UNIVERSITY OF CALIFORNIA, SAN DIEGO

Structural analysis of a forkhead-associated domain from the type III secretion system protein YscD

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Chemistry

by

Alicia Margarita Gamez

Committee in charge:

Professor Partho Ghosh, Chair
Professor Michael Burkart
Professor Daniel Donoghue
Professor Victor Nizet
Professor Susan Taylor

2011

The Dissertation of Alicia Margarita Gamez is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2011

DEDICATION

This thesis is dedicated to my family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Partho Ghosh for being the best Ph.D. advisor, mentor, boss, and supporter. I will forever appreciate my time in his laboratory, for both the scientific and personal growth he has fostered in me. His scientific enthusiasm and vast knowledge always encouraged me to learn new things. And most importantly, he has challenged me professionally. I always appreciated his direct and honest approach during those times, and I know that I will continue to grow because of it.

I would like to thank my committee members, in particular Dr. Donoghue, for their suggestions throughout my graduate school career.

I am extremely grateful to members of the Partho Ghosh lab. I have learned something from each of them, scientific or not. I am particularly grateful to Kryztof Bzymek for everything he has done for me over the years. He is a great scientist and wonderfully opinionated. I am thankful to Brent, Sanghamitra, and Yvonne for their advice, proof-reading, and scientific discussions. I will always value the wine tastings, picnics, and dinners our lab had together. In particular, the lab's patience with my country and love song radios choices was appreciated. And to the faux lab member, Raj Krishnan, who kept me company in lab at odd hours, I will always be grateful. Raj has been the one person I have been able to count on through graduate school, and I could not have asked for a better friend.

My thesis is dedicated to the people who have allowed me to get this far in my education, my family. My mother, Karen Offerdahl, encouraged me to get involved in everything and was always a great supporter. My father, Oscar Gamez, is my role

model for hard work and determination and has been very supportive and understanding throughout graduate school. My brother, Rick, who understands my feelings, shares my love of fashion and politics, has also been supportive. And my sister, Aliya, is the greatest friend I could ever have. She has been there for me all my life, with understanding, logical advice, patience, and love.

This thesis is also dedicated to my newest family members: Justin and Sheelo. Sheelo, you were my first love, the best dog I could ever have. Your positive disposition always cheered me up, no matter how my experiments went. You were always there to make me smile. Justin, my future husband, I have relied on you so much in the past two years, I know I would have been a mess without you. Your patience with my schedule has meant the world to me. I am so happy for our loving future together.

The text of chapter 1, in part, is currently being prepared for submission for publication. The dissertation author was the primary researcher. Partho Ghosh contributed and supervised the research which forms the basis for this chapter.

VITA

2004	B.S., Chemistry, California State Polytechnic University, Pomona
2011	Ph.D., Chemistry, University of California, San Diego

ABSTRACT OF THE DISSERTATION

Structural analysis of a forkhead-associated domain from the
type III secretion system protein YscD

by

Alicia Margarita Gamez

Doctor of Philosophy in Chemistry

University of California, San Diego, 2011

Professor Partho Ghosh, Chair

The type III secretion system, T3SS, is used by many bacterial pathogens to evade the host immune response. The T3SS utilizes a multiprotein complex, the injectisome, to transport virulence proteins directly into the host cytosol. Injectisome proteins extend from the bacterial cytosol, across the bacterial inner and outer membranes, and into the extracellular space. A crucial injectisome protein from the *Yersinia pseudotuberculosis* T3SS is YscD. YscD is an inner membrane protein that has both cytoplasmic and periplasmic domains. The X-ray crystallographic structure of the cytoplasmic domain of YscD, YscDc, is presented here. YscDc has a forkhead-

associated (FHA) domain fold. Based on structural comparisons between the FHA domain of YscDc and FHA domains of known function, two regions of possible importance were identified within YscDc: loops β3β4 and β4β5. Type III secretion in *Y. pseudotuberculosis* was abolished when loop β3β4 and β4β5 amino acids were substituted by alanines. The FHA domain of YscD was determined to be crucial for the function of the T3SS in *Y. pseudotuberculosis.*

The FHA domain of YscDc is highly similar in sequence to YscD homologs of other T3SS-encoding bacteria. This domain may be significant for injectisome structural integrity or may be engaged in critical protein-protein interactions with cytosolic T3SS proteins. Understanding the role that this FHA domain plays in the T3SS will provide insight into the role of a protein crucial for injectisome assembly and T3SS function.

I.

Structural analysis of YscDc

Abstract

YscD is an inner membrane protein of the *Yersinia pseudotuberculosis* type III secretion system (T3SS). The N-terminal domain, YscDc, is cytosolic and has an unknown function in the T3SS. The X-ray crystallographic structure of YscDc (residues 1-108) was determined to 2.5 Å resolution limit. YscDc has a forkhead-associated (FHA) domain fold. This domain was determined to be necessary for type III secretion. To identify which regions of the YscD FHA domain are important for secretion, *yscD* was deleted in *Y. pseudotuberculosis*. Type III secretion was abrogated in the *yscD* Based on structural comparisons between YscDc with other FHA domains, two specific loops in YscDc were targeted for mutagenesis. Residues belonging to loop β3β4 (SDPLQ) and loop β4β5 (SEIA) were substituted with alanines, and these YscD mutants were expressed in *Y. pseudotuberculosis ΔyscD.* YscD loop mutants were not able to rescue type III secretion. Individual amino acids from loop β4β5 were subsituted with alanines and tested in the same secretion assay. This experiment led to the identification of S40A on loop β4β5as being essential for secretion. Loops β3β4 and β4β5 may be involved in the formation of the T3SS apparatus, or may engage in essential interactions with cytosolic proteins.

Introduction

Three species of the gram-negative bacterium *Yersinia* are pathogenic to humans: *Y. pseudotuberculosis, Y. enterocolitica, and Y. pestis. Y. pseudotuberculosis and Y. enterocolitica* are enteropathogenic and are found in contaminated meat products, milk, and water. *Y. pestis* is the causative agent of the plague, which has caused numerous epidemics throughout history and has killed over 200 million people (Perry, 1997). It is estimated that *Y. pestis* evolved from *Y. pseudotuberculosis* 1,500-20,000 years ago (Achtman, 1999). Although improvements in sanitation and healthcare access have decreased the incidence of pandemics, the bubonic form of the plague is still endemic in many countries in Africa, Asia, South America and in the southwestern United States (Franz, 2001). Due to the ease of dissemination along with the potential to cause widespread panic, *Y. pestis* is characterized by the Centers for Disease Control and Prevention (CDC) as a potential biological weapon (Inglesby *et al.*, 2000).

Yersiniae have succeeded as pathogens by evolving numerous ways to alter the host environment and evade the immune system. During bacterial infection in healthy individuals, proteins of the innate immune system called pattern recognition receptors (PRRs) detect pathogen components by their pathogen-associated molecular patterns (PAMPs). This recognition leads to phagocytosis by activated macrophages and ultimately pathogen degradation. Activated macrophages secrete proinflammatory cytokines such as interleukins and tumor necrosis factor (TNF) to illicit an inflammatory response. Inflammatory cytokine production involves the mitogen-

activated protein kinases (MAPKs) and the nuclear factor-κB (NF-κB) signaling pathways. *Yersiniae* produce virulence proteins that are antiphagocytic and function to down-regulate these immune signaling pathways. The genes which encode for these virulence effector proteins, *Yersinia* outer membrane proteins (Yops), are located on the 70 kb virulence plasmid pYV (Gemski, 1980). The Yop proteins YopE, YopT, YpkA, YopH, YopJ, YopK, and YopM are crucial for host cell survival and proliferation (Heesemann, 1986, Lian, 1987, Rosqvist R, 1990).

In order for phagocytosis to occur, the actin cytoskeleton must be reorganized, which occurs under the control of the Rho family of small GTPases. YopE, YopT, and YpkA affect host RhoGTPase functions, leading to the disruption of the actin cytoskeleton and phagocytosis. In addition to inhibiting phagocytosis, Yops target the host innate immune system's signaling pathways. YopJ inhibits MAPK and NF-κB signaling, which affects the production of proinflammatory cytokines and causes macrophage apoptosis.

Yops are actively transported from the bacterial cytosol into the host cell by a pYV-encoded type III secretion system, T3SS (Figure 1.1). Currently there are 25 bacteria known to employ a T3SS, which can be grouped into the following families (based on sequence analysis of members of the YscV membrane protein family): *Yersinia* Ysc-Yop (which includes *Pseudomonas aerginosa*, *Aeromonas salmonicida*), Inv-Mxi-Spa (*Shigella flexneri*, *Salmonella enterica*), Esc (*Escherichia coli)*, Hrp1 (*Erwinia*), and Hrp2 (*Xanthomonas*, *Chlamydia*, and *Rhizobiaceae*) (Cornelis, 2002, Cornelis, 2010). The T3SS transports effectors into the host cell using an injectisome, a multi-protein complex which spans the membranes of the bacterial envelope,

terminating with an extracellular appendage resembling a needle (Kubori, 1998). This needle is utilized to export translocon proteins to the surface of host cells, where such proteins appear to form pores in host cell membranes (Neyt, 1999). It has been proposed that effector proteins are translocated from the bacterial cytosol via the injectisome needle and into the host cell through the translocon protein pores.

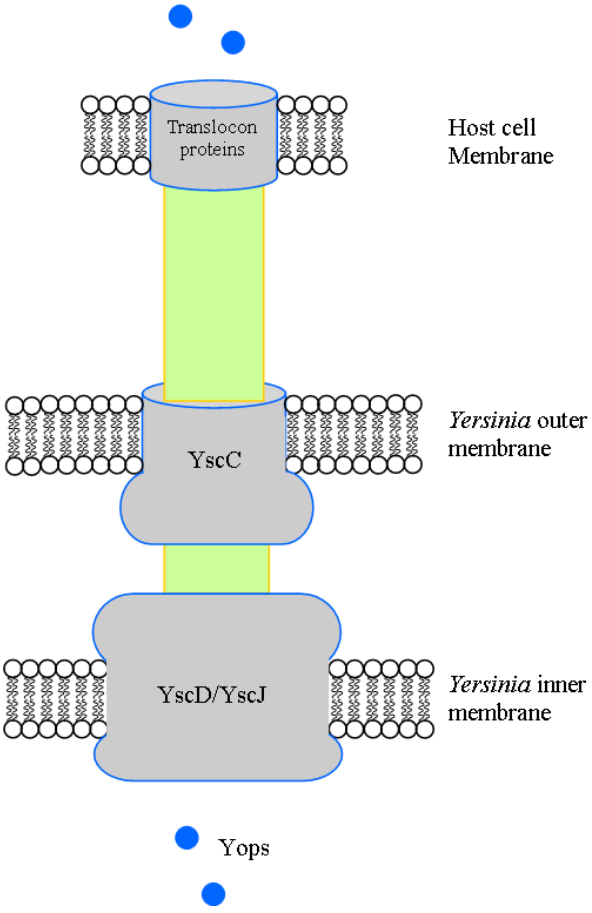

Figure 1.1: The T3SS in *Yersinia pseudotuberculosis.*

Assembly of the injectisome occurs in a regulated and hierarchical manner with sequential export of needle proteins, translocon proteins, and then effector proteins by the T3SS. Recent evidence suggests that the first needle complex protein

to assemble in *Y. enterocolitica* is the outer membrane protein YscC, followed by two of the inner membrane proteins, YscD and YscJ (Diepold, 2010). YscC belongs to the secretin protein family and forms an oligomer consisting of ~13 monomers within the outer membrane that extends into the periplasmic space (Koster, 1997, Burghout, 2004). After incorporation of YscC into the outer membrane, YscD and YscJ oligomerize in the inner membrane. The periplasmic domains of YscD and YscJ interact with the periplasmic domain of YscC, forming the needle complex scaffold. YscJ is an inner membrane lipoprotein with an N-terminal periplasmic domain and a C-terminal membrane anchor.

YscD is a 46.9 kDa protein with an N-terminal cytoplasmic domain, a single-pass membrane spanning region, and a C-terminal periplasmic domain (Figure 1.2). The *yscD* gene is located on the *virC* operon of the pYV plasmid of *Y. pseudotuberculosis*. Strains having in-frame deletions of *yscD* fail to secrete effector proteins (Michiels, 1991, Plano, 1995). Besides its localization in the inner membrane, very little is known about the role of YscD in the T3SS. Different functions can be hypothesized from studies on homologs of YscD.

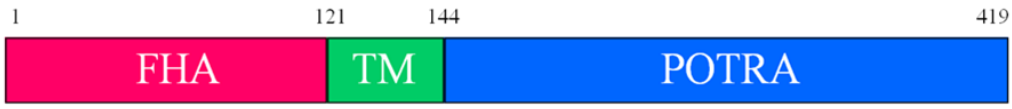

Figure 1.2: Schematic of YscD. The transmembrane region (TM) was predicted using TMHMM (Krogh *et al.*, 2001). The N-terminal and C-terminal domains are predicted by secondary structure to adopt folds similar to FHA and POTRA domains (BLAST).

The *E. coli* homolog of YscD, EscD (also known as PAS, protein associated with secretion), is required for both secretion of effector proteins (Esp proteins) and the formation of the T3SS apparatus in the *E. coli* T3SS. EscD interacts with the outer membrane protein EscC (homolog of YscC) and the needle protein EscF (homolog of the needle protein YscF) (Kresse, 1998, Ogino, 2006). EscD interacts additionally with the effector proteins EspE/Tir, and SepL (Kresse, 1998). Tir is injected by the T3SS into epithelial cells and inserts into the eukaryotic cell membrane, from which it serves as a receptor for the bacterial outer membrane protein intimin. SepL (homolog of YopN) binds to Tir and functions to block the export of effector proteins while the needle is being assembled (Wang, 2008). There is no homolog of Tir in *Yersinia*, however YscD and EscD may function to recognize a subset of secreted proteins in their respective T3SSs.

The YscD homolog from *S. enterica,* PrgH, is the most studied of the T3SS inner membrane proteins. PrgH multimerizes into a tetramer and differs from YscD in that it can oligomerize with PrgK (homolog of YscJ) in the absence of the outer membrane protein InvG (homolog of YscC) (Kimbrough, 2000). In contrast, as noted above, assembly of YscD and YscJ require the outer membrane protein YscC. The N-terminal region of PrgH localizes to the cytoplasm, while the C-terminus localizes to the periplasm and interacts with the periplasmic domain of InvG (Schraidt, 2010, Sanowar, 2010). The X-ray crystallographic structure of residues 170-362 of PrgH, which corresponds to the periplasmic region, was determined (Spreter, 2009). This region of PrgH contains a conserved structural motif found in the soluble portions of many membrane proteins, the polypeptide transport-associated (POTRA) domain.

The T3S periplasmic proteins from *E. coli*, EscJ and EscC, also contain POTRA domains in their tertiary structure (Yip, 2005, Spreter, 2009). The POTRA domain is comprised of a three-stranded β-sheet overlaid with a pair of α-helices. This domain has been characterized in other membrane proteins involved in protein export and assembly of macromolecular structures, including Sec-pathway proteins.

MxiG from *S. flexneri* is an inner member protein with similar topology to YscD and a homolog of YscD. Recently the structure of MxiG was determined by nuclear magnetic resonance to reveal a FHA domain (McDowell *et al.*, 2011). MxiG interacts with the inner membrane protein MxiJ (homolog YscJ) to form a ring complex in the inner membrane composed of an estimated 24 units from each protein (Hodgkinson, 2009, McDowell et al., 2011). MxiG also binds the outer membrane protein MxiD (homolog of YscC) and the cytosolic protein Spa33 (homolog of YscQ) (Morita-Ishihara, 2006, Zenk, 2007). In *Yersinia*, YscQ is part of an assembly that forms at the base of the inner membrane of the needle complex. This assembly includes the ATPase YscN, the ATPase negative regulator YscL, and YscK, a protein of unknown function (Diepold, 2010). Another YscD homolog, CdsD from *Chlamydophila pneumonia*, has been found to associate with the homologous cytosolic assembly. CdsD interacts with the T3SS proteins CdsN (homolog of YscN), CdsQ (homolog of YscQ) and CdsL (homolog of YscL) (Johnson, 2008, Stone, 2008). These data suggest that a multiprotein complex composed of the cytosolic proteins YscN, YscL, YscK, and YscQ may be a binding target for the cytoplasmic domain of YscD.

The cytoplasmic region of *C. pneumonia* CdsD contains two FHA domains. Both CdsD FHA domains can be phosphorylated *in vitro* on serine and tyrosine residues by the membrane-associated kinase PknD (Johnson, 2007). Inhibition of this phosphorylation by PknD affects *C. pneumonia* replication (Johnson, 2009). YscD is predicted to have one FHA domain of undetermined function in the cytosolic region of *Yersinia*. Phosphorylation on the FHA domain of YscD has not been reported and there is no homolog of PknD in *Yersinia*.

FHA domains were initially characterized in forkhead transcription factors and are now known to be in thousands of different proteins from mammals to bacteria (Pfam database) (Hofmann, 1995). FHA domains are small (less than 100 amino acids) and characterized by two β-sheets that form a β-sandwich. FHA domains are diverse in sequence, with only seven residues conserved among most FHA domains (Li, 2000). Structural diversity is found in the loops between the β-strands, and in some FHA domains these loops have additional secondary structure, such as α-helices. Diversity in the sequence of FHA domains is congruent with the various functions of FHA domain-containing proteins. For example, FHA domains are found in kinases, kinesins, phosphatases, ABC-transporters, and transcription regulator proteins among others. A function of many FHA domain-containing proteins involves cell signaling via interactions with proteins containing phosphothreonine residues. However, this function is not absolutely conserved among FHA domain proteins, and even among phosphothreonine-binding FHA domains, there is diversity in the phosphothreonine-binding regions. Bacterial proteins containing FHA domains are also quite functionally diverse, being involved in protein transport (type IV secretion system in

Psuedomonas aeruginosa) (Mougous, 2007), TCA cell cycle control in *Corynebacterium glutamicum* (the FHA domain of Odh1 is phosphorylated to reduce TCA cycle inhibition) (Bott, 2007), the Esx secretion system (EssC in *Staphylococcus aureus*) (Tanaka, 2007), and the infectious capability of *C. pneumoniae* (dependent on the phosphorylation of FHA-2 from CdsD by PknD) (Johnson, 2009).

The goal of the research presented in this dissertation is to understand the role of the YscD cytosolic domain, YscDc. The periplasmic domain of YscD is involved in structural formation of the T3SS injectisome, where it oligomerizes with other structural components. We identified the cytoplasmic domain to be crucial for secretion of virulence proteins. The X-ray crystallographic structure of YscDc was determined to adopt a FHA domain fold. We defined two loops in this domain which are involved in secretion. Based on the analysis of YscDc homologs, we propose that this FHA domain is a conserved feature of the T3SS.

Materials and Methods

Cloning of YscD expression constructs

The coding sequence of *yscD* was amplified by PCR from the pYV plasmid of *Y. pseudotuberculosis* 126 (Bolin, 1982). PCR products encoding YscD, YscDc (residues 1-121), or YscDp (residues 138-419) were cloned into the pET28b expression vector (Novagen) to yield pET28b-yscD plasmids. These constructs included a His-tag followed by a PreScission protease cleavage site at the N-terminus of YscD. YscD was also cloned into the pBAD expression vector (Invitrogen) to yield pBAD-yscD; this construct also included an N-terminal His-tag. Mutants of YscD were constructed by strand overlap extension PCR (Higuchi, 1988). The integrity of all constructs was verified by DNA sequencing.

