard Hang, Seth Darst, and Carl Nathan for their enthusiast, engaged participation on my Faculty Advisory Committee, and Dr. David Thanassi for his service as my external thesis examiner.

A number of individuals must be recognized for their intellectual contributions to this work: All nuclear magnetic resonance (NMR) spectroscopy was performed in close collaboration with Dr. Shibani Bhattacharya (New York Structural Biology Center; NMR operation, spectral processing, and confirmation of my resonance assignments). Drs. Deena Oren (Rockefeller University), Alison North (Rockefeller University), and Wuxian Shi (Brookhaven National Laboratory) are thanked for their technical assistance and training. I am grateful to the members of the Laboratory of Structural Microbiology (Rockefeller University), Drs. Matthew Lefebre (Yale University) and F. Nina Papavasiliou (Rockefeller University) for thoughtful discussions. Dr. Jorge Galán (Yale University) graciously donated plasmids and *S. typhimurium* strains.

This work was funded in part by National Institutes of Health (NIH) grants to C. Erec Stebbins (AI52182 and AI093704) and research funds from the Rockefeller University. I was supported by the Hearst Foundation and by a Medical Scientist Training Program grant from the National Institute of General Medical Sciences of the NIH under award number T32GM07739 to the Weill Cornell/Rockefeller/Sloan-Kettering Tri-Institutional MD-PhD Program. X-ray

diffraction data for this study were measured at beamline X29A of the National Synchrotron Light Source; financial support comes principally from the Offices of Biological and Environmental Research and of Basic Energy Sciences of the US Department of Energy, and from the National Institute of General Medical Sciences (P41GM103473) and the National Institute of Biomedical Imaging (P30-EB-00998). Preliminary X-ray diffraction experiments performed on the Rigaku/ MSC microMax 007HF in the Rockefeller University Structural Biology Resource Center were made possible by Grant Number 1S10RR022321-01 from the National Center for Research Resources of the NIH. The NMR data was collected at the New York Structural Biology Center, made possible by a grant from NYSTAR and ORIP/NIH facility improvement grant CO6RR015495. The 900 MHz NMR spectrometers were purchased with funds from NIH grant P41GM066354, the Keck Foundation, New York State Assembly, and U.S. Dept. of Defense.

The content of this document is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

LIST OF TABLES

CHAPTER 1:

INTRODUCTION

Portions of this chapter have been accepted for publication in the forthcoming 5th edition of "Virulence Mechanisms of Bacterial Pathogens" (ASM Press).

The clinical application of antibacterial small molecules ("antibiotics") has revolutionized infectious disease medicine over the past six decades; however, the rise of antibiotic resistance has generated new challenges in the effective treatment of bacterial infections (1, 2). The United States Centers for Disease Control and Prevention estimates that 2 million Americans are infected by antibiotic resistant bacteria annually, resulting in more than 23,000 deaths (3). Hypothetically, therapeutic targeting of the machinery responsible for bacterial virulence in the human host might afford treatment strategies with a lower likelihood of evolution of resistance and decreased bystander damage to the host microflora (2, 4). To facilitate the discovery or design of such therapeutics, many have endeavored to understand the molecular basis for bacterial virulence, including the bacterial systems responsible for the delivery of toxic biomolecules from the bacterium to the host.

1.1 Type III secretion systems

Type III secretion systems (T3SS) afford gram-negative bacteria a most intimate means of altering the biology of their eukaryotic hosts — the direct delivery of

effector proteins from the bacterial cytoplasm to that of the eukaryote (5, 6). T3SS utilize a conserved set of homologous gene products to assemble the nanosyringe “injectisomes” capable of traversing the three lipid membranes, peptidoglycan layer and extracellular space that form a barrier to the direct delivery of proteins from bacterium to host. While the injectisome is architecturally similar across disparate gram-negatives, its applications are a study in diversity: T3SS are employed by both symbionts and pathogens; they target animals, plants, and protists; and they are used to manipulate a wide array of cellular activities and pathways.

T3SS have attracted intense scientific interest since the seminal work documenting their discovery was published over two decades ago (7-9). Given their role in the virulence of several human and plant pathogens (e.g. *Salmonella enterica, Shigella flexneri, Yersinia spp.,* pathogenic *Escherichia coli, Vibrio spp., Pseudomonas spp., Chlamydia spp.*), T3SS are attractive targets for the discovery or design of novel anti-infective agents and vaccine approaches that disrupt toxin delivery. Conversely, *harnessing* T3SS function might also be of human benefit: as T3SS accomplish the biophysical feat of protein transduction across multiple membranes, their re-engineering for *in vivo* delivery of therapeutic proteins or *in vitro* production of protein reagents provides exciting prospects for future biomedical application. In either case, the manipulation of T3SS for human benefit will require highly refined mechanistic models of T3SS function. Drawing on research from multiple disciplines and employing complementary techniques, such models are beginning to emerge. In particular,

the application of structural biochemical approaches to the T3SS has provided numerous insights into the assembly and function of this system.

The focus of this chapter will be on T3SS function at the structural level; I will summarize the core findings that have shaped our understanding of the structure and function of these systems and highlight recent developments in the field. In turn, I will describe the T3SS secretory apparatus, consider its engagement with secretion substrates, and discuss the post-translational regulation of secretory function. Lastly, I will close with a discussion of the future prospects for the interrogation of structure-function relationships in the T3SS, and highlight those that will be the focus of the remaining chapters of this thesis.

1.2 Architecture of a nanosyringe

The genomic islands and virulence plasmids that support T3SS encode proteins of four broad classes: the components of the secretory system itself, the effector substrates, their chaperones, and transcriptional regulators. Working in concert, these components form a complete secretory system that de-chaperones and secretes substrates in a defined hierarchy and delivers them to the host cytoplasm. The repertoire of effector proteins secreted by a given T3SS is species-specific, as is the transcriptional network regulating T3SS expression (10). A discussion of these elements is beyond the scope of this document and has been expertly reviewed elsewhere (11-20). In contrast to the diverse, species-specific catalog of effector proteins and transcriptional regulators, the nanosyringe-like secretory machinery is well-conserved across species, and

advances in our mechanistic understanding of one species' injectisome are often applicable to others.

The core secretion machinery of the T3SS comprises a homologous set of approximately two-dozen gene products. Because of the high degree of homology of some components of the system, a universal nomenclature was previously suggested to facilitate cross-species comparisons (21), and recently others in the field have endorsed this naming system (22, 23). Similarly, I will employ this nomenclature (Table 1.1) wherever possible in this chapter.

A subset of these proteins have conserved homologues in the flagellar apparatus (Table 1.1), which uses its own T3SS machine to assemble the flagellar filament (24). The flagellar apparatus employs a number of proteins for the flagellar-specific function of torque generation, and not surprisingly these components lack homologues in the injectisome. However, a set of inner membrane and cytoplasmic proteins thought to be involved in the targeting and secretion of substrates is conserved between the two systems. Given that these T3SS subtypes appear to have diverged from a common ancestor some hundreds of millions of years ago (25), their conservation is noteworthy. While the focus of this chapter and the majority of this thesis is the injectisome T3SS subtype, I will draw on the flagellar literature where it offers insights into injectisome function, and a subset of these flagellar homologues will be revisited in depth in Chapter 4.

Table 1.1: A unified nomenclature for the homologous core components of the T3SS. Based on the nomenclature proposed by Hueck (21), with modifications and additions from (20, 22, 23, 26).

Notes	Universal nomencla-ture	*Salmonella SPI-1*	*Shigella*	*EPEC*	*Yersinia spp.*	*Pseudomonas aeruginosa*	*Flagellar Appparatus*
Basal Body	**SctC**	InvG	MxiD	EscC	YscC	PscC	
	SctD	PrgH	MxiG	EscD	YscD	PscD	
	SctJ	PrgK	MxiJ	EscJ	YscJ	PscJ	
Pilotin		InvH	MxiM		YscW	ExsB	
Inner Rod	**SctI**	PrgJ	MxiI	EscI	YscI	PscI	
Needle Fliament	**SctF**	PrgI	MxiH	EscF	YscF	PscF	
Needle Length Regulator	**SctP**	InvJ	Spa32	EscP (Orf16)	YscP	PscP	FliK
Inner Membrane Integral Components	**SctV**	InvA	MxiA	EscV	YscV (LcrD)	PcrD	FlhA
	SctR	SpaP	Spa24	EscR	YscR	PscR	FliP
	SctS	SpaQ	Spa9	EscS	YscS	PscS	FliQ
	SctT	SpaR	Spa29	EscT	YscT	PscT	FliR
	SctU	SpaS	Spa40	EscU	YscU	PscU	FlhB
Needle Tip and Translocon		SipB	IpaB	EspD	YopB	PopB	
		SipC	IpaC	EspB	YopD	PopD	
		SipD	IpaD	EspA	LcrV	PcrV	
ATPase	**SctN**	InvC	Spa47	EscN	YscN	PscN	FliI
Coiled Coil Linker	**SctO**	InvI	Spa13	EscO (Orf15, EscA)	YscO	PscO	FliJ
Sorting Platform	**SctQ**	SpaO	Spa33	SepQ	YscQ	PscQ	FliM/FliN
	SctK	OrgA	MxiK		YscK		
	SctL	OrgB	MxiN	EscL	YscL	PscL	FliH
Export Regulator	**SctW**	InvE	MxiC	SepL/ SepD	YopN/ TyeA	PopN	

The core conserved proteins of the T3SS form a double-membrane-spanning syringe-like structure (27, 28), including its extracellular needle-like appendage, and the associated cytoplasmic and membrane-integral secretion machinery (Figure 1.1). These components are collectively responsible for the delivery of effector proteins into the cytosol of the eukaryotic host cell (29, 30), and their structural and biochemical characterization has yielded significant insights into the processes of machine assembly and substrate secretion.

The basal body. The bacterial double membrane and peptidoglycan layer are spanned by a stack of protein annuli known as the basal body (Figure 1.1). It comprises an outer membrane-anchored layer (SctC) and an inner membrane-anchored layer (SctD and SctJ) that interface at a "neck" (31, 32). In electron microscopic (EM) reconstructions of the injectisome, SctC forms two distinct outer rings (OR1 and OR2), SctD and SctJ together form the distal inner ring (IR1), and the cytoplasmic amino terminus of SctD forms the innermost ring (IR2) (31). The highest resolution cryo-EM models of the *Salmonella* basal body reveal an overall three-fold rotational symmetry, with a resultant symmetry mismatch between the inner and outer layers: each basal body contains 24 SctD molecules, 24 SctJ molecules, and 15 SctC molecules (32). While the 24-fold symmetry of the inner membrane rings appears conserved across T3SS (33, 34), the stoichiometry of the the SctC outer membrane ring may vary between species (12-15 molecules per basal body), such that some systems have an overall 12-fold rotational symmetry (34, 35).

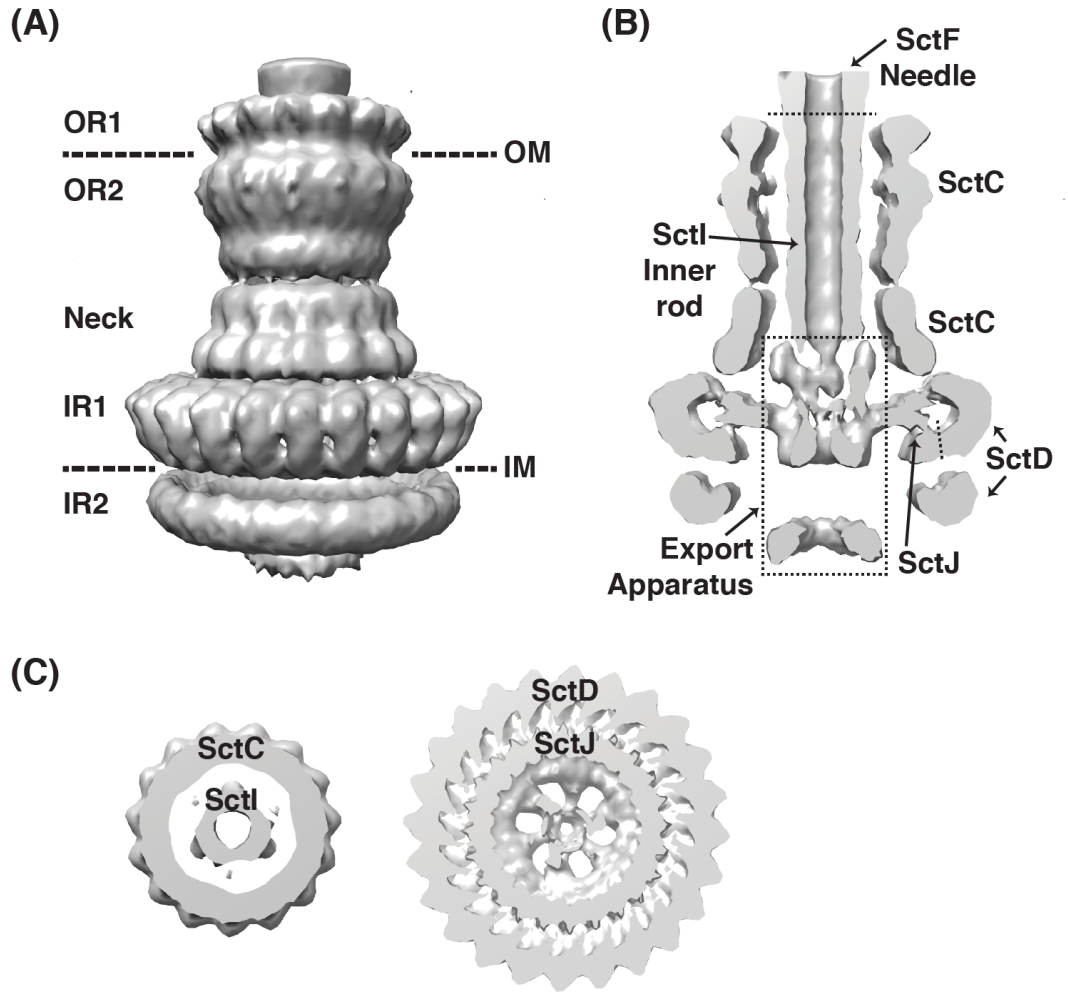


Figure 1.1: Gross architecture of the T3SS. Cryo-EM reconstruction of the *Salmonella typhimurium* injectisome basal body at subnanometer resolution reveals its overall architecture. (A) Surface representation of the highest resolution cryo-EM map (EMD 1875, contour level 0.0233) published by Marlovits and colleagues (32). Dashed lines indicate the positions of bacterial membranes *in vivo*. Abbreviations used: OR, outer ring; IR, inner ring; OM, outer membrane; IM, inner membrane. (B) An axial section through the map in (A). (C) Transverse sections through the map in (A) at the level of the neck (left) and IR1 (right).

SctC is homologous to the Type II secretion system secretins (9, 36), and like other secretin family members requires a pilotin lipoprotein for its optimal localization and assembly (37-39). The membrane-embedded, β-rich region at the SctC carboxy terminus can be isolated and has been visualized by EM (35), but it has yet to be characterized at moderate or high-resolution. The periplasmic amino terminus of SctC contains a modular domain architecture (40) that interacts with the inner membrane ring (41, 42).

SctD and SctJ form the inner membrane rings (31). Each is anchored to the membrane by a single transmembrane helix, and SctJ is additionally lipidated near its amino terminus (43). Like the amino-terminal periplasmic region of SctC, the periplasmic domains of SctD and SctJ comprise a modular multidomain architecture (40). Despite differences in connectivity and little sequence homology, the mixed α/β domains of SctC, SctD, and SctJ show a similar three-dimensional structure: two α-helices pack against the same face of a three strand β-sheet (40, 44). Superhelical crystal packing of the *E. coil* SctJ periplasmic region provided initial insights into the mechanism of inner membrane ring assembly (43), and the modular arrangement of these domains seems to promote oligomerization (45); however, none of these domains have been shown to clearly form annuli in solution, suggesting that additional constraints (e.g. protein-protein interactions or lipid membrane planarity) are critical for ring formation. Similarly, despite their 1:1 stoichiometry in the basal body, the periplasmic domains of SctD and SctJ have not been crystalized in complex.

Might there be a specific functional advantage for the modular domain architecture common to the SctCDJ periplasmic regions? Recent *in situ* electron tomography of the *Yersinia* and *Shigella* injectisomes shows that the basal body has the ability to stretch in response to osmotic expansion of the periplasmic space (34). This resilience could be of potential importance for the maintenance of intact T3SS injectisomes under physiologic stresses and membrane deformations (34). Molecular dynamic simulations suggest that relative motions of the SctD periplasmic domains could account in part for this flexibility (34), but this hypothesis requires empiric support.

The amino-terminal cytoplasmic domain of SctD forms the innermost ring of the T3SS basal body. High-resolution structural analyses have determined this domain to have a forkhead-associated fold that interacts with cytoplasmic components of the T3SS (46-49). Forkhead-associated domains are β-sandwiches that typically serve as phosphothreonine binding scaffolds, suggesting a means of signal-dependent recruitment of the cytoplasmic secretory apparatus to the basal body. However, the potential phosphopeptide binding residues are not conserved among T3SS (48), and the precise nature of the interactions between SctD and the cytoplasmic apparatus remains to be determined.

Within the central lumen of the basal body annuli, SctI is believed to form a cylindrical “inner rod” (31) structure that may support the extracellular needle filament. Computational methods have suggested a predominantly α-helical structure for SctI similar to that of the needle filament protomer (below); however,

structural analyses of *Salmonella* and *Shigella* SctI in solution showed little secondary or tertiary structure (50). It remains to be determined whether SctI can adopt a stable fold within the confines of the basal body. The functional significance of the inner rod in the regulation of needle length and secretion substrate switching will be discussed below.

Direct structural characterization of the T3SS basal body epitomizes the challenges associated with the interrogation of high-molecular weight macromolecular machines: the assembly spans two membranes and a layer of peptidoglycan, and *in situ* electron tomographic analyses suggest that the basal body is capable of substantial conformational dynamism (34). As direct, high-resolution structural characterization of the assembled T3SS basal body has not yet been possible, a multidisciplinary approach integrating cryo-EM maps, X-ray crystallographic domain structures, biochemical analyses and computer modeling has yielded a high-probability static model for injectisome architecture (51) (Figure 1.2). Such “hybrid” models will allow the testing of molecular-level hypotheses about ring assembly until the structure of the T3SS basal body has been determined at high-resolution *in toto.*

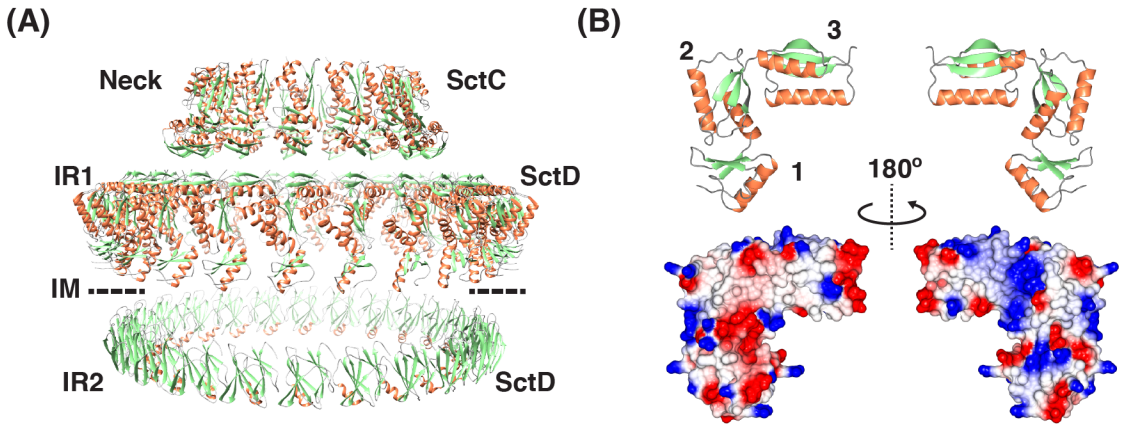


Figure 1.2: Hybrid models of basal body structure. (A) Computational modeling of the neck (SctC, PDB 3J1V), IR1 (SctD, PDB 3J1X), and IR2 (SctD, PDB 3J1W) annuli of the *Salmonella typhimurium* basal body. No high resolution structural information is available for the basal body above the neck. (B) In this model, complementary electrostatic surfaces support ring building, as shown for the SctD periplasmic domains. Note the modular domain architecture (enumerated 1, 2, 3) for $SctD_{periplasmic}$.

The inner membrane machinery. Five highly conserved inner membrane proteins (SctRSTUV) are necessary for the function of the pathogenic T3SS; however their individual functions are unclear. It is worth noting these proteins show a high degree of sequence homology to components of the evolutionarily related flagellar T3SS (Table 1), and may represent a functional core, serving critical chemical roles in initiating or powering protein secretion.

Among the SctRSTUV cohort, SctU and SctV possess cytoplasmic domains in addition to their transmembrane helices, and these domains have been best characterized to date. The cytoplasmic region of SctV contains a modular array of small domains (52-54) and crystallographic data suggest SctV may nonamerize (54). Intriguingly, such an oligomer is well suited to fit in a torus of SctV-associated density observed 5-10 nm beneath the basal body in EM reconstructions of the T3SS (34, 54). Lea and colleagues (54) have forwarded the hypothesis that this SctV homo-oligomer may serve as a “cage” to facilitate the complete unfolding of folded or partially unfolded secretion substrates, but this possibility has not yet been experimentally validated. The cytoplasmic region of SctU is considerably smaller than that of SctV and contains an autoprotease (55); its potential role in the regulation of secretion will be discussed below.

SctRSTUV are important in the organized, stepwise assembly of the T3SS basal body. Galán and colleagues (56) employed a combination of genetic and structural approaches to show that SctRSTUV help to organize the SctDJ inner membrane rings. Subsequently, the SctC ring and cytoplasmic machinery (below) are recruited, and the inner rod and needle polymers assembled (23). While

SctRSTUV are individually not strictly necessary for the formation of the SctCDJ basal body, the efficiency of basal body assembly is significantly decreased in their absence (56).

The needle filament. A needle-like filament tens of nanometers in length protrudes from the extracellular face of the T3SS basal body (27, 57). The needle is formed by a helical assembly of the protein SctF, with an outer diameter of 8 nm and an inner pore diameter of 2.5 nm (58). The apparent similarity of the T3SS basal body and needle filament to a macroscopic syringe makes it tempting to speculate that the T3SS directly injects its substrates into the host cell cytoplasm, with the needle filament serving as a conduit for the passage of partially unfolded effector proteins. Until recently, this hypothesis lacked direct empirical support, and alternative "non-injectisome" models for T3SS effector delivery had been proposed (59). Analyzing substrate-trapped injectisomes by cryo-EM, Marlovits and colleagues (60) and Kolbe and colleagues (61) demonstrated the presence of additional density in the lumen of the T3SS needle filament, consistent with the passage of partially unfolded substrate molecules through the needle.

High-resolution structures of monomeric SctF mutants (62) or chaperone-bound SctF (63, 64), have allowed the characterization of the needle protomer fold. SctF is a hairpin of alpha helices with an intervening conserved PXXP motif. The oligomeric nature of the needle filament had posed a practical barrier to high-resolution structure determination for the intact assembly. Recent hybrid

approaches combining cryo-EM with solid state nuclear magnetic resonance spectroscopy (NMR) and computational modeling have since afforded such a model for both the *Salmonella* (58) and *Shigella* (65) needle filaments. In these models, the SctF amino terminus is oriented towards the convex needle exterior, the carboxy terminus towards the lumen, and the apex loop connecting the two alpha helices points away from the bacterium. It should be noted that this arrangement is in contrast to prior lower resolution models (66), which oriented the SctF amino terminus towards the needle lumen. This correction is significant: the orientation of SctF protomers in the solid state NMR models is such that the lumen walls are formed by highly conserved residues, consistent with the passage of secretion substrates through the lumen (58, 65).

Assembly of needle filaments of a given length is necessary for the proper infectivity of T3SS-bearing pathogens, possibly matching the dimensions of host-pathogen adhesion complexes (67). How, though, is the length of the needle filament controlled? SctP regulates the length of the needle filament in several species (68-70). In *Yersinia spp.,* the number of residues in SctP correlates with needle filament length, leading to the hypothesis that SctP functions as a "molecular ruler" (70). That is, SctP might attach at one end to the basal body or cytoplasmic apparatus and at the other end to the growing needle filament, and once SctP was stretched beyond a given length, it would signal to the secretion apparatus to change substrates.

In contrast to the molecular ruler model, work in *Salmonella* suggest that SctP regulates needle length through control of inner rod assembly (57).

Salmonella lacking SctP show decreased density in the inner rod-supporting socket region, lack a polymerized inner rod, and generate long needles (57, 71). Accordingly, Galán and colleagues hypothesized that completion of inner rod assembly terminates needle growth in a SctP-dependent fashion. Consistent with this "timer" model of length control (23), overexpression of SctF or SctI leads to longer or shorter needles, respectively (57). Moreover, alanine-scanning mutagenesis analyses of the inner rod protein SctI in *Salmonella* revealed numerous point mutations that increased needle length without compromising secretory function, perhaps by slowing the rate of inner rod polymerization (71). Intriguingly, the elongated needles generated by most of these mutants remained attached to the basal body (71), in contrast to those produced by *sctP* deletion mutants, which are easily sheared off (57). This observation is consistent with the hypothesis that the polymerized inner rod joins with the needle filament, anchoring it to the basal body (71).

The needle tip and translocon pore. At the tip of the T3SS needle filament is a pentameric cap formed by the hydrophilic translocator protein (72, 73). The needle tip is believed to interact directly with the host cell surface to facilitate the insertion of a multimeric pore (74, 75), thus completing the cytoplasm-to-cytoplasm protein conduit. The structure and function of the needle tip is of particular biomedical interest, as the hydrophilic translocator protein is a protective antigen in anti-*Yersinia* vaccine formulations (76) and is a target of

recently developed passive immunization strategies for the treatment of *Pseudomonas aeruginosa* infections (77).

X-ray crystallographic analysis of the monomeric tip protein from several species reveals some conserved architectural features (78, 79): the tips of all species show an elongated coiled coil region and a central mixed α/β subdomain. The overall structure of the tip protein shows some interspecies variation though, with *Salmonella/Shigella* SipD/IpaD possessing an amino-terminal autochaperoning subdomain (79) lacked by the *Yersinia/Pseudomonas* LcrV/PcrV tip proteins (23, 75). High-resolution models of the pentameric needle tip are not available, but low resolution negative stain EM models have offered some insight into the organization of the needle tip. While both appear pentameric, the *Shigella* tip complex is narrow and elongated relative to that of the *Yersinia* tip (72, 73), and fitting the *Shigella* IpaD monomeric crystal structure into the EM map required a significant rearrangement of the mixed α/β domain (73), suggesting substantial conformational changes upon incorporation into the tip complex.

Attempts to model the tip protein-needle filament interaction at high-resolution using NMR or X-ray crystallographic data have so far proven challenging, with incompatibilities arising between the proposed models and other data sets (80-82). The crystal structure of a *Salmonella* SctF-SipD fusion protein identified a potential binding mode for the tip with the needle (80); however, modeling the fusion structure onto the solid state NMR model of the needle filament (58) resulted in steric clashes, suggesting that artifactual

constraints imposed by the protein fusion strategy biased the architecture of the complex (82). Regardless, a synthesis of the available data shows that the needle filament interacts with the tip protein at least in part through its elongated coiled coil, a motif observed in all T3SS tip proteins described to date.

EM and biochemical analyses have shown that the *Shigella* tip complex is actually a heteropentamer containing four copies of the hydrophilic translocator IpaD and one copy of the hydrophobic translocator IpaB (81). A refined tip model incorporating this insight is similar to previous models, in that the amino and carboxy termini of IpaD are oriented towards the needle filament, with portions of the coiled coil region contacting the needle. However, the refined model presents an IpaD orientation consistent with antibody binding data and the heteropentameric architecture explains the transition from a helical needle filament to a nearly flat-topped tip complex (81).

In contrast to the annular, pentameric tip complexes observed for other T3SS, enteropathogenic *E. coli* possess a long filamentous needle accessory comprised of EspA, the SipD/IpaD/LcrV/PcrV homologue (83). EspA forms a helical filament similar to the SctF needle, with an internal diameter of approximately 2.5 nm (84), suggesting that it functions to extend the T3SS transport conduit. Filling a functional niche similar to needle filament length control in other T3SS, *E. coli* EspA polymers may adapt the injectisome to reach the target cell membrane beneath the intestinal glycocalyx (75).

The translocon permeating the host cell membrane is formed by the hydrophobic translocators SipB/SipC in *Salmonella* and their homologues (Table

1). Experiments in red blood cell membranes have shown that the needle tip is crucial for the insertion of the hydrophobic translocators into the host membrane and/or the organization of inserted translocators into functional pores (74, 85). While numerous experimental approaches have been employed to characterize the pore diameter and structure of the translocon (75), direct structural interrogation of native translocons is lacking. Intriguingly, the amino-termini of *Salmonella* SipB and *Shigella* IpaB contain extended coiled coils reminiscent of the colicin family of bacteriocins (86), which are known to function in the delivery of protein toxins across bacterial membranes. However, the precise mechanisms of host cell recognition, membrane insertion, pore formation, and protein translocation remain unclear for the T3SS translocon.

1.3 Substrate recruitment and secretion

T3SS secrete only a small fraction of the proteins present in the bacterial cytosol (87), and do so in a defined hierarchy. How does the T3SS select its substrates, how is their secretion hierarchy maintained, and what is the mechanism of secretion? A combination of genetic, biochemical, and structural data provide insight into the role of cytoplasmic injectisome-associated proteins in these processes.

Secretion chaperones. The amino-terminal ~100 amino acids of T3SS substrates possess two secretion signals: an unstructured extreme amino terminus followed by a chaperone-binding region. While the extreme amino-termini of T3SS

substrates are highly variable, computational approaches have identified commonalities in the chemical composition of this secretion signal (88-90): the first ~15 amino acids in an effector sequence are enriched in serine, threonine, isoleucine, and proline. Indeed, an effector with a synthetic, amphipathic poly-serine/isoleucine secretion signal was secreted by *Yersinia* even in the absence of the second secretion signal (its secretion chaperone) (91). However, other studies have shown that chaperone-substrate interactions are necessary for targeting substrates specifically to the injectisome T3SS, as the extreme amino-terminal secretion signal sequence can facilitate injectisome substrate export through the flagellar apparatus in the absence of a chaperone-binding domain (92). This finding suggests that the extreme amino-terminal secretion signal is evolutionarily "ancient" and shared by both types of T3SS (92).

Downstream of the amino-terminal secretion signal, each secretion substrate is recognized by a specific chaperone protein that maintains the bound region of the substrate in a partially unfolded state (93). It is hypothesized that this nonglobular conformation primes the substrate for secretion through the narrow aperture of the T3SS conduit (94). Chaperones can be classified by their structure and substrate type (26), as follows: Class IA chaperones are mixed α/β homodimers that bind to one effector; Class IB are structurally similar to IA but bind multiple effectors; Class II are alpha helical tetratricopeptide repeat (TPR) proteins that bind to translocon proteins; and Class III are heterodimeric TPR proteins that bind SctF needle filament protomers. It should be noted that while the premature polymerization of SipD/IpaD tip proteins is prevented by their

autochaperoning domain (79), tip proteins in *Yersinia* and *Pseudomonas* utilize a unique chaperone (LcrG/PcrG) to stabilize their tip protein monomers (95).

Like the extreme amino-terminal secretion signal, the chaperone-binding sequences of secretion substrates are variable. However, Class I chaperone recognition of a conserved β-strand(s) is a common feature of a diverse array of T3SS effectors (96), and it is conserved from animal to plant pathogens (97) (Figure 1.3). Indeed, the sparse sequence conservation associated with the "β-motif" (96) can be used to recognize previously unknown T3SS effector proteins (98). In contrast, Class II and III TPR chaperones can recognize substrate sequences in either extended unstructured or α-helical conformations; the commonality here is that the TPR concavity is used to bind the substrate (26) (Figure 1.3).

The aforementioned substrate secretion signals are not only necessary for protein secretion through the T3SS, they are sufficient. Fusion of the secretion signal and chaperone-binding domain from endogenous T3SS substrates to heterologously expressed proteins results in their secretion through the T3SS (99), provided they can be properly unfolded for transit through the needle (60, 61). While this observation is noteworthy for its mechanistic insight into the targeting of virulence factors for secretion, it has allowed the benevolent re-engineering of the system to deliver protective antigens in vaccine design (100) and the large scale production of challenging protein reagents, like spider silk (101).

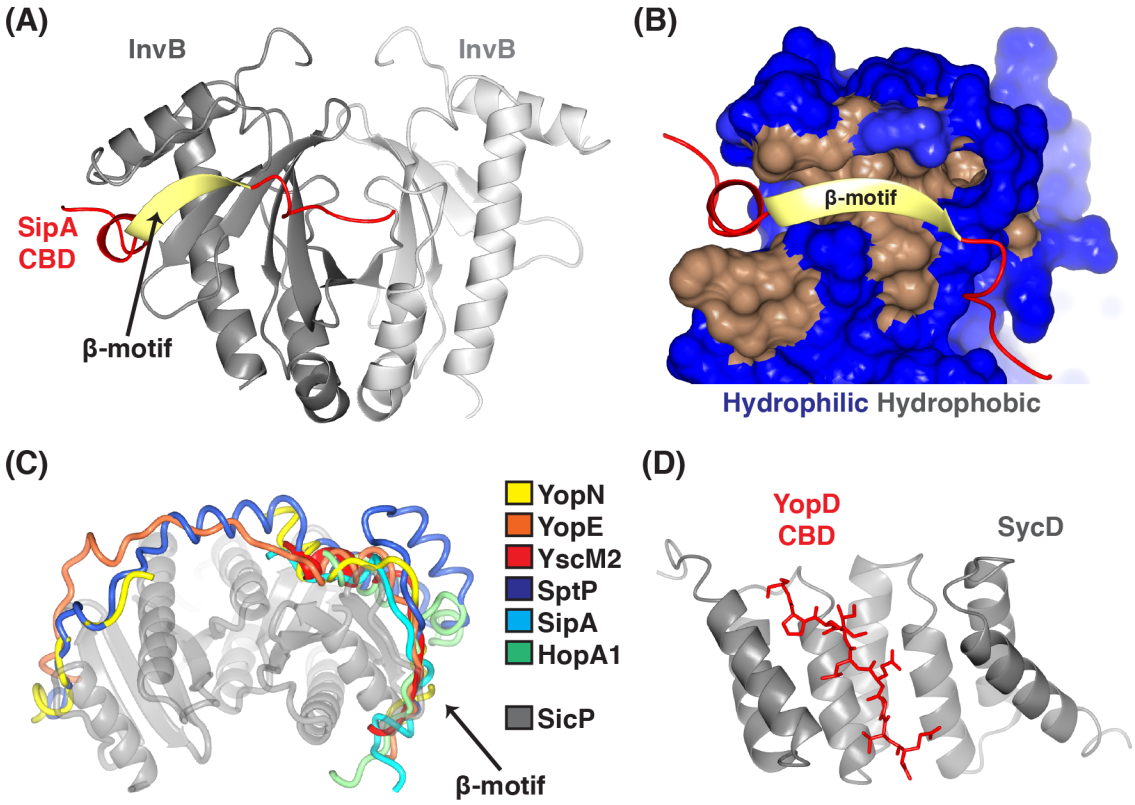


Figure 1.3: Chaperone-substrate interactions. Structural distinctions between effector-chaperone and translocator-chaperone complexes. (A) The structure of the *Salmonella* effector SipA chaperone-binding domain (CBD, red and yellow) in complex with the Class IB chaperone InvB (2 protomers: dark gray and light gray). PDB 2FM8 (52). The structurally conserved β-motif is highlighted in yellow. (B) The SipA β-motif is bound by a hydrophobic (gray) patch on the InvB surface (blue/gray). (C) Superposition of the CBDs from effectors from multiple species shows a common binding mode marked by the structurally conserved β-motif. The prototypical Class I chaperone SicP is shown in place of the various chaperones. PDB codes: YopN, 1XKP (102); YopE, 1L2W (103); YscM2, 1TTW (104); SptP-SicP, 1JYO (93); SipA, 2FM8 (52); HopA1, 4G6T (97). (D) The *Yersinia* translocator YopD CBD (red) lacks secondary structure and is bound by the concave cleft of the Class II chaperone SycD (gray). PDB 4AM9 (105).

The ATPase. Both the injectisome and flagellar T3SS include an ATPase with notable sequence homology to the β subunit of the F_0F_1 ATPase (106). For the injectisome, this ATPase is SctN; for the flagellar apparatus, FliI (Table 1.1). High-resolution structural analysis of the *E. coli* SctN catalytic domain showed similarities to V- and F-type ATPases and confirmed that SctN hexamerization would be required for efficient ATP hydrolysis (107). Both FliI (108, 109) and SctN (106, 110) form such oligomers, but neither has been characterized structurally as a catalytically active hexamer.

Interaction of chaperone-substrate complexes with the T3SS ATPase SctN causes the dechaperoning and unfolding of the substrate in an ATP hydrolysis-dependent fashion (111). Given that disruption of tertiary structure is necessary to fit protein substrates in the 2.5 nm conduit of the needle filament (60, 61), it is not surprising that loss of function mutations in SctN cause the near complete abrogation of T3SS function (106, 111). Similarly, genomic deletion of *fliI* severely compromises flagellar filament assembly (112-115). Consistent with the role of SctN in preparing substrates for export, SctN/FliI-dependent density is observed directly beneath the T3SS basal body by EM (116, 117). The partial structural similarity of SctO/FliJ with the γ-subunit of F-type ATPases and the ability of FliJ to stimulate FliI hexamerization (109) has led to the hypothesis that this coiled coil containing protein (118) might connect oligomeric SctN to the SctV export gate, thus linking the sub-basal body toruses of electron density (54). However, the mere presence of a coiled coil is insufficient evidence to ascribe a γ-like function to SctO (23), and its precise role remains to be determined.

Given the ATPase activity of SctN/FliI, it is tempting to speculate that ATP hydrolysis provides the free-energy for protein secretion; however, chemical and genetic analyses show that SctN/FliI is not the sole energizer of the T3SS. Experiments with the protonophore carbonyl cyanide *m*-chlorophenyl hydrazone have shown that the inner membrane proton motive force is necessary for secretion by both injectisome (119) and flagellar (112, 113) T3SS. Additionally, the flagellar assembly defect of ATPase deletion mutants can be at least partially corrected by mutations that alter the export apparatus, increase substrate levels, or increase the magnitude of the proton motive force (112-114). A similar result was recently reported for SPI-1 T3SS in *Salmonella* (114). These results suggest that under sufficiently permissive conditions, the actual transit of substrates into and/or through the conduit is powered by the proton motive force. However, one must interpret these results with some caution, as ionophores can significantly perturb cellular physiology (120) and SctN-independent injectisome secretion of a substrate requiring dechaperoning has not yet been demonstrated (114). A reasonable synthesis of the available data might surmise that both ATP hydrolysis and the proton motive force are important for energizing T3SS: the SctN/FliI ATPase functions to dechaperone and begin unfolding the secretion substrates with optimal efficiency under (non-ideal) physiologic conditions, while the proton motive force is responsible for the apical transit of the nonglobular substrate (114).

Regardless of the quantitative contributions of either energy source, the mechanics of secretion remain poorly understood. There are no available high-

resolution structural models of the interaction between chaperone-substrate complexes and the SctN ATPase. One computational model suggests a mode of ATPase-chaperone interaction based on structural similarities between Class I chaperones and the F-type ATPase γ-subunit (121). While this model and the accompanying biochemical data are consistent with the observation that relatively carboxy-terminal residues of SctN interact with chaperones (121), its structural accuracy lacks empiric support. Recent small angle X-ray scattering (SAXS) data suggest an alternative model for the interaction of substrates, chaperones, and the ATPase. Complexes of the *Salmonella* effector-chaperone pair SopB-SigE are able to hexamerize in a concentration dependent manner with dimensions comparable to the hexameric models of the ATPase (122). While it is too early to say whether other chaperone-effector complexes can oligomerize (123), whether these oligomers can interact with SctN, or whether such interactions — even if physically possible — are physiologically relevant, these results raise the possibility of alternate ATPase-cargo stoichiometries.

Recent solution NMR analyses suggest an interesting role for chaperone structure in the targeting of substrates to the T3SS. In solution, the *E. coli* chaperone CesAB is a partially folded molten globule (124) that does not interact with the hexameric SctN (125). However, upon binding to its substrate — the EspA tip filament protein — CesAB becomes fully structured and is able to bind SctN (125). The binding site for SctN was mapped onto the CesAB-EspA heterodimer, where it covered regions of CesAB unstructured in the absence of substrate (125). Consistent with the hypothesis that substrate-induced folding of

the chaperone allows for targeting to the ATPase, a mutation-stabilized, structured CesAB homodimer was able to bind SctN in the absence of substrate (125). While it remains to be determined whether similar disorder-to-order transitions effect SctN binding for other chaperone classes, these results are consistent with the role of chaperone-substrate interactions in targeting substrates for secretion. Additionally, the observation that substrate-chaperone complexes are recognized by hexameric SctN, but not its amino-terminally truncated monomeric form, suggests that ATPase hexamerization is critical for both hydrolytic catalysis and substrate recognition (125).

The sorting platform. Located at the peripheral cytoplasmic face of the flagellar basal body is a ring of robust density in EM reconstructions (117). Known as the "C-ring," this annulus is composed of the flagellar proteins FliM, FliN, and FliG, and it plays a role in flagellar motor function (torque generation) and rotational switching (24). FliM and FliN have an injectisome homologue with some conserved domains (SctQ), but torque generation by the injectisome is controversial (126) and a robust C-ring is absent in tomographic reconstructions of the injectisome basal body (34, 117). However, immuno-EM analysis of purified *Shigella* injectisomes shows localization of SctQ to the cytoplasmic face (127), suggesting that it plays some role in protein secretion or the regulation of the secretory process. Indeed, recent cryo-electron tomographic analyses of the *Shigella* injectisome have identified six SctQ-dependent "pods" of density

proximal to the cytoplasmic face of the basal body, forming a structure distinct from that of the flagellar C-ring (128).

Seminal biochemical and genetics work by Galán and colleagues revealed that SctQ forms a critical “sorting platform” for the T3SS (129). Affinity purification of SctQ from secretion-competent *Salmonella* produces high molecular weight complexes containing the SctN ATPase, regulatory proteins, chaperones and secretion substrates (129). Most notably, the sorting platform plays a role in the hierarchical secretion of substrates, queuing substrates in their appropriate order. For example, in *Salmonella* with assembled injectisomes, the sorting platform was predominantly occupied by translocon proteins, but genomic deletion of the translocators allowed the next tier of substrates (effector proteins) to access the sorting platform (129).

In addition to SctQ, formation of the sorting platform required the proteins SctK and SctL (129). While the role of SctK is at present unclear, biochemical analyses of the flagellar apparatus shed light on the potential function of SctL. SctL is a homologue of the flagellar protein FliH (Table 1.1). The SctL/FliH family is predicted to have a conserved domain architecture: an amino-terminal disordered region is followed by a coiled coil and then a mixed α/β domain (130). The carboxy terminus of FliH interacts with the amino-terminal oligomerization domain of FliI (130), inhibiting its ATPase activity (131). While sequence similarities between the FliH carboxy-terminal domain and the F-type ATPase δ-subunit suggest a role for FliH as a “stator” (132), it is not obvious that FliH interacts with oligomeric FliI, and the structural details of this interaction are not

yet known. The amino terminus of FliH interacts with FliN (133), and this FliH-FliN interaction is important (134) — if not absolutely necessary (135) — for the recruitment of FliI to the export apparatus. Given that the homologous injectisome complex (SctQ-SctL-SctN) forms a portion of the sorting platform and that chaperone-substrate complexes interact with the ATPase, these data suggest that one function of the SctQ sorting platform could be to localize chaperone-effector-ATPase complexes to the injectisome export apparatus. Indeed, Minamino, Namba and colleagues have hypothesized that the ATPase exists in two forms: an ATP-hydrolyzing hexamer and a dynamic substrate-carrying monomer bound to FliH and the C-ring (136). Similarly, SctQ-injectisome interactions are dynamic in *Yersinia*, as injectisome-associated SctQ exchanges with a cytoplasmic pool with a half-time of approximately one minute (137).

Structural models of SctQ are a work in progress, and have focused to date on the carboxy-terminal third of the molecule. In *Pseudomonas syringae*, SctQ is spread over two open reading frames (*hrcQA* and *hrcQB*), much like FliM and FliN in the flagellar system. The structure of the carboxy-terminal domain of HrcQB is quite similar to that of the carboxy-terminal domain of FliN (138, 139); both domains are homodimers of the "**S**urface **P**resentation **O**f **A**ntigens" (SPOA) fold, which appears to be unique to T3SS (Figure 1.4). The folded core of each protomer is an antiparallel β-sheet, and a loop from each protomer containing a β-strand and α-helix wraps around the β-sheet core of the other protomer. Like two left hands grasping one another, an antiparallel beta-sheet "palm" of each protomer is grasped by the "fingers" of the other, with a "thumb" protruding from

the top of the palm and strands from each protomer forming an anti-parallel beta sheet on the “floor” of the assembly.

