universität wien

D I P L O M A R B E I T

Titel der wissenschaftlichen Arbeit

Structure and Function of the Type 3 Secretion System in *Salmonella typhimurium*

Verfasser

Wolfgang H. Schmied

zur Erlangung des akademischen Grades

Magister der Naturwissenschaften
(Mag. rer. nat.)

Matrikelnummer: 0201599
Studienrichtung: Molekulare Biologie (A490)
Betreuer: Dr. Thomas Marlovits

Wien, im März 2010

ACKNOWLEDGEMENTS

Even though molecular machines tend to arrive fully assembled, there is hardly ever a manual to be found. Gathering the clues, closely looking at details and finally writing a few new lines into the elusive user guide to the type III secretion system has been both an exciting research project as well as a personal challenge.

I can only offer sincere thanks to Thomas Marlovits for his supervision, thoughts and support to use his lab as a scientific playing ground. Between his goodwill towards often improbable pet projects and his well-placed criticism, there was plenty to learn. More often than not, those seemingly unlikely ideas were cooked up over a shared pot of tea with Oliver Schraidt. The discussions and the exhilarating adventures indoors and out on the mountains and lakes will be something to look back upon, and helped to keep everything in perspective. On the same note, I want to thank the whole lab – Wolfgang Weber, Julia Radics, Matthias Brunner, Lisa Königsmaier, Jesús Fernandez Rodriguez and Agata Kosarewicz – for whether it was to share a croissant in Paris or a long evening in the lab, it was always great to have you around.

Since these are the last lines to be written in this thesis, my studies in Vienna are soon going be to be formally concluded. The final words of thanks therefore go to my close friends, for everything, and my family, for their continued support and trust.

Wien, 10.3.2010 Wolfgang H. Schmied

CONTENTS

ALLGEMEINE ZUSAMMENFASSUNG

Der Typ 3 Sekretionsapparat ist eine molekulare Maschine, die in zahlreichen bakteriellen Krankheitserregern nachgewiesen wurde. Es handelt sich um einen Proteinkomplex in der bakteriellen Membran, der in einem gerichteten Prozess Toxine vom Bakterium in das Zytoplasma einer Wirtszelle transportiert. Er ermöglicht dadurch zahlreiche Strategien, welche die Infektion des Wirtes und komplexe Überlebensstrategien des Bakteriums erlauben. Die Erreger von bedeutenden Krankheiten wie Pest, Typhus oder Bakterienruhr bedienen sich dieses Mechanismus.

Das bisher am besten untersuchte Merkmal des Sekretionsapparates ist der so genannte Nadelkomplex. Er ähnelt äußerlich und funktionell einer Spritze, und erstreckt sich mit einer in der inneren und äußeren Membran verankerten Basis, die das Periplasma quert, und einer extrazellulären Nadelstruktur über etwa 80 nm. Der maximale Durchmesser der unregelmäßig zylindrischen Struktur beträgt etwa 21 nm. Der gesamte Sekretionsapparat besteht aus mehr als zwanzig Proteinen, bisher konnten fünf strukturelle Proteine im Nadelkomplex nachgewiesen und teilweise lokalisiert werden.

Der Aufbau des Komplexes erfolgt schrittweise und ist von der Sekretion spezifischer struktureller Proteine in aufeinander folgenden Stadien geprägt. Dies wird durch einen Wechsel im Erkennungsmuster der sekretierten Proteine erreicht. Nach Abschluss des Aufbaus wiederholt sich dieses Phänomen in Form der gestaffelten Sekretion von Effektorproteinen, welche im fortschreitenden Krankheitsverlauf unterschiedliche Funktionen erfüllen.

In dieser Studie wurde das Protein PrgH, ein Bestandteil des Nadelkomplexes im Humanpathogen *Salmonella typhimurium,* untersucht. Anfängliche elektronen-mikroskopische Untersuchungen zielten darauf ab, das Protein mittels exakt lokalisierter Goldpartikel im Komplex zu markieren, im weiteren Verlauf wurde die Rolle von PrgH im Sekretionsprozess näher analysiert. Mittels genetischer Methoden wurde eine Bibliothek von Mutanten generiert, deren Phänotyp hinsichtlich erfolgreichem Aufbaus des Nadelkomplexes und dessen Sekretionsaktivität biochemisch untersucht wurde.

Die erhalten Daten zeigen, dass PrgH den äußeren Bereich des inneren Ringes des Nadelkomplexes aufbaut, und seine N-terminale Domäne in das Zytoplasma ausgerichtet ist. Durch die Untersuchung der Mutanten wurde mehrere Bereiche von PrgH identifiziert, welche eine kritische Rolle in der Sekretion von Effektorproteinen spielen.

ABSTRACT

Type 3 secretion systems are conserved molecular machines commonly found in Gram-negative bacterial pathogens, such as *Salmonella, Shigella, Yersinia or* enteropathogenic *Escherichia coli* (EPEC), and play essential roles in the development of the diseases typhoid fever, dysentery or bubonic plaque, among others. In these organisms, at least one pathogenicity island encodes this highly evolved secretion apparatus, which allows bacteria to modify host cell behavior in order to allow efficient infection and long-term survival.

The large core structure of the type 3 secretion system is commonly referred to as the needle complex. It consists of at least five proteins with varying copy numbers, assembling into a syringe-shaped configuration, which allows directed transport of proteins from the bacterial cytoplasm to a host cell. Assembly of the complex happens stepwise, with specific structural proteins being secreted per stage of the assembly process. Similarly, effector proteins are secreted in a hierarchical fashion tailored to their role in the pathogenicity process.

Despite the importance of type 3 secretion systems, the topology of the needle complex and the structural basis for substrate recognition is poorly understood. Here, *Salmonella typhimurium* was used as a model organism to study the localization of PrgH and its relevance to the assembly and secretion process.

Using electron microscopy and site-specific labeling methods, PrgH was found to compose the outer areas of the inner ring, with the N-terminal domain located towards the cytoplasm. By screening a library containing mutated variants of PrgH using two novel ELISAs to assess successful assembly and secretion activity, the N-terminal domain was found to contain areas critical for the secretion of the effector protein SptP. Interestingly, secretion of the pore-forming proteins SipB and SipC was less affected, suggesting that PrgH is involved in the substrate switching process between pore-forming proteins and effector proteins, or early and late effector proteins.

INTRODUCTION

The co-evolution of pathogenic bacteria and their hosts has led to a number of remarkable adaptations. Several species of Gram-negative bacteria [1] use a complex molecular machine to induce changes in the behavior of eukaryotic host cells to aid their own survival and proliferation. This so-called type three secretion system consists of a syringe-like structure in the bacterial membrane (needle complex, injectisome), a cytoplasmic bulb (c-ring) and a number of effector proteins. By creating a secure transport pathway from the bacterial cytoplasm through two bacterial membranes into the cytoplasm of a neighboring host cell, they open a directed communication channel, through which effector proteins are transported in an orchestrated manner [2]. The courses of diseases caused by human or plant pathogens such as *Salmonella*, *Yersinia*, *Chlamydia*, *Pseudomonas*, or *Erwinia* are thereby critically influenced by this pathogenicity factor.

A model pathogen: Salmonella enterica

Salmonella enterica strains are enteric pathogens causing a variety of diseases, ranging in severity from typhoid fever to gastroenteritis. Both serovar *typhi* (*S. typhi*) and serovar *typhimurium* (*S. typhimurium*) contain several pathogenicity islands, although only *S. typhi* can cause systemic infections in human. Infection with *S. typhimurium* causes a usually relatively benign gastroenteritis in humans. In contrast to *S. typhi* it can also infect a wide range of animals, including mice, where it causes systemic infection similar to typhoid fever in humans.

Even before the complete sequence of *Salmonella enterica* serovar *Typhimurium LT2* had been available [3], five pathogenicity-associated loci had been identified using genetic techniques and screening methods [4]. They have been termed salmonella pathogenicity islands (SPIs), with a sixth added more recently [5]. SPI-1 and SPI-2 each encode a fully functional type three secretion system. SPI-1 stretches over roughly 40kb and contains at least 29 genes, used for the regulation and expression of proteins constituting a functional type three secretion system. SPI-1 and 2 are structurally and functionally distinct, and show homology to different archetypes [6]. SPI-1 is involved in the invasion of endothelial cells of the small intestine, by injecting effector proteins through the needle complex into the

neighboring host cell, thereby causing disturbances in the actin cytoskeleton. SPI-2 genes are quickly upregulated after bacterial entry into endothelial cells or macrophages and stay active during the course of infection. Its function has been linked to extended survival within macrophages in mouse models. While SPI-1 was described as necessary for bacterial invasion in 1989 [7, 8], SPI-2 was only found and characterized six years later [9-11].

General structure of the SPI-1 injectisome

Since the first visualization of a needle complex on the surface of salmonella cells [12], reconstructions of the three dimensional structure of the basal body or the entire needle complex [13-17] were obtained from *S. typhimurium*, *Shigella flexneri* and enteropathogenic *Escherichia coli*. The overall structure appears to be conserved [6], and consists of a generally cylindrical complex with distinct inner and outer rings (see Figure 1) and a maximum diameter of approximately 21 nm [15]. In S. *typhimurium*, the major structural proteins are InvG in the outer ring, belonging to the secretin family, with PrgH and PrgK in the inner ring.

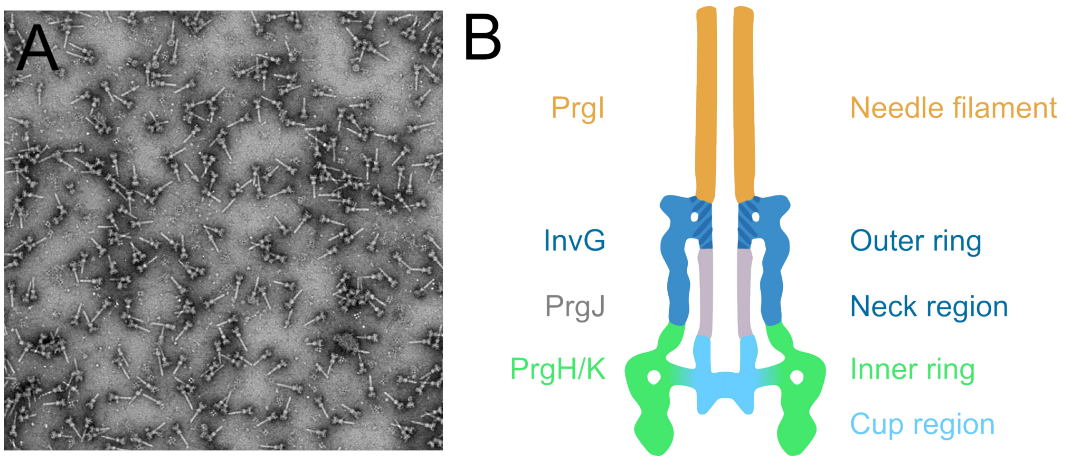


Figure 1: (A) Purified needle complexes (B) Map of the needle complex

The radial symmetry is still under debate, with proposed values ranging from 12-fold upwards, limiting the usefulness of proposed topology models [13, 14, 18]. The needle filament is extending outwards from the basal body, a helical structure polymerizing from PrgI subunits [19]. Within the basal structure, a so-called inner

rod, a cup and a socket region have been identified. Strains lacking InvJ are deficient for the socket region, while PrgJ has been suggested to form the inner rod [14]. Additionally, strains lacking any of the proteins OrgA, OrgB or several Spa-proteins [20] successfully form the basal body, but fail to assemble the needle filament. As of now, the functional role of most of these proteins is unclear, although SpaP, Q, R and SpaS are expected to localize within the inner ring, based on homologies to the flagellar system [21].

The ATPase InvC is said to energize the secretion process and might connect to the basal body through homologs of the c-ring proteins [6]. It could already be shown that InvC is required to release effector proteins from their cognate chaperones [22, 23]. In this unfolded state, the proteins would then be able to pass through the central pore, with an estimated width of 2.5 nm.

Assembly of the injectisome

Similar to the flagellar system, assembly of the needle complex is considered a strictly linear process [6]. Membrane insertions of PrgK and InvG are dependent on the general secretory pathway (Sec), while other components of the needle complex are secreted by themselves in a highly regulated fashion, suggesting the existence of several checkpoints and changes in the specificity for secretion substrates. As proposed here [2], two switching events occur during the assembly process. The first would occur once the needle filament reaches its full length, the second one on cell contact, if the translocators are in place (Figure 2).

InfJ, PrgJ and PrgI are usually considered early substrates, are secreted through the basal body and polymerize to form the socket region, the inner rod and the needle filament, respectively. In *S. typhimurium*, needle length control was described to depend on the relative amounts of PrgI and PrgJ [14]. C-ring related proteins and InvC-deficient strains are unable to form the needle filament [20], suggesting recruitment before secretion of the first substrates takes place.

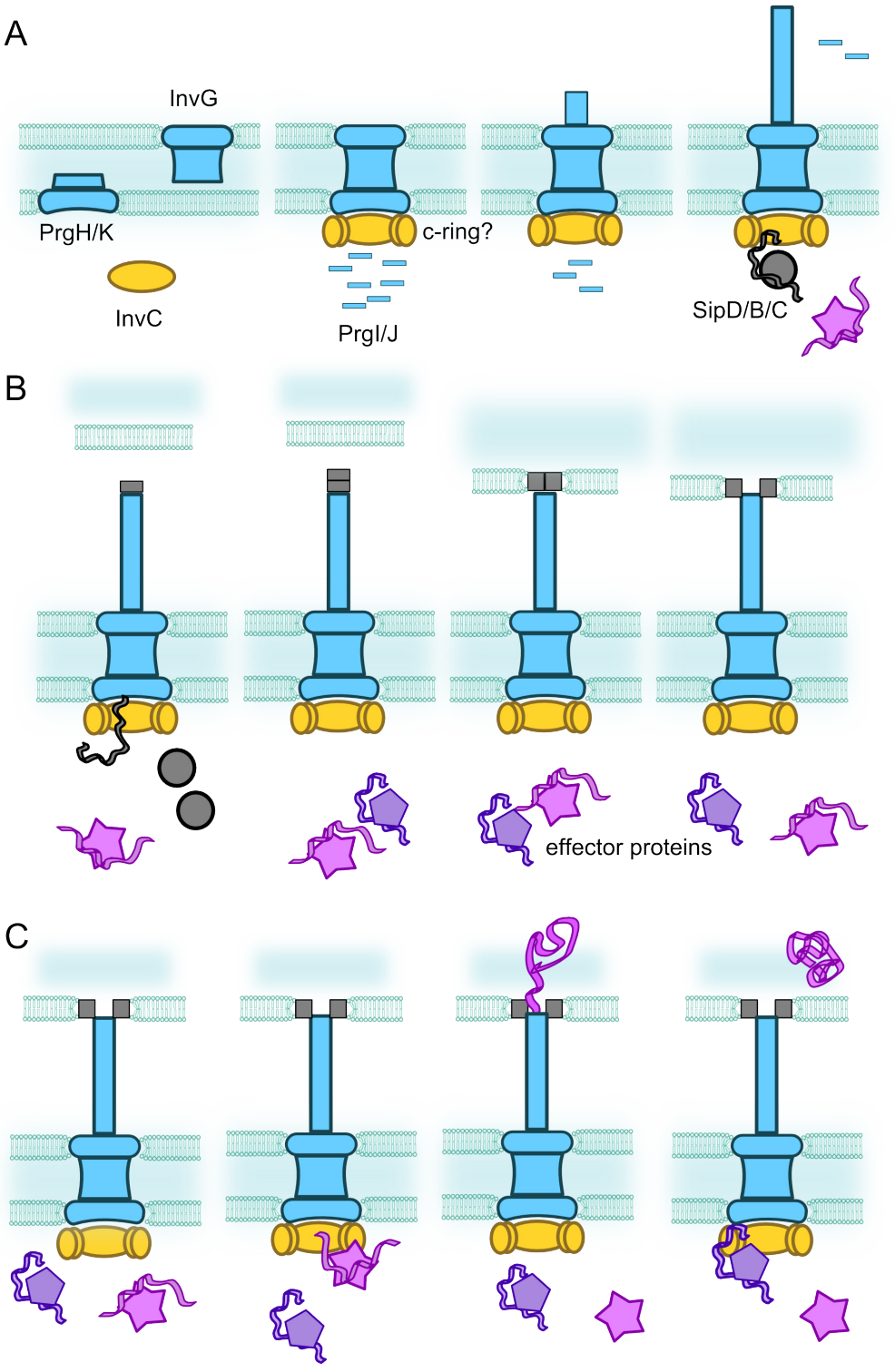


Figure 2: Illustration showing assembly of the needle complex, with substrate switching after PrgI/J and SipD/B/C. (A) Assembly of the needle complex (B) Secretion of translocators and assembly of tip complex/pore (C) Secretion of effector proteins [6, 19, 24]

Once the needle filament is assembled, the middle substrates, translocator proteins SipB, SipC and SipD are targeted to the complex, creating a secretion-blocked "stand-by" situation in Yersinia and Shigella [2]. Once a host cell is detected, possibly by the tip structure, a pore can be formed in the host cell membrane and protein translocation [2, 25] can occur.

Effector proteins of SPI-1

While S. typhimurium shows secretion of effector proteins to the culture supernatant without cell contact, the start of secretion is highly regulated in other organisms [6]. SipA, SopB, SopE, SipE as well as AvrA and SptP among other effector proteins are transported through the needle complex. Using fluorescence microscopy, the process could be visualized and appeared to take place within less than 10 minutes after cell contact [26].

Their combined effect is a drastic, although temporary, reshaping of the host cell membrane (ruffling) at the contact point between endothelial cell and bacterium, followed by internalization of the bacterium. While early effector proteins induce membrane ruffling and promote bacterial engulfment, late effector proteins reverse these effects and allow the cell to regain its original shape. A distinct hierarchy has to exist in the secretion of effectors through the type three secretion system of *S. typhimurium*, to avoid the presence of effector proteins with antagonizing effects in the host cell at the same time [27, 28].

Research questions

In order to better understand the role of proteins suspected to lie at the interface between the needle complex and cytoplasmic components of the type three secretion system, different lines of research were followed to answer the following questions:

What is the exact localization and organization of PrgH and PrgK, and can they be directly visualized using site-specific labels?

Is PrgH purely a scaffolding component of the needle complex, or is it directly interacting with cytoplasmic or membrane associated proteins beyond the core complex?

Site specific labeling

Mapping the subunits of an intricate protein assembly as the needle complex is challenging. While the structure can by now be routinely imaged at low resolution, the identification and localization of distinct subunits would require ideally near-atomic resolution. While this goal can in principle be achieved based on electron microscopy data alone [29], no structure of the type three secretion system with a sufficiently high resolution has so far been published.

By using x-ray crystallography, protein structures are routinely solved at atomic resolutions, although the process becomes more complex with increasing size of the target. While large assemblies over 1 megadalton (MDa) in size are now in reach of the technique [30], smaller structures are more readily available.
The strengths of both techniques can be combined when high resolution structural data of subunits is used to reconstruct the overall topology within a global, low resolution map. This modeling process depends on the availability of restraints in order to position individual subunits with a high degree of confidence. Restraints are derived from multiple sources, including molecular modeling, information on the copy number or quaternary structures of subunits or anchor points created by identifying the subunit on the overall structure.

Since atomic structures of several core proteins, or domains of core proteins, became available, a number of incompatible models for the topology of the needle complex were proposed [13, 18]. In this study, polyhistidine tags in multiple positions on the protein complex were used as site-specific binding partners for 1.4 nm small Ni-NTA gold particles. When a gold particle is bound to a needle complex, it can be recognized as an additional electron density on electron micrographs. Particles of this size offer an acceptable compromise between ease of visualization, and accuracy of the localization.

Screening for functional domains

PrgH (392 amino acids) is a predicted transmembrane protein located in the inner ring of the needle complex, with at least two separate domains in the periplasmic and the cytoplasmic space. It is essential for and recruited early [20, 31] in the assembly process of the needle complex, suggesting a potential scaffolding role for later assembly steps. Additionally, the cytoplasmic domain could represent an interaction site for secretion substrates.

No information on the role of PrgH in either the assembly process, or the secretion process was available so far. The experimental setup therefore aimed to create a library with mutation sufficient to disrupt structural elements, and screen these mutants for failures in the assembly process and the ability to secrete proteins.

The generation of mutant forms a protein in order to understand its function is a common approach. Between removing large parts of a protein, to substituting individual amino acids, a wide range of commonly used strategies exists. The decision for a specific type of mutagenesis must therefore be carefully matched to the situation at hand, weighing the information gained from more subtle mutations against the complexity of dealing with an increasing number of constructs.

By creating extensive deletion mutants [32], the entire sequence can be covered with few constructs. The limited number of strains can then be screened for changes in the phenotype using even elaborate manual methods. The resulting map of the protein will necessarily be coarse, but drastic changes in the protein architecture can be expected and understood.

On the other extreme end, the creation of an exhaustive library of single amino acid substitutions will usually require hundreds or thousands of mutated variants screened for potentially subtle changes in the phenotype, prohibiting the use of complex or manual screening methods [23]. If suitable screening or selection methods are available, the resulting data set can offer a very detailed view on the functional relevance of small parts of the protein.

In this study, a transposon-based mutagenesis kit was used to create a library of PrgH variants, each containing an insert of five amino acids. A mutation of this size will in most cases not alter the tertiary structure drastically, but will readily disrupt secondary structures and possibly affect other interaction sites.

The number of strains to achieve full coverage was estimated at around 80. While it would be possible to perform traditional secretion assays on this scale, imaging all strains by electron microscopy to identify phenotypes relevant to the assembly process would be extremely time consuming. To avoid this bottle neck, two ELISA-based assay methods were established.

In a typical ELISA, a target molecule is fished from a sample and bound to the reaction chamber by antibodies. In sequential steps the fixed molecule is then bound by a primary antibody, which is recognized by an enzyme-coupled secondary antibody. The enzyme facilitates conversion of a specific substrate into a detectable product, with the speed of the reaction being dependent upon the amount of bound target molecules.