Expression and purification of YscDc and YscDp

YscDc and YscDp were expressed from pET28b in *Escherichia coli* BL21 (DE3). Bacteria were grown at 37 ºC in LB media supplemented with 50 mg/L kanamycin to an OD_{600} of 0.5, at which point expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The bacteria were then grown for 16 h at 20 ºC, after which point they were harvested by centrifugation (5,800 x *g*, 10 min, 4º C). The bacterial pellet was resuspended in 1/100th volume of the original bacterial culture of buffer A (500 mM NaCl, 50 mM sodium phosphate buffer, pH 8.0, 10 mM β-mercaptoethanol (βME)) supplemented with EDTA free protease cocktail inhibitor

(Roche) (one tablet per 2 L of bacterial culture). Resuspended bacteria were lysed using an Emulsiflex-C5 (Avestin) homogenizer with three passes at 15,000 psi, and the lysate was clarified by centrifugation (14,000 x *g*, 10 min, 4 ºC). The supernatant was applied to a Ni^{2+}-nitrilotriacetic acid (NTA) agarose column (Sigma), the column was washed with 25 column volumes of buffer A containing 7 mM imidazole, and bound protein eluted from the column with three column volumes of buffer A containing 500 mM imidazole. The eluted fractions were concentrated by ultrafiltration using a YM-3 Centricon (Amicon), and further purified by size-exclusion chromatography (16/60 Superdex 200, GE Healthcare) in buffer B (50 mM NaCl, 10 mM HEPES, pH 7.5, and 10 mM βME). Purified YscDc and YscDp were cleaved at a 50:1 YscD:PreScission protease mass ratio to remove the His-tag, and the cleaved samples were applied to a Ni^{2+}-NTA agarose column in buffer B. Cleaved YscDc was isolated from the flow-through of the column, and concentrated by ultrafiltration using a YM-3 Centricon to 7.5 mg/mL; the concentration of YscDc was determined using the calculated molar ε_{280} of 12,490 M^{-1} (Gasteiger, 2003). Cleaved YscDp was isolated from the flow-through and subjected to molecular mass analysis by MALDI-TOF.

Selenomethionine was incorporated into YscDc (Semet-YscDc) as described previously (Van Duyne, 1993), and SeMet-labeled YscDc was purified as above.

Expression and Purification of YscD

YscD was expressed, harvested, and lysed in the same manner as YscDc and YscDp from *E. coli*. After lysis by homogenization, bacterial membranes were pelleted by ultracentrifugation (95,000 x *g*, 4 h, 4°C). Membranes were solubilized in 1/100th volume of the original bacterial culture in buffer A and 20 mM lauryldimethylamine oxide (LDAO). Membranes were applied to a Ni^{2+}-NTA column and the column was washed with 25 column volumes of buffer A containing 7 mM imidazole and 5 mM LDAO. Bound protein was eluted from the column with three column volumes of buffer A containing 500 mM imidazole and 5 mM LDAO. The eluted fractions were concentrated by ultrafiltration using a YM-10 Centricon and further purified by size-exclusion chromatography (10/30 Superose 6, GE Healthcare) in buffer C (50 mM NaCl, 10 mM NaPi, pH 8.0, 5 mM LDAO, and 10 mM βME).

Crystallization and Data collection

Crystals of unlabeled and SeMet-labeled YscDc were grown by the sitting-drop, vapor diffusion method at 20 ºC by mixing 40% 7.5 mg/mL YscDc in buffer B and 60% 3.5 M NaHCOO in a drop size of 0.5 μL. The drop was dispensed by an Oryx8 crystallization robot into a CrystalClearDuo crystallization plate (Hampton), which contained 80 μl of 3.5 M NaHCOO in the well. Crystals were cryoprotected in 15% glycerol and 3.5 M NaHCOO, mounted in 50 μm loops (Hampton), and flash cooled in liquid N_2.

Structure Determination

Diffraction data were collected from native YscDc crystals and single wavelength anomalous dispersion (SAD) diffraction data were collected from SeMet-labeled YscDc crystals at beamline 23 I-B (Advanced Photon Source, Argonne National Laboratory).

Data from native and Se-Met crystals were indexed and scaled using HKL2000 (Otwinowski, 1997). Two molecules of YscDc were located in the asymmetric unit. Four selenomethionine sites (Met1, Met50 for each chain) were located using the hybrid substructure search (HYSS) from the program Autosol in the Phenix suite (Adams, 2010). Using these selenomethionine sites, phases were calculated and refined using Phenix. Automated model building was done using Phenix Autobuild to produce a trace of residues 1-105. Refinement was done against the native data set (Table 1.1). σ_A-weighted 2mFo-DFc and mFo-DFc maps were visualized by crystallographic object-oriented toolkit (COOT) and used to improve the model before subsequent rounds of refinement (Emsley, 2004). Each refinement cycle consisted of three rounds of bulk-solvent correction, anisotropic scaling, and refinement of atomic model parameters. Each refinement cycle used default parameters with tight NCS restraints set at 0.2. Waters were added in the later stages of refinement using PhenixRefine and manually into ≥3σ Fo-Fc density. The electron density for the main chain was unbroken and modeled from residues 1-108. The electron density of all side chains was visible, except for Glu54, in both chains A and B. Structure validation was done using MolProbity (Table 1.4) (Chen, 2010).

Generation of *Y. pseudotuberculosis* (*ΔyscD*)

To generate *Y. pseudotuberculosis* (*ΔyscD*), the entirety of *yscD*, except for 15 bp at the 3' end, was substituted in-frame with *aph* (kanamycin resistance) by homologous recombination with a PCR fragment, as described previously (Datsenko KA, 2000). The 15 bp at the 3' end were left intact as they contain a predicted ribosome-binding site for *yscE*. The PCR fragment contained 500 bp of the pYV sequence upstream of *yscD*, followed by *aph*, the terminal 15 bp of *yscD*, and 500 bp of the pYV sequence downstream of *yscD*. Briefly, 600 ng of this fragment were transformed into competent *Y. pseudotuberculosis* harboring the plasmid pWL204 by electroporation (Conchas, 1990, Lathem *et al.*, 2007). Bacteria were grown in brain heart infusion (BHI) medium at 26 ºC for 2 h, centrifuged (3,000 x *g*, 5 min, 25 °C), and the pellet was resuspended in BHI supplemented with 3 mM $CaCl_2$ and 50 mg/L kanamycin. Bacteria were grown further for 2 h at 26 °C in these conditions, and were transferred to BHI agar plates containing 3 mM $CaCl_2$ and 50 mg/L kanamycin. Plates were grown overnight at 26 °C. The integrity of the allelic replacement was verified by sequencing a PCR fragment resulting from primers that anneal 650 bp upstream and downstream of the *yscD* locus. To select for loss of the pWL204 plasmid, *ΔyscD* was grown on agar plates as above but supplemented with 2% sucrose. Loss of pWL204 was confirmed by sensitivity to ampicillin. *Y. pseudotuberculosis* (*ΔyscD*) was complemented with wild-type or mutant *yscD* expressed from the pBAD plasmid, as described above.

Secretion Assay

Y. psuedotuberculosis was grown in BHI media overnight at 26 ºC. The overnight culture was diluted to an OD_{600} of 0.1 in 4 mL of BHI containing 10 mM ethylene glycol tetraacetic acid (EGTA), 10 mM $MgCl_2$, and appropriate antibiotics. If *Y. pseudotuberculosis* was harboring pBAD-yscD, then cultures were supplemented with 0.002% arabinose once they reached an OD_{600} of 0.6; 0.002% arabinose was subsequently added every two hours during growth to account for the metabolic depletion of arabinose. Cultures were shifted to 37 ºC at an OD_{600} of 0.6 and grown an additional 4 h. For western blot analysis of expressed proteins, 4 mL of bacterial culture at an OD600 of 1.0 were harvested by centrifugation (3,800 x *g*, 5 min, 4 °C). Pelleted bacteria was resuspended in 40 µL of SDS-PAGE sample buffer and boiled for 5 min. Whole cell lysate proteins were then separated by SDS-PAGE, transferred to PVDF membranes, and subject to western blot analysis using antibodies against YscD. To visualize secreted proteins, the bacterial supernatant was precipitated overnight by the addition of 10% trichloroacetic acid (TCA). Precipitated samples were centrifuged (5,000 x *g*, 5 min, 4 °C) and the pellet washed with ice-cold acetone. Pelleted proteins were air-dried, resuspended in SDS-PAGE sample buffer, boiled for 5 min, and separated by SDS-PAGE.

Table 1.1: Data collection for YscDc and SeMet YscDc.

	YscDc	SeMet YscDc
Space Group	P432	P432
Cell dimensions		
a, b, c (Å)	117.07, 117.07, 117.07	116.78, 116.78, 116.78
α,β,γ (°)	90, 90, 90	90, 90, 90
Wavelength (Å)	0.97961	0.97949
Resolution	50.0-2.50	41.3-2.85
Highest shell	(2.59-2.50)	(2.90-2.85)
R_{merge}	57.8 (7.6)	37.0 (7.6)
I/σI	46.8 (8.4)	39.9 (11.0)
Completeness	98.1 (80.6)	99.3 (99.8)

Table 1.2: Refinement statistics for SAD YscDc structure.

Resolution (Å)	35.3-2.52
No. unique reflections	17,603
R_{work}	0.1963
R_{free}	0.2553
R.m.s. deviations	
Bond length (Å)	0.008
Bond angles (°)	1.245

Table 1.3: YscDc structure validation using Molprobity.

No. of steric overlaps (>0.4 Å) per 1000 atoms	14.45 (88th percentile)
Poor rotamers	0.00 %
Ramachandran outliers	0.95 %
Ramachandran favored	97.14 %
Residues with bad bonds	0.00 %
Residues with bad angles	0.93 %
MolProbity score	1.82 (98th percentile)

Results

Purification of YscD, YscDc, and YscDp

YscD is predicted to contain a single-pass transmembrane α-helix between residues 121 and 144, and two soluble domains, each located on opposite sides of the membrane. Full length histidine-tagged YscD (48.7 kDa) was expressed and purified from *E. coli* membranes by Ni^{2+}-NTA affinity and size exclusion chromatography. YscD was not homogeneous and had a broad gel filtration profile on a Superose 6 size exclusion column (Figure 1.3). Three predominant species of YscD were identified by SDS-PAGE. The sizes of these products correspond to the molecular masses of full-length YscD, the periplasmic domain of YscD, and the cytoplasmic domain of YscD. To facilitate crystallization of YscD, the cytoplasmic and periplasmic domains of YscD were cloned, expressed, and purified separately.

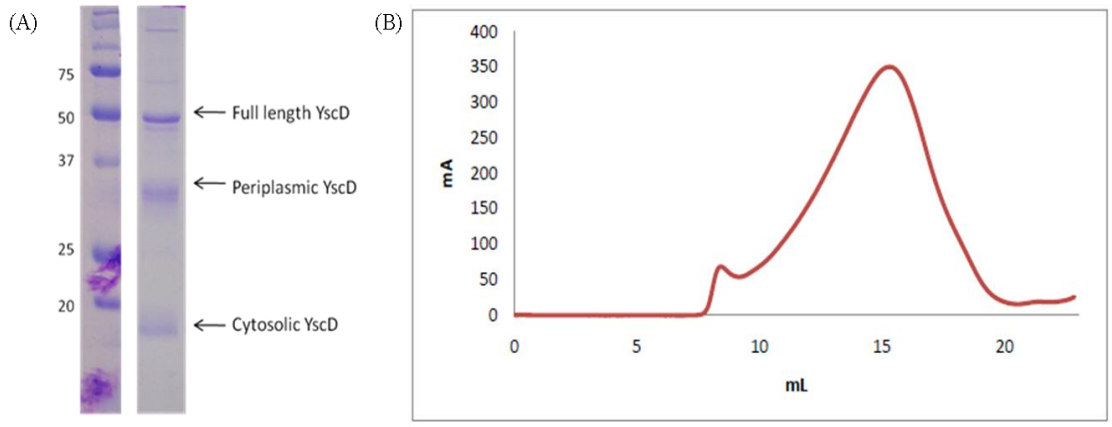

Figure 1.3: Purification of YscD. (A) SDS-PAGE gel of purified YscD after size exclusion chromatography. The presence of several proteins of different molecular mass are shown with arrows. Full length YscD is 48.7 kDa, the periplasmic domain is 31.2 kDa, and the cytoplasmic domain has a mass of 13.4). (B) Chromatogram of YscD on a 10/30 Superose 6 column. The column volume for this column is 23 mL, with a void volume at 8 mLs. The profile for YscD is extended, corresponding to a heterogenous population with molecular masses spanning between 5 kDa and 0.5 MDa.

The cytoplasmic and periplasmic domains of YscD, YscDc (13.4 kDa) and YscDp (31.2 kDa), were purified by Ni^{2+}-NTA affinity and size exclusion chromatrography from *E. coli*. The size exclusion profile of each protein was consistent with a monomeric state (Figures 1.4 and 1.5). The histidine-tag was cleaved from YscDc and YscDp using histidine-tagged PreScission protease. The cleaved proteins were reapplied to the Ni^{2+}-NTA column and unbound sample was collected. YscDc and YscDp were further purified by size exclusion chromatography.

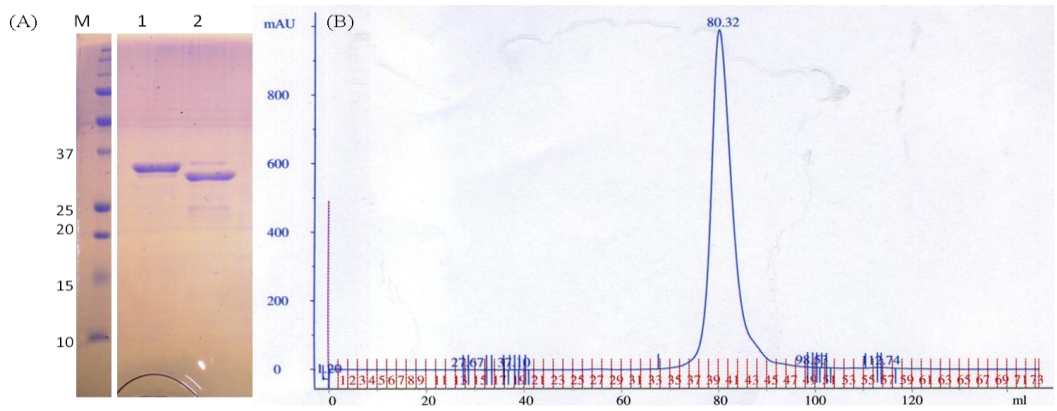

Figure 1.4: Purification of YscDp. (A) SDS-PAGE gel of YscDp. Lane M corresponds to the molecular weight marker, lane 1 corresponds to histidine tagged YscDp after Ni^{2+}-NTA purification, and lane 2 corresponds to YscDp after histidine tag removal. (B) Size exclusion chromatogram of cleaved YscDp corresponding to a monomer.

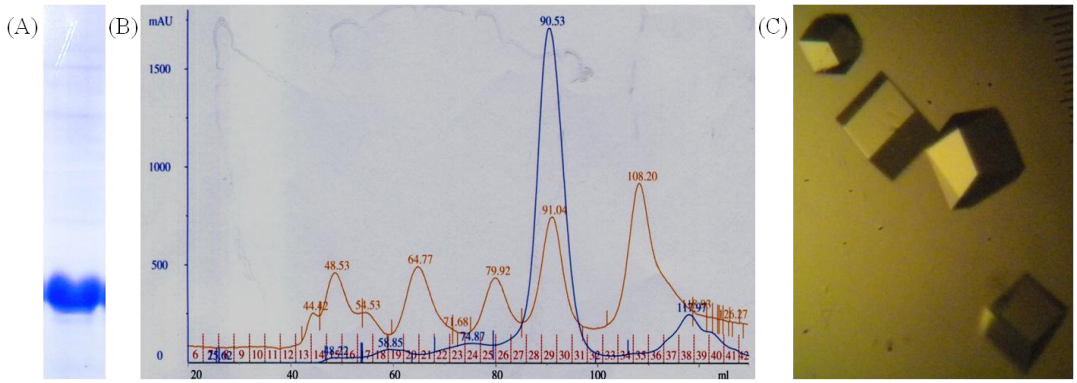

Figure 1.5: Purification of YscDc. (A) SDS-PAGE gel showing YscDc after purification and removal of the histidine tag. (B) Chromatogram of YscDc on a Superdex 200 column. The blue trace is YscDc, the red trace corresponds to globular molecular mass protein standards. (C) Crystals of YscDc formed in NaHCOO.

The periplasmic domain of YscD was susceptible to degradation after purification, with a lower molecular mass product apparent at 23 kDa, as determined by MALDI-TOF (Figure 1.6). An anti-histidine-tag western blot confirmed that this

fragment still contained the N-terminal histidine-tag. A shorter construct of YscDp was cloned containing residues 144-349 (YscDp is 144-419), called YscDp205. YscDp205 was purified in similar manner to YscDp. Two peaks of YscDp205 were obtained, one with a size corresponding to a molecular mass greater than the protein standard at 660 kDa and another corresponding to monomeric YscDp205 (Figure 1.7). Peak fractions corresponding to the monomer were pooled and concentrated to 5 mg/mL and subjected to vapor-diffusion crystallization trials. No protein crystals were observed.

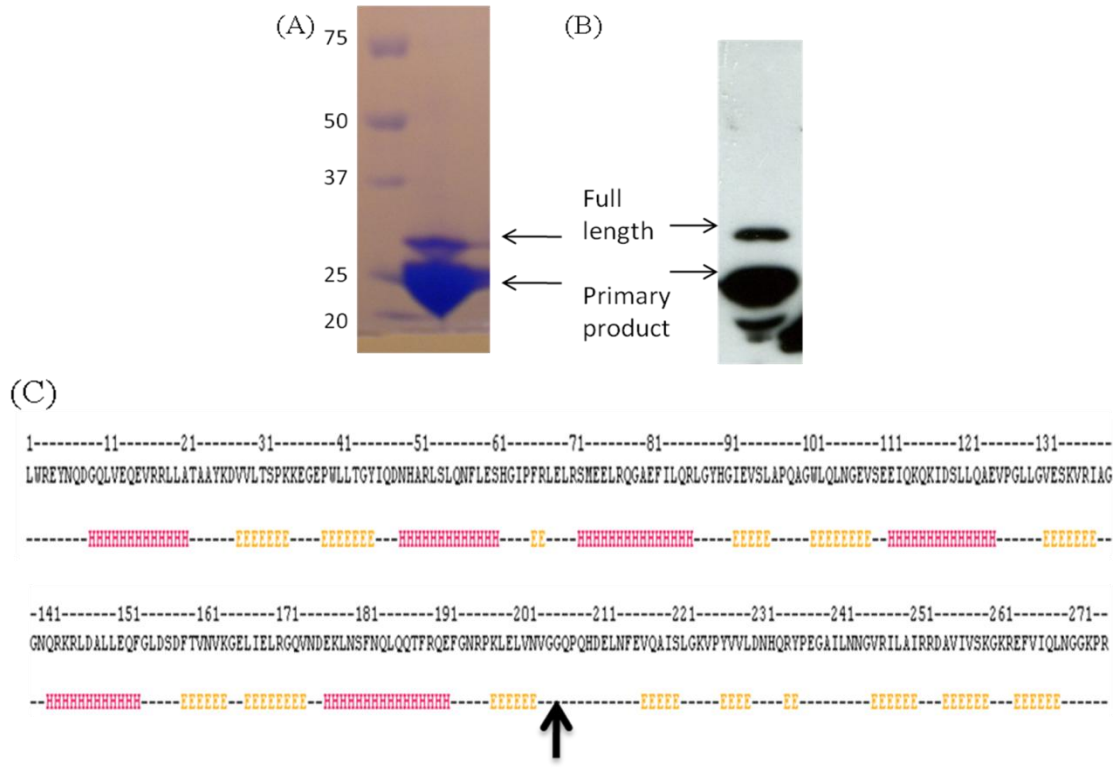

Figure 1.6: YscDp forms a lower molecular mass product. (A) SDS-PAGE gel showing two forms of YscDp at different molecular masses. (B) An anti-histidine-tag western blot showing that both species of YscDp still have the N-terminal histidine tag present. (C) Secondary structure prediction of YscDp by Jpred3 (Cole *et al.*, 2008). The black arrow indicates where cleavage could occur based on the size of the lower molecular mass product (MALDI-TOF).

