# STRUCTURAL CHARACTERIZATION OF THE TYPE 3 SECRETION SYSTEM NEEDLE COMPLEX BY SINGLE PARTICLE CRYOGENIC ELECTRON MICROSCOPY

by

JINHONG HU

B.Sc., The University of Calgary, 2014
M.Sc., The University of Calgary, 2016

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Biochemistry and Molecular Biology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

December 2021

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Structural characterization of the type 3 secretion system needle complex by single particle cryogenic electron microscopy

submitted by Jinhong Hu in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Biochemistry and Molecular Biology

**Examining Committee**:

Dr. Natalie C. J. Strynadka, Biochemistry and Molecular Biology, UBC

Supervisor

Dr. Brett B. Finlay, Microbiology and Immunology, UBC

Supervisory Committee Member

Dr. Sriram Subramaniam, Biochemistry and Molecular Biology, UBC

University Examiner

Dr. Steve Plotkin, Physics and Astronomy, UBC

University Examiner

**Additional Supervisory Committee Members:**

Dr. Calvin K. Yip, Biochemistry and Molecular Biology, UBC

Supervisory Committee Member

## Abstract

The bacterial type III secretion system, or injectisome, is a syringe-shaped nanomachine essential for the virulence of many pathogenic Gram-negative bacteria. A major functional subcomplex of the injectisome, the needle complex, is a 3.5MDa complex formed by more than ten unique proteins. The needle complex forms a continuous channel spanning both the inner and outer membranes of Gram-negative pathogens, created by three highly oligomerized inner and outer membrane hollow rings and a polymerized helical needle filament. The effector proteins secreted through this channel, which vary amongst different bacterial species, are essential for subsequent pathogenicity. Thus, structural studies of this complex can provide important atomic level information for understanding complex assembly and function of the injectisome as well as potentially development of new antivirulence drugs or vaccines to combat infections in susceptible human, animal and plant hosts pathogens.

The first high-resolution needle complex structures determined by cryogenic electron microscopy (cryo-EM) here shows the atomic details of the inner and outer membrane protein complex and the needle filaments. The outer membrane component of the needle complex belongs to the secretin family, a giant necessarily gated pore common and essential to other bacterial secretion systems but which had remained largely uncharacterized at the atomic level until recent work including major contributions as outlined in this thesis. Further, the structures of the dual nested rings that form the major inner membrane structural component of the needle complex shows remarkable similarity regardless of the assembly stages, inferring a highly stable foundation for the other components of the system to pack and function within.

The snapshots of multiple needle complexes at different assembly stages revealed multiple new structures, the dynamics and the sequence of the assembly. The structural information also answered several long-standing additional questions, such as the mystery of the apparent symmetry mismatch between the inner and outer membrane complex of prior structures, and the fold, span and functional role of the historically named “inner rod” protein which this thesis works shows is not a rod but an adaptor to set the needle helicity and anchor it to the stable inner membrane platform.

## Lay Summary

Antibiotic resistance has become a global health threat. This is because pathogens have acquired resistance to the commonly used antibiotics, which leads to complicated and costly treatment of common infectious diseases.

One of the ways to combat antibiotic resistance is to develop new drugs that do not kill the bacteria but rather weaken its virulence. Antivirulence drugs, such as those targeting the type III secretion system, would not kill the bacteria, leads to less drug resistance compared to the classically used bactericidal antibiotics. The bacterial type 3 secretion system functions as a molecular syringe to deliver effector proteins into the host cell, a process essential to pathogenicity and which is relatively conserved amongst many important Gram-negative pathogens. Structural studies of this system can therefore aid in the development of new drugs and vaccines to help treat drug resistant infections now prevalent globally in the clinic and community.

## Preface

Chapter 1 (from page 14 - page 27) is from a published review article (Hu, J., Worrall, L. J. & Strynadka, N. C. Towards capture of dynamic assembly and action of the T3SS at near atomic resolution. *Curr. Opin. Struct. Biol.* **61**, 71–78 (2020)). I prepared the text and figures from chapter 1, with revision by my supervisor Natalie Strynadka.

Chapter 2 is a published article (Hu, J. *et al.* Cryo-EM analysis of the T3S injectisome reveals the structure of the needle and open secretin. *Nat. Commun.* **9**, 1–11 (2018)). I performed protein expression and purification with help from Maria Vuckovic. The needle complex cryo-EM data is collected by Claire Atkinson. I processed the data with help from Liam Worrall. The needle filament data is collected and processed by Hong Chuan and Zhiheng Yu. The secretion assay and knock-out strain were done by Deng Wanyin. The model is built and refined by Liam Worrall and me. I wrote the manuscript and made the figure with help from Liam Worrall and revisions by Natalie Strynadka.

Chapter 3 is a published article (Hu, J. *et al.* T3S injectisome needle complex structures in four distinct states reveal the basis of membrane coupling and assembly. *Nat. Microbiol.* **4**, 2010–2019 (2019)). I performed protein expression and purification with help from Maria Vuckovic. The cryo-EM data is collected by Hong Chuan and Zhiheng Yu. I processed the data with help from Liam Worrall. The secretion assay and knock-out strain were done by Deng Wanyin. The model is built and refined by Liam Worrall and me. I wrote the manuscript and made the figure with help from Liam Worrall and revisions by Natalie Strynadka.

Chapter 4 was written by me and revised by Natalie Strynadka.