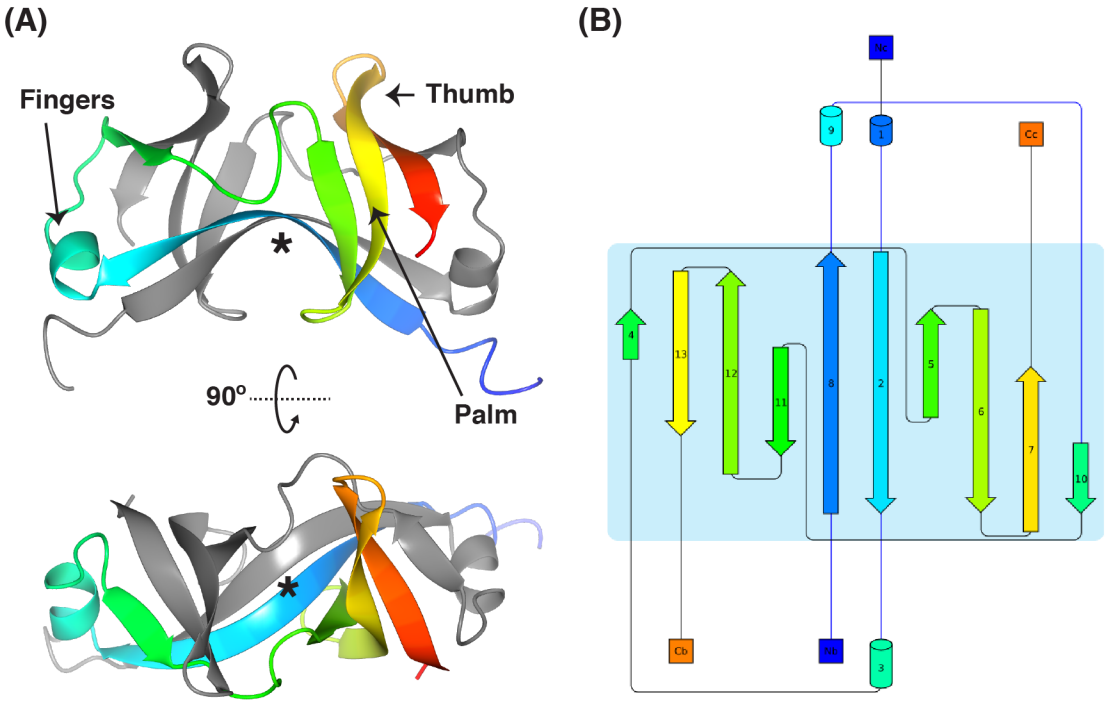


Figure 1.4: The SPOA fold. (A) Ribbon model of the SPOA homodimer from *Pseudomonas syringae* HrcQB (PDB 1O9Y) (138). One protomer is colored as a rainbow from amino- (blue) to carboxy- (red) termini. Asterisk indicates the location of the β-sheet floor. In keeping with the “two left hands” architectural analogy, the “fingers,” “thumb,” and “palm” of one protomer are labeled. (B) Topology diagram of the HrcQB SPOA homodimer; generated using Pro-Origami (140). The rainbow pattern is not aligned between (A) and (B).

In *Yersinia,* SctQ is the product of a single open reading frame (as in most injectisomes), but the carboxy-terminal SPOA domain is duplicitously translated from an internal translation start site (141). Structurally, this homodimer is architecturally similar to its *Pseudomonas* and flagellar homologues. The homodimer produced by this translation product is able to interact with full-length SctQ and, at least in the *Yersinia* system, is necessary for secretion *in vivo.* In both the flagellar (142, 143) and injectisome (144) systems, this SPOA domain tetramerizes as a dimer of dimers, but appears to do so in different orientations in each system. Cross-linking analyses suggest that the FliN SPOA tetramers form a “doughnut” at the base of the C-ring (142), but high-resolution support for this arrangement is lacking.

Despite the progress that has been made, numerous structural questions remain unanswered for the SctQ sorting platform. The function of the SPOA domain is unclear, SctQ-SctQ(SPOA) interactions have yet to be structurally characterized, and the structural basis for the interaction of SctL with SctQ has not yet been determined. Moreover, while the amino-terminal domains of FliM have well characterized functions in the regulation of flagellar rotation switching (145, 146), these motor functions are likely flagella-specific and involve interactions with partners not conserved from the flagellar apparatus to the injectisome (e.g. FliG and CheY). Thus, the function of the SctQ amino terminus is also unclear. Lastly, how and when SctQ or its soluble interaction partners interface with the basal body or export apparatus remains to be determined.

Substrate switching. T3SS substrates are secreted in a defined order that is necessary for the proper assembly and function of the system (147, 148): secretion of the needle filament (SctF) and inner rod (SctI) is followed by secretion of the needle tip protein and translocon pore proteins, which is followed by the secretion of effector proteins. Thus, it seems that there are several sequential substrate "switching" events that must occur for the hierarchical secretion of substrates to be maintained (148).

The first such switching event halts the extension of the growing needle filament and allows for secretion of the needle tip protein. As discussed above, the length of the needle filament is controlled by the assembly of the inner rod in a SctP-dependent fashion (57, 71). Full deletion of *sctP* locks the T3SS into a mode of exclusive SctF filament secretion; that is, deletion of *sctP* results in not only elongated needles, but a lack of translocon and effector secretion (68, 70, 149, 150). Indeed, deletion of *sctP* results in the absence of translocon components from the SctQ sorting platform (129). However, small deletions in the amino-terminal regions of SctP alter needle length without compromising translocon secretion, suggesting that some portion of SctP performs a crucial switching function (150). Deletions within the conserved mixed α/β region at the carboxy terminus of the protein compromise translocon secretion (in addition to disrupting needle length regulation), and this presumptive domain has been termed the "type III secretion substrate specificity switch" (T3S4) domain (150).

The three-dimensional structure of an injectisome T3S4 domain has not yet been determined, but the flagellar FliK T3S4 domain has been solved by

NMR (151). The carboxy-terminal domain of FliK possesses two α-helices folded against a four-strand β-sheet (151), and the predicted structural conservation of these secondary structural elements in SctP suggests that this model may be generalizable to the injectisome. While it is still unclear at the molecular level how SctP functions to promote specificity switching, its interaction partners suggest some viable hypotheses. For example, the SctP T3S4 domain interacts with the SctO protein (152), suggesting that it may be able to transmit regulatory information to the SctN ATPase or the SctV export gate. Moreover, the T3S4 domain interacts with the cytoplasmic autoprotease domain of SctU (153). Like SctP, SctU regulates the secretion of the the inner rod protein (154). The interaction between these two proteins is intriguing given that SctU interacts with components of the SctQ sorting platform and the SctV export apparatus (155), again suggesting mechanisms for the relay of switching information throughout the secretory apparatus.

The second major switching event distinguishes between translocon components and effector proteins (129). Deletion of translocon components allows for the localization of effector proteins to the sorting platform, consistent with a model where a gradient of substrate and/or chaperone affinities for the sorting platform controls the hierarchy of secretion (129). The identification of several classes of secretion apparatus mutants that can secrete effectors but not translocon proteins offers some insights into the establishment of secretion hierarchy. Deletion of *sctW* in *Salmonella* results in the specific loss of translocon component secretion (147). SctW binds the translocon proteins and their

chaperone in *Salmonella* (SicA) (147), and it is necessary for translocator binding to the SctQ sorting platform (129). These observations are consistent with SctW enhancing the affinity of translocon-containing complexes for the sorting platform. However, recent genetics data suggest the mechanism of hierarchy control for SctW may be more complex. A subset of the SctI alanine mutants identified by Lefebre and Galán (71) have normal needle lengths but phenocopy *sctW* deletion, and an interaction between SctW and SctI was recently reported in *Shigella* (156). Together, these data raise the possibility of SctW binding not only the sorting platform but also portions of the basal body.

Further clouding the role of SctW in T3SS is the observation of species-specific effects of *sctW* mutation. In *Yersinia* and *Shigella*, SctW is secreted and *sctW* deletion does not specifically impair translocon protein secretion (157, 158). Moreover, the *Yersinia* SctW protein pair YopN/TyeA is part of a complex calcium response apparatus in the bacterial cytosol (159) that involves several *Yersinia*-specific proteins (160, 161). While the structures of the *Shigella* and *Yersinia* SctW homologues have been determined (102, 162), a fuller understanding of SctW function (and its species-specific nuances) will require structural characterization in complex with other injectisome components.

In addition to its role in the first switching event, SctU is also involved in the second switch. The cytoplasmic domain of the SctU family autocatalyzes cleavage between the asparagine and proline residues of its conserved NPTH cleavage site (148). Alanine mutations on either side of the cleavage site cause aberrant specificity switching: translocon proteins are no longer secreted but

effector secretion remains intact (55, 157). An amphipathic linker connects the SctU transmembrane region to the cytoplasmic autoprotease, and this linker undergoes a disorder-to-order transition in the presence of anionic lipids (163). Introducing charge-altering mutations in the linker impaired T3SS function, suggesting that the ordering of the SctU linker against the bacterial inner membrane is crucial, perhaps favorably orienting the autoprotease domain for interactions with other members of the export apparatus (163). As mentioned above, SctU interacts with multiple members of the sorting platform, but the bases for these interactions — and the mechanisms by which they would effect specificity switching — are unclear.

Control of secretion. The T3SS can assemble a basal body, needle and tip, then pause in a "primed" state until the relevant stimulus arrives and secretion resumes. This strategy prevents the wanton waste of translocon and effector proteins. Interrogating this additional level of complexity is important to our full understanding of the pathobiology of T3SS, and may suggest routes to anti-virulence compounds that prevent the activation of otherwise structurally competent injectisomes.

Bile salts play a regulatory role in the T3SS used by several enteric pathogens. The interaction of bile salts with the *Shigella* tip complex promotes IpaB recruitment to the tip, forming the heteropentameric tip complex described above (81, 164). In contrast, bile salts suppress *Salmonella* SPI-1 T3SS function (165). These observations provide an intriguing correlation between host

gastrointestinal physiology and pathogen virulence that ties environmental factors to the species-specific adaptation of the T3SS. Despite reports describing the interaction of bile salts with monomeric *Shigella* IpaD (166) and *Salmonella* SipD (80, 167), the structural basis for bile salt interaction with the intact tip complex has yet to be determined in either species, and so the mechanism of its regulatory activity remains unclear.

Contact with host cells stimulates the activation of T3SS in several species (168-170). In *Salmonella*, contact with target cells stimulates the secretion of the translocon proteins SipB and SipC (171), and in *Shigella,* interaction of the IpaD-IpaB tip with liposomes resembling host cell membranes induces IpaC secretion (172). It is tempting to speculate that contact of the needle tip with the host cell sends a mechanical signal to the basal body and/or export apparatus that reinitiates secretion (10, 23, 173). As the connecting factor between the host cell surface and the basal body, the needle filament itself is a promising candidate for force transduction. Specific needle filament protein mutations can trap the *Shigella* T3SS in a constitutively active secretion mode, and one might hypothesize that these mutations stabilize needle filaments in a post-contact activated conformation (174). However, the filaments formed by these mutants do not exhibit the gross conformational changes one might expect if the needle filament architecture were transducing this signal (174). Alternatively, local changes in the tip environment may permit the secretion of substrates trapped within the needle by a closed tip, restarting secretion without requiring signal transduction to the bacterial cytoplasm or basal body (173).

Work from the *Salmonella* SPI-2 T3SS suggests a tantalizing third (and non-mutually exclusive) possibility, that the needle is not only a conduit for protein secretion, but a passageway for the diffusion of chemical signals (175). *Salmonella* makes use of two T3SS: broadly, the SPI-1 T3SS promotes cell invasion and subsequently the SPI-2 T3SS facilitates the formation of the *Salmonella* Containing Vacuole (SCV), an intracellular environment for *Salmonella* survival and replication (176). Holden and colleagues noted that priming of the SPI-2 T3SS requires exposure of the bacteria to low pH (as would be experienced in the endosomal compartment), but that triggering of effector secretion required a return to neutral pH (175). It is noteworthy that the SPI-2 SctW protein was required for this transition (175), consistent with the apparent role of SctW in translocon-to-effector specificity switching in other systems (above). However, it is most intriguing that this switch required intact translocon components, suggesting that the neutral pH signal may be transduced from the host cell cytosol, through the translocon and needle, to the basal body and/or export apparatus (175).

1.4 Summary

In the past 20 years, our models of T3SS structure and function have evolved substantially, from the first visualization of the injectisome architecture, to the high-resolution structural interrogation of many of its individual components. Combining these insights with a plethora of genetic and biochemical data, the molecular mechanics of this astounding secretory nanomachine are coming into

focus. However, numerous questions remain — the answers to which are critical to our understanding of bacterial virulence, the design of new therapeutics, and the imaginative re-engineering of the system.

Despite the improvements in cryo-EM models of the injectisome, the precise architecture of the membrane embedded components of the T3SS is still unclear, as is the structural basis for their interactions with the soluble components of the system. The native structures of the filament-bound needle tip and the translocon in the host membrane must be determined to understand how the extracellular environment regulates secretion, how proteins penetrate to the host cytosol, and how to rationally design secretion-blocking vaccines. Although the constituents of the cytoplasmic sorting platform have been identified, the structural bases for their interactions are unknown: how the sorting platform assembles, how substrate-chaperone complexes engage the system, and how the numerous regulatory elements interact to govern a secretory hierarchy all remain to be determined.

Answering these questions is likely to require a hybrid approach, characterizing local interactions and large assemblies alike, and employing a range of structural and molecular techniques. However, it is clear that high-resolution models of intact macromolecular assemblies (e.g. basal body, needle tip, translocon pore) would greatly advance the field. Much like the role that atomic models of the ribosome have played in the interrogation of its multiple functional states, one can imagine the watershed of insight that would come from successful visualization of the injectisome or sorting platform in each of their

several forms: needle-assembling, translocator-secreting, and effector-secreting. Ideally, these mechanistic insights will allow the uncoupling of some pathogenic gram-negative bacteria from virulence and/or the re-engineering of the nanosyringe for the benefit of biotechnology.

The work described in this thesis makes strides towards the mechanistic understanding of T3SS sorting platform assembly by applying structural, biochemical, and genetic techniques to the characterization of SctQ/FliM/FliN and their interactions with other sorting platform components. I will present the molecular structures of distinct homotypic and heterotypic SPOA-SPOA interactions in the *Salmonella typhimurium* SPI-1 sorting platform protein SpaO (Chapter 2). In Chapter 3, I structurally characterize the interaction of the heterotypic SPOA complex with a regulator of the SPI-1 ATPase and demonstrate the necessity of the interaction for T3SS secretory function. Then, I will present the homologous structures from the *S. typhimurium* flagellar apparatus and compare and contrast them with their SPI-1 homologues, providing an explanation for the observed subtype specificity in sorting platform assembly (Chapter 4). In Chapter 5, biochemical evidence for an interaction of the SpaO amino-terminal domain(s) with the homotypic SPOA complex is presented. The implications of this work and potential future directions are discussed in Chapter 6. Lastly, the materials and methods used in this study are described (Chapter 7).

CHAPTER 2:

SPOA-SPOA INTERACTIONS IN SpaO

Portions of Chapters 2-7 have been published previously (177):

Notti et al. *(2015). A common assembly module in injectisome and flagellar type III secretion sorting platforms. Nature Communications* ***6****, 7125.*

To date, the T3SS sorting platform has been most thoroughly characterized in *S. typhimurium* (129). This species contains three T3SS: two injectisomes (SPI-1 and SPI-2, as discussed in Chapter 1) and a flagellar apparatus. As the genetic and biochemical characterization of the sorting platform explored the SPI-1 constituents, I chose to begin my efforts there. For clarity and consistency with the primary literature, I will now refer to the relevant SPI-1 gene products by their *Salmonella*-specific names.

2.1 Expression and purification of SpaO

The major constituent of the SPI-1 sorting platform is a 303-residue protein SpaO (SctQ) (129). SpaO was robustly heterologously expressed as a double hexahistidine fusion in *E. coli* and yielded soluble material by nickel-affiinity purification (Figure 2.1C). Like its *Yersinia* homologue, full length SpaO co-affinity purified with a small fragment (~14 kDa). In *Yersinia,* the ~11 kDa YscQ fragment is translated from an internal translation start site at M218 and is necessary for T3SS function (141). While SpaO lacks a potential initiator methionine codon in

this region of the gene, I hypothesized that a nearby valine GTG codon (Val_{GTG}203, arrow in Figure 2.1A) with a putative upstream ribosomal binding site (Figure 2.1B) might be able to function as a cryptic translation start site. Indeed, mutation of the Val203 codon from GTG to GTT resulted in the loss of fragment expression and co-purification (Figure 2.1C). Intriguingly, in the absence of this carboxy-terminal fragment, the soluble yield of full length SpaO protein was greatly diminished, despite an identical amino acid sequence and overall similar expression levels (Figure 2.1C).

The *Yersinia* YscQ carboxy-terminal fragment contains a SPOA domain as previously seen in *Pseudomonas* HrcQB and the flagellar protein FliN. Preliminary bioinformatic analyses — specifically, sequence homology and secondary structure prediction — suggested the presence of two putative SPOA domains in the carboxy-terminal half of SpaO, which we prospectively denote SPOA1 and SPOA2 (Figure 2.1A). The presumptive SPOA domain located within the cryptically translated carboxy-terminal fragment and at the carboxy terminus of the full length SpaO protein was 34% and 24% identical to its *Y. pseudotuberculosis* and *P. syringae* homologues, respectively, by Clustal Omega alignment (178). (These two structurally characterized SPOA domains are 26% identical to one another.) Putative SPOA1 showed lower sequence identity to the structurally characterized SPOAs from *Yersinia* and *Pseudomonas* (18% and 19%, respectively), but showed 24% identity to *S. typhimurium* SPOA2 and secondary structure prediction was consistent with a SPOA fold.

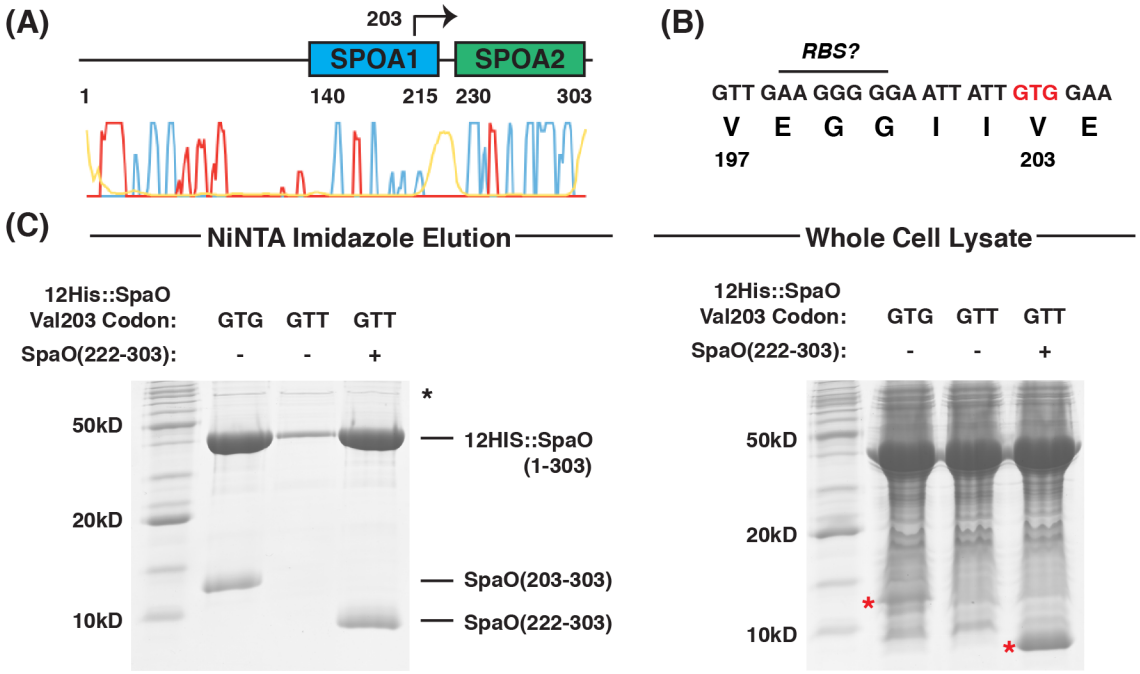


Figure 2.1: Bioinformatic analysis and heterologous expression of SpaO. (A) Secondary structure prediction and sequence homology suggest the presence of two putative SPOA domains in SpaO. Approximate domain boundaries are indicated beneath the block diagram. PSIPRED probability of helical character is plotted in red, strand in blue, and disorder in yellow. The arrow represents a predicted $Val_{GTG}203$ internal translation start site. (B) SpaO genomic sequence and amino acid translation in the vicinity of codon 203. Select codon numbers appear at bottom. The putative cryptic translation initiation codon is highlighted in red. A purine rich region 8bp upstream of $Val_{GTG}203$ is denoted as the hypothetical ribosomal binding site (RBS). (C) Heterologous expression of SpaO in *E. coli*. Coomassie-stained sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) gels showing the imidazole elution or whole cell lysates for recombinant expression experiments. The genotype of SpaO at codon 203 is shown over each lane and coexpression of the SPOA2 containing construct SpaO(222-303) is indicated with a "+". Presumptive chaperone contamination is indicated by a black asterisk. Red asterisks in the whole cell lysate gel indicate the bands corresponding to SpaO(203-303) and SpaO(222-303) (left and right, respectively).

I hypothesized that the SPOA2 region of the cryptically expressed carboxy-terminal SpaO fragment might mediate its interaction with the full length protein. Indeed, coexpression of double hexahistidine tagged SpaO(1-303, Val_{GTT}203) with the SPOA2 construct SpaO(222-303) resulted in co-affinity purification of the two polypeptides over NiNTA resin (Figure 2.1C). Furthermore, coexpression of this SPOA2 containing fragment with the SpaO(1-303, Val_{GTT}203) rescued the protein's solubility defect, restoring it to wild-type levels (Figure 2.1C) and implicating the SPOA2 domain specifically in this phenotype.

Co-affinity purified complexes of full length SpaO with SPOA2-containing constructs remained partially intact over anion-exchange chromatography, with a fraction of SPOA2 separating from the full length protein (Figure 2.2A, compare i, ii, and iii). However, intact complexes remained associated over size exclusion chromatography (Figure 2.2B, see ii) and I endeavored to crystallize this complex for structure determination by X-ray diffraction.

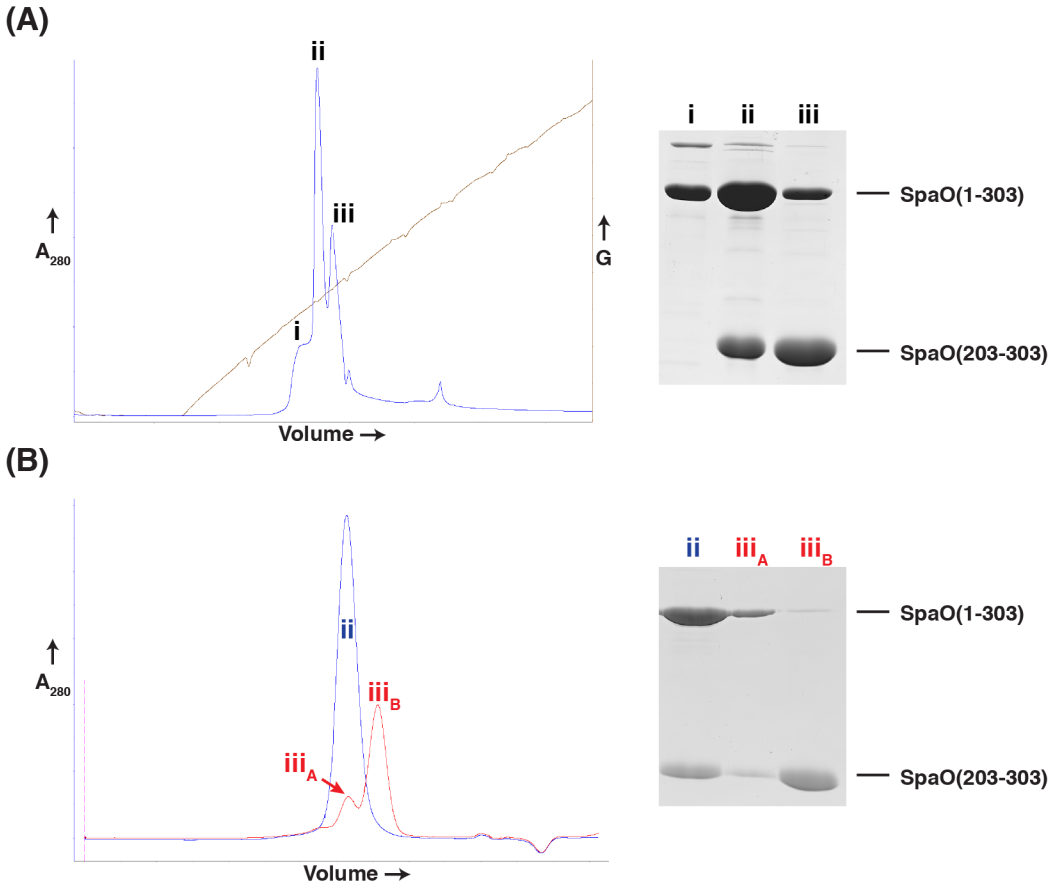


Figure 2.2: Purification of the SpaO-SPOA2 complex. (A) Full length SpaO and its co-purifying carboxy-terminal fragment were subjected to anion exchange chromatography (left). The blue trace shows UV absorbance at 280 nm (A_{280}) and the brown trace shows conductance (G). At right, samples of the elution at three points (i, ii, and iii) are shown on a coomassie-stained SDS-PAGE gel. (B) Fractions of peaks "ii" and "iii" were further purified by size exclusion chromatography (left), and samples of the elution are shown at right processed and stained as in (A).

2.2 Diffraction analysis of macromolecules: A brief overview

Before proceeding with further experimental description, it is prudent to briefly discuss the methodology of X-ray crystallography, as it will be central to much of the work presented here. (For a full resource, see Drenth and Mesters, ref. 180.) Over the past century, X-ray crystallography has revolutionized our understanding of small molecule and macromolecular structure by allowing the determination of the positions of individual atoms within a macromolecule (179). While it is now complemented by NMR spectroscopy and cryo-EM, X-ray crystallography was the first technique to allow the visualization of macromolecular structure at high (atomic or near-atomic) resolution.

In order to determine the structure of a protein or protein complex by X-ray diffraction, the target must be packed into an ordered, three dimensional crystal lattice by exposure to an empirically determined solution of precipitants, buffers, salts, and additives. Exposure of the crystal to bright, collimated beams of X-ray photons results in interactions of the protein's electrons with the incoming photons, changing their direction as they pass through the crystal. As a product of the repeating nature of the crystal lattice, patterned constructive and destructive interference of the scattered X-ray waves generates a stereotyped three dimensional array of spots measurable by X-ray detection equipment (a "diffraction pattern," Figure 2.3). Individual spots are identified by integer indices h, k, and l, and their intensities can be measured with high accuracy.

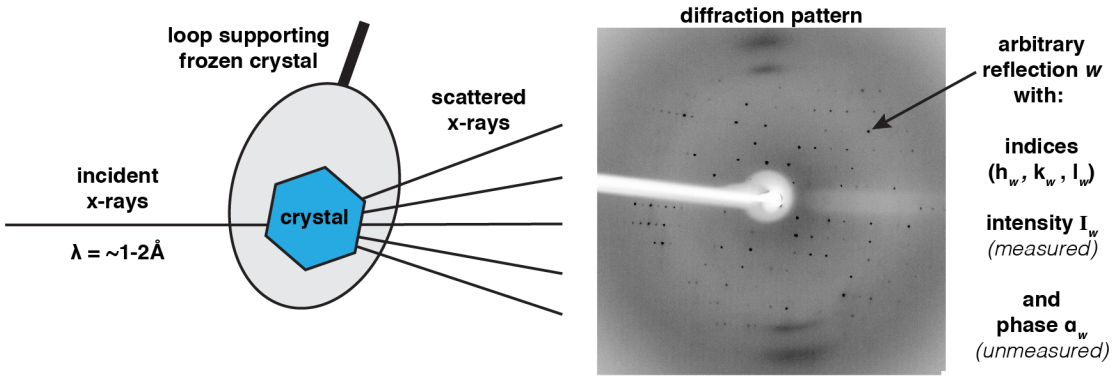


Figure 2.3: X-ray diffraction analysis of macromolecules. Exposure of crystalline macromolecules generates a diffraction pattern of "reflections." The intensity of these spots can be used to provide structural information about the macromolecule crystallized.

From these intensities, one can calculate a map of the electron density within the crystal lattice and build a molecular model of the crystallized protein's structure (180). The electron density (**ρ**) at a given point (x,y,z) within the crystal's repeating unit cell is related to the diffraction pattern by Fourier transformation:

$$\rho(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l}\left|F(h,k,l)\right| e^{-2\pi i(hx+ky+lz)+i\alpha(h,k,l)}$$

where V is the volume of the crystalline unit cell, F(h,k,l) is derived from the measured X-ray intensity at a given index (h,k,l), and α(h,k,l) is the phase angle of the wave at that index. While the X-ray intensities can be measured directly, the phase angles are not measured by the detector. This missing information creates the "phase problem" of X-ray crystallography: without the phase data, the electron density map cannot be determined from the X-ray diffraction pattern.

Phase information can be calculated by several methods, allowing for the successful determination of molecular structure from diffracted X-ray intensities. In this work, novel structures were solved by selenomethionine (SeMet) single wavelength anomalous diffraction (SAD). Many proteins produced in *E. coli* are compatible with the biosynthetic replacement of methionine sulfur atoms with selenium. Protein crystals derived from this selenium-doped material generate X-ray diffraction patterns with selenium position-dependent asymmetries ("anomalous" diffraction). Whereas the intensities I(h,k,l) and I(-h,-k,-l) are normally equivalent, this is no longer the case in the presence of a structured anomalous scatterer, and these measurable intensity differences can be used to determine Se positions and starting estimates of phase angles. For derivative structures in this work, known partial structures were used to determine starting phase estimates; this method is known as molecular replacement. In both cases, an iterative approach is used refine the molecular model and inferred phase angles against the empirical diffraction data until the fit between them has been optimized.

2.3 Structure of the SpaO SPOA2 homodimer

Full length SpaO in complex with separately translated SPOA2 was subjected to standard sparse matrix crystallization screening to generate crystals for diffraction analysis. While several conditions grew tetragonal crystals, these crystals were composed solely of the SPOA2 polypeptide and diffracted poorly. No crystals containing the full length SpaO protein were identified.

In order to facilitate structural characterization of SpaO, I attempted to break SpaO into smaller domain constructs that might be more amenable to crystallization. Ideally, structural characterization of small modules within the full length protein might shed light on higher order structures or functions. Given the inherent crystallizability of SPOA2, I decided to begin with this domain.

Using the aforementioned bioinformatic information and preliminary diffraction data from SpaO(222-303), the SPOA2 containing construct was truncated to SpaO(232-297). This construct was solubly expressed in *E. coli*, and purified by Ni-affinity, anion exchange, and size exclusion chromatography (Figure 2.4). Ghosh and colleagues previously used multiangle light scattering to determine that the *Y. pseudotuberculosis* YscQ carboxy-terminal fragment exists as a dimer in solution (141). Similarly, I subjected SpaO(232-297) to dynamic light scattering (DLS) and found its hydrodynamic radius (20 Å ± 1 Å) and resulting estimated mass (16 kDa ± 1 kDa) to be consistent with a dimer (expected mass: 16 kDa) over a range of concentrations (Figure 2.4).

Crystals of SpaO(232-297) were obtained that routinely diffracted to high resolution (all crystallization and cryoprotection conditions are available in Chapter 7). SeMet incorporated SpaO(232-297) crystallized under the same conditions as native material and diffraction data were collected to 1.35 Å resolution (Figure 2.5 and Table 2.1). Given the high resolution data obtained from SeMet crystals, the model was refined against the SeMet data set.

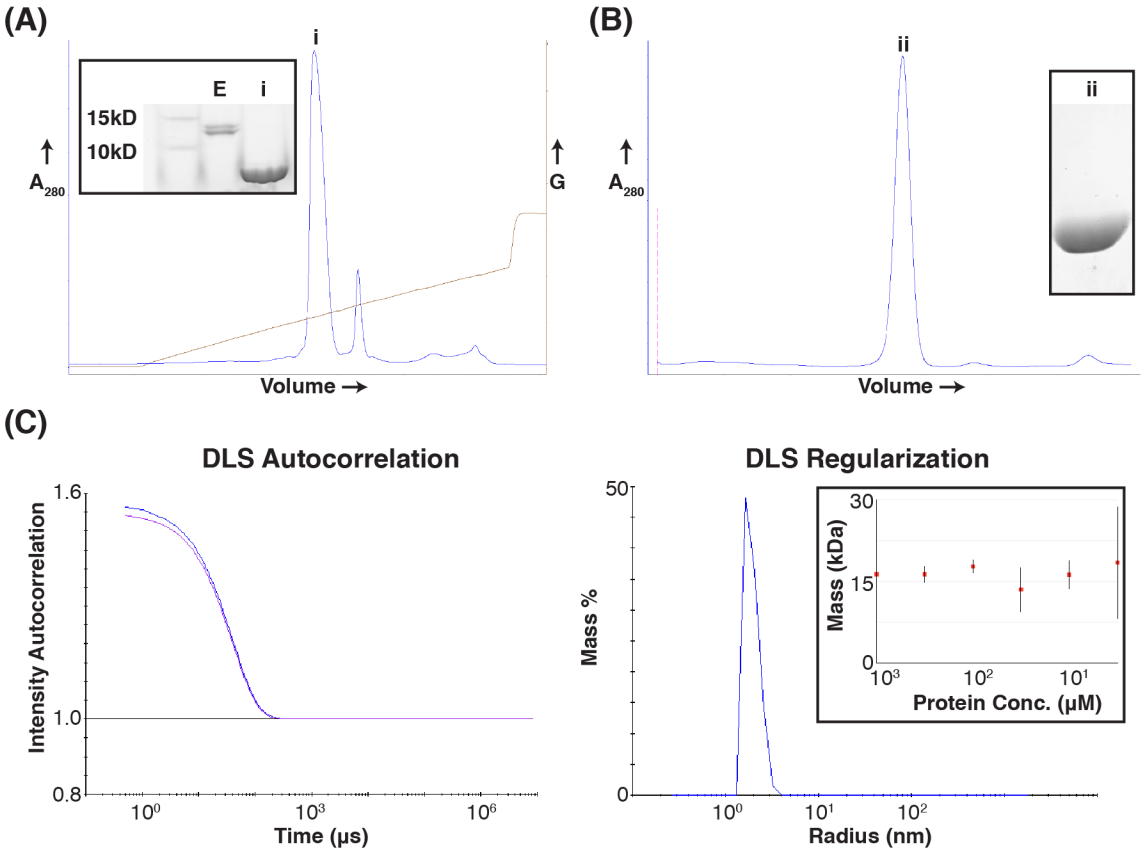


Figure 2.4: Purification and solution behavior of SpaO(232-297). (A) Anion exchange chromatographic purification of SpaO(232-297). The blue trace shows UV absorbance at 280 nm and the brown trace shows conductance. Inset shows coomassie-stained gel of the Ni-NTA resin imidazole elution (E) and the main SourceQ elution peak (i). Note that the affinity tag was removed with human rhinovirus 3C protease prior to anion exchange chromatography. (B) Size exclusion chromatography purification of SpaO(232-297) subsequent to anion exchange chromatography. Inset shows coomassie-stained gel of the main peak (ii). (C) DLS data for SpaO(232-297). The data for 1mM protein is shown in the autocorrelation and regularization graphs. In the autocorrelation graph, the raw data is shown in blue and the regularization fit is shown in violet. Inset shows the estimated particle mass over a range of concentrations.

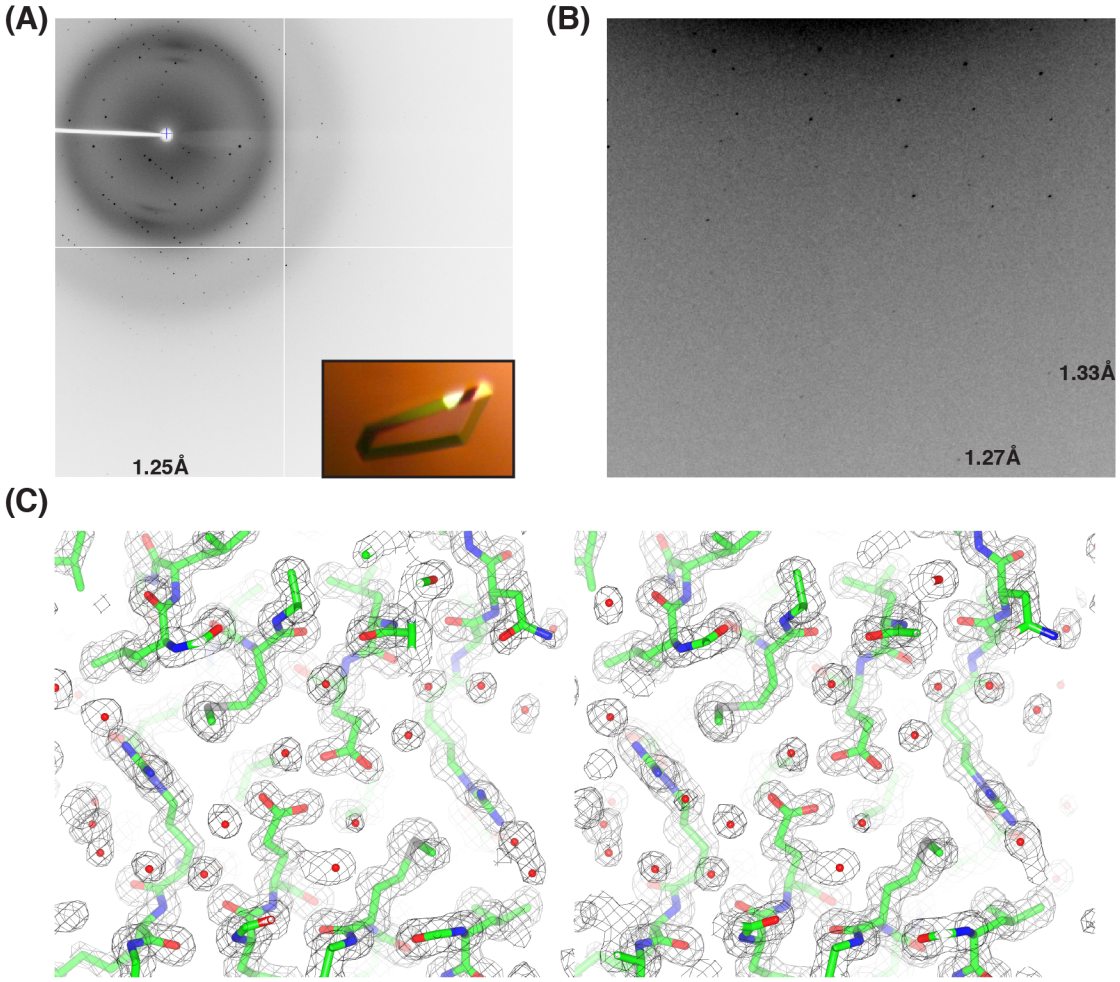


Figure 2.5: Diffraction analysis of SpaO(232-297). (A) Representative diffraction image for SpaO(232-297, SeMet). Image is truncated for fit. Edge resolution is 1.25 Å. Inset shows an example of the crystal form. (B) Zoomed view of the diffraction image in (A) highlighting high resolution spots. Resolution values shown correspond to spots immediately to their left. (C) Stereoimage of a selected region of the 2Fo-Fc electron density map for the 1.35 Å resolution SpaO(232-297) structure. The map is contoured to 1σ.

Table 2.1: Diffraction and refinement statistics for SpaO(232-297).

		SeMet
PDB ID		4YX1
Data Collection	**Space group**	$P2_1$
	Cell: a, b, c (Å)	35, 41.27, 48
	Cell: α, β, γ (°)	90, 103.92, 90
	Resolution (Å)	31.26–1.35 (1.37–1.35)
	R_{merge}	0.146 (1.281)
	I/σI	8.6 (2.1)
	$CC_{1/2}$	0.994 (0.750)
	Completeness (%)	99.7 (100)
	No. of reflections	29,246
	Redundancy	7.0 (7.1)
Refinement	**R_{work}/R_{free}**	0.1724/0.2053
	No. of atoms	1,286
	Protein	1,062
	Ligand/ion	2
	Water	222
	B factors: Protein	14.70
	B factors: Ligand/ion	14.20
	B factors: Water	32.10
	Geometry (r.m.s.d.)	
	Bond lengths (Å)	0.007
	Bond angles (°)	1.09
	[Φ,Ψ] Favored (%)	98
	[Φ,Ψ] Allowed (%)	100
	[Φ,Ψ] Outliers (%)	0

The structure of SpaO(232-297) confirms the fold of the presumptive SPOA2 domain as a bona fide SPOA fold (Figure 2.6). The construct forms a dimer in the crystallographic asymmetric unit, in which each protomer interfaces with the other in a “fingers-to-palm” arrangement, as described for the other structurally characterized SPOA homodimers (Figure 1.4). The SPOA2-SPOA2 interface buries 1898 $Å^2$ from each protomer.

The conserved arrangement of β-strands is striking when comparing the topology diagrams of SpaO SPOA2 (Figure 2.6) with its *P. syringae* homologue (Figure 1.4). Indeed, the SpaO SPOA2 homodimer superposes on its *Yersinia* and *Pseudomonas* homologues with 2.24 Å and 3.05 Å r.m.s.d., respectively. The structural similarity is readily apparent, especially when viewing the structures from the “top” (Figure 2.7A, right). Comparison of the SpaO SPOA2 sequence with that of *S. flexneri* Spa33, *Yersinia enterocolica* YscQ, and *P. aeruginosa* PscQ (Figure 2.7E) reveals that the greatest sequence homology maps to the hydrophobic core and dimerization interface of the structure, as opposed to solvent exposed surfaces (Figure 2.7B,C). Thus it seems that dimer formation is likely to be a conserved structural feature of the SPOA2 domain, consistent with all described crystal structures. The dimer architecture creates clefts on the “top” and “bottom” of the assembly, and it is noteworthy that these crevices are somewhat structurally conserved, displaying electronegative surfaces in *S. typhimurium* and its homologues (Figure 2.7D, *Y. pseudotuberculosis* shown), despite inconsistent sequence homology.

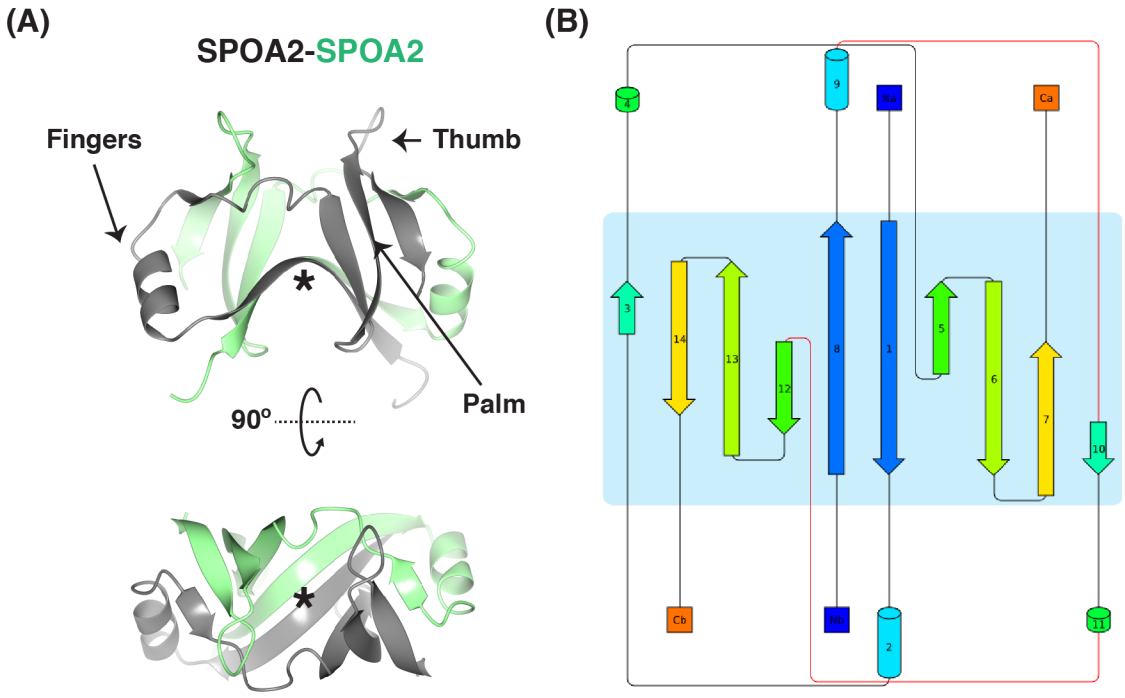
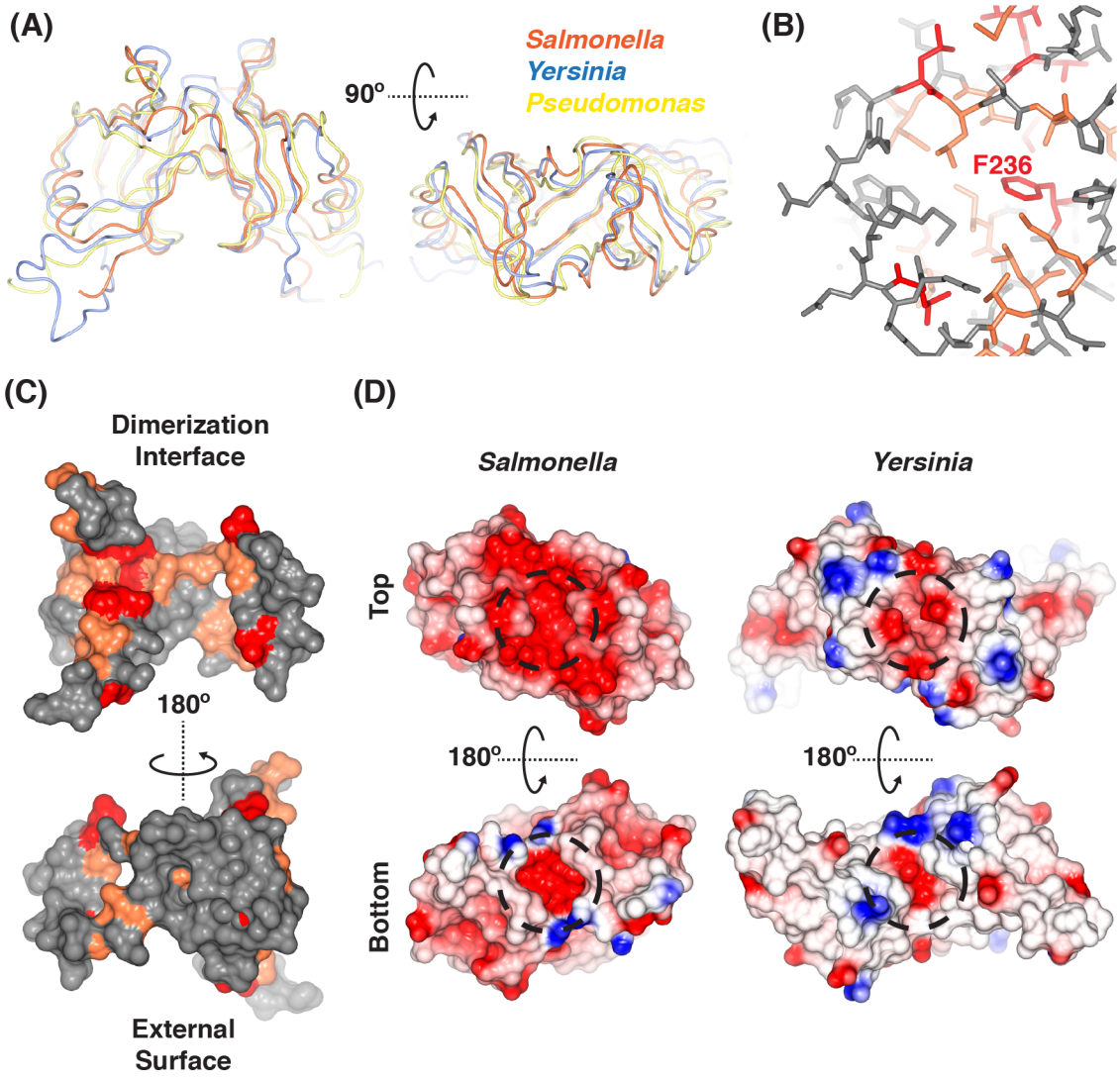


Figure 2.6: Structure of the SpaO SPOA2 homodimer. (A) Ribbon diagram of the SpaO(232-297) crystal structure. One protomer is shown in gray and the other in green. The structure is annotated as SPOA diagrams previously shown; asterisk indicates the β-sheet "floor." (B) Topology diagram for the structure in (A) generated using Pro-Origami.