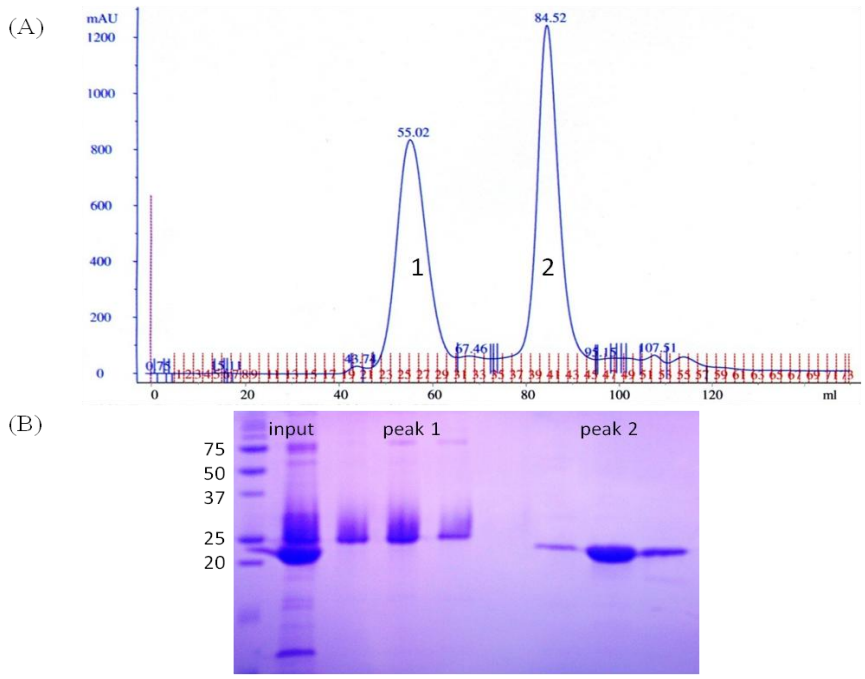

Figure 1.7: Purification of YscDp205. (A) Size exclusion chromatography profile of YscDp205. (B) SDS-PAGE gel showing the YscDp205 input onto the chromatography column and the proteins corresponding to fractions from each peak. Peak 1 has a molecular mass higher than the largest protein standard (>660 kDa). Peak 2 is consistent with monomeric YscDp205 based on the molecular masses of protein standards.

Purified YscDc was concentrated to 7.5 mg/mL and crystallized with NaHCOO as the precipitant by the vapor diffusion method (Figure 1.5). SelenomethionineYscDc was purified and crystallized in the same conditions as native YscDc. The X-ray crystallographic structure of YscDc was determined to 2.5 Å resolution limit. Two copies of YscDc were found in the asymmetric unit, with a R.M.S.D. of 0.347.

Structure of YscDc

YscDc has a typical FHA domain fold. It has a β-sandwich composed of 10 β-strands that form two sheets, as found in other FHA domains (Figure 1.8). The two sheets, one six-stranded (β2β1β10β9β9β8) and the other four-stranded (β4β3β5β6), are connected to each other by a long loop that has interspersed within it the very short β7, β8, and β9 strands. The β4 strand is also very short and runs parallel to β3, whereas all the other stands are antiparallel. Between strands β1 and β2 is a short 3_{10}-helix. The two β-sheets interact through hydrophobic side chains, which is a conserved feature of the FHA domains of other proteins. The N-terminus is well ordered, whereas the C-terminal residues 109-121, which connect to the transmembrane region of YscD, appear to be unstructured. This raises the possibility that the linker between the cytoplasmic domain and the transmembrane region is flexible, enabling freedom in interactions between the YscD cytoplasmic domain and membrane-associated or cytoplasmic components of the T3SS apparatus.

Figure 1.8: X-ray crystallographic structure of YscDc. Coloring denotes secondary structure with beta strands in purple and loops in pink. Figure is made with PyMOL (Schrödinger).

FHA domains are found in a variety of proteins across multiple kingdoms, including *Mycobacteria tuberculosis* Embr, *Saccharomyces cerevisiae* Rad53, and *Homo sapiens* Chfr. The most structurally similar proteins to YscDc are *Chlamydia trachomatis* CT664 (RMSD of 1.974, 93 C-alphas), *M. tuberculosis* Embr, (RMSD of

2.054 , 92 C-alphas) *M. tuberculosis* TB39.8 (RMSD of 2.175, 93 C-alphas), and *H. sapiens* kinesin KIF13 (RMSD of 2.096 , 93 C-alphas) (Figure 1.9). However, YscDc has little sequence similarity (~5 % identity) to these or other structurally characterized FHA domain proteins (Figures 1.10). Most notably, the identity of the side chains that make up the hydrophobic core differ between YscDc and these other FHA domain proteins, suggesting convergent evolution of the FHA fold in YscDc (Figure 1.11). There are two residues that are found in these FHA domains and in other confirmed FHA domains: a glycine at the beginning of loop β3β4 (Gly27 in YscDc) and a histidine at the end of loop β4β5 (His46 in YscDc). His46 engages in electrostatic interactions with backbone atoms of Ala43 in loop β4β5 and Val86 in strand β9. The amino acids located between these conserved Gly and His residues are known to be crucial for function in FHA domains.

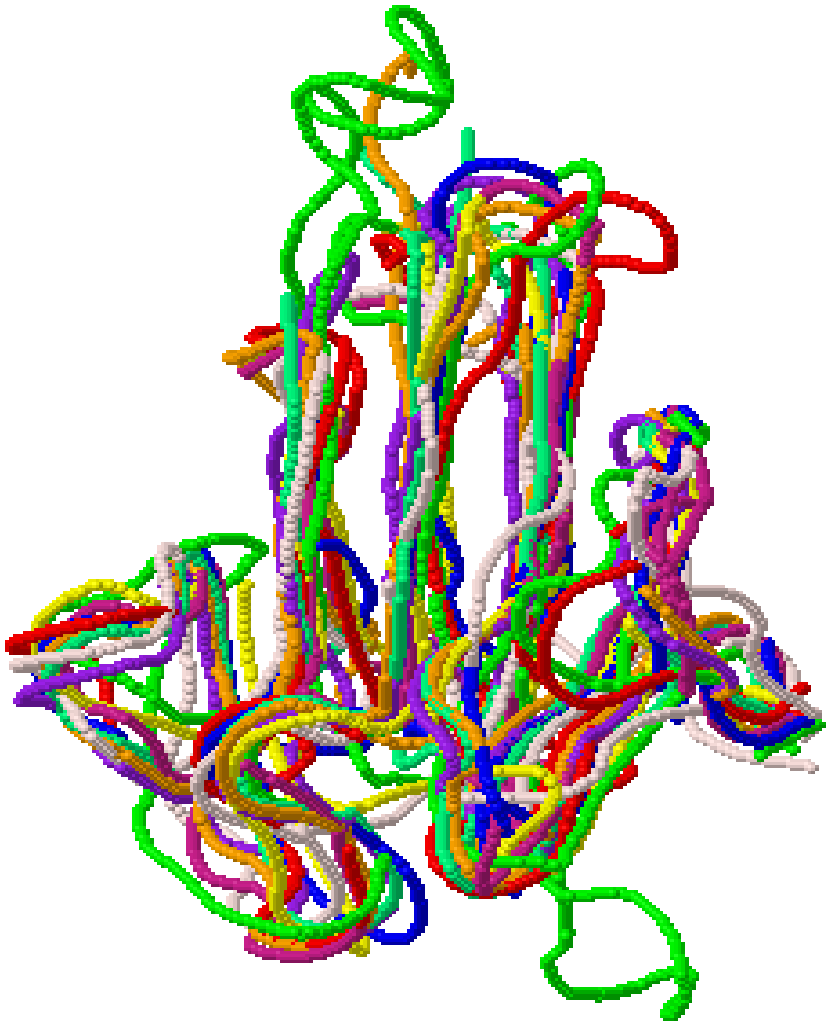

Figure 1.9: Overlay of FHA domains. YscDc (light pink), Ct664 (1.974, dark purple), NIPP1 (2.468, red), KIF13 (2.096, blue), EmbR (2.054, orange), OdhI (2.228, yellow), TB398 (2.175, cyan), Ki67 (2.310, light purple), and Rad53 (2.733, green) were overlaid using Flexible structure AlignmenT by Chaining Aligned fragment pairs allowing Twists (FATCAT) (Ye & Godzik, 2004). The RMSD between 93 C-alphas of YscDc and each FHA domain is shown in parenthesis.

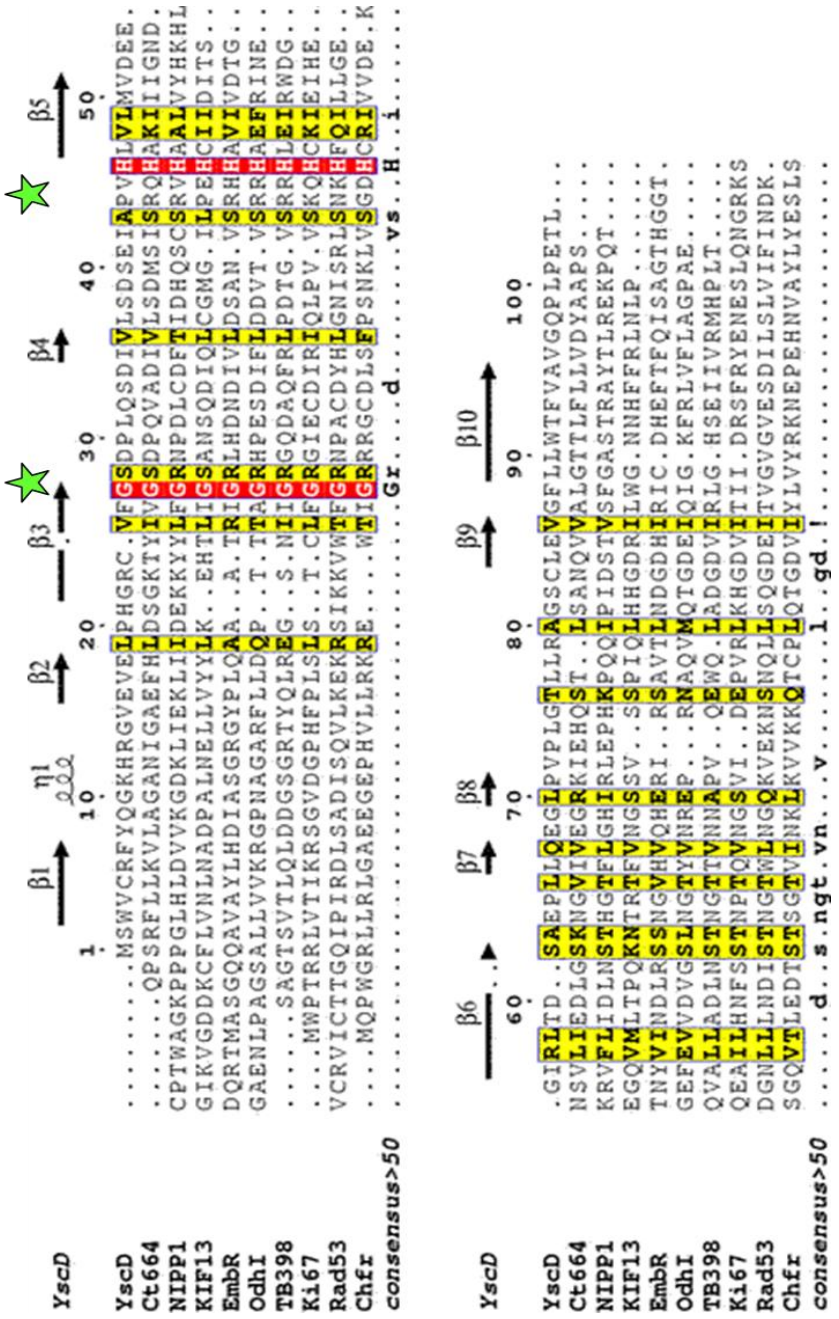

Figure 1.10: Alignment of FHA-domain containing proteins of known structure with YscDc. The PDB codes for each protein: Ct664 (3GQS), NIPP1 (2JPE), KIF13 (3FM8), EmbR (2FEZ), OdhI (2KB3), TB398 (3PO8), Ki67 (2AFF), Rad53 (1G3G), and Chfr (1LGQ). Sequence alignment was done using Multialin (Corpet, 1988), and theYscDc secondary structure denotation above the alignment was made with ESPript (Gouet *et al.*, 2003). Typical FHA domain residues are denoted with a star: GR and SxxH.

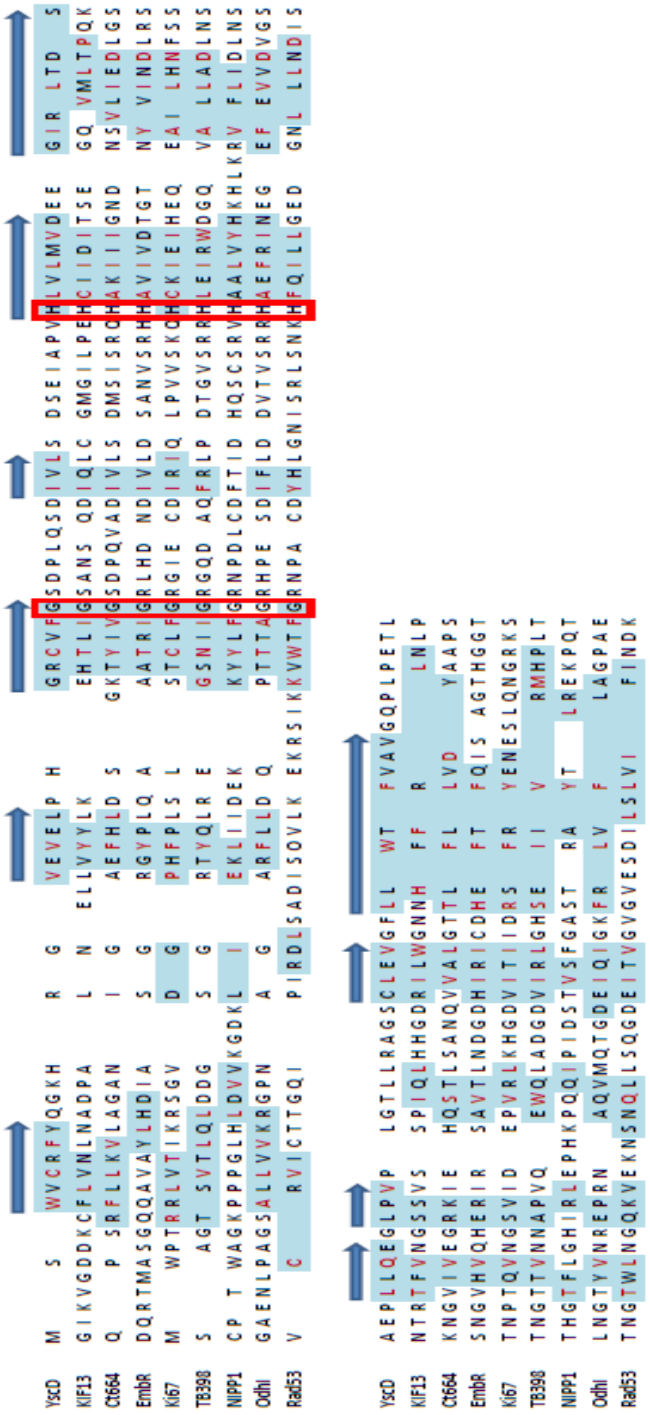

Figure 1.11: Alignment of FHA domains illustrating hydrophobic core residues. Arrows above the sequence alignment denote the secondary structure of YscDc. The residues highlighted in blue indicate β-strands in the protein secondary structure. Residues colored in red comprise the hydrophic core of each FHA domain. Alignment was done using Multialin (Corpet, 1988). Secondary structure assignment was done using DSSP and colored manually.

Most but not all FHA domains bind phosphothreonine residues. The interaction with phosphothreonine occurs primarily through a serine located on the β4β5 loop and secondarily through an arginine found in either loops β3β4 or β4β5. There is little overall sequence conservation in these loops among FHA domains, reflecting the differing binding specificities of FHA domains. In proteins which bind phosphothreonine residues, the surface of loops β3β4 and β4β5 is positively charged. However, the equivalent surface in YscDc is negatively charged (Figure 1.12). Furthermore, YscDc lacks the conserved Ser on the β4β5 loop (two residues prior to the conserved His47 in YscD) and the Arg on the β3β4 or β4β5 loops. These pieces of evidence indicate that YscDc is unlikely to bind proteins in a phosphothreonine-dependent manner. However, this same region may be involved in phosphorylation-independent protein-protein interactions, as has been observed for the FHA domain protein *H. sapiens* KIF13. The KIF13 FHA domain lacks the conserved Ser and Arg, similar to YscDc, and has an uncharged electrostatic surface in the β3β4 and β4β5 loop regions (Nott, 2009).

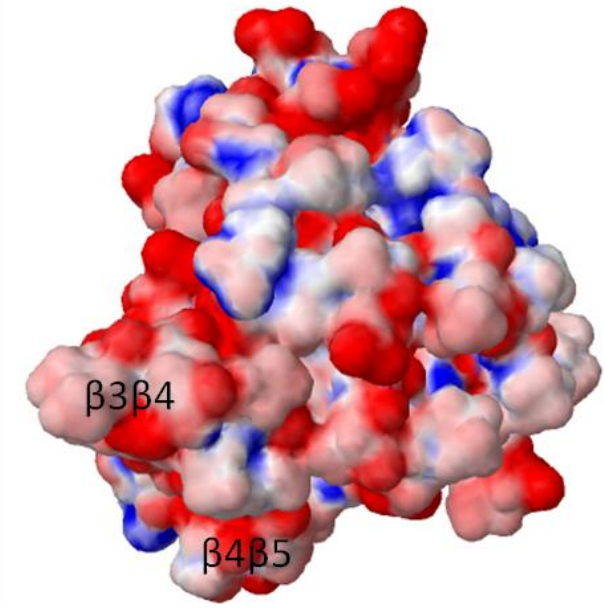

Figure 1.12: Electrostatic potential of YscDc. Red is negative (-10 kT) and blue is positive (+10 kT). The region surrounding loops 3 and 4 have a negative surface charge. This figure was made using Adaptive Poisson-Boltzmann Solver (APBS) from PyMOL (Baker *et al.*, 2001).

The FHA domain of YscD has significant sequence similarity to orthologs in the Ysc family of T3S systems. These include *Aeromonas hydrophila* AscD, *Pseudomonas aeruginosa* PscD, and *Photorhabdus luminescens* SctD, with which YscDc has an average of 29% sequence identity. The FHA domain of YscD has more distant but recognizable sequence similarity to the cytoplasmic domain of orthologs in the SPI-1 family (*Salmonella enterica* PrgH, 18% identity), SPI-2 family (*Escherichia coli* EscD, 24% identity), Hrp2 family (*Xanthomonas campestris* HrpD5, 24% identity), Hrp1 family (Erwinia amylovora HrpQ, 24% identity), and the *Chlamydiales* family (*Chlamydia trachomatis* CT664, 21% identity) (Figure 1.13). The cytoplasmic domains of these orthologs are predicted to have an FHA fold, as are the more closely related Ysc family proteins. The hydrophobic core in the Ysc proteins is well

conserved (e.g., Leu49, Val51, Ile56, Leu58), suggesting divergent evolution of the T3S orthologs (Figure 1.14). As with YscDc, the putative β3β4 and β4β5 loops in the T3S orthologs are missing the conserved Ser and Arg associated with phosphothreonine binding. For the orthologs in the Ysc family, the small β4 strand and the surrounding β3β4 and β4β5 loops are the most strongly conserved, consistent with a functional role for this region. This same region is divergent in PrgH and EscD, and indeed these proteins are lacking the conserved His in the β4β5 loop. This suggests that possible YscD interactions are conserved within the Ysc family but likely to differ from that of PrgH and EscD.