## Table of Contents

## List of Tables

## List of Figures

## List of Abbreviations

| | |
|---|---|
| AHL | Amphipathic helical loop |
| ATP | Adenosine triphosphate |
| BAM | β-barrel assembly machinery |
| CTF | Contrast transfer function |
| DDM | n-Dodecyl-β-maltoside |
| DNA | Deoxyribonucleic acid |
| EDTA | Ethylenediaminetetraacetic acid |
| EHEC | Enterohemhorragic *Escherichia Coli* |
| EM | Electron microscopy |
| EMDB | Electron microscopy data bank |
| EPEC | Enteropathogenic *Escherichia coli* |
| ET | Electron tomography |
| FSC | Fourier shell correlation |
| GDN | Glyco-diosgenin |
| GPF | Green fluorescent protein |
| IM | Inner membrane |
| LDAO | Lauryldimethylamine oxide |
| MS | Mass spectrometry |
| NC | Needle complex |
| NMR | Nuclear Magnetic Resonance |

| | |
|---|---|
| OD | Optical density |
| OM | Outer membrane |
| PCR | Polymerase chain reaction |
| PDB | Protein data bank |
| PM | Host plasma membrane |
| PMF | Proton motif force |
| RBM | Ring building motif |
| RMSD | Root-mean-square deviation |
| RNA | Ribonucleic acid |
| SCV | *Salmonella*-containing vacuole |
| SDS | Sodium dodecyl sulfate |
| SEC | Secretory |
| SPA | Single particle analysis |
| SPI | *Salmonella* pathogenicity island |
| ssNMR | Solid-state nuclear magnetic resonance |
| T3SS | Type 3 secretion system |
| T4PS | Type IV pilus system |
| TAT | Twin-arginine translocase |
| TEM | Transmission electron microscopy |
| TM | Transmembrane |

## Acknowledgements

There are so many people who made my Ph.D. journey as wonderful as it is, and I cannot list all the names, but I'll try to name a few:

First of all, I'd like to thank my amazing supervisor Dr. Natalie Strynadka. It is such a pleasure to be able to work in your lab with all the amazing projects. You are always so positive and supportive whether the project is working well or not. During the five years, I learned a lot of scientific knowledge as well as the altitude to work with the science with your valuable mentorship. I do not know what makes you take me as your graduate student at that time when I was clearly not the best candidate but thank you for putting your trust in me over the years and help me to succeed. The knowledge and the experience I learned in this lab will be such a valuable fortune for the rest of my life.

I also want to say thank you to my committee members, Dr. Brett Finlay and Dr. Calvin Yip. I'm so thankful for your support over the past five years.

Next, I'd like to thank my family members, my parents, and, in particular, my wife, XiaoShan Lai. You left your home and come to Vancouver with me with an unknown future ahead of you. You are always so supportive and encourage me to work at night and weekend despite you actually want me to spend more time with you. You also helped me on multiple projects on the lonely weekend in the lab and helped me prepare buffers, media, gels. As Marija always mentioned, she missed you gel you cast!

Next, I'd like to thank all the lab members. You all helped me so much along this five-year journey. I will not go through a long list of every one of you, but I do want to say thank you for your companionship, support, and mentorship.

In particular, I'd like to give a shoutout to Dr. Liam Worrall. All the projects would not be happening without your help. The initial introduction and help of the needle complex project are so valuable to a newbie in the lab who have no idea where to start. I have really enjoyed your mentorship over the past five years. You taught me a lot about programming, modeling, and writing, besides biological science.

I also like to give a shoutout to "lab mom" Marija Vuckovic. All the projects would not be happening without your help as well. Thank you for taking all the hard cloning orders and going through hundreds of liters of stink salmonella cells with me for many months, and more recently hear all the complaints and provide all valuable suggestions about the newborn baby.

Lastly, I'd like to thank all the collaborators. It's a pleasure to work with all the brilliant scientists in Canada and across the world. In particular, I'd like to thank Dr. Finlay and Dr. Deng for helping out on the microbiology, which I still know very little, and Dr. Zhiheng and Dr.Chuan on the cryo-EM when I only knew little about it. You all not only taught me the science but also how grateful to work as a collaborative team.

# Chapter 1: Introduction

## 1.1 Antibiotic resistance and virulence factors

Antibiotic resistance is one of the biggest health threats globally[1]. According to review on antimicrobial resistance report[2], it is estimated 700,000 people are affected by antibiotic resistance each year, and it is estimated that by 2050 this number will increase to 10 million[2]. With a growing number and spread of bacterial strains harboring antibiotic resistance mechanisms, an increasing number of hospital and community acquired infections have become harder and harder to treat as the antibiotics used to previously treat the diseases caused have become less efficient[1].

Antibiotic resistance is a result of the typically bactericidal way in which classical antibiotics work, for example disrupting bacterial reproduction or reducing survivability[3]. Unfortunately, this action often leads to the death of most but not all bacteria and leaves those resistant to that specific antibiotic to survive and massively reproduce, eventually leading to that type of bacteria gaining resistance to particular antibiotics[3].

Numerous efforts have been invested in combating antibiotic resistance[2,4]. In addition to establishing better regulation and reducing the usage of antibiotics in the clinical setting, much scientific research has been done to help combat this issue as well. As a result, a very small scattering of new antibiotics have been developed in recent years as well as synergistic drug

combinations that promote the efficacy of existing antibiotics[2]; however, the relatively slow speed of new drug development has been outpaced by the rapid growth of antibiotic resistance, demanding a new way to combat this threat. One of these new approaches is to target the virulence factors of bacteria, so called antivirulence drug strategies.