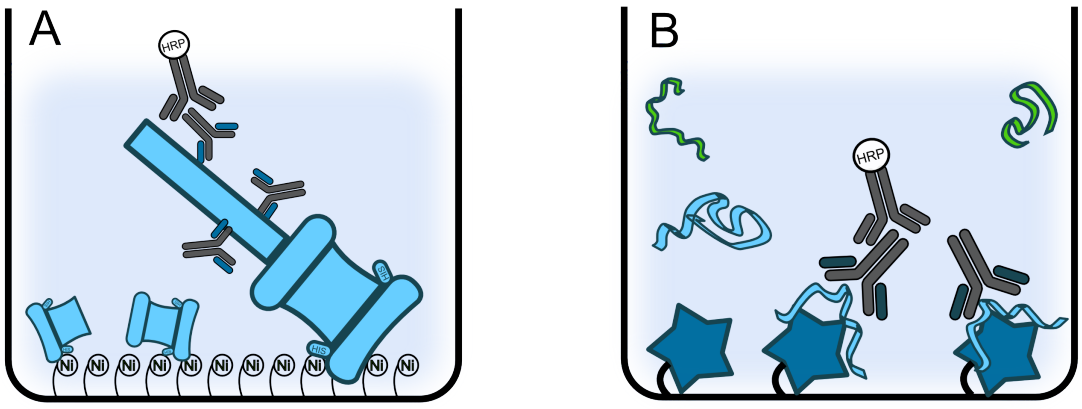


Figure 3: Schematics for (A) Structure ELISA and (B) Secretion ELISA

Overall, the first two binding steps act as selectivity filters, while the catalytic activity of the enzyme allows amplification of even weak signals stemming from minute concentrations of the target molecule. The methods described here are variations on this concept.

The structure ELISA (StrucELISA, Figure 3A) takes advantage of the sequential assembly of the needle complex. By binding the primary inner ring subunit PrgH with a polyhistidine tag to Ni-groups on the multiwell plate, detection of other, later added subunits becomes a possibility. Here, the needle filament subunit, PrgI, was detected. As this is the last structure formed during the assembly process, it serves as a marker for the presence of completely assembled needle complexes.

For the secretion ELISA (SecELISA, Figure 3B), chaperones with high binding affinities are used instead of a capturing antibody to bind corresponding effector proteins. By using two highly selective binding steps, effector proteins could be detected directly from the culture supernatant.

METHODS

DNA work

Plasmids were constructed using standard laboratory procedures. Kits routinely used for DNA purification and extraction include the QIAquick® PCR Purification Kit, Nucleotide Removal Kit, Gel Extraction Kit,QIAprep® Miniprep and QIAfilter® Maxiprep Kit, all used according to the manufacturers instructions. Polymerase chain reactions (PCR) were performed using Pfu or Pfu turbo polymerase (Fermentas, USA) according to the supplied manual and with the set of primers listed for each plasmid. Restriction enzymes were obtained from Fermentas (USA) or New England Biolabs (USA) used in combination with their respective buffers and under the recommended reaction conditions. Dephosphorylation of DNA fragments was performed using calf intestinal alkaline phosphatase (CIAP), to promote formation of specific ligation products.

All constructs could either be created through attaching restriction sites at both ends of an insert or by fusing two primary PCR fragments with overlapping ends in a second PCR step ([33], p.208). The resulting fragment could then be digested with the appropriate restriction enzymes and ligated into the target plasmid. The second technique offered advantages when fusing the upstream promoter region to a construct without inserting a restriction site in between, while avoiding the use of inefficient blunt end ligation.

If a large number of base pairs had to be attached to a protein, multiple PCR steps were used, each using the prior product as a template and a primer binding to the newly created overhang.

Plasmids and DNA fragments were visualized by electrophoresis (90 V) in 1% agarose gels, prepared in Tris-acetate buffer with EDTA (TAE buffer), stained with Sybr® green dye. Images were taken using BioRads Universal Hood II and the Quantity One / Gel Doc XR software package. Finished constructs were routinely sequenced by the in-house service department and analyzed using 4peaks (mekentsoj.com) and DNAman (Lynnon Corp.).

Transformation of DH5α cells

Amplification of plasmids was done by transformation in *Escherichia coli (E. coli)* DH5α cells. Heat shock competent cells were supplied by the in-house service department and stored at -80°C. 150 µl of these cells were thawed on ice for 10 minutes, up to 100 ng of DNA were added and carefully mixed. After keeping the mixture on ice for 30 minutes, cells were heat shocked for 30 seconds and left to recover on ice for 5 minutes. 950 µl room temperature SOC medium were added and the cells grown at 37°C with strong shaking for at least 60 minutes. 10 – 200 µl were plated on pre-warmed LB agar plates with necessary antibiotics. Colonies could be picked after growth over night.

For blunt end cloning of PCR fragments into a SmaI-digested bluescript vector (pBSK II+, Stratagene, USA), samples were ligated over night at 20°C, 40 µl XGal and 40 µl Isopropyl-β-D-thiogalactopyranosid (IPTG) [34] were spread an hour before plating the cells. Blue colonies indicated no insert, while white colonies contained plasmids with ligated inserts.

Transformation of *Salmonella typhimurium*

To achieve higher transformation rates, Salmonella strains were routinely transformed by electroporation. The following protocol describes both the preparation of electrocompetent cells and the transformation itself.

100 ml of LB liquid culture with suitable antibiotics were inoculated and the culture grown over night at 37°C and 220 rpm. On the following morning this pre-culture could be diluted in 1000 ml LB medium plus suitable antibiotics, and grown at 37°C and 180 rpm to an OD_{600} of 0.5 – 0.7. The culture was cooled down quickly in ice cold water for up to 30 minutes and harvested in a cooled centrifuge (4250g, 15 min, 4°C, Beckmann Coulter Avanti J26-XP, rotor JLA 9.100, 7250 k-factor). The cells needed to be resuspended in 500 ml of 4°C cold sterilized water, and pelleted again. This washing step was repeated an additional time with 200 ml of cold water (200 ml), followed by washing with 10% cold glycerol, first 200 ml, then 30 ml. An additional washing step was done in 50 ml falcon tubes (rotor #3057, Multifuge 3SR, Heraeus, UK), after which the cells were resuspended in 1 ml 10% glycerol,

aliquoted (80 µl each) and stored at -80°C. Fresh cells could be kept on ice and offered higher transformation efficiency than frozen cells. The initial culture and the subsequent washing steps could be scaled up or down, depending on the number of aliquots required.

The electroporation was done using 80 µl of frozen cells with 50-200 ng of DNA, filled up to 100 µl total volume with cold water. To optimize transformation rates, the volume of the added DNA solution was kept as small as possible. Minimizing the salt concentration in the samples gave higher transformation rates, therefor DNA for electroporation was eluted in water instead of the supplied elution buffer. All electroporation equipment (cuvette and holder) was cooled to 4°C. Using a 2510 electroporator (Eppendorf, USA) and cuvettes with 0.1 cm distance between electrodes, the voltage was set to 1800 V. Directly following electroporation, 900 µl of SOC medium were added to the cuvette, and the cell suspension transferred into a 1.5 ml tube. Good recovery rates were achieved by growing the cells for 60 minutes at 37°C at 650 rpm in a thermo mixer (Eppendorf, USA). Longer recovery times of up to 120 minutes were used if the transformed plasmid contained a chloramphenicol acetyltransferase (CAT) resistance marker. If a sufficient amount of DNA was used, plating 200 µl of the mixture on a LB plate with suitable selection markers was sufficient to obtain 20 to 100 separate colonies. If the amount of DNA was low, all cells were pelleted, resuspended in 100 µl LB medium and spread on a agar plate. Plates incubated at 37°C usually displayed at least small visible colonies over night.

Linker Scanner Mutagenesis

Purified target plasmid (pWS006) was subjected to the GPS-LS reaction using the supplied pGPS5 plasmid (Kan^r) as donor, transformed into commercial highly efficient E. coli cells (NEB, US) and grown on LB agar plates with kanamycin and ampicillin, selecting target plasmids with transprimer insertions. In order to archive a suitable coverage several thousand individual colonies were calculated to be necessary. Colony numbers on several plates were estimated and pooled by scraping them from a plate. This pool of cells was grown for shortly at 37°C, followed by DNA extraction. To remove plasmids with insertions in the backbone, the promoter-PrgH insert was cut out and religated into pwsk29. The resulting plasmid library was

again transformed into commercial highly efficient E. coli cells (NEB, US), grown on LB agar plates under selection, pooled and the DNA extracted. To remove the 1.7 kb transprimer insertion carrying Kan^r, the plasmid pool was digested with PmeI and religated, resulting in a pool of plasmids carrying 15 bp insertions only in PrgH and the upstream promoter (Figure 4).

This library was transformed into commercial highly efficient E. coli cells (NEB, US), about 1200 single colonies were picked, grown and the insert sequenced using a set of three primers. The resulting dataset was analyzed using a custom set of tools provided by the IMBA/IMP bioinformatics department, which automatically aligned multiple sequencing runs and annotated the insert. A set of plasmids carrying well distributed mutations all over PrgH was put together. Insertions resulting in a new stop codon, plasmids carrying more than one mutation or mutations in the promoter region were not considered. The final 87 plasmids ("original") were transformed into ΔprgH-Salmonella (SB906+pHilA(Cm^r) and screened for the presence of needle complex and secretion activity using the StrucELISA and SecELISA protocols. Each strain was tested at least three times from independently grown cultures. To confirm both the sequence, the structural and secretion data, the 87 plasmids were transformed separately into DH5α cells, the DNA extracted and sequenced again. Glycerol stocks of these strains were stored at -80°C. The plasmids ("control") were also transformed into SB906+pHilA(Cm^r) again, and tested for the presence of structure and secretion activity. During these additional confirmation experiments, a number of strains was identified that contained stop codons or redundant inserts, leaving 75 independent strains for analysis.

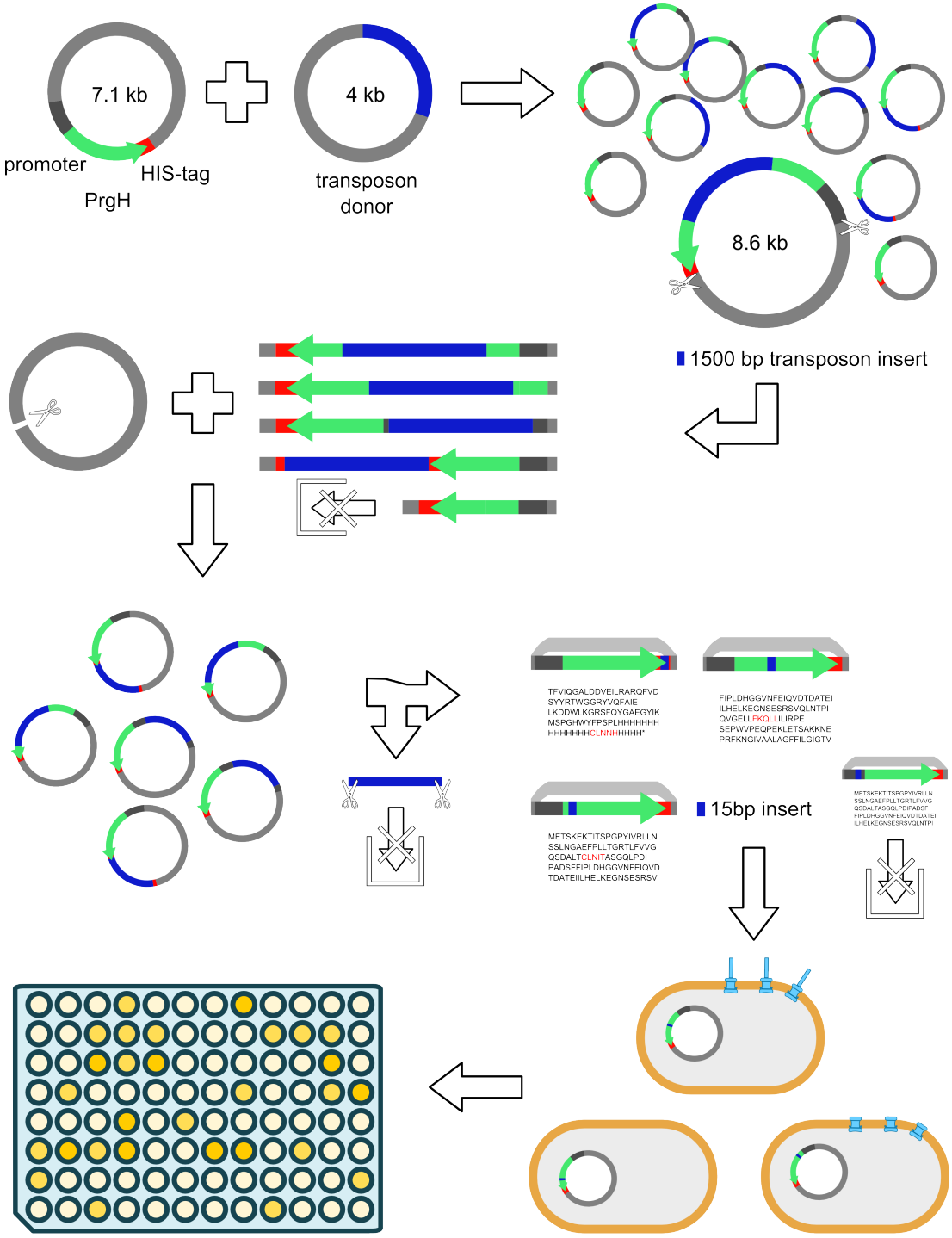


Figure 4: Workflow for creating a library of mutants based on the GPS™-LS Linker Scanning Mutagenesis Kit.

Standard needle complex purification

The commonly referred to preparation of a needle complex involves three major sections. First a growth period, in which formation of the needle complex by a bacterial strain is encouraged. Secondly, a lysis step that aims at gently dissociating the needle complex from the bacterial membranes it is embedded in. Lastly, a series of purification steps, ranging from isopycnic and isokinetic ultracentrifugation to pelleting steps, is necessary in order to obtain a sufficiently pure and dispersed sample. The original protocol [14] was modified by Oliver Schraidt, and further adapted as necessary.

The yield of a needle preparation depended to a large extent on the modifications imposed upon the proteins (tags, etc) making up the needle complex. For the salmonella strain SB905 (considered wild-type), with HilA-based expression amplification, 1 liter of culture gave a fairly large number of complexes, easily sufficient for most imaging purposes. The yield of other strains with modified needle complexes could be drastically reduced, requiring culture volumes of 10 liters and beyond. Even though the protocol here states 1 liter culture volume, the lysis protocol was successfully applied to bacterial cultures of 3 liters with a final OD_{600} of about 1.8. This appeared to be sufficient for most purposes. If an even larger culture was required, the lysis volume could be scaled up accordingly.

A 200 ml preculture of the required strain was grown over night in liquid LB medium, with 300 mM NaCl and the appropriate antibiotics, here commonly ampicillin (AMP, 100 mg/ml), chloramphenicol (CMP, 25 mg/ml) and/or kanamycin (KAN, 20 mg/ml). The incubator was set to 37°C and 220 rpm, 2 liter Erlenmeyer flasks were used.

On the following morning, the preculture was diluted to an OD_{600} of 0.1 to 0.3 with 1 liter of preheated (37°C) LB medium, containing 300 mM NaCl. The resulting culture was split up into 500-600ml aliquots and transferred to 2 liter culture flasks, to be incubated at 37°C and 180 rpm. Antibiotics could be used for stringent selection, but prolonged the growth phase and draw out the entire protocol. Once the culture reached an OD_{600} of about 0.5, induction of HilA expression using the pSB667 or pSB1418 plasmid took place by adding arabinose to a final concentration of 0,012% (0,3 ml 20% arabinose solution per 500ml). 4.5 to 5 hours

after induction the culture reached stationary phase and the fully formed needle complex was present on the cells surface. Recording a growth curve allowed following the maturation of the culture and is recommended. Most strains used for this thesis reached a final OD_{600} of 1.5 to 1.8, depending on the OD_{600} at the time of induction.

If a culture was harvested prematurely, e.g. after 3 to 3.5 hours, a larger portion of the isolated needle complexes displayed shorter needles or no needle filament at all. Cells were harvested by centrifugation at 4250g for 15 min (Beckmann Coulter Avanti J26-XP, rotor JLA 9.100, 11410 k-factor) and the supernatant discarded. The lysis procedure started by resuspending the cell pellet in 100 ml 0.5M sucrose and 0.15M Tris buffer at pH 8.0. Ions were chelated by adding 2.5 ml 0.5 M EDTA solution. 5 ml lysozyme solution (30 mg/ml, 98381 IU/mg) were added, and the suspension slowly stirred on ice for 30 to 60 minutes. Afterwards, a protease inhibitor was added (Complete® Protease Inhibitor, 1 tablet) and the temperature raised to 37°C by placing the flask in a water bath for 15 to 30 minutes. After adding 4 to 5 ml 10% Lauryldimethylamine-oxide (LDAO), the viscosity of the solution increased. Adding 11 ml of 5 M NaCl cleared the solution, and 2 ml 1 M $MgCl_2$ reactivated DNAses, liquefying the solution within 5 to 10 minutes. After this, the sample had to be put on ice again and could be stirred for 30 to 60 minutes. The first centrifugation (Beckmann Coulter Avanti J26-XP, 250ml tubes, rotor JLA-16.250; 9000g, 20min, 4°C; 1400 k-factor) helped to clear large cell debris from the solution. Dhe pellet was discarded and the supernatant split into equal parts for the following ultracentrifugation. In this step, the needle complex and similar structures were pelleted (Beckman UC, 70ml PC tubes, rotor Ti 45; 40000 rpm, 150 min, 4°C, 168.1 k-factor). The supernatant could be discarded, the pellet needed to be carefully resuspended in 2 or 4 ml of resuspension buffer 1 (10mM Tris, 0.5M NaCl, , 0.5% LDAO, pH 8.0). The use of pipet tips with a wide opening worked well to reduce mechanical stress. After careful rinsing of the wall above and the pellet itself, the sample was placed on a shaker for 60 minutes. A slow shaking setting was used to avoid the formation of foam.

For the subsequent isopycnic centrifugation, cesiumchloride (CsCl) was added to the resuspended sample. A stock solution of 52,5% CsCl in resuspension buffer 1 was prepared. 4 ml of 52,5% CsCl solution were added to 200 µl 10% LDAO and 3.8 ml

of resuspended sample, resulting in a final concentration of 27,5% CsCl. The gradient formed during ultracentrifugation over night (Sorvall 90SE, rotor TH-660, 50 krpm, 12 to 15 h, 5°C), maximum acceleration, minimal deceleration, 64 k-factor) in 4.4 ml PA thinwall tubes (Thermo Scientific, US).

If 6 fractions of 660 µl each were taken using tips with wide openings, membranes and byproducts were found in fraction 1, complexes without needle filaments were found in fraction 2, while fully formed needle complexes were located in fraction 3 and 4, if counted from the top. To check for the presence of the needle complex, see the imaging protocol.

Fractions containing the desired structure were pooled. To remove the CsCl, the sample was diluted with resuspension buffer 2 (10mM Tris, 0.5M NaCl, , 0.1% LDAO, pH 8.0) and pelleted. 1.2 ml sample were filled up with 1.7 ml of resuspension buffer 2, then centrifuged (Sorvall RC M150 GX, rotor S100-AT4-463, 90000 rpm, 30min, 5°C, 19.7 k-factor) in 3 ml polycarbonate thickwall tubes (Hitachi Koki, JP). The pellet was resuspended in 70 to 100 µl of resuspension buffer 2, again care was taken to avoid the formation of foam. The wall of the tube was carefully rinsed (~10x) and the pellet itself gently resuspended. The tube was slowly shaken for 30 minutes and the amount of needle complex determined using the electron microscope.

To further purify the sample and to remove needle complex aggregates, an isokinetic centrifugation was used. The necessary continuous 10-25% sucrose gradient was prepared on a linear gradient maker. Both the 10% and 25% sucrose solution (10 mM Tris or PBS based buffer, 500 mM NaCl, 5 mM EDTA, 0,1% LDAO) was diluted from a 50% stock solution at room temperature, LDAO was freshly added. A 4.4 ml PA thinwall tube (Thermo Scientific, US) was marked using the measuring device belonging to the gradient mixer and filled up to this mark with 10% sucrose solution. The 25% solution was added from below with a spinal needle, to create two distinctly separated layers, which fill the tube completely. The short plug was put on the tube without enclosing any air, and the tube rotated using the “10-25% short” program.

The finished gradient was cooled to 4°C for 15 minutes, the plug taken off and up to 150 µl of sample were applied on top. After the centrifugation (Sorvall 90SE, rotor TH-660. 50,000rpm, 4°C, 210 min, 64,5 k-factor) fractions were taken. If fractions

of 2x1.1 ml and 4x 450µl are taken from top, wild type needle complexes were usually located in fractions 3 to 5.

The chosen fractions could be pooled and diluted with suitable resuspension buffer (10mM Tris or PBS, 0,5M NaCl, 5mM EDTA, 0,1% LDAO, pH 8) and pelleted (Sorvall RC M150 GX, rotor S100-AT4-463, 90000 rpm, 30min, 5°C, 19.7 k-factor) in 3 ml polycarbonate thickwall tubes (Hitachi Koki, JP). Resuspending the sample was done using tips with a wide opening in resuspension buffer 3, or other suitable buffers as required.

The presence of a polyhistidine tag on a needle complex tended to lead to aggregated samples, which could be counteracted by the addition of Imidazol (20mM final concentration) and ethylenediaminetetraacetic acid (EDTA, 20mM final concentration) both after the lysis step, and in the resuspension buffers used during subsequent steps. No structural changes were observed due to treatment with EDTA or Imidazol at this concentration and within the period of time necessary to complete the purification.

Two different buffers were commonly used for storing the final sample. For day-to-day imaging, buffers based on tris(hydroxymethyl)aminomethane (TRIS) were used. If a specific labeling reaction impeded the use of TRIS, posphate buffered saline (PBS) buffers were a suitable alternative. For complete removal of TRIS, the change of buffers happened with the resuspension step after the CsCl-gradient, and all subsequent buffers had to be changed to PBS as well.

Ni-NTA-NanoGold® labelling of polyhistidine tags

Purified needle complex was incubated with with 1 mM N-ethylmaleimide at 4 °C over night to block free thiols and to prevent unspecific NanoGold interactions with cysteine residues. To further suppress unspecific labelling, imidazol was added (20 mM). Incubation with 5 µM nickel-nitrilotriacetic acid (Ni^{2+}-NTA) NanoGold (Nanoprobes, USA) at room temperature for 10 minutes lead to binding of the gold label. The resulting sample was separated from unbound Ni-NTA-NanoGold particles by gel filtration (Sepharyl 300 column) and used for cry electron imaging (Schraidt et al 2010).

Long needle filament purification

Purifying a needle complex with a extended needle filament is taxing, due to the sheer size of the complex and the relative sensitivity regarding shear forces. The protocol described here is a heavily modified version of a normal needle complex preparation, with a focus on avoiding pelleting and resuspending steps, and gradient concentrations adapted to the density of the needle filament. The primary growth period needed to be extended to allow full formation of the filament, whereas the lysis could proceed according to standard procedures.

The purification steps were adapted to avoid the preclearing step, although the exact consequences of this decision were not tested. Additionally, two CsCl gradient centrifugation were performed, one being a short CsCl cushion to gently move the NC out of the lysed sample while keeping cell debris afloat, and a second one over night to remove additional impurities, followed by a single pelleting and resuspension step.