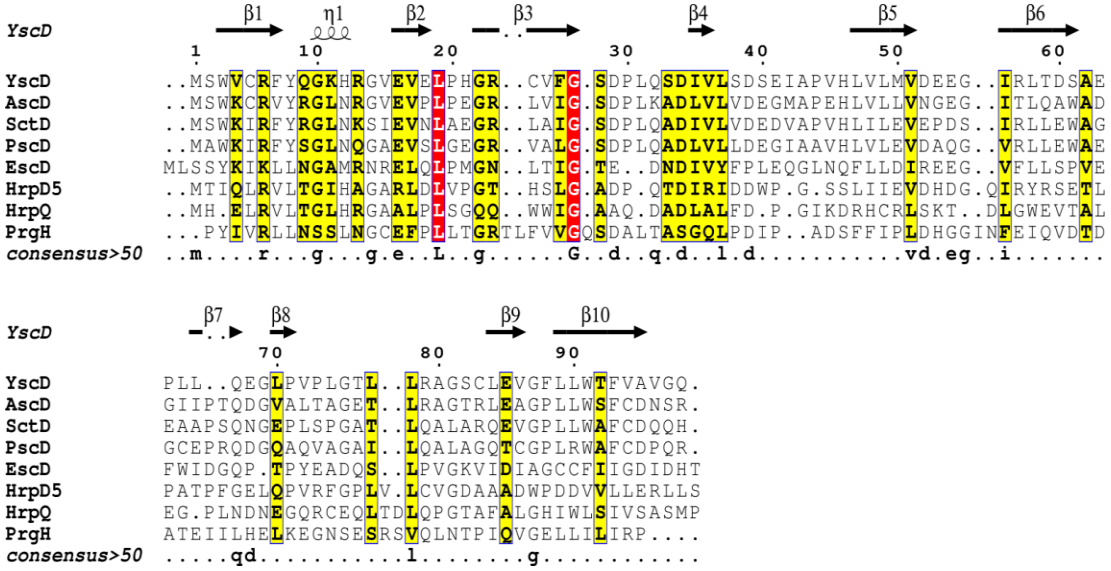

Figure 1.13: Alignment of YscDc with T3SS homologs. Conserved residues are highlighted in red. Consesrvatively substituted residues with common chemical properties are highlighted in yellow. Sequences were aligned using Multialin (Corpet, 1988). The secondary structure of YscDc above the alignment was done using EsPript (Gouet et al., 2003).

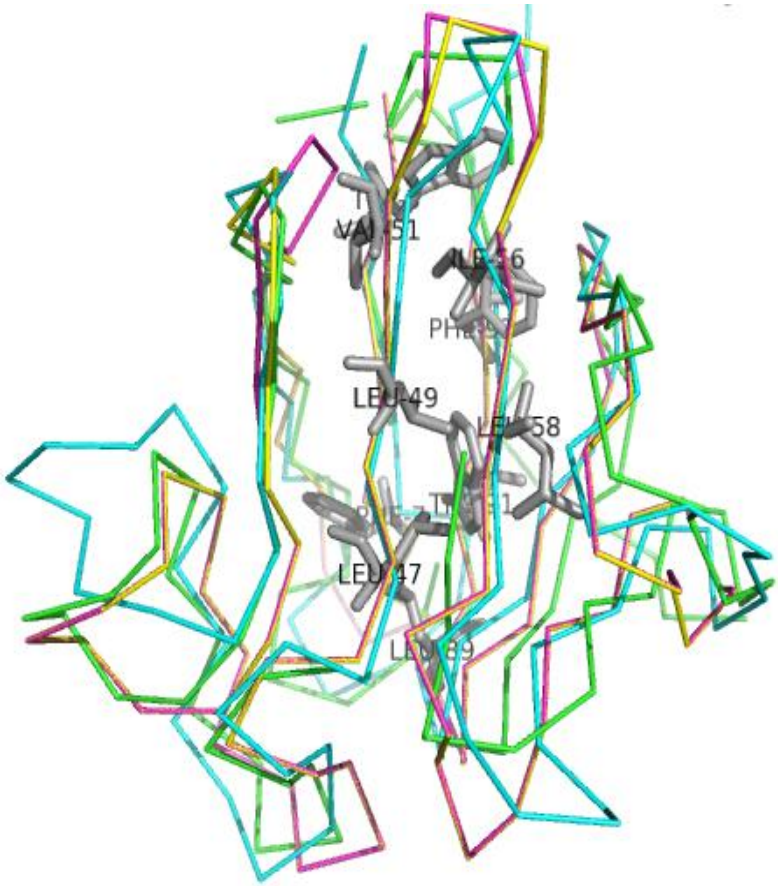

Figure 1.14: Overlay of the cytoplasmic domains from YscD familiy members AscD, PscD, and SctD with YscDc. The structures of the cytoplasmic domains of AscD, PscD, and SctD were modeled using PhyreBeta (Kelley, 2009). Conserved hydrophobic residues located in strands 1, 5, 6, and 10 are shown in grey sticks.

Recently the FHA domain from the *Shigella flexneri* T3SS protein, MxiG, was determined by nuclear magnetic resonance, NMR (McDowell et al., 2011). MxiG is in the Inv-Mxi-Spa T3SS family along with *S. enterica* PrgH. MxiG is an ortholog to YscD, with a sequence identity of 11 %, which is significantly lower when compared to the other YscD orthologs described above (Figure 1.15). The most sequence divergent amino acids are located in loops β3β4 and β4β5. The structural overlay

between these FHA domains (RMSD of 2.309) reveals that these loops are similarly structurally divergent (Figure 1.16). As with YscD, MxiG loops β3β4 and β4β5 are candidates for T3SS function. Residues R39 (loop β3β4), S61 (loop β4β5), and S63 (loop β4β5) were subjected to mutagenesis and strains carrying these mutations were analyzed for their ability to secrete effector proteins. These amino acids were not found to be essential for secretion.

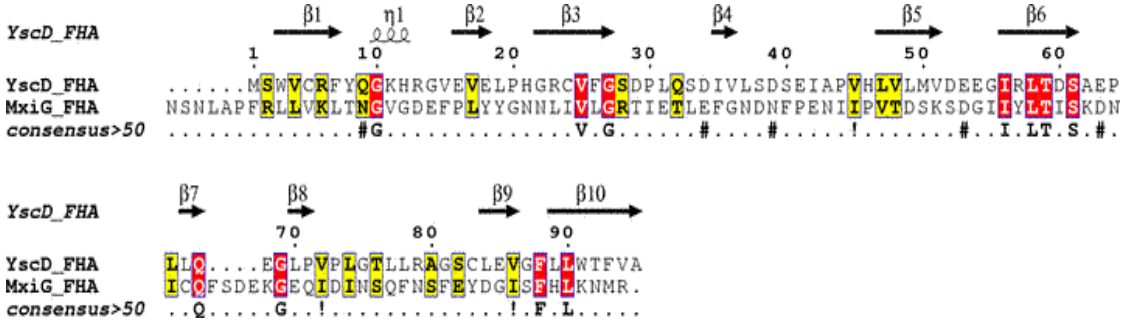

Figure 1.15: Sequence alignment between the FHA domains of YscD and MxiG. The sequence alignment was done using Multialin (Corpet, 1988). The secondary structure of YscDc was drawn above the alignment using ESPript (Gouet et al., 2003). Residues identical between proteins are highlighted in red. Residues highly similar in side chain properties are highlighted in yellow.

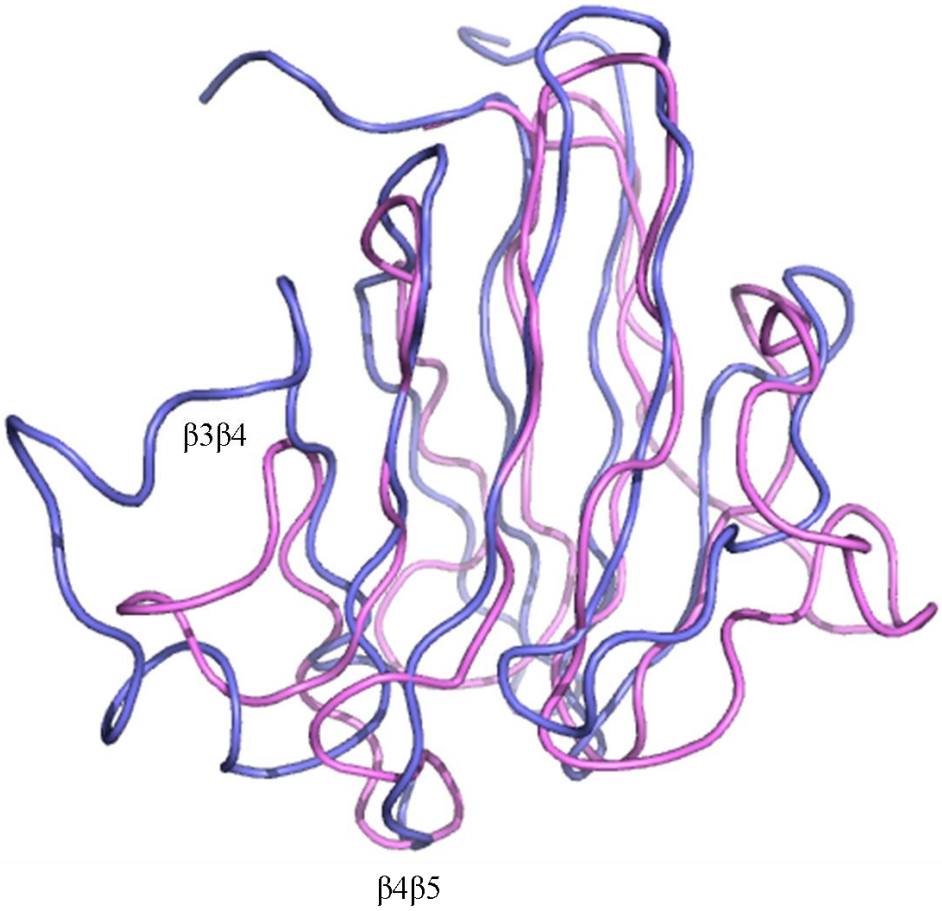

Figure 1.16: Structural alignment between the FHA domains of YscD and MxiG. YscDc is colored purple and MxiG blue. The RMSD between 8 Cα's is 2.309. Loop β3β4 in MxiG contains a 3_{10}-helix.

Regions of functional importance in the YscDc structure

FHA domain containing proteins have been shown to engage in protein-protein or protein-peptide interactions at loops β3β4, β4β5 and often β6β7 of the FHA domain. The FHA domain of EmbR from *M. tuberculosis* interacts with a peptide derived from kinase PknH at FHA loops β3β4, β4β5, and β6β7 (Alderwick, 2005). The FHA domain of Ki67 *from H. sapiens* also recognizes a phosphothreonine peptide

from the Ki67 binding partner hNIFK at FHA loops β3β4 β4β5, and β6β7. The FHA domain of RNF8 from *H. sapiens* binds to a phosphopeptide at loops β3β4 and β4β5 as does the FHA domain of NIPP1 from *Mus musculus* (Kumeta, 2008, Huen, 2007). The FHA domain from the *Corynebacterium glutamicum* protein OdhI binds to a phosphorylated region upstream of its own FHA domain through loops β3β4 and β4β5 (Barthe, 2009). These loops in YscDc, β3β4 (SDPLQ) and β4β5 (SDSEIAPV) are surface exposed and represent possible binding sites between YscDc and T3S proteins (Figure 1.8).

For many FHA-protein or FHA-peptide interactions, there is a conserved serine residue on loop β4β5 of the FHA domain which binds to phosphothreonine residues; this serine is located two residues upstream of the conserved histidine. EmbR, Ki67, RNF8, NIPP1, and OdhI all bind to phosphorylated threonine residues from their respective protein partners via this conserved loop β4β5 serine. In YscDc, this conserved serine is replaced by an alanine (A43), as is the case with many T3S system homologs to YscD (Figure 1.13). The kinesin KIF13, which engages in phopho-independent interactions, also has an alanine in place of this serine residue (Nott, 2009). YscDc does have two serine residues in loop β4β5, Ser38 and Ser40. Ser38 faces towards the β-sandwich interface while S40 faces outward.

To address the importance of loops β3β4 and β4β5, we constructed a *Y. pseudotuberculosis* strain in which *yscD* was deleted, *ΔyscD*. This *yscD* knockout was then complemented with wild-type *yscD* and mutants of *yscD* expressed inducibly from a plasmid to address which regions of the cytoplasmic domain of YscD are involved in the secretion of Yop proteins. *Y. pseudotuberculosis* (*ΔyscD*) was

constructed by replacing *yscD* with kan^R using homologous recombination. This knockout failed to secrete Yop proteins. Alanine substitution mutants were constructed for loops β3β4 (SDPLQ) and β4β5 (DSEI). We confirmed that wild-type *yscD* complemented secretion in the knockout background, but found that neither *yscD*-L3 nor *yscD*-L4 did (Figure 1.17). Additional alanine substitutions were made for the two serines located on loop β3β4, *yscD*-S38A and *yscD*-S40A. Ser40 was found to be essential for Yop secretion, while Ser38 was not essential to secretion.

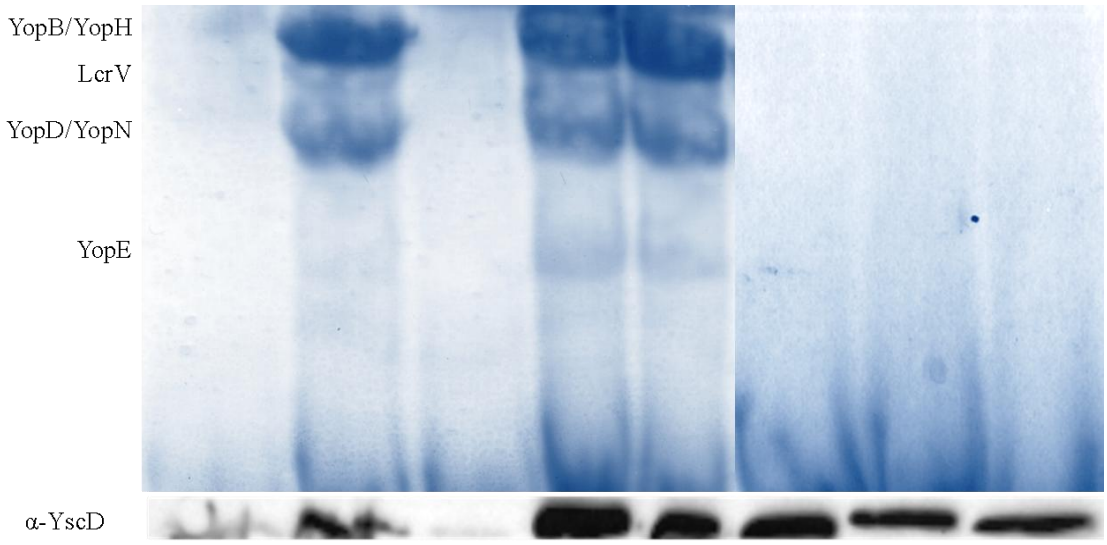

Figure 1.17: Secretion of effector proteins by the T3SS in *Y. pseudotuberculosis.* (A) SDS-PAGE showing protein secretion. (B) Anti-YscD western blot showing expression of YscD from pBAD. Lanes correspond to wild-type grown without $CaCl_2$, wild-type, *ΔyscD*, *ΔyscD* + pBAD-yscD, *ΔyscD* + pBAD-S38A, *ΔyscD* + pBAD-S40A, *ΔyscD* + pBAD-yscDL3, and *ΔyscD* + pBAD-yscDL4.

Discussion

In *Yersinia*, YscD is located in the inner membrane where it is an integral part the T3SS injectisome. Recombinant YscD was solubilized from *E. coli* membranes with detergents into micelles. This artificial membrane environment, which also lacks the other injectisome proteins, may not be suitable for YscD stability, as YscD was not purified to homogeneity. Purified recombinant YscD yielded three predominant species of different molecular mases. These fragments corresponded in molecular mass to full-length YscD, the cytoplasmic domain, and the periplasmic domain. Therefore the cytoplasmic and periplasmic domains, YscDc and YscDp, respectively, were expressed and purified from *E. coli* separately.

YscDp was purified as a single monomeric species. However, after storage at 4 °C for a few days, the presence of a smaller molecular mass species was observed by SDS-PAGE. This YscDp fragment was identified to contain the first 205 amino acids of the periplasmic domain, YscDp205. The secondary structure of YscDp205 is predicted to contain a mix of α-helices and β-strands to comprise two POTRA domains. In the periplasmic region of the YscD ortholog protein, PrgH, there are three POTRA domains which were determined by X-ray crystallography (Spreter, 2009). YscDp205 is predicted to form these domains as well, which are implicated in membrane oligomerization. YscDp205 was cloned, expressed, and purified from *E. coli*. Crystallization trials were conducted with YscDp205. However, these trials did not result in protein crystals. At the C-terminus of YscDp205, there are additional

amino acids not predicted to contain a POTRA motif, these could be removed for subsequent crystallization trials.

YscDc was purified from *E. coli,* crystallized by vapor diffusion, and its X-ray crystallographic structure was determined to 2.5 Å resolution limit. YscDc has a FHA domain fold, a domain identified in numerous proteins of various functions. One of these functions involves signaling via interactions with phosphothreonine residues involving certain loops of the FHA domain. Based on sequence and structural analysis between phosphothreonine-binding FHA domains and YscDc, we determined that YscDc does not have the phosphopeptide binding signature. Recently the structure of the FHA domain from MxiG (YscD ortholog in *S. flexneri*) was determined by NMR (McDowell et al., 2011). A NMR titration with phosphothreonine gave no indication of an interaction with MxiG. YscDc may thus have the same phosphorylation-independent function in the T3SS as MxiG. Additionally T3SS orthologs to YscD and MxiG were studied by secondary structure prediction and sequence analysis. We propose that these proteins contain an FHA domain, identifying this domain as a common feature of the T3SS.

YscD was analyzed for its role in the T3SS. When *yscD* was deleted from *Y. pseudotuberculosis,* no type III secretion was observed. This knockout was transformed with *yscD* encoded on the inducible plasmid, which allowed for YscD expression in *ΔyscD* and the rescue of secretion to wild-type levels. YscD mutants were designed at regions of FHA domains which are known to be important in protein function: loops β3β4 and β4β5. Amino acids from these loops were substituted with

alanines, the resulting mutants were expressed in *ΔyscD,* and analyzed for their ability to rescue secretion. Neither loop mutant was competent for type III secretion, indicating their importance to the T3SS. In addition, two serine residues from loop β4β5 were similarly studied. The mutant S38A was able to rescue secretion while S40A was not, indicating the importance of Ser40 and the non-essentiality of Ser38.

YscDc has been identified to be a crucial part of the T3SS, with specific loops involved in T3SS function. It is possible that the FHA domain is important for the structural integrity of the injectisome in *Y. pseudotuberculosis*. This hypothesis can be addressed by the fusion of GFP to YscD mutants and visualization of punctate fluorescence patterns arising proper assembly of the T3SS, as described by Diepold et al.

The text of chapter 1, in part, is currently being prepared for submission for publication. The dissertation author was the primary researcher. Partho Ghosh contributed and supervised the research which forms the basis for this chapter

References

Achtman, M., K. Zurth, G. Morelli, G. Torrea, A. Guiyoule, and E. Carniel, (1999) *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis Proceedings of the National Academy of Sciences* **96**: 14043-14048.

Adams, P. D., P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echoo ls, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart., (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* **66**: 213-221.

Alderwick, L. J., V. Molle, L. Kremer, A. J. Cozzone, T. R. Dafforn, G. S. Besra, and K. Futterer, (2005) Molecular structure of EmbR, a response element of SerThr kinase signaling in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences* **103**: 2558-2563.

Baker, N. A., D. Sept, S. Joseph, M. J. Holst & J. A. McCammon, (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences* **98**: 10037-10041.

Barthe, P., C. Roumestand, M. J. Canova, L. Kremer, C. Hurard, V. Molle and M. Cohen-Gonsaud, (2009) Dynamic and Structural Characterization of a Bacterial FHA Protein Reveals a New Autoinhibition Mechanism. *Structure* **17**: 568-578.

Bolin, I. a. H. W.-W., (1982) Temperature-inducible outer membrane protein of Yersinia pseudotuberculosis and Yersinia enterocolitica is associated with the virulence plasmid. *Infection and Immunity* **37**.

Bott, M., (2007) Offering surprises: TCA cycle regulation in Corynebacterium glutamicum. *Trends in Microbiology* **15**: 417-425.

Burghout, P., R. van Boxtel, P. Van Gelder, P. Ringler, S. A. Müller, J. Tommassen, and M. Koster, (2004) Structure and electrophysiological properties of the YscC secretin from the type III secretion system of Yersinia enterocolitica. *Journal of Bacteriology* **186**: 4645-4654.