Unlike the classical antibiotics that typically kill the bacteria, creating a strong selective pressure that promotes development or acquisition of antibiotic resistance mechanisms, targeting virulence rather than an essential mechanism for viability, is thought to less likely generate the selective pressure and drug resistance phenomenon[5]. Furthermore, some virulence factors are significantly conserved across many pathogens, thus making them a compelling candidate for the new development of drugs and vaccines. Bacterial secretion systems, several of which are virulence factors, are one of the great potential candidates in this regard.

### 1.2 Bacterial secretion systems

Bacteria have evolved sophisticated systems to secret various substrates such as small molecules, protein or DNA for their ultimate survivability, pathogenicity and adaptation[6]. Those substrates are either secreted into the periplasmic space, extracellular space, or host cell. In gram-negative bacteria, the secretion nanomachines can be broadly divided into two categories: single membrane-spanning systems and dual membrane-spanning systems[6]. The one membrane-spanning systems, such as the SecYEG translocon[7] or twin-arginine translocon[8], transports signal peptide tagged substrates from their point of synthesis in the cytosol into the periplasmic space or inner membrane, depending on their specific signal peptide signatures. In

contrast, the dual membrane-spanning systems, unique to duodermic Gram negative bacteria, directly transport a variety of substrates from the cytosol across both inner and outer membranes as well as the intervening periplasmic space and peptidoglycan layer[6]. Most of the dual membrane-spanning systems can only transport unfolded or partially folded substrates, and the secretion system assembly and transport process therein is under strict cellular regulation. There are so far nine different dual-membrane spanning systems identified (Type I to Type IX). They all perform critical roles in bacterial survival and reproduction. For example, the Type I secretion system secretes various sizes of products that are important for nutrient acquisition and virulence[9], while the Type II secretion system secretes folded substrates such as hydrolyzing enzymes and toxins which are important for survival and growth[10]. Most of the dual-membrane secretion systems use a one-step secretion method, directly transporting substrates from their point of synthesis in the cytoplasmic space of bacteria into the extracellular space or the host cell. The Type III secretion system is one prominent and widely dispersed example of the latter[11].

### 1.3 Type 3 secretion system

The Type III secretion system (T3SS) is one of the most sophisticated and well-studied bacterial secretion systems due to its central role in the pathogenicity of many gram-negative bacteria including enteropathogenic *Escherichia coli* (EPEC), *Salmonella enterica* and *Pseudomonas aeruginosa* amongst several others, that cause a wide spectrum of potentially devastating diseases, for example, typhoid fever, whooping cough, diarrheal disease, and sexually transmitted infections[12].

The T3SS encompasses two cellular apparatus, the flagellum and injectisome. Although these two systems have different functions, the core of the secretion machinery, particularly that centered at the inner membrane. is similar. In the flagellum, the T3SS is used to export and assemble flagellar monomers to build extracellular filaments that facilitate essential movement of the bacteria. In the needle like injectisome, on the other hand, the T3SS is used to secrete virulence effector proteins into host cells to aid pathogenicity [13]. The injectisome and flagellar match most closely when considering components of the IM rings and some cytosolic components. However, the OM components have no similarity; the injectisome uses a secretin giant gated pore, whereas the flagellar does not. In this thesis, I will only focus on the structure and function of the T3SS injectisome critical to pathogenicity and disease in many Gram negative bacteria.

The *Salmonella* T3SS was first isolated more than 20 years ago[14,15]. Previously determined sub-nanometer cryogenic electron microscopy (cryo-EM) structures and a 17Å in situ cryogenic electron tomography (cryo-ET) structure clearly show the general architecture of this system[16–18]. Prior to this study, many components of the T3SS were determined using various biophysical methods. The main structural component is a syringe-like nanomachine that is ~3.5MDa. It has three major components: the extracellular needle filaments, the needle complex base which spans both bacteria inner and outer membrane ring, and the cytosolic complex (Fig 1.1, Table 1.1). The needle filament and the needle complex base collectively are referred to as the needle complex (NC). The needle filament is comprised of an extended needle polymer, formed by a coiled coild helical repeat of type III secreted monomers with lengths dependent

upon the bacterial species. The needle polymer is capped by a poorly characterized needle tip oligomer that presumably acts as an adaptor to in turn anchor the terminating pore-forming complex (translocon) that inserts into the host cell membrane. The first *Salmonella* needle polymer structure was determined to 7Å by cryo-EM[19], and subsequently, an atomic model was generated by solid state nuclear magnetic resonance (ssNMR) spectroscopy[20]. To date however, there is limited atomic information defining the molecular features of the needle tip oligomer or the pore-forming complex. The needle polymer is anchored within the needle complex base at the inner membrane, specifically through interactions facilitated by the core export apparatus and a coiled coil adaptor (historically termed the inner rod). The three centromeric rings of the basal body surround and structurally support the export apparatus and the filaments in their span between the two bacteria membranes. Interestingly, the export apparatus was quite unambiguously predicted by sequence analysis to be a multiple transmembrane complex. But a cryo-EM structure of the isolated flagellar export apparatus (a T3S homolog to that of injectisome's) led to the proposal that it rather may sit on top of the inner membrane (IM)[21]. The structures of the two IM ring monomers were determined individually by X-ray crystallography and NMR[22–25], with Rosetta based modelling and various validations/restraints used to predict features of the resulting ring oligomer. The structure of the outer membrane (OM) channel, a member of the previously uncharacterized secretin family of giant gated pores, as well as the three rings in an assembled form (the basal body) was determined at near atomic resolution by cryo-EM[26]. The cytosolic complex houses an export gate complex, a set of scaffold proteins as well as an ATPase/central stalk complex, which are essential for the substrate recruiting and unfolding. The c-terminus of the export gate was