The protocol here describes a rather large preparation based on 12 liter of culture, which will yield an amount of needle complex exceeding more than sufficient for standard imaging procedures.

A swab from a glycerol stock was diluted in 10 ml of LB medium 10 minutes before the inoculation. 12 liter of 4°C cold LB liquid medium (LB, 0.3 M NaCl, suitable antibiotics, 0,02% Arabinose) were split into 22 2 liter culture flasks, containing 600 ml each, and were inoculated with the preculture. Cells were allowed to grow over night (37°C, 180 rpm), the culture was stopped when the OD_{600} reached 1.5, usually after 17-18 hours. The lysis protocol is similar to the normal preparation, but using three times as much volume.

Cells were harvested by centrifugation at 4250g for 15 min (Beckmann Coulter Avanti J26-XP, rotor JLA 9.100, 11410 k-factor) and the the supernatant discarded.

Cell pellets were resuspended in 300 ml 0.5 M sucrose and 0.15 M Tris buffer at pH 8.0. Ions were chelated by adding 7.5 ml 0.5 M EDTA solution, 15 ml lysozyme solution (30 mg/ml, 98381 IU/mg) were added, and the suspension was slowly stirred on ice for 30 to 60 minutes. Protease inhibitor was dissolved (Complete®

Protease Inhibitor, 3 tablets) and the temperature raised to 37°C by placing the flask in a water bath for 15 to 30 minutes. After adding 15 ml 10% Lauryldimethylamine-oxide (LDAO), the viscosity of the solution increased to a honey-like appearance. Adding 33 ml of 5 M NaCl cleared the solution, and 9 ml 1 M $MgCl_2$ reactivated DNAses, liquefying the solution within 5 to 10 minutes. After this, the sample was put on ice again and was stirred for 30 to 60 minutes. Western analysis of the preclearing pellet suggested extensive losses of needle filament even at slow centrifugation speeds, therefore no preclearing was done here. A side effect is a thick layer of cell debris on the subsequent CsCl cushion.

28 ml aliquots of the lysed cell sample were taken and transferred into 36 ml PA tubes (Sorvall centrifuges. US). A 3 ml layer of 36% CsCl solution (36% CsCl in 10mM Tris, 500mM NaCl, 5mM EDTA, 0.5% LDAO, pH 8) was put under 28 ml of sample, using a spinal needle (BD Spinal Needle 18GA 3,50 IN 1,2x90mm). 6 tubes were centrifuged at each run, the second batch was stored at 4°C in between. (Sorvall Discovery 90SE, rotor AH-629, 28000 rpm, 5°C, 3 hours, k-factor 260k).The supernatant was removed down to the floating debris layer, which was then either carefully removed or dissected and moved to the side, depending on the texture.

0.5 ml fractions were taken starting from the bottom and using a spinal needle (BD Spinal Needle 18GA 3,50 IN 1,2x90mm) on a 3 ml syringe. 1:30 diluted samples were checked for the presence of needle filaments. Those fractions (usually lowest 1-3 fractions) were pooled. The density of the pooled fractions had to be estimated and was adjusted to be above 35%. Here, the lowest 4 fractions were pooled (24 ml total) and 2 ml of 60% CsCl stock solution added.

The sample was split into 6 thinwall tubes (4.4 ml PA, Thermo Scientific, US) and centrifuged over night (Sorvall 90SE, rotor TH-660. 40,000rpm, 12h, 5°C, k-factor xxxk,. Acceleration 9, Deceleration 0). The needle complexes formed a distinctly visible band in the upper third part of the tube. This fraction was taken in as little volume as possible and pelleted using small 200µl thickwall tubes (Sorvall RC M150 GX, rotor S100-AT3, 90 kprm, 30min, 5°C).

The pellet was resuspended with resuspension buffer (10mM Tris, 0,5M NaCl, 5mM EDTA, 0,1% LDAO, pH 8) and appeared to give a high yield, although the filaments tended to clump and form rather thick layers.

Alternatively, the sample was resuspended with TE buffer and 0,1% LDAO, which gave a far lower concentration of needle filaments as seen on the EM, but with a much better distribution and a less obvious tendency to aggregate. TE buffer is therefore preferred for preparing samples for imaging purposes.

Transmission Electron Microscopy (TEM)

In order to image macromolecular structures using a traditional TEM, the sample must be dried, fixed to a sample holder and the contrast enhanced.. The Morgagni electron microscope (FEI company, USA) with an 11 megapixel digital camera offers sufficient resolution to work on the needle complex, if the sample is coated with a heavy metal film by so called “negative staining”. Using the protocols outlined below, imaging is possible during all purification steps, ranging from the visualization of the needle complex in the cell membrane of intact cells, screening of various fractions after centrifugation steps and a final assessment of a purified sample. If done properly, the resolving power is sufficient to identify at least large structural changes. The necessary protocols are outlined below.

Ultimately, more advanced sample preparation and imaging techniques are necessary to create representations of the needle complex that even allow the unambiguous localization of structures less than 2nm in size, such as Nanogold® (Nanoprobes, USA). The recently installed FEI Tecnai Polara at 300 kV and a Gatan Ultrascan 4000 UHS CCD camera open this possibility. Complemented and enhanced by automated image acquisition and processing techniques, it is possible to collect a dataset sufficient for two-dimensional single particle reconstruction within less than a single day, if a high quality sample is supplied.

In cryo-EM, the sample is embedded in a thin layer of vitreous ice, by freezing in liquid ethane, and imaged at low temperatures. This reduces the chemical and mechanical stress exerted during negative staining, and avoids deformation of complexes by contact and binding to the carbon film or other surfaces. The added complexities of working on the Polara cryo-EM prohibited that the work was done in person. Plunge-freezing the sample was mostly done in cooperation with Oliver Schraidt, who also handled subsequent steps. The relevant protocols are outlined in [35] and the thesis of Oliver Schraidt (in preparation).

Negative stain EM sample preparation

Carbon-coated CuPd grids with a 400 µm mesh width were used as supplied by the IMP/IMBA electron microscopy service facility.

Glow discharge

The surface of a carbon film is modified to improve the distribution of the later applied aqueous solution, and to facilitate binding of protein structures to the surface. Protocol according to IMP/IMBA Electron Microscopy Service, GR/TM (V1.1, June 2006).
Carbon coated grids were placed in batches of 3 to 6 on a glass plate and placed between the electrodes of the Bal-Tec SCD 005 sputter coater. Coating was done for 30 seconds at 21 mA at less than 10 mbar air pressure. Glow discharging a grid had to be done directly before applying a sample for negative staining.

Negative staining

Protein complexes are bound to a carbon surface, dried und coated with phosphotungstic acid, to create a electron-dense layer, that can easily be imaged. Protocol according to IMP/IMBA Electron Microscopy Service, GR/TM (V1.0, May 2007).
A freshly glow-discharged grid was taken with a forceps at the outer rim and 5 µl sufficiently diluted sample were applied for 45 seconds on the carbon-coated side. By touching it slightly against filter paper, the liquid sample was taken off. Without allowing the grid to dry out, 5 µl of 2% phosphotungstic acid (PTA, filtered, pH 7 with KOH) were applied, taken off instantly and another 5 µl PTA applied. This represents the washing and the coating step. Depending on the thickness of the desired stain, PTA was taken off with filter paper from the edge after about 30 seconds. To avoid salt artifacts, both sides and the forceps had to be thoroughly dried.
When looking for the needle complex in salt-rich buffers (e.g. directly from CsCl-gradient fractions), dilutions of 1:20 or higher were necessary, and very careful removal of all liquid during the first washing step is advisable. Even under these

conditions, EM samples prepared directly from gradient fractions generally gave low quality results, preventing any quantitative predictions, although still being useful to quickly localize relevant fractions without pelleting steps.

Direct imaging of cells

In order to identify surface structures as the needle filament, or even needle complexes, whole cells could be imaged. Acceptable results were obtained by diluting a culture shortly before harvest 1:5 or 1:10 in TE buffer. The cells were applied and stained with PTA as described earlier, although the initial settling time was raised to 1 minute.

Electron Microscope

Negative stain images were taken on a FEI Morgagni transmission electron microscope (FEI company, USA) at 80 kV, using an 11 megapixel CCD and the iTEM software.

Secretion profiles: protein precipitation

To visualize which proteins a strain is able to transport across its membranes to its surroundings, the proteins have to be concentrated, which is done here by precipitation with trichloroacetic acid (TCA). Adding TCA to the culture supernatant lowers the pH to the point at which most proteins cannot be solubilized anymore, and form a precipitate. Of course, not all proteins identified in a solution like this are transported through the type III secretion system – proteins transported by other secretion systems, surface proteins or membrane-permeable proteins will be identified as well. All strains tested here are deficient for the flagellar assembly, thereby removing a close relative of the type III secretion system.
A shorter growth period appeared to favour proteins secreted by the needle complex, although extended growth tended to increase total signal strength. Lysis of cells and thereby unspecific release of type IIII effector proteins appeared not to be an issue for the culture conditions described here. The optimal growth period must therefore be chosen with the application in mind: for the unambiguous identification of

secretion positive or negative strains by immunoblotting, a long growth period might be desirable. For secretion profiles, a shorter growth period might help to limit the number of nonrelevant proteins in the supernatant. The size of the culture also depends on the application, ranging from 20 ml for western blotting and 100 ml for secretion profiles.

A salmonella culture was grown under conditions promoting needle complex formation. This was achieved by using culture conditions similar for needle complex preparation, with an uninduced preculture over night followed by 5 to 10 hours of induced growth, or with direct inoculation and growth over night in liquid LB (0.3 M NaCl, 0,012% arabinose, suitable antibiotics), both at 37°C.

The culture was cleared of cells by centrifugation (5000g, 15 min, 4°C), in tubes suitable for the amount of culture volume. Cells of a defined amount of culture (1 ml) were usually pelleted as an expression control and the supernatant discarded. The supernatant of the main culture was passed through a 0.45 µm filter, either with a syringe or through a bottle-top filter (Nalgene, US), again depending on the culture volume at hand.

1.5 ml of trichloroacetic acid (TCA, 100% w/v stock solution) were added to 10 ml of protein sample and stored at 4°C for at least 60 min. For the following steps, centrifuges were cooled to below 0°C. Acetone used for washing the sample was stored at -18°C and kept on ice on the bench.

The precipitate was pelleted by centrifugation in falcon tubes at 8000 rpm, for 40 min (rotor #3057, Multifuge 3SR, Heraeus, UK), the supernatant was carefully discarded and the pellet washed with 1.8 ml cold acetone. The sample could then be transferred to 2 ml plastic tubes and was stored at -20°C for 60 minutes or over night. Proteins were pelleted in a refrigerated table top centrifuge at maximum speed (13 krpm, 10 min), the supernatant removed and 1.8 ml of cold acetone (-20°C) added. The last stop was repeated, the supernatant taken off and the pellet dried on a heating block set to 95°C for a few minutes. The pellet was resuspended in 40 µl H_2O and 60µl 5x Lämmli sample buffer. The buffer usually kept its blue colour, if the washing steps were executed properly. If the colour shifts to a greenish or yellow hue, the pH could be adjusted to compensate for residual TCA. The sample was heated to 98°C for 5 min on a heating block before loading on a polyacrylamid gel.

Identification of single protein bands was performed by the in-house protein chemistry department, using mass spectroscopy.

Protein separation: polyacrylamide gels

SDS Polyacrylamide (PA) gels were prepared according to standard laboratory procedures [34]. Commonly used gels include 10 or 12% PA gels, which were prepared in advance and stored under humid conditions at 4°C. A 4% loading gel was added directly before usage, with 1.5 mm / 10 or 15 lane combs. Gradient gels were prepared by combining 5 and 20% PA solutions in a gradient mixer, using a econo pump (Bio-Rad, USA). The comb was directly inserted into the gradient and the gel polymerized for 60 minutes. Polymerized gels could be stored with or without comb at 4°C under humid conditions. For larger gels, the procedure was simply upscaled.

Protein samples were mixed with 5x Lämmli buffer and heated to 99°C for 5 minutes before loading. Gels were run using the Bio-Rad Mini-PROTEAN equipment at 20-40 mA/gel in SDS running buffer. PageRuler™ prestained protein ladder or PageRuler™ prestained protein ladder + (Fermentas, USA) were used as a molecular weight indicator.

Coomassie Brilliant Blue staining

A gel with separated protein samples was removed from the glass chamber, washed by placing it in a plastic container with 100 ml of H_2O, heating it for 1 minute in a microwave oven and letting it slowly shake for 5 minutes. The dH_2O was replaced and the procedure repeated two more times. The gel was then stained by covering it with PageBlue™ Protein Staining Solution (Fermentas, USA) and heating it for 30 seconds without boiling in the microwave oven. To enhance the staining, the solution was left on the gel for 20 minutes while slowly shaking. The gel was then washed multiple times with dH_2O, preferentially over night to reduce background staining.

Immunodetection of proteins (Western Blot)

A gel with separated protein samples was removed from the glass chamber, washed by placing it in a plastic container with 100 ml of H_2O and gently shaken for 5 minutes at room temperature. The H_2O was then replaced by protein transfer buffer (PTB) [34]., and left shaking for another 10 minutes

A PVDF membrane (Millipore, Billerica, USA) and a suitable amount of filter paper (2x three layers of Whatman® 3MM Chr filter paper) were cut to size. The membrane was activated by wetting it with 100% methanol, then washed for 5 minutes in dH_2O, followed by 10 minutes equilibration in 1x PTB. The filter paper was wetted using 1x PTB. The transfer of proteins from the gel to the membrane was done by using Bio-Rads Trans-Blot SD Semi-Dry Transfer Cell. The components were stacked as follows, from bottom to top: three layers of filter paper, the membrane, the gel and three layers of filter paper. Trapped air was removed by gently rolling a tube over the stack and excess liquid removed from the electrodes. Blotting was done at a constant 12 V for 75 minutes.

After blotting, the membrane was carefully placed in a suitable plastic container, in which all incubation steps took place while slowly shaking. Additional binding sites were blocked by soaking the membrane in 5% non-fat dried milk solved in 0.1% Tween 20 in PBS (PBS-T) for 1h at room temperature or over night at 4°C. The membrane was rinsed twice with PBS-T and then placed in a dilution of the primary detection antibody in PBS-T. The membrane was incubated for one hour at room temperature, or over night at 4°C. Two rinsing steps and three washing steps, lasting each 5 minutes in PBS-T, were done before the diluted (PBS-T) secondary antibody was added for 1 to 2 hours at room temperature. The membrane was again rinsed twice and washed three times before the detection reagents (Amersham™ ECL Western Blotting Detection Reagents; GE Healthcare, USA) were applied. Excessive liquid was carefully removed and the 1 ml of detection reagent pipetted and spread over each membrane. After one minute of incubation the membrane was placed in transparent film, placed in a x-ray cassette and the protein band could be detected by film exposure (Amersham™ Hyperfilm™, GE Healthcare, USA) for 10 seconds to 30 minutes. The exposed film was developed using a Colenta MP 900 F film developer.

Enzyme-linked immunosorbent assay (ELISA)

Detecting secretion of SptP by ELISA

The so called secretion ELISA (SecElisa) is an innovative method which allows detection of effector molecules directly from culture supernatant. The primary advantage is the possibility to test up to 96 samples per multiwell plate, including both culture growth and effector detection within 24 hours. It has been successfully used so far to test 384 samples per man-day, and is routinely used instead of TCA precipitation and western blotting. The SecELISA is established using the effector/chaperone pairs SptP/SicP and SipA-Flag/InvB. As only the first pair is routinely used, it is the focus of the protocol.

GST-SicP was expressed in *E. coli*, the cells lysed and the protein purified by high pressure liquid chromatography (HPLC) on a GST-trap column. The final protein concentration of SicP for precoating was 60 µg/ml in TBS. 96 well flat bottom multiwell plates (300µl, Microtest™ 96-well ELISA plates clear, BD Falcon, US) were washed with 100 µl TBS per well and precoated with 100 µl of SicP in solution for either 2 hours at room temperature, or at 4°C over night while slowly shaking. All liquid was removed and the plates stored at -80°C. Successful ELISAs were performed with plates stored for up to a year.

Bacteria were grown in 2 ml deepwell plates (Riplate® sw 2 ml, Ritter) in 1.8 ml LB growth medium (0,3 M NaCl, 0,012% arabinose, antibiotics) for at least 16 hours. Culture supernatant was obtained by pelleting cells at 4000g for 15 minutes (Multifuge 3 SR, Heraeus, UK).

Precoated plates were wrapped in paper to avoid condensation and warmed to room temperature. During incubation steps, the plate was covered with aluminium foil and placed on a shaker, the protocol is best performed at room temperature. A washing step consists of adding and removing 200 µl of PBS-T per well. Unspecific binding was minimized by blocking with 300 µl 3% bovine albumine serum (BSA) in PBS for one hour, after which the blocking solution was removed without a washing step. 100 µl of culture supernatant were loaded for one hour to capture the effector protein, followed by three washing steps. 100 µl of the first antibody dilution (1:3000 monoclonal α-SptP from mouse in PBS-T) were added for

one hour, followed by three washing steps. 100 µl of the second antibody dilution (1:15000 α-mouse from rabbit, coupled to horse radish peroxidase (HRP)) were added for one hour, followed by three washing steps. For the final detection step, 100 µl of freshly prepares substrate were applied. 3,3´,5,5´-tetramethylbenzidin (TMB) was prepared as a 100x stock solution by dissolving 10mg/ml in dimethylsulfoxide (DMSO). The working solution is prepared by mixing 100 µl of 100x stock solution, 2µl of 30% H_2O_2 and 10 ml 0.1 M citrate buffer (pH 6), and should be at room temperature when applied.

While a blue color was developing, the reaction kinetic could be measured using a plate reader (GENios Pro, Tecan, US) at 640 nm. As soon as a distinct blue color was displayed, the reaction was stopped using 25 µl of 2 M sulfuric acid (H_2SO_4) and the absorbance was measured at 450 nm.

Detecting the presence of needle filament by ELISA

The nonflagellated *Salmonella typhimurium* strain SB906 [20] carrying the transcriptional regulator hilA under the araBAD promoter was complemented by a plasmid carrying PrgH with a C-terminal poly-histidine tag (His_{18}). Bacteria were grown in 2 ml deepwell plates (Riplate® sw 2 ml, Ritter) in 1.8 ml LB (0,3 M NaCl, 0,012% arabinose, antibiotics) for at least 16 hours. Cells were harvested by pelleting at 4000g for 15 minutes (Multifuge 3 SR, Heraeus, UK), the supernatant could be used to perform the SecELISA. The cell pellet was resuspended in 400 µl lysis buffer (0.5 M Sucrose in 0.15 M Tris buffer, 0.48 mg/ml lysozyme, 5 mM EDTA) and incubated at 14°C for 45 minutes in a thermoshaker. The tubes were then incubated at 37°C for 20 minutes at 37°C. This step could be done either in a thermoshaker or in a waterbath, depending on the number of samples. 100 µl of LDAO/salt mixture (6 parts 10% LDAO, 44 parts 5 M NaCl) were added to each sample, followed by another 5 minute incubation step at 37°C. 4 µl of $MgCl_2$ are added, followed by a similar incubation period. The finished cell lysate was then kept at 4°C for 10-20 minutes.

20 µl of the cell lysate were diluted with 80 µl resuspension buffer (PBS, 0.5M NaCl, 3% BSA, 0.5% LDAO, pH 8.0). The resulting solution was loaded onto pre-wetted (resuspension buffer) multiwell plates with nickel-coating (Ni-NTA Hisorb

plates, Qiagen, USA), incubated for 1 hour, followed by three washing steps. A washing step included adding 200 µl resuspension buffer, slowly shaking the plate for 5 minutes and removing the solution. For this protocol, all incubations were performed at room temperature on a slowly shaking plate, covered by aluminum foil.

Antibodies were diluted in resuspension buffer and left on the multiwell plate for 1 hour. αPrgI (1:2000) was added first and removed by three washing steps after one hour incubation time. α-rabbit peroxidase-coupled antibodies (1:10000) were used as secondary antibodies, again incubated for one hour and washed off by three washing steps. The final detection was performed by adding 100-150 µl of 3,3 - Tetramethylbenzidine supersensitive (Sigma, US). As soon as a distinct blue color was displayed, the reaction was stopped using 20 µl of 2 M sulfuric acid and the absorbance was measured at 450 nm.

Needle complex MiniPrep

It is possible to image needle complexes from culture volumes as small as 2-4 ml. The lysis protocol described for the structure ELISA can be used to prepare samples for EM imaging.

The resulting cell lysate was cleared by a centrifugation step (7000g, 30 min) in a tabletop centrifuge. The needle complexes were pelleted by ultracentrifugation (Sorvall RC M150 GX, rotor S100-AT3, 50000 rpm, 30min, 5°C) and resuspended in 12-40 µl resuspension buffer 2 (10mM Tris, 0.5M NaCl, , 0.1% LDAO, pH 8.0). The sample was then used pure or diluted (1:10) for the negative staining. Usually, several needle complexes were found, however significant amounts of additional proteins will prevent high quality imaging.

RESULTS

Topology of PrgH and PrgK: site specific labeling

PrgH C-terminus can be visualized at the interface to InvG

Initial experiments focused on labeling a his_6 tag at the C-terminal end of PrgH. Only a small subset of the imaged particles displayed an additional density, located in the lower neck region of the base (Figure 1). The interaction between the Ni-NTA NanoGold™ and the his_6 tag appeared to be specific, but happened either too infrequent, or too weak to persist during the final purification and sample preparation steps. The low proportion between labeled and unlabeled complexes made a reliable 2D reconstruction unfeasible (data not shown).

The length of the tag was increased to 18 histidines to reinforce the binding behavior of the gold particles. The functionality of the complex was not impaired by either of the tags according to secretion tests [35], but the yield of complexes after a routine purification appeared lower compared to a wild-type strain. Gold particles bound readily to the his_{18} tagged needle complexes, resulting in a prominent additional density. A number of complexes showed multiple labeled copies of PrgH, resulting in a ring-like label around the complex. By using single particle reconstruction [35], this ring became easy to identify and positions the C-terminal domain of PrgH in the lower neck region of the needle complex (Figure 5D).

PrgH N-terminus can be visualized at the cytoplasmic tips of the needle complex

Based on experiences with the C-terminal label, the corresponding N-terminal his tag was initially designed with a length of 18 histidines. The resulting strain showed secretion activity, but hardly any complexes were obtained after purification. The few complexes actually observed had extensive, albeit smeared densities below the cytoplasmic face of the structure (data not shown). Given the unusual length of the his_{18} tag, strong interactions between tagged PrgH subunits were assumed. Aggregates would then be either formed between PrgH monomers in the cytoplasm or the membrane, or between fully formed, tagged complexes during the purification. Large protein aggregates would tend to pellet faster, and would therefore be lost

during centrifugation steps. To reduce the strength of these interactions and thereby reduce aggregation, gradually shorter tags were introduced.