Figure 2.7: Comparison of SpaO SPOA2 with known homologues. (A) The *S. typhimurium*, *Y. pseudotuberculosis*, and *P. syringae* SPOA2 homodimers are shown as backbone worm diagrams, superposed, and colored as indicated. (B) 100% conserved (red) or highly conserved (orange) residues in the hydrophobic core of the SPOA2 homodimer. A shell of highly conserved residues surrounds F236. (C) Comparison of the degree of sequence conservation on the solvent exposed (bottom) and dimerization interface (top) surfaces of a protomer in the SPOA2 homodimer. Colored as in (B). (D) Electrostatic surface diagrams for the *S. typhimurium*, and *Y. pseudotuberculosis* SPOA2 homodimers highlighting structurally conserved electronegative clefts on the “top” and “bottom” of the dimer (dashed circles). (E) Excerpt of the M-COFFEE alignment of SpaO, *S. flexneri* Spa33, *Y. enterocolica* YscQ, and *P. aeruginosa* PscQ is shown beginning with SpaO V232. Asterisk denotes identity; colon, high conservation; period, low conservation.

Figure 2.7: Comparison of SpaO SPOA2 with known homologues.



(E)

```
SpaO     VKLEFVLYRKNVTLAELEAMGQQQLLSLPTNAELNVEIMANGVLLGNGELVQMNDTLGVEIHEWLSESGNGE
Spa33    VKVDFILLEKKMTINELKMYVENELFKFPDDIVKHVNIKVNGSLVGHGELVSIEDGYGIEISSWMVK----E
YscQ     VQVSFEVGRQILDWHTLTSLEPGSLIDLTTPVDGEVRLLANGRLLGHGRLVEIQGRLGVRIERLTEVTI--S
PscQ     IPVSFEVGRRTLDLHTLSTLQPGSLLDLDCALDGEVRILANQRCLGIGELVRLQDRLGVRVTRLFGHDE--A

cons     : :.* : .: :    *      .*:.:      .*.: .*   :* *.** ::.  *:.:
```

2.4 Structural basis for SpaO SPOA1-SPOA2 interactions

Constructs containing only the presumptive SPOA1 domain were insoluble and could not be refolded after affinity purification under denaturing conditions (Figure 2.8A,B). However, constructs containing both presumptive SPOA1 and SPOA2 (residues 140-297) were stable and soluble (Figure 2.8A). As this construct contains the $Val_{GTT}203$ mutation, no free SPOA2 is translated. Ni-affinity purified SpaO(140-297) could be purified by subsequent anion exchange and size exclusion chromatography (Figure 2.9); however, this construct did not yield crystals when subjected to sparse matrix crystallization screening.

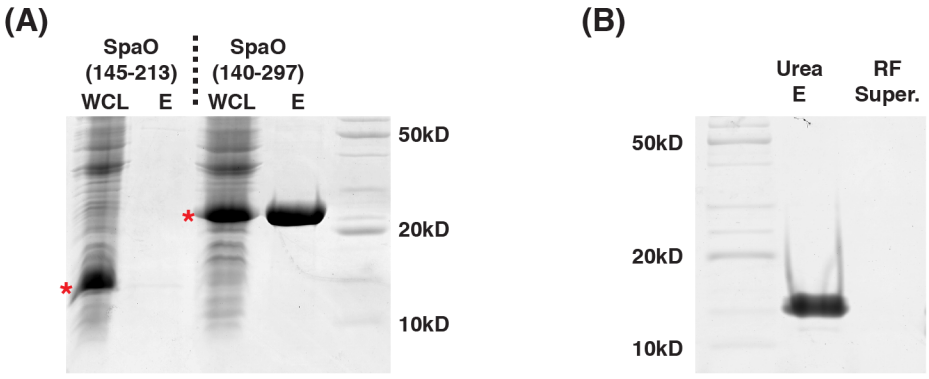


Figure 2.8: The putative SPOA1 of SpaO is stabilized by SPOA2. (A) Heterologous expression of double hexahistidine-tagged SpaO SPOA1 (145-213) and SPOA1-SPOA2 (140-297) constructs in *E. coli.* A coomassie-stained gel of the whole cell lysate (WCL) and soluble NiNTA imidazole elution (E) are shown for each construct. Red asterisks denote expression of the intended constructs as visualized in the WCL. (B) Attempted refolding of SpaO(145-213). NiNTA elution under denaturing conditions (Urea E) and the supernatant after refolding by dialysis and filtration of the insoluble material (RF Super.).

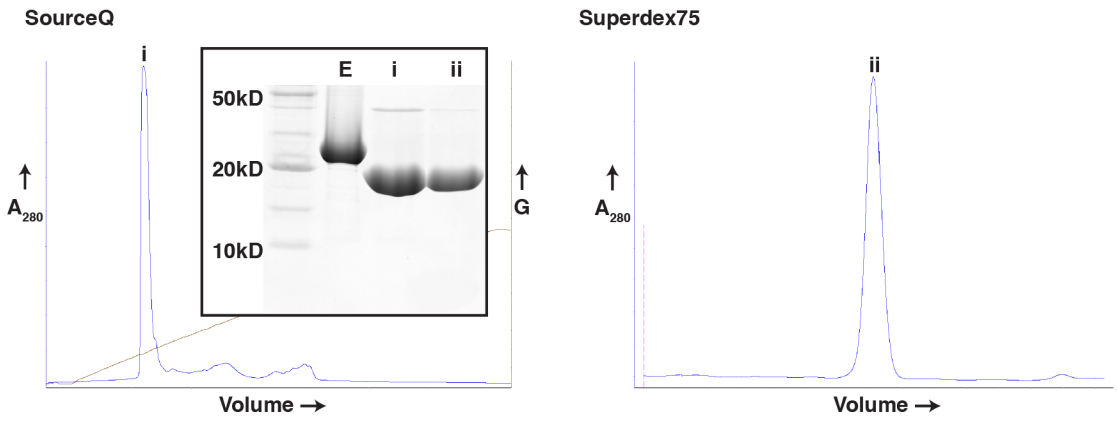


Figure 2.9: Expression and purification of SpaO(140-297). Anion exchange (SourceQ, left) and size exclusion (Superdex75, right) chromatographic purification of SpaO(140-297). The blue trace shows UV absorbance at 280 nm and the brown trace shows conductance. Inset shows coomassie-stained gel of the Ni-NTA resin imidazole elution (E) and the main anion exchange and size exclusion chromatography elution peaks (i and ii, respectively). Note that the affinity tag was removed with human rhinovirus 3C protease prior to anion exchange chromatography (compare E and i).

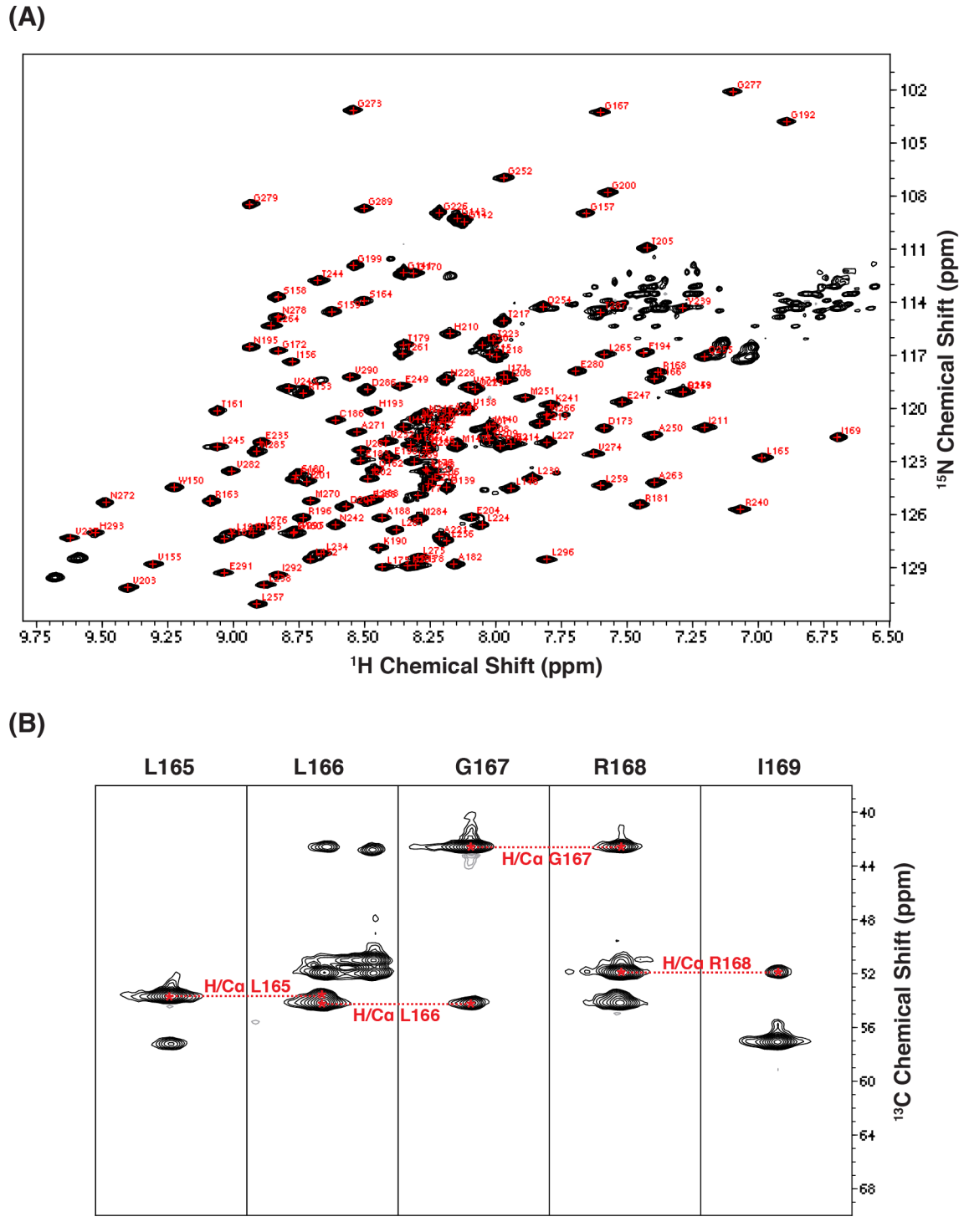
To further characterize SpaO(140-297) in solution and refine expression constructs for crystallography, I analyzed its backbone amide resonances by solution NMR. Heteronuclear NMR experiments allow the identification of resonances corresponding to individual backbone amide protons in a protein, all of which can be simultaneously visualized in heteronuclear single quantum coherence spectroscopy (HSQC) or transverse relaxation-optimized spectroscopy (TROSY) spectra. As these resonances are very sensitive to the immediate chemical environment of a given amide proton, a fully assigned HSQC

or TROSY serves as the chemical “fingerprint” of a protein and offers insight into local secondary structure and likelihood of disorder for a given residue (181).

SpaO(140-297) was analyzed by solution NMR and using a standard suite of backbone assignment experiments, and 150 backbone amide proton resonances were assigned out of a possible 155 (Figure 2.10). An example set of linked, assigned amide resonances from one experiment in this suite (TROSY-HNCA) is shown in Figure 2.10B. The overall impression from the ^{15}N-TROSY is one of a well dispersed pattern of amide resonances, consistent with a structured, folded globular protein. A set of resonances with sharp peaks and long relaxation times clustered around δ^{1}H=8.25 is consistent with a region or regions of disorder (e.g. flexible linker, extended surface loops, or disordered termini). Chemical shift deviation analysis of backbone amide resonances suggested a secondary structure pattern similar to that predicted by bioinformatic analyses: two SPOA domains connected by a flexible linker (Figure 2.11).

Figure 2.10: NMR analysis of SpaO(140-297). (A) ^{15}N-TROSY for SpaO(140-297) annotated with backbone amide resonance assignments in red. (B) Annotated strip-view of a portion of the TROSY-HNCA experiment used as part of a suite of backbone assignment experiments. Linked resonances are indicated with red asterisks connected with dotted lines.

Figure 2.10: NMR analysis of SpaO(140-297)



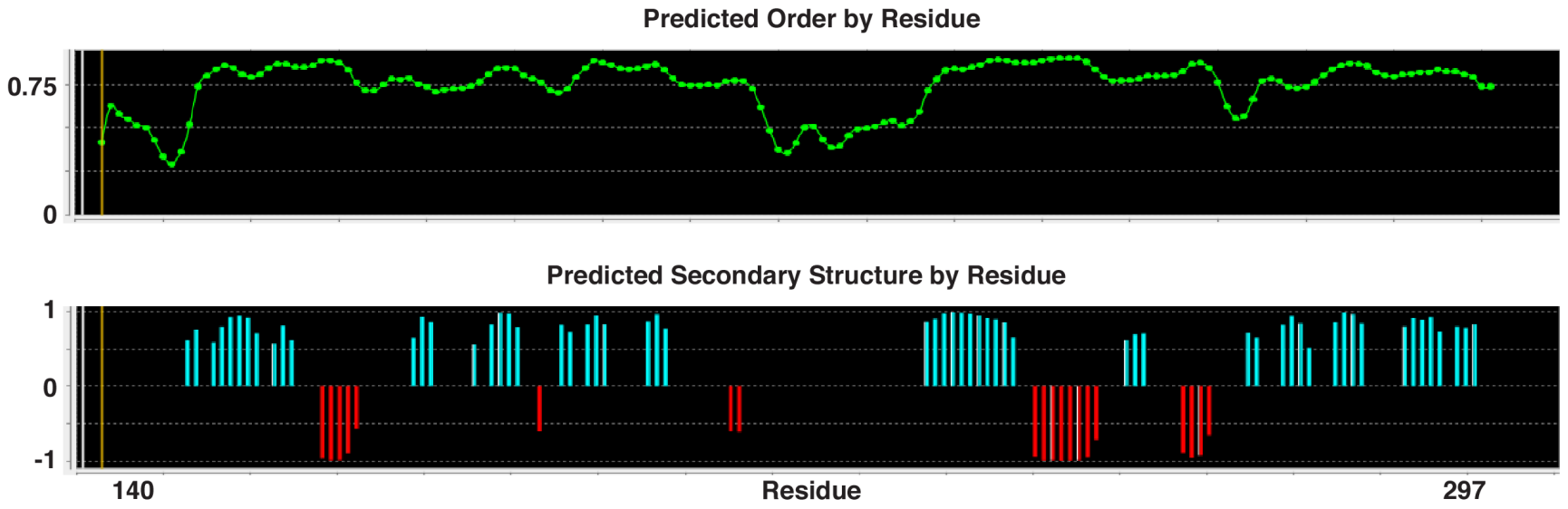


Figure 2.11: Chemical shift deviation analysis of SpaO(140-297). TALOS+ analysis of SpaO(140-297) amide proton chemical shifts. Predictions of residue order (top; green) and secondary structure (bottom; blue, strand; red, helix). The overall impression is one of two domains connected by a flexible linker. Secondary structure in each domain is grossly similar to bioinformatic predictions.

The putative SPOA1 domain might be solubilized by the presence of SPOA2 (see Figure 2.8) in one of two ways: the two regions might interact directly, burying otherwise aggregation prone surfaces, or the two regions might not interact directly, and the presence of SPOA2 solubilizes SPOA1 simply by virtue of its own high solubility. I hypothesized the former to be the case — that SPOA2 directly interacts with and solublizes SPOA1, perhaps in a manner structurally analogous to SPOA2 homodimer formation. Consistent with this hypothesis, the SPOA1-containing construct SpaO(145-213) could be co-refolded with SPOA2 (SpaO(232-297), Figure 2.12). Anion-exchange chromatography allowed the separation of the SPOA1-SPOA2 complex (Figure 2.12, peak “i”) from excess free dimeric SPOA2 (Figure 2.12, peak “ii”), and the complex remained stable over size exclusion chromatography, despite lacking the polypeptide linker that normally connects the two regions covalently.

One might hypothesize that the disordered region between the SPOA domains (predicted by bioinformatics and supported by NMR chemical shift analysis) might have previously been hindering crystallization, by virtue of its high conformational entropy. Thus, I hypothesized that this linker-free SPOA1-SPOA2 complex might be more amenable to crystallization than SpaO(140-297). Indeed, the SpaO(145-213) + SpaO(232-297) complex crystalized, and its structure was determined to 2.9 Å resolution (Figure 2.13, Table 2.2). Compared to the SPOA2 homodimer crystals that diffracted to high resolution, the more weakly diffracting SPOA1-SPOA2 crystals had a much higher solvent content (44% and 62%).

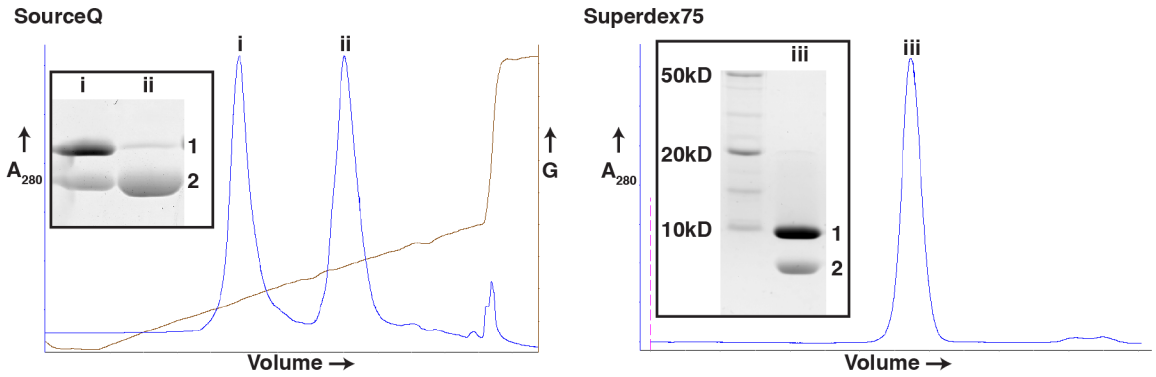


Figure 2.12: Purification of SpaO(145-213) + SpaO(232-297). Anion exchange (SourceQ, left) chromatographic purification of co-refolded SpaO(145-213) + SpaO(232-297). The blue trace shows UV absorbance at 280 nm and the brown trace shows conductance. Inset shows a coomassie-stained gel of the main elution peaks (i, and ii). "1" indicates SpaO(145-213) and "2" indicates SpaO(232-297). Peak "i" from the SourceQ elution was isolated from peak "ii" and further purified by size exclusion chromatography (Superdex75, right). Inset abbreviations are the same as those in the inset at left.

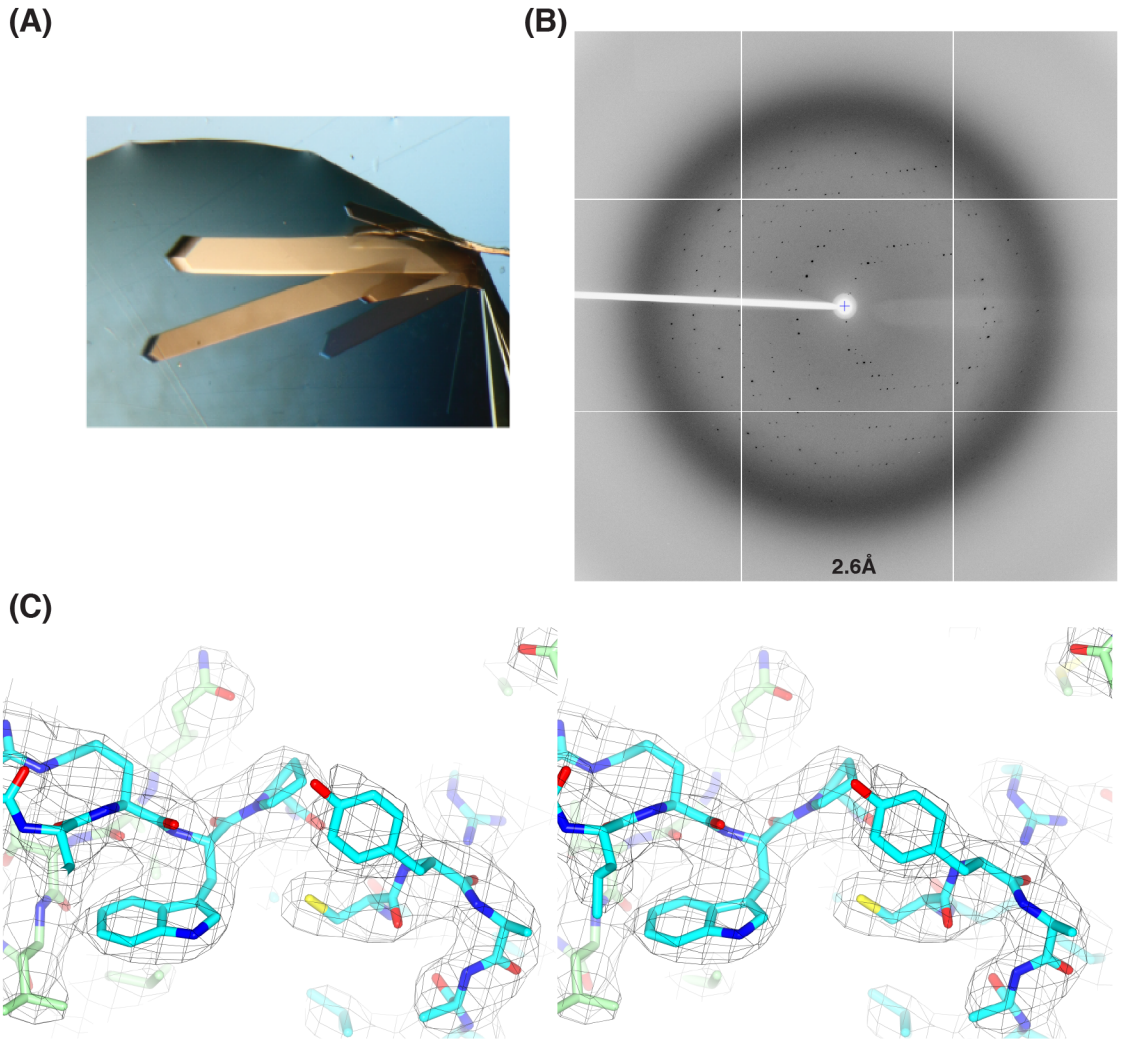


Figure 2.13: Diffraction analysis of SpaO(145-213) + SpaO(232-297). (A) Crystals of SpaO(145-213) + SpaO(232-297). (B) Representative diffraction image of SpaO(145-213) + SpaO(232-297) crystals with the edge resolution indicated. (C) Stereoimage of a selected region of the 2Fo-Fc electron density map for the 2.9 Å SpaO(145-213) + SpaO(232-297) structure. The map is contoured to 1σ and clipped to within 2 Å of the peptide for clarity.

Table 2.2: Diffraction and refinement statistics for SpaO(145-213) + SpaO(232-297).

		SeMet	Native
PDB ID			4YX5
Data Collection	**Space group**	$P4_12_12$	$P4_12_12$
	Cell: a, b, c (Å)	66.38, 66.38, 95.21	65.76, 65.76, 95.65
	Cell: α, β, γ (°)	90, 90, 90	90, 90, 90
	Resolution (Å)	46.94–3.00 (3.18–3.00)	38.68–2.9 (3.08–2.9)
	R_{merge}	0.221 (1.463)	0.166 (1.447)
	I/σI	11.8 (3.1)	14.4 (2.7)
	$CC_{1/2}$	0.989 (0.839)	0.996 (0.856)
	Completeness (%)	100 (100)	99.3 (99.3)
	No. of reflections	4,653	4,964
	Redundancy	25.6 (27.2)	24.6 (26.2)
Refinement	**R_{work}/R_{free}**		0.2085/0.2795
	No. of atoms		1,024
	Protein		1,023
	Ligand/ion		1
	Water		0
	B factors: Protein		74.20
	B factors: Ligand/ion		105.00
	B factors: Water		
	Geometry (r.m.s.d.)		
	Bond lengths (Å)		0.010
	Bond angles (°)		1.33
	[Φ,Ψ] Favored (%)		89
	[Φ,Ψ] Allowed (%)		97
	[Φ,Ψ] Outliers (%)		3

Structure determination confirmed the presumptive SPOA1 region as having a bond fide SPOA fold. SPOA1 and SPOA2 form a distinct, heterotypic SPOA-SPOA interaction with an overall topology similar to that observed in SPOA2 homodimers and a similar interfacial area (1799 $Å^2$). The SPOA1 backbone follows that of the prototypical SPOA fold, retaining the antiparallel beta-sheet floor and fingers-to-palm architecture (Figure 2.14A,B).

SPOA1-SPOA2 and the SPOA2 homodimer superpose with 2.47 Å r.m.s.d. (Figure 2.14C). The overall topology of SPOA2 in association with SPOA1 is grossly similar to that seen in the homodimer (Figure 2.13D, r.m.s.d. 1.67 Å), but with some conformational alteration. In context, each thumb in the SPOA1-SPOA2 complex is rotated clockwise towards its adjacent fingers relative to the thumb positions in the SPOA2 homodimer (Figure 2.14E,F), creating a shallower, less tall structure. To accommodate the SPOA1 thumb in this orientation, the SPOA2 fingers are displaced relative to their position in the homodimer. This displacement is the largest for the amide nitrogen of L256 (5.7 Å, Figure 2.14F).

To confirm that SPOA1 and SPOA2 interact outside of the solid-phase crystalline environment, I re-evaluated previously acquired NMR data in the context of the crystal structure. A *post hoc* analysis of the ^{15}N-NOESY-HSQC spectrum for SpaO(140-297) revealed long-range amide proton correlations between SPOA1 and SPOA2 (Figure 2.15), consistent with association of the two SPOAs in solution.

Figure 2.14: Structure of the SpaO SPOA1-SPOA2 interaction. (A) Ribbon diagram of the SpaO(145-213) + SpaO(232-297) crystal structure. SpaO(145-213) is shown in cyan and SpaO(232-297) is shown in green. The structure is annotated as SPOA diagrams previously shown; asterisk indicates the β-sheet "floor." (B) Topology diagram for the structure in (A) generated using Pro-Origami. (C) SpaO(145-213, cyan) + SpaO(232-297, green) superposed on the SpaO(232-297, gray) homodimer. (D) SpaO SPOA2 from the SPOA1-SPOA2 model (green) superposed in isolation on one protomer of the SPOA2 homodimer model (gray). (E,F) Worm diagrams derived from the alignment in (C) highlighting the displacement of the SPOA1 and SPOA2 thumbs relative to the thumb positions in the homodimer (orange arrows). The maximum displacement in the fingers region of SPOA2 is highlighted in red.

Figure 2.14: Structure of the SpaO SPOA1-SPOA2 interaction.

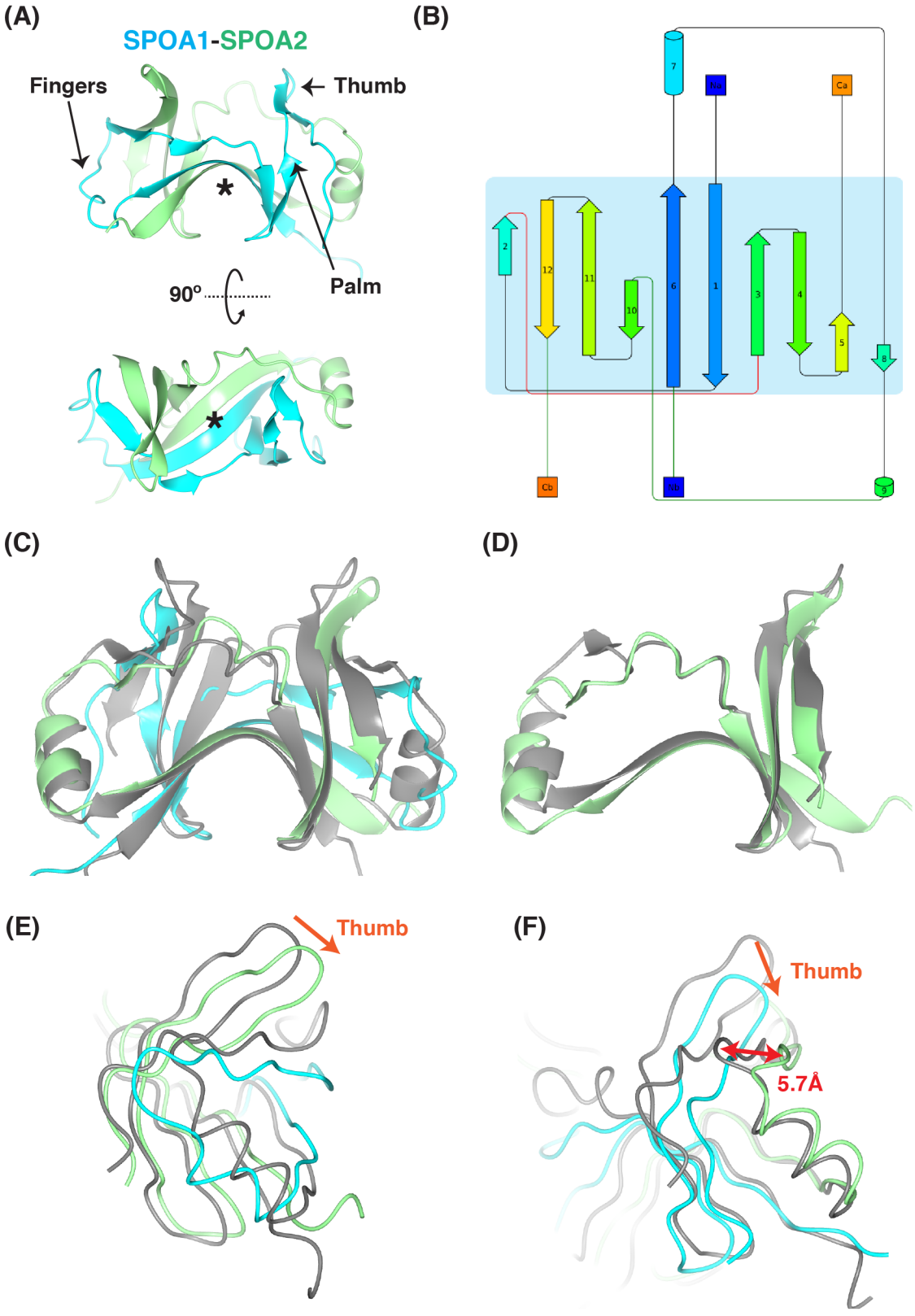
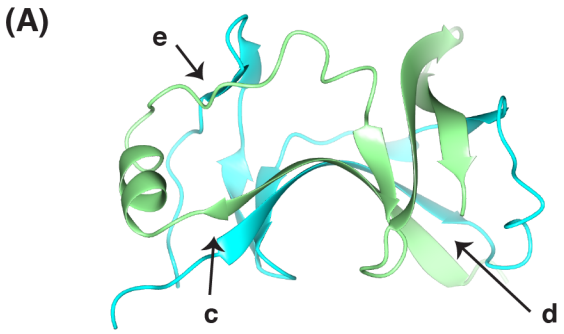


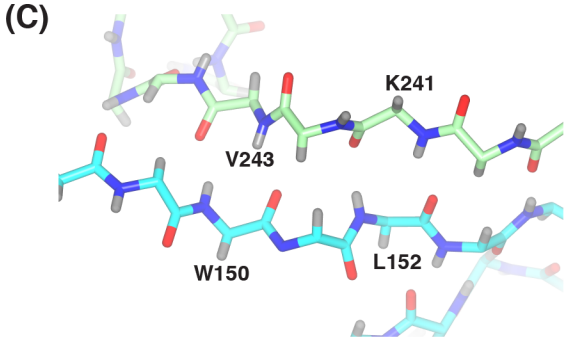
Figure 2.15: NOESY data support a SPOA1-SPOA2 interaction. (A) Schematic illustration of the SPOA1-SPOA2 crystal structure with regions for further analysis indicated. (B) Protons were modeled on the SPOA1-SPOA2 structure using ReadySet in Phenix. Inter-proton distances are specified for selected inter-domain long-range amide proton correlations observed in the ^{15}N-NOESY-HSQC spectrum. (C-E) Protonated backbone models for the regions specified in (A) are shown at left with residues of interest noted. In the ^{15}N-NOESY-HSQC spectra at right, the diagonal-peaks are indicated by red asterisks and the cross-peaks by red dotted lines.

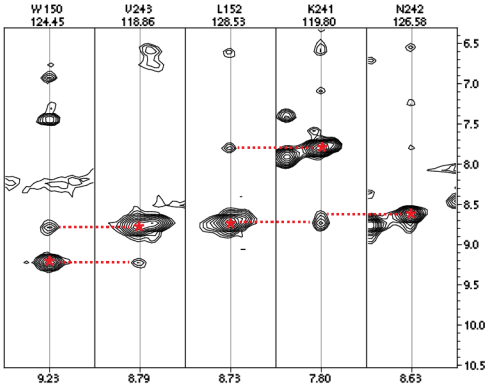
Figure 2.15: NOESY data support a SPOA1-SPOA2 interaction.

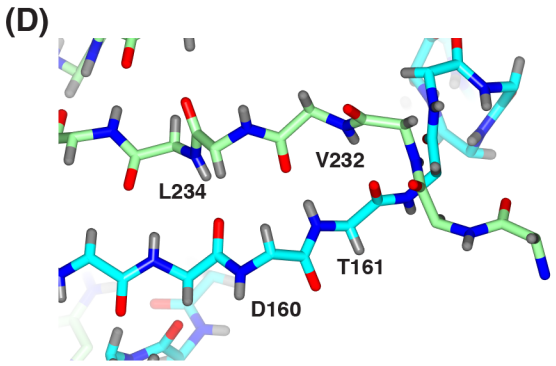


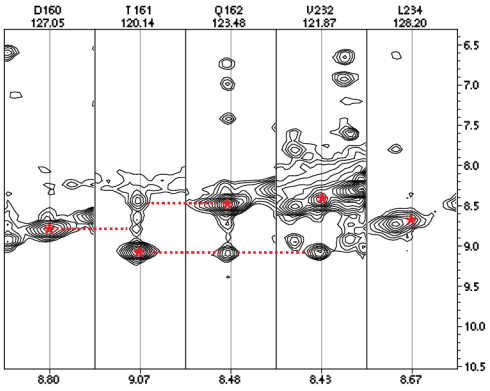
(B)

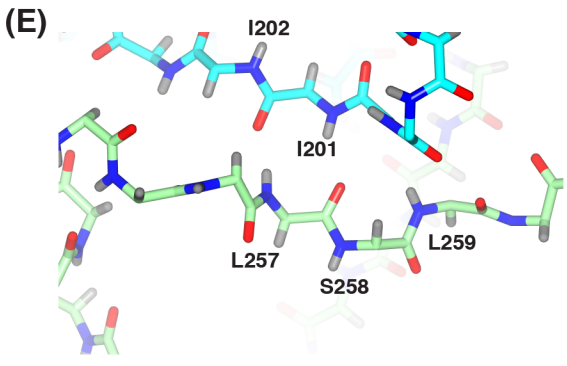
SPOA1	SPOA2	Dist.(Å)
W150	V243	2.7
L152	K241	3.1
T161	V232	3.5
I201	L257	2.8
I201	L259	3.9

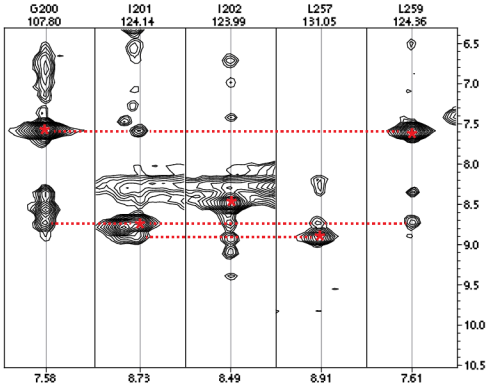












2.5 Do SPOA1-SPOA2 interactions facilitate SpaO oligomerization?

The structural characterization of the SPOA1-SPOA2 interaction suggests a potential mechanism for the higher order assembly of a high molecular weight sorting platform: interacting protein regions covalently linked in a polypeptide chain can interact in *cis* (and remain monomeric) or can form higher order oligomers by interacting in *trans* (Figure 2.16). Given the less than 20 residue linker connecting SPOA1 and SPOA2, they would experience a low millimolar-range relative concentration and would likely interact in an intramolecular fashion. However, at high local SpaO concentrations in association with the T3SS, intermolecular heterotypic SPOA interactions might dominate, perhaps forming fibrils or annuli and explaining the apparent oligomeric nature of the sorting platform. Indeed, a similar model of intermolecular domain swapping was recently suggested for the ring-forming injectisome protein PrgK (45).

Hypothetical SpaO oligomerization driven by intermolecular heterotypic SPOA interactions would be dependent on the covalent linkage of SPOA1 and SPOA2. Dynamic light scattering analysis of co-refolded SPOA1 and SPOA2 revealed a hydrodynamic radius (22 Å ± 1 Å) and predicted molecular weight (21 kDa ± 3 kDa) most consistent with a single heterodimer (expected molecular weight: 16 kDa), as would be expected in the absence of a covalent linker (Figure 2.16). Similarly, even at 1mM concentration, SpaO(140-297) was monodispersed and non-oligomeric (hydrodynamic radius: 24 Å ± 1 Å; calculated mass: 26 kDa ± 3 kDa; expected mass: 18 kDa), suggesting that the presence of the linker does not promote intermolecular SPOA-SPOA interactions and

oligomerization (Figure 2.16). The small increase in hydrodynamic radius relative to co-refolded SPOA1-SPOA2 is most likely the result of the unstructured linker projecting from the globular core, and is not consistent with high-order oligomerization, where one would expect a very large increase in hydrodynamic radius and calculated molecular weight. This finding is also consistent with the size exclusion chromatography profile of SpaO(140-297), which shows a lack of aggregation (Figure 2.9).

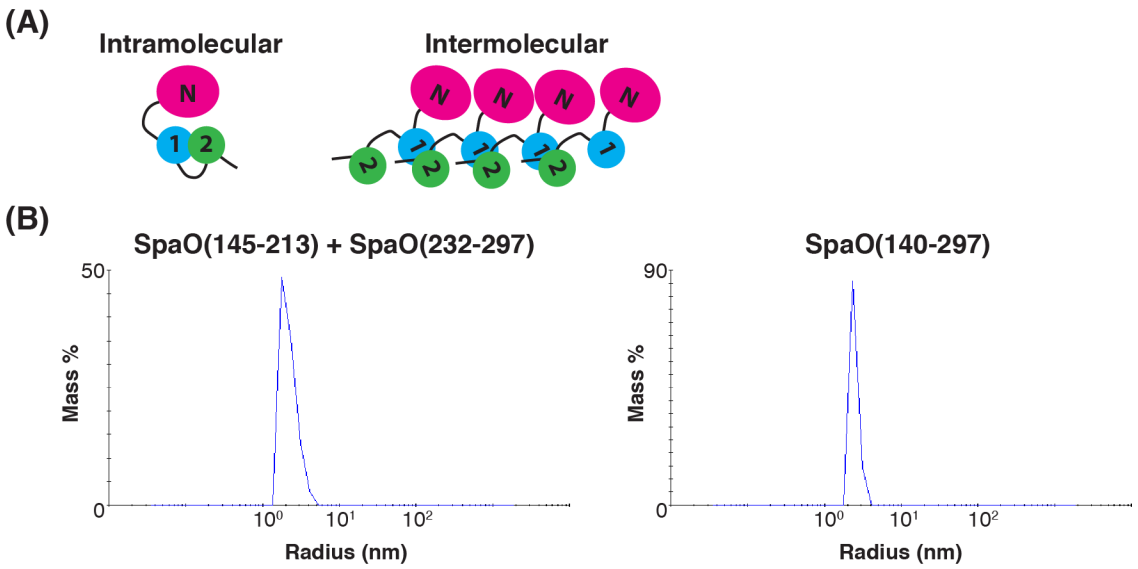


Figure 2.16: Hypothetical models of SPOA interactions in *cis* and *trans*. (A) Schematic models for putative intra- and intermolecular SPOA1-SPOA2 interactions and their implications for the SpaO oligomerization state. 1 and 2 indicate the SpaO SPOA1 and SPOA2, respectively, and N indicates the SpaO amino-terminal domain(s). (B) DLS analysis of SPOA1-SPOA2 with (right) and without (left) the inter-SPOA linker. Regularization graph shown.

However, one might counter that perhaps some *in vivo* factor promotes intermolecular SPOA-SPOA interactions (e.g. association with the T3SS basal body or export apparatus). Thus, I tested whether genomic deletion of the SpaO amino-terminal domain and SPOA1 can be complemented in trans, as assayed by *S. typhimurium* T3SS secretory function. When grown under T3SS stimulating conditions, the culture supernatant of *S. typhimurium* has a stereotyped protein composition, consisting of both flagellar and injectisome secretion substrates (Figure 2.17; secreted proteins are annotated as per Aizawa and colleagues (87)). Deletion of spaO results specifically in the loss of injectisome-dependent secretory products from the culture supernatant, and deletion of spaO codons 1-203 phenocopies spaO deletion, indicating that the SpaO amino-terminal domain(s) and/or SPOA1 are necessary for T3SS function (Figure 2.17). Because SpaO(1-219) is able to complement the deletion of spaO codons 1-203 (Figure 2.17, red asterisks), the covalent linkage of SPOA1 and SPOA2 is not necessary for T3SS function. While this does not rule out the presence of intermolecular heterotypic SPOA interactions *in vivo*, if they do occur, they are not explicitly necessary for secretion. It should be noted that SpaO(1-219) does not complement a full genomic deletion of spaO, demonstrating that SPOA2 is also necessary for T3SS function (Figure 2.17). Similarly, insertion of a double stop codon after spaO codon 219 abrogates T3SS (Figure 2.17). The successful expression of all SpaO genomic domain deletions and complementation plasmids was confirmed by western blot (Figure 2.18).

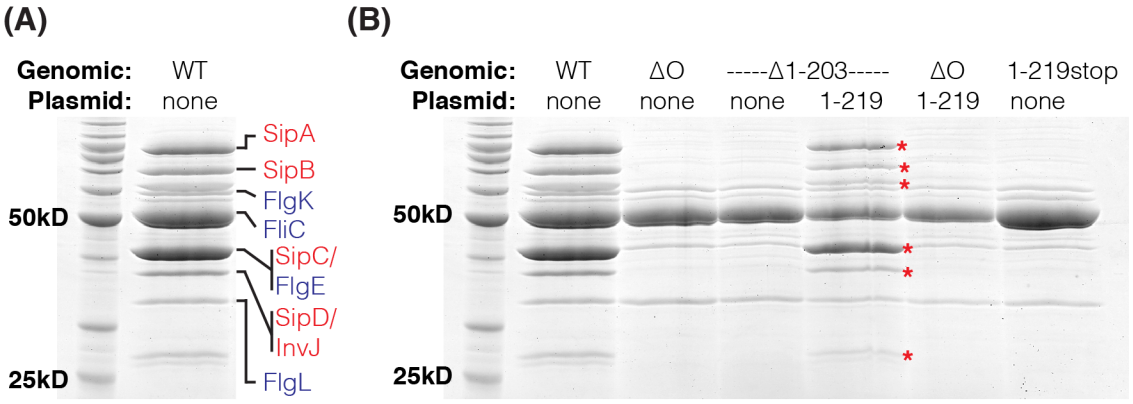


Figure 2.17: Covalent linkage of SPOA1 and SPOA2 is not necessary for T3SS function. (A) Coomassie-stained SDS-PAGE of *S. typhimurium* culture supernatants grown under T3SS stimulating conditions (0.3M NaCl, strain SB1741). Bands previously identified by Aizawa and colleagues (87) are noted and color-coded by T3SS subtype – injectisome in red, flagellar in blue. (B) Coomassie-stained SDS-PAGE of *S. typhimurium* culture supernatants grown under T3SS stimulating conditions. Red asterisks indicate injectisome-specific secretion substrates. Abbreviations: WT, wild-type; ΔO, deletion of spaO; Δ1-203, deletion of spaO codons 1-203; 1-219, complementation with SpaO(1-219); 1-219stop, insertion of two stop codons following spaO codon 219. SpaO was 3xFLAG tagged at its amino terminus in each *S. typhimurium* strain (except ΔO) and complementation construct. The first two lanes of (A) and (B) are from the same image, truncated in (A) to allow annotation.