determined first by X-ray crystallography and subsequently by cryo-EM, and the structure of the ATPase with its central stalk was determined by cryo-EM as well[27–31]. The T3SS specific substrates, also referred to as effectors, often have an N-terminal signaling peptide and bind to T3S specific chaperones that maintain them in a partially unfolded state, as evidenced by prior crystallographic structures of these complexes[32,33]. The substrate dissociates from their cognate chaperone and is partially unfolded in the cytoplasmic complex before it enters the channel at the inner lumen of the NC, which is comprised of the export apparatus core (SctRSTU) and associated SctV IM pore, through which early substrates (the "rod" and needle filaments) followed by late substrates (virulence effectors destined to the host cell cytoplasm through the pore-forming complex). The early effectors are therefore building blocks of the T3SS itself and the late effectors can alter host cell functions through targeted interactions to promote pathogenicity[11]. Although the effectors vary significantly amongst different bacterial types due to their host targets and functional roles, the secretion system components themselves show a high level of similarity between the species at the sequence and structural level[11]. Such conservation has made the T3SS a compelling antivirulence drug and vaccine candidate amongst both the academic and biotechnology sectors.

To date, several different chemical classes of T3SS specific inhibitors have also been developed including salicylidene acylhydrazides, benzimidazoles, coiled-coil peptides, and various natural products. These compounds indicate promise in this approach for the treatment of various bacterial infections reliant on the T3SS[4]. In addition, the extracellular protrusions of the T3SS provide a compelling scenario for vaccine strategies with several proof of concept results

approved (EPEC cattle vaccine for example)[34] or under development. Improving the efficacy of these antivirulence strategies includes a concerted effort to understand the molecular details of their effects on T3SS structure/function, with major efforts underway over the past 2 decades to pave the way for current atomic level understanding.

### 1.4 The biological role of T3SS

The ultimate actions of T3SS delivered effectors on the host involves manipulation of common cellular processes, such as host immune responses, cytoskeletal dynamics, vesicle transport and signal transduction pathways[35]. There are two types of pathogens that require the T3SS to facilitate their invasion, the extracellular pathogens such as enteropathogenic *Escherichia coli* (EPEC) and enterohaemorrhagic *E. coli* (EHEC) as well as *Citrobacter rodentium*, and intercellular pathogens such as *Salmonella*, *Shigella*, and *Chlamydia*[36,37]. The extracellular pathogens use the T3SS to secret effector proteins to modulate the host cytoskeleton and immune responses so that they can attach to the intestinal epithelium of the host during invasion. Attachment triggers cytoskeletal rearrangement and weakens the epithelial barrier of the host cell. The effector proteins are also used to avoid the immune response by interrupting nuclear factor-κB and mitogen-activated protein kinase activation[38,39]. In the case of extracellular pathogens, they use the T3SS to invade host cells and establish their life cycles therein. For example, *Salmonella* spp. has two distinct T3SSs, the *Salmonella* pathogenicity island 1 (SPI-1) and *Salmonella* pathogenicity island 2 (SPI-2). SPI-1 is used to invade the host cell, create the Salmonella-containing vacuole (SCV), and cause inflammation, while SPI-2 is key to survival and reproduction of the *Salmonella* spp.[35,36]. The inflammation caused by the

effectors secreted through the T3SS allows the *Salmonella* spp. to reside in the intestinal space and secure both a carbon source and electron acceptors, which are essential for its metabolism and replication, key to its pathogenicity[40]. Due to its central role in pathogenicity, this machinery has drawn much attention since its discovery 30 years ago. Many efforts have been put into identifying and understanding the assembly, structure, and function of this T3SS. The recent advances in cryo-EM allows us to capture unprecedented molecular details of this sophisticated nanomachine[11].

### 1.5 Cryo-EM and data processing

Cryo-EM is a high-resolution imaging method to acquire structural information from biological samples, such as proteins, nucleic acids, macromolecular complexes, viruses, and even cells[41]. A temporal and spatial coherent electron beam generated by either tungsten filaments, lanthanum hexaboride, or a field emission gun is used to create 2D projection images of these samples made possible from the resultant elastic scattering when electrons interact with the biological sample with no energy loss. The images formed are mostly from the phase contrast, with around 10% from the amplitude contrast[42]. The sample been imaged is embedded in vitreous ice to preserve its native structures[43]. This is done typically by rapid plunge freeze of a thin layer of buffer containing the desired biological sample in liquid ethane that is subsequently cooled to the temperature of liquid nitrogen during the plunge. Vitrification is necessary in order to optimally preserve the sample against electron beam induced radiation damage during the extended data collection times. In addition, the electron beam of the microscope has to be confined to a high vacuum environment due to its strong scattering; without vitrification the sample would be

destroyed by dehydration in this environment [44]. The elastically scattered electrons are focused by the magnetic lens systems and recorded on a film or more commonly in modern microscopes used for high resolution data collection, a direct electron detector (DDD). The recorded sample data, i.e. its 2D projections, is convoluted by a contrast transfer function (CTF), which is a sine function that oscillates in a frequency-dependent manner that provides both phase and amplitude of the resultant image. CTF is partially determined by the level of focus as the image is recorded, also referred to as the defocus value. Close to focus will retain the high-resolution information but suffer the loss of low-resolution information, which can make it challenging to identify the biological sample[45]. Thus, a range of defocus values is used during the data collection. This strategy also avoids the 0 value at the CTF curve. The resulting 2D images are then aligned to estimate the 3D volume of such samples[46]. The most common method of alignment involves use of the Fourier slice theorem, which means the 2D projection of the 3D object in real space is equivalent to a slice through the centre of the Fourier transformed 3D object. Thousands of images with multiple views of the projections need to be aligned properly to generate a high-resolution 3D map[47]. The final resolution of these reconstructions are estimated by Fourier shell correlation (FSC) which measures the correlation of the signal in two half data sets[48]. It is important to note that these resolution estimates, although useful as a generalized relative indicator of data quality, do not reflect the effective resolution at all parts of the reconstruction which may vary substantially due to for example disorder, dynamics and strategy in data processing.