Strains carrying a his_{14} and his_{12} tag on the N-terminus of PrgH were constructed, which both exhibited secretion activity. When complexes with a his_{12} tag were purified, unusual aggregation patterns were observed. A rough visual inspection showed that hardly any single needle complex was recovered. Most complexes interacted via their cytoplasmic side with other complexes, often forming symmetric dimers on the EM grid (data not shown). Base structures (without a needle filament) appeared to be less prone to aggregate. While a base would readily bind to a needle complex, base-base dimers were less commonly observed, possibly related to structural changes during needle complex assembly [15].

The addition of imidazol, ethylenediaminetetraacetic acid (EDTA) or nickel ions, which affect the binding behavior of histidine groups, either during the purification process, or to the purified complexes, remained without effect. Similarly, treatments with methanol or ethanol tended to destroy complexes before the aggregates could be broken up (data not shown).

A strain carrying an N-terminal his_6 tag finally appeared to aggregate only lightly, and showed secretion activity. Binding of Ni-NTA-NanoGold™ was robust enough to yield a strong signal, although smeared around the innermost tips of the needle complex (Figure 5F).

PrgH M267 is located at the widest point of the inner ring

Based on the crystal structure of $PrgH_{170–362}$, secondary structure predictions and the mutagenesis screen, a poly-histidine patch was inserted into PrgH after Met_{267}, a position known to tolerate a five amino acid long insert (Table 3). Two constructs, adding either a his_{12} or a longer GSG-HHHHHH-GSG sequence abolished secretion activity, while a short stretch of his_6 was tolerated, resulting in functionally active and only lightly aggregating complexes (data not shown). Gold particles could be imaged bound to the widest part of the lower ring (Figure 5E), demonstrating the relative location of PrgH and offering a valuable reference point for modeling the overall topology.

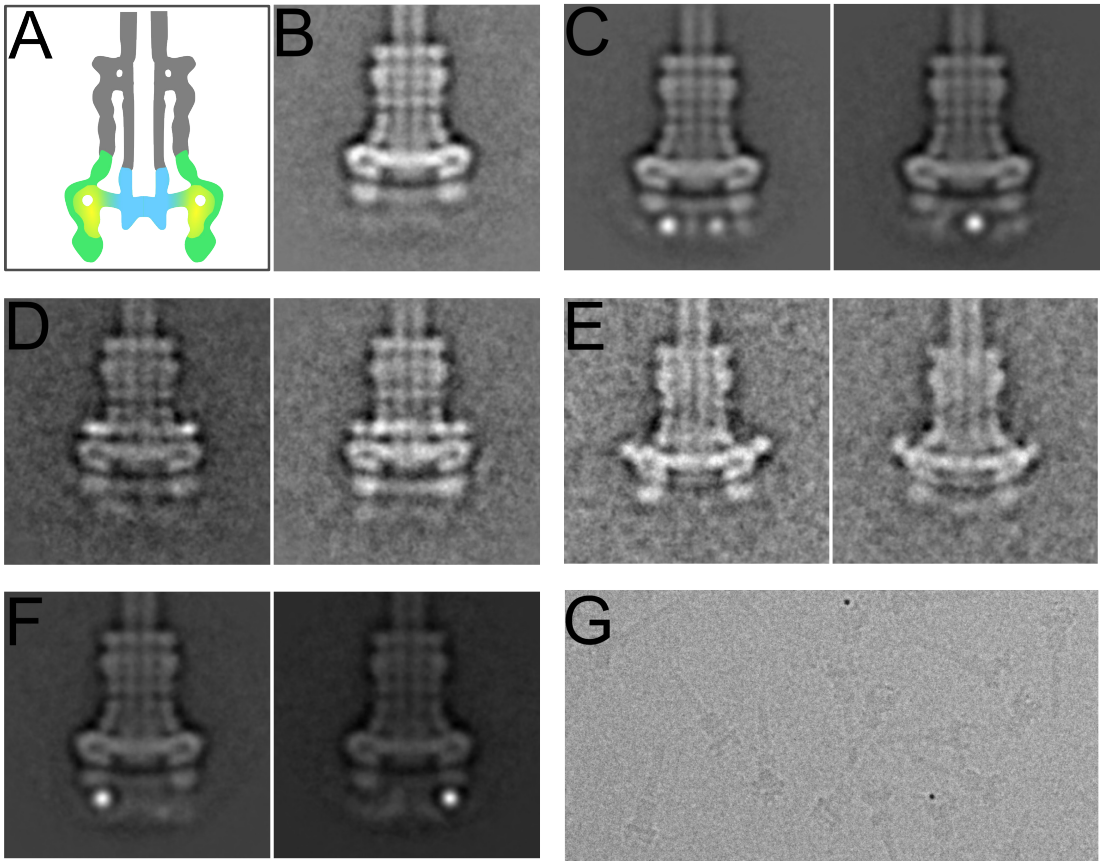


Figure 5: Various reconstructions from electron micrographs. Each image was created by averaging multiple similar images in order to increase contrast. (A) map of the needle complex, selected class averages of (B) wild type needle complex (C) gold label at C-terminus of PrgK (D) gold label at C-terminus of PrgH (E) gold label at position M267 of PrgH (F) gold label at the N-terminus of PrgH, (G) single needle complexes at imaged by cryoEM. Images and reconstructions by Oliver Schraidt.

PrgK C-terminus is located more inwards of the N-terminus of PrgH

A his_{18} tag at the C-terminal domain of PrgK (Galán Lab, Yale University, USA) prevented successful purification, possibly similar to the problems encountered with the N-terminal label of PrgH. The label was therefore exchanged for a his_6 tag. The resulting strain showed secretion activity and only lightly aggregating complexes. The final gold label was located in a comparable position as the N-terminal label of PrgH at the cytoplasmic face of the complex, although with a more restricted freedom of movement and more confined to the inner opening of the ring (Figure 5C, [35]).

PrgH can support large N-terminal tags or fusion proteins

Since tags of varying length at the N-terminus of PrgH were tolerated, an eGFP-PrgH fusion protein was constructed. The strain was surprisingly able to assemble needle complexes and displayed secretion activity (Figure 6A). Purified complexes obtained from the strain tended to aggregate, so far only a few complexes were visualized (Figure 6C). Confocal fluorescence microscopy was performed on bacterial cells in cooperation with Martin Breuss to localize or count needle complexes in the bacterial membrane. The resolving power of the convocal microscope was insufficient to distinguish individual bacterial features, even after applying deconvolution methods. Most of the GFP appeared to be localized at the bacterial membrane (Figure 6B).

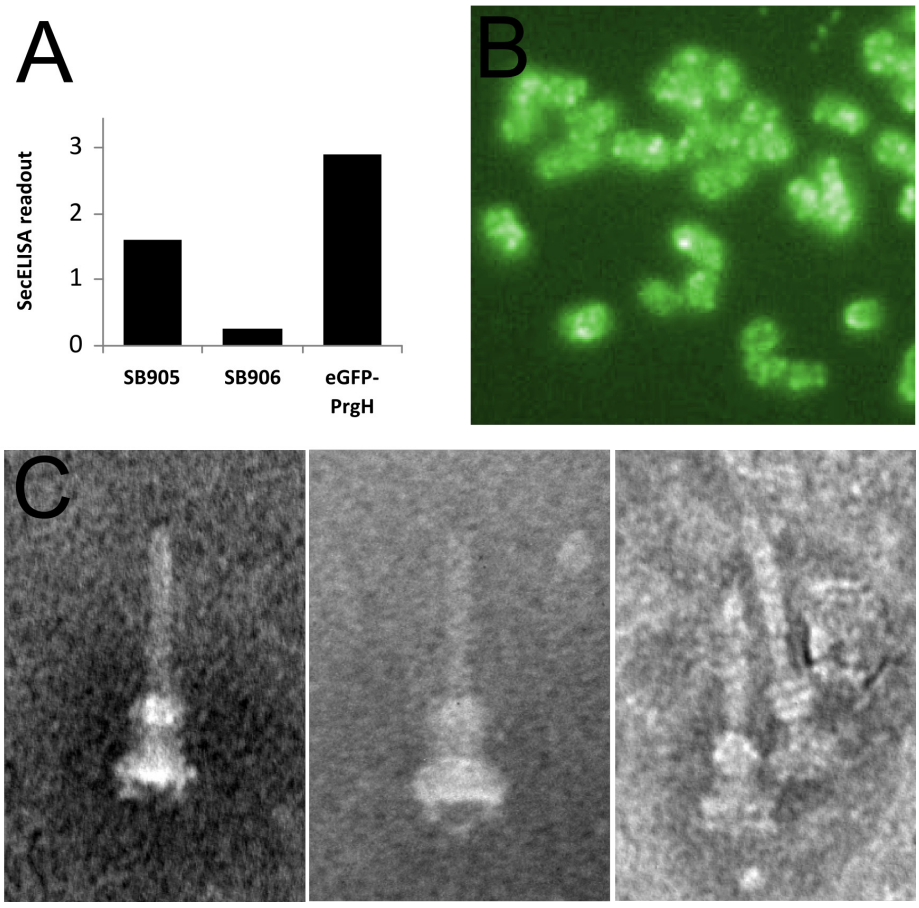


Figure 6: Summary of strain expressing eGFP-PrgH data. (A) Secretion activity, low readout for SB905 due to different growth conditions. (B) eGFP-Prgh expressing *S. typhimurium*, imaged by fluorescence microscopy/deconvolution (Martin Breuss). (C) Single needle complexes obtained from an eGFP-PrgH expressing strain

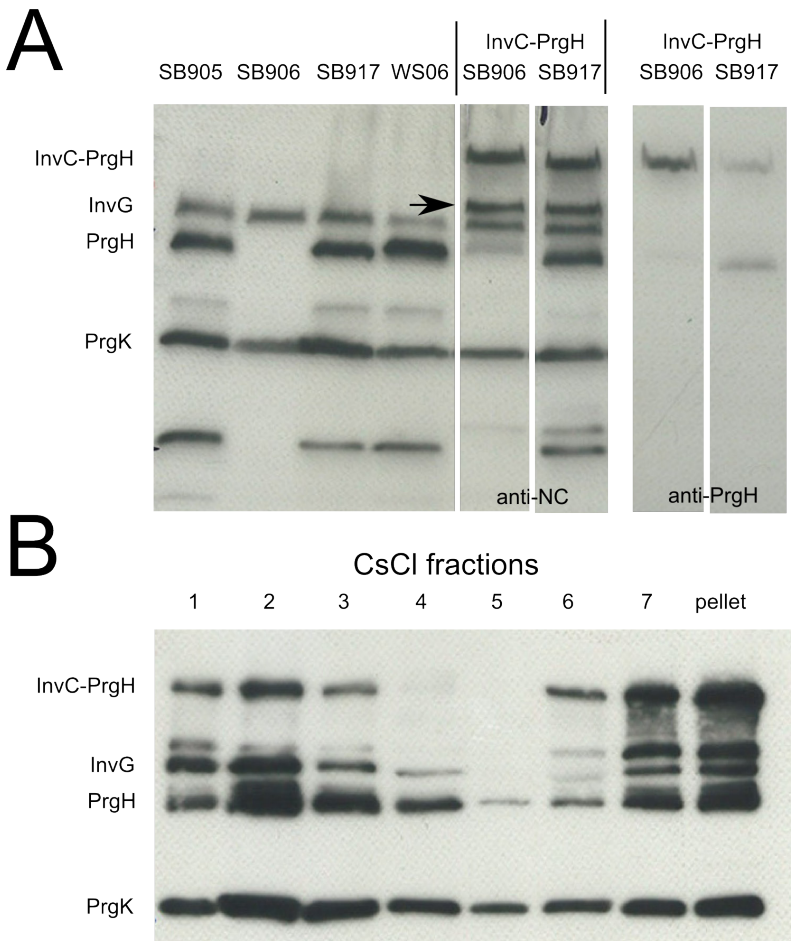


Figure 7: Purification of an InvC-PrgH expressing strain. Western blots showing (A) expression patterns in various strains (B) Localization of InvC-PrgH on a CsCl gradient. Arrow marks potential degradation product.

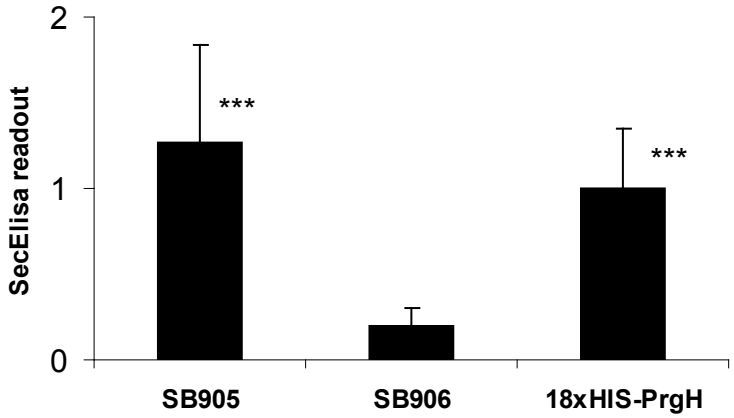
A similarly constructed InvC-PrgH fusion protein was successfully expressed in *Salmonella* (SB917 – ΔinvC and SB906 – ΔPrgH) and showed secretion activity in first trials in the SB917 background. An attempt to purify the strain showed that InvC-PrgH was translated, but degradation products were visible on western blots (Figure 7A). The full length protein co-localized with needle complex base proteins PrgK and InvG on a cesium chloride gradient during purification and therefore has to be assembled into the needle complex (Figure 7B). As no needle filament could be detected, and secretion activity could only be seen in the SB917 background, it is unclear how InvC-PrgH affects the secretion functionality. It seems possible that cytosolic InvC-PrgH can rescue the ΔInvC strain, but that the ΔPrgH strain is blocked by the addition of only InvC-PrgH. This strain will have to be characterized in more detail in the future. It could be a useful tool for the structure determination of

the ATPase InvC, as the needle complex might act as a soluble scaffold for the assembly of a multimeric InvC complex.

Mutagenesis screen of PrgH

Establishing ELISAs to assess the assembly and secretion functionality

The secretion ELISA in its current form, based on the SptP/SicP effector/chaperone pair, routinely delivered highly reproducible results. It offered advantages in terms of handling and throughput over the immunoblotting based methods [36] used so far. The signal obtained from a secretion positive strain was 8 to 15 times higher compared to a secretion negative strain (Figure 8). Signal variation between ELISAs performed on comparable culture supernatants and pre-coated SicP plates was low.



	average readout	standard deviation	n wells	n cultures	t-test (vs 906)	
SB905	1,27	0,57	35	8	<0,001	***
SB906	0,19	0,11	62	15	1,000	
PrgH-His_{18}	1,01	0,34	46	22	<0,001	***

Figure 8: Summary of Secretion ELISA data of multiple cultures.

The protocol lined out in the methods worked well for strains commonly used, but has been designed with a limited supply of primary antibodies in mind. Raising the concentration of the primary antibody from 1:3000 to 1:2000 or higher increased the

sensitivity, which might be required for testing strains with hardly any secretion activity. Growth conditions and duration of Salmonella culture appeared to influence signal intensity drastically. In general, 96-well 3 ml deep well plates offered higher throughput for large numbers of samples, but cultures tended to grow to a lower optical density and delivered a lower readout compared to cultures grown in larger, e.g. 10 ml, culture tubes (data not shown, [37]). Results obtained by the structure ELISA followed this pattern closely.

For routine operation, note should be taken of the stability of chaperone pre-coated plates. The oldest plates tested so far had been wrapped in aluminum foil and stored at -80°C for over a year without apparent degradation.

Details on the structure ELISA are described in the diploma thesis of Julia Radics [37]. Conceptionally, assembling the needle complex is a strictly stepwise process, with each successive step relying on the completion of the prior step. By screening assembled complexes for the presence of stage-specific proteins, the consequences of mutations on the assembly process can be inferred. Here, the presence of the needle filament protein PrgI was used as a marker for a fully assembled needle complex, which would structurally be able to transport proteins into the culture supernatant.

Initial experiments used a C-terminal his_6 tag on PrgH to capture the needle complex and were only of limited success. Similar to the improved binding of an extended C-terminal his_{18} tag to Ni-NTA gold particles seen in the topology studies, a strain carrying a C-terminal his_{18} tag also gave higher readings in the structural ELISA. Further improvements of the protocol as well as using a more sensitive substrate for the detection reaction finally allowed successful detection of the finished assembly.

Mutagenesis screen

By using the GPS®-LS Linker Scanning System (NEB, US) a primary library of plasmids with random 15 base pair (bp) long insertions within the insert (Promoter and PrgH-his_{18}) of the plasmid used was produced. 5 bp were duplicated from the site of insertion, the remaining 10 bp encoded, depending on the reading frame, a cysteine, lysine or stop codon, surrounded by other amino acids.

From roughly 1200 sequenced plasmids a library of initially 87 unique plasmid was selected, which was reduced to 75 plasmids during the course of the experiment due

to sequencing errors, phage contaminations or irreproducible results for certain strains. Each plasmid contained a mutation resulting in a 5 amino acid insert in the coding sequence of PrgH. On average, these mutations were spaced 16.2 base pairs apart (±16.3 STD).

The primary library of 1200 plasmids also contained variants with mutations in the 500 bp upstream promoter sequence, one or more stop codons within PrgH or more than one mutation within the insert. No systematic analysis was done on these additional strains, but PrgH variants with stop codons close to the C-terminus were observed to affect needle complex stability (pers. comm. Julia Radics) and allowed to pinpoint interactions between PrgH and InvG [35].

The library was screened both for the presence of needle filament as a marker for completion of the assembly process, and secretion of SptP by ELISAs. Each test was performed at least three times from independent cultures, plus additional experiments if conflicting results were obtained. A strain was considered to deliver a positive signal if the readout of the ELISA exceeded a threshold calculated at the average of the negative control plus 3 times the corresponding standard deviation.

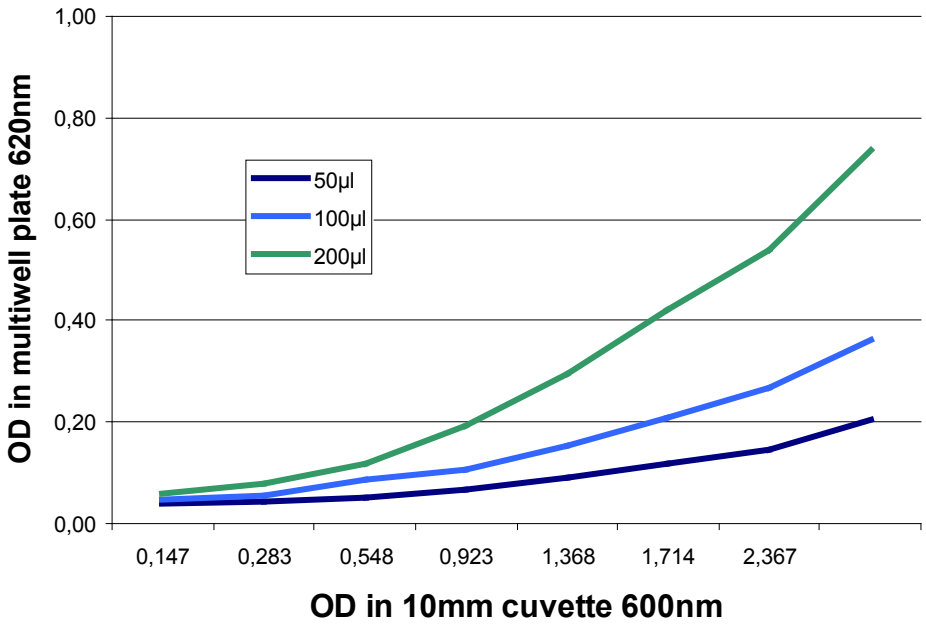


Figure 9: Conversion graph to estimate the optical density of a bacterial culture. Tables 1-4 give the OD value as obtained from a multiplate reader, and have to be converted to conform to values obtained from commonly used 10mm cuvettes.