Chen, V. B., W. B. Arendall III, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson and D. C. Richardson, (2010)

MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D* **66**: 12-21.

Cole, C., J. D. Barber & G. J. Barton, (2008) The Jpred 3 secondary structure prediction server. *Nucleic Acids Research* **36**: W197-W201.

Conchas, R. F. a. E. C., (1990) A highly efficient electroporation system for transformation of Yersinia. *Gene* **87**: 133-137.

Cornelis, G. R., (2002) The Yersinia Ysc-Yop 'Type III' weaponry. *Nature Reviews Molecular Cell Biology* **3**: 742-754.

Cornelis, G. R., (2010) The type III secretion injectisome, a complex nanomachine for intracellular 'toxin' delivery. *Biological Chemistry* **391**: 745-751.

Corpet, F., (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research* **16**: 10881-10890.

Datsenko KA, W. B., (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci USA* **97**.

Diepold, A., M. Amstutz, S. Abel, I. Sorg, U. Jenal and G. R. Cornelis, (2010) Deciphering the assembly of the Yersinia type III secreetion injectisome. *The EMBO Journal* **29**: 1928-1940.

Emsley, P. a. K. C., (2004) Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D* **60**: 2126-2132.

Franz, D. R., P. B. Jahrling, D. J. McClain, D. L. Hoover, W. R. Byrne, J. A. Pavlin, G. W. Christopher, T. J. Cieslak, A. M. Friedlander, and E. M. Eitzen Jr., (2001) Clinical recognition and management of patients exposed to biological warfare agents. *Clin. Lab. Med.* **21**: 435-473.

Gasteiger, E., A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel, and A. Bairoch (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* **31**: 3784-3788.

Gemski, P., J. R. Lazere, T. Casey, and J. A. Wohlmieter, (1980) Presence of a virulence-associated plasmid in *Yersinia pseudotuberculosis*. *Infection and Immunity* **28**: 1044-1047.

Gouet, P., X. Robert & E. Courcelle, (2003) ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Research* **31**: 3320-3323.

Heesemann, J., U. Gross, N. Schmidt, R. Laufs, (1986) Immunochemical analysis of plasmid-encoded proteins released by enteropathogenic Yersinia sp. grown in calcium-deficient media. *Infection and Immunity* **54**: 561-567.

Higuchi, R., B. Krummel, and R. K. Saiki, (1988) A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Research* **16**: 7351-7361.

Hodgkinson, J. L., A. Horsley, D. Stabat, M. Simon, S. Johnson, P. C A da Fonseca, E. P. Morris, J. S. Wall, S. M. Lea and A. J. Blocker, (2009) Three-dimensional reconstruction of the Shigella T3SS transmembrane regions reveals 12-fold symmetry and novel features throughout *Nature Structural & Molecular Biology* **16**: 477-485.

Hofmann, K. a. P. B., (1995) The FHA domain: a putative nuclear signalling domain found in protein kinases and transcription factors. *Trends in Biochemical Sciences* **20**: 347-349.

Huen, M. S. Y., R. Grant, I. Manke, K. Minn, X. Yu, M. B. Yaffe, and J. Chen, (2007) RNF8 Transduces the DNA-Damage Signal via Histone Ubiquitylation and Checkpoint Protein Assembly. *Cell* **131**: 901-914.

Inglesby, T. V., D. T. Dennis, D. A. Henderson, J. G. Bartlett, M. S. Ascher, E. Eitzen, A. D. Fine, A. M. Friedlander, J. Hauer, J. F. Koerner, M. Layton, J. McDade, M. T. Osterholm, T. O'Toole, G. Parker, T. M. Perl, P. K. Russell, M. Schoch-Spana & K. Tonat, (2000) Plague as a Biological Weapon. *JAMA: The Journal of the American Medical Association* **283**: 2281-2290.

Johnson, D. L., C. B. Stone, and J. B. Mahony, (2008) Interactions between CdsD, CdsQ, and CdsL, Three Putative Chlamydophila pneumoniae Type III Secretion Proteins *Journal of Bacteriology* **190**: 2972-2980.

Johnson, D. L., C. B. Stone, D. C. Bulir, B. K. Coombes, and J. B. Mahony, (2009) A novel inhibitor of Chlamydophila pneumoniae protein kinase D (PknD) inhibits phosphorylation of CdsD and suppresses bacterial replication. *BMC Microbiology* **9**.

Johnson, D. L. a. J. B. M., (2007) Chlamydophila pneumoniae PknD Exhibits Dual Amino Acid Specificity and Phosphorylates Cpn0712, a Putative Type III Secretion YscD Homolog *Journal of Bacteriology* **189**: 7549-7555.

Kelley, L. A. a. M. J. E. S., (2009) Protein structure prediction on the web: a case study using the Phyre server. *Nature Protocols* **4**: 363-371.

Kimbrough, T. G. a. S. I. M., (2000) Contribution of Salmonella typhimurium type III secretion components to needle complex formation. *Proceedings of the National Academy of Sciences* **97**: 11008-11013.

Koster, M., W. Bitter, H. de Cock, A. Allaoui, G. R. Cornelis, and J. Tommassen, (1997) The outer membrane component, YscC, of the Yop secretion machinery of Yersinia enterocolitica forms a ring-shaped multimeric complex. *Molecular Microbiology* **26**: 789-797.

Kresse, A. U., K. Schulze, C. Deibel, F. Ebel, M. Rohde, T. Chakraborty, and C. A. Guzmán, (1998) Pas, a novel protein required for protein secretion and attaching and effacing activities of enterohemorrhagic Escherichia coli. *Journal of Bacteriology* **180**: 4370-4379.

Krogh, A., B. Larsson, G. von Heijne & E. L. L. Sonnhammer, (2001) Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *Journal of Molecular Biology* **305**: 567-580.

Kubori, T., Y. Matsushima, D. Nakamura, J. Uralil, M. Lara-Tejero, A. Sukhan, J. E. Galán and S.-I. Aizawa, (1998) Supramolecular Structure of the Salmonella typhimurium Type III Protein Secretion System. *Science* **280**: 602-605.

Kumeta, H., K. Ogura, S. Adachi, Y. Fujioka, K. Tanuma, K. Kikuchi and F. Inagaki, (2008) The NMR structure of the NIPP1 FHA domain. *Journal of Biomolecular NMR* **40**: 219-224.

Lathem, W. W., P. A. Price, V. L. Miller & W. E. Goldman, (2007) A Plasminogen-Activating Protease Specifically Controls the Development of Primary Pneumonic Plague. *Science* **315**: 509-513.

Li, J., G. Lee, S. R. Van Doren, and J. C. Walker (2000) The FHA domain mediates phosphoprotein interactions. *Journal of Cell Science* **113**: 4143-4149.

Lian, C. J., W. S. Hwang, and C. H. Pai, (1987) Plasmid-mediated resistance to phagocytosis in *Yersinia enterocolitica Infection and Immunity* **55**: 1176-1183.

McDowell, M. A., S. Johnson, J. E. Deane, M. Cheung, A. D. Roehrich, A. J. Blocker, J. M. McDonnell & S. M. Lea, (2011) Structural and functional studies on the N-terminal domain of the Shigella type III secretion protein MxiG. *Journal of Biological Chemistry*.

Michiels, T., J. C. Vanooteghem, C. Lambert de Rouvroit, B. China, A. Gustin, P. Boudry, and G. R. Cornelis, (1991) Analysis of virC, an operon involved in the secretion of Yop proteins by Yersinia enterocolitica. *Journal of Bacteriology* **173**: 4994-5009.

Morita-Ishihara, T., M. Ogawa, H. Sagara, M. Yoshida, E. Katayama, and C. Sasakawa, (2006) Shigella Spa33 is an essential C-ring component of type III secretion machinery. *Journal of Biological Chemistry* **281**: 599-607.

Mougous, J. D., C.A. Gifford, T.L. Ramsdell and J.J. Mekalanos, (2007) Threonine phosphorylation post-translationally regulates protein secretion in Pseudomonas aeruginosa. *Nature Cell Biology* **9**: 797 - 803.

Neyt, C. a. G. R. C., (1999) Insertion of a Yop translocation pore into the macrophage plasma membrane by Yersinia enterocolitica: Requirement for translocators YopB and YopD, but not LcrG. *Molecular Microbiology* **33**: 971-981.

Nott, T. J., G. Kelly, L. Stach, J. Li, S. Westcott, D. Patel, D. M. Hunt, S. Howell, R. S. Buxton, H. M. O'Hare, and S. J. Smerdon, (2009) An Intramolecular Switch Regulates Phosphoindependent FHA Domain Interactions in Mycobacterium tuberculosis. *Science Signaling* **2**: ra12.

Ogino, T., R. Ohno, K. Sekiya, A. Kuwae, T. Matsuzawa, T. Nonaka, H. Fukuda, S. Imajoh-Ohmi, and A. Abe, (2006) Assembly of the Type III Secretion Apparatus of Enteropathogenic Escherichia coli *Journal of Bacteriology* **188**: 2801-2811.

Otwinowski, Z. a. W. M., (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode *Macromolecular Crystallography* **276**: 307-326.

Perry, R. D. a. J. D. F., (1997) *Yersinia pestis* - etiologic agent of plague. *Clinical Microbiology Reviews* **10**: 35-66.

Plano, G. V. a. S. C. S., (1995) Mutations in yscC, yscD, and yscG prevent high-level expression and secretion of V antigen and Yops in Yersinia pestis. *Journal of Bacteriology* **177**: 3843-3854.

Rosqvist R, F. A., Rimpilainen M, Bergman T, Wolf-Watz H, (1990) The cytotoxic protein YopE of *Yersinia* obstructs the primary host defense. *Mol Microbiol* **4**: 657-667.

Sanowar, S., P. Singh, R. A. Pfuetzner, I. André, H. Zheng, T. Spreter, N. C. Strynadka, T. Gonen, D. Baker, D. R. Goodlett, and S. I. Miller, (2010) Interactions of the Transmembrane Polymeric Rings of the Salmonella enterica Serovar Typhimurium Type III Secretion System. *mBio* **1**: e00158-00110.

Schraidt, O., M. D. Lefebre, M. J. Brunner, W. H. Schmied, A. Schmidt, J. Radics, K. Mechtler, J. E. Galán, and T. C. Marlovits, (2010) Topology and Organization of the Salmonella typhimurium Type III Secretion Needle Complex Components. *PLoS Pathogens* **6**: e1000824.

Schrödinger, L., The PyMOL Molecular Graphics System.

Spreter, T., C. K. Yip, S. Sanowar, I. André, T. G. Kimbrough, M. Vuckovic, R. A. Pfuetzner, W. Deng, A. C. Yu, B. B. Finlay, D. Baker, S. I. Miller, and N. C. Strynadka, (2009) A conserved structural motif mediates formation of the periplasmic rings in the type III secretion system. *Nature Structural & Molecular Biology* **16**: 468-476.

Stone, C. B., D. L. Johnson, D. C. Bulir, J. D. Gilchrist, and J. B. Mahony, (2008) Characterization of the Putative Type III Secretion ATPase CdsN (Cpn0707) of Chlamydophila pneumonia. *Journal of Bacteriology* **190**: 6580-6588.

Tanaka, Y., M. Kuroda, Y. Yasutake, M. Yao, K. Tsumoto, N. Watanabe, T. Ohta, and I. Tanaka, (2007) Crystal structure analysis reveals a novel forkhead-associated domain of ESAT-6 secretion system C protein in Staphylococcus aureus. *Proteins: Structure, Function, and Bioinformatics* **69**: 659–664.

Van Duyne, G. D., R. F. Standaert, P. A. Karplus, S. L. Schreiber and J. Clardy, (1993) Atomic Structures of the Human Immunophilin FKBP-12 Complexes with FK506 and Rapamycin. *Journal of Molecular Biology* **229**: 105-124.

Wang, D., A. J. Roe, S. McAteer, M. J. Shipston, and D. L. Gally, (2008) Hierarchal type III secretion of translocators and effectors from Escherichia coli O157:H7 requires the carboxy terminus of SepL that binds to Tir. *Molecular Microbiology* **69**: 1499-1512.

Ye, Y. & A. Godzik, (2004) FATCAT: a web server for flexible structure comparison and structure similarity searching. *Nucleic Acids Research* **32**: W582-W585.

Yip, C. K., T. G. Kimbrough, H. B. Felise, M. Vuckovic, N. A. Thomas, R. A. Pfuetzner, E. A. Frey, B. B. Finlay, S. I. Miller, and N. C. Strynadka, (2005) Structural characterization of the molecular platform for type III secretion system assembly. *Nature Reviews Molecular Cell Biology* **435**: 702-707.

Zenk, S. F., D. Stabat, J. L. Hodgkinson, A. K. Veenendaal, S. Johnson, and A. J. Blocker, (2007) Identification of minor inner-membrane components of the Shigella type III secretion system ‘needle complex’ *Microbiology* **153**: 2405-2415.

II.

Appendix:

Purification and characterization of

YscO and YscP

Abstract

Yersiniae express a type III secretion system (T3SS) which is necessary for pathogenesis. This system utilizes a transport apparatus, the injectisome, to translocate virulence effector proteins into host cells. The injectisome spans the membranes of the bacterial envelope, terminating with an extracellular appendage resembling a needle. Formation of the T3SS needle is regulated by a size-controlling mechanism that is crucial for bacterial survival in the host. After the completion of the needle complex, a switching event occurs, allowing for the transport of effector proteins into the host cell. The *Yersinia* protein responsible for the control of both needle length and substrate switching is YscP. To understand the mechanisms by which YscP exhibits this control, crystallization of this protein was pursued. Recombinant YscP was overexpressed and purified for crystallization screening. Additional experiments examining the potential interaction of YscP with *Yersinia* T3SS proteins were conducted. One of the *Yersinia* proteins that is potentially involved in YscP function, YscO, was also overexpressed and purified for structural studies.

Introduction

The *Yersinia* T3SS injectisome is used to introduce effector proteins into the host cell, where these proteins function to disrupt the host immune response. Indirect evidence suggests that these bacterial effector proteins are translocated through the needle and into the host cell by way of a pore formed in the host cell membrane by bacterial translocon proteins. Assembly of this needle complex occurs in a regulated and hierarchical manner so that there is sequential export of needle proteins, translocon pore-forming proteins, and then effector proteins by the T3SS (Sukhan, 2001). As with many multiprotein structures, polymerization of the T3SS needle is regulated by a size-controlling mechanism, which is crucial for bacterial survival in the host. Abrogation of needle length control suppresses secretion of effector proteins (Mota, 2005). Upon completion of needle polymerization, a switching event occurs, allowing for the transport of translocon and effector proteins into the host cell.

A protein essential to the process of needle length control and the switching between needle formation and effector transport in *Yersinia* is YscP. Without substrate switching, the length of the needle polymer is uncontrolled and effector secretion is low (Payne & Straley, 1999). Needle length control has been mapped to segments of *Y. enterocolitica* YscP called the ruler region: deletions and insertions within the central amino acid residues of YscP produce shorter and longer needles, respectively (Journet, 2003). These needles still grow to a controlled length whereas deletions in the N- and C-terminus of YscP result in a complete loss of needle length control. This suggests that the terminal residues may serve to anchor the protein at the

needle tip and injectisome base, while the central amino acid residues act as a molecular ruler, directing the needle length (Journet, 2003). The C-terminal amino acid residues are termed the type III secretion substrate specificity switch or T3S4. Loss of the T3S4 in *Y. enterocolitica* results in an effector protein secretion defect (Agrain, 2005). A schematic of *Y. pseudotuberculosis* YscP is shown in Figure 2.1.

Figure 2.1: Schematic of functionally important regions of YscP from *Y. pseudotuberculosis.* There are two regions important for YscP secretion located at the N-terminus, deletion of either region results in a loss of YscP protein secretion. The ruler region is predicted to have no regular secondary structure. The ruler can be shortened or lengthened to decrease or increase the injectisome needle size. The Type III Secretion Substrate Specificity Switch (T3S4) located at the C-terminus of YscP is important for the secretion of effector proteins.

A role for YscP in the switching mechanism has been supported by experiments with the *Yersinia* T3SS bacterial inner membrane protein YscU. Cleavage at the C-terminus of YscU is essential for translocon but not effector transport through the T3SS (Sorg, 2007). Mutations in the C-terminus of YscU can partially suppress a *yscP* mutant, which has a phenotype of uncontrolled needle polymerization (Edqvist, 2003). These mutations in YscU result in greater levels of effector secretion in the *yscP* mutant while reducing export of the needle protein (Edqvist, 2003). Additional information about needle protein export control is found in studies of the flagellar T3SS of *Salmonella enterica.* Deletion of FliK, the homolog

of YscP, produces uncontrolled elongation of the hook appendage, and the interaction of FliK with FlhB (homolog of YscU) results in substrate specificity switching from hook proteins to late secretion proteins (Moriya, 2006).

The *yscP* gene is located in the *virB* operon of the *Y. pseudotuberculosis* virulence plasmid pYV. The gene upstream of *yscP* is *yscO*, which encodes an 18.9 kDa protein. Deletion of *yscO* suppresses the secretion of effector proteins, similar to *yscP* knockouts (Payne, 1998, Payne & Straley, 1999). YscO and YscP were first hypothesized to interact when it was found that the overexpression of *yscO* rescued secretion inhibition resulting from the overexpression of *yscP* (Payne & Straley, 1999). This interaction was confirmed when glutathione S-transferase tagged YscP was found to bind YscO from *Y. enterocolitica* lysates (Riordan, 2008). YscO is thus a possible candidate for assisting in YscP-dependent control of effector secretion.

The capability of *Yersinia* to control the length of the injectisome apparatus is not unique. Along with the needle-length control of T3SS-harboring bacteria, the hooks of bacterial flagella and the tails of bacteriophages are similarly controlled by a sizing mechanism (Cornelis, 2006). Mapping the process by which Yersinia needle length is regulated through YscP will contribute to defining the mechanisms of structural control in various microbial machines.

Materials and Methods

Expression and purification of YscP, Δ46-96, Δ46/Δ222, and T3S4

The coding sequences for *yscP* and the T3S4 domain of *yscP* were amplified by PCR from the pYV plasmid of *Y. pseudotuberculosis* 126. The PCR products were ligated into the pET28b expression vector. These constructs include the coding sequence for the PreScission Protease cleavage site (LEVLFQ/GP, with the slash indicating the cleavage site) followed by six histidines at the C-terminus of *yscP* and the T3S4 domain. The integrity of the resulting vectors was verified by DNA sequencing. The truncation constructs of YscP, *Δ46* and *Δ46/Δ222,* were constructed by strand overlap extension PCR (Stratagene).

YscP, Δ46, Δ46/Δ222, and T3S4 were expressed from pET28b in *E. coli* BL21 (DE3). Bacteria were grown at 37 ºC in LB media supplemented with 50 mg/L kanamycin to an OD_{600} of 0.5, at which point expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The bacteria were grown further for 16 h at 20 ºC. Bacteria were then harvested by centrifugation (5,800 x *g*, 10 min, 4 ºC). The bacterial pellet was resuspended in 1/100 of the culture growth volume in buffer A (500 mM NaCl, 50 mM sodium phosphate, pH 8.0, 10 mM β-mercaptoethanol) supplemented with 1 tablet of EDTA free protease cocktail inhibitor (Roche) per 2 L of culture growth. Resuspended bacteria were lysed using an Emulsiflex-C5 homogenizer with three passes at 15,000 psi, and the lysate was clarified by centrifugation (14,000 x *g*, 10 min, 4 ºC). The supernatant was applied to a Ni^{2+}-nitrilotriacetic acid (Ni-NTA) agarose column, the column was washed with 25

column volumes of buffer A containing 10 mM imidazole, and bound protein eluted from the column with three column volumes of buffer A containing 500 mM imidazole. The eluted fractions were concentrated by ultrafiltration using a YM-10 Centricon and further purified by size-exclusion chromatography (16/60 Superdex 200, GE Healthcare) in buffer B (20 mM NaCl, 20 mM HEPES, pH 7.5, and 10 mM β-mercaptoethanol). YscP was cleaved at a 50:1 protein:protease molar ratio with PreScission Protease to remove the His-tag, and the cleaved sample was applied to a Ni-NTA agarose column in buffer B. Cleaved YscP was isolated from the flow-through of the column and concentrated by ultrafiltration using a YM-3 Centricon.