In reality, particularly before the resolution revolution, many aspects required to generate a high-resolution 3D map can be lacking. The stability of the beam, sample motions induced by the beam, slow throughput of the film (and hard to record low-frequency signals) or poor detection quantum efficiency (DQE) charge-coupled devices (CCD)[49], and low accuracy and reliability of algorithms used for aligning 3D particles[44] are all contributing factors to ultimate resolutions obtained. The development of a more stable and better vacuum microscope, and more importantly, the direct electron detector (DDD), which can detect charges generated by the strike of an electron directly rather than the indirect coupling requiring first conversion to photons as per the older generation CCD versions, significantly increase the DQE and allows the retention of the high-resolution signal[50]. Also, being able to record images as movies due to the more rapid readout speed afforded by the modern DDDs allows for correction of the otherwise unpredictable sample motions caused by the electron beam, thus further improving the ultimate signal to noise and quality of data[51,52]. On the other front, the data processing algorithms are also greatly improved with stochastic gradient descent[53], maximum likelihood [54] as well as Bayesian statistical approaches[55]. In addition, many constantly evolving software packages with user-friendly GUIs, such as RELION[56], Cryosparc[53], help guide less expert users to process their data to high-resolution. Together, the advancement of these techniques overcomes many hurdles of determining high-resolution cryo-EM data; however, challenges remain perhaps the most confounding still the production of stabile samples that allow for sufficient conformationally homogeneous particle populations needed for high resolution reconstructions, determination of optimal vitrification conditions and finally, in data collection/processing, overcoming the low signal-to-noise issue.

Due to the sensitivity of many biological samples, only a low dose of the electrons can be applied to the sample, leading to a low signal-to-noise ratio for the resulting image. The low signal-to-noise ratio issue in turn severely hinders the alignment needed for generating accurate 3D volumes[57]. To overcome this hurdle, various data processing strategies have been developed.

The general idea of these strategies is to obtain a high number of homogeneous sample projections and to average them computationally. Essentially the goal here is to effectively increase the signal of the true protein sample and reduce the noise level. One of the ways to increase the number of sample projections is to take advantage of the symmetry that the sample has. In that way, one sample can likely provide multiple homogenous projections. Also, the various views of projections critical to a 3D reconstruction can be obtained by applying symmetry. However, the symmetry within the sample can also be a disadvantage. The part that does not obey the dominant symmetry can often have a lower local resolution or even be smeared with no features, which causes difficulties in interpreting the map and extrapolating the structural information. This is referred to as the symmetry mismatch issue.

Symmetry mismatch can happen for multiple reasons. First, conformational dynamics or flexibility in a particular region can break the overall symmetry assumption and lead to a lower local resolution[58]. For example, the nuclear pore complex overall obeys eight-fold symmetry. However, due to local flexibility, the resolution of many regions are still relatively low, despite the core reaching 4.5Å resolution[59]. Second, partial occupancy of a substrate or small binding partner may not resolve well with symmetry applied[60] as not all binding region are

homogenous. One such case is a virus capsid with icosahedral symmetry which can bind up to 60 subunits, but frequently not all potential binding sites are occupied[61]. Third, subcomplexes may exhibit different symmetry compared to the overall symmetry components, which leads to the smear of that subcomplex. This is a common phenomen in many bacterial secretion systems[62]. For example, the T3SS has various closely abutted components of varying symmetries, ranging from no symmetry to helical symmetry to different centromeric symmetries[63].

Many methods have been developed to overcome the symmetry mismatch issue. The first one is asymmetric refinement[64]. If all the different symmetry components are locked in a fixed orientation with regards to each other, and the data quality is sufficiently good to detect the basic symmetry components, then the asymmetric refinement can work well. For example, in the case of the T3SS needle complex with substrate bound in the center lumen, 15,000 micrographs were collected, providing lots of particles to overcome the difficulty of aligning the low signal of the non-symmetric component, and led to a high-resolution un-symmetrized map[65]. On the other hand, suppose the signal of the basic symmetric component is not good enough. In that case, one can always apply the dominant symmetry to generate the model and use that model for asymmetric refinement. In the recent work of determining the structure of single-stranded-RNA in viruses, the author first applied the symmetry to the icosahedral virus during the refinement and then resolved the single-stranded-RNA with asymmetric refinement[66]. In addition to the asymmetric refinement, combining asymmetric refinement and symmetry relaxation could generate a better-resolved map quickly. This approach limits the search parameters based on the

dominant symmetry component and only searches angles based on the symmetry operators. For example, in the case of the icosahedrally symmetric capsid, one can only search another 59 symmetry-related angles based on the initial angle, which significantly improves the search speed[60].