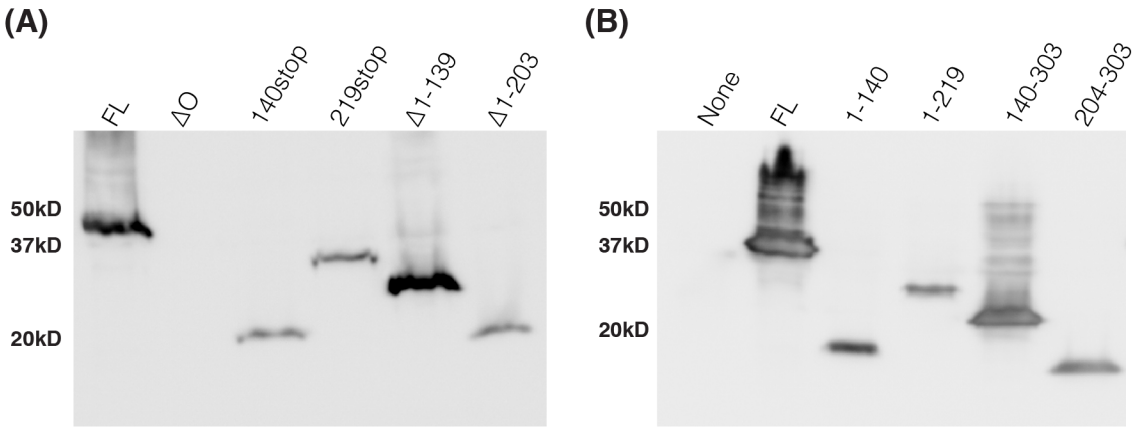


Figure 2.18: SpaO deletion mutants and complementation plasmids. (A, B) Anti-FLAG western blots of SDS-PAGE separated *S. typhimurium* whole cell lysates. (A) Strains carrying amino-terminal 3xFLAG fusions to SpaO with the indicated stop codon insertions or domain deletions. ΔO denotes a strain carrying a full deletion of SpaO, which lacks a 3xFLAG insertion and serves as a negative control. (B) ΔO strain complemented with the indicated SpaO expression plasmids. All expression constructs are amino-terminal 3xFLAG fusions.

CHAPTER 3:

SpaO SPOA1-SPOA2 FORMS A PLATFORM FOR THE OrgB APAR

If heterotypic SPOA1-SPOA2 interactions do not function to create higher order SpaO oligomers, what might be the function of this novel type of SPOA-SPOA interaction? I hypothesized that this interaction might function in sorting platform assembly by interacting with other sorting platform components. This chapter describes the identification and structural characterization of a binding partner for SpaO SPOA1-SPOA2.

3.1 Partial reconstitution of the SPI-1 sorting platform

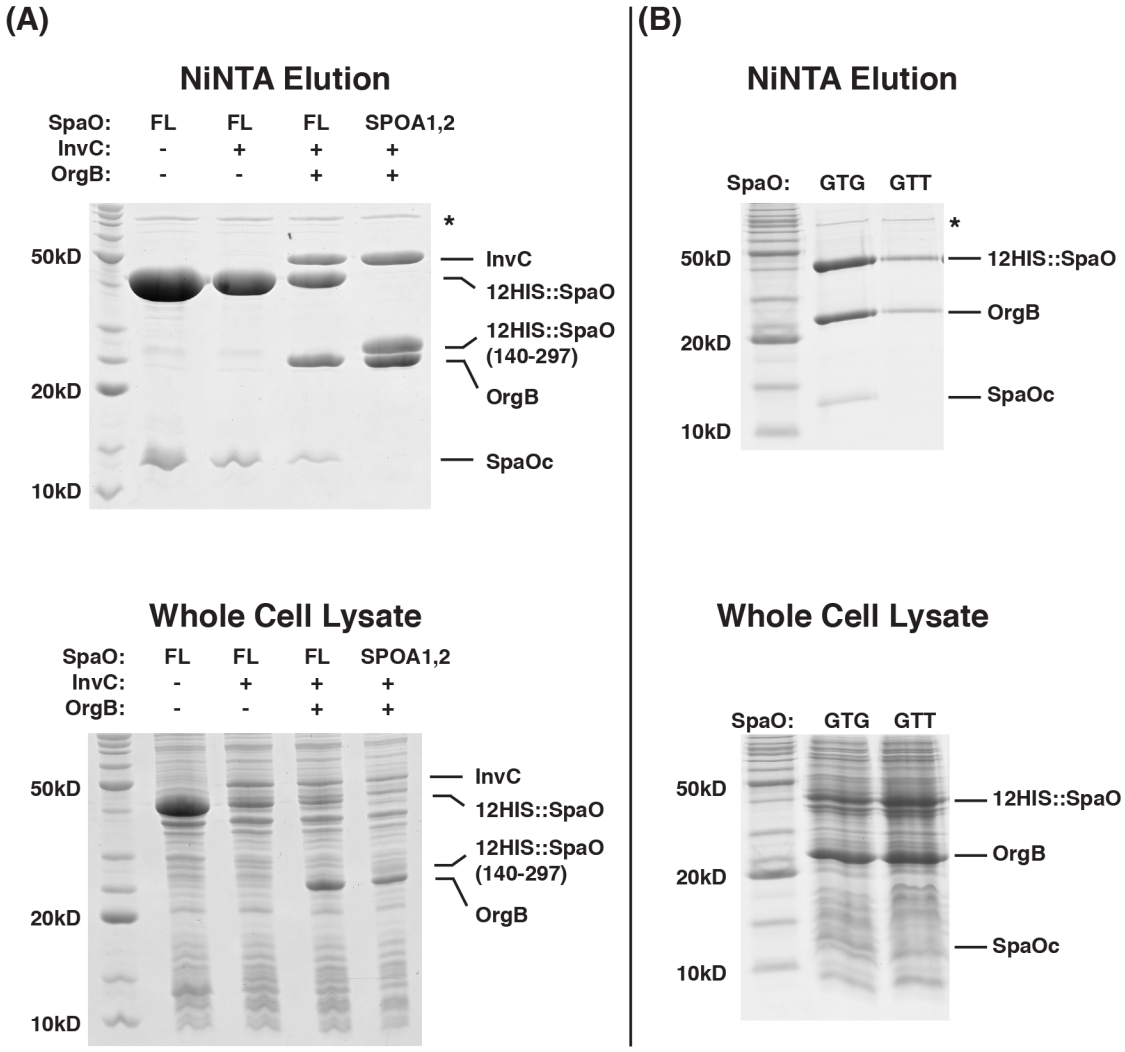
Seminal work by Galán and colleagues (129) combined co-immunoaffinity purification of SpaO and its binding partners from *S. typhimurium* with mass spectrometric proteomic analysis to identify the majority constituents of the SPI-1 T3SS sorting platform. The proteins InvC (SctN, ATPase), OrgB (SctL, “ATPase regulator”), and OrgA (SctK, function unknown) robustly co-purified with SpaO. To probe these interactions more precisely, I attempted to reconstitute the core sorting platform by heterologous coexpression in *E. coli.*

Double hexahistidine-tagged SpaO is able to co-affinity purify the sorting platform components OrgB and InvC when co-expressed in *E. coli* (Figure 3.1A). Formation of the SpaO-OrgB-InvC termary complex is OrgB-dependent, as SpaO alone is insufficient to co-affinity purify InvC (Figure 3.1A). In fact, SpaO interacts directly with OrgB, as InvC is dispensible for SpaO-OrgB complex formation

(Figure 3.1B). Expression of the cryptically translated SpaO carboxy-terminal fragment is not necessary for SpaO-OrgB interaction, as SpaO(Val_{GTT}203) was sufficient to co-affinity purify OrgB (Figure 3.1B). Interestingly, as with full length SpaO alone, loss of the free, SPOA2-containing carboxy-terminal fragment resulted in substantially lower soluble yields.

Figure 3.1: SpaO interacts directly with OrgB to form a SpaO-OrgB-InvC ternary complex. (A) Coomassie-stained gel of protein elution from NiNTA resin (top) following coexpression of double hexahistidine (12HIS)-tagged SpaO ("FL," full length) or SpaO(140-297, Val_{GTT}203) with the indicated sorting platform constituents (InvC and/or OrgB). Below, whole cell lysates from the same preparations allow confirmation of the intended constituents' expression. Asterisk denotes nonspecific co-purifying E. coli proteins, likely chaperones. SpaOc indicates the cryptically expressed SPOA2-containing carboxy-terminal fragment. (B) Coomassie-stained gel of protein elution from NiNTA resin (top) following coexpression of OrgB and double hexahistidine (12HIS)-tagged, full length SpaO with the indicated codon 203 genotype. Below, whole cell lysates from the same preparations allow confirmation of the intended constituents' expression. Asterisk denotes nonspecific co-purifying E. coli proteins, likely chaperones. SpaOc indicates the cryptically expressed SPOA2-containing carboxy-terminal fragment.

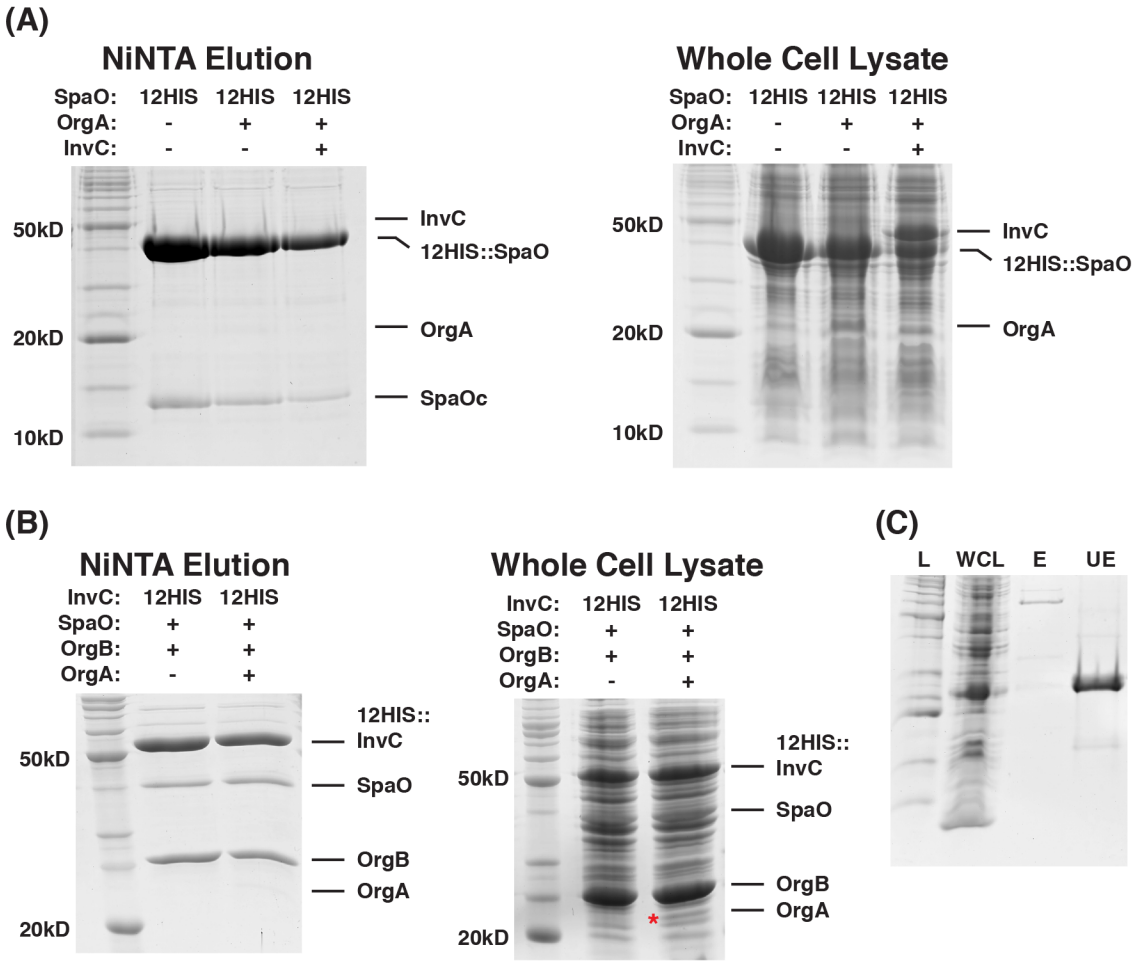
Figure 3.1: SpaO interacts directly with OrgB to form a SpaO-OrgB-InvC ternary complex.



Coexpression of double hexahistidine-tagged SpaO with OrgA did not result in co-affinity purification of the two species and presumptive OrgA expression was particularly low (Figure 3.2A). Additional coexpression of InvC or InvC-OrgB did not enhance OrgA expression levels or result in co-affinity purification of OrgA (Figure 3.2B). OrgA could be robustly expressed in *E. coli* as a double hexahistidine fusion and purified under denaturing conditions (Figure 3.2C), but could not be solubly refolded.

Figure 3.2: Attempts to incorporate OrgA into SpaO containing complexes were unsuccessful. (A) Coomassie-stained gel of protein elution from NiNTA resin (left) following coexpression of double hexahistidine (12HIS)-tagged SpaO with the indicated sorting platform constituents (InvC and/or OrgA). At right, whole cell lysates from the same preparations allow confirmation of the intended constituents' expression. SpaOc indicates the cryptically expressed SPOA2-containing carboxy-terminal fragment. (B) Coomassie-stained gel of protein elution from NiNTA resin (left) following coexpression of double hexahistidine (12HIS)-tagged InvC with the indicated sorting platform constituents. At right, whole cell lysates from the same preparations allow confirmation of the intended constituents' expression. Red asterisk denotes presumptive OrgA expression. (C) OrgA can be robustly overexpressed in *E. coli* as a double hexahistidine fusion, but it is not soluble when purified under native conditions. L, ladder; WCL, whole cell lysate; E, NiNTA elution after purification under native conditions; UE, NiNTA elution after purification under denaturing conditions.

Figure 3.2: Attempts to incorporate OrgA into SpaO containing complexes were unsuccessful.

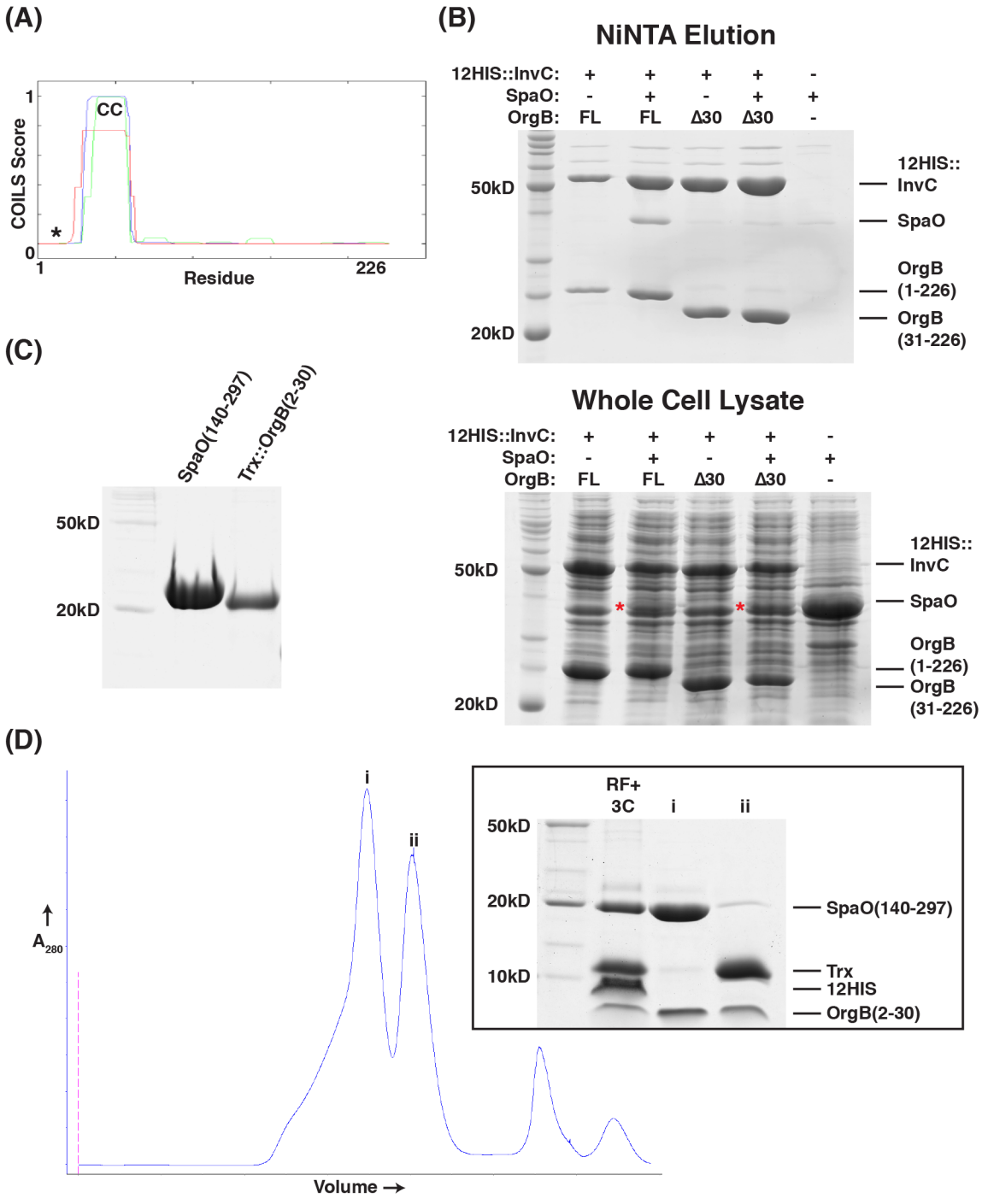


Given the biochemical intractability of OrgA, I decided to focus my efforts on the SpaO-OrgB-InvC subset of the sorting platform. I hypothesized that the SPOA1-SPOA2 module might serve as a scaffold for the interaction of SpaO with OrgB-InvC. Indeed, SpaO(140-297) is sufficient to co-affinity purify OrgB-InvC (Figure 3.1A). This construct also contains the $Val_{GTT}203$ mutation, further demonstrating that the SPOA2 homodimer is dispensable for SpaO-OrgB-InvC complex formation.

OrgB and its homologues are predicted to share a common amino-terminal organization: a disordered region followed by a coiled coil (Figure 3.3A). In the flagellar system, the unstructured region at the amino terminus of the OrgB homologue FliH is necessary for its interaction with the SpaO homologues FliM and FliN (133). Similarly, deletion of the pre-coiled-coil residues (1-30) of OrgB prevented the formation of the SpaO-OrgB-InvC ternary complex (Figure 3.3B). Of note, these residues were dispensable for the interaction of OrgB with InvC (Figure 3.3B), consistent with the work of Dr. Mirjana Lilic (Laboratory of Structural Microbiology, Rockefeller University) and previously published yeast two hybrid data from *Shigella* (182). The amino-terminal 30 residues of OrgB were by themselves sufficient to interact with SPOA1-SPOA2 of SpaO. OrgB(1-30) could be expressed as a fusion to human thioredoxin and co-refolded with SpaO(140-297). After cleavage of the thioredoxin carrier from the OrgB peptide, OrgB(2-30) co-eluted with SpaO(140-297) from a size exclusion column (Figure 3.3D).

Figure 3.3: The OrgB amino terminus is necessary and sufficient to mediate the interaction with SpaO SPOA1-SPOA2. (A) COILS analysis of OrgB. Likelihood of coiled coil formation is assessed by the program COILS. Different scanning window sizes are indicated by line color: 14 residues, green; 21 residues, blue; 28 residues, red. CC indicates the presumed coiled coil region. Asterisk highlights the region amino-terminal to the CC. (B) Coomassie-stained gel of protein elution from NiNTA resin (top) following coexpression of double hexahistidine (12HIS)-tagged InvC with the indicated sorting platform constituents. FL denotes full length OrgB(1-226) and Δ30 denotes OrgB(31-226). Below, whole cell lysates from the same preparations allow confirmation of the intended constituents' expression. Red asterisk denotes SpaO expression. (C) SpaO(140-297) and Trioredoxin(Trx)::OrgB(2-30) purified under denaturing conditions for co-refolding. (D) Co-refolded, 3C protease-cut SpaO(140-297) in complex with OrgB(2-30) can be separated from free Trx and affinity tags by size exclusion chromatography. RF+3C indicates co-refolded material after 3C cleavage, but before further purification. The main size exclusion chromatography peaks are denoted i and ii.

Figure 3.3: The OrgB amino terminus is necessary and sufficient to mediate the interaction with SpaO SPOA1-SPOA2.



3.2 Structure of the SpaO(SPOA1-SPOA2)-OrgB(APAR) complex

Herein, I will refer to the SpaO-binding region at the amino terminus of OrgB and its homologues as the **A**daptor **P**eptide of the **A**TPase **R**egulator (APAR). What is the structural basis for this interaction? In order to characterize the SPOA1-SPOA2-APAR interaction by X-ray crystallography, I attempted to crystallize complexes of the OrgB APAR peptide and SpaO(SPOA1-SPOA2), but they were resistant to crystallization.

Fusion of readily crystallizable proteins to target proteins has been used to enhance target crystallization (183), most notably in the case of T4 lysozyme fusions to various GPCRs (184, 185). To explore whether this strategy might promote crystallization of the SpaO-OrgB complex, I fused the pre-coiled-coil region of OrgB (residues 1-30) to T4 lysozyme and found that it still bound to SpaO(140-297) (Figure 3.4A). Consistent with the finding that the SPOA2 homodimer was not necessary for the SpaO-OrgB interaction, OrgB(1-30)::T4 lysozyme did not pull down SpaO SPOA2 homodimers (Figure 3.4A). The Org(1-30)::T4 lysozyme fusion formed a complex with SpaO(145-213) + SpaO(232-297) that was stable over size exclusion chromatography (Figure 3.4B), and this complex was subjected to crystallization screening. The SpaO(SPOA1-SPOA2)-OrgB(APAR)::T4 lysozyme complex was crystallized, its structure solved by molecular replacement using the SpaO(145-213) + SpaO(232-297) model and a known T4 lysozyme model (PDB 2LZM), and the structure refined to 2.0 Å resolution (Figure 3.5 and Table 3.1).

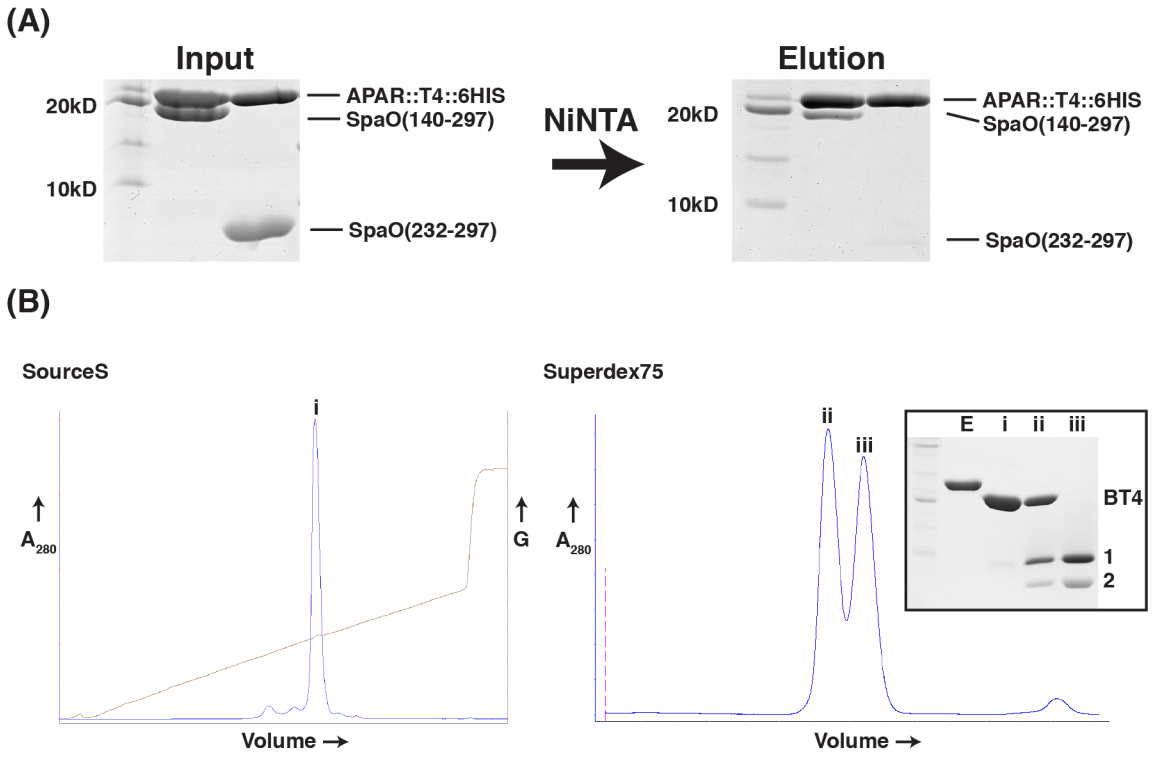


Figure 3.4: Expression and purification of SpaO(145-213) + SpaO(232-297) + OrgB(1-30)::T4lysozyme. (A) OrgB(1-30) fused to hexahistidine-tagged T4 lysozyme (APAR::T4::6HIS) was mixed with the indicated SpaO constructs and passed over NiNTA resin. Coomassie-stained gels of the input and imidazole elution are shown. (B) Purification of the SpaO(145-213) + SpaO(232-297) + OrgB(1-30)::T4 lysozyme complex. OrgB(1-30)::T4 lysozyme (BT4) purified under denaturing conditions (E) was refolded and purified by cation exchange chromatography (i) after removal of affinity tags by 3C protease. Cation exchange purified material was complexed with an excess of SpaO(145-213) + SpaO(232-297) ("1" and "2", respectively) and the ternary complex (ii) separated from excess SpaO(145-213) + SpaO(232-297) (iii) by size exclusion chromatography.

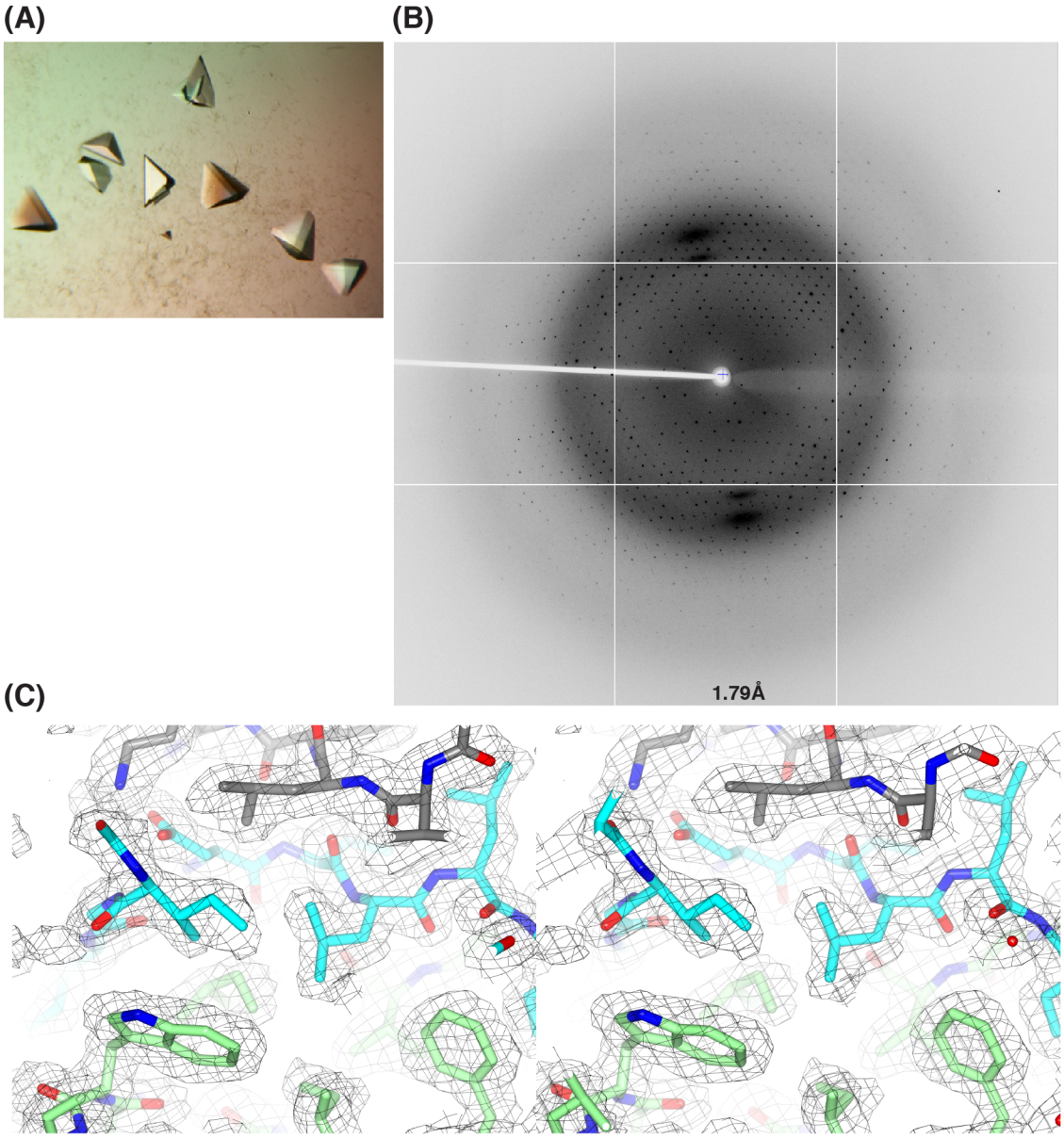


Figure 3.5: Diffraction analysis of SpaO(145-213) + SpaO(232-297) + OrgB(1-30)::T4lysozyme. (A) Crystals and (B) diffraction pattern for SpaO(145-213) + SpaO(232-297) + OrgB(1-30)::T4lysozyme. Edge resolution for diffraction image is indicated. (C) Stereoimage of a selected region of the 2Fo-Fc electron density map for the 2.0 Å SpaO(145-213, cyan) + SpaO(232-297, green) + OrgB(1-30)::T4 lysozyme (gray) structure. The map is contoured to 1σ.

Table 3.1: Diffraction and refinement statistics for SpaO(145-213) + SpaO(232-297) + OrgB(1-30)::T4lysozyme.

		Native	SeMet SpaO Native OrgB::T4
PDB ID		4YX7	4YXA
Data Collection	**Space group**	$P2_1$	$P2_1$
	Cell: a, b, c (Å)	62.092, 89.07, 62.092	62.88, 88.5, 63.32
	Cell: α, β, γ (°)	90, 114.94, 90	90, 116.07, 90
	Resolution (Å)	47.59–2.0 (2.05–2.0)	45.8–2.35 (2.43–2.35)
	R_{merge}	0.102 (0.530)	0.088 (0.617)
	I/σI	10.5 (3.2)	12.8 (2.6)
	$CC_{1/2}$	0.994 (0.828)	0.997 (0.816)
	Completeness (%)	99.5 (99.8)	99.0 (99.6)
	No. of reflections	41,183	25,740
	Redundancy	5.1 (5.2)	6.6 (6.6)
Refinement	**R_{work}/R_{free}**	0.1571/0.2096	0.1984/0.2618
	No. of atoms	5,769	4,940
	Protein	5,112	4,818
	Ligand/ion	0	0
	Water	657	122
	B factors: Protein	33.10	46.90
	B factors: Ligand/ion		
	B factors: Water	39.60	45.60
	Geometry (r.m.s.d.)		
	Bond lengths (Å)	0.008	0.011
	Bond angles (°)	1.16	1.46
	[Φ,Ψ] Favored (%)	94	89
	[Φ,Ψ] Allowed (%)	99.2	98.2
	[Φ,Ψ] Outliers (%)	0.8	1.8

The crystallographic asymmetric unit contained two copies of the complex in nearly identical conformations (superposed with 0.81 Å r.m.s.d., Figure 3.6), so only one constituent complex will be described in the following analyses. A cursory analysis of the crystal architecture highlights the importance of the construct modifications used to enhance crystallization. Deletion of the inter-SPOA linker peptide appears to have allowed favorable packing within the asymmetric unit: the two SPOA1-SPOA2 dimers pack against each other's "bottom" face, where the linker peptide might normally be positioned. Furthermore, the T4 lysozyme fusion chaperones form multiple packing interfaces with both SpaO and other T4 lysozyme molecules (Figure 3.6)

Within each constituent SpaO-OrgB complex, the OrgB APAR forms a lariat-like structure, contacting the thumb of SPOA2 and fingers of SPOA1 (Figure 3.7). OrgB makes substantial contact with both SPOA1 and SPOA2 of SpaO, burying 570 $Å^2$ against SPOA1 and 470 $Å^2$ against SPOA2, consistent with the finding that both SPOAs are required for the SpaO-OrgB(APAR) interaction (e.g. that SPOA2 alone is insufficient, Figure 3.4A). In the APAR-bound structure, there is little change in the conformation of SpaO (Figure 3.7, 1.01 Å r.m.s.d.). To confirm the crystallographic model, crystals grown with SeMet-doped SpaO (but native OrgB::T4 lysozyme) were subjected to diffraction analysis (Table 3.1). The structure was solved by molecular replacement using the all-native structure and anomalous scatterer positions were determined using ANODE (186). Six pronounced anomalous density peaks were observed in the asymmetric unit, spatially corresponding to the three structured Met residues in

each SPOA2 (Figure 3.8), further confirming the arrangement and stoichiometry of SPOA domains in the structure.

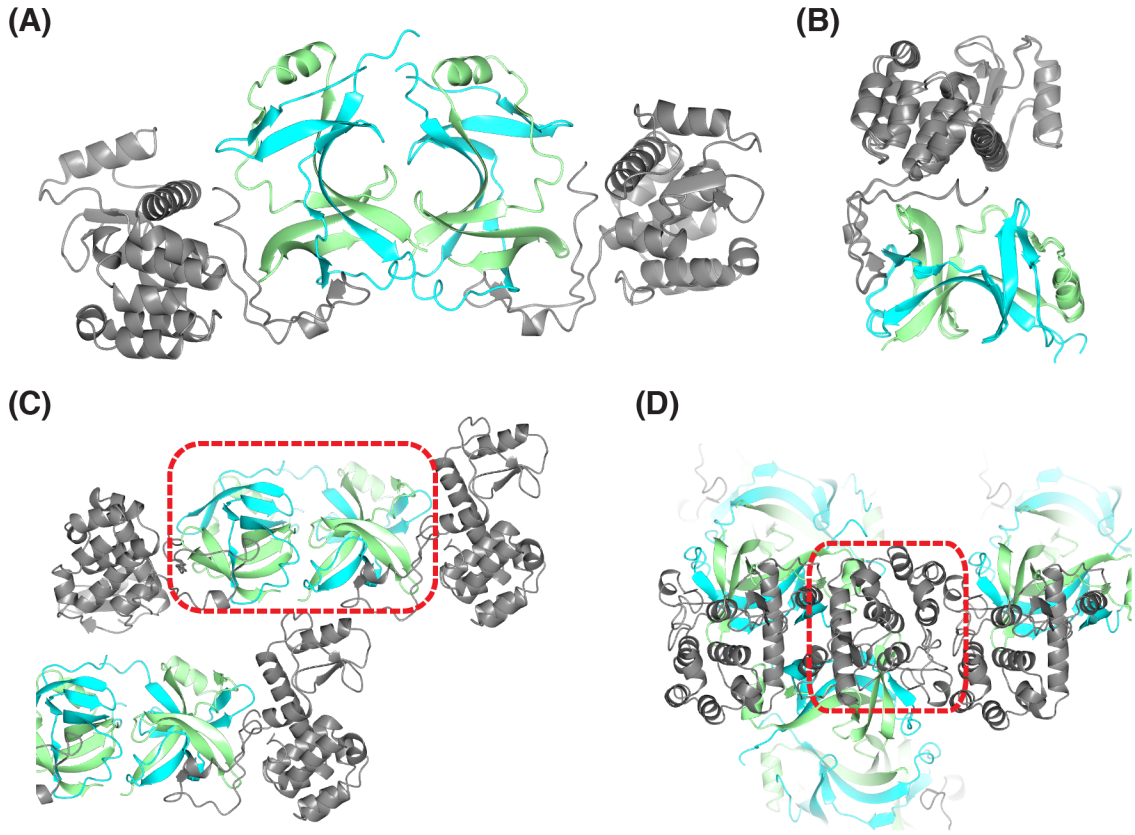


Figure 3.6: Asymmetric unit and crystal packing of the SpaO(SPOA-SPOA2)-OrgB(APAR) complex. (A) The asymmetric unit of the SpaO(145-213, cyan) + SpaO(232-297, green) + OrgB(1-30)::T4 lysozyme (gray) crystal contains two copies of the SpaO-OrgB complex. (B) Superposition of the two constituent complexes in the asymmetric unit of the crystal showing their architectural similarity (0.81 Å r.m.s.d.). (C,D) Qualitative analysis of crystal packing reveals that T4 lysozyme is involved in contacts with SpaO (C) and with itself (D). (C, red dashed box) Two copies of SpaO surrounded by T4 lysozyme molecules (3 shown). (D, red dashed box) T4 lysozyme molecule sandwiched between two other T4 lysozyme molecules.

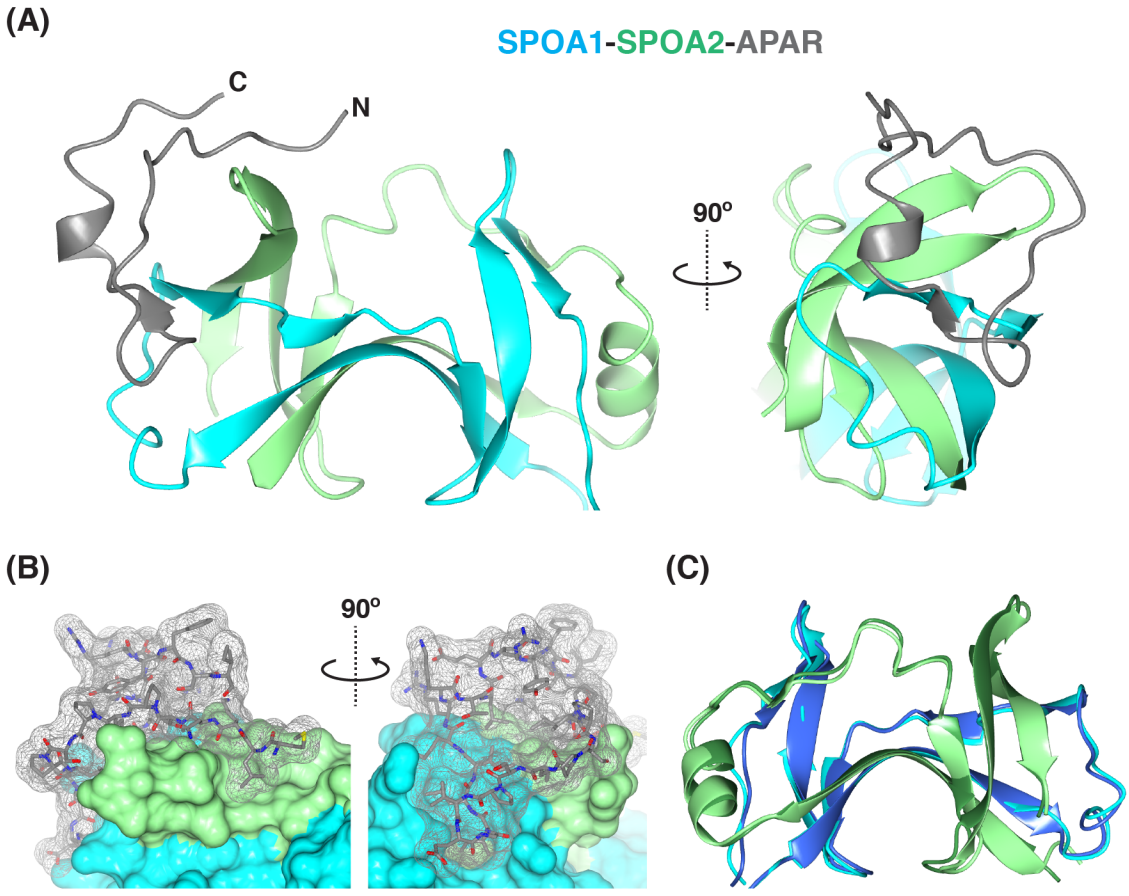


Figure 3.7: Crystal structure of the SpaO(SPOA1-SPOA2)-OrgB(APAR) complex. (A) Ribbon diagram of the SpaO-OrgB crystal structure. For simplicity, the T4 lysozyme crystallization chaperone has been omitted and only one of the two constituent complexes from the crystallographic asymmetric unit is shown. The amino- and carboxy-termini of the OrgB APAR are denoted as “N” and “C,” respectively. (B) Surface representation of the complex in (A). The OrgB APAR (gray mesh) contacts both SpaO SPOA1 (cyan) and SPOA2 (green). (C) Superposition of the apo- (cyan/light green) and APAR-bound (blue/dark green) forms of SpaO (1.01 Å r.m.s.d.).

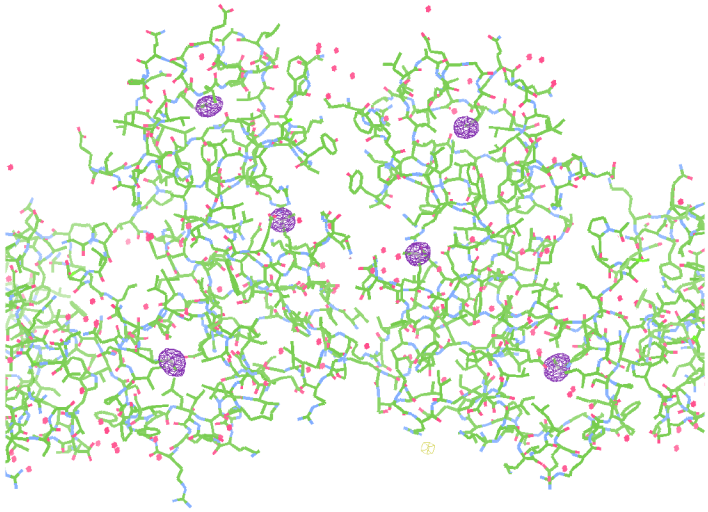


Figure 3.8: ANODE confirmation of SPOA positioning and stoichiometry in the SpaO(SPOA1-SPOA2)-OrgB(APAR) complex. ANODE analysis of the anomalous signal in the SpaO(SPOA1-SPOA2)-OrgB(APAR)::T4 lysozyme complex generated by SeMet substitution in SpaO (but not OrgB::T4 lysozyme). The anomalous signal (violet, contoured at 8 e/Å^3) localizes to the three structured SeMet in each copy of SpaO(232-297), providing empirical confirmation of the SpaO-OrgB model coordinates and stoichiometry. The image was created in COOT. Red crosses are structured solvent.

Independent NMR analyses of SpaO(140-297) bound to OrgB(APAR) in solution are consistent with the interface defined in the crystal (Figure 3.9). 151 out of 155 possible amide proton resonances were assigned for APAR-bound SpaO(140-297). Changes in amide proton chemical shift upon APAR binding were calculated as the magnitude of weighted Euclidean vectors (see Chapter 7). Compared to apo-SpaO, the largest chemical shift deviations (CSD) of backbone amide resonances in SpaO-OrgB occur in two major clusters along the SpaO

primary structure — one each in SPOA1 and SPOA2 (Figure 3.9). These residues map on the crystal structure as a single surface patch at the SpaO-OrgB interface (Figure 3.9). It should be noted that the NMR data suggest that the inter-SPOA linker is not directly involved APAR-binding (Figure 3.9), an assertion that was impossible to make using the linker-free crystallographic model.

Additional NMR analyses allowed me to rule out interactions between SpaO(SPOA1-SPOA2) and the APAR-adjacent coiled coil region of OrgB. The HSQC of SpaO(140-297) bound to OrgB(1-70), which contains both the APAR and coiled coil regions, was grossly similar to that of SpaO bound to the OrgB APAR alone (Figure 3.10A), suggesting that the OrgB coiled coil does not make additional contacts with SpaO upon APAR binding. The few differences between the OrgB(2-30)- and OrgB(1-70)-bound SpaO HSQCs map to the region of SpaO binding the OrgB amino terminus, which differed between these two constructs because of divergent cloning scars (Figure 3.10B). Indeed, the HSQC of SpaO bound to OrgB(1-70) was indistinguishable from that of SpaO bound to OrgB(1-30) fused to an inert carrier (the cyan fluorescent protein variant Cerulean), in which both OrgB amino-termini were identical Figure 3.10C).