Clone Number	Insert Position	Reading Frame	Amino acid	Plate Code	Translation	OD density (620nm)	STD optical density	Average StrucELISA	STD strucELISA	Average SecELISA	STD SecELISA
478	18	1	6	1G5	METSKE**CLNKE**KTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,215	0,02	1,14	0,18	0,88	0,20
334	19	3	6	1F3	METSKE**MFKQE**KTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,24	0,02	0,81	0,12	0,83	0,13
61	28	3	9	1G1	METSKEKTI**MFKQI**TSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,296	0,07	0,38	0,05	0,20	0,09
690	33	1	11	1B8	METSKEKTITS**CLNTS**PGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,254	0,04	0,90	0,09	0,81	0,14
372	37	6	12	1D4	METSKEKTITSP**VFKHP**GPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,209	0,03	1,00	0,10	0,81	0,06
638	43	3	14	1E7	METSKEKTITSPGP**LFKQP**YIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,24	0,01	0,53	0,05	1,19	0,18
955	49	6	16	1D10	METSKEKTITSPGPYIV**FKHIV**RLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,274	0,09	0,17	0,02	0,14	0,05
762	58	3	19	1E8	METSKEKTITSPGPYIVRL**LFKQL**LNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,22	0,01	0,16	0,05	0,10	0,01
812	73	3	24	1H8	METSKEKTITSPGPYIVRLLNSSL**MFKQL**NGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,256	0,06	0,53	0,06	0,11	0,01
675	82	6	27	1H7	METSKEKTITSPGPYIVRLLNSSLNGA**VFKHA**EFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,26	0,07	0,28	0,04	0,13	0,02
608	102	1	34	1B7	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTG**CLNTG**RTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,206	0,02	1,02	0,24	0,15	0,03
115	118	6	39	1B2	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFV**VFKHV**VGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,295	0,05	1,09	0,20	0,14	0,02
407	141	1	47	1D5	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALT**CLNIT**ASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,213	0,01	1,67	0,22	0,47	0,00
562	150	1	50	1F6	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGCL**NTGQL**PDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,221	0,03	0,76	0,18	0,71	0,18
362	184	6	61	1B4	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFF**MFKHF**IPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,241	0,04	2,03	0,26	0,13	0,03
429	193	3	64	1E5	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPL**VFKQL**DHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,193	0,01	0,40	0,34	0,29	0,06
304	223	6	74	1D3	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQV**FKHQV**DTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,243	0,03	1,06	0,15	0,77	0,12
821	231	1	77	1B9	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDT**CLNNT**DATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,241	0,06	1,12	0,14	0,95	0,08
938	234	1	78	1C10	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTD**CLNTD**ATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,279	0,09	0,53	0,05	0,34	0,20
48	237	1	79	1C1	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDA**CLNNA**TEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,313	0,05	0,79	0,06	0,44	0,18
910	246	1	82	1A10	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEICL**NKIIL**HELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,255	0,08	0,73	0,08	1,30	0,03
196	253	3	84	1C3	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIIL**LFKQL**HELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,287	0,05	0,19	0,01	0,11	0,00
365	255	1	85	1C4	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILH**CLNIH**ELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,237	0,03	0,80	0,15	0,27	0,04
23	256	6	85	1A1	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILH**VFKHH**ELKEGNSESRSVQLNTPIQVGELLILIRPESEPWVPEQPEKLETS	0,285	0,08	1,14	0,29	0,37	0,28
985	309	1	103	1H10	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPI**CLNTI**QVGELLILIRPESEPWVPEQPEKLETS	0,301	0,09	0,17	0,02	0,11	0,00

Table 1: Summary of sequence information, OD, SecELISA and StrucELISA data for all strains in the mutagenesis screen

Clone Number	Insert Position	Reading Frame	Amino acid	Plate Code	Translation	OD density (620nm)	STD optical density	Average StrucELISA	STD strucELISA	Average SecELISA	STD SecELISA
991	324	1	108	1A11	PIQVGELCLNKLLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,275	0,08	0,18	0,02	0,12	0,02
979	325	3	108	1F10	PIQVGELLFKQLLLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,249	0,09	0,17	0,02	0,09	0,01
460	352	6	117	1F5	PIQVGELLILIRPESELFKHEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,18	0,03	1,03	0,14	0,73	0,06
839	361	3	120	1E9	PIQVGELLILIRPESEPWVLFKQVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,22	0,05	0,78	0,47	0,56	0,13
354	403	3	134	1A4	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNVFKQNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,249	0,02	1,09	0,23	0,74	0,16
722	439	3	146	1C8	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALVFKQLAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,271	0,04	0,58	0,04	0,46	0,21
831	450	1	150	1D9	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFCLNIFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,25	0,05	0,34	0,04	0,43	0,16
789	462	1	154	1G8	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGICLNRIIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,239	0,03	0,64	0,10	0,63	0,26
677	481	3	160	1A8	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLLFKQLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,262	0,04	0,84	0,08	0,87	0,10
543	505	3	168	1D6	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRLFKQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,242	0,01	0,73	0,17	0,57	0,14
342	513	1	171	1G3	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAACLNTAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,257	0,02	0,53	0,08	0,95	0,15
59	532	3	177	1F1	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLVFKQLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,305	0,05	0,40	0,04	0,38	0,12
1030	624	1	208	1D11	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARCLNTRQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,282	0,07	0,18	0,02	0,10	0,01
982	640	3	213	1G10	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARVFKQRGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,266	0,07	0,49	0,06	0,31	0,07
403	649	6	216	1C5	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYVFKHYDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,248	0,02	0,96	0,16	0,91	0,05
772	664	3	221	1F8	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDKNARVFKQRVINENEENKRISIWLDTYYPQLAYYRIHF	0,252	0,02	0,18	0,01	0,25	0,10
1038	675	1	225	1E11	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDKNARVINECLNNENEENKRISIWLDTYYPQLAYYRIHF	0,277	0,08	0,16	0,06	0,11	0,02
378	676	6	225	1F4	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDKNARVINEMFKHENEENKRISIWLDTYYPQLAYYRIHF	0,244	0,03	0,19	0,02	0,17	0,03
127	679	3	226	1C2	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDKNARVINENVFKQNEENKRISIWLDTYYPQLAYYRIHF	0,306	0,05	0,20	0,02	0,12	0,02
377	690	1	230	1E4	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDKNARVINENEENKCLNNKRISIWLDTYYPQLAYYRIHF	0,251	0,02	0,17	0,04	0,11	0,01
964	699	1	233	1E10	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDKNARVINENEENKRISCLNISIWLDTYYPQLAYYRIHF	0,277	0,08	0,15	0,05	0,10	0,00
996	700	6	233	1B11	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDKNARVINENEENKRISMFKHSIWLDTYYPQLAYYRIHF	0,282	0,07	0,20	0,02	0,12	0,02
618	709	3	236	1C7	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDKNARVINENEENKRISIWLVFKQLLDTYYPQLAYYRIHF	0,274	0,04	0,21	0,02	0,11	0,00
390	715	6	238	1A5	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDKNARVINENEENKRISIWLDTLFKHTYYPQLAYYRIHF	0,271	0,03	0,19	0,02	0,16	0,03
1082	717	1	239	1G11	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDKNARVINENEENKRISIWLDTYCLNTYYPQLAYYRIHF	0,28	0,08	0,17	0,02	0,10	0,01

Table 2: Summary of sequence information, OD, SecELISA and StrucELISA data for all strains in the mutagenesis screen (cont.).

Clone Number	Insert Position	Reading Frame	Amino acid	Plate Code	Translation	OD density (620nm)	STD optical density	Average StrucELISA	STD strucELISA	Average SecELISA	STD SecELISA
506	765	1	255	1H5	RK**CLNSK**PVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,284	0,06	0,21	0,02	0,12	0,01
570	801	1	267	1A7	RKPVFWLSRQRNTM**CLNTM**SKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,267	0,01	0,24	0,02	0,27	0,20
314	802	3	267	1E3	RKPVFWLSRQRNTM**MFKQM**SKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,259	0,03	0,91	0,62	0,61	0,17
509	855	1	285	1A6	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPY**CLNTY**ADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,257	0,02	0,18	0,01	0,12	0,04
401	868	3	289	1B5	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSV**MFKQV**NITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,266	0,03	0,19	0,01	0,13	0,03
345	901	3	300	1H3	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAA**VFKQA**GQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,292	0,05	0,20	0,03	0,14	0,01
142	907	6	302	1F2	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQ**VFKHQ**AEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,296	0,02	0,22	0,02	0,13	0,03
379	922	3	307	1G4	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGL**MFKQL**KQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,253	0,02	0,20	0,01	0,13	0,02
870	987	1	329	1G9	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQG**CLNKG**ALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,249	0,06	0,45	0,08	0,41	0,10
81	994	3	331	1H1	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGAL**VFKQL**DDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,304	0,03	0,37	0,04	0,25	0,13
655	1009	3	336	1F7	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEIL**FKQIL**RARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,245	0,04	0,17	0,01	0,10	0,01
723	1050	1	350	1D8	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTW**CLNTW**GGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,251	0,03	0,19	0,02	0,11	0,01
168	1075	6	358	1A3	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFA**MFKHA**IELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,288	0,05	0,20	0,04	0,12	0,03
187	1090	3	363	1B3	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKD**VFKQD**DWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,239	0,03	0,48	0,06	0,64	0,09
1013	1108	3	369	1C11	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGR**LFKQR**SFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,285	0,10	0,35	0,03	1,29	0,04
91	1114	3	371	1A2	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSF**LFKQF**QYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,258	0,04	0,41	0,09	0,43	0,22
569	1119	1	373	1H6	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQY**CLNKY**GAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,218	0,06	0,20	0,02	0,19	0,02
143	1131	1	377	1G2	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEG**CLNKG**YIKMSPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,256	0,01	0,47	0,04	0,40	0,16
905	1143	1	381	1H9	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKM**CLNKM**SPGHWYFPSPLHHHHHHHHHHHHHHHHHH*	0,226	0,05	0,21	0,03	0,35	0,14
566	1158	1	386	1G6	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHW**CLNNW**YFPSPLHHHHHHHHHHHHHHHHHH*	0,204	0,02	0,21	0,03	0,24	0,03
133	1168	6	389	1E2	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFP**MFKHP**SPLHHHHHHHHHHHHHHHHHH*	0,268	0,04	0,48	0,26	0,77	0,17
165	1183	6	394	1H2	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLH**HLFKH**HHHHHHHHHHHHHHHHH*	0,281	0,04	0,54	0,06	0,81	0,20
628	1203	1	401	1D7	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHH**CLNNH**HHHHHHHHH*	0,233	0,03	0,51	0,05	0,74	0,07
389	1209	1	403	1H4	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHH**HCLNNH**HHHHHHHH*	0,263	0,07	0,65	0,10	0,96	0,18
672	1218	1	406	1G7	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHH**CLNNH**HHHH*	0,242	0,03	0,19	0,02	0,59	0,22

Table 3: Summary of sequence information, OD, SecELISA and StrucELISA data for all strains in the mutagenesis screen (cont.).

additional data				Translation	OD density (620nm)	STD optical density	Average StrucELISA	STD strucELISA	Average SecELISA	STD SecELISA
blank					0,073	0,05	0,13	0,03	0,11	0,00
WS006					0,259	0,04	1,15	0,13	1,19	0,18
SB906					0,347	0,08	0,21	0,02	0,17	0,04
baseline				calculated from negative signal + 3x STD			0,27		0,30	
conflicting data										
922	870	1	1B10	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNCLNINITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHH*	0,26	0,10	0,66		0,58	0,38
stop codon mutants										
825	356	2	1C9	METSKEKTITSPGPYIVRLLNSSLNGAEFPLLTGRTLFVVGQSDALTASGQLPDIPADSFFIPLDHGGVNFEIQVDTDATEIILHELKEGNSESRSVQLNTPIQVGELLILIRPESEPCV*TPWVPEQPEKLETS	0,293	0,07	0,21	0,03	0,11	0,01
554	428	2	1E6	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVV*TIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,248	0,02	0,16	0,05	0,09	0,00
1072	722	2	1F11	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGGDYDKNARVINENEENKRISIWLDTYYPV*TYPQLAYYRIHF	0,282	0,08	0,19	0,02	0,11	0,02
536	755	2	1C6	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGGDYDKNARVINENEENKRISIWLDTYYPQLAYYRIHFDV*TDEPRKPVF	0,28	0,02	0,17	0,01	0,10	0,02
54	1004	2	1D1	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDV*TVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHH*	0,341	0,05	0,21	0,03	0,15	0,05
24	1088	2	1B1	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDV*TKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHH*	0,303	0,06	0,21	0,02	0,13	0,04
813	1112	2	1A9	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFV*TSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHHHHHHH*	0,217	0,07	0,23	0,02	0,13	0,01
129	1208	2	1D2	RKPVFWLSRQRNTMSKKELEVLSQKLRALMPYADSVNITLMDDVTAAGQAEAGLKQQALPYSRRNHKGGVTFVIQGALDDVEILRARQFVDSYYRTWGGRYVQFAIELKDDWLKGRSFQYGAEGYIKMSPGHWYFPSPLHHHHHHHHHHV*THHHHHHHHH*	0,238	0,04	0,23	0,02	0,86	0,15
redundant strains										
857	462	1	1F9	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGICLNRIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,222	0,07	0,53	0,03	0,46	0,14
56	624	1	1E1	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARCLNTRQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,321	0,05	0,20	0,07	0,16	0,07
510	624	1	1B6	PIQVGELLILIRPESEPWVPEQPEKLETSAKKNEPRFKNGIVAALAGFFILGIGTVGTLWILNSPQRQAAELDSLLGQEKERFQVLPGRDKMLYVAAQNERDTLWARCLNTRQVLARGDYDDKNARVINENEENKRISIWLDTYYPQLAYYRIHF	0,239	0,02	0,18	0,01	0,13	0,04

Table 4: Summary of sequence information, OD, SecELISA and StrucELISA data for all strains in the mutagenesis screen (cont.).

Of seventy-five strains tested, forty-four were found to completely assemble needle complexes and thirty-seven were actively secreting SptP (Table 1-4). Looking at the cytoplasmic, N-terminal domain (1-140 AA) and the periplasmic, C-terminal domain (140-392 AA, including a predicted transmembrane helix) separately allows a more detailed analysis. The density of mutations was slightly higher for the shorter N-terminal domain (30 strains) compared to the C-terminal domain (45 strains).

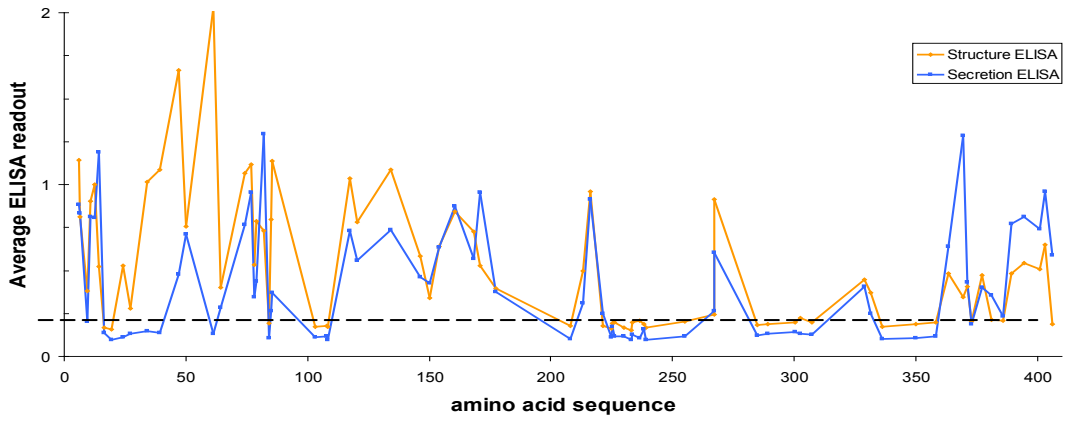


Figure 10: Diagram showing structure and secretion ELISA results for each of the 75 mutations. The dashed line represents the lower level (0.27) above which a strain was considered to form needle filament. For the secretion ELISA the corresponding level was calculated at 0.30, but not drawn here (table 4).

Mutations on the C-terminal domain of PrgH

Of the forty-five strains in the C-terminal domain, twenty were found to assemble needle complexes compared to twenty-five which failed to do so (Table 1-2). Twenty-one strains were secretion competent, and twenty-four failed to secrete SptP. Three strains were found in which a positive signal of one ELISA was not mirrored by the other ELISA.

Two strains, one containing the insert within the his_{18}-tag, one eleven amino acids upstream of the tag, were found to secrete SptP, but no signal could be detected in the structural ELISA. As the length of the his-tag was found to be important early during the development of the structural ELISA, these strains most likely produced

needle complexes unable to bind sufficiently strong to the ELISA plates to allow detection.

One strain (insert Leu331) appeared at first glance to be positive for needle complex assembly, but the readout for the SecELISA remained slightly below the threshold. The strain showed some variation in the secretion activity, as indicated by the corresponding standard deviation. Data obtained from secretion profiles and western blot analysis suggests that the strain is secretion competent, although at a very low level. Given that both StrucELISA and SecELISA results are rather low, the strain most likely only produces a minute amount of needle complexes, but is in principle secretion competent.

Based on the recent, more precise [19, 35] docking of PrgH and PrgK into the overall electron density map, the loop containing the mutation protrudes into the central core of the basal body. With the limited resolution of the currently available model any explanation is bound to be vague, but it appears likely that additional amino acids within the structure could easily cause steric disturbances, apart from changes in local chemical properties of the side-chains. It might be interesting to see if Leu331 of PrgH interacts solely with PrgH, or if the protein is also involved in stabilizing the inner rod or cup region of the complex.

Visualizing the mutations of the on the recently published [18] crystal structure of the truncated C-terminal domain (AA 170-362) showed that those insertions, that still allow needle complex assembly, map to connecting loops. Mutations resulting in a loss of function map onto either alpha-helices or beta-sheets.

Mutations on the N-terminal domain of PrgH affect secretion of SptP

Of the thirty mutant strains covering the N-terminal domain, twenty-four assembled needle complexes while six were unable to do so (Table 2-3). Secretion activity was only found for sixteen of the strains, while for fourteen strains no secretion was detectable. The mutations in the eight strains, which appear to be able to assemble non-functional needle complexes, map to four different areas. One large stretch was located between Leu24 to Val39, a short stretch between Phe61 to Leu64 and two hotspots around Ile9 and His85. Here, at His85, two different inserts showed a

slightly different behavior: CLNIH abolished secretion, while VFKHH still allows very weak secretion to take place.

Similar to the C-terminal domain, mutations not affecting structure assembly or secretion activity mapped to areas between predicted secondary structure elements (data not shown).

Mutations differ in their effect on secretion of SptP, SipB and SipC

To verify the results obtained by the secretion ELISA for SptP, and to test the influence of the mutations on other secreted proteins, secretion profiles were obtained. The complexity of the sample prevented any detailed analysis based on coomassie protein gels and protein fingerprinting by mass spectroscopy (Figure 11), therefore western blotting was used to test for the presence of secreted proteins in the culture supernatant. All strains with mutations located between AA14 to 363 and producing fully assembled complexes according to the StrucELISA, were grown under conditions promoting formation of needle complexes. Proteins from the culture supernatants were precipitated. Western analysis was done using monoclonal antibodies against SptP, SipB and SipC. The readout obtained for SptP mostly confirmed the readout of the SecELISA, with three exceptions that will have to be verified. The strains in question are AA14AA85 and AA117, all of which show no secretion of SptP according to immunodetection. At position AA85, two different inserts were available. One strain contained a CLNIH insertion, the other VFKHH. Originally, both mutations were found to be secretion negative, in later retests the strain containing a VFKHH insert secreted SptP, although very slowly. Similarly, the strain carrying a mutation at position AA14 was originally classified secretion negative by SecELISA data, but later confirmed to be secretion active (Figure 12).

As predicted by the data collected through the SecELISA, SptP secretion was abolished by several mutations between AA24 to 39, AA61-64 and AA 85-117. Surprisingly, other secreted proteins were not affected to the same degree by the mutations.

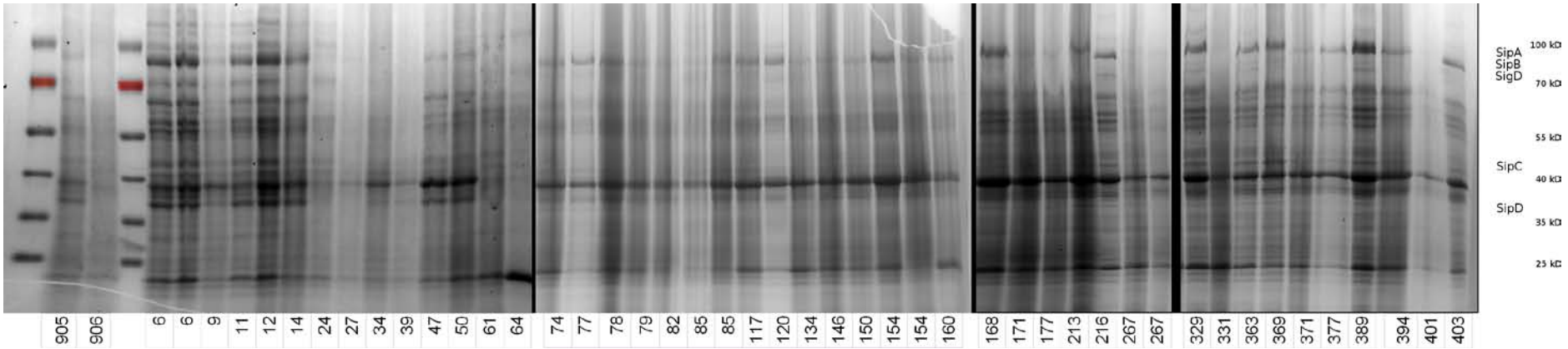


Figure 11: Secretion profiles of all strains capable of assembling the needle complex, according to StrucELISA data. Label: SB905 as positive control, SB906 as negative control, remaining strains are marked by the amino acid position of the insert. Strain 922 (AA280) removed due to conflicting data.

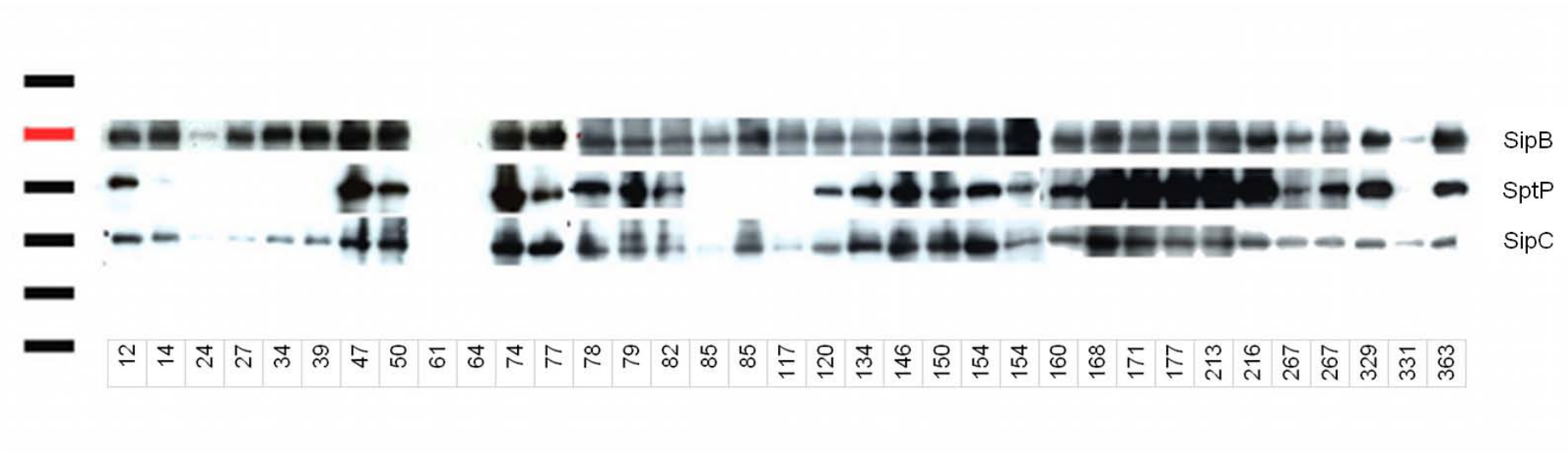


Figure 12: Immunoblots against SipB, SipC and SptP performed on the secretion profiles of all strains capable of assembling the needle complex, according to StrucELISA data. Label: amino acid position of the insert. Strain 922 (AA280) removed due to conflicting data.

SipB could be secreted by all but two strains, although some slight variations in the amount of SipB in the supernatant were detected, judged from the relative intensity of the corresponding bands. The two exceptions were those containing inserts after Phe61 and Leu64. The second strain later showed no assembly of needle filament, thereby explaining the inability to secret any of the three proteins. Given this information, the reason for the high readout of the StrucELISA is unclear.

SipC was similarly secreted by all strains except for those with inserts at Phe61 and Leu64. In contrast to SipB, the amount of secreted SipC varied greatly between different strains, and was hardly detectable in two strains.

The area between Leu 24 and Val 39, as well as the hotspot around His 85 already identified by the ELISA based screen was thereby verified to play a role in secretion. Overall, the mutations as generated here affect the three secreted proteins differently. SipB secretion appears to be fairly robust, whereas SptP appears highly sensitive to mutations, with SipC showing intermediate sensitivity.