Another strategy that has proven useful in many complex assemblies is focused refinement on the mismatched symmetry component. The common issue of asymmetric refinement is that the highly symmetric components dominate the alignment and cause the smaller sub-complex with different or no symmetry to align incorrectly and leads to a low local resolution[64]. To overcome this issue, refinement can be done just on the symmetry mismatched sub-complex. By subtracting the symmetry dominant component from the 3D map and in the 2D projected images, the refinement is more sensitive to detection of the true symmetry of the smaller sub-complex. This in turn can potentially facilitate the correct alignments needed to achieve a higher resolution sub-complex structure[67]. The subtraction of the 3D map can be done by applying a binary mask in the 3D volume. The subtraction of the 2D projection image can be done by removing the unwanted part in 3D volume first and then correlate the same signal in the 2D projections by comparing the angle information from the alignment in the 3D volume[67]. More importantly, this sub-complex structure, calculated from subtracted projections, could generate a true asymmetrized map with all the symmetry components to high-resolution if all the sub-complexes are fixed in the same orientation relative to each other. The work presented in this thesis took advantage of this approach[63]. By determining the structure of the asymmetrized T3SS export apparatus first, and applying this angle to the whole needle complex, we successfully aligned the

entire needle complex, which we could not align properly when only the asymmetric refinements were done.

The last common strategy is a localized refinement or "sub-particle" refinement. Although the focused refinement can be powerful to align a small part of the complex, the range of the search can be vast/prohibitive, with the mass of the center often not at the region of interest[68]. Therefore, it is often advantageous to generate sub-particles. The original particle is re-extracted in a smaller box with the center of mass at the region of interest[62,69,70]. The alignment parameters can be obtained from the asymmetric refinement. This way, it decreases the search time and increases the sensitivity of alignment because the search range is much smaller, usually deviating a few degrees from the initial angle, and the search is done in a much smaller box size. In the case of the reconstruction of native COPII cages, the author used this method to improve the local resolution of the vertices from 35Å to 12Å[69].

In summary, symmetry can be beneficial to the image processing pipeline but can also lead to difficulties. The variety of reasons for symmetry mismatch pose different types of challenges in data processing. Several image processing techniques have been developed to overcome the symmetry mismatch challenges in different cases. Often, these techniques can also be used in conjunction to obtain maps with better local resolution and reduce the time of processing. Ultimately, our understanding of those symmetry mismatched complexes can be enhanced with the increased level of detail in all regions.

## 1.6 Structural components of T3SS

Historically, complimentary hybrid methods have been employed with high-resolution techniques including X-ray crystallography and NMR used to study the structure of individual components, traditionally low-resolution methods such as electron microscopy used to study the overall structural envelopes, and bioinformatic methods used to combine the information resulting in increasingly more detailed molecular models[17,19,23–25,71–75]. However, the dynamic and transient nature of these complexes and assemblies, with many of the high-resolution structures representing only monomers belonging to undefined higher order homo- and hetero- oligomers, and the low-resolution data with unclear component boundaries or stoichiometries, resulted in incomplete and/or inaccurate models. The recent technological advances in the field of cryo-EM have opened the door to remarkable and exciting advances in our structural understanding of the injectisome. Both single particle analysis (SPA) and cryo-ET [76] – are contributing complimentary structural information such that it seems feasible that a complete molecular understanding of the entire system can be achieved in the near future. In this thesis, we have aimed to summarize the recent advances in our structural understanding of the T3SS as a result of this resolution revolution in particular focussing on how these structures illuminate the pathway of injectisome assembly.

### 1.6.1 Overall structure of the injectisome

The T3S injectisome constitutes a large nanomachine spanning three membranes with associated cytoplasmic components. The characteristic structural unit, stable on extraction from its native

membranous environment, is the syringe shaped needle complex (NC), spanning both bacterial membranes with an attached hollow needle-like projection that provides the delivery conduit for secreted effectors[14] (see Fig. 1.1, Table 1.1).

| **Unified nomenclature** | **Proposed function** | ***Salmonella* spp. (SPI-1)** | **Symmetry** |
|---|---|---|---|
| **Needle complex** | | | |
| SctC | Secretin, outer membrane ring | InvG [78] | C15/C16 |
| SctD | Inner membrane ring | PrgH [17] | C24 |
| SctJ | Inner membrane ring | PrgK [17] | C24 |
| SctF | Needle | PrgI [17] | Filaments |
| SctI | Inner rod | PrgJ [63] | C1 |
| SctU | Autoprotease | SpaS [71] | C1 |
| **Export apparatus** | | | |
| SctV | Export gate | InvA [79] | C9 |
| SctR | Export apparatus | SpaP [63] | C1 |
| SctS | Export apparatus | SpaQ [63] | C1 |
| SctT | Export apparatus | SpaR [63] | C1 |
| **Cytosol complex** | | | |
| SctQ | Cytoplasmic ring | SpaO [84] | C6 |
| SctN | ATPase | InvC [28] | C6 |
| SctL | Stator | OrgB [18] | C6 |
| SctO | Stalk | InvI [28] | C1 |
| SctK | ATPase cofactor | OrgA [18] | C6 |
| **Regulators** | | | |
| SctP | Needle-length regulator | InvJ [11] | C1 |
| SctW | Switch regulator | InvE [11] | C1 |
| **Translocators** | | | |

| | | | |
|---|---|---|---|
| SctB | Translocation pore | SipC [81] | C8 |
| SctE | Translocation pore | SipB [81] | C8 |
| SctA | Needle tip | SipD [11] | Filaments |