Expression and purification of YscO

The coding sequence of *yscO* was amplified by PCR from the pYV plasmid of *Y. pseudotuberculosis* 126. The PCR product was ligated into the pET28b expression vector. This construct includes the coding sequence for six histidines followed by a thrombin cleavage site (LVPR/GS) prior to the N-terminus of YscO. The integrity of the resulting vector, pET28b-yscO, was verified by DNA sequencing.

YscO was expressed from pET28b-yscO in *E. coli* BL21 (DE3). Bacteria were grown at 37 °C in LB media supplemented with 50 mg/L kanamycin to an OD_{600} of 0.6, at which point expression was induced with 0.5 mM IPTG. The bacteria were grown further for 4 h at 37 °C. Bacteria were then harvested by centrifugation (5,800 x *g*, 10 min, 4 °C). The bacterial pellet was resuspended in 1/100 of the culture growth volume in buffer A and supplemented with 10 mM phenylmethylsulfonyl fluoride (PMSF), 20 μg/ml DNase, and 1 mg/ml lysozyme. Resuspended bacteria

were lysed by sonication on ice. Inclusion bodies, which contained YscO, were pelleted from soluble protein by centrifugation (14,000 x *g*, 10 min, 4 ºC). Inclusion bodies were washed with buffer A supplemented with 0.1 % Triton X-100 and centrifuged (14,000 x *g*, 10 min, 4 ºC); this wash was repeated seven times to remove soluble protein. Inclusion bodies were resuspended in denaturing buffer (100 mM NaCl, 8 M guanidinium hydrochloride, 50 mM Tris, pH 8.0, 10 mM βME) to a protein concentration of 10 mg/mL as determined by Bradford Assay. The sample was ultracentrifuged (50,000 x *g*, 1 hr, 24 °C) to remove debris, and the supernatant containing YscO was stored at -80 °C.

YscO was refolded by rapidly diluting the sample to 2 μM in refolding buffer (750 mM arginine, 25% glycerol, 1% PEG 4000, 30 mM NaCl, 50 mM Tris, pH 8.5, 1 mM dithiothreitol) with stirring. The refolding mixture was incubated at 16 °C for 16 h. Refolded YscO was filtered and loaded onto a Ni-NTA agarose column. The column was washed with five column volumes of renaturation buffer (500 mM NaCl, 50 mM sodium phosphate, pH 8.5, 10 mM β-mercaptoethanol) containing 7 mM imidazole. Refolded YscO was eluted in buffer C (300 mM NaCl, 50 mM sodium phosphate, pH 8.5, 10 mM β-mercaptoethanol) containing 500 mM imidazole. Eluted fractions were concentrated by ultrafiltration using a YM-3 Centricon and further purified by size-exclusion chromatography (10/30 Superdex 75, GE Healthcare) in buffer C.

YscO and YscP binding assays

Y. pseudotuberculosis lysates were prepared from 2 L of culture grown at 26 °C in BHI media. At an OD_{600} of 0.5, the bacterial culture was supplemented with 10 mM $MgCl_2$ and 10 mM $Na_2C_2O_4$. Bacteria were grown further at 37 °C for 3.5 h and then pelleted by centrifugation (5,800 x *g*, 10 min, 4 °C). Bacteria were resuspended in 15 mL of lysis buffer (25 mM sodium phosphate, pH 8.0, 100 mM NaCl, 5% glycerol, 1% Triton X-100, 1 mM PMSF, 10 mM β-mercaptoethanol). Bacteria were lysed by sonication, and insoluble debris was removed by centrifugation (14,000 x *g*, 10 min, 4 °C). The clarified lysate was used immediately for binding assays.

Binding assays were performed using 25 μL Ni-NTA agarose in 1 mL microcentrifuge tubes. The nickel beads were pre-equilibrated with binding buffer (100 mM NaCl, 50 mM sodium phosphate, pH 8.5, 10 mM β-mercaptoethanol), then incubated with 50 μg of bait protein for 1 h at 4 °C (His-YscO or His-YscP). Unbound protein was removed after centrifugation (4,000 x *g*, 1 min, 25 °C) and the nickel beads were washed two times in 500 μL binding buffer. To determine if lysate proteins bound to the nickel beads, a control assay was performed with the bait protein omitted. The nickel beads were then incubated in 500 μL of *Y. pseudotuberculosis* lysate for 30 min at 4 °C. Unbound proteins were removed after centrifugation (4,000 x *g*, 1 min, 25 °C), and the nickel beads were washed 10x in 500 μL of binding buffer containing 20 mM imidazole. Samples were removed from the nickel beads by incubating in 30 μL SDS-PAGE loading buffer and boiling for 5 min. Proteins bound to the nickel beads were visualized by SDS-PAGE.

YscP crystallization trials

Purified YscP constructs (YscP, Δ46, Δ46/Δ222, and T3S4) were set up in vapor-diffusion crystallization trials. The YscP buffer contained 20 mM HEPES, pH 7.5, 20 mM NaCl, and 10 mM β-mercaptoethanol. Proteins, at concentrations ranging between 5-10 mg/mL, were mixed with precipitant at 1:2, 1:1, and 2:1 ratios. Conditions were screened both by hanging and sitting drop methods. Precipitant solutions were from Hampton and Jena Bioscience.

Additional crystallization trials were carried out using YscP proteins (YscP, Δ46, Δ46/Δ222) with reductively methylated lysines (Walter, 2006). The success of the methylation procedure was verified by MALDI mass spectrometry (Scripps Center for Metabolomics and Mass Spectrometry), which demonstrated the expected increase in molecular mass indicating 100% methylation. Lysine-methylated proteins were purified by gel filtration chromatography (Superdex 200, GE) and set up in crystallization trials at concentrations between 10-20 mg/mL, as described above.

Results

Sequence analysis of YscP

YscP is a soluble protein with a molecular weight of 50.4 kDa. Initial bioinformatic characterization of YscP was performed using the secondary structure prediction server PredictProtein (Rost, 2004). This server calculated a significant portion of the N-terminus of YscP to have no regular secondary structure (Figure 2.2). A BLAST search of these N-terminal amino acids gives no sequence similarity among YscP homologs, possibly indicating species-dependent interactions at this terminus. The C-terminus, however, is predicted to be structured and contains a stretch of 110 amino acid residues that is 35 % identical to the C-terminal portions of the YscP homologs AscP, PscP, and LscO (Figure 2.3). These amino acids comprise the Type Three Secretion Substrate Specificity Switch, T3S4, which is essential for secretion of effector proteins after needle and translocon export is complete. Only deletions in the T3S4 region of *yscP* result in effector secretion deficiencies (Agrain, 2005).

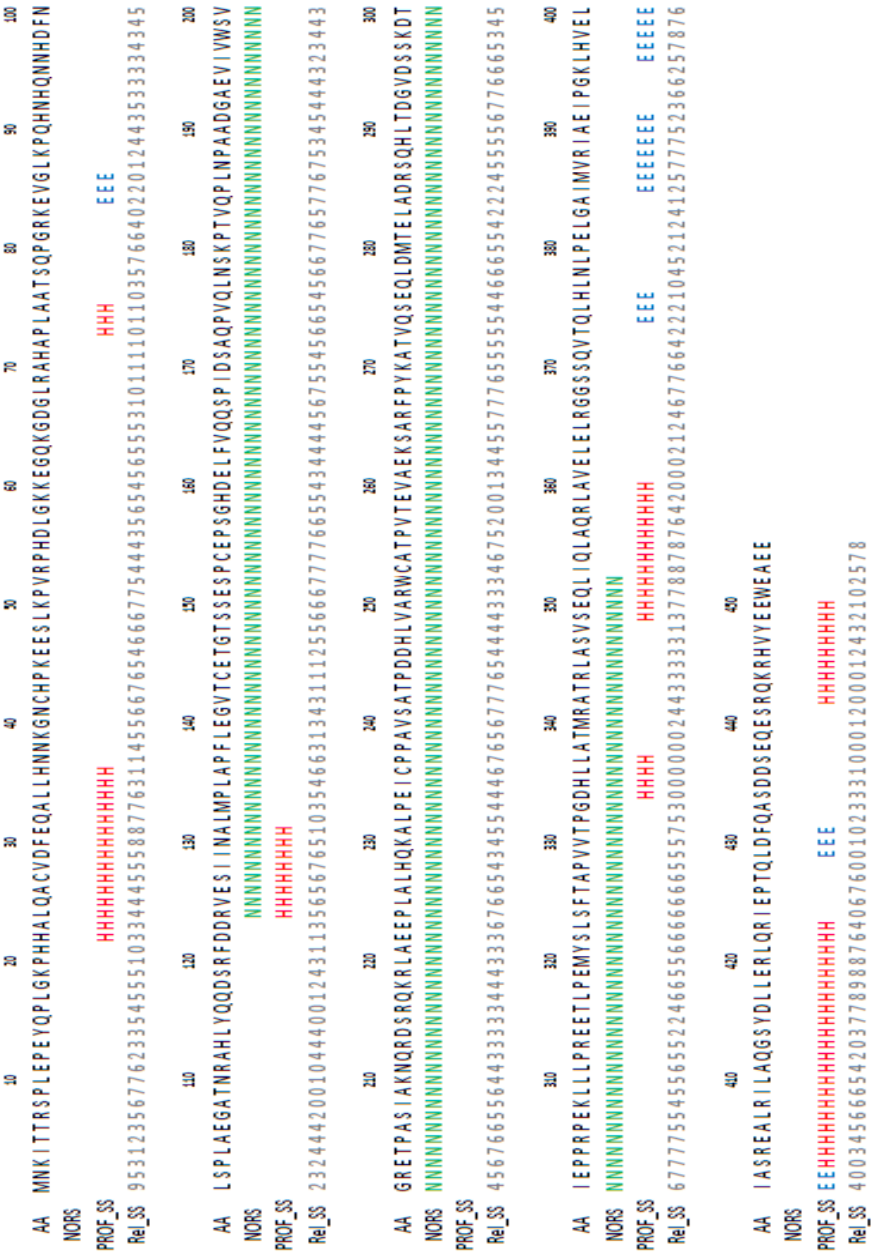

Figure 2.2: Secondary structure prediction of YscP by PredictProtein. The green Ns correspond to regions which are predicted to have no regular secondary structure (NORS: Non-Ordinary Secondary Structure). E represents a β-strand and H represents an α-helix (PROF_SS: Profile network prediction HeiDelberg). The confidence for this prediction is shown in the Rel_SS row (Reliability for PROFsec prediction, 0=low to 9=high).

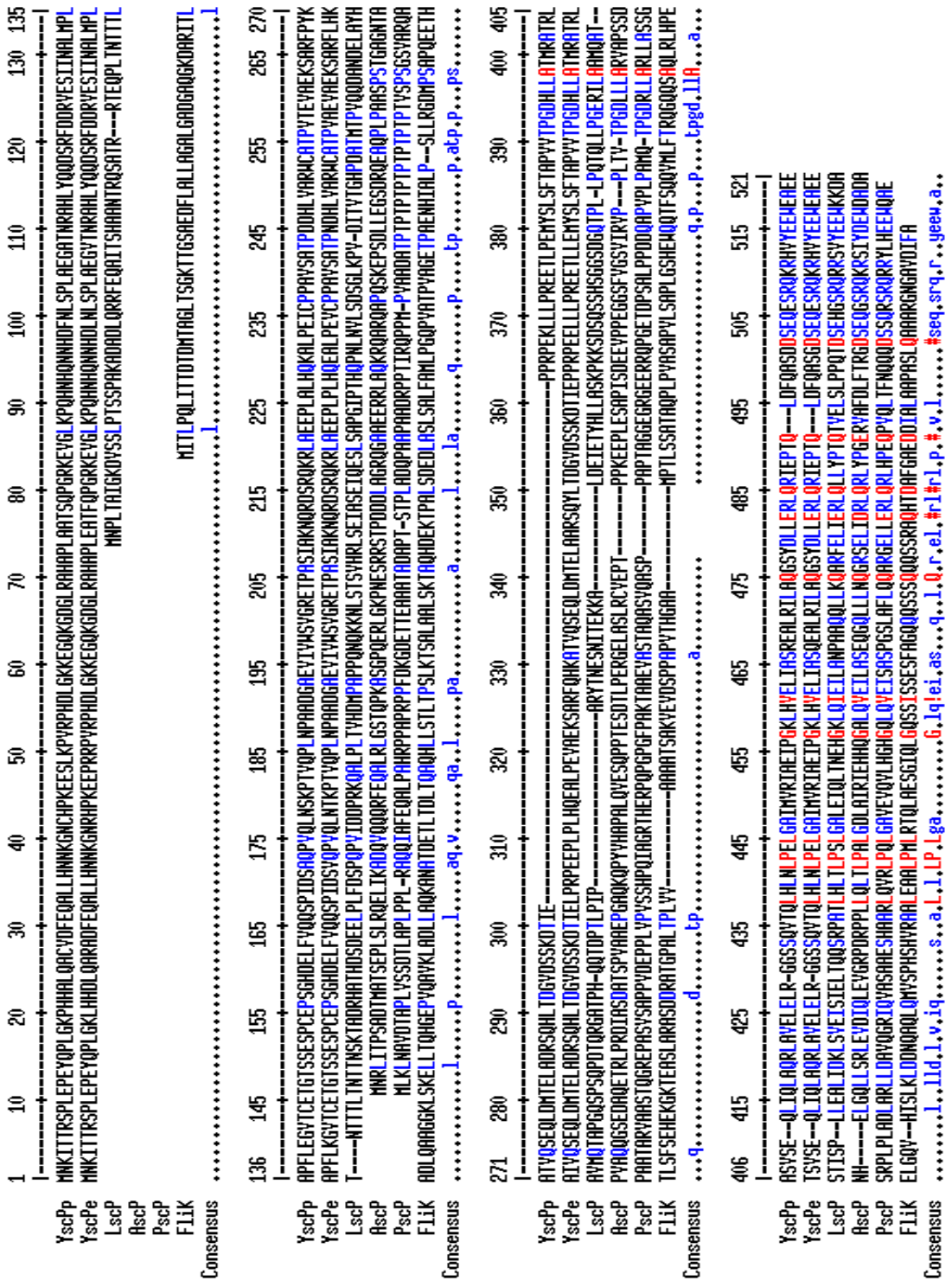

Figure 2.3: Alignment of YscP with T3SS homologs. Subscripts denote differing species: $YscP_p$, *Y. pseudotuberculosis*; $YscP_e$, *Y. enterocolitica*. *Y. enterocolitica* has a longer sequence located in the ruler region. Alignment done with Multialin (Corpet, 1988).

Recently the structure of the flagellar homolog of YscP, FliK, was determined by nuclear magnetic resonance (NMR). This structure includes the T3S4 region (Ibuki, 2011). The T3S4 of FliK has the same secondary structure as the predicted secondary structure of YscP, despite a sequence identity of only 19% (Figure 2.4). A model of the T3S4 region of YscP has been modeled from the FliK structure using the Phyre2 server (Figure 2.5) (Kelley, 2009). The Qmean score for this model is 0.411 (total range is 0-1, where a value of 0 represents a poor model and a value of 1 represents a good model), which gives a Z-score of -2.62 when related to Qmean scores of reference PDB structures The Z-score represents how close the statistics for the model align with structures solved using experimental data (Benkert, 2008). Experimental structures have Qmean scores closer to 1, indicating that this model may not be the best fit for the T3S4. This secondary structure is also predicted in another molecular ruler protein, the bacteriaphage gpH. GpH has a ruler function similar to the one proposed for YscP (Abuladze, 1994). The three dimensional structure of the T3S4 may crucial for effector secretion in *Yersinia*. The T3S4 of YscP can be swapped with the sequence for AscP or PscP while maintaining secretion. The T3SS either recognizes the structure of the T3S4 or a sequence of the T3S4 which is shared among this YscP family (Agrain, 2005). It is possible that one of the inner membrane components of the injectisome could recognize this T3S4 fold during YscP secretion.

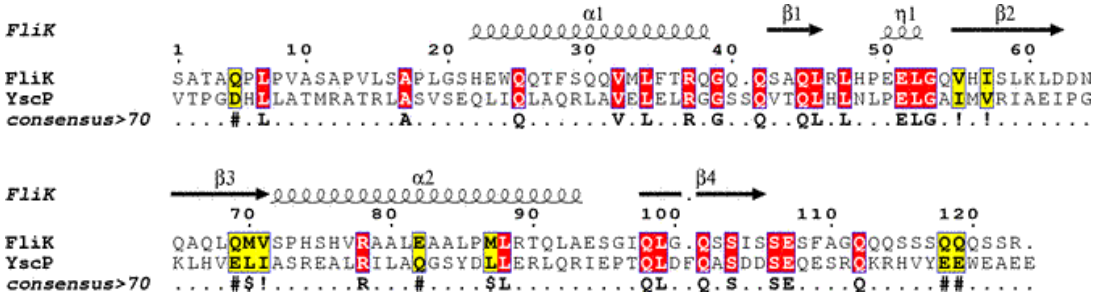

Figure 2.4: Alignment of the T3S4 regions of YscP and FliK. Sequence alignment was done using Multialin (Corpet, 1988). The secondary structure of FliK was done using ESPript (Gouet *et al.*, 2003). Amino acids identical between the proteins are highlighted in red. Amino acids with sides of similar properties are highlighted in yellow.

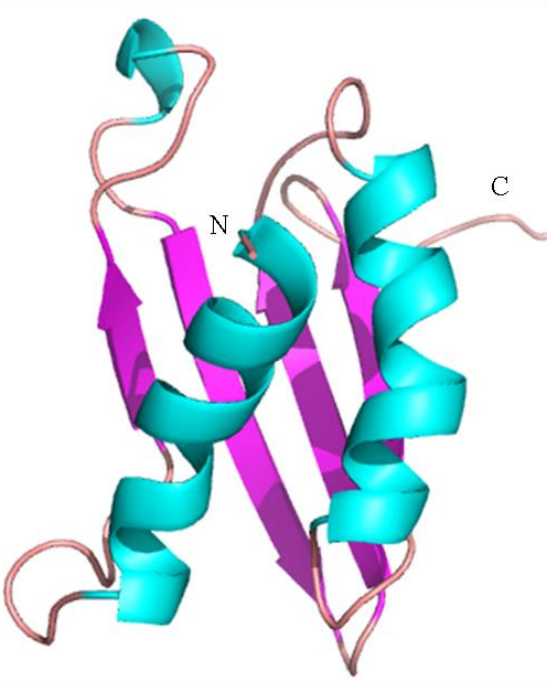

Figure 2.5: Model of the YscP T3S4 region using FliK as a template. This model was created using PhyreBeta and covers amino acids 350-436 of YscP from *Y. pseudotuberculosis.*

Sequence analysis of YscO

YscO is an 18.9 kDa protein which is localized in both the cytosolic and membrane-associated portions of *Y. pseudotuberculosis* (Payne, 1998). The localization of YscO is corroborated from studies conducted on the YscO homolog, FliJ, from the flagellar T3SS. FliJ acts as a chaperone transporter, shuttling chaperones away from the injectisome after release of the chaperone-effector complex (Evans, 2006). Based on amino acid sequence alignment, YscO is 14% identical and 26% similar to FliJ (Figure 2.6). The structure of FliJ reveals a coiled coil which is also predicted to form in YscO (Ibuki, 2011). YscO may function in chaperone shuttling since YscO is able to bind to a translocon chaperone SycD. This selective binding to translocon chaperone suggests that YscO may also play a role in the hierarchy of effector secretion (Evans, 2006; Evans, 2009).

Figure 2.6: Sequence alignment of YscO with FliJ. The secondary structure of FliJ is shown above the alignment. All alignments were done using the Multialign server. The secondary structure of FliJ was added to the alignment using ESPript.