DISCUSSION

Fitting PrgH & PrgK into an experimentally verified model of the needle complex

In recent years, a number of 3D structures related to type three secretion systems have become available. Structural data at atomic resolution is available for truncated subunits (as PrgI [38]) and a number of domains derived from PrgH [18], the InvG homolog EscC [18] and the PrgK homolog EscJ [39]. In addition, the structures of needle complexes from *Salmonella typhimurium* [15] and *Shigella flexneri* [13] were published at lower resolutions. Since the needle complex itself is considered to be highly conserved across a number of organisms [40], the attempt to integrate these proteins into a single model despite the varied origin is well accepted. Due to inconclusive [13, 15, 18] data on the copy number of various subunits respectively the overall symmetry of the needle complex, restraints for the model building process have been missing, resulting in a number of conflicting suggestions for the placement of subunits.

Needle complexes obtained from ΔInvG mutant lack the outer ri ng structure, thereby allowing placement of PrgH and PrgK in the remaining, prominent inner ring [35]. By studying the surface accessibility of lysines and through structural restraints created from cross-linking experiments, the topology map for *Salmonella* has already been refined with relative positional data [41]. This allowed locating the C-terminus of PrgH close to the N-terminus of InvG, and the C-terminus of PrgK close to the N-terminus of PrgH. Additionally, and in contrast to PrgK, the surface of PrgH appeared to be accessible for solvents. With the data presented in Figure 5, the orientation of PrgH relative to InvG and PrgK as indicated by cross-linking experiments can be confirmed.

With the orientation and maximum extension of PrgH towards the neck region given by the C-terminal gold label, and a specific label for the loop region around M267, the crystal structure of the C-terminal domain of PrgH can now be fitted into the existing electron density map with high confidence. Using gold labels at the C-terminal end of PrgK to determine the relative orientation and additional data gained from top-views of the inner rings, PrgK could be docked in a ring-like structure

within the encompassing PrgH ring [35]. One of the so far less favored models [18] can therefore be verified and expanded upon, as first published in [35] shown in Figure 13.

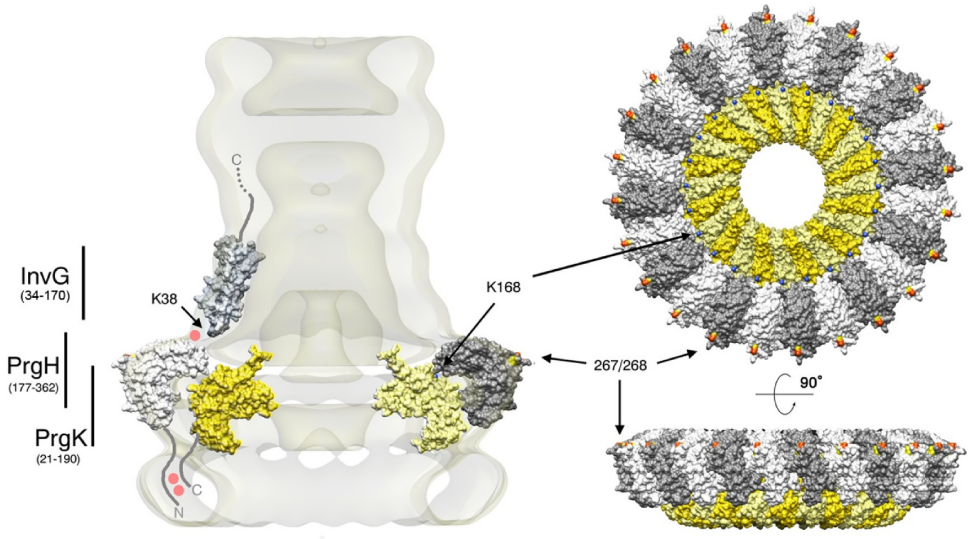


Figure 13: Topology model of the needle complex. PrgH forms the large outer ring and encloses the smaller ring created by PrgK. InvG interacts with its N-terminal domain with the C-terminus of PrgH. Originally published in [35].

Mutagenesis of PrgH

Ratios of ELISA-readouts allow identifying distinct domains in PrgH

With the topology of PrgH firmly established, the effects of short inserts on its various domains could be visualized. For the initial analysis, strains were simply evaluated as showing positive or negative results depending on a defined threshold level, following a nomenclature of (+/+), (-/-) and (+/-) with the first symbol corresponding to the StrucELISA result, and the second to the SecELISA result. (+/-) strains would be fully assembled complexes, with mutations preventing successful secretion of SptP. Tables 1-4 indicate this nomenclature by green (+) and red (-) background of the respective ELISA readout.

Here, the ratio of the StrucELISA to each SecELISA result per strain was calculated and plotted over the protein sequence for all strains with a StrucELISA signal above the threshold level. Incidentally, inclusion of StrucELISA negative strains would not change the result, as the ratio for (-/-) strains varies within a very narrow corridor, as negative results from either ELISA display little variation.

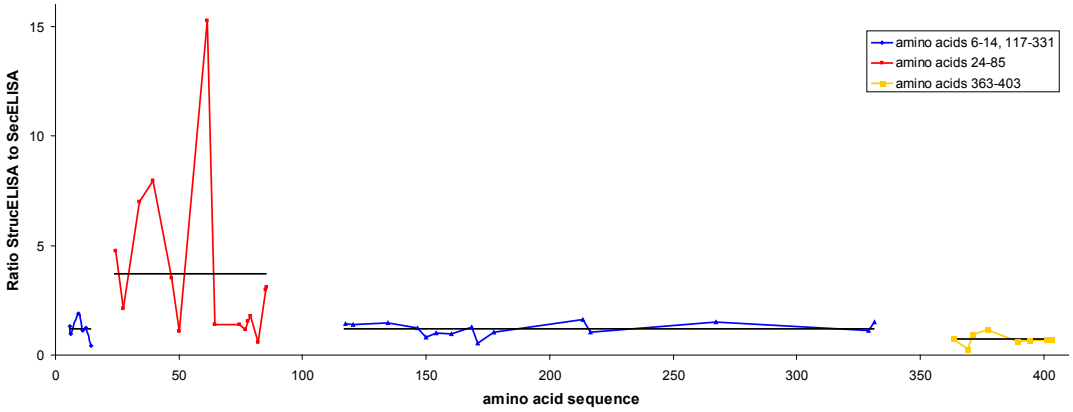
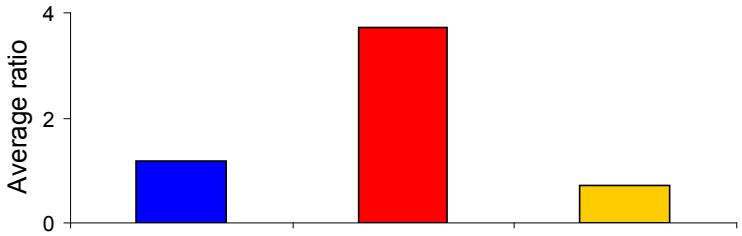


Figure 14: Diagram plotting the ratio of structure ELISA to secretion ELISA. Only strains positive for needle complex assembly were used for the calculation. Outliers were used to mark the beginning and the end of four sequence ranges.



Sequence range	6-14,117-331	24-85	363-403
Average ratio	1,19	3,71*	0,72***
t-Test	1,000	0,024	0,001

Figure 15: Diagram showing the average ratio for each of the four sequence ranges. The first and third sequence ranges were consolidated, considered as baseline and a t-test performed against the second and forth sequence range.

While the absolute value of the ratio is necessarily arbitrary, and cannot be used to compare data collected from independently performed experiments, it can be used to compare strains within this closed data set. A plot showing the calculated ratios over the sequence of PrgH is given in Figure 14. The advantage of this visualization lies in the idea that it removes variations that affect both the SecELISA and the StrucELISA, such as a lower culture density compared to other strains, or complexes that are inefficiently assembled but show normal secretion activity once completed. This type of analysis is commonly used for other assays.

A ratio lower than average can be caused by failure to detect needle complexes, e.g. through mutations close to the his_{18} tag, shorter needle filaments or increased secretion of SptP. A ratio higher than average can be caused by needle complexes binding with higher affinity to the Ni-coated plates, a longer needle filament, unspecific binding of PrgI due to conformational changes or a reduced amount of SptP in the culture supernatant.

Analyzing individual ratios would be unwarrantable, due to the inherent noise. By looking at averages over fairly large areas of the protein, three very distinct domains could nevertheless be identified, with on average significantly higher and lower ratios Figure 15. The start and end of each domain was determined by the first outlier above or below average ratios, and corresponds very well with domains identified in the literature.

The very C-terminal domain beyond residue 362 has been identified as a linker domain before, interacting with N-terminus of InvG in order to connect the inner and outer rings of the base structure [35]. The highly significant drop in the readout for the StrucELISA observed in this domain is therefore most likely related to needle complexes being destabilized [35] by mutants, or changes in the accessibility of the his tag.

The N-terminal domain appears prone to mutations decreasing or abolishing secretion activity, and thereby shifting the ratios on average higher.

Mutations in the N-terminal domain of PrgH affect secretion

Core proteins of the needle complex (InvG, PrgH, PrgK, PrgI, PrgJ) or their respective homologs are regularly recovered from purified needle complexes [31]

and can in principle be considered conserved across multiple organisms. Looked upon in detail, the connection between PrgH and its homologs (YscD, EscD, MxiG, CdsD) appears less clear [6]. While the general organization of the protein with an N-terminal Forkhead-associated-domain (FHA) [42, 43], a transmembrane helix, a C-terminal ring-forming motif [18] and a putative phospholipid-binding domain (BON) [42] appears to be similar, overall sequence conservation is fairly low.

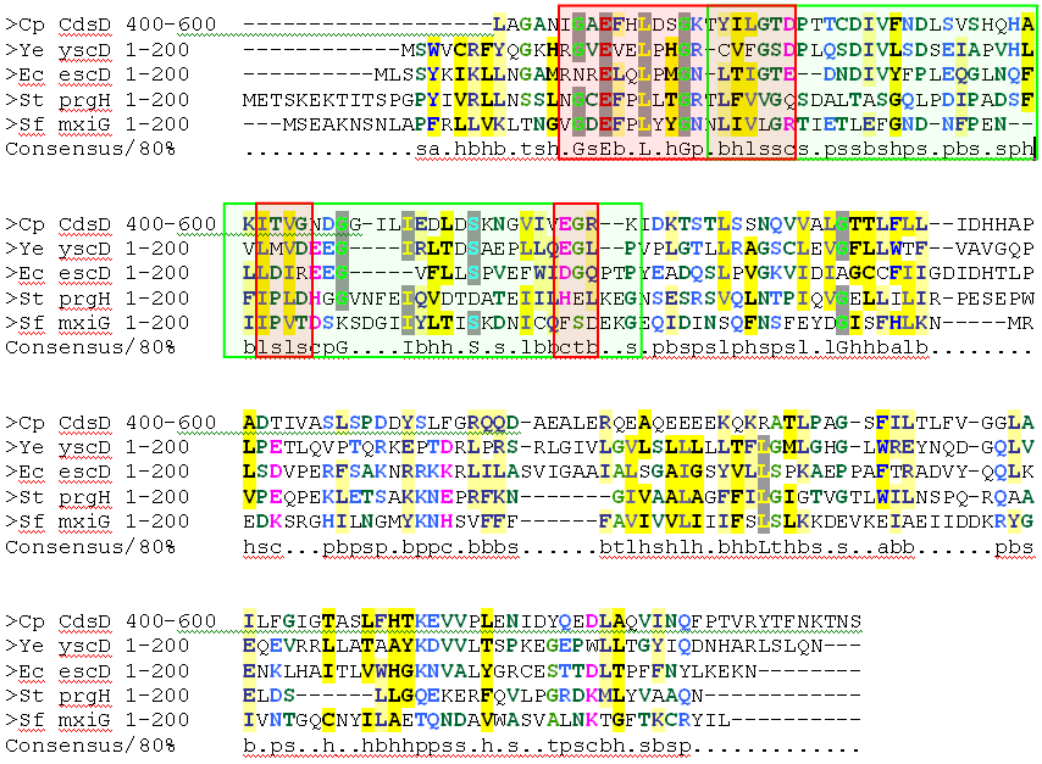


Figure 16: Alignment of the n-terminal domains of various PrgH homologs. Green boxes mark the predicted FHA domain [42], red boxes mark regions likely to produce secretion-negative needle complexes identified in the mutagenesis screen. Sequences were aligned using the ClustalW2 [44] web service provided by EMBL-EBI and formatted using default CHROMA [45] settings: Consensus abbreviations (amino acids): a, aromatic (FHWY, blue lettering on a dark yellow background); b, big (EFHIKLMQRWY, blue on light yellow); h, hydrophobic (ACFGHILMTVWY, black on dark yellow); l, aliphatic (ILV, grey on dark yellow); p, polar (CDEHKNQRST, blue on white); s, small (ACDGNPSTV, dark green on white); t, tiny (AGS, light green on white); -, negatively-charged (DE, red on white); and, +, positively-charged (KR, blue on white), c, charged (DEKRH, pink on white. Organism and gene name abbreviations used: Sf: *Shigella flexneri MxiG (*P0A221), Yp: *Yersinia enterocolitica* YscD (Q93KT0), Ec: *Escherichia coli* EscD (O86193), St: *Salmonella typhimurium* PrgH (P41783), Cp: *Chlamydia pneumoniae* CdsD (Q9Z7J3)

FHA domains were originally mapped to 50-70 amino acids long protein sequences, the domain has since been revised to contain 120-140 amino acids, which would easily include the whole N-terminal domain of PrgH. FHA proteins canonically bind to phosphorylated serine, threonine or tyrosine, with five essential conserved amino acids described, localized in loops between the beta-sheets of the domain [46]. So far, no data on phosphorylated proteins in the type three secretion system of *Salmonella* is available.

In the mutagenesis screen described here, short insertions in the N-terminal domain, specifically in the predicted FHA region of PrgH affected the transport of both proteins considered [2] middle (translocators SipB, SipC) and late (SptP) secretion substrates, although to a different extent. Multiple strains unable to secrete SptP still transported SipB into the culture supernatant, and only showed slightly reduced secretion of SipC (Figure 12).

All of the mutations resulting in (+/-) strains mapped either directly upstream of the predicted FHA domain into a highly conserved region, or directly into the domain itself (Figure 16). Since five amino acid long inserts were used for the mutagenesis presented here, it is difficult to assess if and to what extent conserved residues of the FHA domain were involved in the observed phenotypes. Additional studies, focusing on creating point mutations in the critical areas identified here, and an atomic structure of the relevant domain will be necessary to uncover the exact interaction site.

Possible roles of the N-terminal domain of PrgH

Variations in the secretion of multiple proteins as observed in the strains obtained from the mutagenesis screen might be explained by direct interactions between the chaperones/effectors and PrgH. The domain was located in the cytoplasm and showed enough freedom of movement to suggest a specific role in supplying bound effectors to a hypothetical ATPase complex nearby. The specific impact of a mutation would then depend on the affinity and binding surface for each protein, with stronger interactions being harder to disrupt than weaker interactions. Variations in the affinity of effectors to the secretion machinery have been observed and modeled [26] to be sufficient to cause hierarchical secretion. Since the translocators

SipB and SipC are secreted earlier than the late effector SptP, this working model would explain the data very well. Unfortunately, preliminary experiments failed to show interaction between a number of effectors and PrgH (data not shown).

An extensive review of the literature did not return any known interactions between homologs of PrgH and effectors or chaperones, although another potential interaction partner was discovered, as discussed below.

An additional model closely fitting the data could be constructed based on the idea of a generally slowed secretion process, without specific effect on any secretion substrate. If proteins were secreted in a hierarchical manner, a secretion apparatus hampered by a severe mutation would not be able to progress towards the later repertoire of effector proteins and would show diminished levels of these proteins in the culture supernatant. In *Salmonella typhimurium*, the effector pool of a cell can be moved through the type three secretion system within minutes of contact with a host cell [26]. Secretion tests were performed over several hours, offering some protection against this problem. At this point, this model cannot be ruled out, until the effect of the mutation against additional secretion substrates can be studied.

C-ring proteins interact with PrgH homologs

The type three secretion system consists of the needle complex and a cytoplasmic bulb, or c-ring [6]. This additional structure contains the ATPase InvC [47] and regulatory proteins such as SpaO, OrgA and OrgB, and is conserved in multiple organisms. Strains lacking InvC or other proteins of the c-ring fail to assemble needle filaments and are unable to secrete early substrates as PrgI [20]. Since most of the (+/-) mutants appeared to be fully assembled (with the exception of mutation AA61), it appeared unlikely that the interaction between the needle complex and the c-ring was completely disrupted.

Figure 17 summarized the data obtained from aggregating multiple yeast2hybrid screens performed in different organisms. In *Shigella flexneri,* interactions between the core c-ring proteins have been mapped exhaustingly. The complete interaction map includes interactions between MxiG (PrgH) and MxiJ (PrgK) to Spa33 (SpaO). The latter one interacts with the ATPase Spa47 (InvC), which is bound by the

regulatory proteins MxiK and MxiN (OrgA, OrgB). Interactions between Spa33 and middle or late effector proteins were also shown [32, 48].

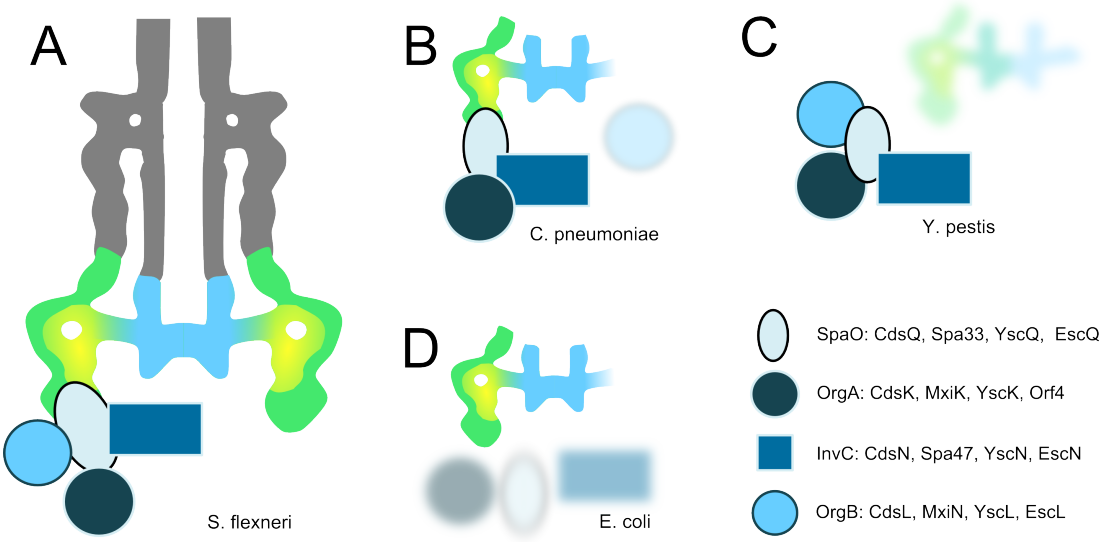


Figure 17: Illustration of interactions between basal and cytoplasmic proteins of the type 3 secretion system according to current literature [32, 48-52]. Published data from yeast2hybrid screens showing interactions between several conserved proteins belonging to the type three secretion system in five different organisms**.**

Studies in other organisms are less comprehensive, but all interactions shown so far follow the pattern laid out in *Shigella flexneri,* and suggest a conserved mechanism. In *Yersinia pestis* interactions of the ATPase YscN with SpaO homolog YscQ, and of OrgA/B homologs YscL/N with YscQ were shown [50]. In *Chlamydophila pneumoniae* interactions between the potential distant PrgH homolog CdsD, CdsQ (SpaO) and CdsL (OrgB) [51] were shown, and additional interactions between those three proteins and the ATPase, specifically the first 105 amino acids of CdsN [52]. Overall, the picture here is less clear, probably related to an additional 400 AA long domain in CdsD, showing a second FHA domain. In *Escherichia coli* and *Salmonella typhimurium*, no interaction between the c-ring proteins and the needle complex were shown so far.

Current models cannot explain the phenotype of FHA domain mutants

The earlier described model on direct interactions between chaperones/effectors with PrgH would readily explain the observed phenotypes, but is most likely incorrect. Given the current understanding of the cytoplasmic, peripheral type three secretion system protein network, no conclusive mechanistic explanation for the interplay between the FHA domain of PrgH and c-ring proteins, or even the C-terminal domain of PrgK can be given. In principle, the three hotspots identified could correspond to binding sites for several separate proteins.

Besides the C-terminus of PrgK, SpaO is the likeliest candidate for direct interactions with PrgH. Mutations preventing binding of SpaO to PrgH should result in a phenotype similar to a knockout of SpaO, with no secretion of either early, middle or late proteins. Despite high readouts from the StrucELISA, strains with inserts at position AA61 and AA65 show no secretion activity for middle to late substrates, and no needle filament could be detected on purified complexes by electron microscopy from strain AA61 (AA65 was not yet analyzed). This might suggest either a critical binding site in this position, or disruption of the protein fold due to the mutation.

The other mutations between AA24-39 and around AA85 appeared to specifically select against secretion of late (effector) substrates, but still allowed secretion of translocators. So far, two proteins were linked to phenotypes relevant to translocator secretion.

Mutants of InvE [53] in *Salmonella* and YscU in *Yersinia [54, 55]* will secrete effector proteins, but not translocators. YscU is an autocatalytic protein, whose cleavage is required to switch the needle complex from secreting early stage proteins to translocators. An inactive $YscU_{N263A}$ strain therefore ignores the secretion switch towards translocators and proceeds directly to secrete late substrates. InvE has been shown to interact with the chaperone-effector complex SicA/SipB and SicA/SipC, and appears to bind to a membrane-associated protein [53].

Given the data presented so far, it might be possible to present a hypothetical model in which two partly independent mechanisms exist for the recognition of middle and late secretion substrates. The translocator/chaperone complexes (SipB and

SipC/SicA) would here be dependent on the activity of the YscU homolog SpaS and binding of InvE to the needle complex or c-ring. This interaction could either occur at independent or only partially overlapping sites with high affinity compared to the recognition of late chaperone/effector, or to separate binding sites formed by as of yet not identified regulatory proteins. InvE only interacts with the chaperone-effector complex, and not with either alone. Hypothetically, this could lead to self-regulation once the translocator pool is depleted, as InvE would run out of binding partners and dissociate from the needle complex.

In this model, the mutations in the FHA domain of PrgH would disrupt the interaction sites of protein/chaperone complexes of late substrates as SptP/SicP on the ATPase or the c-ring in general, possibly by not allowing interactions with regulatory proteins or through a general destabilization of the c-ring. The successful secretion of translocator proteins on the other hand would be caused by the use of separate recognition sites or through stabilization of the interaction between the c-ring and the needle complex by additional membrane associated proteins.