**Table 1.1 Core structural components and functions of T3SSs**

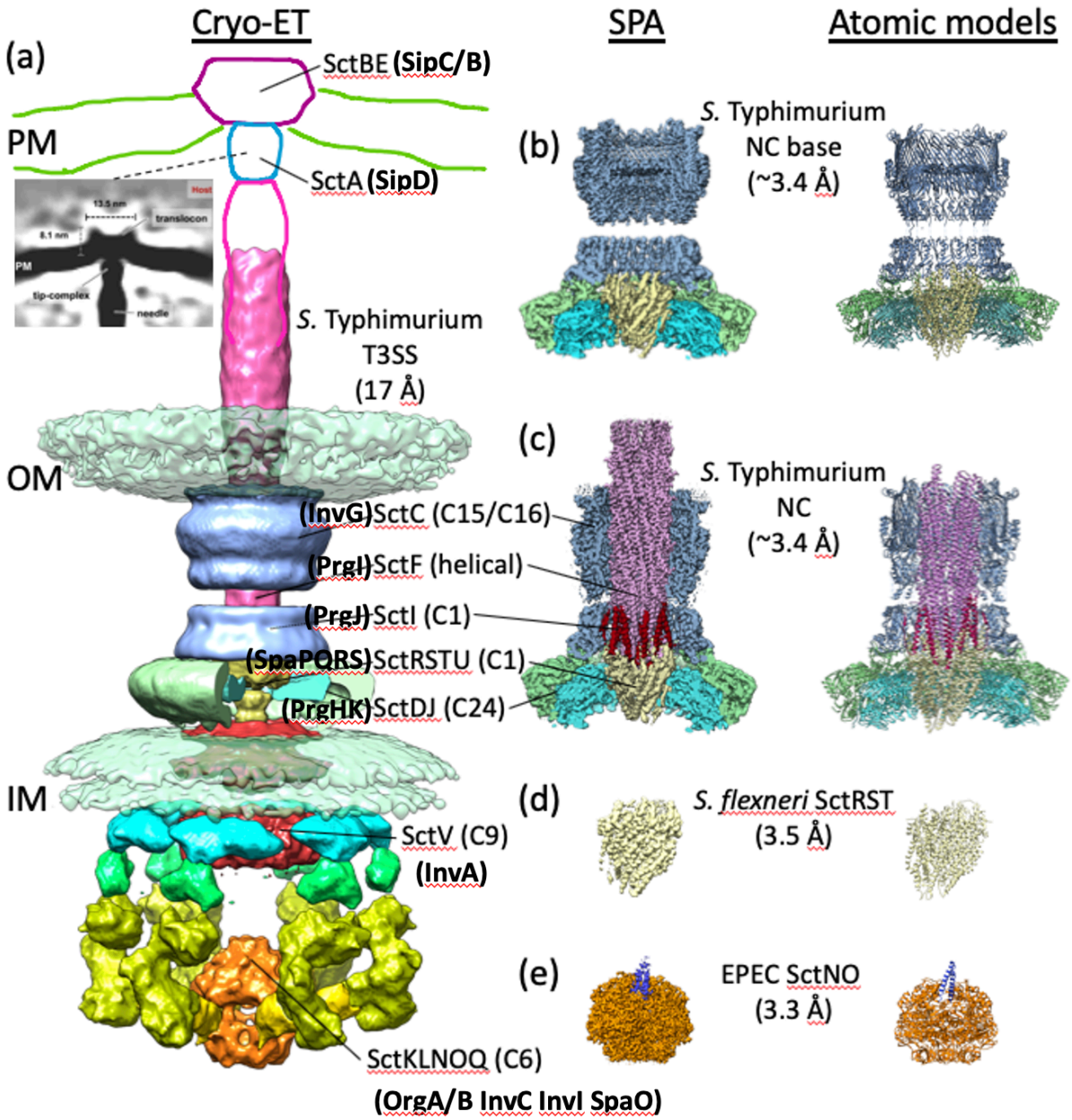


**Figure 1.1: Cryo-ET and SPA reconstructions of the T3SS. (a) In situ cryo-ET structure of the *Salmonella* Typhimurium SPI-1 T3SS showing overall architecture (EMDB-8544). The tip and translocon are outlined based on the cryo-ET visualization of the host-cell interface shown in inset (figure reproduced in accordance with the CC-BY licence). Components and their symmetries labelled. IM – bacterial inner membrane, OM – bacterial outer membrane, PM – host plasma membrane. (b)-(e) SPA reconstructions (left) and corresponding atomic models (right) of the NC base with no needle and secretin periplasmic gate closed (b; Composite of EMDB-20315 and EMDB-20316; PDB 6Q14), NC with assembled needle and secretin periplasmic gate open (c; composite of EMD-20556, EMD-8914, EMD-8924; PDB 6Q15), isolated flagellar export apparatus complex (d; EMDB-4173; PDB 6F2D), and ATPase-central stalk complex (e; EMDB-9391; PDB 6NJP).**

Despite its size, the NC outer scaffold is surprisingly made up of just three highly oligomeric ring forming proteins: PrgH (SctD) and PrgK (SctJ) (universal nomenclature), which form concentric periplasmic rings tethered to the IM[17], and InvG (SctC), a member of the secretin family of OM pores broadly conserved in the type II secretion system (T2SS), type IV pilus system (T4PS) and filamentous phage[78]. Together, these three proteins form a stable housing in which the helical needle filament – comprised of homologous proteins rod PrgJ (SctI) and needle PrgI (SctF) – are polymerized onto the IM abutting "export apparatus" complex SpaPQRS (SctRSTU) assembled within the IM rings. Localized to the IM patch within the NC and proximal cytoplasmic region are several proteins typically lost during purification[74]. These include the export gate protein InvA (SctV) with large membrane embedded N-terminus and soluble C-terminal domain shown to form a nonomer in solution[79], and a collection of soluble proteins termed the sorting platform which include the T3S ATPase InvC (SctN) with central

InvI (SctO) and peripheral OrgB (SctL) stalks (homologous to the rotary ATPase subunits) and two further proteins SpaO and OrgA (SctQ and SctK) which anchor the sorting platform to the injectisome base. Collectively, the export gate and sorting platform have central roles in the regulation and energization of effector secretion. At the distal end of the filament, the tip and translocon terminate the needle and form a pore in the host cell membrane respectively, completing the direct delivery channel between bacteria and host.