In addition to chaperone shuttling, YscO may have other functions in the T3SS. YscO localizes at the base of the injectisome where it interacts with another

crucial T3SS protein, the ATPase YscN (Riordan, 2008). Similarly, FliJ interacts with the coiled coil domain of the flagellar ATPase FliI (Ibuki, 2011). This interaction also occurs between a T3SS homolog of YscO (15% identity), Cpn0706, and the ATPase CdsN in *Chlamydophila pneumonia* (Stone, 2008). The function of the YscO-YscN association in *Yersinia* is unknown, however FliJ is known to increase the ATPase activity of FliI (Evans, 2009). In *Yersinia*, YscN function is essential for effector transport into the host cell and YscO may assist in this process.

Interactions between T3SS coiled coil containing proteins is not only known to occur between YscO and YscN homologs, but also between the homologs of YscO and YscL in the *Pseudomonas syringae* T3SS. HrpO (homolog to YscO with 13% identity) interacts with the predicted coiled coil of HrpE (homolog to YscL). This association can be tied to ATPase function since HrpE is the negative regulator of the *P. syringae* ATPase. This F_0F_1 type ATPase has three regions: a membrane bound F_0 region, a stalk region, and the F_1 region containing the catalytic site. HrpE, the YscL homolog, is similar in sequence and predicted secondary structure to the stalk component of the *P. syringae* ATPase (Gazi, 2008). These data further implicate a role for YscO in facilitating ATPase-dependent T3SS.

Proteins which are the most similar in sequence to YscO are members of the Ysc family of T3SSs. There is 25% sequence identity between the family members AscO (*Aeromonas hydrophila*), PscO (*Pseudomonas aeruginosa*), and LscO (*Photorhabdus luminescens*) (Figure 2.7). A more divergent T3SS homolog to YscO is CT670 from *Chlamydia trachomatis*, which is 18% identical to YscO (Figure 2.8). CT670 adopts a coiled coil similar in structure to FliJ, with a backbone RMSD of 3.21

Å as calculated from the Dali server (Holm, 2010). The secondary structures of CT670 and FliJ are also similar to that predicted for YscO (Rost, 2004, Lorenzini, 2010). A model of YscO using FliJ as a template was created using EsyPred3D and is shown in Figure 2.9 (Lambert, 2002). The Qmean score for this model is 0.584 (scale of 0-1), with a Z-score compared to a set of PDB structures at -1.76 (Benkert, 2008). Among YscO homologs there are two sequence stretches located at either end which are conserved (Figure 2.7). In the YscO model this corresponds to one end of the coiled coil, where the N- and C-termini meet. Within these YscO homologs, eight of the conserved residues are glutamic acids: E14, E17, E36, E95, E119, E140, E141, E145 and E146. This creates a charged region which may be important for YscO binding to T3SS components.

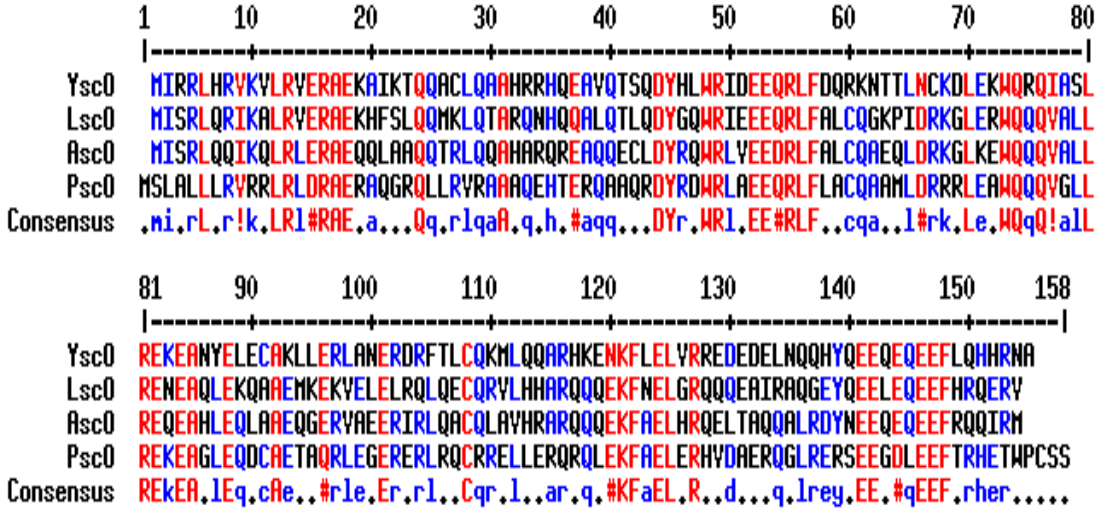

Figure 2.7: YscO alignment with Ysc family members. Highly conserved residues (90%) are highlighted in red with similar residues (50%) colored in blue.

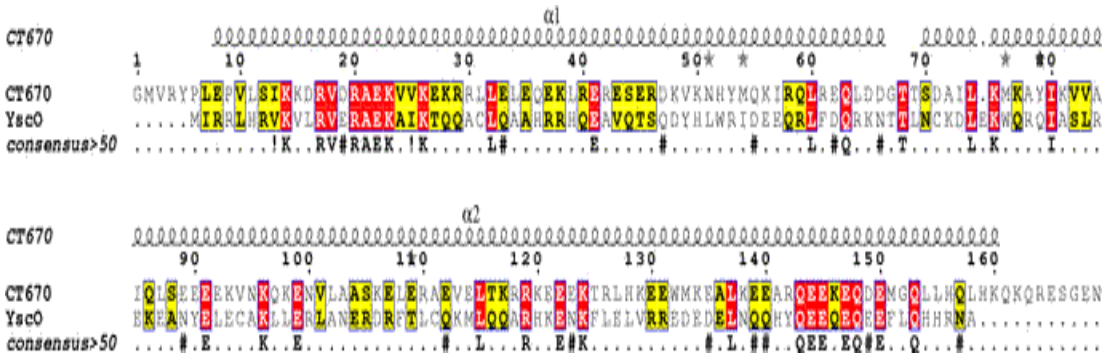

Figure 2.8: Sequence alignment of YscO with *Chlamydia* CT670. The secondary structure of CT670 is shown above the alignment. This figure was made using Multialin and ESPript (Corpet, 1988, Gouet et al., 2003).

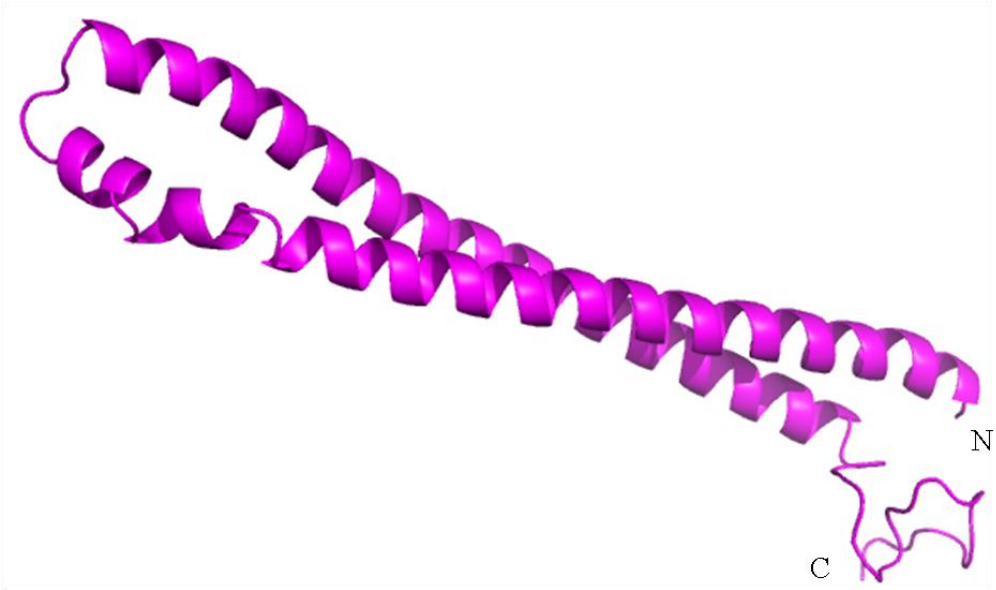

Figure 2.9: Model of YscO structure using CT670 as a template. The model was created using the EsyPred3D server (Lambert, 2002).

Purification of YscP

YscP was purified by nickel affinity and gel filtration chromatography (Figures 2.10, 2.11). Two species were evident from gel filtration, with retention volumes indicative of molecular weights of 360 and 141 kDa, respectively (Figure 2.11). These molecular weight determinations are based on globular protein standards and may be misleading since YscP presumably has an extended structure. Early retention volumes may represent structural elongation or a multimeric complex. A sample of YscP that eluted at an apparent molecular weight of 141 kDa was re-introduced onto the gel filtration column to determine if the two species of YscP were in equilibrium. The chromatogram showed a single peak, indicating two independent, non-exchanging species of YscP. Only fractions containing YscP from the predominant peak at 141 kDa were collected and subsequently used for crystallization trials.

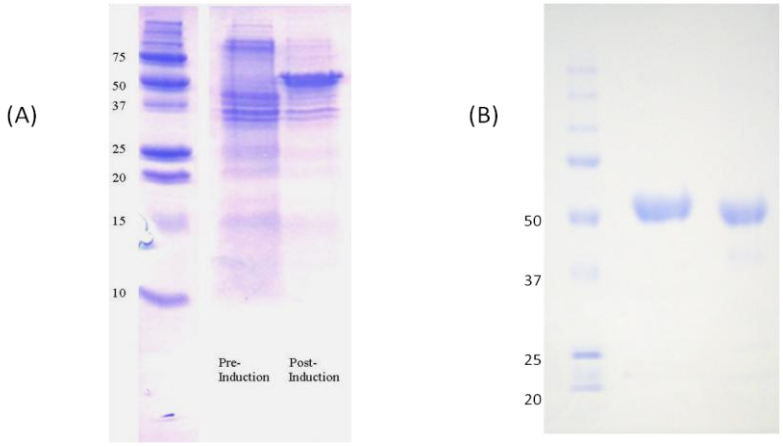

Figure 2.10: Expression and purification of YscP from *E. coli*. (A) SDS-PAGE visualization of YscP expressed in *E. coli*. Lanes correspond to protein marker, total bacterial proteins prior to IPTG induction, and 16 hours after induction. (B) SDS-PAGE visualization of purified YscP before and after removal of the histidine tag by PreScission protease. Lanes correspond to protein marker, YscP after Ni^{2+}-NTA purification, and YscP after PreScission protease digestion. YscP has a molecular weight of 50.4 kDa.

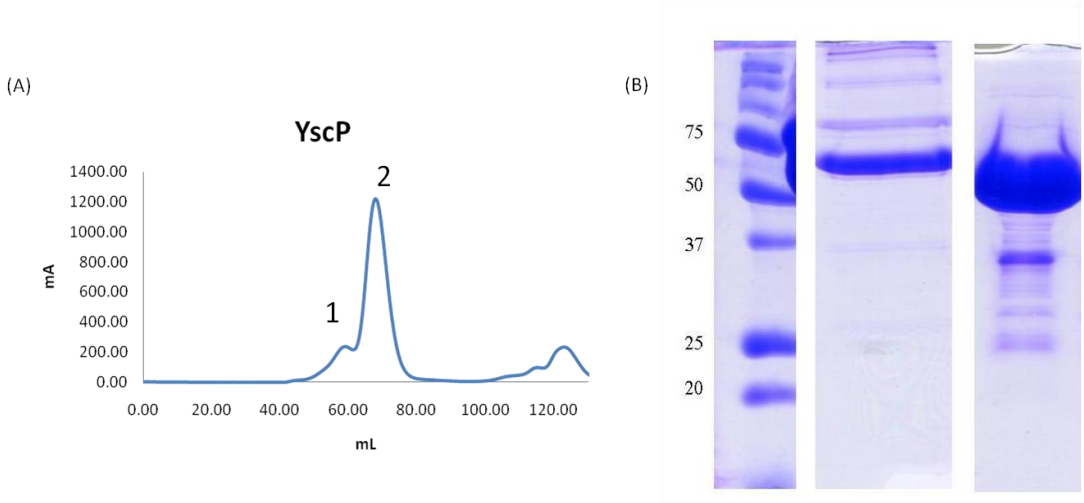

Figure 2.11: Gel filtration analysis of purified YscP. (A) Gel filtration profile of YscP. YscP runs at sizes corresponding to molecular masses of 360 and 141 kDa. (B) The two major peaks from gel filtration, were visualized by SDS-PAGE. Lanes correspond to the protein marker, peak 1, and peak 2.

YscP is predicted to contain large regions lacking secondary structure (Figure 2.2). Purification of YscP yielded degradation products. An anti-histidine western blot against YscP indicated that these products are specific to histidine tagged YscP (Figure 2.12). To generate a version of YscP more likely to crystallize, smaller constructs were created to delete two of the large regions predicted to contain no regular secondary structure: Δ46 (46-96) and Δ46/Δ222 (46-96, 222-306). These regions are analogous to regions in *Y. enterocolitica*, which when deleted, resulted in a functional T3SS with shorter needles (Journet, 2003). These constructs contain C-terminal histidine tags for purification. The deletion constructs, Δ46 and Δ46/Δ222, were expressed and purified by nickel affinity followed by gel filtration (Figures 2.13 and 2.14). The elution profile for Δ46 showed three species corresponding to the void volume, 306 kDa, and 74 kDa, respectively. The predicted molecular weight of Δ46 is

44.9 kDa, again suggesting elongation or multimer formation. The gel filtration profile for Δ46/Δ222 had two peaks with molecular masses predicted to be 60 kDa and 30 kDa, respectively, with the calculated molecular mass being 35.7 kDa. This YscP truncation fragment may be more globular and the two fractions may represent dimeric and monomeric species of the protein.

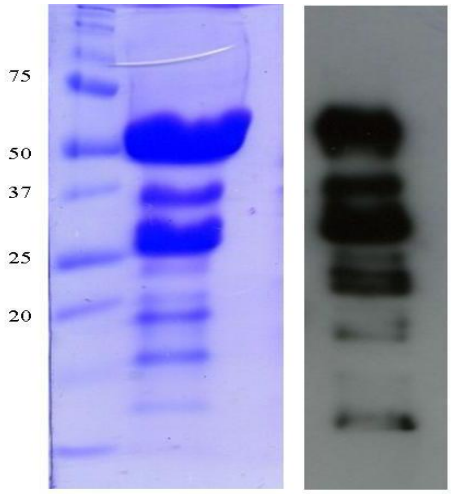

Figure 2.12: YscP is susceptible to proteolysis. SDS-PAGE showing YscP after purification (left). With time, YscP degrades into smaller fragments as visualized by western blot against the C-terminal histidine tag (right).

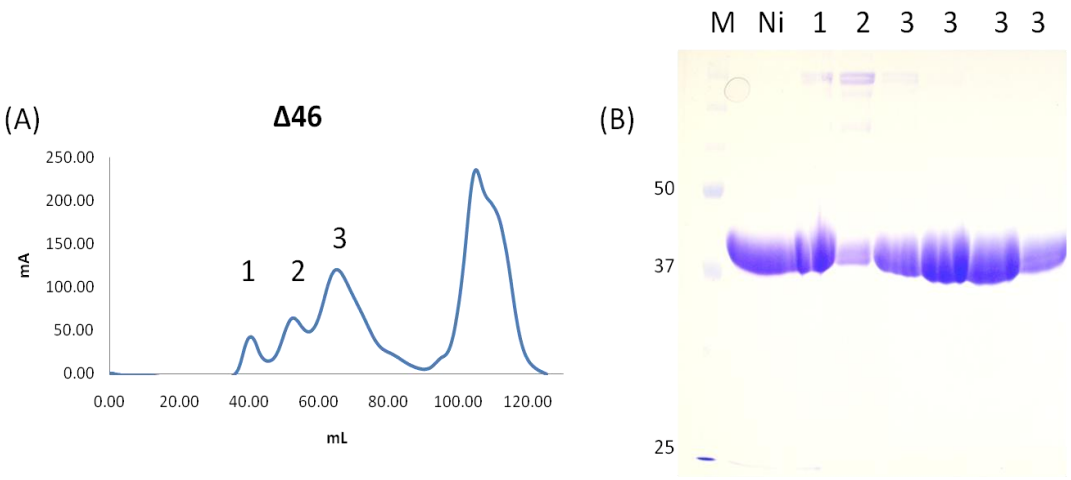

Figure 2.13: Purification of Δ46 from *E. coli*. (A) Gel filtration chromatogram of Δ46. (B) SDS-PAGE visualization of Δ46 purified from *E. coli* using Ni-NTA affinity and gel filtration chromatography. Lanes correspond to protein marker (M), Ni-NTA purification (Ni), and peak fractions from the gel filtration column.

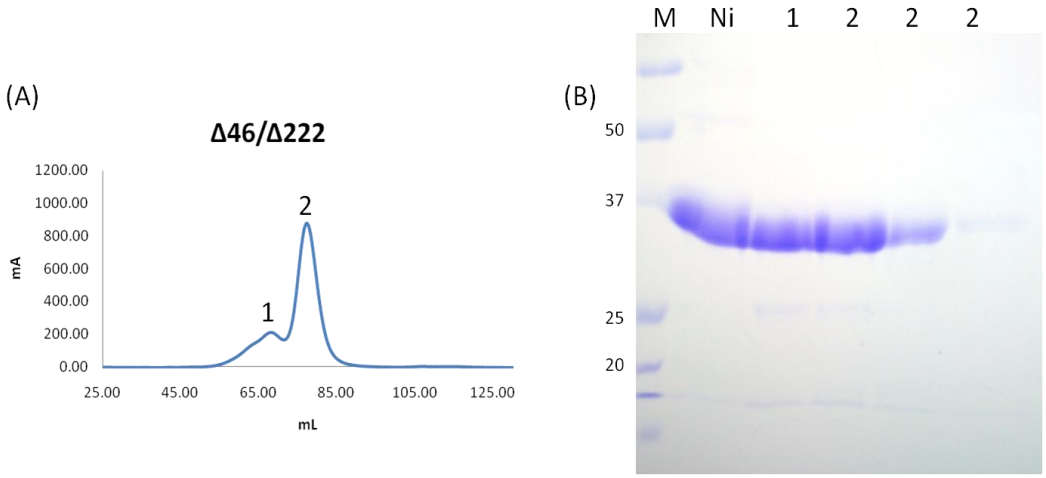

Figure 2.14: Purification of Δ46/Δ222 from *E. coli*. (A) Gel filtration chromatogram of Δ46/Δ222. (B) SDS-PAGE visualization of Δ46/Δ222 purified from *E. coli* using Ni-NTA affinity and gel filtration chromatography. Lanes correspond to protein marker (M), Ni-NTA purification (Ni), and peak fractions from the gel filtration column.

In addition to deletions in the ruler region, a smaller construct of YscP was constructed to only contain the T3S4 region. T3S4 was also purified by nickel affinity and gel filtration chromatography. T3S4 eluted as an elongated monomer or dimer at 19 kDa, as compared to the calculated molecular mass of 11 kDa (Figure 2.15).

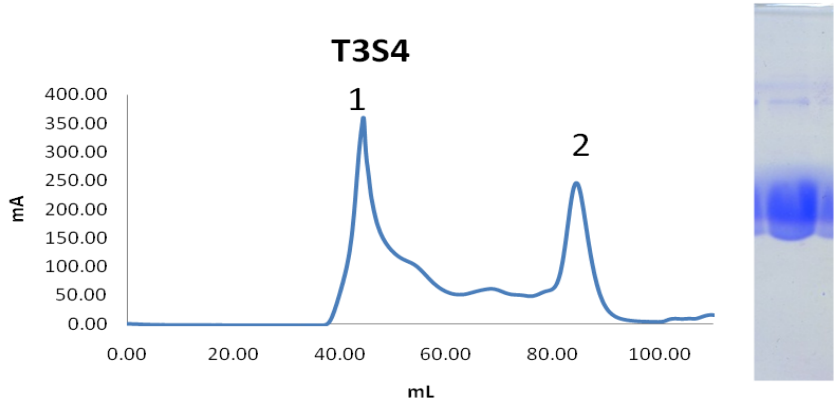

Figure 2.15: Purification of YscP T3S4. Gel filtration profile of T3S4 reveals two major peaks, one in the void volume and one corresponding to soluble T3S4. T3S4 from peak 2 is shown by SDS-PAGE (right).