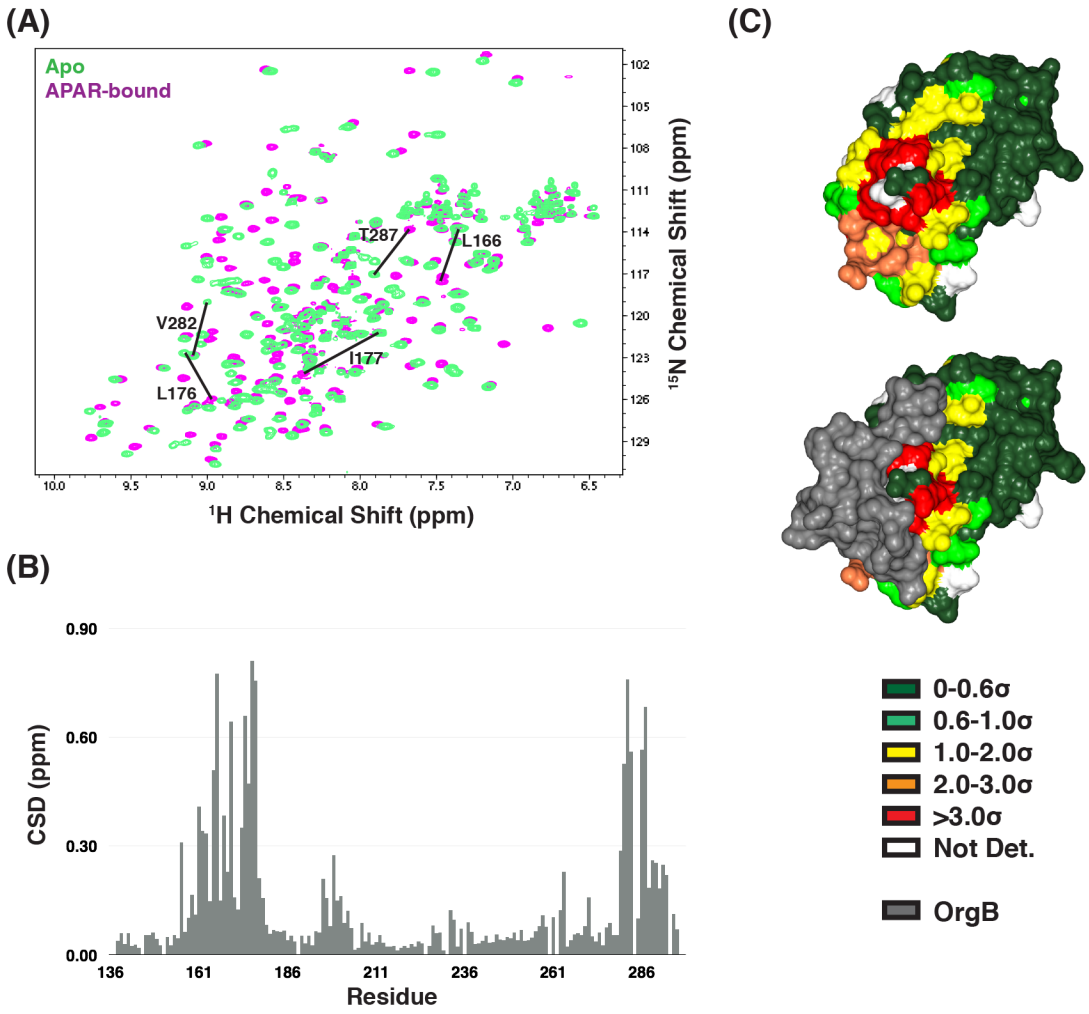


Figure 3.9: Solution NMR data support the crystallographic model of the SpaO(SPOA1-SPOA2)-OrgB(APAR) complex. (A) Overlayed ^{15}N-HSQC of apo- (green) and APAR-bound- (violet) SpaO(140-297). The five largest peak shifts are noted. (B) Chemical shift deviations (CSD) for apo- vs. APAR-bound-SpaO(140-297). Note that the unstructured linker is residues 214-231. (C) The solution interaction data from (A, B) are mapped onto the SpaO-OrgB crystal structure. Surface residues are color coded by the size of their weighted CSD in units of standard deviation. Residues not assigned an amide resonance in one of the two data sets are left white. The same view of SpaO is shown without (top) and with (bottom) the OrgB APAR (gray surface).

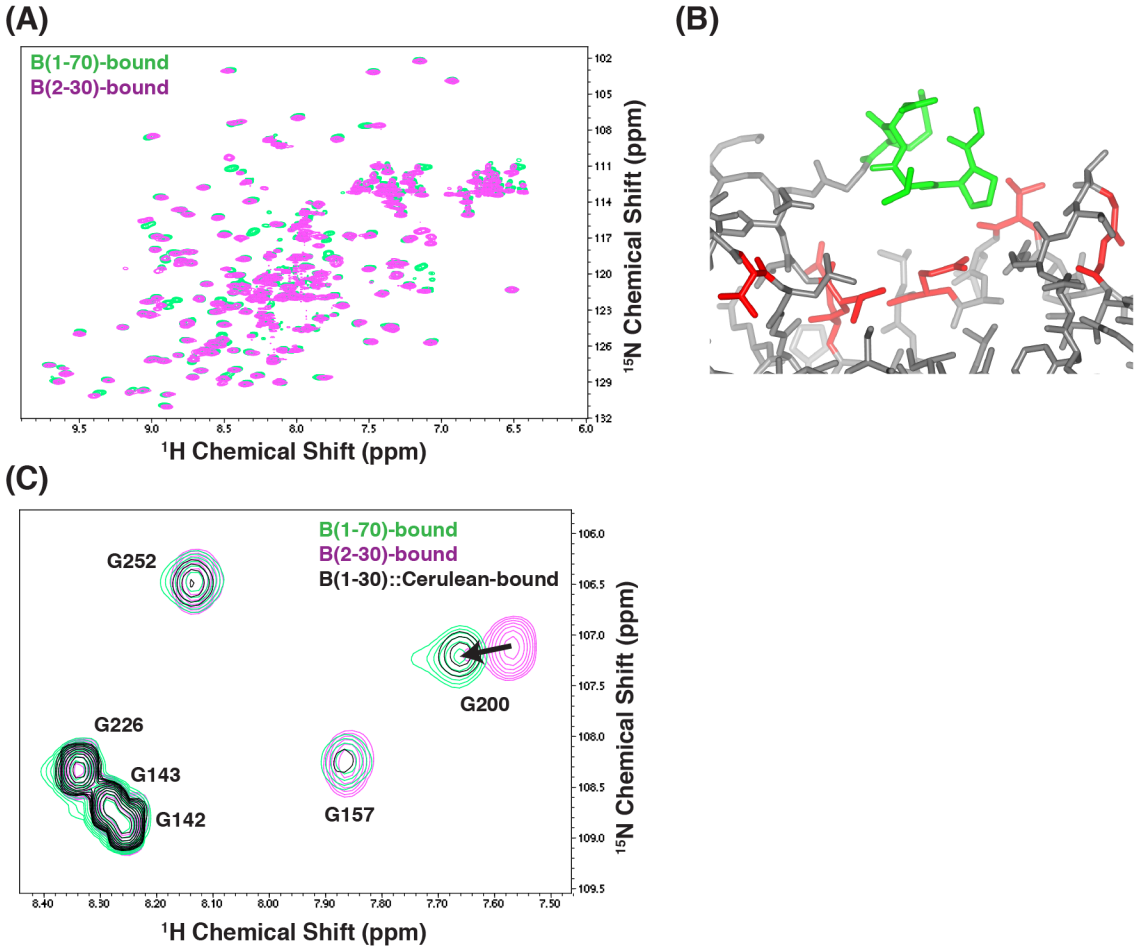
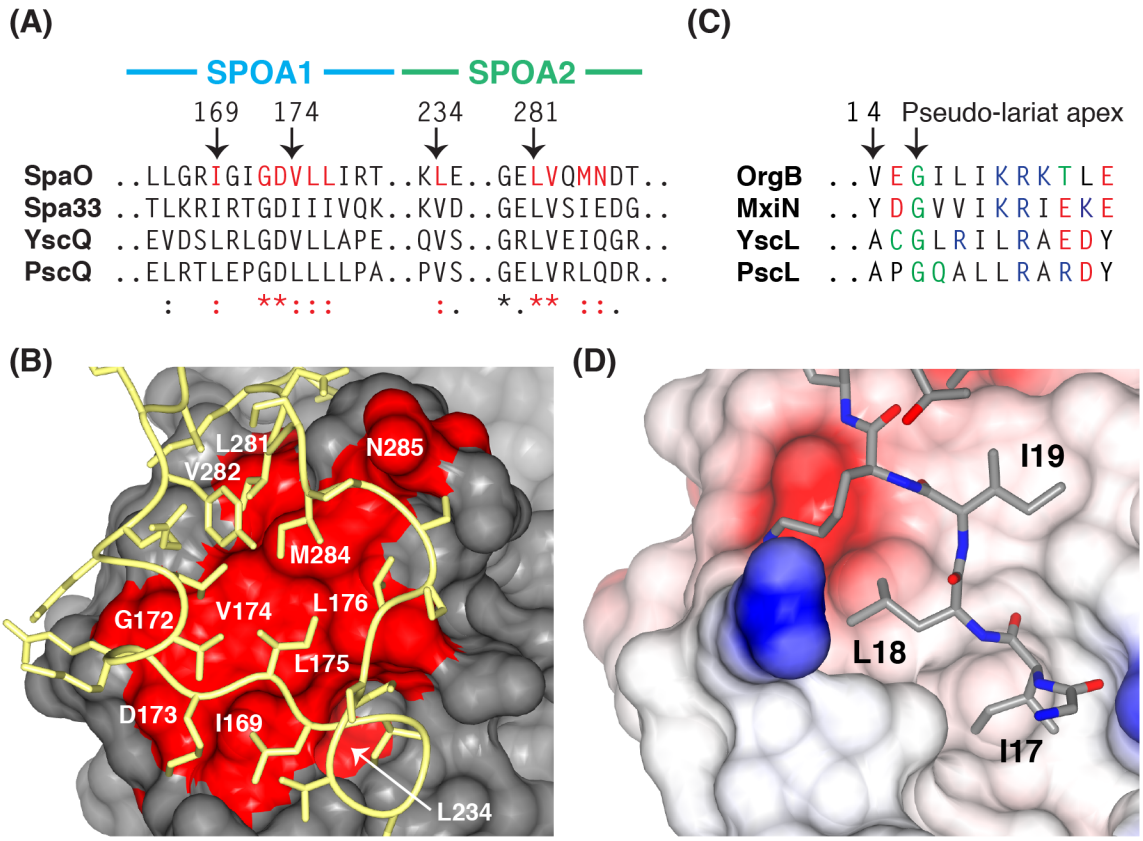


Figure 3.10: The OrgB coiled coil region does not interact with SpaO(SPOA1-SPOA2). (A) Overlayed ^{15}N-HSQC of APAR-coiled-coil-bound- (B(1-70), green) and APAR-bound- (B(2-30), violet) SpaO(140-297). (B) The 7 residues (red) with the most perturbed resonances in (A) are clustered around the OrgB amino-terminal 5 residues (green) in the SpaO-OrgB crystal structure. (C) The amide proton resonances of SpaO(140-297) bound to OrgB(1-30):: Cerulean is indistinguishable from that bound to OrgB(1-70). Zoomed ^{15}N-HSQC highlighting concordance at HN/G200 is shown. HN/G200 was one of the 7 resonances discordant between OrgB(2-30)- and OrgB(1-70)-bound SpaO.

The SpaO residues that make up the core of the OrgB-binding interface are highly conserved across both the *Salmonella/Shigella* and *Yersinia/ Pseudomonas* T3SS clades (Figure 3.11A,B). In the crystal, a subset of these residues form the hydrophobic docking site for the aliphatic side chains of OrgB residues Ile17, Leu18, and Ile19 (Figure 3.11D). The OrgB surface area buried by these three residues (360 $Å^2$) accounts for approximately one third of the APAR's total buried area. Here, the APAR shows noteworthy sequence homology: immediately following a conserved glycine (Gly16, pseudo-lariat apex) is a string of aliphatic and basic amino acids in each homologue (Figure 3.11C). The sequence conservation on both sides of the SpaO-OrgB interface suggests that this mode of binding is generally applicable to other injectisome systems.

Figure 3.11: Generalizability of the SPOA1-SPOA2-APAR model. (A) The SpaO residues at the APAR interaction site are highly conserved across homologues in other species. Excerpts of the M-COFFEE alignment of SpaO, *S. flexneri* Spa33, *Y. enterocolica* YscQ, and *P. aeruginosa* PscQ are shown with conserved APAR-interacting residues highlighted in red. Asterisk denotes identity; colon, high conservation; period, low conservation. (B) A surface representation of SpaO with the 100% or highly conserved interfacial residues identified in (A) colored red and the OrgB APAR backbone is yellow. (C) Clustal Omega alignment of the APAR regions of OrgB, *S. flexneri* MxiN, *Y. enterocolica* YscL, and *P. aeruginosa* PscL. Color code: Aliphatic/aromatic, black; basic, blue; acidic, red; polar/gylcine, green. (D) The binding site for OrgB(17-19, gray) is shown as an electrostatic surface. OrgB(1-15) have been removed for clarity.

Figure 3.11: Generalizability of the SPOA1-SPOA2-APAR model.



To quantitatively assess the thermodynamic parameters of the SpaO-OrgB(APAR) interaction, the binding of SpaO(140-297) to OrgB(1-30)::Cerulean was probed by isothermal titration calorimetry (ITC). At 25°C, SpaO(140-297) bound to the APAR peptide with a dissociation constant of 2.0 μM ± 0.5 μM at 1.04 ± 0.01 sites (Figure 3.12). No binding was observed between SpaO(140-297) and Cerulean without the APAR fusion. Providing additional support for the necessity of SPOA1 in APAR binding, no specific binding was observed between the SPOA2 homodimer and OrgB(1-30)::Cerulean (Figure

3.12). The interaction between the APAR and SpaO(140-297) was endothermic and entropy driven (ΔH = 1602 ± 71 cal/mol; ΔS = 31.5 ± 0.3 cal/mol/deg), consistent with the importance of hydrophobic interactions at the interface.

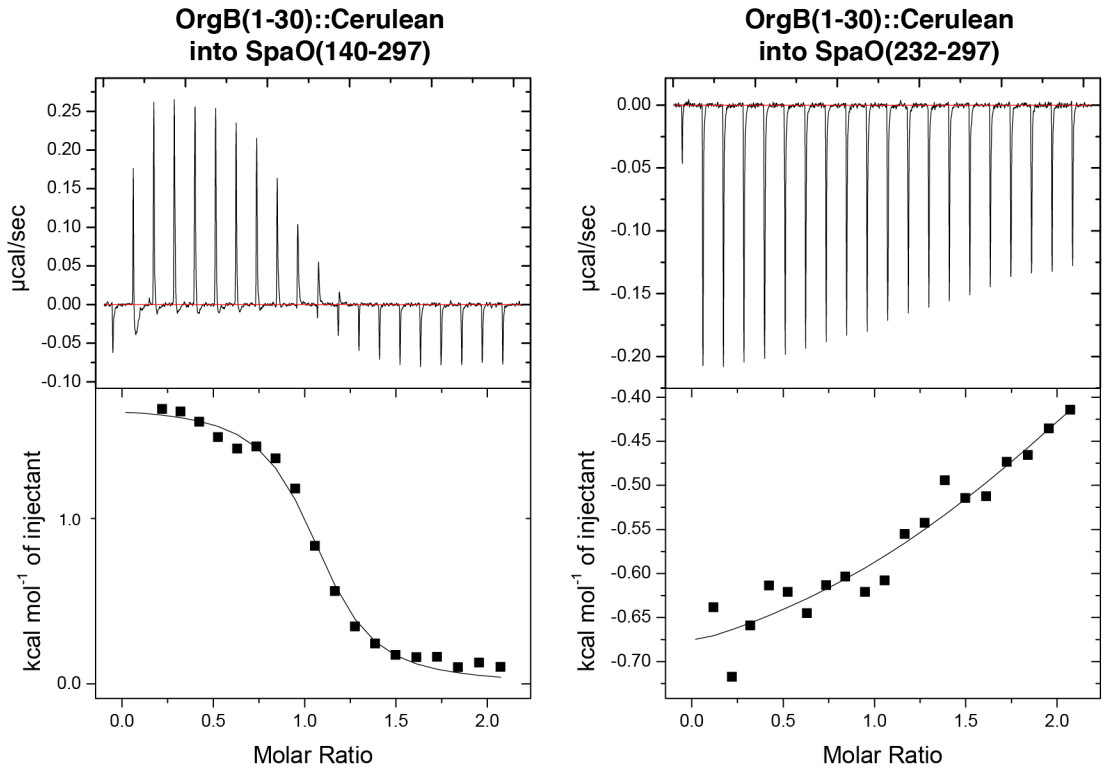


Figure 3.12: Quantitative assessment of SpaO(SPOA1-SPOA2)-OrgB(APAR) binding thermodynamics. Representative ITC data for the OrgB APAR interacting with SpaO SPOA1-SPOA2 (left) and SPOA2 alone (right). The uncorrected power differential between the reference and sample cells is shown on top; the background subtracted heats are shown on the bottom.

3.3 SPOA-APAR interaction is necessary for T3SS function

To test whether the SPOA1-SPOA2-APAR interaction per se is necessary for T3SS function, I sought to construct an APAR mutant that would disrupt SpaO binding. The crystallographic model and binding thermodynamics suggested that mutation of OrgB(17-19) from aliphatic to charged residues (I17D,L18D,I19D) would disrupt the the SpaO-OrgB interface. As predicted, SpaO failed to co-affinity purify OrgB(I17D,L18D,I19D)-InvC when co-expressed in *E. coli* (Figure 3.13A). Notably, the triple mutation did not impair OrgB solubility or affect InvC binding (Figure 3.13B). *In vivo*, the aspartate triple mutation completely abolished T3SS secretory function (Figure 3.13C), consistent with the hypothesis that the SpaO(SPOA1-SPOA2)-OrgB(APAR) interaction is necessary for T3SS function.

Fluorescence microscopic analyses of the *Yersinia* SpaO homologue have shown it to localize in discrete perimembranous punctae (137). Might the SPOA1-SPOA2-APAR interaction function to localize SpaO to the bacterial inner membrane? In an otherwise wild-type genomic background, an EGFP::3xFLAG::SpaO fusion exhibits punctate, perimembranous localization, consistent with its recruitment to injectisome basal body channels (Figure 3.13D). Deletion of *orgB* disrupts proper SpaO localization, producing a more diffuse, cytoplasmic pattern, and the asparate triple mutation was sufficient to phenocopy the *orgB* deletion mutant (Figure 3.13D). Together, these data suggest that the SPOA1-SPOA2-APAR assembly is necessary for the proper localization of SpaO to discrete perimembranous puncta, and that this arrangement is required for T3SS function.

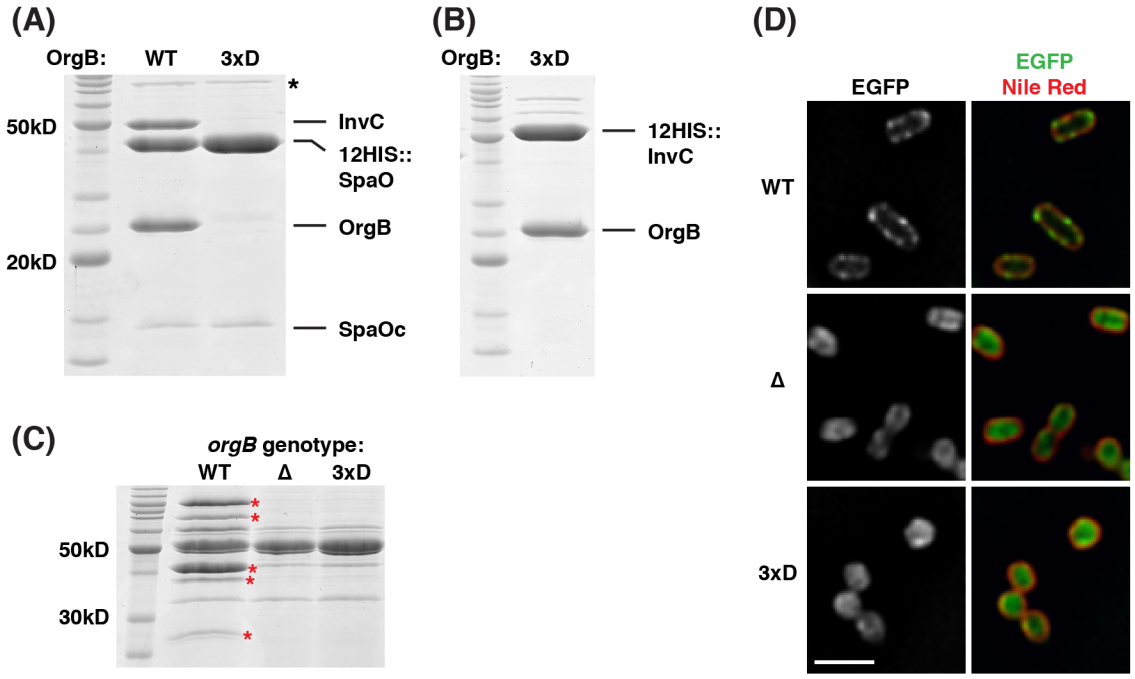


Figure 3.13: Structure based disruption of the SpaO(SPOA1-SPOA2)-OrgB(APAR) interaction disrupts T3SS function and SpaO subcellular localization. (A) Co-affinity purification of double hexahistidine-tagged SpaO with InvC and OrgB (WT, wildtype; 3xD, I17D/L18D/I19D) coexpressed in *E. coli*. Coomassie-stained gel is shown. Asterisk denotes nonspecific co-purifying E. coli proteins, likely chaperones. SpaOc indicates the cryptically expressed SPOA2-containing carboxy-terminal fragment. (B) Co-affinity purification of double hexahistidine-tagged InvC and 3xD OrgB. (C) Coomassie-stained culture supernatant from wild-type (WT, strain SB1741), *orgB* deletion(Δ) and *orgB(I17D,L18D,I19D)* (3xD) *S. typhimurium* shows loss of injectisome substrate (red asterisks) secretion in the mutants, while flagellar secretion remains intact. (D) Widefield microscopic imaging of fixed S. typhimurium shows exclusive perimembranous localization of EGFP::3xFLAG::SpaO in the wild type background but cytoplasmic localization in the *orgB* mutants (scale bar is 2 *μ*m, single z-slices shown). Nile red is used as a membrane stain.

Studies of the flagellar apparatus suggest an intriguing alternate function for the APAR region of OrgB. In the flagellar system, the extreme amino terminus of the OrgB homologue FliH has been shown to be necessary for proper localization of the FliH-FliI (InvC homologue) complex to the flagellar T3SS export apparatus (133). Photocrosslinking studies place the FliH amino terminus in close proximity to the export apparatus protein FlhA (InvA homologue), but evidence for a direct interaction is weak (187). Might the APAR of OrgB interact with the cytoplasmic domain of InvA? One might imagine a scenario where APARs extending from OrgB-InvC complexes are able to interact with both SpaO and InvA, bridging the sorting platform and export apparatus. However, while double hexahistidine-tagged SpaO(140-297) could co-affinity purify OrgB(1-30)::Cerulean, no interaction between double hexahistidine tagged InvA(cytoplasmic domain) and OrgB(1-30)::Cerulean was observed (Figure 3.14).

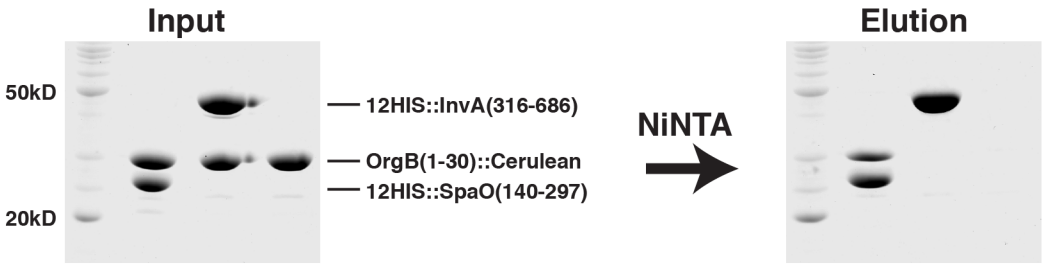


Figure 3.14: The OrgB APAR does not interact with the InvA cytoplasmic domain. Co-affinity purification of OrgB(1-30)::Cerulean with the indicated double hexahistidine tagged proteins shows that the OrgB APAR does not interact with the cytoplasmic domain of InvA in isolation.

CHAPTER 4:

SPOA-APAR INTERACTIONS IN THE FLAGELLAR T3SS

4.1 SPOA-APAR interactions are T3SS subtype specific

As described above, *S. typhimurium* contains three T3SS: two injectisomes and a flagellar apparatus. The flagellar structural analog of the the sorting platform is a robust annulus of density at the cytoplasmic face of the flagellar basal body. This "C-ring" is primarily composed of three proteins: FliM, FliN, and FliG (188). The SpaO homologues FliM and FliN are predicted to contain one SPOA domain each, which I prospectively designate as SPOA1 and SPOA2, respectively (Figure 4.1). Paralleling the injectisome, FliN is known to interact with the OrgB homologue FliH (133). The evolutionary relationship between injectisomes and flagella creates a practical conundrum: how are homologous T3SS components segregated to their corresponding secretion systems within a common cytoplasmic milieu?

To qualitatively assess the subtype specificity of SPOA-APAR interactions, I co-affinity purified a panel of *Salmonella* SPOA domains with hexahistidine-tagged APAR::T4 lysozyme fusions (Figure 4.2). Indeed, the OrgB and FliH APARs robustly co-affinity purify their cognate SPOA1-SPOA2 proteins but not that of the other T3SS subtype (Figure 4.2). Neither SpaO nor FliM-FliN are pulled down by the APAR from a second pathogenic T3SS found in *S. typhimurium* (SPI-2 SsaK).

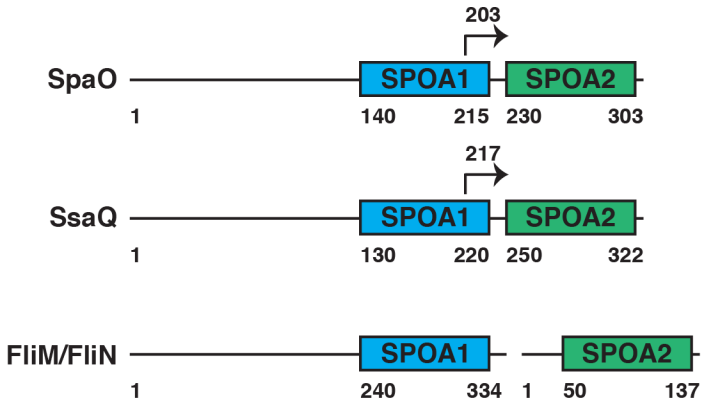


Figure 4.1: SPOA domains in the *S. typhimurium* genome. Prospective annotation of the six SPOA domains in the *S. typhimurium* genome based on the SpaO structures presented here and those of its injectisome and flagellar homologues. Note that the prediction of an internal translation start site at M217 in SsaQ is based on the work of Ghosh and colleagues (141). FliM and FliN are expressed from tandem open reading frames in the same operon.

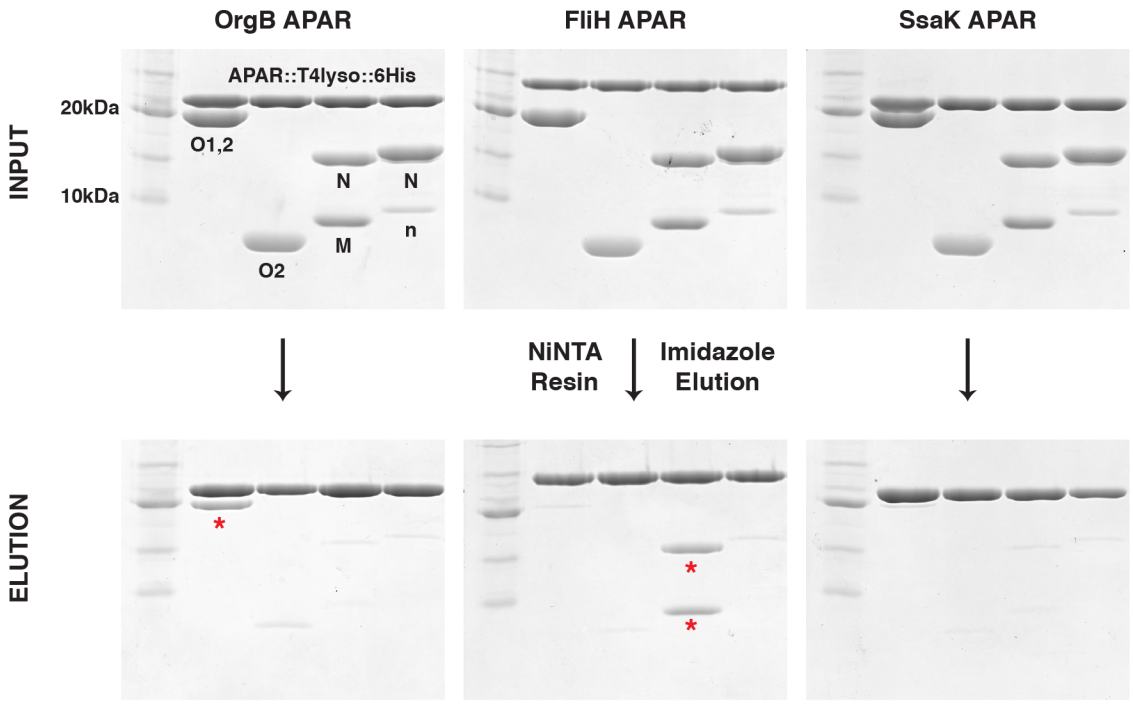


Figure 4.2: SPOA-APAR interactions are T3SS subtype specific. Coomassie-stained gels showing the input and imidazole elution for APAR-SPOA co-affinity purification experiments. Red asterisks indicate the cognate SPOA1-SPOA2 band(s). Abbreviations: T4lyso: T4 lysozyme; O1,2: SpaO(140-297); O2: SpaO(232-297); M: FliM(245-320); N: FliN (1-137); n: co-purifying amino-terminal FliN degradation product.

4.2 Structural basis for the interaction of FliM(SPOA) with FliN(SPOA)

I hypothesized that divergence of the SPOA1-SPOA2-APAR assembly architecture contributes to proper component segregation among T3SS subtypes, and sought to structurally characterize the flagellar SPOA-APAR interactions. While complexes of the FliM SPOA and FliN were stable, they were resistant to crystallization. Interestingly, FliM and FliN can be fused and still support flagellin secretion (Figure 4.3) and some swarming motility (189). In a sense, such FliM-FliN fusions are analogous to SpaO(Val_{GTT}203), in that a single polypeptide contains both SPOA domains and no free SPOA2 is produced.

A construct containing the SPOA of FliM (residues 245-334) fused to FliN(5-137) was expressed in *E. coli* and purified by anion exchange and size exclusion chromatography (Figure 4.3). This construct readily crystallized, and SeMet-doped material gave rise to crystals with diffraction characteristics superior to native protein. The FliM(245-334)::FliN(5-137) structure was solved by SeMet SAD and refined to 2.56 Å using the SeMet data set (Figure 4.4 and Table 4.1).

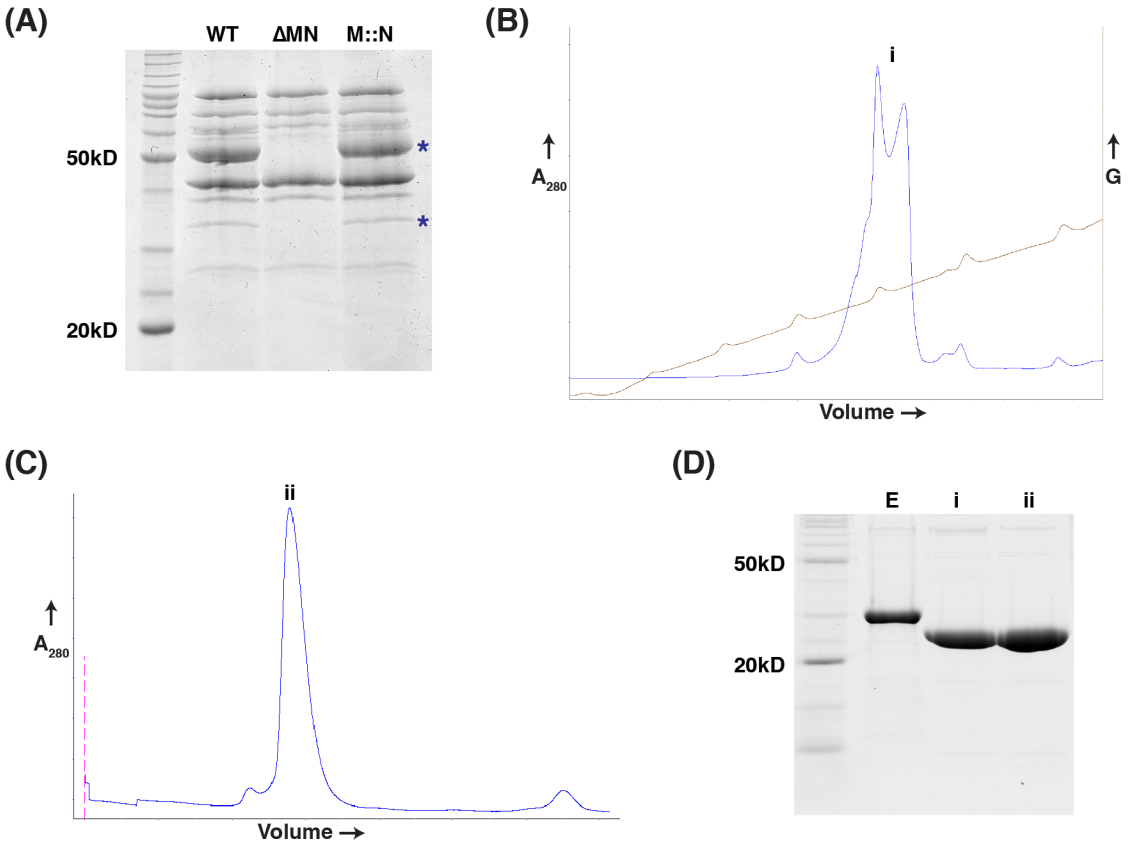


Figure 4.3: Expression, purification, and function of FliM-FliN fusion protein. (A) FliM-FliN fusion is compatible with flagellar secretory function. Coomassie-stained PAGE of culture supernatants from *S. typhimurium* with the indicated genotype (WT, wild-type; ΔMN, deletion of *fliM* and *fliN*; M::N, FliM-FliN(5-137) fusion). Flagellar specific secretory products (lost in the *fliM/fliN* deletion background) are present in the FliM-FliN fusion strain (blue asterisks highlight FliC and FlgL). The FliM::FliN fusion brings FliN residue 5 in frame with full-length FliM; both Met1 and Met4 of FliN are deleted to prevent spurious translation of free FliN. (B-D) Purification of FliM(245-334)::FliN(5-137). NiNTA purified protein ("E") was cut with 3C and subjected to anion exchange chromatography (B). The main peaks (i) were then further purified by size exclusion chromatography (C), peak (ii). (D) Coomassie-stained gel of E, i, ii.

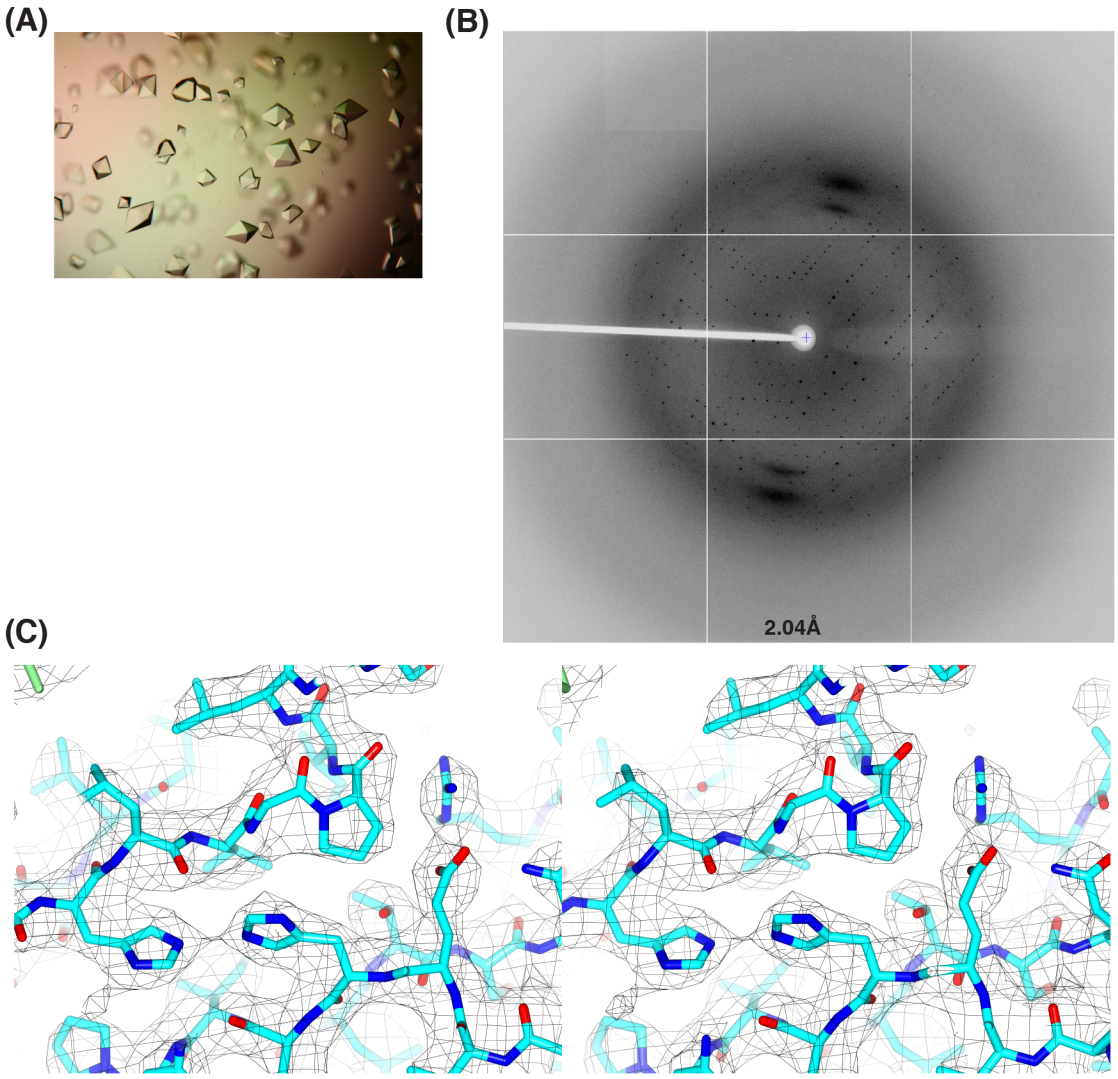


Figure 4.4: Diffraction analysis of FliM(245-334)::FliN(5-137). (A) Crystals and (B) diffraction image of the FliM-FliN fusion protein (SeMet incorporated). (C) Stereoimage of a selected region of the 2Fo-Fc electron density map for the 2.56 Å FliM(245-334, cyan)::FliN(5-137, green) structure. The map is contoured to 1σ.

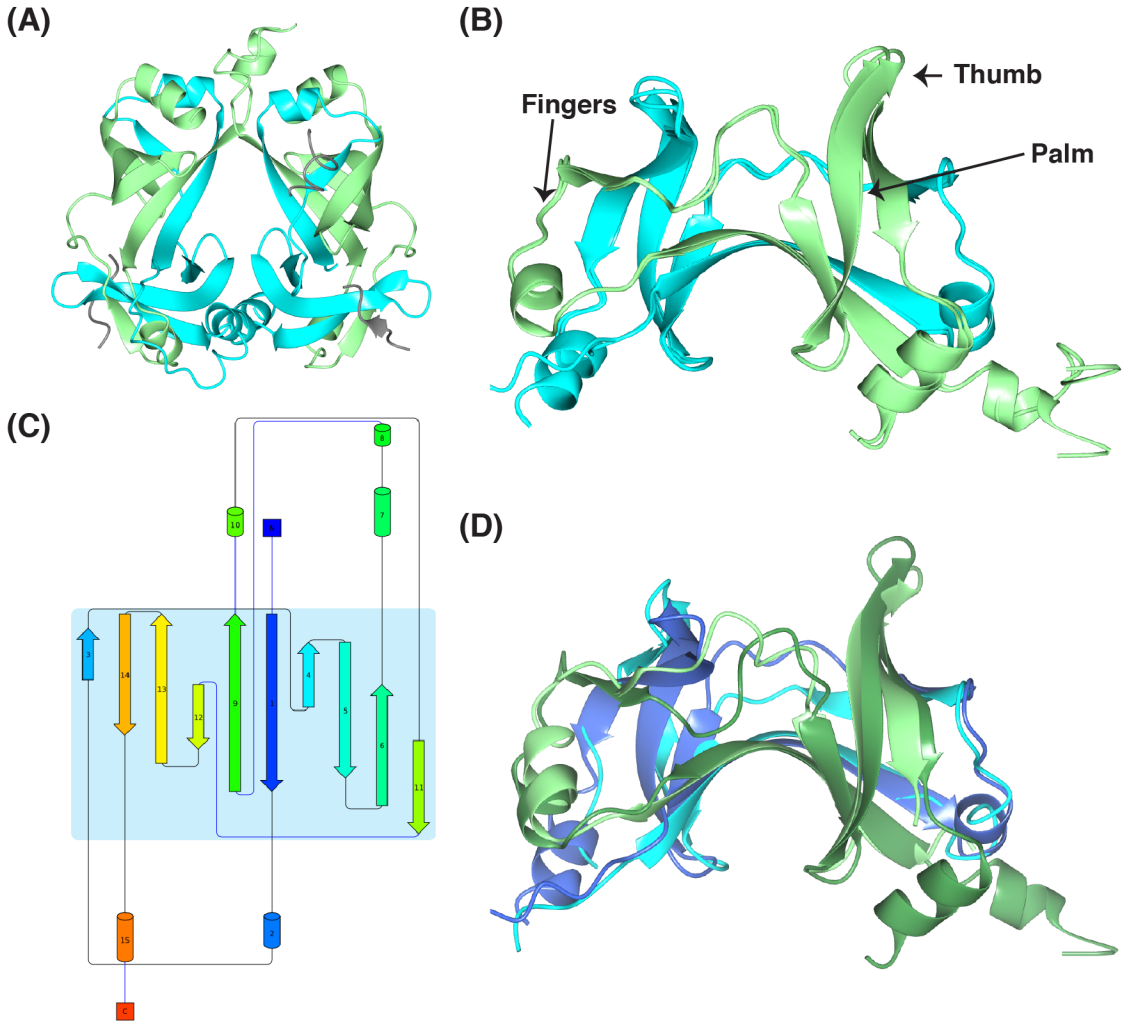
Table 4.1: Diffraction and refinement statistics for FliM(245-334):: FliN(5-137).

		SeMet
PDB ID		4YXB
Data Collection	**Space group**	$P2_12_12_1$
	Cell: a, b, c (Å)	75.15, 81.50, 89.96
	Cell: α, β, γ (°)	90, 90, 90
	Resolution (Å)	57.67–2.56 (2.67–2.56)
	R_{merge}	0.097 (1.215)
	I/σI	18.5 (2.7)
	$CC_{1/2}$	0.999 (0.814)
	Completeness (%)	100 (100)
	No. of reflections	18,372
	Redundancy	13.8 (14.3)
Refinement	**R_{work}/R_{free}**	0.2175/0.2593
	No. of atoms	2,633
	Protein	2,605
	Ligand/ion	5
	Water	23
	B factors: Protein	68.30
	B factors: Ligand/ion	73.40
	B factors: Water	64.50
	Geometry (r.m.s.d.)	
	Bond lengths (Å)	0.010
	Bond angles (°)	1.31
	[Φ,Ψ] Favored (%)	92
	[Φ,Ψ] Allowed (%)	99.1
	[Φ,Ψ] Outliers (%)	0.9

In the FliM(245-334)::FliN(5-137) structure, only the SPOA domains were clearly structured (FliM: 258-330; FliN: 53-137). Architecturally, the FliM(SPOA1)-FliN(SPOA2) interaction is similar to that of SpaO (Figure 4.5, 2.28 Å RMSD), with the exception of additional helices present at the carboxy terminus of each SPOA, as observed in FliN homodimers from *Thermotoga maritima* (139). Additional weak density was observed in contact with the surface of the SPOA-SPOA structure that was strong enough in segments to be modeled as peptide. I hypothesize that this additional density corresponds to the ~50 residues at the amino terminus of FliN linking the two SPOA domains in this artificial fusion. While it is tempting to speculate that these contacts are physiologically significant, it has not yet been determined that they are anything other than an artifact of the fusion strategy and/or crystal packing and thus will not be discussed further here.

Figure 4.5: Structure of FliM(SPOA)-FliN(SPOA). (A) Asymmetric unit of the FliM(245-334, cyan)::FliN(5-137, green) crystal. Noncontiguous modeled peptide fragments are shown in gray. (B) Superposition of the two constituents of the asymmetric unit (0.56 Å r.m.s.d) reveals a stereotypical SPOA-SPOA interaction. (C) The topology diagram is consistent with a *bona fide* SPOA fold for both FliM and FliN. (D) Superposition of apo-SpaO SPOA1-SPOA2 (light blue and green, respectively) with apo-FliM::FliN (dark blue and green, respectively; 2.28 Å r.m.s.d.).

Figure 4.5: Structure of FliM(SPOA)-FliN(SPOA).



The FliM(245-334)::FliN(5-137) fusion protein eluted from size exclusion chromatography as a much larger particle than would be expected for a monomer of its size (Figure 4.3). Dynamic light scattering analysis of gel filtration purified protein suggested the presence of a relatively monodispersed entity with a 36 Å ± 1 Å hydrodynamic radius and a predicted molecular mass of 65 kDa ± 3

kDa (Figure 4.6). Interestingly, the crystallographic asymmetric unit contained a dimer of fusion proteins with a diameter of 50-55 Å (mass: 50 kDa; Figure 4.5). Previous studies have noted the ability of FliN to tetramerize into a similar doughnut-like annulus (139) and have speculatively assigned this structure to the distal cytoplasmic aspect of the C-ring cryo-EM density, because of its quasi-toroid appearance. I posit that while neither the FliN homotetramer nor FliM-FliN heterotetramers ($FliM_2$-$FliN_2$ or FliM-$FliN_3$) can be definitively ascribed to the density in question using the data available, the FliM-FliN heterotetramers should be considered alongside the FliN homotetramer as a candidate for the composition of this structure.

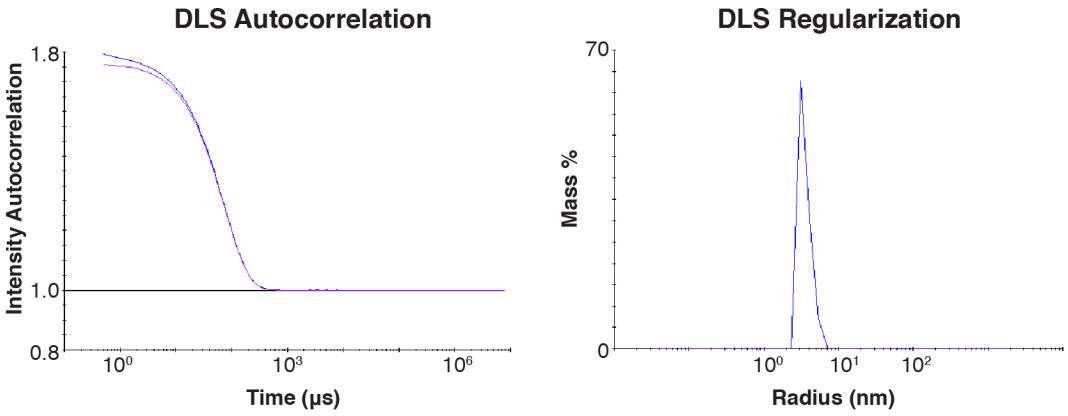


Figure 4.6: Solution behavior of FliM(245-334)::FliN(5-137). The data for 40 μM protein is shown in the autocorrelation and regularization graphs. In the autocorrelation graph, the raw data is shown in blue and the regularization fit is shown in violet.