Outlook

To further understand the complex regulation of the type three secretion process, it will be necessary to approach a number of research questions opened up by the work presented here.

So far, no interaction of the N-terminal domain of PrgH with a cytoplasmic protein was shown. The most promising approach appears to be identifying interactions in vitro, most likely based on the candidates discussed here. Once an interaction partner is identified, the interaction will have to be characterized in detail, to answer the questions in regard to the identified hotspots representing independent binding sites. While the usage of short, five amino acid long insert proved to be a very useful compromise to keep the number of strains low while maintaining good coverage, the effects of the inserts on the protein fold shouldn't be neglected. It would be preferable to understand the role of individual amino acid residues by screening a library of point mutations, especially since the areas of interest have already been identified and the necessary tools for high throughput screening are in place.

As the topology of PrgH in the needle complex is now firmly established, its potential capacity to link the basal body to the complex cytoplasmic machinery is a surprising discovery. So far, structural studies have focused on the prominent needle complex, with hardly any available structural data on proteins forming the c-ring. Ultimately, in order to understand the complex machinery evolved to aid the survival of bacteria in a hostile environment, these gaps will have to be filled.

PLASMIDS AND STRAINS

Strain	Plasmids	Construct	backbone	AB	cloning sites
SB906	pWS006	PrgH, 500bp upstream, 18xHis C-terminal, C27A, BsaHI site upstream TATA	pwsk29	AMP	HindIII, XmaI
SB906	pWS007	PrgH, 500bp upstream, 12xHis N-terminal	pwsk29	AMP	PstI, NotI
SB906	PWS008	PrgH, 500bp upstream, 14xHis N-terminal	pwsk29	AMP	PstI, NotI
SB906	pWS009	PrgH, 500bp upstream, 18xHis N-terminal	pwsk29	AMP	PstI, NotI
SB906	pWS014	PrgH, 500bp upstream, 6xHIS N-terminal	pwsk29	AMP	PstI, NotI
SB906	pWS016	PrgH, 500bp upstream. NdeI & BamHI site for cloning of upstream elements	WS200s	AMP	PstI, NotI
SB906	pWS017	eGFP + linker cloned between NdeI/BamHI of WS016	pWS016	AMP	PstI, NotI
SB906	pWS018	PrgH, 500bp upstream, 6xHIS-PrgH after M267	pwsk29	AMP	PstI, NotI
SB906	pWS019	PrgH, 500bp upstream, GSG-6xHIS-GSG-PrgH after M267	pwsk29	AMP	PstI, NotI
SB906	pWS020	PrgH, 500bp upstream, 12xHIS-PrgH after M267	pwsk29	AMP	PstI, NotI
SB906	pWS021	InvC-PrgH-fusion, 500bp upstream, cloned into pWS016	pwsk29	AMP	PstI, NotI
SB917	pWS021	InvC-PrgH-fusion, 500bp upstream, cloned into pWS016	pwsk29	AMP	PstI, NotI
SB905	pPrgI5	PrgI, ~900bp upstream	pwsk29	AMP	PstI, NotI

Table 5: Overview of all strains and plasmids used in this study. SB905: non-flagellated wild-type, SB906: non-flagellated ΔprgH, SB917: non-flagellated ΔinvC [20]. AMP : Ampicillin. All strains carried an additional plasmid containing the transcriptional regulator hilA under the araBAD promoter with an chloramphenicol (CMP) resistance marker.

REFERENCES

1. Hueck, C.J., *Type III protein secretion systems in bacterial pathogens of animals and plants.* Microbiol Mol Biol Rev, 1998. **62**(2): p. 379-433.
2. Deane, J.E., et al., *Timing is everything: the regulation of type III secretion.* Cell Mol Life Sci, 2009.
3. McClelland, M., et al., *Complete genome sequence of Salmonella enterica serovar Typhimurium LT2.* Nature, 2001. **413**(6858): p. 852-6.
4. Marcus, S.L., et al., *Salmonella pathogenicity islands: big virulence in small packages.* Microbes Infect, 2000. **2**(2): p. 145-56.
5. Zhao, Y., *Virulence factors of Salmonella enterica serovar Enteritidis*, in *Department of Infectious Diseases and Immunology*. 2002, Universiteit Utrecht: Utrecht. p. 90.
6. Cornelis, G.R., *The type III secretion injectisome.* Nat Rev Microbiol, 2006. **4**(11): p. 811-25.
7. Galan, J.E. and R. Curtiss, 3rd, *Cloning and molecular characterization of genes whose products allow Salmonella typhimurium to penetrate tissue culture cells.* Proc Natl Acad Sci U S A, 1989. **86**(16): p. 6383-7.
8. Mills, D.M., V. Bajaj, and C.A. Lee, *A 40 kb chromosomal fragment encoding Salmonella typhimurium invasion genes is absent from the corresponding region of the Escherichia coli K-12 chromosome.* Mol Microbiol, 1995. **15**(4): p. 749-59.
9. Hensel, M., et al., *Simultaneous identification of bacterial virulence genes by negative selection.* Science, 1995. **269**(5222): p. 400-3.
10. Shea, J.E., et al., *Identification of a virulence locus encoding a second type III secretion system in Salmonella typhimurium.* Proc Natl Acad Sci U S A, 1996. **93**(6): p. 2593-7.
11. Ochman, H., et al., *Identification of a pathogenicity island required for Salmonella survival in host cells.* Proc Natl Acad Sci U S A, 1996. **93**(15): p. 7800-4.
12. Kubori, T., et al., *Supramolecular structure of the Salmonella typhimurium type III protein secretion system.* Science, 1998. **280**(5363): p. 602-5.
13. Hodgkinson, J.L., et al., *Three-dimensional reconstruction of the Shigella T3SS transmembrane regions reveals 12-fold symmetry and novel features throughout.* Nat Struct Mol Biol, 2009. **16**(5): p. 477-85.
14. Marlovits, T.C., et al., *Assembly of the inner rod determines needle length in the type III secretion injectisome.* Nature, 2006. **441**(7093): p. 637-40.
15. Marlovits, T.C., et al., *Structural insights into the assembly of the type III secretion needle complex.* Science, 2004. **306**(5698): p. 1040-2.
16. Sani, M., et al., *Structural organization of the needle complex of the type III secretion apparatus of Shigella flexneri.* Micron, 2007. **38**(3): p. 291-301.
17. Ogino, T., et al., *Assembly of the type III secretion apparatus of enteropathogenic Escherichia coli.* J Bacteriol, 2006. **188**(8): p. 2801-11.

18. Spreter, T., et al., *A conserved structural motif mediates formation of the periplasmic rings in the type III secretion system.* Nat Struct Mol Biol, 2009. **16**(5): p. 468-76.

19. Galkin, V.E., et al., *The Structure of the Salmonella typhimurium Type III Secretion System Needle Shows Divergence from the Flagellar System.* J Mol Biol, 2010. **396**(5): p. 1392-1397.

20. Sukhan, A., et al., *Genetic analysis of assembly of the Salmonella enterica serovar Typhimurium type III secretion-associated needle complex.* J Bacteriol, 2001. **183**(4): p. 1159-67.

21. Moraes, T.F., T. Spreter, and N.C. Strynadka, *Piecing together the type III injectisome of bacterial pathogens.* Curr Opin Struct Biol, 2008. **18**(2): p. 258-66.

22. Akeda, Y. and J.E. Galan, *Chaperone release and unfolding of substrates in type III secretion.* Nature, 2005. **437**(7060): p. 911-5.

23. Akeda, Y. and J.E. Galan, *Genetic analysis of the Salmonella enterica type III secretion-associated ATPase InvC defines discrete functional domains.* J Bacteriol, 2004. **186**(8): p. 2402-12.

24. Galan, J.E. and H. Wolf-Watz, *Protein delivery into eukaryotic cells by type III secretion machines.* Nature, 2006. **444**(7119): p. 567-73.

25. Deane, J.E., et al., *Crystal structure of Spa40, the specificity switch for the Shigella flexneri type III secretion system.* Mol Microbiol, 2008. **69**(1): p. 267-76.

26. Winnen, B., et al., *Hierarchical effector protein transport by the Salmonella Typhimurium SPI-1 type III secretion system.* PLoS One, 2008. **3**(5): p. e2178.

27. Collazo, C.M. and J.E. Galan, *Requirement for exported proteins in secretion through the invasion-associated type III system of Salmonella typhimurium.* Infect Immun, 1996. **64**(9): p. 3524-31.

28. Fu, Y. and J.E. Galan, *A salmonella protein antagonizes Rac-1 and Cdc42 to mediate host-cell recovery after bacterial invasion.* Nature, 1999. **401**(6750): p. 293-7.

29. van Heel, M., et al., *Single-particle electron cryo-microscopy: towards atomic resolution.* Q Rev Biophys, 2000. **33**(4): p. 307-69.

30. Chandran, V., et al., *Structure of the outer membrane complex of a type IV secretion system.* Nature, 2009. **462**(7276): p. 1011-5.

31. Kimbrough, T.G. and S.I. Miller, *Assembly of the type III secretion needle complex of Salmonella typhimurium.* Microbes Infect, 2002. **4**(1): p. 75-82.

32. Morita-Ishihara, T., et al., *Shigella Spa33 is an essential C-ring component of type III secretion machinery.* J Biol Chem, 2006. **281**(1): p. 599-607.

33. Mülhardt, C., *Molekularbiologie/Genomics*. 4. ed. Der Experimentator. 2003, München: Elsevier GmbH. 280.

34. Balkwill, F., J. Roskams, and L. Rogers, *Lab Ref: A Handbook of Recipes, Reagents, and Other Reference Tools for Use at the Bench.* 2002: Cold Spring Harbor Laboratory Press.

35. Schraidt, O., et al., *Topology and organization of the Salmonella typhimurium type III secretion needle complex components.* PloS Pathogens, 2010. **in press**.

36. Kaniga, K., J.C. Bossio, and J.E. Galan, *The Salmonella typhimurium invasion genes invF and invG encode homologues of the AraC and PulD family of proteins.* Mol Microbiol, 1994. **13**(4): p. 555-68.

37. Radics, J., *The role of the integral protein PrgH within the Salmonella SPI-1 TIIISS*. 2009, Universität für Bodenkultur: Vienna.

38. Wang, Y., et al., *Differences in the electrostatic surfaces of the type III secretion needle proteins PrgI, BsaL, and MxiH.* J Mol Biol, 2007. **371**(5): p. 1304-14.

39. Yip, C.K., et al., *Structural characterization of the molecular platform for type III secretion system assembly.* Nature, 2005. **435**(7042): p. 702-7.

40. Pallen, M.J., S.A. Beatson, and C.M. Bailey, *Bioinformatics, genomics and evolution of non-flagellar type-III secretion systems: a Darwinian perspective.* FEMS Microbiol Rev, 2005. **29**(2): p. 201-29.

41. Brunner, M.J., *Structural studies of the Salmonella SPI-1 type III secretion apparatus*. 2009, University Vienna: Vienna.

42. Pallen, M.J., S.A. Beatson, and C.M. Bailey, *Bioinformatics analysis of the locus for enterocyte effacement provides novel insights into type-III secretion.* BMC Microbiol, 2005. **5**: p. 9.

43. Pallen, M., R. Chaudhuri, and A. Khan, *Bacterial FHA domains: neglected players in the phospho-threonine signalling game?* Trends Microbiol, 2002. **10**(12): p. 556-63.

44. Thompson, J.D., T.J. Gibson, and D.G. Higgins, *Multiple sequence alignment using ClustalW and ClustalX.* Curr Protoc Bioinformatics, 2002. **Chapter 2**: p. Unit 2 3.

45. Goodstadt, L. and C.P. Ponting, *CHROMA: consensus-based colouring of multiple alignments for publication.* Bioinformatics, 2001. **17**(9): p. 845-6.

46. Li, J., et al., *The FHA domain mediates phosphoprotein interactions.* J Cell Sci, 2000. **113 Pt 23**: p. 4143-9.

47. Galan, J.E., *Energizing type III secretion machines: what is the fuel?* Nat Struct Mol Biol, 2008. **15**(2): p. 127-8.

48. Jouihri, N., et al., *MxiK and MxiN interact with the Spa47 ATPase and are required for transit of the needle components MxiH and MxiI, but not of Ipa proteins, through the type III secretion apparatus of Shigella flexneri.* Mol Microbiol, 2003. **49**(3): p. 755-67.

49. Creasey, E.A., et al., *Yeast two-hybrid system survey of interactions between LEE-encoded proteins of enteropathogenic Escherichia coli.* Microbiology, 2003. **149**(Pt 8): p. 2093-106.

50. Jackson, M.W. and G.V. Plano, *Interactions between type III secretion apparatus components from Yersinia pestis detected using the yeast two-hybrid system.* FEMS Microbiol Lett, 2000. **186**(1): p. 85-90.

51. Johnson, D.L., C.B. Stone, and J.B. Mahony, *Interactions between CdsD, CdsQ, and CdsL, three putative Chlamydophila pneumoniae type III secretion proteins.* J Bacteriol, 2008. **190**(8): p. 2972-80.

52. Stone, C.B., et al., *Characterization of the putative type III secretion ATPase CdsN (Cpn0707) of Chlamydophila pneumoniae.* J Bacteriol, 2008. **190**(20): p. 6580-8.

53. Kubori, T. and J.E. Galan, *Salmonella type III secretion-associated protein InvE controls translocation of effector proteins into host cells.* J Bacteriol, 2002. **184**(17): p. 4699-708.

54. Sorg, I., et al., *YscU recognizes translocators as export substrates of the Yersinia injectisome.* EMBO J, 2007. **26**(12): p. 3015-24.

55. Wiesand, U., et al., *Structure of the type III secretion recognition protein YscU from Yersinia enterocolitica.* J Mol Biol, 2009. **385**(3): p. 854-66.

APPENDIX

The Structure of the Salmonella typhimurium Type III Secretion System Needle Shows Divergence from the Flagellar System.

Galkin* VE, Schmied* WH, Schraidt O, Marlovits TC, Egelman EH

*equal contribution

Reprinted from Journal of Molecular Biology, 2010 Mar 12;396(5):1392-7, "The structure of the Salmonella typhimurium type III secretion system needle shows divergence from the flagellar system.", Galkin* VE, Schmied* WH, Schraidt O, Marlovits TC, Egelman EH, Copyright (2010), with permission from Elsevier.

doi:10.1016/j.jmb.2010.01.001 J. Mol. Biol. (2010) 396, 1392–1397

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

The Structure of the *Salmonella typhimurium* Type III Secretion System Needle Shows Divergence from the Flagellar System

Vitold E. Galkin[1]†, Wolfgang H. Schmied[2,3]†, Oliver Schraidt[2,3], Thomas C. Marlovits[2,3]* and Edward H. Egelman[1]*

[1]*Department of Biochemistry and Molecular Genetics, University of Virginia Medical School, Charlottesville, VA 22908-0733, USA*

[2]*Research Institute of Molecular Pathology, Dr. Bohr Gasse 7, A-1030 Vienna, Austria*

[3]*Institute of Molecular Biotechnology, Austrian Academy of Sciences, Dr. Bohr Gasse 3-5, A-1030 Vienna, Austria*

Received 18 November 2009; received in revised form 22 December 2009; accepted 1 January 2010
Available online 11 January 2010

Edited by R. Craig

The type III secretion system (T3SS) is essential for the infectivity of many pathogenic Gram-negative bacteria. The T3SS contains proteins that form a channel in the inner and outer bacterial membranes, as well as an extracellular needle that is used for transporting and injecting effector proteins into a host cell. The homology between the T3SS and the bacterial flagellar system has been firmly established, based upon both sequence similarities between respective proteins in the two systems and the structural homology of higher-order assemblies. It has previously been shown that the *Shigella flexneri* needle has a helical symmetry of ~5.6 subunits/turn, which is quite similar to that of the most intensively studied flagellar filament (from *Salmonella typhimurium*), which has ~5.5 subunits/turn. We now show that the *Sa. typhimurium* needle, expected by homology arguments to be more similar to the *Sa. typhimurium* flagellar filament than is the needle from *Shigella*, actually has ~6.3 subunits/turn. It is not currently understood how host cell contact, made at the tip of the needle, is communicated to the secretory system at the base. In contrast to the *Sa. typhimurium* flagellar filament, which shows a nearly crystalline order, the *Sa. typhimurium* needle has a highly variable symmetry, which could be used to transmit information about host cell contact.

Keywords: cryo-electron microscopy; helical reconstruction; protein polymers; bacterial pathogenesis; structural polymorphism

Introduction

Many pathogenic Gram-negative bacteria contain type III secretion systems (T3SSs) that are essential for infectivity.[1] These T3SSs serve as both a secretion system and an injection system for bacterial effector proteins that need to be passed through the inner and outer bacterial membranes and into a host cell. The strong morphological and sequence similarities between the T3SS and the bacterial flagellar system gave rise to the suggestion that these systems are homologs, having common origins.[2,3] Consistent with this, it was shown[4] that the needle of the *Shigella flexneri* T3SS has approximately the same helical symmetry (5.6 subunits/turn of a 24-Å-pitch helix) as the best-characterized flagellar filament[5]—that from *Salmonella typhimurium* (~5.5 subunits/turn of a 26-Å-pitch helix). Since contact between the bacterium and the host cell has been shown to switch the T3SS into a secretory state, attempts have been made to find changes in the helical structure of the *Sh. flexneri* needle with mutations in the needle protein MxiH, which lock the T3SS into a constitutive "on" state.[6] Given the homology between the T3SS and the flagellar system, it was thought that a switching of the state of the needle might occur similarly to the switching of the flagellar filament from a left-handed state to a right-handed state of protofilaments.[7] These studies[6] failed to find any

**Corresponding authors.* E-mail addresses: marlovits@imp.ac.at; egelman@virginia.edu.

† V.E.G. and W.H.S. contributed equally to this work.

Abbreviations used: T3SS, type III secretion system; cryo-EM, cryo-electron microscopy; IHRSR, iterative helical real space reconstruction; STEM, scanning transmission electron microscopy; EDTA, ethylenediaminetetraacetic acid; LDAO, lauryldimethylamine oxide.

changes in average helical parameters, leading to the suggestion that such changes in helical symmetry might be too small to be detected by the analysis methods used, or that signal transduction in this system occurs through some novel mechanism that has not yet been characterized.

Results

Sa. typhimurium needles have been imaged by both negative staining (Fig. 1d) and cryo-electron microscopy (cryo-EM) (Fig. 1a and g). Fields of needles imaged by cryo-EM (Fig. 1a) were used to generate an averaged power spectrum (Fig. 1b) that is quite similar to a fiber diffraction pattern, as it shows no sampling due to an ordered arrangement of filaments. This power spectrum shows only three layer lines: one at ~1/(26 Å) arising from a 1-start helix, a more diffuse one at ~1/(38 Å), and a very diffuse one at ~1/(85 Å). The broadening of these other layer lines is a characteristic feature of variability in the twist of a helical structure,[8] where helical symmetry is not exactly conserved between every adjacent subunit. As a result, deviations in twist accumulate along a filament, and long-range order does not exist. Indexing these two more diffuse layer lines was problematic, as multiple solutions consistent with the diameter of these particles are possible. The layer lines at 1/(38 Å) and 1/(85 Å) could have any Bessel order less than 8. The power spectrum at this resolution simply does not uniquely specify the helical symmetry due to the fact that, for many structures at limited resolution, there are intrinsic ambiguities in determining helical symmetry.[9] Similarly, multiple solutions can be obtained with the iterative helical real space reconstruction (IHRSR) method.[10] For example, a scheme with 1/(38 Å) being $n=-3$ and with 1/(85 Å) being $n=4$ (3.7 units/turn), and a scheme with 1/(38 Å) being $n=7$ and with 1/(85 Å) being $n=-6$ (6.3 units/turn) both led to stable solutions, with reconstructions generating power spectra indistinguishable from the power spectra generated from

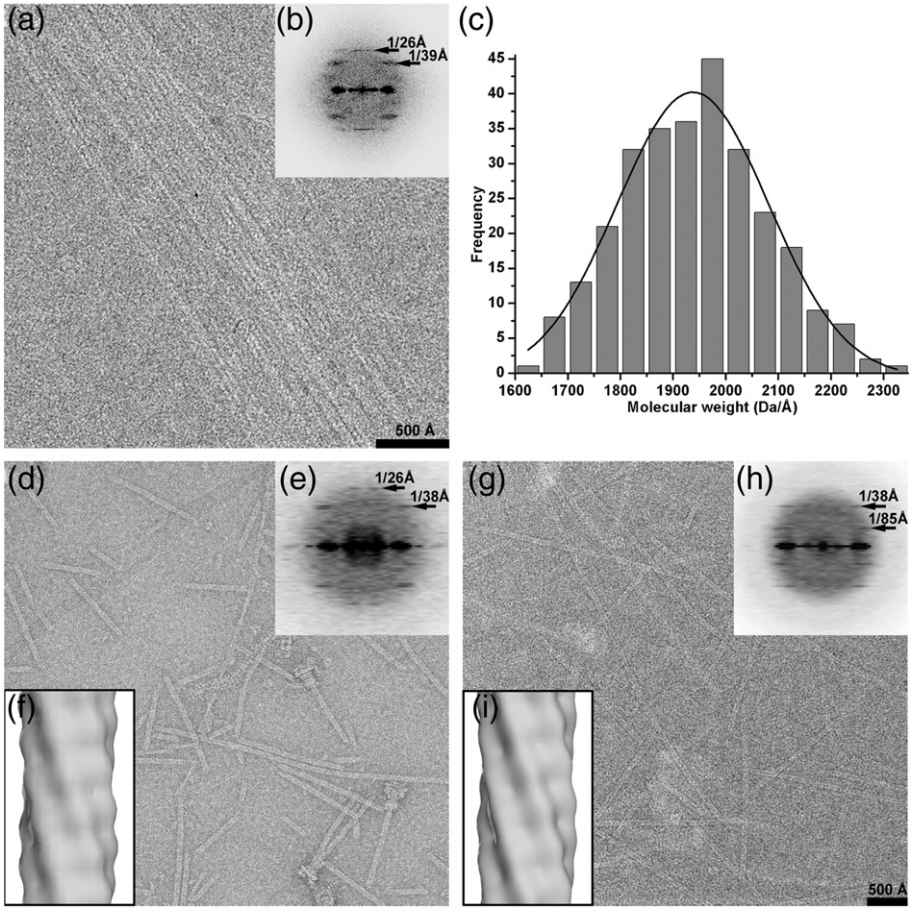


Fig. 1. (a) Cryo-EM image of bundles of needles. (b) Averaged power spectrum generated from 14 nonoverlapping boxes (512×512 pixels, 5.9 Å/pixel) taken from bundles such as in (a). (c) STEM mass histogram, with a Gaussian curve fitted. (d) Negatively stained needles. (e) Averaged negative-stain power spectrum from 259 nonoverlapping boxes, each 480 pixel long (1.33 Å/pixel). (f) Overall IHRSR three-dimensional negative-stain reconstruction derived from 5420 segments (120 pixels long, 5.32 Å/pixel), filtered to 24 Å resolution. This reconstruction converged to a helical symmetry of 57° rotation and an axial rise of 4.2 Å, corresponding to 6.3 subunits/turn of a 26.5-Å-pitch helix. (g) Cryo-EM of isolated filaments. (h) Averaged cryo-EM power spectrum from 850 nonoverlapping boxes, each 480 pixels long (1.33 Å/pixel). (i) Overall IHRSR cryo-EM three-dimensional reconstruction derived from 22,834 segments (120 pixels long, 2.66 Å/pixel), filtered to 24 Å resolution. This converged to the same symmetry as the negative-stain reconstruction (57°/4.2 Å).

the images. We therefore used scanning transmission electron microscopy (STEM) to obtain an accurate estimate of mass per unit length in these needles.[11] STEM measurements indicated that the average mass per length was ~2.0 kDa/Å (Fig. 1c). Since the component protein of the *Sa. typhimurium* needle, PrgI, contains 80 residues and is 8.4 kDa, this would suggest that the axial rise per subunit is, on average, ~4.2 Å, and that there are, on average, ~6.2 subunits/turn of the 26-Å-pitch 1-start helix.