Purification of YscO

YscO was expressed from pET28b in *E. coli* and refolded from inclusion bodies (Figure 2.16). After refolding, YscO was maintained at 4°C in storage buffer (300 mM NaPi, pH 8.5, 300 mM NaCl, 5 mM βME) but precipitated within days. This precipitation increased at protein concentrations above 2 mg/mL. YscO was stable to freezing and thawing, and purified YscO was prepared to be stored in this manner. If the storage buffer contained less than 300 mM NaCl, precipitation occurred rapidly. The storage buffer was maintained at pH 8.5; when the buffer was at pH 8.0, most of YscO precipitated. A rationale for why the higher pH of the buffer was crucial for protein stability is that the pI of YscO is calculated to be 7.75.

Additional buffers with higher pH values were not tested for solubility improvements. YscO also contains cysteines and requires a reducing agent. For future work with YscO it is suggested that a suitable buffer be determined to enhance stability. Additionally, some mutants or deletion constructs could be made based on the structures of FliJ and CT670.

YscO refolded as a dimer, as observed by gel filtration chromatography. By comparison with globular protein standards, YscO migrated with a molecular weight of 44 kDa (Figure 2.17). This dimerization is in agreement with the dimer formed by the homologous protein CT670. However, this dimerization does not occur with the flagellar T3SS protein FliJ (Lorenzini, 2010).

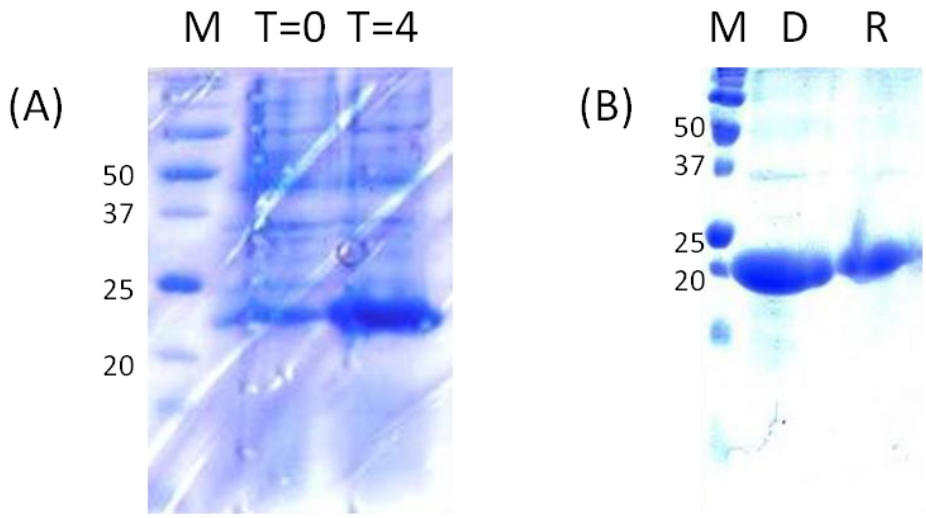

Figure 2.16: Expression, purification, and refolding of YscO from *E. coli*. (A) SDS-PAGE visualization of YscO expressed as inclusion bodies in *E. coli*. Lanes correspond to protein marker (M), total bacterial proteins prior to IPTG induction (T=0), and 4 hours after induction (T=4). (B) SDS-PAGE visualization of YscO refolding and purification. Lanes correspond to protein marker (M), denatured inclusion bodies (D), and refolded proteins after Ni^{2+}-NTA purification (R).

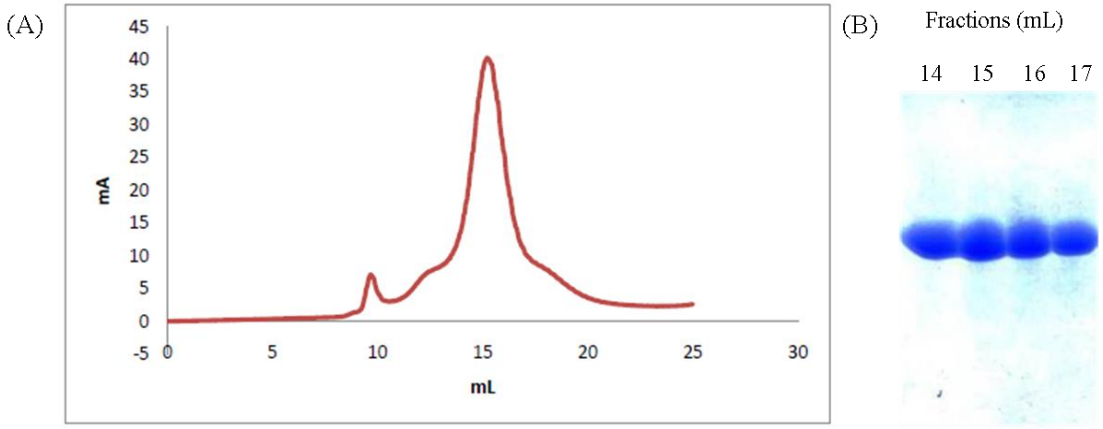

Figure 2.17: Gel filtration analysis of refolded YscO. (A) Gel filtration profile of refolded YscO. YscO runs at a molecular weight of 44 kDa, corresponding to a dimer. (B) The major peak fractions from gel filtration were visualized by SDS-PAGE.

YscO Binding Assay

To identify proteins that bind to YscO, refolded histidine-tagged YscO was bound to nickel beads and used as bait against *Y. pseudotuberculosis* lysates. To control for proteins that non-specifically bind to nickel beads, lysates were also loaded onto nickel beads lacking YscO. Many proteins were found to bind to both YscO and the nickel beads by themselves, with no significant differences were noticeable by Coomassie-stained SDS-PAGE gel (Figure 2.18). Interactions with YscO may be transient, as predicted by the shuttling of chaperone cargo in FliJ, so these interactions may not be captured via a coprecipitation assay. Cross-linking could address this issue. Other possible limiations to this experiment include the high background binding of lysates to the nickel beads, the limited sensitivity of a Coomassie-stained

gel, and the presence of the histidine tag which may have compromised the availability of YscO binding sites.

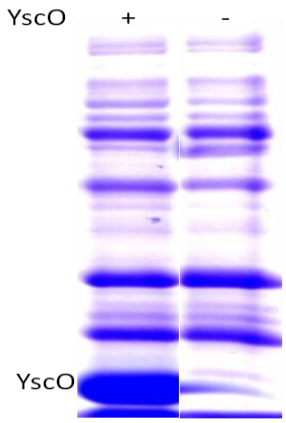

Figure 2.18: YscO and *Yersinia* lysate coprecipitation. Lane 1 corresponds to the bound fraction of YscO incubated with bacterial lysate (boiled off of nickel beads). Lane 2 corresponds to bound fraction of bacterial lysates, in the absence of YscO, incubated with nickel beads (boiled off nickel beads).

YscP Binding Assays

YscP was subjected to lysate coprecipitations in the same manner as YscO. A significant amount of lysate proteins bound to nickel beads alone (Figure 2.19A). A possible reason for this is that the *Y. pseudotuberculosis* proteins are not stable in the binding buffer utilized, and therefore precipitated but were solubilized by SDS during sample preparation. The coprecipitation assay was repeated with the addition of 2% glycerol and 0.1% Triton X-100 in the binding and wash buffers. This decreased the amount of background interaction with the nickel beads (Figure 2.19B). To further minimize non-specific interactions, the beads were pre-incubated with BSA before the addition of lysate. This technique was combined with binding and washing buffers

containing glycerol and Triton X-100. Lysate proteins still bound non-specifically to the nickel beads, but not as much as in initial experiments (Figure 2.19). Further optimization of these coprecipitation assays were not conducted because many of the proteins visualized on the SDS-PAGE gel were identified to be degradation products of YscP. This degradation of histidine-tagged YscP was visualized as multiple fragments on an anti-histidine western blot (Figure 2.12). This degradation made it difficult to determine which proteins on the SDS-PAGE gel were YscP degradation products and which were putative binding partners.

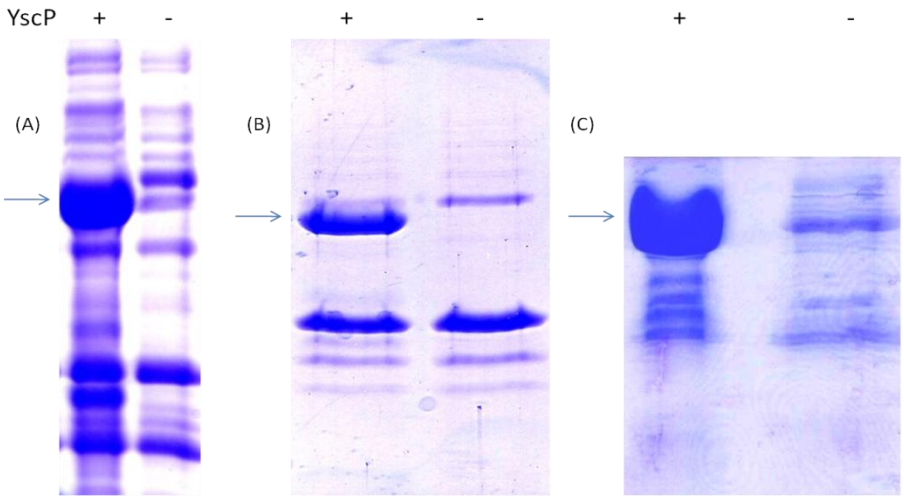

Figure 2.19: YscP and *Yersinia* lysate coprecipitation. YscP was bound to nickel beads. *Yersinia* lysate was applied to both YscP-bound nickel beads and nickel beads lacking YscP. (A) Bound fraction in an experiment in which unbound proteins were removed by washing in binding buffer. (B) Bound fraction in an experiment in which unbound proteins were removed by washing in binding buffer, 2% glycerol, and 0.5% Triton X-100. (C) Bound fraction in an experiment in which the nickel beads were preincubated with bovine serum albumin (BSA) before addition of lysates. After removal of BSA, the nickel beads were incubated with lysate proteins. Washes were done with buffer used for experiment B. Arrows indicates YscP.

YscP Crystallization Trials

Purified YscP, Δ46, Δ46/Δ222, and T3S4 were subjected to crystallization trials by the vapor diffusion method. Protein concentrations varied between 5 and 20 mg/mL. Crystallization trials were conducted robotically (Oryx) and manually; no protein crystals were observed.

YscP is predicted to have many potentially flexible residues, which can hinder crystallization. To immobilize some of these side chains, surface exposed lysines were methylated. This procedure has proven to enhance the crystallizability of several proteins and thus was employed for the surface-exposed lysines of YscP (there are 20 total lysines in YscP) (Walter, 2006). Each YscP construct was reductively methylated, purified by size exclusion chromatography as before, and crystallization trials were repeated. No protein crystals resulted from these attempts, however the solubility of YscP did increase. Additional regions which could be altered include sites which were identified by the Surface Entropy Reduction prediction (SERp) server (Goldschmidt, 2007). These sites were not subject to mutation but are suitable candidates for additional YscP constructs.

Discussion

The conformation of YscP prior to secretion and during its subsequent role in needle elongation is unknown. Bioinformatic analysis suggested that a great portion of amino acid residues comprising the N-terminus of YscP is devoid of regular secondary structure. This region may be stabilized before secretion and then unfolded to function inside the polymerizing needle for length control. Additionally, it is unknown how the C-terminus of YscP is involved in the switching of substrate export from structural to effector proteins. YscP was purified and subjected to crystallographic trials to provide knowledge on how YscP functions as a molecular ruler.

None of the constructs of YscP resulted in protein crystals. Full length YscP is not homogeneous and is subject to degradation over time. Construction of the truncation fragments Δ46 and Δ46/Δ222 was carried out to see if removing flexible regions could help deter degradation. Both constructs were indeed more homogeneous, but neither yielded crystals. Reductive methylation of lysines was helpful in improving solubility but did not result in crystallization.

A promising region of YscP for crystallization is the T3S4. Although the T3S4 truncation fragment did not yield crystals, information from the FliK structure and YscP secondary structure prediction reveal that the most C-terminal amino acids of YscP are likely to be unstructured. These C-terminal residues could be deleted from the T3S4 construct to improve globularity and crystal packing.

YscO was purified from inclusion bodies and refolded. YscO refolded into a dimer as determined by gel filtration. This dimerization has been noted before in the YscO homologs Cpn0706 and HrpO (Stone, 2008, Gazi, 2008), which contrasts with the monomeric state of FliJ.

YscO is predicted to be α-helical coiled coil, a feature common in the T3SS. The YscO homologs CT670, HrpO, and FliJ have coiled-coil characteristics as well (Gazi, 2008). The ATPase γ-subunit of FliI forms an α-helical coiled coil (Ibuki, 2011). The YscN ATPase regulator YscL is also predicted to have a coiled-coil region. Coiled coils are predicted in the translocation regulator protein LcrV, LcrG, the needle protein YscF, and the translocon proteins YopB and YopD (Derewenda, 2004, Hamad, 2007, Sun, 2008).

In addition to crystallization attempts, YscO and YscP were studied in binding assays with *Y. pseudotuberculosis* proteins. Proteins which specifically bound to either YscO or YscP were not identififed by SDS-PAGE visualization when compared to control experiments. There may certainly be *Y. pseudotuberculosis* proteins which bound to these bait proteins, but the interactions may be weak, transient, or the protein levels were too low to be visualized by coomassie. In fact, both YscO and YscP have been shown in *Yersinia* to associate with many T3SS proteins. Whether those interactions are direct or indirect however has not been addressed. Future experiments with YscO and YscP must take in consideration that these proteins are likely parts of a larger macromolecular complex, possibly including YscN, YscL, and YscU.

References

Abuladze, N. K., M. Gingery, J. Tsai, and F. A. Eiserling, (1994) Tail Length Determination in Bactiophage T4. *Virology* **199**: 301-310.

Agrain, C., I. Callebaut, L. Journet, I. Sorg, C. Paroz, L. J. Mota, and G. R. Cornelis, (2005) Characterization of a type III secretion substrate specificity switch (T3S4) domain in YscP from Yersinia enterocolitica. *Molecular Microbiology* **56**: 54-67.

Benkert, P., Tosatto, S.C.E. and Schomburg, D., (2008) QMEAN: A comprehensive scoring function for model quality assessment. *Proteins: Structure, Function, and Bioinformatics* **71**: 261-277.

Cornelis, G. R., C. Agrain, and I. Sorg, (2006) Length control of extended protein structures in bacteria and bacteriophages. *Current Opinion in Microbiology* **9**: 201-206.

Corpet, F., (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research* **16**: 10881-10890.

Derewenda, U., A. Mateja, Y. Devedjiev, K. M. Routzahn, A. G. Evdokimov, Z. S. Derewenda and D. S. Waugh, (2004) The Structure of Yersinia pestis V-Antigen, an Essential Virulence Factor and Mediator of Immunity against Plague. *Structure* **12**: 301-306.

Edqvist, P. J., J. Olsson, M. Lavander, L. Sundberg, A. Forsberg, H. Wolf-Watz, and S. A. Lloyd, (2003) YscP and YscU regulate substrate specificity of the Yersinia type III secretion system. *Journal of Bacteriology* **185**: 2259-2266.

Evans, L. D., G. P. Stafford, S. Ahmed, G. M. Fraser, and C. Hughes, (2006) An escort mechanism for cycling of export chaperones during flagellum assembly. *Proceedings of the National Academy of Sciences* **103**: 17474-17479.

Evans, L. D. a. C. H., (2009) Selective binding of virulence type III export chaperones by FliJ escort orthologues InvI and YscO. *FEMS Microbiology Letters* **293**: 292-297.

Gazi, A. D., M. Bastaki, S. N. Charova, E. A. Gkougkoulia, E. A. Kapellios, N. J. Panopoulos and M. Kokkinidis, (2008) Evidence for a Coiled-coil Interaction Mode of Disordered Proteins from Bacterial Type III Secretion Systems. *Journal of Biological Chemistry* **283**: 34062-34068.

Goldschmidt, L., D. Cooper, Z. Derewenda, and D. Eisenberg, (2007) Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. *Protein Science* **16**: 1569-1576.

Gouet, P., X. Robert & E. Courcelle, (2003) ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Research* **31**: 3320-3323.

Hamad, M. A. a. M. L. N., (2007) Structure-Function Analysis of the C-Terminal Domain of LcrV from Yersinia pestis. *Journal of Bacteriology* **189**: 6734-6739.

Holm, L. a. P. R., (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Research* **38**: W545-549.

Ibuki, T., K. Imada, T. Minamino, T. Kato, T. Miyata and K. Namba, (2011) Common architecture of the flagellar type III protein export apparatus and F- and V-type ATPases. *Nature Structural & Molecular Biology* **18**: 277–282.

Journet, L., C. Agrain, P. Broz and G. R. Cornelis, (2003) The needle length of bacterial injectisomes is determined by a molecular ruler. *Science* **302**: 1757-1760.

Kelley, L. A. a. M. J. E. S., (2009) Protein structure prediction on the web: a case study using the Phyre server. *Nature Protocols* **4**: 363-371.

Lambert, C., N. Léonard, X. De Bolle and E. Depiereux, (2002) ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics* **18**: 1250-1256.

Lorenzini, E., A. Singer, B. Singh, R. Lam, T. Skarina, N. Y. Chirgadze, A. Savchenko, and R. S. Gupta, (2010) Structure and Protein-Protein Interaction Studies on Chlamydia trachomatis Protein CT670 (YscO Homolog). *Journal of Bacteriology* **192**: 2746-2756.

Moriya, N., T. Minamino, K. T. Hughes, R. M. Macnab, and K. Namba, (2006) The Type III Flagellar Export Specificity Switch is Dependent on FliK Ruler and a Molecular Clock. *Journal of Molecular Biology* **359**: 466-477.

Mota, L. J., L. Journet, I. Sorg, C. Agrain and G. R. Cornelis, (2005) Bacterial Injectisomes: Needle Length Does Matter. *Science* **307**: 1278.

Payne, P. L. & S. C. Straley, (1999) YscP of Yersinia pestis Is a Secreted Component of the Yop Secretion System. *Journal of Bacteriology* **181**: 2852-2862.

Payne, P. L. a. S. C. S., (1998) YscO of Yersinia pestis Is a Mobile Core Component of the Yop Secretion System. *Journal of Bacteriology* **180**: 3882-3890.

Riordan, K. E., J. A. Sorg, B. J. Berube, and O. Schneewind, (2008) Impassable YscP Substrates and Their Impact on the Yersinia enterocolitica Type III Secretion Pathway. *Journal of Bacteriology* **190**: 6204-6216.

Rost, B., G. Yachdav, and J. Liu, (2004) The PredictProtein server. *Nucleic Acids Research* **32**: W321-326.

Sorg, I., S. Wagner, M. Amstutz, S. A. Müller, P. Broz, Y. Lussi, A. Engel, and G. R. Cornelis, (2007) YscU recognizes translocators as export substrates of the Yersinia injectisome. *The EMBO Journal* **26**: 3015-3024.

Stone, C. B., D. L. Johnson, D. C. Bulir, J. D. Gilchrist, and J. B. Mahony, (2008) Characterization of the Putative Type III Secretion ATPase CdsN (Cpn0707) of Chlamydophila pneumonia. *Journal of Bacteriology* **190**: 6580-6588.

Sukhan, A., T. Kubori, J. Wilson, and J. E. Galán, (2001) Genetic Analysis of Assembly of the Salmonella enterica Serovar Typhimurium Type III Secretion-Associated Needle Complex. *Journal of Bacteriology* **183**: 1159-1167.

Sun, P., B. P. Austin, J. E. Tropea, and D. S. Waugh, (2008) Structural characterization of the Yersinia pestis type III secretion system needle protein YscF in complex with its heterodimeric chaperone YscE/YscG. *Journal of Molecular Biology* **377**: 819-830.

Walter, T. S., C. Meier, R. Assenberg, K.-F. Au, J. Ren, A. Verma, J. E. Nettleship, R. J. Owens, D. I. Stuart and J. M. Grimes, (2006) Lysine Methylation as a Routine Rescue Strategy for Protein Crystallization. *Structure* **14**: 1617-1622.