4.3 Structure of the FliM(SPOA)-FliN(SPOA)-FliH(APAR) complex

The similarity of the SpaO and FliM/FliN SPOA1-SPOA2 structures is consistent with a model for SPOA1-SPOA2 interactions being generalizable across T3SS subtypes. Is the mechanism of APAR binding also structurally similar across T3SS subtypes?

To elucidate the mechanism of FliH-specific assembly with FliM-FliN, I sought to co-crystalize the SPOAs of FliM and FliN in complex with the FliH APAR as a T4 lysozyme fusion. FliM(245-320), FliN(1-137), and FliH(1-43)::T4 lysozyme formed a stable complex over size exclusion chromatography (Figure 4.7); however this complex was resistant to crystallization. As one would predict from the FliM-FliN fusion protein's function *in vivo*, FliM(245-334)::FliN(5-137) was able to bind the FliH APAR (residues 1-43) fused to lysozyme (Figure 4.7). The APAR construct fused to lysozyme was further truncated to residues 1-18, which still formed a complex with FliM(245-334)::FliN(5-137) (Figure 4.8). This complex crystallized, its structure was solved to 2.30 Å resolution by molecular replacement using the FliM(245-334)::FliN(5-137) model and T4 lysozyme (Figure 4.9, Table 4.2)

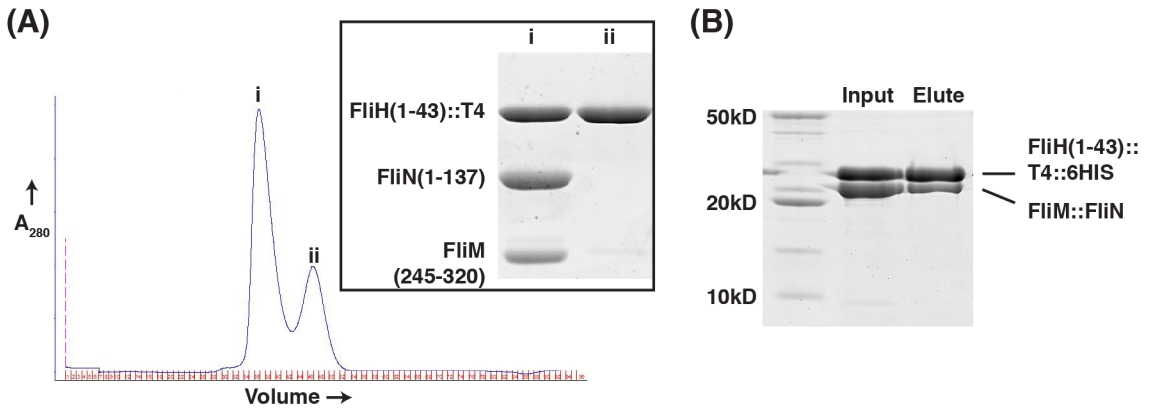


Figure 4.7: FliM-FliN-FliH complexes. (A) Separation of excess FliH(1-43)::T4 lysozyme fusion from the FliM-FliN-FliH::T4 lysozyme complex by size exclusion chromatography. (B) Co-affinity purification of FliM(245-334)::FliN(5-137) with hexahistidine-tagged FliH(1-43)::T4 lysozyme over NiNTA resin. Gels in (A) and (B) are coomassie-stained.

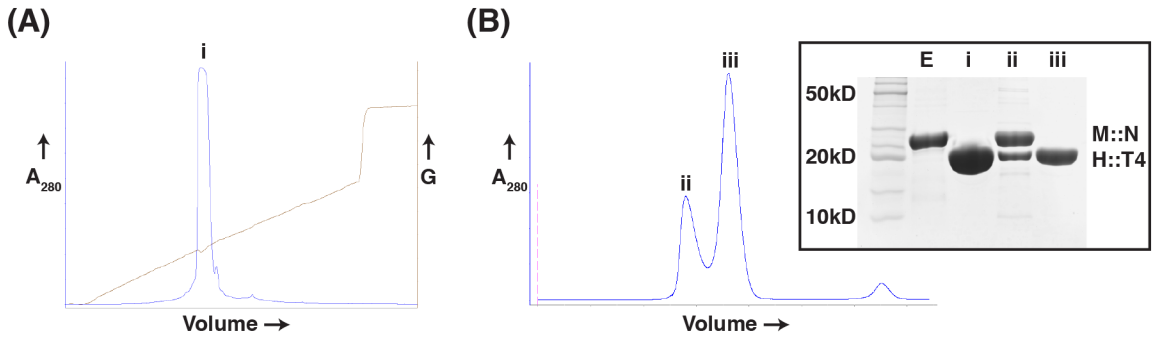


Figure 4.8: Expression and purification of FliM(245-334)::FliN(5-137) + FliH(1-18)::T4 lysozyme. FliH(1-18)::T4 lysozyme fusion (H::T4) was purified under denaturing conditions (E), refolded, and affinity tags were removed with 3C protease. Cleaved protein was further purified by cation-exchange chromatography (A), mixed with FliM(245-334)::FliN(5-137) (M::N), and excess H::T4 (iii) was separated from the complex (ii) by gel filtration chromatography (B). Inset, Coomassie-stained gel of E and the main chromatography peaks.

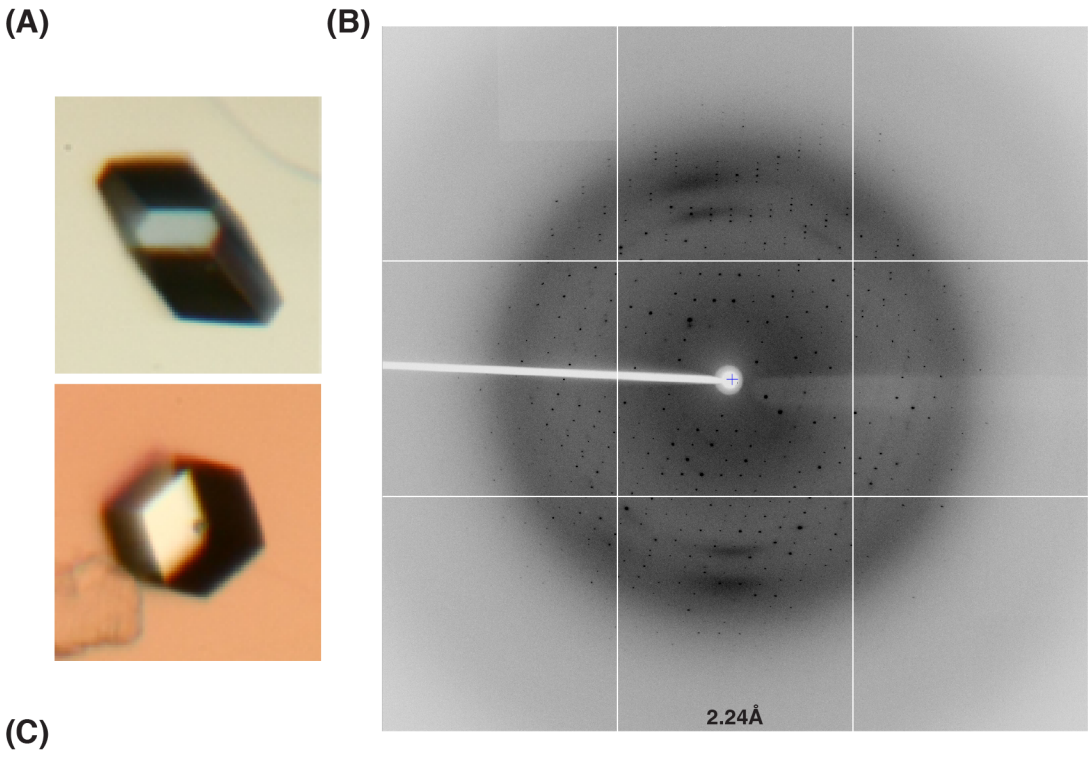


Figure 4.9: Diffraction analysis of FliM(245-334)::FliN(5-137) + FliH(1-18)::T4lysozyme. (A) Representative crystals and (B) diffraction pattern for the FliM(245-334)::FliN(5-137) + FliH(1-18)::T4 lysozyme complex. (C) Stereoimage of the 2Fo-Fc electron density map for the 2.30 Å FliM(245-334, cyan)::FliN(5-137, green) + FliH(1-18)::T4 lysozyme (gray) structure. The map is contoured to 1σ and highlights the FliH W7 binding region.

Table 4.2: Diffraction and refinement statistics for FliM(245-334)::FliN(5-137) + FliH(1-18)::T4lysozyme.

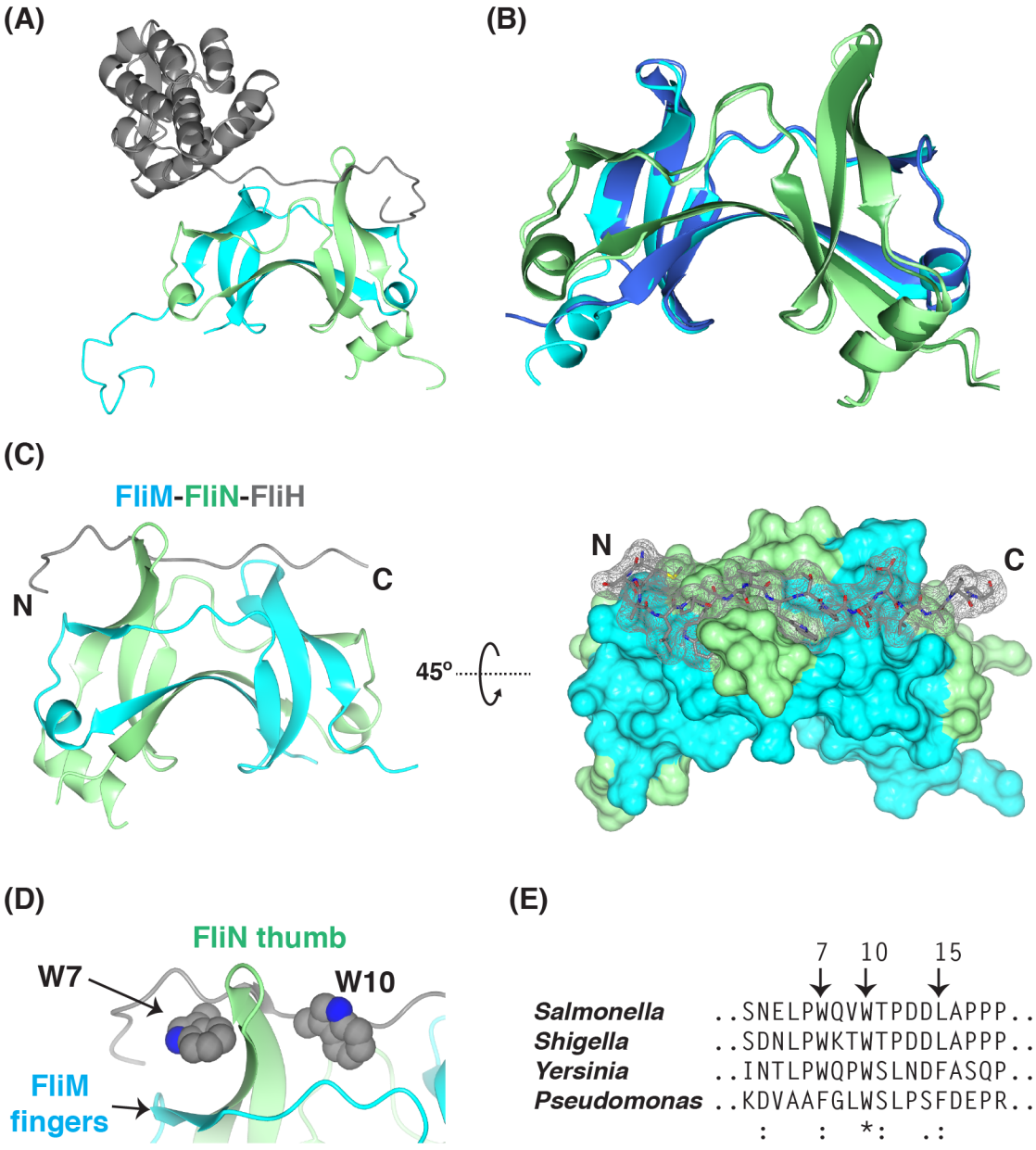
		Native
PDB ID		4YXC
Data Collection	**Space group**	$P2_12_12_1$
	Cell: a, b, c (Å)	43.21, 76.37, 119.4
	Cell: α, β, γ (°)	90, 90, 90
	Resolution (Å)	64.33–2.30 (2.38–2.30)
	R_{merge}	0.070 (0.923)
	I/σI	20.2 (2.6)
	$CC_{1/2}$	0.999 (0.811)
	Completeness (%)	99.8 (99.8)
	No. of reflections	18,174
	Redundancy	12.9 (12.9)
Refinement	**R_{work}/R_{free}**	0.1967/0.2620
	No. of atoms	2,739
	Protein	2,668
	Ligand/ion	0
	Water	71
	B factors: Protein	69.70
	B factors: Ligand/ion	
	B factors: Water	65.00
	Geometry (r.m.s.d.)	
	Bond lengths (Å)	0.009
	Bond angles (°)	1.15
	[Φ,Ψ] Favored (%)	94
	[Φ,Ψ] Allowed (%)	99.1
	[Φ,Ψ] Outliers (%)	0.9

As with its injectisome counterparts, the FliM-FliN SPOA1-SPOA2 did not undergo large conformational changes upon APAR binding (Figure 4.10B, 1.11 Å r.m.s.d.); however, the binding mode for the FliH APAR is radically different. In contrast to the OrgB pseudo-lariat, the FliH APAR adopts a near-linear conformation along the "top" of FliM-FliN (Figure 4.10C). As observed in the SpaO-OrgB assembly, the FliH APAR makes extensive contact with both SPOA1 and SPOA2 (Figure 4.10C), supporting the observation that the FliH APAR interacts more strongly with the FliM-FliN heterodimer than the FliN homodimer (Figure 4.2).

The FliM-FliN-FliH assembly is characterized by the burial of several highly conserved hydrophobic FliH side-chains. Two tryptophan side-chains form an aromatic "clamp," which binds hydrophobic pockets on opposite faces of the FliN thumb (Figure 4.10D). Alanine scanning mutagenesis of the FliH amino terminus had previously identified these tryptophans as critical for flagellar function (133), and they are highly conserved (Figure 4.10). Similarly, the bulky side-chain of FliH Leu15 fills a hydrophobic pocket on the thumb of FliM (Figure 4.11). The binding interfaces for these three residues are formed by both FliM and FliN and are highly conserved across species (Figure 4.11E). This structure presents a conserved model for FliH-FliM-FliN interaction, which is distinct from that of SpaO-OrgB.

Figure 4.10: Structure of the FliM(SPOA)-FliN(SPOA)-FliH(APAR) complex. (A) The crystallographic asymmetric unit for FliM(245-334, cyan)::FliN(5-137, green) + FliH(1-18)::T4 lysozyme (gray). (B) Superposition of apo- (cyan/green) and APAR-bound (blue/dark green) FliM::FliN (1.11 Å r.m.s.d.). (C) Ribbon diagram (left) and surface representation (right) of the FliM-FliN-FliH structure. T4 lysozyme has been omitted. N and C indicate the amino- and carboxy-termini of the FliH APAR, respectively. (D) A zoomed view of the FliH aromatic clamp, with the side-chain atoms of FliH W7 and W10 represented as spheres. (E) Excerpted M-COFFEE alignment of FliH with its homologues from *S. flexneri, Y. enterocolica,* and *P. aeruginosa*. Highly conserved residues of interest are noted (*S. typhimurium* numbering shown).

Figure 4.10: Structure of the FliM(SPOA)-FliN(SPOA)-FliH(APAR) complex.



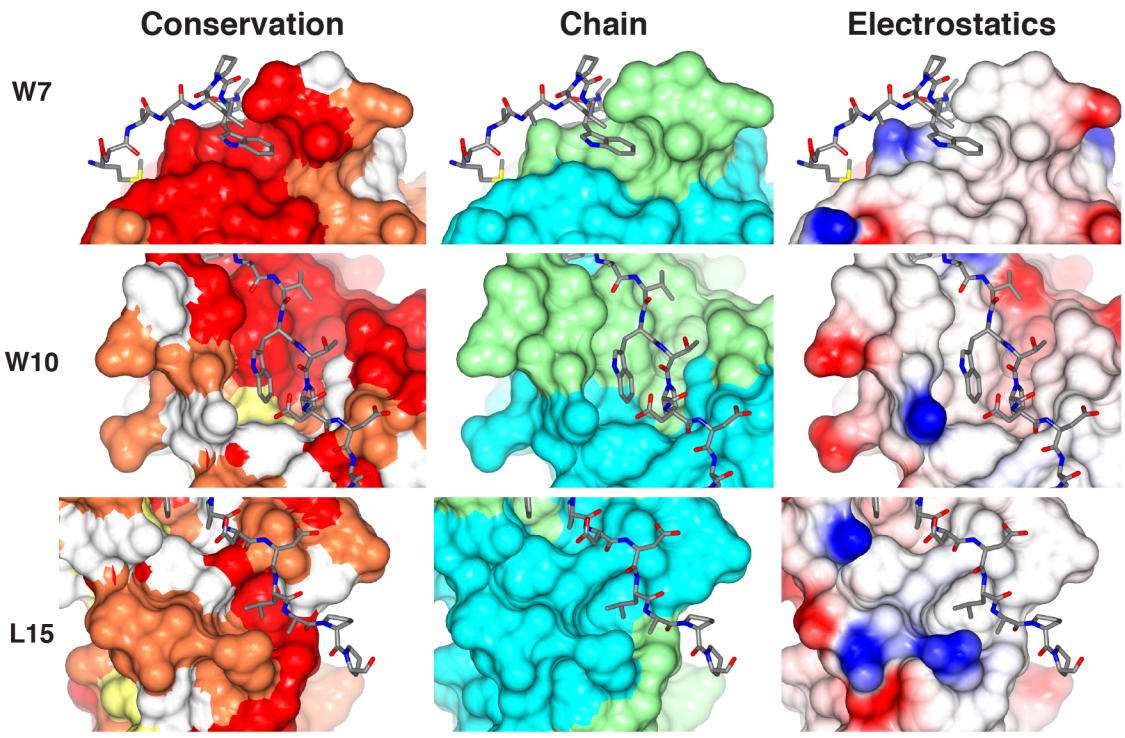


Figure 4.11: Surface characteristics of the FliH(APAR) binding site. The binding pockets for FliH W7, W10, and L15 on FliM-FliN are shown as surface renderings. On the left, the surface is color coded by residue conservation across *S. typhimurium, S. flexneri, Y. enterocolica,* and *P. aeruginosa*. Red indicates 100% identity. Orange, yellow, and white indicate high, medium, and low conservation by M-COFFEE multiple sequence alignment. In the middle, the same surface is shown but color-coded by chain: FliM is cyan, FliN is light green. On the right, the surface is color coded by electrostatic potential (red, negative; blue, positive).

CHAPTER 5:

THE SpaO AMINO-TERMINAL DOMAIN

The preceding chapters have presented structures for SPOA1-SPOA2 and SPOA2-SPOA2 complexes. In the case of the former, this architecture is conserved from injectisomes to the flagellar apparatus and functions as a scaffold for the SctL/FliH APAR. What, though, is the function of the SPOA2 homodimer? Similarly, my structural dissection of SpaO and its homologues has thus far neglected its 140 amino-terminal residues: what might be their structure and function?

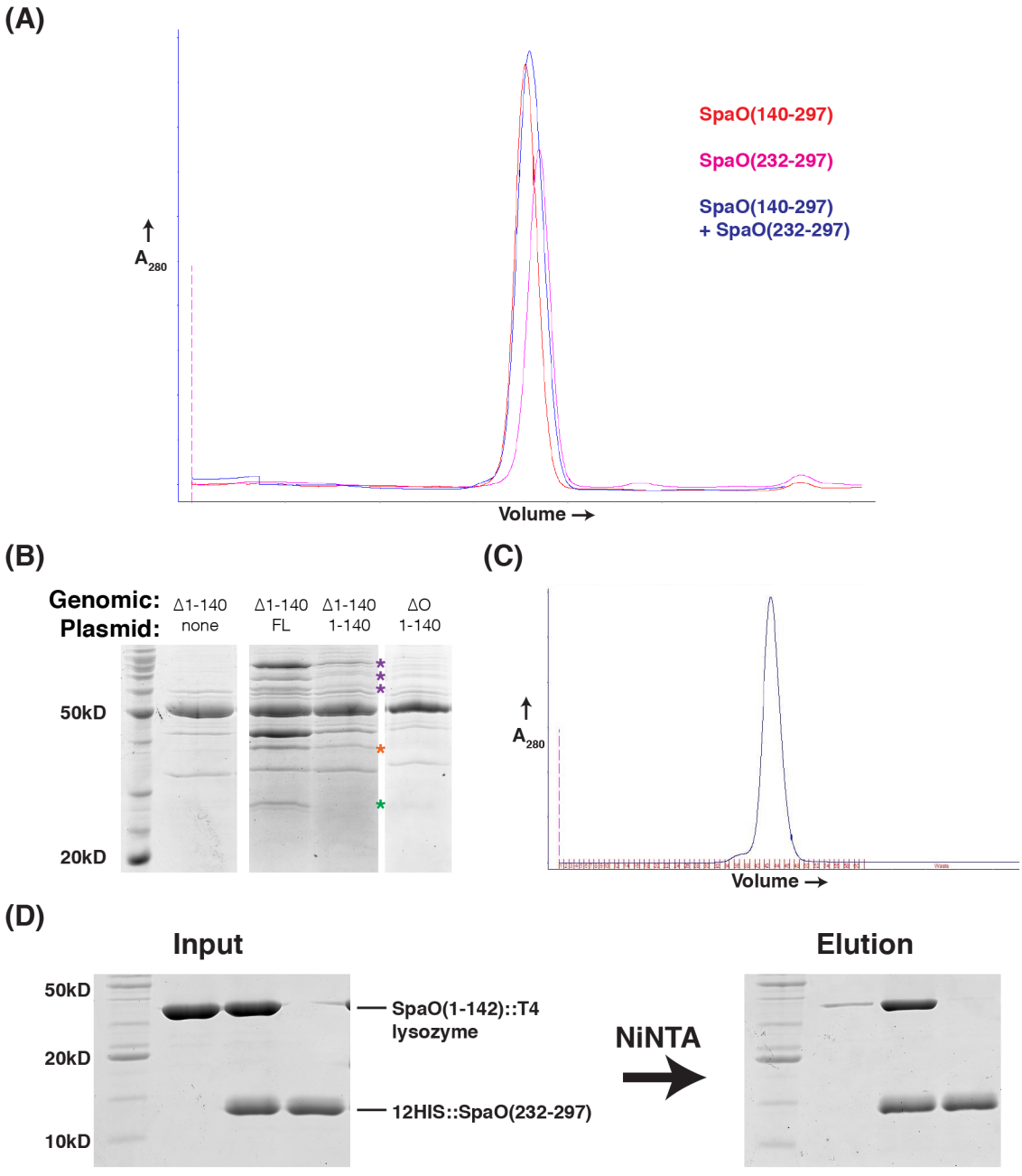
In Chapter 2, I presented data showing that free SPOA2 binds to full length SpaO and remains as a complex through purification (Figures 2.1 and 2.2). Thus, I hypothesized that free SPOA2 might be able to interact with either the SPOA1-SPOA2 complex (residues 140-297) or with the SpaO NTD in isolation. To test the former possibility, SpaO(140-297) was mixed with SpaO(232-297) and passed over a size exclusion chromatography column. There was no shift in elution volume for either component, indicating a lack of complex formation (Figure 5.1A).

The amino-terminal domain (NTD) of SpaO is necessary for T3SS, as its genomic deletion completely disrupts secretory function (Figure 5.1B). While T3SS function is fully complemented by SpaO(1-303) in this mutant, it is only partially complemented by the deleted domain (residues 1-140) (Figure 5.1B). While secretion of some substrates is clearly restored (orange asterisk in Figure

5.1B), secretion of other substrates is only partially restored (purple asterisks) or is not restored at all (green asterisk), suggesting that full SpaO NTD function requires covalent linkage to SPOA1-SPOA2. While biochemical analysis or (ideally) structure determination might shed light on the intriguing function of the NTD, such efforts have been impeded by its poor solubility. Inspired by my success at producing the OrgB and FliH APAR peptides as T4 lysozyme fusions, I fused SpaO(1-142) to lysozyme and found that the fusion protein could be solubly refolded (Figure 5.1C). In contrast to SPOA1-SPOA2, the SpaO NTD::T4 lysozyme construct interacted with free SPOA2, as determined by co-affinity purification (Figure 5.1D).

Figure 5.1: The SpaO NTD interacts with SPOA2. (A) Mixing SpaO(140-297) with SpaO(232-297) does not result in a complex with a smaller elution volume on size exclusion chromatography. (B) Coomassie-stained SDS-PAGE of *S. typhimurium* culture supernatants grown under T3SS stimulating conditions. Abbreviations: FL, full length SpaO; ΔO, deletion of spaO; Δ1-140, deletion of spaO codons 1-140; 1-140, complementation with SpaO(1-140). SpaO was 3xFLAG tagged at its amino terminus in each *S. typhimurium* strain (except ΔO) and complementation construct. Asterisks highlight bands of interest, as described in the text. (C) Size exclusion chromatography of refolded SpaO(1-142)::T4 lysozyme is consistent with a monodispersed sample. (D) Co-affinity purification of SpaO(1-142)::T4 lysozyme with double hexahistidine tagged SpaO(SPOA2) is robust and substantially greater than background binding of SpaO(1-142)::T4 lysozyme to NiNTA resin in the absence of a hexahistidine-tagged binding partner.

Figure 5.1: The SpaO NTD interacts with SPOA2.



To confirm and expand on the finding that free SPOA2 interacts with the SpaO NTD and not SPOA1-SPOA2, these interactions (or lack thereof) were probed by isothermal titration calorimetry (Figure 5.2). Free SPOA2 bound to SpaO(1-142)::T4 lysozyme with a dissociation constant of 20.0 μM ± 0.22 μM at 25°C. The binding isotherm was consistent with the NTD interacting with dimeric SPOA2, as the calculated binding stoichiometry was 1:1.05 (±0.02) when the SPOA2 concentration was input as the concentration of SPOA2 dimers. Binding was exothermic and completely enthalphy driven (ΔH = -14500 ± 36 cal/mol; ΔS = -27.1 ± 0.1 cal/mol/deg). No specific interaction with free SPOA2 was detected for SpaO(140-297).

Given the interaction of SPOA2 homodimers with the SpaO NTD, I hypothesized that the NTD might be stably co-refolded with SPOA2. Indeed, SpaO(1-140) could be refolded with SpaO(232-297) and the complex was stable over size exclusion chromatography (Figure 5.3). Thus far, the complex has resisted crystallization, but it is hoped that further screening and construct re-engineering might yield diffraction quality crystals of this complex.

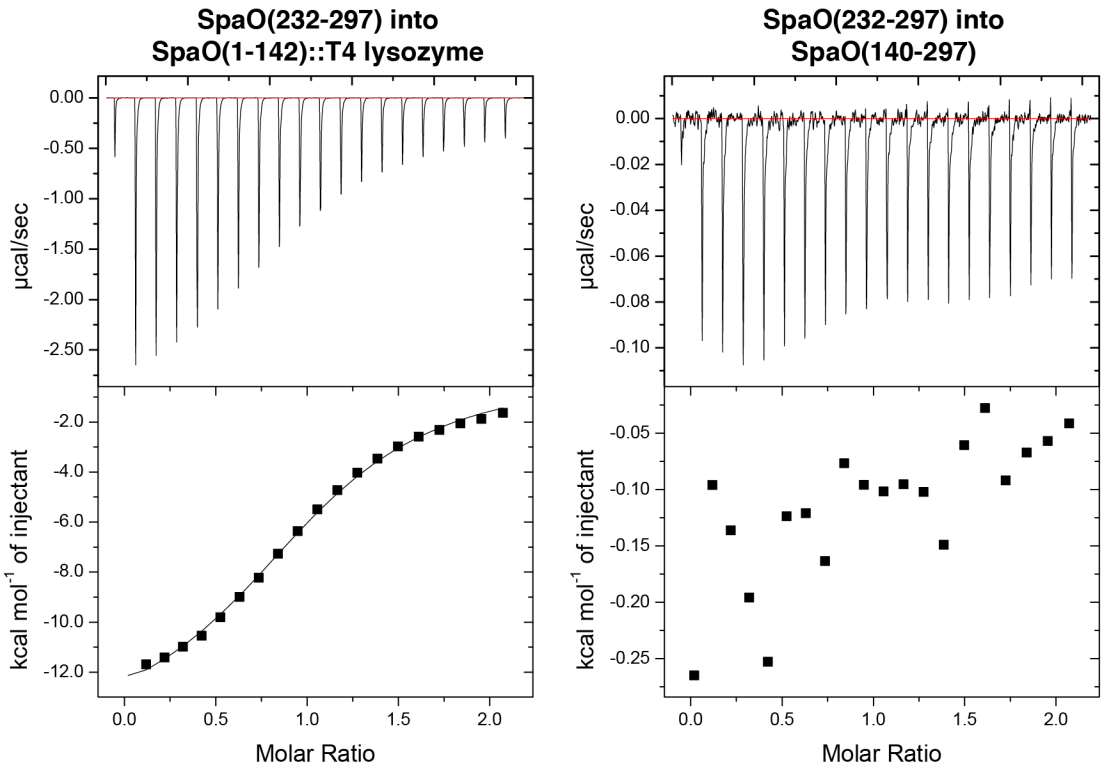


Figure 5.2: Quantitative analysis of the SpaO NTD-SPOA2 interaction. Representative ITC data for the SpaO SPOA2 interacting with the SpaO NTD (left) and SPOA1-SPOA2 (right). The uncorrected power differential between the reference and sample cells is shown on top; the background subtracted heats are shown on the bottom. The concentration of SPOA2 dimers in solution was used to calculate the molar ratio.

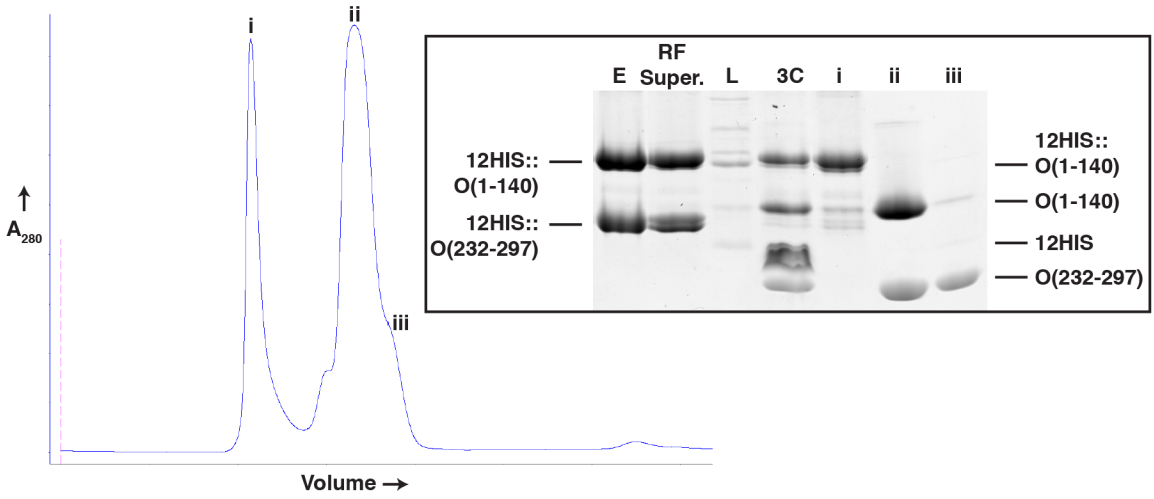


Figure 5.3: The SpaO NTD co-refolds with SPOA2. A SpaO (O) NTD construct can be purified under denaturing conditions and co-refolded with SPOA2 (denaturing NiNTA elution shown, E). After refolding and filtration (RF Super.), soluble material was liberated from its affinity tags (12HIS) by 3C protease and soluble aggregate (uncleaved, i) separated from the NTD-SPOA2 complex (ii). Some excess free SPOA2 remained as a chromatographic shoulder (iii).

CHAPTER 6:

CONCLUSIONS AND FUTURE DIRECTIONS

6.1 SPOA domain structure

I have presented here a series of structures and supporting biochemical and cell biological data that yield critical mechanistic insights into T3SS sorting platform assembly across multiple species and secretion subtypes. The existence of heterotypic SPOA interactions provides a structural explanation for the previously reported 1:2 stoichiometry of full length protein to free SPOA2 in SpaO homologues (141). While two of these SPOA2 domains could be accounted for by a homodimer interacting with full length SpaO, the conformation of the third SPOA2 (located in the full length protein) was unclear. Previous reports had proposed the existence of an alternate autostabilizing conformation for the third SPOA2, wherein the fingers of the SPOA domain in the full length protein "snap back" against its own palm to bury the surfaces located at the dimerization interface in SPOA homodimers (141). We show here that this third SPOA2 can be stabilized by a SPOA1-SPOA2 interaction, and that the SPOA2 homodimer interacts with the SpaO NTD (Figure 6.1).

What might be the function of the SPOA2 homodimer-NTD interaction in SpaO and its SctQ-family homologues? Studies in *Yersinia* argue that the cryptically translated SPOA2 fragment is necessary for T3SS function (141) and formation of perimembranous SctQ-containing puncta (137). However, preliminary data and communications with others in the field indicate that this is

not the case in *S. typhimurium.* Perhaps the SPOA2 homodimer functions as a chaperone, stabilizing the SpaO NTD until it can interact with some components of the T3SS basal body or export gate. This would be consistent with the marked decrease in solubility of full length SpaO when the internal translation start site is mutated (Figures 2.1, 3.1), and might explain why the SPOA2 homodimer is unnecessary for T3SS function under the "forcing" conditions of high salt stimulation. Further biochemical analysis of the SctQ NTD and its binding partners is needed to address this question, and whether the SPOA2 homodimer is necessary for *S. typhimurium* T3SS function under physiologic conditions (i.e. infection) should be determined.

The ratio of FliM to FliN *in situ* is estimated to be 1:3 (188). In the context of our FliM-FliN structure, this suggests a model for FliM-FliN interaction similar to that of SpaO. FliM(SPOA1) would engage FliN(SPOA2) in a heterotypic SPOA-SPOA interaction, much like the interaction between SPOA1 and SPOA2 in full length SpaO. Additional homodimeric FliN would interact with FliM-FliN in an as of yet undetermined fashion, analogous to the SpaO SPOA2 homodimer interaction with the SpaO NTD (Figure 6.1). However, several lines of evidence suggest that FliN homo-oligomers might not interact with the regions of FliM amino-terminal to its SPOA domain in a manner homologous to the SpaO NTD-$(SPOA2)_2$ interaction. First, the SpaO NTD and the pre-SPOA regions of FliM do not share noteworthy sequence homology, and portions of this region of FliM are known to interact with flagellar-specific proteins (e.g. FliG, CheY) involved in torque generation and flagellar rotation switching (145, 146). Second, FliN homo-

oligomers may interact with the FliM(SPOA)-FliN region instead, as co-affinity purification of hexahistidine-tagged FliM(SPOA) with FliN results in a co-purifying stoichiometric excess of FliN that separates from 1:1 FliM-FliN on ion-exchange chromatography (not shown). Moreover, reports of FliN tetramerization and FliM:FliN ratios between 1:3 and 1:4 suggest that more complicated higher order structures may be utilized by the flagellar apparatus (139).

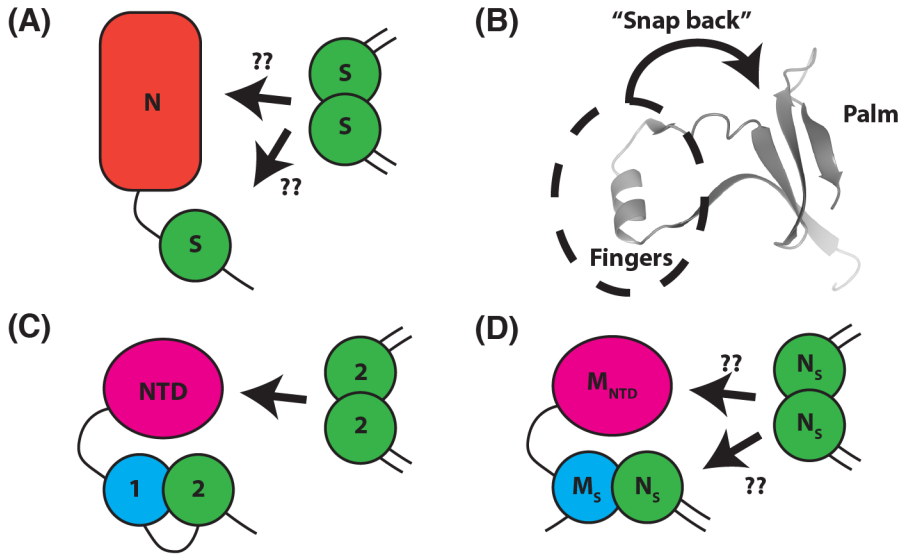


Figure 6.1: Overview of domain interactions within SPOA-containing proteins. (A) Schematic of the domain interactions for the SctQ family prior to this study. S indicates SPOA domains; N, the domains amino-terminal to the SPOA; ??, uncertain targets of interaction. (B) The model in (A) requires an autostablizing conformation of the SPOA domain, like the "snap back" model shown here. (C) Schematic of SpaO domain interactions described in this work. NTD, amino-terminal domain; 1, SPOA1; 2, SPOA2. (D) Schematic of FliM-FliN domain interactions, both hypothetical and demonstrated. M_{NTD} indicates the regions of FliM amino-terminal to the SPOA; M_S, FliM SPOA; N_S, FliN SPOA.

6.2 The SPOA-APAR interaction

The structures presented here suggest a partial model for the subtype-specific assembly of the T3SS sorting platforms: the SPOA1-SPOA2 module within a given T3SS subtype functions as an adaptor for the ATPase and its regulator through interaction with the APAR peptide (Figure 6.2). It should be noted that while previous investigations of the flagellar T3SS have focused on the interaction between FliH and FliN specifically (133), my structures and biochemical data show that the FliH APAR more strongly interacts with the FliM-FliN complex than with FliN alone, suggesting that the FliM-FliN complex is the physiologically relevant binding partner for FliH.

It is intriguing that disruption of the SPOA-APAR interaction prevented membrane localization of the SpaO sorting platform. What might be the mechanism for sorting platform targeting? One might imagine four non-mutually exclusive scenarios in which the SPOA-APAR interaction promotes SpaO membrane localization (ii-iv are diagrammed in Figure 6.3):

(i) InvC-OrgB has constitutive membrane localizing activity, and the SPOA-APAR interaction simply facilitates the trafficking of SpaO as inert cargo.

(ii) At baseline, the APAR binds to a region of InvC or OrgB preventing InvC-OrgB from localizing to the membrane (either by direct competition for a binding site or allosterically). SpaO binding of the APAR releases the inhibition, allowing membrane localization of the complex.

(iii) At baseline, the SpaO NTD is sequestered by interacting with SPOA1-SPOA2. Binding to the APAR releases the NTD, allowing it to interact with

membrane lipid or injectisome-associated proteins, driving membrane targeting.

(iv) The SPOA-APAR assembly might function as a scaffold module, inducing proximity between sorting platform components whose interactions then drive membrane localization.

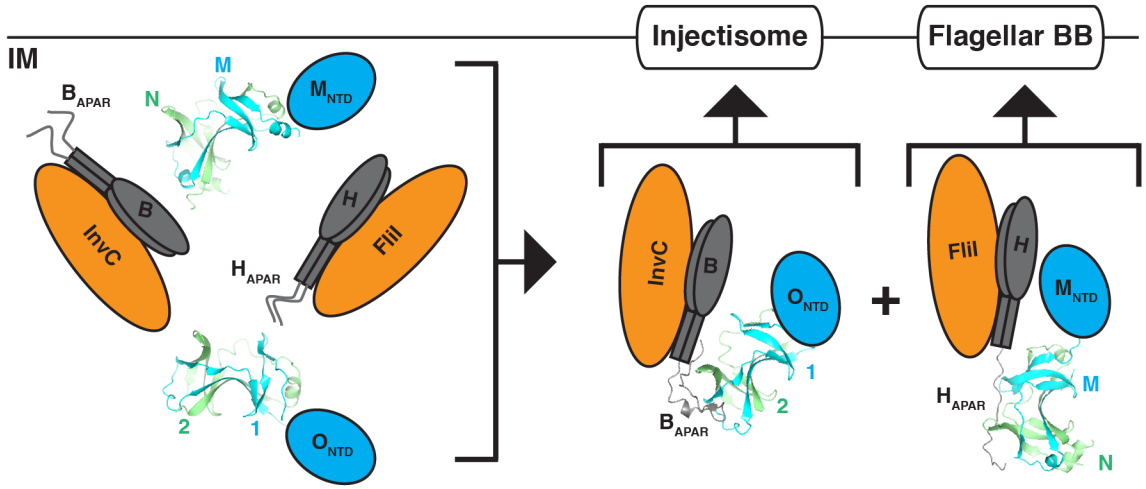


Figure 6.2: Segregation of homologous sorting platform components by subtype-specific SPOA-APAR interactions. Schematic illustration of the proposed role for the SPOA1-SPOA2-APAR assembly in organizing and localizing the T3SS sorting platforms in a subtype-specific fashion. IM indicates the inner membrane; O_{NTD}, the SpaO amino-terminal domain(s); 1 and 2, SpaO SPOA1 and SPOA2; B, OrgB; H, FliH; M_{NTD}, the FliM amino-terminal domains; M and N, the SPOA domains of FliM and FliN; Injectisome, the membrane integral components of the pathogenic T3SS; Flagellar BB, the flagellar basal body and associated integral membrane components.

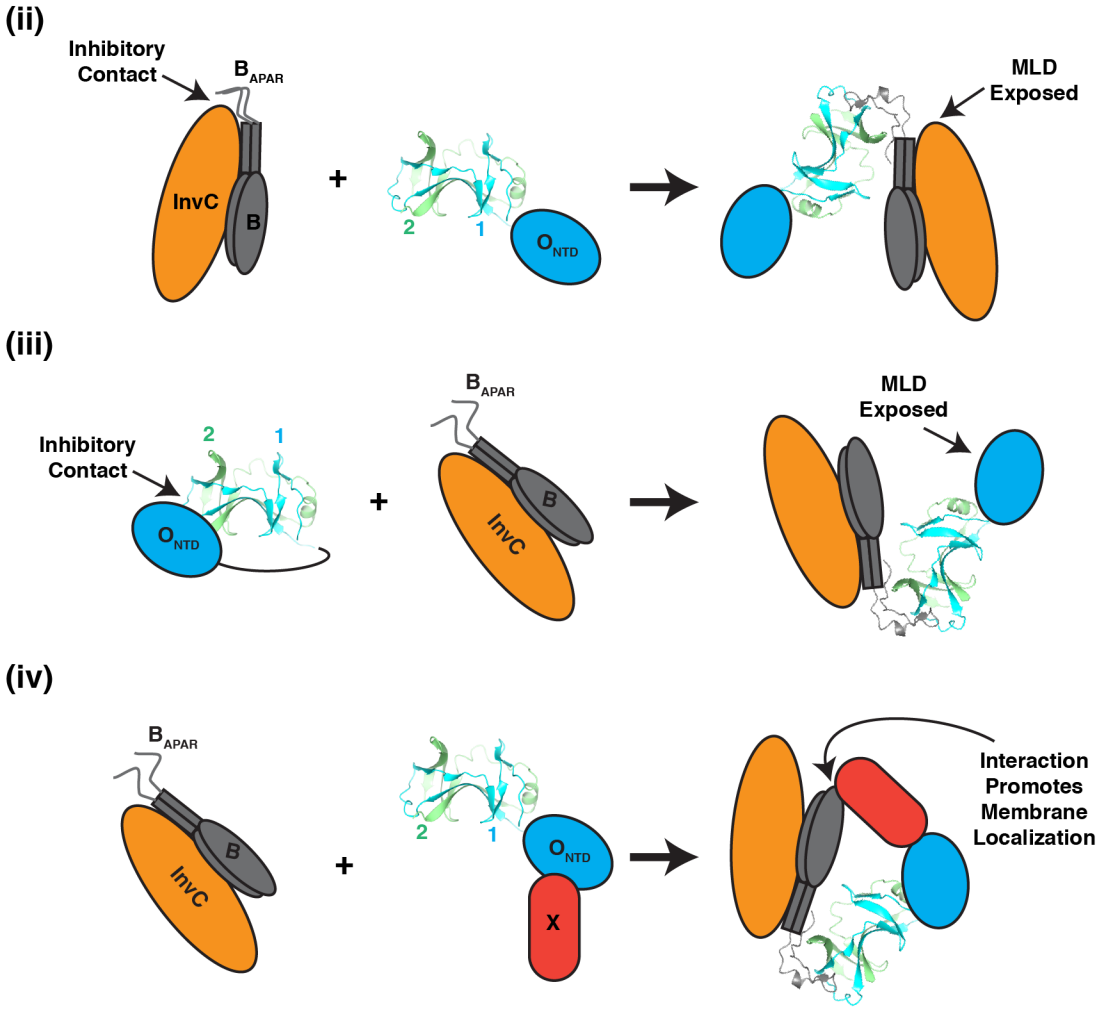


Figure 6.3: Hypothetical mechanisms of sorting platform targeting to the injectisome. Schematic representation of three hypothetical mechanisms for SPOA-APAR interaction-induced membrane localization of the sorting platform (ii, iii, and iv). Abbreviations are the same as in Figure 6.2, except MLD denotes membrane localization domain. X represents some hypothetical interacting protein or even a portion of the SpaO NTD.