This knowledge of mass per length was then used to generate overall reconstructions of the needles from both negative staining (Fig. 1f) and cryo-EM (Fig. 1i). The failure of the IHRSR method to converge to the same solution from different starting symmetries, as well as the broad layer lines seen in the averaged power spectrum from bundles (Fig. 1b), strongly suggested that these samples were heterogeneous with respect to helical parameters. Power spectra computed from subsets of the filaments after sorting (Supplementary Movie 1) show the very large variability in twist, with a range that extends beyond 6.2–6.5 subunits/turn. Supplementary Movie 1 displays a fixed $n=7$ layer line at ~1/(38 Å), with a very variable $n=-6$ layer line.

The best reconstruction (Fig. 2a) was obtained from the central bin (6.3 subunits/turn) of the cryo-EM filaments, using 9366 segments from a total of 22,834 segments. The most conservative estimate[14] of the resolution of this reconstruction is 19.5 Å, using a Fourier shell correlation of 0.5, and two completely independent reconstructions were generated from different starting symmetries and using completely different images. This actually provides an estimate of the resolution for a reconstruction containing half of the total images and must therefore be an underestimate of the resolution of the full data set. Filtering an atomic model of the two-helix bundle within the PrgI subunit[12] (residues 18–60; Fig. 2a) to 18 Å resolution provides a good

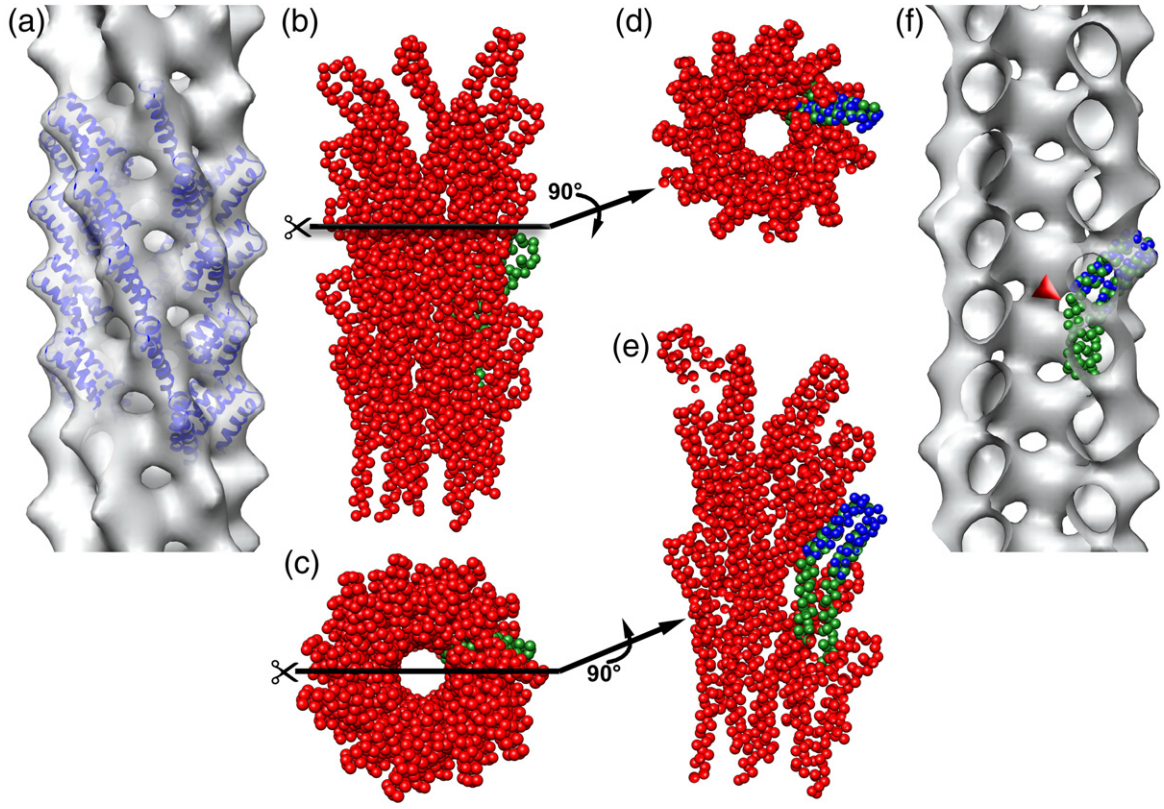


Fig. 2. (a) Reconstruction of the frozen hydrated filaments filtered to 18 Å resolution, using 9366 segments from a total of 22,834 segments. Sorting was based on variable twist and axial rise while keeping the $n=+7$ layer line fixed at 1/(38 Å). This class from the center of the distribution yielded a symmetry of 57°/4.2 Å. Residues 18–60 (blue ribbons) from the solution structure[12] of the *Sa. typhimurium* needle subunit (Protein Data Bank entry 2JOW) were manually docked into the reconstruction. (b) Side view of the atomic model[13] of the *Sh. flexneri* T3SS needle (Protein Data Bank entry 2V6L), with one of the protomers shown in green, and (c) top view of this model. (d) The atomic model of the *Sh. flexneri* T3SS needle in (b) cut along the plane perpendicular to the helical axis (black line). One of the protomers (green) was aligned with residues 18–60 of the *Sa. typhimurium* needle subunit (blue), positioned as fitted to the reconstruction shown in (a). Despite the difference in symmetry, the paired helices of the needle protomers from both bacteria are located at the same radius and have the same orientation. (e) The atomic model of the *Sh. flexneri* T3SS needle in (c) cut along the plane parallel with the helical axis (black line). It can be seen in this view as well that the paired helices in the models for both bacteria are located at the same radius and have the same orientation. (f) The reconstruction in (a) cut along the plane parallel with the helical axis. When the subunit from the atomic model[13] of the *Sh. flexneri* needle (green) is docked into the reconstruction of the *Salmonella* needle, its N-terminal part sticks out from the volume (red arrowhead).

match to the reconstruction and suggests that this is a better estimate of the resolution. This resolution is insufficient to determine the absolute hand of the reconstruction, as the two enantiomorphic volumes that can be produced are fitted equally well by the PrgI two-helix bundle that projects out. We have therefore adopted the same hand (a right-handed 1-start helix) assumed for the *Sh. flexneri* needle,[4,13] although the hand was not determined for that needle and was assumed to be the same as that of the *Sa. typhimurium* flagellar hook and filament.

The outer diameter of the reconstructed volume is ~85 Å, and the central lumen has a diameter of ~25 Å. Docking the two-helix bundle (residues 18–60) from the center of the PrgI sequence into the reconstructed volume (Fig. 2a) shows a remarkably good agreement (Fig. 2d and e) with the corresponding portion of the atomic model[13] proposed for the *Sh. flexneri* needle. Unfortunately, the limited resolution of our present reconstruction, combined with the structural plasticity of the PrgI subunit outside of the fixed two-helix bundle,[12] prevents us from generating a pseudo-atomic model for the remainder of the PrgI subunit that must be surrounding the lumen. It is quite likely that the residues on the N-terminal side of the two-helix bundle are involved in crucial subunit–subunit interactions, as deletion of the first five residues renders the PrgI subunit incapable of polymerization.[12,15]

Discussion

An atomic model of the *Sa. typhimurium* flagellar filament[5] shows that the lumen in that hollow assembly is surrounded by coiled coils (D0 domain) formed by the N-terminus and the C-terminus of the component subunit. These coiled coils generate 11 nearly vertical protofilaments. The near-atomic resolution for this system that was achieved by electron microscopy was possible due to the almost crystalline order of these filaments. In contrast, we have observed that the needle from these same bacteria is quite disordered and structurally heterogeneous. The average symmetry that we observe (~6.3 subunits/turn of the 26-Å-pitch helix) would generate 11-start helices with a very short pitch (~60 Å). Although we lack the resolution needed to directly observe the packing of subunits surrounding the lumen, the symmetry that exists is incompatible with long-pitch 11-start helices. It has recently been shown that although the *Sa. typhimurium* flagellar hook protein forms α-helical coiled coils, the packing of these coiled coils surrounding the lumen of the hook is very different from the packing in the flagellar filament, even though the helical symmetries of both are nearly identical.[16] We have previously shown[17] that the flagellin subunits of *Sa. typhimurium* and *Campylobacter jejuni*, while highly conserved at the sequence level in the D0 and D1 coiled-coil domains responsible for filament polymerization in *Sa. typhimurium*, assemble into flagellar filaments with very different packing and helical symmetry. Thus, the ability of the extracellular flagellar and needle proteins to adopt different packing schemes around the lumens likely reflects diversity in the physical properties of these different polymers. While the packing of subunits surrounding the lumen of the needle remains unknown, there is a remarkably good match between the two-helix bundle that we can fit into our reconstruction and a model that has been previously proposed for the *Sh. flexneri* needle.[13]

How information about host cell contact is transmitted to turn on the secretory machinery located at the base of the T3SS remains a mystery. The likely role of the needle in transmitting such information was shown by the fact that mutants of the needle protein that induce a constitutively "on" secretion could be found in *Shigella*.[18] However, the needles formed from the mutant proteins unexpectedly showed no overall changes in the average helical symmetry.[6] The structural heterogeneity that we observe for the *Sa. typhimurium* needle raises new possibilities for how the needle might transmit information about host cell contact. The large variability in twist present in the *Sa. typhimurium* needles means that there must be a correspondingly large plasticity in how subunits are locally packed around the lumen. Such plasticity could be used to transmit information along the needle. A testable prediction of this hypothesis, beyond the scope of the present work, is that agents (such as cross-links) that might lock the needle into a fixed symmetry should abrogate host cell sensing.

Materials and Methods

Needle complex preparation

To facilitate the formation of elongated needle structures, we constructed a pWSK29-based plasmid containing *prgI*, including a 900-bp upstream region. The plasmid was transformed into a *Sa. typhimurium* strain (SB905, described elsewhere) carrying a plasmid containing an inducible *hilA* gene.[19] The strain was grown under conditions promoting formation of the T3SS and induction of the regulatory protein HilA overnight (LB, 300 mM NaCl, suitable antibiotics, 0.02% arabinose, 37 °C, and 180 rpm).

Needle complexes were purified as described previously,[20] with minor modifications. In brief, cells from a 12-liter culture volume were harvested, resuspended in 300 ml of 500 mM sucrose in 150 mM Tris buffer (pH 8.0), and incubated after the addition of 7.5 ml of 500 mM ethylenediaminetetraacetic acid (EDTA) solution and 15 ml of lysozyme solution (30 mg/ml, 98,381 IU/mg; Sigma-Aldrich) on ice for 60 min. Protease inhibitor was added (three tablets of Complete® Protease Inhibitor; Roche), and the temperature was raised to 37 °C for 30 min. Successively, 15 ml of 10% lauryldimethylamine oxide (LDAO), 33 ml of 5 M NaCl, and 9 ml of 1 M $MgCl_2$ solution were added before the sample was incubated on ice for 60 min.

Aliquots (28 ml) were transferred into 36-ml ultracentrifugation tubes (Sorvall Centrifuges, USA), and a 3-ml layer of 36% wt/vol CsCl solution [36% wt/vol CsCl in 10 mM Tris, 500 mM NaCl, 5 mM EDTA, and 0.5% LDAO

(pH 8)] was created below each sample. Solubilized needle complexes were centrifuged into the lower fractions of the CsCl cushion (Sorvall Discovery 90SE, rotor AH-629, 28,000 rpm, 5 °C, and 3 h) to avoid pelleting directly onto the tube wall and damaging the filaments. Fractions containing needle filaments were pooled, and the density was adjusted to equal a CsCl concentration of 35–36% wt/vol. The sample was split into six thinwall tubes (4.4 ml of PA; Thermo Scientific, USA) and centrifuged overnight (Sorvall 90SE, rotor TH-660, 40,000 rpm, 12 h, and 5 °C). The needle complexes formed a visible band in the upper third of the tube. This fraction was collected and pelleted (Sorvall RC M150 GX, rotor S100-AT4, 90,000 rpm, 30 min, and 5 °C). The pellet was resuspended in 10 mM Tris buffer containing 500 mM NaCl, 5 mM EDTA, and 0.1% LDAO (pH 8).

Electron microscopy and image processing

A purified needle complex sample was applied to glow-discharged carbon-coated 400-mesh hexagonal Cu/Pd grids. For negative-stain images, 5 μl of the sample was placed on the grid and incubated for 40 s. The grid was washed and subsequently stained with 2% phosphotungstic acid at neutral pH for 25 s. For cryo-EM, a 5-μl sample was applied to the glow-discharged grids, the sample was allowed to settle for 40 s, and excess liquid was removed by blotting with double-layered filter paper (Whatman 4) before vitrifying the sample by plunge freezing in liquid ethane. Low-dose data were collected with a FEI Tecnai Polara at 300 kV using a Gatan Ultrascan 4000 UHS charge-coupled device camera (16 megapixels, 4000 × 4000, 15 μm pixel size). Images were acquired at a magnification of 112,968× (1.33 Å/pixel at the level of the specimen), with underfocus values ranging from 1.2 to 3.0 μm. Images were subsequently decimated to 2.66 Å, and boxes of 120 pixel length (320 Å) were used for IHRSR.

Scanning transmission electron microscopy

STEM experiments were performed at the Brookaven National Laboratory using tobacco mosaic virus as internal standard for mass-per-length measurements. The preparation of samples and analysis of images were performed as described previously.[11]

Acknowledgements

We thank Joe Wall and Martha Simon for the use of the STEM facility at Brookhaven National Laboratory. This work was supported by National Institutes of Health grant EB001567 (to E.H.E.) and the Zentrum für Innovation and Technology through the Center of Molecular and Cellular Nanostructure, Vienna (to T.C.M.).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2010.01.001

References

1. Galan, J. E. (2001). *Salmonella* interactions with host cells: type III secretion at work. *Annu. Rev. Cell Dev. Biol.* **17**, 53–86.
2. Desvaux, M., Hebraud, M., Henderson, I. R. & Pallen, M. J. (2006). Type III secretion: what's in a name? *Trends Microbiol.* **14**, 157–160.
3. Blocker, A., Komoriya, K. & Aizawa, S. (2003). Type III secretion systems and bacterial flagella: insights into their function from structural similarities 1. *Proc. Natl Acad. Sci. USA*, **100**, 3027–3030.
4. Cordes, F. S., Komoriya, K., Larquet, E., Yang, S., Egelman, E. H., Blocker, A. & Lea, S. M. (2003). Helical structure of the needle of the type III secretion system of *Shigella flexneri*. *J. Biol. Chem.* **278**, 17103–17107.
5. Yonekura, K., Maki-Yonekura, S. & Namba, K. (2003). Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature*, **424**, 643–650.
6. Cordes, F. S., Daniell, S., Kenjale, R., Saurya, S., Picking, W. L., Picking, W. D. *et al.* (2005). Helical packing of needles from functionally altered *Shigella* type III secretion systems. *J. Mol. Biol.* **354**, 206–211.
7. Yamashita, I., Hasegawa, K., Suzuki, H., Vonderviszt, F., Mimori-Kiyosue, Y. & Namba, K. (1998). Structure and switching of bacterial flagellar filaments studied by X-ray fiber diffraction. *Nat. Struct. Biol.* **5**, 125–132.
8. Egelman, E. H. & DeRosier, D. J. (1982). The Fourier transform of actin and other helical systems with cumulative random angular disorder. *Acta Crystallogr. Sect. A*, **38**, 796–799.
9. Egelman, E. H. (2007). The iterative helical real space reconstruction method: surmounting the problems posed by real polymers. *J. Struct. Biol.* **157**, 83–94.
10. Egelman, E. H. (2000). A robust algorithm for the reconstruction of helical filaments using single-particle methods. *Ultramicroscopy*, **85**, 225–234.
11. Wall, J. S. & Hainfeld, J. F. (1986). Mass mapping with the scanning transmission electron microscope. *Annu. Rev. Biophys. Biophys. Chem.* **15**, 355–376.
12. Wang, Y., Ouellette, A. N., Egan, C. W., Rathinavelan, T., Im, W. & De Guzman, R. N. (2007). Differences in the electrostatic surfaces of the type III secretion needle proteins PrgI, BsaL, and MxiH. *J. Mol. Biol.* **371**, 1304–1314.
13. Deane, J. E., Roversi, P., Cordes, F. S., Johnson, S., Kenjale, R., Daniell, S. *et al.* (2006). Molecular model of a type III secretion system needle: implications for host-cell sensing. *Proc. Natl Acad Sci. USA*, **103**, 12529–12533.
14. Galkin, V. E., Orlova, A., Cherepanova, O., Lebart, M. C. & Egelman, E. H. (2008). High-resolution cryo-EM structure of the F-actin–fimbrin/plastin ABD2 complex. *Proc. Natl Acad. Sci. USA*, **105**, 1494–1498.
15. Darboe, N., Kenjale, R., Picking, W. L., Picking, W. D. & Middaugh, C. R. (2006). Physical characterization of MxiH and PrgI, the needle component of the type III secretion apparatus from *Shigella* and *Salmonella*. *Protein Sci.* **15**, 543–552.
16. Fujii, T., Kato, T. & Namba, K. (2009). Specific arrangement of alpha-helical coiled-coils in the core domain of the bacterial flagellar hook for the universal joint function. *Structure*, **17**, 1485–1493.
17. Galkin, V. E., Yu, X., Bielnicki, J., Heuser, J., Ewing, C. P., Guerry, P. & Egelman, E. H. (2008). Divergence of quaternary structures among bacterial flagellar filaments. *Science*, **320**, 382–385.

18. Kenjale, R., Wilson, J., Zenk, S. F., Saurya, S., Picking, W. L., Picking, W. D. & Blocker, A. (2005). The needle component of the type III secreton of *Shigella* regulates the activity of the secretion apparatus. *J. Biol. Chem.* **280**, 42929–42937.
19. Kubori, T., Sukhan, A., Aizawa, S. I. & Galan, J. E. (2000). Molecular characterization and assembly of the needle complex of the *Salmonella typhimurium* type III protein secretion system. *Proc. Natl Acad. Sci. USA*, **97**, 10225–10230.
20. Marlovits, T. C., Kubori, T., Sukhan, A., Thomas, D. R., Galán, J. E. & Unger, V. M. (2004). Structural insights into the assembly of the type III secretion needle complex. *Science*, **306**, 1040–1042.

curriculum vitae

Personal Details

Wolfgang Hellmut Schmied

Böcklinstrasse 98/8
1020 Vienna
Austria
schmied@gmail.com

Date of Birth	30 March 1983 in Wels, Austria
Gender	Male
Nationality	Austrian

Education

2008-2009	**Diploma thesis** in the group of **Thomas Marlovits**, Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna, Austria Functional and structural studies on the type three secretion system of *Salmonella typhimurium*
2002-2010	**Molecular Biology**, **University of Vienna**, Austria focus on cell biology, immunology/microbiology and ecology
1997-2002	**HTL for Food Technology**, Wels (secondary school, offering professional training and university entry qualification)
1995-1997	Secondary School, branch with emphasis on science, BRG Brucknerstraße, Wels
1993-1995	Secondary School, BRG Brucknerstraße, Wels

Publications

1. **Schmied WH**, Takken W, Killeen GF, Knols BG, Smallegange RC. **Evaluation of two counterflow traps for testing behaviour-mediating compounds for the malaria vector Anopheles gambiae s.s. under semi-field conditions in Tanzania**. Malaria journal. 2008
2. Schraidt O, Lefebre M, Brunner M, **Schmied WH**, Schmidt A, Radics J, Mechtler K, Galan J, Marlovits TC. **Topology and organization of the Salmonella typhimurium type III secretion needle complex components**. PloS Pathogens, in press, 2010.
3. Galkin* VE, **Schmied* WH**, Schraidt O, Egelman EH, Marlovits TC. **The Structure of the Salmonella typhimurium Type III Secretion System Needle Shows Divergence from the Flagellar System**. JMB 2010.
   * equally contributed

Work Experience

05/2009 – 12/2009	**Research Technician at the Medical University of Vienna,** based at the **Institute of Molecular Biotechnology (IMBA)**, Vienna, AUT
10/2007 – 01/2008	**Internship/Thesis – Thomas Marlovits Group, IMBA,** Vienna, AUT
02/2008 – 04/2009	Determining the function, structure and organisation of subunits in the type three secretion system of *Salmonella typhimurium* using electron microscopy, nuclear magnetic resonance spectroscopy, biochemical and ELISA-based methods.
02/2007 – 06/2007	**Internship Wageningen University**, Entomology Lab. Wageningen, NL Microbial production of mosquito kairomones and identification of potential key substrates, olfactometer and large cage experiments.
03/2006 – 10/2006	**Internship Ifakara Health Research and Development Center (IHRDC)**, Entomology Unit. Ifakara, Tanzania Working on behavioural assays regarding mosquito olfaction under semi-field conditions within a Grand Challenges in Global Health - Project.
02/2004 - 12/2006	**Technical Assistant, Food Research Institute Vienna** 8 hours/week, working on clean room technology in the food industry: hygienic and microbiological monitoring, interpretation and presentation.
07-08/2005	**Internship International Atomic Energy Agency (IAEA), Entomology Unit**, Mosquito Group Developing a PCR/RFLP based assay to distinguish two *Anopheles arabiensis* strains, various morphologic screens, establishing new colonies from rare phenotypes.
07/2003	**Internship Food Research Institute Vienna** Conducting chemical analyses for external companies, investigations in scientific databases, researching regulations for the food- and pharmaceutics industry