One might distinguish among these hypothetical mechanisms with a series of genetic experiments utilizing genomic encoded GFP fusions. For example, hypothesis (i) would predict that deletion of neither *spaO* nor the *orgB* APAR would affect GFP::InvC localization to perimembranous puncta. However, hypothesis (ii) would predict that deletion of *spaO* would impair GFP::InvC localization to the membrane, and that phenotype would be rescued by deletion of the *orgB* APAR. Both hypotheses (i) and (ii) would predict that GFP::SpaO(SPOA1-SPOA2) would localize to the membrane in an OrgB-dependent fashion, but that GFP::SpaO(NTD) would show cytoplasmic localization. By contrast, hypothesis (iii) would predict that GFP::SpaO(NTD) would be membrane localized, regardless of *invC* or *orgB* genotype, and that GFP::SpaO(SPOA1-SPOA2) would be cytoplasmic. Lastly, hypothesis (iv) would predict that both GFP::SpaO(NTD) and GFP::SpaO(SPOA1-SPOA2) would be localized to the cytoplasm.

Hypothesis (iii) — that the SpaO NTD has membrane localizing activity and that this activity is sequestered prior to SPOA-APAR interaction — is particularly intriguing on several levels. First, early preliminary data suggests that GFP::SpaO(NTD) does localize to the membrane (not shown). Second, this hypothesis would imply a testable function for the NTD interaction with the inner membrane or some membrane associated T3SS component (discussed further in Section 6.3, below). Third, this hypothesis implies an interaction between the NTD and the SPOA1-SPOA2 module prior to APAR binding. Preliminary data do not demonstrate the existence of such an interaction when the modules are not

covalently linked (pull down experiments, not shown). However, in the context of the full length protein, these domains would be constrained by their covalent linkage, and so even low affinity, transient interactions between the domains might be functionally relevant. Of course, capturing full length SpaO in a crystal for diffraction analysis remains a priority for the field; however, solution NMR may offer some insight into whether these presently hypothetical interactions exist and are worth pursuing. For example, if the TROSY “fingerprint” for SpaO(140-297) (Figure 2.10) is altered in the context of the full length protein, it would suggest that the NTD makes contact with the SPOA1-SPOA2 module. In particular, one might imagine a case where the APAR-binding region shows signs of perturbation, consistent with some portion of the NTD docking in that site.

Regardless of the underlying mechanism, the finding that the SPOA-APAR interaction is necessary for T3SS function makes it an interesting target for the development of potential anti-virulence compounds. Unlike other potential anti-virulence targets in the T3SS that have human homologues (e.g. the SctN ATPase), the SPOA fold appears to be unique to the T3SS, with no reported structures possessing this fold in other systems, raising the prospect that therapeutics targeting this interaction could have little off-target toxicity to the host. In particular, the high degree of sequence (and presumably structural) conservation among the flagellar SPOA-APAR pairs suggests that it might even be possible to target the flagellar apparatus with “broad spectrum” efficacy.

The design of therapeutics targeting the SPOA-APAR interaction would not be without major challenges, though. While it is tempting to use the APAR

peptide structures themselves as models for the development of “APAR-mimetics” that could compete with the endogenous peptide for SPOA-binding, the relatively low affinity (K_d=2.0 μM) of the interaction is not within the nanomolar range typically seen as attractive in a pharmacologic lead compound. Indeed, the APAR peptides in both the SPI-1 T3SS and flagellar apparatus bind large exposed protein surfaces, not the deep pockets that are most amenable to tight, “lock and key” fits of small molecules.

Nonetheless, I have developed a Förster resonance energy transfer (FRET)-based assay for the high throughput screening of small molecule disruptors of the SPOA-APAR interaction. The architecture of the SPOA-APAR complexes (both flagellar and SPI-1) place APAR and SPOA domain termini within the Förster radius of the modified yellow fluorescent protein-cyan fluorescent protein pair, Venus-Cerulean (R_0=52 Å, ref. (190); Figure 6.4C,D), suggesting that fusions to these fluorescent proteins at the indicated termini might generate a detectable FRET signal *in vitro.* Carboxy-terminal fusion of Venus to SpaO(115-297) is compatible with protein expression and soluble purification from *E. coli.* In the presence of the previously characterized OrgB(1-30)::Cerulean fusion, substantial FRET is observed (Figure 6.4A,B). Neither fusion protein exhibits FRET in the presence of the other fluorescent protein without the cognate interaction partner, ruling out nonspecific interactions between either SpaO or OrgB and the fluorescent proteins as a spurious cause of FRET (Figure 6.4B).

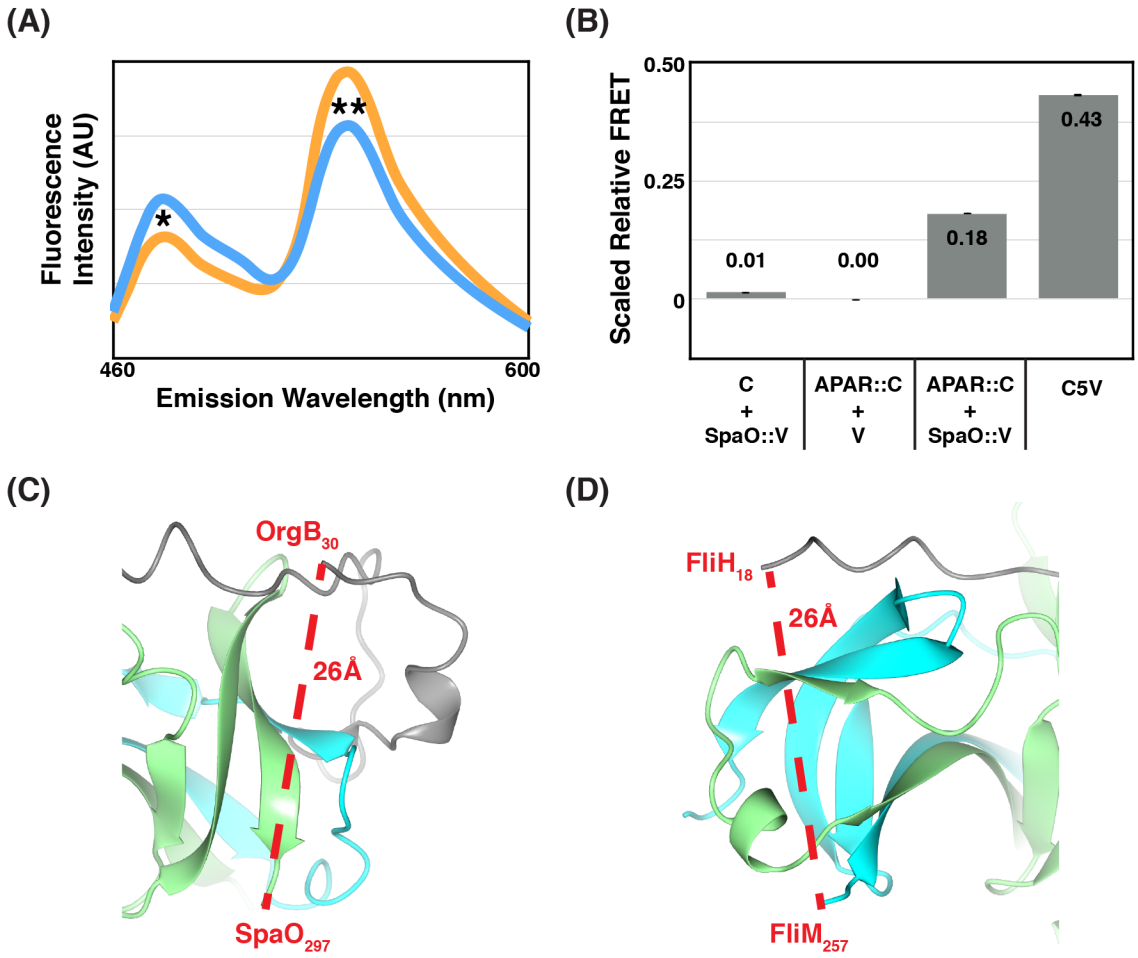


Figure 6.4: A FRET-based assay for high throughput drug screening. (A) Fluorescence emission scans of 100 μM each Cerulean and Venus (blue) or OrgB(1-30)::Cerulean and SpaO(115-297)::Venus (orange). Excitation wavelength fixed at 435 nm. Single asterisk highlights donor quenching, and double asterisk highlights acceptor stimulated emission. (B) Relative FRET efficiencies, scaled to Cerulean + Venus = 0. Abbreviations: C, Cerulean; V, Venus; APAR::C, OrgB(1-30)::Cerulean; SpaO::V, SpaO(115-297)::Venus; C5V, Cerulean and Venus connected by a 5 amino acid linker. Each mixture contained 100 μM of each fluor. Values are the mean of five technical replicates; error bars are +/- the standard deviation. (C, D) Structural basis for the design of fluorescent protein fusion FRET probes for the SPI-1 (C) and flagellar (D) SPOA-APAR interaction.

6.3 Higher order interactions in the T3SS sorting platform

A number of questions remain regarding the higher order architecture of the sorting platform *in situ.* I hypothesize that the puncta formed by SpaO *in vivo* represent the high-molecular weight sorting platforms described by Galán and colleagues (129). Armitage and colleagues have quantified the stoichiometry and dynamics of these puncta in *Yersinia*, showing them to possess ~22 copies of YscQ (SpaO homologue) per punctum and to be in dynamic exchange with the cytoplasm (137). In contrast, Liu and colleagues' recent tomographic reconstruction of *Shigella* injectisomes revealed the presence of only six SpaO homologue-dependent pods of density beneath the injectisome, and their localization was OrgB homologue-independent (128). Taken together with my findings, these results suggest that there may be two subpopulations of SpaO *in vivo*: one stably associated with the injectisome, and a second dynamic population in exchange with the cytoplasm, requiring the SPOA-APAR interaction to form high molecular weight, perimembranous sorting platforms. Recent analyses of FliI ATPase dynamics by Minamino and colleagues suggest a similar two population model, which they hypothesize functions to deliver secretion substrates to the assembling flagella (136).

Integrating SpaO structures into such models of sorting platform assembly function will require further exploration of the interactions between constituents of the large complex reported by Galán and colleagues (129). For example, the structure of the sorting platform component OrgA remains to be determined, and its role in platform assembly is similarly opaque. Its apparent insolubility and

previous reports of self association in yeast two hybrid studies (182) suggest that it may play a role in high molecular weight oligomerization of the sorting platform. While OrgA lacks a clear flagellar homologue, the flagellar protein FliG (which itself lacks a clear injectisome homologue) self associates to organize C-ring assembly (191), and it is temping to speculate that these non-homologous proteins might be functional analogues.

Similarly, the biochemical basis for sorting platform association with the basal body (e.g. SctD), the export apparatus (SctV), and/or their associated regulators (e.g. SctP, SctW, SctO, SctU) remains to be determined and may offer further insights into the mechanism of sorting platform localization to the injectisome superstructure. For example, previous work in *Shigella* identified a peptide from SctQ that interacts with the cytoplasmic domain of the basal body protein SctD (47). My structures would place this peptide in the “fingers” region of SPOA1, a plausible site for an interaction interface. However, I have been unable to reproduce an interaction between SpaO and PrgH (SctD) *in vitro* (not shown), and others have questioned the physiologic relevance of this interaction (48).

Lastly, the implications of sorting platform assembly for substrate recruitment, dechaperoning, and secretion remain unclear, and will require further studies of interactions between substrate-chaperone pairs and the sorting platform, the ATPase, and the export gate. Indeed, analysis of these complexes may even offer insight into the recruitment of the sorting platform to the injectisome, as interactions with substrate-chaperone pairs have been hypothesized to bridge the sorting platform and the injectisome basal body (48).

However, with few exceptions (125), such complexes of substrates, chaperones, and components of the sorting platform or export apparatus have not proven sufficiently stable for rigorous structural interrogation.

Ideally, micro-level insights from structural characterization of the aforementioned complexes will be paralleled with improvements in the electron microscopic (or, potentially, X-ray crystallographic) maps of the holo-injectisome and flagellar apparatus, allowing the incorporation of precise structural details into larger scale structural models of T3SS structure and function.

6.4 Technical considerations

The use of T4 lysozyme as a peptide stabilizing agent and crystallization chaperone merits brief technical consideration. For both the SpaO-OrgB and FliM-FliN-FliH structures, T4 lysozyme fusions were used to solubilize peptides with considerable hydrophobic character and also supplied additional crystallogenic packing interfaces. While T4 lysozyme fusions have received attention for their ability to facilitate the crystallization of membrane proteins, their applicability as general crystallization chaperones is less well studied. Other large protein fusions have been used to facilitate the crystallization of globular proteins (183), and I propose that T4 lysozyme might be similarly useful. However, I did note that the intrinsic flexibility between the amino- and carboxy-terminal lobes of T4 lysozyme resulted in substantially less order in the amino-terminal lobe and model building for this region was challenging (Figure 6.5). Kobilka and colleagues also noted this phenomenon in their G-protein coupled

receptor::T4 lysozyme structures and have engineered T4 lysozyme variants to overcome it (192). These re-engineered crystallization chaperones may prove useful to the crystallization of challenging globular proteins or protein-peptide interactions.

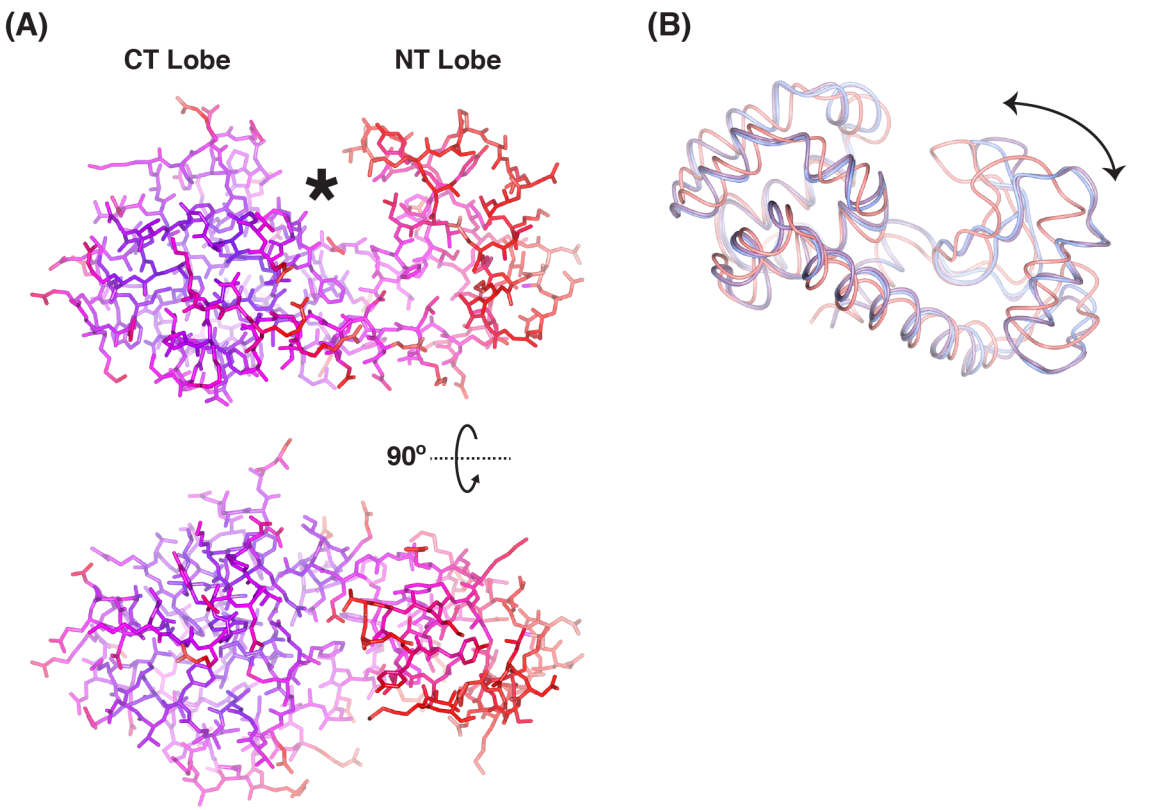


Figure 6.5: Disorder of the T4 lysozyme amino-terminal lobe complicates model building. (A) Model of one T4 lysozyme molecule from the SpaO-OrgB crystallographic asymmetric unit color coded by increasing (purple to red) B-factor. The amino- and carboxy-terminal lobes are indicated as “NT” and “CT,” respectively. Asterisk indicates the substrate binding pocket. (B) Comparison of the T4 lysozyme molecules from SpaO-OrgB (blue, purple) and FliN-FliM-FliH (pink) aligned globally. Note the greater divergence of the amino-terminal lobe between structures (arrow).

CHAPTER 7:

METHODS

7.1 Bioinformatics

Sequence alignments were performed using Clustal Omega (178) or M-COFFEE (193). Secondary structure and disorder predictions were performed using the PSIPRED server (194). Coiled coil prediction was preformed using the COILS server (195).

7.2 Molecular biology and general microbiology

Polymerase chain reaction (PCR) was performed using OneTaq (New England Biolabs), Phusion (New England Biolabs), or PfuTurbo (Agilent) as per manufacturer guidelines with oligonucleotides purchased from Integrated DNA Technologies. All mutations or gene fusions were created by overlap extension PCR. Gene sequences from *S. typhimurium* were PCR amplified from the T3SS-competent strains SB300 (wild-type; gift of J. Galán) or SB1741 (3xFLAG::SpaO, silent SpaO $L79_{CTG}$ to $L79_{CTA}$ variant; gift of J. Galán) (129). The T4 lysozyme (C54T, C97A) sequence was obtained from Addgene plasmid 18111. An additional mutation (D20N) in T4 lysozyme was made to decrease toxicity in *E. coli* (196), and the terminal three residues were mutated to alanines to decrease conformational entropy.

Plasmid DNA was purified using silica spin columns (Epoch Life Sciences) and the manufacturer recommended buffers. Standard molecular biology

protocols were followed to clone sequences of interest into modified pCDFduet or pETduet vectors for expression in *E. coli* or pBAD for expression in *S. typhimurium*. Restriction enzymes and Quick Ligase (New England Biolabs) were used as per manufacturer specifications. *E. coli* were rendered chemically competent for transformation with DNA by growing cells to an OD_{600} of 0.3 and washing cells 4 times with 10 mM tris.Cl pH=7.0, 60 mM $CaCl_2$ in 15% glycerol, 85% water at 4°C and snap freezing cells in liquid nitrogen. Cells were thawed and incubated with DNA for 0.5 h on ice, heat shocked at 42°C for 60 s, recovered on ice for 2-5 minutes, and then grown in LB (formulation below) for 1 h at 37°C prior to plating on LB agar with antibiotic selection.

LB broth (Miller formulation: 1% w/v tryptone, 0.5% w/v yeast extract, 1% w/v NaCl; Boston BioProducts) or LB agar (Lenox formulation: 1% w/v tryptone, 0.5% w/v yeast extract, 0.5% w/v NaCl 1.5% w/v agar; Boston BioProducts) were used to culture bacteria unless otherwise specified. Antibiotics were used a the following concentrations: ampicillin, 100 μg/mL; streptomycin, 10 0μg/mL; kanamycin, 50μg/mL; tetracycline, 10 μg/mL.

S. typhimurium genomic mutants were produced using homologous recombination from SacB-expressing suicide plasmids (36). SM10λpir *E. coli* were used to carry suicide plasmids and deliver them to *S. typhimurium* by conjugation. Plasmid carrying SM10λpir (tetracycline resistant) and target *Salmonella* (streptomycin resistant) were grown overnight on LB agar plates, resuspended in LB broth, and pelleted by centrifugation for 5 minutes at 2300 g. The *E. coli* and *Salmonella* were resuspended together in 50-75 μL LB broth,

spotted to antibiotic-free LB agar plates, incubated at 37°C overnight, and spread on LB agar containing streptomycin and tetracycline. Exconjugates were restreaked on the same media. Individual colonies were picked and grown for 2 h in 1 mL LB with streptomycin and tetracycline, diluted 100-fold into a fresh 1 mL LB containing only streptomycin, grown for 2 h, and 1-10 μL were plated to L-agar plates (1% w/v tryptone, 0.5% w/v yeast extract, 0.05% w/v NaCl, 1.5% agar) supplemented with 10% w/v sucrose and grown overnight at 30°C. Double recombinants were confirmed by PCR and tetracycline sensitivity. All *spaO* and *orgB* mutants were prepared on the SB1741 background. *fliM* and *fliN* mutants were prepared on the SB300 background.

S. typhimurium were rendered electrocompetent for transformation with DNA by growing cells to an OD_{600} of 0.3 and then washing 4 times with 10% glycerol at 4°C. Cells were incubated with salt-free DNA for 5 minutes before electroporation in 0.2 cm cuvettes with a Bio-Rad MicroPulser using the preprogrammed settings for gram-negative bacteria. Cells were recovered at 37°C for 1 h prior to plating on LB agar with selective antibiotics.

Bacterial strains carrying plasmids (Table 7.1) or genomic modifications (Table 7.2) were stored by diluting an aliquot of overnight culture in LB 1:1 with a solution of 65% glycerol, 20 mM tris.Cl pH=8.0, 100 mM $MgCl_2$ and frozen at -80°C.

Table 7.1: Plasmids used in this study. Listed in order of appearance in the figures.

Plasmid ID	*E. coli* Strain	Description
1366	BL21DE3Gold	12HIS(3C)::SpaO(1-303) in Stebbins modified pCDF-Duet
7000	BL21DE3Gold	12HIS(3C)::SpaO(1-303, Val_{203}GTT) in Stebbins modified pCDF-Duet
3013	BL21DE3Gold	12HIS(3C)::SpaO(1-303, Val_{203}GTT) in Stebbins modified pCDF-Duet position 1. Position 2 contains SpaO(222-303).
3483	BL21DE3Gold	12HIS(3C)::SpaO(232-297) in Stebbins modified pCDF-Duet
3804	BL21DE3Gold	12HIS(3C)::SpaO(145-213) in Stebbins modified pCDF-Duet
3702	BL21DE3Gold	12HIS(3C)::SpaO(140-297, Val_{203}GTT) in Stebbins modified pCDF-Duet
5860	BL21DE3Gold	3xFLAG::SpaO(1-219) in pBAD
6044	BL21DE3Gold	3xFLAG::SpaO(1-303) in pBAD
6046	BL21DE3Gold	3xFLAG::SpaO(1-303, stop after codon 140) in pBAD
5862	BL21DE3Gold	3xFLAG::SpaO(140-303) in pBAD
5213	BL21DE3Gold	3xFLAG::SpaO(204-303) in pBAD
5977	SM10λpir	Suicide plasmid for *spaO(Δ1-202)*
5973	SM10λpir	Suicide plasmid for *spaO(Δ1-139)*
5976	SM10λpir	Suicide plasmid for *spaO(Stop after codon 140)*
5983	SM10λpir	Suicide plasmid for *spaO(double stop after codon 219)*
5858	BL21DE3Gold	InvC (untagged) in modified pET-Duet
5412	BL21DE3Gold	12HIS(3C)::SpaO(1-303) in Stebbins modified pCDF-Duet position 1. OrgB is in position 2.
3796	BL21DE3Gold	12HIS(3C)::SpaO(140-297, Val_{203}GTT) in Stebbins modified pCDF-Duet position 1. OrgB is in position 2.
1658	BL21DE3Gold	12HIS(3C)::SpaO(1-303) in Stebbins modified pCDF-Duet position 1. OrgA is in position 2.
1835	BL21DE3Gold	12HIS(3C)::InvC in modified pET-Duet position 1. OrgB is in position 2.
1917	BL21DE3Gold	SpaO(1-303) (untagged) in modified pCDF-Duet position 1, SpaO(203-303) in position 2.

Table 7.1: Plasmids used in this study (continued).

Plasmid ID	*E. coli* Strain	Description
1930	BL21DE3Gold	OrgA (untagged) in modified pCOLA-Duet position 1, SpaO(203-303) in position 2.
3006	BL21DE3Gold	12HIS(3C)::OrgA in Stebbins lab modified pCDF-Duet position 1, SpaO(203-303) in position 2.
3599	BL21DE3Gold	12HIS(3C)::InvC in modified pET-Duet position 1. OrgB(31-226) is in position 2.
3474	BL21DE3Gold	12HIS(3C)::Thioredoxin(3C)::OrgB(2-30) in Stebbins modified pCDF-Duet
5180	BL21DE3Gold	OrgB(1-30)::T4 lysozyme::6HIS in modified pCDF-Duet
3848	BL21DE3Gold	12HIS(3C)::OrgB(1-30)::T4 lysozyme in Stebbins modified pCDF-Duet
5631	BL21DE3Gold	12HIS(3C)::OrgB(1-30)::Cerulean in Stebbins modified pCDF-Duet
3648	BL21DE3Gold	12HIS(3C)::OrgB(1-70) in Stebbins modified pCDF-Duet
5873	BL21DE3Gold	12HIS(3C)::SpaO(1-303) in Stebbins modified pCDF-Duet position 1. OrgB (I17D, L18D, I19D) is in position 2.
5868	BL21DE3Gold	12HIS(3C)::InvC in modified pET-Duet position 1. OrgB (I17D, L18D, I19D) is in position 2.
5946	SM10λpir	Suicide plasmid for *orgB(I17D, L18D, I19D)*
5392	SM10λpir	Suicide plasmid for *EGFP::3xFLAG::spaO*
5102	BL21DE3Gold	12HIS(3C)::InvA(316-686) in Stebbins modified pCDF-Duet
5152	BL21DE3Gold	FliH(1-43)::T4 lysozyme::6HIS
5178	BL21DE3Gold	SsaK(1-32)::T4 lysozyme::6HIS
3987	BL21DE3Gold	12HIS(3C)::FliM(245-320) in Stebbins modified pCDF-Duet position 1. FliN(1-137) is in position 2.
5968	BL21DE3Gold	12HIS(3C)::FliN(1-137) in Stebbins modified pCDF-Duet
5032	SM10λpir	Suicide plasmid for *ΔfliMN*
5962	SM10λpir	Suicide plasmid for *fliM::fliN(5-137)*
5100	BL21DE3Gold	12HIS(3C)::FliM(245-334)::FliN(5-137) in Stebbins modified pCDF-Duet
3879	BL21DE3Gold	12HIS(3C)::FliH(1-43)::T4 lysozyme in Stebbins modified pCDF-Duet

Table 7.1: Plasmids used in this study (continued).

Plasmid ID	*E. coli* Strain	Description
5449	BL21DE3Gold	12HIS(3C)::FliH(1-18)::T4 lysozyme in Stebbins modified pCDF-Duet
4086	BL21DE3Gold	12HIS(3C)::SpaO(1-142)::T4 lysozyme in Stebbins modified pCDF-Duet
3675	BL21DE3Gold	12HIS(3C)::SpaO(1-140) in Stebbins modified pCDF-Duet
5671	BL21DE3Gold	12HIS(3C)::SpaO(115-297)::Venus in Stebbins modified pCDF-Duet

Table 7.2: *S. typhimurium* strains used in this study. Listed in order of appearance in the figures.

Strain ID	Description
SB300	Wild-type. Gift of J. Gálan.
SB1741	*3xFLAG::spaO*. Gift of J. Gálan.
SB2130	*ΔspaO*. Gift of J. Gálan.
5161	*3xFLAG::spaO(Δ1-202)*
6011	5161 carrying plasmid 5860
6033	SB2130 carrying plasmid 5860
6008	*3xFLAG::spaO(double stop after codon 219)*
5995	*3xFLAG::spaO(Δ1-139)*
5999	*3xFLAG::spaO(stop after codon 140)*
6036	SB2130 carrying plasmid 5213
6035	SB2130 carrying plasmid 5862
6054	SB2130 carrying plasmid 6044
6056	SB2130 carrying plasmid 6046
SB2136	*3xFLAG::spaO, ΔorgB*. Gift of J. Gálan.
5683	*3xFLAG::spaO, orgB(I17D, L18D, I19D)*
5965	*EGFP::3xFLAG::spaO*
5958	*EGFP::3xFLAG::spaO, ΔorgB*
5925	*EGFP::3xFLAG::spaO, orgB(I17D, L18D, I19D)*
5057	*ΔfliMN*
5993	*fliM::fliN(5-137)*
6061	5995 carrying plasmid 6044
6062	5995 carrying plasmid 6046

7.3 Protein expression and purification

Constructs were transformed into BL21(DE3)Gold *E. coli* for heterologous expression and protein expression induced mostly as described (197). Specifically, bacteria were grown to an OD600 of 0.5-0.6 at 37°C in LB medium, the cultures were cooled to 18°C, induced with 250 μM IPTG, and grown overnight at 18°C. Selenomethionine (SeMet) substituted protein was produced in the methionine auxotrophic *E. coli* B834(DE3) grown in methionine-free media supplemented with SeMet (197). Uniformly labeled $^{15}N/^{13}C$- or $^{2}H/^{15}N/^{13}C$- protein samples were produced by overexpression in isotopically enriched minimal media. Deuterium oxide, ^{15}N-ammonium chloride, and ^{13}C-glucose were obtained from Cambridge Isotope Labs.

After induction overnight at 18°C, cells were harvested by centrifugation and resuspended in lysis buffer (200 mM NaCl, 20 mM Tris(hydroxymethyl)-aminomethane (tris).Cl pH=8.0, 5% v/v glycerol, 3 mM imidazole.Cl pH=8.0, 5 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 1 mM PMSF, and 0.1 mg/mL DNaseI). Cells were lysed by 1-2 passes through a mechanical homogenizer (Avestin C5) at 4°C.

Proteins were purified from *E. coli* cell lysates under native or denaturing conditions (as indicated for each downstream application below) and affinity purified on NiNTA agarose resin (Qiagen). For purification under native conditions, all steps were performed at 4°C. Lysate was clarified by centrifugation for 0.5 h at 30,000 g and loaded onto NiNTA resin by gravity flow. The column was washed with 5-10 volumes of wash buffer (200 mM NaCl, 20 mM tris.Cl

pH=8.0, 5%v/v glycerol, 30 mM imidazole.Cl pH=8.0) and then eluted in elution buffer (200 mM NaCl, 20 mM tris.Cl pH=8.0, 5%v/v glycerol, 360 mM imidazole.Cl pH=8.0). The elution was supplemented with 1 mM EDTA and dialyzed overnight against 200 mM NaCl, 20 mM tris.Cl pH=8.0, 1 mM dithiothreitol. Affinity tags were removed by cleavage with HRV 3C protease.

For purification under denaturing conditions, guanidinium chloride was added to the lysate to a final concentration of 6 M. The post extraction lystate was clarified by centrifugation at 30,000 g for 0.25 h at 4°C and loaded onto NiNTA resin in batch at 25°C. Still at 25°C, the resin was washed with denaturing wash buffer (8 M urea, 500 mM NaCl, 20 mM tris.Cl pH=8.0, 30 mM imidazole.Cl pH=8.0) and eluted in denaturing elution buffer (8 M urea, 200 mM NaCl, 20 mM tris.Cl pH=8.0, 360 mM imidazole.Cl pH=8.0). Urea containing solutions were stabilized by the addition of 75 mM ammonium sulfate. The elution was supplemented with 5 mM EDTA and 5 mM DTT and protein refolded by dialysis against 200 mM NaCl, 20 mM tris.Cl pH=8.0, 1 mM DTT (3-4 changes, dialysis time of 24 h total, 4°C). For T4 lysozyme fusions, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES).Na pH=7.0 was substituted for tris.Cl pH=8.0. Insoluble material was removed by centrifugation or filtration and affinity tags were removed by cleavage with HRV 3C protease.

Affinity purified proteins were further purified by ion exchange chromatography using an AKTA FPLC and the following columns (GE Healthcare): T4 lysozyme fusions were purified by cation exchange on a SourceS column; all other constructs were purified by anion exchange on a

SourceQ column. For cation exchange chromatography, proteins were loaded in batch in 10 mM HEPES.Na pH=7.0, 50-100 mM NaCl and eluted by a NaCl gradient (from 0 to 1000 mM) in the same buffer. For anion exchange, proteins were loaded in batch in 20 mM tris.Cl pH=8.0, 50-100 mM NaCl and eluted by a NaCl gradient (from 0 to 1000 mM) in the same buffer.

Prior to crystallography, ion exchange purified proteins were further purified by gel filtration chromatography on a Superdex 75 column (GE Healthcare) in final buffer (200 mM NaCl, 20 mM tris.Cl pH=8.0, 2 mM DTT) and concentrated using centrifugal concentrators (Amicon). To form the SpaO-OrgB::lysozyme complex for crystallization, cation exchange purified OrgB(1-30)::T4 lysozyme was mixed with an excess of anion exchange purified SpaO(145-213) + SpaO(232-297) and allowed to incubate overnight at 4°C. The SpaO-OrgB::lysozyme complex was then purified by gel filtration chromatography. To form the FliM::FliN-FliH::lysozyme complex for crystallization, anion exchange purified FliM(245-334)::FliN(5-137) was mixed with an excess of cation exchange purified FliH(1-18)::T4 lysozyme and allowed to incubate overnight at 4°C. The FliM::FliN-FliH::lysozyme complex was then purified by gel filtration chromatography.

Protein samples were analyzed by SDS-PAGE on 15% polyacrylamide gels. Samples were run in Laemmli running buffer (192 mM glycine, 25 mM tris, 0.1% w/v SDS) unless a component of interest was less than 15 kDa, in which case tris-tricine running buffer was used instead (100 mM tris, 100 mM tricine, 0.1% w/v SDS). Gels were stained with coomassie (2 mg/mL coomassie dye in

50% methanol, 40% water, 10% acetic acid) and destained in 30% methanol, 60% water, 10% acetic acid).

7.4 Crystallization

All proteins were crystallized by hanging drop vapor diffusion with 1:1 and 2:1 ratios of protein (in final buffer) to precipitant at 25°C (except where noted). For crystallization, SpaO(232-297) and FliM(245-334)::FliN(5-137) were purified under native conditions; SpaO(145-213) and SpaO(232-297) were purified under denaturing conditions and co-refolded; the T4 lysozyme fusions were purified under denaturing conditions, refolded, and mixed with their cognate SPOA1-SPOA2 as described above. The protein concentrations, crystallization buffers, and cryoprotection conditions for each protein or complex are as follows:

SpaO(232-297) was concentrated to 8 mg/mL and crystallized with 35% PEG400, 200 mM calcium acetate, 100 mM sodium acetate pH=5.0. Crystals were cryoprotected in the mother liquor. Microseeding was employed to enhance crystal uniformity and diffraction. Briefly, crystals to be seeded were harvested in precipitant solution and vortexed in a microfuge tube with a small stir bar for ~60 s. The slurry of microseeds was serially dilluted (5-10-fold steps) in precipitant solution and 5 selected microseed-precipitant mixtures were mixed with fresh protein as in a normal hanging drop experiment.

SpaO(145-213) + SpaO(232-297) was concentrated to 12 mg/mL and crystallized with 25% PEG400, 10% isopropanol, 100 mM sodium citrate pH=5.6 at 4°C. Microseeding (as above) was employed to enhance crystal uniformity and

diffraction. Crystals were cryoprotected in mother liquor with the PEG400 concentration raised to 37.5%.

SpaO(145-213) + SpaO(232-297) + OrgB(1-30)::T4 lysozyme was concentrated to 18.5 mg/mL and crystallized with 25% PEG3350, 200 mM ammonium formate, 100 mM sodium acetate pH=5.0. Microseeding (as above) was employed to enhance crystal uniformity and diffraction. Crystals were cryoprotected in 30% PEG3350, 10% glycerol, 200 mM ammonium acetate, 100 mM sodium acetate pH=5.0.

SpaO(145-213, SeMet) + SpaO(232-297, SeMet) + OrgB(1-30)::T4 lysozyme (native) was concentrated to 18 mg/mL, supplemented with 50 mM maltose, and crystallized with 25% PEG3350, 200 mM ammonium formate, 100 mM sodium acetate pH=5.0. Microseeding (as above) was employed to enhance crystal uniformity and diffraction. Crystals were cryoprotected in 25% PEG3350, 10% ethylene glycol, 200 mM ammonium formate, 100 mM sodium acetate pH=5.0, 50 mM maltose.

FliM(245-334)::FliN(5-137) was concentrated to 7.5 mg/mL and crystallized with 2.2 M NaCl, 100 mM imidazole.Cl pH=8.0. Crystals were cryoprotected with 2 M NaCl, 100 mM imidazole.Cl pH=8.0, 30% glycerol.

FliM(245-334)::FliN(5-137) + FliH(1-18)::T4 lysozyme was concentrated to 17 mg/mL and crystallized with 11% PEG400, 100 mM sodium potassium phosphate pH=6.5. Crystals were cryoprotected with 40% PEG400, 200 mM sodium potassium phosphate pH=6.5.

7.5 Structure determination

Data were collected at the National Synchrotron Light Source (Brookhaven National Laboratory) beamline X29A at a temperature of -173°C using the following X-ray wavelengths: 0.979 Å for SeMet crystals, 1.075 Å for native crystals. Diffraction data sets were indexed and integrated in iMOSFLM (198) and scaled and reduced with AIMLESS (199). Data sets were truncated at I/σI>2.0, and all sets were determined to have a $CC_{1/2}$>0.7 in the outermost resolution shell (200).

The PHENIX program suite (201) was used to solve the crystallographic phase problem. SpaO(232-297), SpaO(145-213) + SpaO(232-297), and FliM(245-334)::FliN(5-137) were solved by SeMet single wavelength anomalous diffraction in Autosol. The SPOA1-SPOA2-APAR::lysozyme structures were solved by molecular replacement in Phaser-MR using the experimentally phased cognate SPOA1-SPOA2 structure and T4 lysozyme (PDB 2LZM). Structures were built in Phenix (Autobuild) with additional manual model building performed in Coot (202). Structures were refined and validated in Phenix. SpaO(145-213) + SpaO(232-297) + OrgB(1-30)::T4 lysozyme crystals exhibited twinning and were refined in Phenix using the twin law l,-k,h. ANODE (186) was used to perform post-hoc analysis of anomalous scatters in SpaO(145-213, SeMet) + SpaO (232-297, SeMet) + OrgB(1-30)::T4 lysozyme crystals, providing additional empirical support for the SpaO-OrgB model coordinates. Except where indicated, all representations of models and maps for figures were produced in QtMG (203).

7.6 NMR spectroscopy

The NMR sample of refolded SpaO(140-297) consisted of 0.3 mM U-^{2}H/^{15}N/^{13}C labeled protein in 10 mM citrate buffer at pH 5.6 with 90%H2O/10%D2O (v/v), 100 mM NaCl, and 1 mM dithiothreitol. For comparison of the apo and APAR-bound forms, ^{15}N/^{13}C-labeled SpaO(140-297) was co-refolded with an excess of unlabeled thioredoxin::OrgB(2-30). The thioredoxin solublization tag was cleaved off by overnight incubation with HRV 3C protease. Protease and affinity tags were removed on NiNTA resin and the SpaO-OrgB complex was separated from the majority of free thioredoxin by Superdex75 gel filtration chromatography. The final concentration of the protein complex was 0.2 mM in 10 mM citrate buffer at pH 5.6 supplemented with 10% v/v deuterium oxide, 100 mM NaCl, and 2 mM dithiothreitol.

The NMR data were collected on Bruker 600, 800, and 900 MHz AVANCE spectrometers equipped with TCI/TXI CryoProbes™ at 20°C for the apo-SpaO and 30°C for the APAR-bound forms. For resonance assignments of apo-SpaO, transverse relaxation optimized (TROSY) triple resonance (204) experiments including trHNCO, trHN(CA)CO, trHNCA, trHN(CO)CA, trHNCACB and trHN(CO)CACB were acquired at 600 and 900 MHz. A ^{15}N-NOESY-HSQC spectrum with 100 ms mixing time was also acquired at 900 MHz. To assign APAR-bound SpaO, a suite of conventional backbone experiments (205) were acquired at 600 and 800 MHz.

The data were processed in Topspin 2.1 spectra and analyzed using the Autolink module in CARA 1.5 (206). The heteronuclear chemical shifts were

analyzed using the TALOS+ (181) database to predict the secondary structure of the protein. The weighted CSD were calculated from amide proton (δH) and nitrogen chemical shifts (δ^{15}N) using the following equation:

$$CSD = \sqrt{(\Delta \boldsymbol{\delta} \mathrm{H})^2 + (\frac{\Delta \boldsymbol{\delta}^{15}\mathrm{N}}{5})^2}$$

7.7 Co-affinity purification assays

For co-affinity purification of the SpaO-OrgB-InvC complex (Figures 3.1, 3.2, and 3.3), the proteins indicated were co-expressed and purified under native conditions as described above.

For the SPOA-APAR::lysozyme pulldown experiment (Figure 4.2), the indicated SPOA-containing proteins were Ni-affinity purified under native conditions, their affinity tags removed by overnight incubation with HRV protease 3C, and they were further purified by anion exchange chromatography (as above). APAR::lysozyme fusions were separately purified under denaturing conditions and were subjected to cation exchange chromatography after refolding (as above). 1 mg of hexahistidine-tagged APAR::lysozyme fusion protein was mixed with 2 mg of the indicated SPOA-containing protein in 0.2 M NaCl, 20 mM tris.Cl pH=8.0 (final volume 4 mL) and incubated on ice for 2 h. The mixture was twice passed over 2 mL of NiNTA resin, washed with 8 mL wash buffer (200 mM NaCl, 20 mM tris.Cl pH=8.0, 5% v/v glycerol, 30 mM imidazole.Cl pH=8.0) and then eluted in 3.5 mL elution buffer (200 mM NaCl, 20 mM tris.Cl pH=8.0, 5% v/v glycerol, 360 mM imidazole.Cl pH=8.0). For Figure 3.4, the same

procedure was followed but glycerol was excluded from the wash and elution buffers, and 2 x 5 mL washes were used instead of 1 x 8 mL, which reduced background binding.

For the InvA-APAR binding experiment (Figure 3.14), all proteins were purified under native conditions, the OrgB(1-30)::Cerulean fusion was cleaved from its affinity tag with 3C protease, and the proteins were further purified by anion exchange chromatography (as above). 125 μM OrgB(1-30)::Cerulean was mixed with 125 μM double hexahistidine-tagged InvA(316-686) or SpaO(140-297) in a final volume of 1 mL (0.2 M NaCl, 20 mM tris.Cl pH=8.0, 5 mM β-mercaptoethanol). After 4 h on ice, mixtures were passed over 1 mL NiNTA resin, washed with 6 mL of wash buffer (200 mM NaCl, 20 mM tris.Cl pH=8.0, 5% v/v glycerol, 30 mM imidazole.Cl pH=8.0) and then eluted in 4 mL of elution buffer (200 mM NaCl, 20 mM tris.Cl pH=8.0, 5% v/v glycerol, 360 mM imidazole.Cl pH=8.0).

For the SPOA2-NTD binding experiment (Figure 5.1), double hexahistidine-tagged SpaO(232-297) was purified under native conditions and anion exchange purified without removal of the affinity tags. SpaO(1-142)::T4 lysozyme was purified under denaturing conditions, refolded, affinity tags removed with 3C protease, and further purified by cation exchange chromatography. 1 mg of double hexahistidine-tagged SpaO(232-297) was mixed with 2 mg of SpaO(1-142)::T4 lysozyme in a total volume of 4 mL buffer (0.2 M NaCl, 20 mM tris.Cl pH 7.0). After 2 h on ice, mixtures were passed over 1.5 mL NiNTA resin, washed with 8 mL of wash buffer (200 mM NaCl, 20 mM

tris.Cl pH=7.0, 30 mM imidazole.Cl pH=7.0), washed again with 5 mL wash buffer and then eluted in 4 mL of elution buffer (200 mM NaCl, 20 mM tris.Cl pH=7.0, 360 mM imidazole.Cl pH=7.0).

7.8 Isothermal titration calorimetry

Natively purified, 3C-cleaved SpaO(140-297), SpaO(232-297), or OrgB(1-30)::Cerulean were anion exchange purified as described above. SpaO(1-142)::T4 lysozyme was purified under denaturing conditions, refolded, cleaved with 3C protease, and cation exchange purified as described above for APAR::T4 lysozyme fusions. For the experiments in Figure 3.12, proteins were dialyzed overnight against 0.2 M NaCl, 20 mM tris.Cl pH=8.0, 5 mM β-mercaptoethanol. SpaO constructs were concentrated to 1 mM and injected into 100 μM OrgB(1-30)::Cerulean (or buffer for background subtraction). For the experiments in Figure 5.2, proteins were dialyzed overnight against 150 mM NaCl, 10 mM HEPES.Na pH=7.0, 5 mM β-mercaptoethanol. SpaO(232-297) concentrated to 1 mM (dimer concentration) was injected into 100 μM SpaO(1-142)::T4 lysozyme or SpaO(140-297) (or buffer for background subtraction). All experiments were performed using an auto-ITC_{200} (GE Healthcare) and data were processed in Origin. Values reported in the text are the mean of 3 technical replicates ± standard deviation.

7.9 *In vitro* secretion assay

S. typhimurium of the indicated genotype were grown for 6 h at 37°C in LB medium with NaCl supplemented to a final concentration of 0.3 M. Cells were pelleted by centrifugation at 3400 g for 0.5-1 h and the supernatants were 0.22 μm filtered. Secreted proteins were precipitated from the filtered supernatants with 15% trichloroacetic acid overnight at 4°C. The precipitate was pelleted by centrifugation at 3400 g for 1 h at 4°C, resuspended in ice cold acetone and transferred to a microfuge tube. After 0.25 h on ice, the precipitate was harvested by centrifugation at 16,000 g for 0.75 h at 4°C and resuspended in 0.2 M tris.Cl pH=8.0, 0.2 M NaCl to neutralize any residual acid before the addition of SDS-PAGE loading buffer. For plasmid complementation analysis, *S. typhimurium* were electroporated with SpaO sequences cloned into the pBAD vector and expression was induced with 0.01% arabinose for the entire duration of the experiment.

7.10 Fluorescence microscopy

S. typhimurium were grown as for the *in vitro* secretion assay. Cells were harvested by centrifugation, washed 3 times in PBS, and fixed overnight with 4% formaldehyde in PBS at 4°C. Cells were again washed 3 times in PBS, counterstained with DAPI (Sigma Aldrich) and/or 10 mM Nile Red (Sigma-Aldrich), and immobilized on poly-L-lysine (P4707, Sigma Aldrich) coated coverslips. [Poly-L-lysine coating: Coverslips were incubated with 400 μL of 0.01% poly-L-lysine for 1-2 h at room temperature, washed by immersion in

distilled water (rapidly, five times) and air dried.] Covers were mounted in Prolong Diamond (Life Technologies) and sealed with nail polish. Slides were imaged on a DeltaVision Image Restoration Microscope with a 100x objective (Applied Precision). Images were deconvoluted in Softworx (Applied Precision) and processed identically in ImageJ (NIH) and Photoshop (Adobe).

7.11 Western blotting

After SDS-PAGE, samples were transferred to nitrocellulose membranes (Amersham Hybond ECL) by electrophoresis overnight with a constant 60 mA current in tris-glycine transfer buffer (20 mM tris, 150 mM glycine in 20% methanol, 80% water). Blots were blocked at room temperature for 1h in TBST (20 mM tris.Cl pH=7.4, 400 mM NaCl, 0.2% v/v Tween-20) supplemented with 5% w/v nonfat milk powder. Blots were incubated with primary antibody (mouse anti-FLAG M2, Sigma F1804, 1:2000) in TBST/5% milk for 4 h at room temperature, washed with TBST four times, and then incubated with secondary antibody (goat anti-mouse IgG conjugated to horseradish peroxidase, Sigma A4416, 1:2000) in TBST/5% milk for 2 h at room temperature. Blots were again washed four times in TBST and then developed using the ECL 2 Western Blotting Substrate kit (Pierce) as per manufacturers guidelines. Chemiluminescence was visualized on a LAS-3000 Intelligent Dark Box (Fuji).

7.12 FRET

Fluorescence readings were performed on a Spectramax Gemini XS plate reader (Molecular Devices). Samples were excited with 435 nm light and emission measurements made at 540 nm (acceptor fluorescence) and 480 nm (donor fluorescence). Relative FRET efficiency was calculated as:

$$E_{FRET} = \frac{I_A}{I_D + I_A}$$

where I_A is the fluorescence intensity of the Venus acceptor stimulated by the Cerulean donor (ex: 435 nm, em: 540 nm), and I_D is the fluorescence intensity of the Cerulean donor stimulated at its own maximum excitation wavelength (ex: 435 nm, em: 480 nm). FRET efficiency was scaled so that the background FRET of an equimolar mixture of Cerulean and Venus ("C+V") was set to 0:

$$E_{scaled} = \frac{E_{sample}}{E_{C+V}} - 1$$

A positive FRET control comprised of Cerulean fused to Venus with a 5 amino acid linker (207) was obtained on plasmid (Addgene #26394).

REFERENCES

1. Salyers AAW, D. D. 2002. Bacterial Pathogenesis: A Molecular Approach, 5th ed. ASM Press, Washington, DC.

2. Clatworthy AE, Pierson E, Hung DT. 2007. Targeting virulence: a new paradigm for antimicrobial therapy. Nat Chem Biol 3:541-548.

3. Prevention CfDCa. http://www.cdc.gov/drugresistance/. Accessed June 2, 2015.

4. Allen RC, Popat R, Diggle SP, Brown SP. 2014. Targeting virulence: can we make evolution-proof drugs? Nat Rev Microbiol 12:300-308.

5. Galan JE, Collmer A. 1999. Type III secretion machines: bacterial devices for protein delivery into host cells. Science 284:1322-1328.

6. Gauthier A, Thomas NA, Finlay BB. 2003. Bacterial injection machines. J Biol Chem 278:25273-25276.

7. Galan JE, Curtiss R, 3rd. 1989. Cloning and molecular characterization of genes whose products allow Salmonella typhimurium to penetrate tissue culture cells. Proc Natl Acad Sci U S A 86:6383-6387.

8. Michiels T, Wattiau P, Brasseur R, Ruysschaert JM, Cornelis G. 1990. Secretion of Yop proteins by Yersiniae. Infect Immun 58:2840-2849.

9. Michiels T, Vanooteghem JC, Lambert de Rouvroit C, China B, Gustin A, Boudry P, Cornelis GR. 1991. Analysis of virC, an operon involved in the secretion of Yop proteins by Yersinia enterocolitica. J Bacteriol 173:4994-5009.

10. Galan JE, Wolf-Watz H. 2006. Protein delivery into eukaryotic cells by type III secretion machines. Nature 444:567-573.

11. Cornelis GR. 2002. Yersinia type III secretion: send in the effectors. J Cell Biol 158:401-408.

12. Francis MS, Wolf-Watz H, Forsberg A. 2002. Regulation of type III secretion systems. Curr Opin Microbiol 5:166-172.

13. Waterman SR, Holden DW. 2003. Functions and effectors of the Salmonella pathogenicity island 2 type III secretion system. Cell Microbiol 5:501-511.

14. Stebbins CE. 2004. Structural insights into bacterial modulation of the host cytoskeleton. Curr Opin Struct Biol 14:731-740.

15. Yahr TL, Wolfgang MC. 2006. Transcriptional regulation of the Pseudomonas aeruginosa type III secretion system. Mol Microbiol 62:631-640.

16. Bhavsar AP, Guttman JA, Finlay BB. 2007. Manipulation of host-cell pathways by bacterial pathogens. Nature 449:827-834.

17. Ellermeier JR, Slauch JM. 2007. Adaptation to the host environment: regulation of the SPI1 type III secretion system in Salmonella enterica serovar Typhimurium. Curr Opin Microbiol 10:24-29.

18. Galan JE. 2007. SnapShot: effector proteins of type III secretion systems. Cell 130:192.

19. Buttner D. 2012. Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. Microbiol Mol Biol Rev 76:262-310.

20. Burkinshaw BJ, Strynadka NC. 2014. Assembly and structure of the T3SS. Biochim Biophys Acta 1843:1649-1663.

21. Hueck CJ. 1998. Type III protein secretion systems in bacterial pathogens of animals and plants. Microbiol Mol Biol Rev 62:379-433.

22. Abrusci P, McDowell MA, Lea SM, Johnson S. 2014. Building a secreting nanomachine: a structural overview of the T3SS. Curr Opin Struct Biol 25:111-117.

23. Galan JE, Lara-Tejero M, Marlovits TC, Wagner S. 2014. Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells. Annu Rev Microbiol 68:415-438.

24. Macnab RM. 2003. How bacteria assemble flagella. Annu Rev Microbiol 57:77-100.

25. Gophna U, Ron EZ, Graur D. 2003. Bacterial type III secretion systems are ancient and evolved by multiple horizontal-transfer events. Gene 312:151-163.

26. Izore T, Job V, Dessen A. 2011. Biogenesis, regulation, and targeting of the type III secretion system. Structure 19:603-612.

27. Kubori T, Matsushima Y, Nakamura D, Uralil J, Lara-Tejero M, Sukhan A, Galan JE, Aizawa SI. 1998. Supramolecular structure of the Salmonella typhimurium type III protein secretion system. Science 280:602-605.

28. Blocker A, Jouihri N, Larquet E, Gounon P, Ebel F, Parsot C, Sansonetti P, Allaoui A. 2001. Structure and composition of the Shigella flexneri "needle complex", a part of its type III secreton. Mol Microbiol 39:652-663.

29. Schlumberger MC, Muller AJ, Ehrbar K, Winnen B, Duss I, Stecher B, Hardt WD. 2005. Real-time imaging of type III secretion: Salmonella SipA injection into host cells. Proc Natl Acad Sci U S A 102:12548-12553.

30. Van Engelenburg SB, Palmer AE. 2010. Imaging type-III secretion reveals dynamics and spatial segregation of Salmonella effectors. Nat Methods 7:325-330.

31. Marlovits TC, Kubori T, Sukhan A, Thomas DR, Galan JE, Unger VM. 2004. Structural insights into the assembly of the type III secretion needle complex. Science 306:1040-1042.

32. Schraidt O, Marlovits TC. 2011. Three-dimensional model of Salmonella's needle complex at subnanometer resolution. Science 331:1192-1195.

33. Hodgkinson JL, Horsley A, Stabat D, Simon M, Johnson S, da Fonseca PC, Morris EP, Wall JS, Lea SM, Blocker AJ. 2009. Three-dimensional reconstruction of the Shigella T3SS transmembrane regions reveals 12-fold symmetry and novel features throughout. Nat Struct Mol Biol 16:477-485.

34. Kudryashev M, Stenta M, Schmelz S, Amstutz M, Wiesand U, Castano-Diez D, Degiacomi MT, Munnich S, Bleck CK, Kowal J, Diepold A, Heinz DW, Dal Peraro M, Cornelis GR, Stahlberg H. 2013. In situ structural analysis of the Yersinia enterocolitica injectisome. Elife 2:e00792.

35. Kowal J, Chami M, Ringler P, Muller SA, Kudryashev M, Castano-Diez D, Amstutz M, Cornelis GR, Stahlberg H, Engel A. 2013. Structure of the dodecameric Yersinia enterocolitica secretin YscC and its trypsin-resistant core. Structure 21:2152-2161.

36. Kaniga K, Bossio JC, Galan JE. 1994. The Salmonella typhimurium invasion genes invF and invG encode homologues of the AraC and PulD family of proteins. Mol Microbiol 13:555-568.

37. Crago AM, Koronakis V. 1998. Salmonella InvG forms a ring-like multimer that requires the InvH lipoprotein for outer membrane localization. Mol Microbiol 30:47-56.

38. Daefler S, Russel M. 1998. The Salmonella typhimurium InvH protein is an outer membrane lipoprotein required for the proper localization of InvG. Mol Microbiol 28:1367-1380.

39. Burghout P, Beckers F, de Wit E, van Boxtel R, Cornelis GR, Tommassen J, Koster M. 2004. Role of the pilot protein YscW in the biogenesis of the YscC secretin in Yersinia enterocolitica. J Bacteriol 186:5366-5375.

40. Spreter T, Yip CK, Sanowar S, Andre I, Kimbrough TG, Vuckovic M, Pfuetzner RA, Deng W, Yu AC, Finlay BB, Baker D, Miller SI, Strynadka NC. 2009. A conserved structural motif mediates formation of the periplasmic rings in the type III secretion system. Nat Struct Mol Biol 16:468-476.

41. Sanowar S, Singh P, Pfuetzner RA, Andre I, Zheng H, Spreter T, Strynadka NC, Gonen T, Baker D, Goodlett DR, Miller SI. 2010. Interactions of the transmembrane polymeric rings of the Salmonella enterica serovar Typhimurium type III secretion system. MBio 1.

42. Schraidt O, Lefebre MD, Brunner MJ, Schmied WH, Schmidt A, Radics J, Mechtler K, Galan JE, Marlovits TC. 2010. Topology and organization of the Salmonella

typhimurium type III secretion needle complex components. PLoS Pathog 6:e1000824.

43. Yip CK, Kimbrough TG, Felise HB, Vuckovic M, Thomas NA, Pfuetzner RA, Frey EA, Finlay BB, Miller SI, Strynadka NC. 2005. Structural characterization of the molecular platform for type III secretion system assembly. Nature 435:702-707.

44. Marlovits TC, Stebbins CE. 2010. Type III secretion systems shape up as they ship out. Curr Opin Microbiol 13:47-52.

45. Bergeron JR, Worrall LJ, De S, Sgourakis NG, Cheung AH, Lameignere E, Okon M, Wasney GA, Baker D, McIntosh LP, Strynadka NC. 2015. The modular structure of the inner-membrane ring component PrgK facilitates assembly of the type III secretion system basal body. Structure 23:161-172.

46. McDowell MA, Johnson S, Deane JE, Cheung M, Roehrich AD, Blocker AJ, McDonnell JM, Lea SM. 2011. Structural and functional studies on the N-terminal domain of the Shigella type III secretion protein MxiG. J Biol Chem 286:30606-30614.

47. Barison N, Lambers J, Hurwitz R, Kolbe M. 2012. Interaction of MxiG with the cytosolic complex of the type III secretion system controls Shigella virulence. Faseb j 26:1717-1726.

48. Gamez A, Mukerjea R, Alayyoubi M, Ghassemian M, Ghosh P. 2012. Structure and interactions of the cytoplasmic domain of the Yersinia type III secretion protein YscD. J Bacteriol 194:5949-5958.

49. Lountos GT, Tropea JE, Waugh DS. 2012. Structure of the cytoplasmic domain of Yersinia pestis YscD, an essential component of the type III secretion system. Acta Crystallogr D Biol Crystallogr 68:201-209.

50. Zhong D, Lefebre M, Kaur K, McDowell MA, Gdowski C, Jo S, Wang Y, Benedict SH, Lea SM, Galan JE, De Guzman RN. 2012. The Salmonella type III secretion system inner rod protein PrgJ is partially folded. J Biol Chem 287:25303-25311.

51. Bergeron JR, Worrall LJ, Sgourakis NG, DiMaio F, Pfuetzner RA, Felise HB, Vuckovic M, Yu AC, Miller SI, Baker D, Strynadka NC. 2013. A refined model of the prototypical Salmonella SPI-1 T3SS basal body reveals the molecular basis for its assembly. PLoS Pathog 9:e1003307.

52. Lilic M, Quezada CM, Stebbins CE. 2010. A conserved domain in type III secretion links the cytoplasmic domain of InvA to elements of the basal body. Acta Crystallogr D Biol Crystallogr 66:709-713.

53. Worrall LJ, Vuckovic M, Strynadka NC. 2010. Crystal structure of the C-terminal domain of the Salmonella type III secretion system export apparatus protein InvA. Protein Sci 19:1091-1096.

54. Abrusci P, Vergara-Irigaray M, Johnson S, Beeby MD, Hendrixson DR, Roversi P, Friede ME, Deane JE, Jensen GJ, Tang CM, Lea SM. 2013. Architecture of the major component of the type III secretion system export apparatus. Nat Struct Mol Biol 20:99-104.

55. Zarivach R, Deng W, Vuckovic M, Felise HB, Nguyen HV, Miller SI, Finlay BB, Strynadka NC. 2008. Structural analysis of the essential self-cleaving type III secretion proteins EscU and SpaS. Nature 453:124-127.

56. Wagner S, Konigsmaier L, Lara-Tejero M, Lefebre M, Marlovits TC, Galan JE. 2010. Organization and coordinated assembly of the type III secretion export apparatus. Proc Natl Acad Sci U S A 107:17745-17750.

57. Marlovits TC, Kubori T, Lara-Tejero M, Thomas D, Unger VM, Galan JE. 2006. Assembly of the inner rod determines needle length in the type III secretion injectisome. Nature 441:637-640.

58. Loquet A, Sgourakis NG, Gupta R, Giller K, Riedel D, Goosmann C, Griesinger C, Kolbe M, Baker D, Becker S, Lange A. 2012. Atomic model of the type III secretion system needle. Nature 486:276-279.

59. Edgren T, Forsberg A, Rosqvist R, Wolf-Watz H. 2012. Type III secretion in Yersinia: injectisome or not? PLoS Pathog 8:e1002669.

60. Radics J, Konigsmaier L, Marlovits TC. 2014. Structure of a pathogenic type 3 secretion system in action. Nat Struct Mol Biol 21:82-87.

61. Dohlich K, Zumsteg AB, Goosmann C, Kolbe M. 2014. A substrate-fusion protein is trapped inside the Type III Secretion System channel in Shigella flexneri. PLoS Pathog 10:e1003881.

62. Deane JE, Roversi P, Cordes FS, Johnson S, Kenjale R, Daniell S, Booy F, Picking WD, Picking WL, Blocker AJ, Lea SM. 2006. Molecular model of a type III secretion system needle: Implications for host-cell sensing. Proc Natl Acad Sci U S A 103:12529-12533.

63. Cordes FS, Komoriya K, Larquet E, Yang S, Egelman EH, Blocker A, Lea SM. 2003. Helical structure of the needle of the type III secretion system of Shigella flexneri. J Biol Chem 278:17103-17107.

64. Poyraz O, Schmidt H, Seidel K, Delissen F, Ader C, Tenenboim H, Goosmann C, Laube B, Thunemann AF, Zychlinsky A, Baldus M, Lange A, Griesinger C, Kolbe M. 2010. Protein refolding is required for assembly of the type three secretion needle. Nat Struct Mol Biol 17:788-792.

65. Demers JP, Habenstein B, Loquet A, Kumar Vasa S, Giller K, Becker S, Baker D, Lange A, Sgourakis NG. 2014. High-resolution structure of the Shigella type-III secretion needle by solid-state NMR and cryo-electron microscopy. Nat Commun 5:4976.

66. Fujii T, Cheung M, Blanco A, Kato T, Blocker AJ, Namba K. 2012. Structure of a type III secretion needle at 7-A resolution provides insights into its assembly and signaling mechanisms. Proc Natl Acad Sci U S A 109:4461-4466.

67. Mota LJ, Journet L, Sorg I, Agrain C, Cornelis GR. 2005. Bacterial injectisomes: needle length does matter. Science 307:1278.

68. Kubori T, Sukhan A, Aizawa SI, Galan JE. 2000. Molecular characterization and assembly of the needle complex of the Salmonella typhimurium type III protein secretion system. Proc Natl Acad Sci U S A 97:10225-10230.

69. Tamano K, Katayama E, Toyotome T, Sasakawa C. 2002. Shigella Spa32 is an essential secretory protein for functional type III secretion machinery and uniformity of its needle length. J Bacteriol 184:1244-1252.

70. Journet L, Agrain C, Broz P, Cornelis GR. 2003. The needle length of bacterial injectisomes is determined by a molecular ruler. Science 302:1757-1760.

71. Lefebre MD, Galan JE. 2014. The inner rod protein controls substrate switching and needle length in a Salmonella type III secretion system. Proc Natl Acad Sci U S A 111:817-822.

72. Mueller CA, Broz P, Muller SA, Ringler P, Erne-Brand F, Sorg I, Kuhn M, Engel A, Cornelis GR. 2005. The V-antigen of Yersinia forms a distinct structure at the tip of injectisome needles. Science 310:674-676.

73. Epler CR, Dickenson NE, Bullitt E, Picking WL. 2012. Ultrastructural analysis of IpaD at the tip of the nascent MxiH type III secretion apparatus of Shigella flexneri. J Mol Biol 420:29-39.

74. Goure J, Broz P, Attree O, Cornelis GR, Attree I. 2005. Protective anti-V antibodies inhibit Pseudomonas and Yersinia translocon assembly within host membranes. J Infect Dis 192:218-225.

75. Mueller CA, Broz P, Cornelis GR. 2008. The type III secretion system tip complex and translocon. Mol Microbiol 68:1085-1095.

76. Lawton WD, Surgalla MJ. 1963. Immunization against Plague by a Specific Fraction of Pasteurella Pseudotuberculosis. J Infect Dis 113:39-42.

77. DiGiandomenico A, Keller AE, Gao C, Rainey GJ, Warrener P, Camara MM, Bonnell J, Fleming R, Bezabeh B, Dimasi N, Sellman BR, Hilliard J, Guenther CM, Datta V, Zhao W, Gao C, Yu XQ, Suzich JA, Stover CK. 2014. A multifunctional bispecific antibody protects against Pseudomonas aeruginosa. Sci Transl Med 6:262ra155.

78. Derewenda U, Mateja A, Devedjiev Y, Routzahn KM, Evdokimov AG, Derewenda ZS, Waugh DS. 2004. The structure of Yersinia pestis V-antigen, an essential virulence factor and mediator of immunity against plague. Structure 12:301-306.

79. Johnson S, Roversi P, Espina M, Olive A, Deane JE, Birket S, Field T, Picking WD, Blocker AJ, Galyov EE, Picking WL, Lea SM. 2007. Self-chaperoning of the type III secretion system needle tip proteins IpaD and BipD. J Biol Chem 282:4035-4044.

80. Lunelli M, Hurwitz R, Lambers J, Kolbe M. 2011. Crystal structure of PrgI-SipD: insight into a secretion competent state of the type three secretion system needle tip and its interaction with host ligands. PLoS Pathog 7:e1002163.

81. Cheung M, Shen DK, Makino F, Kato T, Roehrich AD, Martinez-Argudo I, Walker ML, Murillo I, Liu X, Pain M, Brown J, Frazer G, Mantell J, Mina P, Todd T, Sessions RB, Namba K, Blocker AJ. 2014. Three-dimensional electron microscopy reconstruction and cysteine-mediated crosslinking provide a model of the type III secretion system needle tip complex. Mol Microbiol doi:10.1111/mmi.12843.

82. Rathinavelan T, Lara-Tejero M, Lefebre M, Chatterjee S, McShan AC, Guo DC, Tang C, Galan JE, De Guzman RN. 2014. NMR model of PrgI-SipD interaction and its implications in the needle-tip assembly of the Salmonella type III secretion system. J Mol Biol 426:2958-2969.

83. Knutton S, Rosenshine I, Pallen MJ, Nisan I, Neves BC, Bain C, Wolff C, Dougan G, Frankel G. 1998. A novel EspA-associated surface organelle of enteropathogenic Escherichia coli involved in protein translocation into epithelial cells. Embo j 17:2166-2176.

84. Daniell SJ, Kocsis E, Morris E, Knutton S, Booy FP, Frankel G. 2003. 3D structure of EspA filaments from enteropathogenic Escherichia coli. Mol Microbiol 49:301-308.

85. Picking WL, Nishioka H, Hearn PD, Baxter MA, Harrington AT, Blocker A, Picking WD. 2005. 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.

86. Barta ML, Dickenson NE, Patil M, Keightley A, Wyckoff GJ, Picking WD, Picking WL, Geisbrecht BV. 2012. The structures of coiled-coil domains from type III secretion system translocators reveal homology to pore-forming toxins. J Mol Biol 417:395-405.

87. Mizusaki H, Takaya A, Yamamoto T, Aizawa S. 2008. Signal pathway in salt-activated expression of the Salmonella pathogenicity island 1 type III secretion system in Salmonella enterica serovar Typhimurium. J Bacteriol 190:4624-4631.

88. Arnold R, Brandmaier S, Kleine F, Tischler P, Heinz E, Behrens S, Niinikoski A, Mewes HW, Horn M, Rattei T. 2009. Sequence-based prediction of type III secreted proteins. PLoS Pathog 5:e1000376.

89. Samudrala R, Heffron F, McDermott JE. 2009. Accurate prediction of secreted substrates and identification of a conserved putative secretion signal for type III secretion systems. PLoS Pathog 5:e1000375.

90. McDermott JE, Corrigan A, Peterson E, Oehmen C, Niemann G, Cambronne ED, Sharp D, Adkins JN, Samudrala R, Heffron F. 2011. Computational prediction of type III and IV secreted effectors in gram-negative bacteria. Infect Immun 79:23-32.

91. Lloyd SA, Norman M, Rosqvist R, Wolf-Watz H. 2001. Yersinia YopE is targeted for type III secretion by N-terminal, not mRNA, signals. Mol Microbiol 39:520-531.

92. Lee SH, Galan JE. 2004. Salmonella type III secretion-associated chaperones confer secretion-pathway specificity. Mol Microbiol 51:483-495.

93. Stebbins CE, Galan JE. 2001. Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type III secretion. Nature 414:77-81.

94. Stebbins CE, Galan JE. 2003. Priming virulence factors for delivery into the host. Nat Rev Mol Cell Biol 4:738-743.

95. DeBord KL, Lee VT, Schneewind O. 2001. Roles of LcrG and LcrV during type III targeting of effector Yops by Yersinia enterocolitica. J Bacteriol 183:4588-4598.

96. Lilic M, Vujanac M, Stebbins CE. 2006. A common structural motif in the binding of virulence factors to bacterial secretion chaperones. Mol Cell 21:653-664.

97. Janjusevic R, Quezada CM, Small J, Stebbins CE. 2013. Structure of the HopA1(21-102)-ShcA chaperone-effector complex of Pseudomonas syringae reveals conservation of a virulence factor binding motif from animal to plant pathogens. J Bacteriol 195:658-664.

98. Costa SC, Schmitz AM, Jahufar FF, Boyd JD, Cho MY, Glicksman MA, Lesser CF. 2012. A new means to identify type 3 secreted effectors: functionally interchangeable class IB chaperones recognize a conserved sequence. MBio 3.

99. Michiels T, Cornelis GR. 1991. Secretion of hybrid proteins by the Yersinia Yop export system. J Bacteriol 173:1677-1685.

100. Chen LM, Briones G, Donis RO, Galan JE. 2006. Optimization of the delivery of heterologous proteins by the Salmonella enterica serovar Typhimurium type III secretion system for vaccine development. Infect Immun 74:5826-5833.

101. Widmaier DM, Tullman-Ercek D, Mirsky EA, Hill R, Govindarajan S, Minshull J, Voigt CA. 2009. Engineering the Salmonella type III secretion system to export spider silk monomers. Mol Syst Biol 5:309.

102. Schubot FD, Jackson MW, Penrose KJ, Cherry S, Tropea JE, Plano GV, Waugh DS. 2005. Three-dimensional structure of a macromolecular assembly that regulates type III secretion in Yersinia pestis. J Mol Biol 346:1147-1161.

103. Birtalan SC, Phillips RM, Ghosh P. 2002. Three-dimensional secretion signals in chaperone-effector complexes of bacterial pathogens. Mol Cell 9:971-980.

104. Phan J, Tropea JE, Waugh DS. 2004. Structure of the Yersinia pestis type III secretion chaperone SycH in complex with a stable fragment of YscM2. Acta Crystallogr D Biol Crystallogr 60:1591-1599.

105. Schreiner M, Niemann HH. 2012. Crystal structure of the Yersinia enterocolitica type III secretion chaperone SycD in complex with a peptide of the minor translocator YopD. BMC Struct Biol 12:13.

106. Akeda Y, Galan JE. 2004. Genetic analysis of the Salmonella enterica type III secretion-associated ATPase InvC defines discrete functional domains. J Bacteriol 186:2402-2412.

107. Zarivach R, Vuckovic M, Deng W, Finlay BB, Strynadka NC. 2007. Structural analysis of a prototypical ATPase from the type III secretion system. Nat Struct Mol Biol 14:131-137.

108. Claret L, Calder SR, Higgins M, Hughes C. 2003. Oligomerization and activation of the FliI ATPase central to bacterial flagellum assembly. Mol Microbiol 48:1349-1355.

109. Ibuki T, Imada K, Minamino T, Kato T, Miyata T, Namba K. 2011. Common architecture of the flagellar type III protein export apparatus and F- and V-type ATPases. Nat Struct Mol Biol 18:277-282.

110. Pozidis C, Chalkiadaki A, Gomez-Serrano A, Stahlberg H, Brown I, Tampakaki AP, Lustig A, Sianidis G, Politou AS, Engel A, Panopoulos NJ, Mansfield J, Pugsley AP, Karamanou S, Economou A. 2003. Type III protein translocase: HrcN is a peripheral ATPase that is activated by oligomerization. J Biol Chem 278:25816-25824.

111. Akeda Y, Galan JE. 2005. Chaperone release and unfolding of substrates in type III secretion. Nature 437:911-915.

112. Minamino T, Namba K. 2008. Distinct roles of the FliI ATPase and proton motive force in bacterial flagellar protein export. Nature 451:485-488.

113. Paul K, Erhardt M, Hirano T, Blair DF, Hughes KT. 2008. Energy source of flagellar type III secretion. Nature 451:489-492.

114. Erhardt M, Mertens ME, Fabiani FD, Hughes KT. 2014. ATPase-Independent Type-III Protein Secretion in Salmonella enterica. PLoS Genet 10:e1004800.

115. Minamino T, Morimoto YV, Kinoshita M, Aldridge PD, Namba K. 2014. The bacterial flagellar protein export apparatus processively transports flagellar proteins even with extremely infrequent ATP hydrolysis. Sci Rep 4:7579.

116. Chen S, Beeby M, Murphy GE, Leadbetter JR, Hendrixson DR, Briegel A, Li Z, Shi J, Tocheva EI, Muller A, Dobro MJ, Jensen GJ. 2011. Structural diversity of bacterial flagellar motors. EMBO J 30:2972-2981.

117. Kawamoto A, Morimoto YV, Miyata T, Minamino T, Hughes KT, Kato T, Namba K. 2013. Common and distinct structural features of Salmonella injectisome and flagellar basal body. Sci Rep 3:3369.

118. Lorenzini E, Singer A, Singh B, Lam R, Skarina T, Chirgadze NY, Savchenko A, Gupta RS. 2010. Structure and protein-protein interaction studies on Chlamydia trachomatis protein CT670 (YscO Homolog). J Bacteriol 192:2746-2756.

119. Wilharm G, Lehmann V, Krauss K, Lehnert B, Richter S, Ruckdeschel K, Heesemann J, Trulzsch K. 2004. Yersinia enterocolitica type III secretion depends on the proton motive force but not on the flagellar motor components MotA and MotB. Infect Immun 72:4004-4009.

120. Galan JE. 2008. Energizing type III secretion machines: what is the fuel? Nat Struct Mol Biol 15:127-128.

121. Allison SE, Tuinema BR, Everson ES, Sugiman-Marangos S, Zhang K, Junop MS, Coombes BK. 2014. Identification of the docking site between a type III secretion system ATPase and a chaperone for effector cargo. J Biol Chem 289:23734-23744.

122. Roblin P, Dewitte F, Villeret V, Biondi EG, Bompard C. 2014. A Salmonella Type Three Secretion Effector/Chaperone Complex adopts Hexameric Ring-Like Structure. J Bacteriol doi:10.1128/JB.02294-14.

123. Tsai CL, Burkinshaw BJ, Strynadka NC, Tainer JA. 2014. Salmonella Type III Secretion System (T3SS) Virulence Effector Forms a New Hexameric Chaperone Assembly for Export of Effector/Chaperone Complexes. J Bacteriol doi:10.1128/JB.02524-14.

124. Chen L, Balabanidou V, Remeta DP, Minetti CA, Portaliou AG, Economou A, Kalodimos CG. 2011. Structural instability tuning as a regulatory mechanism in protein-protein interactions. Mol Cell 44:734-744.

125. Chen L, Ai X, Portaliou AG, Minetti CA, Remeta DP, Economou A, Kalodimos CG. 2013. Substrate-activated conformational switch on chaperones encodes a targeting signal in type III secretion. Cell Rep 3:709-715.

126. Ohgita T, Hayashi N, Hama S, Tsuchiya H, Gotoh N, Kogure K. 2013. A novel effector secretion mechanism based on proton-motive force-dependent type III secretion apparatus rotation. Faseb j 27:2862-2872.

127. Morita-Ishihara T, Ogawa M, Sagara H, Yoshida M, Katayama E, Sasakawa C. 2006. Shigella Spa33 is an essential C-ring component of type III secretion machinery. J Biol Chem 281:599-607.

128. Hu B, Morado DR, Margolin W, Rohde JR, Arizmendi O, Picking WL, Picking WD, Liu J. 2015. Visualization of the type III secretion sorting platform of Shigella flexneri. Proc Natl Acad Sci U S A 112:1047-1052.

129. Lara-Tejero M, Kato J, Wagner S, Liu X, Galan JE. 2011. A sorting platform determines the order of protein secretion in bacterial type III systems. Science 331:1188-1191.

130. Gonzalez-Pedrajo B, Fraser GM, Minamino T, Macnab RM. 2002. Molecular dissection of Salmonella FliH, a regulator of the ATPase FliI and the type III flagellar protein export pathway. Mol Microbiol 45:967-982.

131. Minamino T, MacNab RM. 2000. FliH, a soluble component of the type III flagellar export apparatus of Salmonella, forms a complex with FliI and inhibits its ATPase activity. Mol Microbiol 37:1494-1503.

132. Lane MC, O'Toole PW, Moore SA. 2006. Molecular basis of the interaction between the flagellar export proteins FliI and FliH from Helicobacter pylori. J Biol Chem 281:508-517.

133. Minamino T, Yoshimura SD, Morimoto YV, Gonzalez-Pedrajo B, Kami-Ike N, Namba K. 2009. Roles of the extreme N-terminal region of FliH for efficient localization of the FliH-FliI complex to the bacterial flagellar type III export apparatus. Mol Microbiol 74:1471-1483.

134. McMurry JL, Murphy JW, Gonzalez-Pedrajo B. 2006. The FliN-FliH interaction mediates localization of flagellar export ATPase FliI to the C ring complex. Biochemistry 45:11790-11798.

135. Minamino T, Gonzalez-Pedrajo B, Kihara M, Namba K, Macnab RM. 2003. The ATPase FliI can interact with the type III flagellar protein export apparatus in the absence of its regulator, FliH. J Bacteriol 185:3983-3988.

136. Bai F, Morimoto YV, Yoshimura SD, Hara N, Kami-Ike N, Namba K, Minamino T. 2014. Assembly dynamics and the roles of FliI ATPase of the bacterial flagellar export apparatus. Sci Rep 4:6528.

137. Diepold A, Kudryashev M, Delalez NJ, Berry RM, Armitage JP. 2015. Composition, formation, and regulation of the cytosolic c-ring, a dynamic component of the type III secretion injectisome. PLoS Biol 13:e1002039.

138. Fadouloglou VE, Tampakaki AP, Glykos NM, Bastaki MN, Hadden JM, Phillips SE, Panopoulos NJ, Kokkinidis M. 2004. Structure of HrcQB-C, a conserved component of the bacterial type III secretion systems. Proc Natl Acad Sci U S A 101:70-75.

139. Brown PN, Mathews MA, Joss LA, Hill CP, Blair DF. 2005. Crystal structure of the flagellar rotor protein FliN from Thermotoga maritima. J Bacteriol 187:2890-2902.

140. Stivala A, Wybrow M, Wirth A, Whisstock JC, Stuckey PJ. 2011. Automatic generation of protein structure cartoons with Pro-origami. Bioinformatics 27:3315-3316.

141. Bzymek KP, Hamaoka BY, Ghosh P. 2012. Two translation products of Yersinia yscQ assemble to form a complex essential to type III secretion. Biochemistry 51:1669-1677.

142. Paul K, Blair DF. 2006. Organization of FliN subunits in the flagellar motor of Escherichia coli. J Bacteriol 188:2502-2511.

143. Paul K, Harmon JG, Blair DF. 2006. Mutational analysis of the flagellar rotor protein FliN: identification of surfaces important for flagellar assembly and switching. J Bacteriol 188:5240-5248.

144. Fadouloglou VE, Bastaki MN, Ashcroft AE, Phillips SE, Panopoulos NJ, Glykos NM, Kokkinidis M. 2009. On the quaternary association of the type III secretion system HrcQB-C protein: experimental evidence differentiates among the various oligomerization models. J Struct Biol 166:214-225.

145. Lee SY, Cho HS, Pelton JG, Yan D, Henderson RK, King DS, Huang L, Kustu S, Berry EA, Wemmer DE. 2001. Crystal structure of an activated response regulator bound to its target. Nat Struct Biol 8:52-56.

146. Vartanian AS, Paz A, Fortgang EA, Abramson J, Dahlquist FW. 2012. Structure of flagellar motor proteins in complex allows for insights into motor structure and switching. J Biol Chem 287:35779-35783.

147. Kubori T, Galan JE. 2002. Salmonella type III secretion-associated protein InvE controls translocation of effector proteins into host cells. J Bacteriol 184:4699-4708.

148. Deane JE, Abrusci P, Johnson S, Lea SM. 2010. Timing is everything: the regulation of type III secretion. Cell Mol Life Sci 67:1065-1075.

149. Magdalena J, Hachani A, Chamekh M, Jouihri N, Gounon P, Blocker A, Allaoui A. 2002. Spa32 regulates a switch in substrate specificity of the type III secreton of Shigella flexneri from needle components to Ipa proteins. J Bacteriol 184:3433-3441.

150. Agrain C, Callebaut I, Journet L, Sorg I, Paroz C, Mota LJ, Cornelis GR. 2005. Characterization of a Type III secretion substrate specificity switch (T3S4) domain in YscP from Yersinia enterocolitica. Mol Microbiol 56:54-67.

151. Mizuno S, Amida H, Kobayashi N, Aizawa S, Tate S. 2011. The NMR structure of FliK, the trigger for the switch of substrate specificity in the flagellar type III secretion apparatus. J Mol Biol 409:558-573.

152. Mukerjea R, Ghosh P. 2013. Functionally essential interaction between Yersinia YscO and the T3S4 domain of YscP. J Bacteriol 195:4631-4638.

153. Botteaux A, Sani M, Kayath CA, Boekema EJ, Allaoui A. 2008. Spa32 interaction with the inner-membrane Spa40 component of the type III secretion system of Shigella flexneri is required for the control of the needle length by a molecular tape measure mechanism. Mol Microbiol 70:1515-1528.

154. Wood SE, Jin J, Lloyd SA. 2008. YscP and YscU switch the substrate specificity of the Yersinia type III secretion system by regulating export of the inner rod protein YscI. J Bacteriol 190:4252-4262.

155. Botteaux A, Kayath CA, Page AL, Jouihri N, Sani M, Boekema E, Biskri L, Parsot C, Allaoui A. 2010. The 33 carboxyl-terminal residues of Spa40 orchestrate the multi-step assembly process of the type III secretion needle complex in Shigella flexneri. Microbiology 156:2807-2817.

156. Cherradi Y, Schiavolin L, Moussa S, Meghraoui A, Meksem A, Biskri L, Azarkan M, Allaoui A, Botteaux A. 2013. Interplay between predicted inner-rod and gatekeeper in controlling substrate specificity of the type III secretion system. Mol Microbiol 87:1183-1199.

157. Sorg I, Wagner S, Amstutz M, Muller SA, Broz P, Lussi Y, Engel A, Cornelis GR. 2007. YscU recognizes translocators as export substrates of the Yersinia injectisome. EMBO J 26:3015-3024.

158. Botteaux A, Sory MP, Biskri L, Parsot C, Allaoui A. 2009. MxiC is secreted by and controls the substrate specificity of the Shigella flexneri type III secretion apparatus. Mol Microbiol 71:449-460.

159. Cheng LW, Kay O, Schneewind O. 2001. Regulated secretion of YopN by the type III machinery of Yersinia enterocolitica. J Bacteriol 183:5293-5301.

160. Rimpilainen M, Forsberg A, Wolf-Watz H. 1992. A novel protein, LcrQ, involved in the low-calcium response of Yersinia pseudotuberculosis shows extensive homology to YopH. J Bacteriol 174:3355-3363.

161. Skryzpek E, Straley SC. 1993. LcrG, a secreted protein involved in negative regulation of the low-calcium response in Yersinia pestis. J Bacteriol 175:3520-3528.

162. Deane JE, Roversi P, King C, Johnson S, Lea SM. 2008. Structures of the Shigella flexneri type 3 secretion system protein MxiC reveal conformational variability amongst homologues. J Mol Biol 377:985-992.

163. Weise CF, Login FH, Ho O, Grobner G, Wolf-Watz H, Wolf-Watz M. 2014. Negatively Charged Lipid Membranes Promote a Disorder-Order Transition in the Yersinia YscU Protein. Biophys J 107:1950-1961.

164. Stensrud KF, Adam PR, La Mar CD, Olive AJ, Lushington GH, Sudharsan R, Shelton NL, Givens RS, Picking WL, Picking WD. 2008. Deoxycholate interacts with IpaD of Shigella flexneri in inducing the recruitment of IpaB to the type III secretion apparatus needle tip. J Biol Chem 283:18646-18654.

165. Prouty AM, Gunn JS. 2000. Salmonella enterica serovar typhimurium invasion is repressed in the presence of bile. Infect Immun 68:6763-6769.

166. Barta ML, Guragain M, Adam P, Dickenson NE, Patil M, Geisbrecht BV, Picking WL, Picking WD. 2012. Identification of the bile salt binding site on IpaD from Shigella flexneri and the influence of ligand binding on IpaD structure. Proteins 80:935-945.

167. Chatterjee S, Zhong D, Nordhues BA, Battaile KP, Lovell S, De Guzman RN. 2011. The crystal structures of the Salmonella type III secretion system tip protein SipD in complex with deoxycholate and chenodeoxycholate. Protein Sci 20:75-86.

168. Menard R, Sansonetti P, Parsot C. 1994. The secretion of the Shigella flexneri Ipa invasins is activated by epithelial cells and controlled by IpaB and IpaD. EMBO J 13:5293-5302.

169. Rosqvist R, Magnusson KE, Wolf-Watz H. 1994. Target cell contact triggers expression and polarized transfer of Yersinia YopE cytotoxin into mammalian cells. EMBO J 13:964-972.

170. Zierler MK, Galan JE. 1995. Contact with cultured epithelial cells stimulates secretion of Salmonella typhimurium invasion protein InvJ. Infect Immun 63:4024-4028.

171. Lara-Tejero M, Galan JE. 2009. Salmonella enterica serovar typhimurium pathogenicity island 1-encoded type III secretion system translocases mediate intimate attachment to nonphagocytic cells. Infect Immun 77:2635-2642.

172. Epler CR, Dickenson NE, Olive AJ, Picking WL, Picking WD. 2009. Liposomes recruit IpaC to the Shigella flexneri type III secretion apparatus needle as a final step in secretion induction. Infect Immun 77:2754-2761.

173. Blocker AJ, Deane JE, Veenendaal AK, Roversi P, Hodgkinson JL, Johnson S, Lea SM. 2008. What's the point of the type III secretion system needle? Proc Natl Acad Sci U S A 105:6507-6513.

174. Cordes FS, Daniell S, Kenjale R, Saurya S, Picking WL, Picking WD, Booy F, Lea SM, Blocker A. 2005. Helical packing of needles from functionally altered Shigella type III secretion systems. J Mol Biol 354:206-211.

175. Yu XJ, McGourty K, Liu M, Unsworth KE, Holden DW. 2010. pH sensing by intracellular Salmonella induces effector translocation. Science 328:1040-1043.

176. Figueira R, Holden DW. 2012. Functions of the Salmonella pathogenicity island 2 (SPI-2) type III secretion system effectors. Microbiology 158:1147-1161.

177. Notti RQ, Bhattacharya S, Lilic M, Stebbins CE. 2015. A common assembly module in injectisome and flagellar type III secretion sorting platforms. Nat Commun 6:7125.

178. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539.

179. Shi Y. 2014. A glimpse of structural biology through X-ray crystallography. Cell 159:995-1014.

180. Drenth J, Mesters J. 2007. Principles of protein x-ray crystallography, 3rd ed. Springer, New York.

181. Shen Y, Delaglio F, Cornilescu G, Bax A. 2009. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J Biomol NMR 44:213-223.

182. Jouihri N, Sory MP, Page AL, Gounon P, Parsot C, Allaoui A. 2003. MxiK and MxiN interact with the Spa47 ATPase and are required for transit of the needle components MxiH and MxiI, but not of Ipa proteins, through the type III secretion apparatus of Shigella flexneri. Mol Microbiol 49:755-767.

183. Smyth DR, Mrozkiewicz MK, McGrath WJ, Listwan P, Kobe B. 2003. Crystal structures of fusion proteins with large-affinity tags. Protein Sci 12:1313-1322.

184. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC. 2007. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science 318:1258-1265.

185. Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, Ijzerman AP, Stevens RC. 2008. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. Science 322:1211-1217.

186. Thorn A, Sheldrick GM. 2011. ANODE: anomalous and heavy-atom density calculation. J Appl Crystallogr 44:1285-1287.

187. Hara N, Morimoto YV, Kawamoto A, Namba K, Minamino T. 2012. Interaction of the extreme N-terminal region of FliH with FlhA is required for efficient bacterial flagellar protein export. J Bacteriol 194:5353-5360.

188. Zhao R, Pathak N, Jaffe H, Reese TS, Khan S. 1996. FliN is a major structural protein of the C-ring in the Salmonella typhimurium flagellar basal body. J Mol Biol 261:195-208.

189. Kihara M, Francis NR, DeRosier DJ, Macnab RM. 1996. Analysis of a FliM-FliN flagellar switch fusion mutant of Salmonella typhimurium. J Bacteriol 178:4582-4589.

190. Markwardt ML, Kremers GJ, Kraft CA, Ray K, Cranfill PJ, Wilson KA, Day RN, Wachter RM, Davidson MW, Rizzo MA. 2011. An improved cerulean fluorescent

protein with enhanced brightness and reduced reversible photoswitching. PLoS One 6:e17896.

191. Paul K, Gonzalez-Bonet G, Bilwes AM, Crane BR, Blair D. 2011. Architecture of the flagellar rotor. EMBO J 30:2962-2971.

192. Thorsen TS, Matt R, Weis WI, Kobilka BK. 2014. Modified T4 Lysozyme Fusion Proteins Facilitate G Protein-Coupled Receptor Crystallogenesis. Structure 22:1657-1664.

193. Wallace IM, O'Sullivan O, Higgins DG, Notredame C. 2006. M-Coffee: combining multiple sequence alignment methods with T-Coffee. Nucleic Acids Res 34:1692-1699.

194. Buchan DW, Minneci F, Nugent TC, Bryson K, Jones DT. 2013. Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Res 41:W349-357.

195. Lupas A, Van Dyke M, Stock J. 1991. Predicting coiled coils from protein sequences. Science 252:1162-1164.

196. Shoichet BK, Baase WA, Kuroki R, Matthews BW. 1995. A relationship between protein stability and protein function. Proc Natl Acad Sci U S A 92:452-456.

197. Bhaskaran SS, Stebbins CE. 2012. Structure of the catalytic domain of the Salmonella virulence factor SseI. Acta Crystallogr D Biol Crystallogr 68:1613-1621.

198. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr 67:271-281.

199. Evans PR, Murshudov GN. 2013. How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr 69:1204-1214.

200. Karplus PA, Diederichs K. 2012. Linking crystallographic model and data quality. Science 336:1030-1033.

201. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Echols N, Headd JJ, Hung LW, Jain S, Kapral GJ, Grosse Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner RD, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2011. The Phenix software for automated determination of macromolecular structures. Methods 55:94-106.

202. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66:486-501.

203. McNicholas S, Potterton E, Wilson KS, Noble ME. 2011. Presenting your structures: the CCP4mg molecular-graphics software. Acta Crystallogr D Biol Crystallogr 67:386-394.

204. Salzmann M, Pervushin K, Wider G, Senn H, Wuthrich K. 1998. TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. Proc Natl Acad Sci U S A 95:13585-13590.

205. Sattler M, Schleucher J, Griesinger C. 1999. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Progress in Nuclear Magnetic Resonance Spectroscopy 34:93-158.

206. Keller RLJ. 2004. The Computer Aided Resonance Assignment Tutorial. CANTINA Verlag, Goldau, Switzerland.

207. Koushik SV, Chen H, Thaler C, Puhl HL, 3rd, Vogel SS. 2006. Cerulean, Venus, and VenusY67C FRET reference standards. Biophys J 91:L99-L101.