SPA and cryo-ET (specifically with subtomogram averaging applied) are now being routinely employed to study purified T3SS complexes to near-atomic resolutions or the entire system to low nm resolutions in its native cellular environment respectively. A significant challenge lies in data processing, which for both techniques relies on correct alignment of the images for reconstruction. The injectisome is made of components with varying symmetries including the 24x IM rings, 15x OM secretin beta-barrel, and internal helical symmetry of the needle, (Fig. 1.1) and alignment of one component can often leave others incorrectly aligned with no interpretable features in the resulting reconstructions. The 24x IM rings specifically typically dominate SPA processing with earlier studies generating high-resolution reconstructions for this region with the remainder of the NC not interpretable requiring either isolation or focussed refinement to further determine the structures of other components[26,80]. By first determining the structure of the internal asymmetrical export apparatus using a focussed refinement approach and subsequently expanding the correctly aligned orientations to the entire NC particles, Hu et al. were recently able to resolve the entire NC revealing the structures, stoichiometries and interactions of all components[63] (Fig. 1.1b-c). Cryo-ET with subtomogram averaging is

similarly starting to resolve the constituent symmetries of the injectisome components with recent studies revealing the hexameric nature of the cytoplasmic sorting platform and the nonomeric membrane embedded export gate[18,81].

### 1.6.2 Assembly of the T3SS

In this thesis, we will present the specific structures in relation to the dynamic process of injectisome assembly. Alternate pathways have been proposed[82]; however, in the prototypical *Salmonella* SPI-1 system discussed here, a consensus is emerging whereby assembly of a T3S competent complex is initiated in the IM.

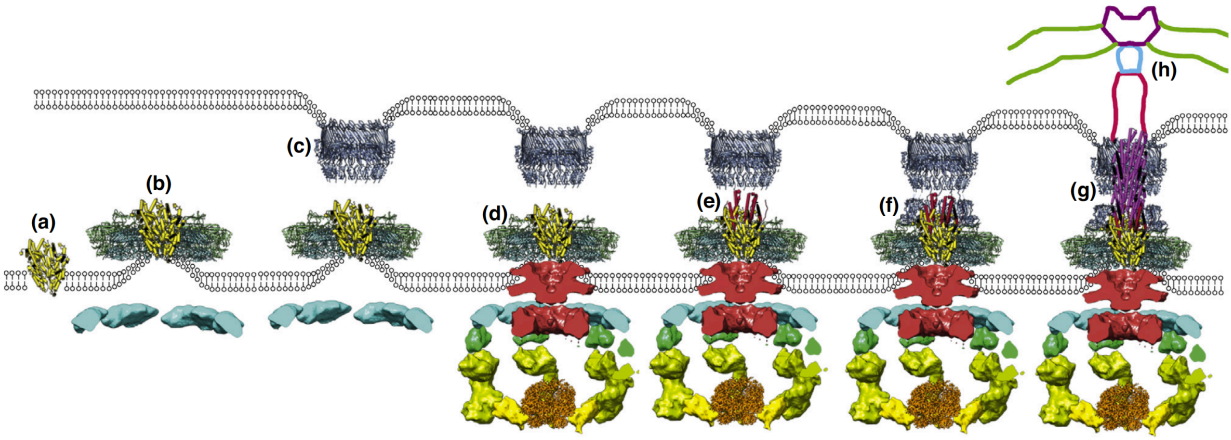



**Figure 1.2: Proposed assembly pathway of the T3S injectisome. Assembly is initiated at the IM with formation of the SctRSTU complex (a). Oligomerization of the IM SctDJ rings is nucleated around the SctRSTU complex (b), facilitating its extraction from the IM. The secretin SctC independently forms a 15-mer pore in the OM (c). Export gate SctV is recruited (note it is not clear if this happens before or after IM ring assembly) and the cytoplasmic sorting platform assembles (d) resulting in a T3S competent assembly. The rod SctI is secreted (e) and is proposed to assist in local peptidoglycan clearing and recruitment of SctC, which locally incorporates an extra N0-N1 domain to facilitate OM/IM component coupling (f). SctF needle**

**secretion and polymerization opens the secretin periplasmic gate (g). On host cell contact, secretion switches to the tip and translocases, which form a pore in the host cell membrane allowing direct delivery of effector proteins (h).**

### 1.6.3 The export apparatus

The IM export apparatus components assemble independently with SctRST, and to a lesser extent SctU, required for subsequent efficient NC assembly[72]. These four proteins have been historically predicted to be integral membrane proteins based on their hydrophobic helical character. Surprisingly, however, in 2018 Lea and colleagues published the structure of a recombinantly expressed, isolated T3SS export apparatus from the orthologous flagellar T3S apparatus showing instead a pseudo-helical assembly with 5:4:1 stoichiometry[21]. Based on shape, it was proposed that the injectisome equivalent would be localized in entirety within the IM rings[21], a region typically unresolved in NC reconstructions due to the challenges of SPA processing. Overcoming these, Hu et al. recently determined the structure of the entire Salmonella SPI-1 NC, resolving the SpaPQR (SctRST) complex in its native environs, intimately packed within the pore formed by 24-mer PrgK (SctJ) and providing an assembly platform for the rod (SctI) and needle (SctF) filament[63] (see below) (Fig. 1.1c). Interestingly, cryo-ET and SPA of a NC IM subcomplex isolated under conditions that maintained the local membrane structure, show the lipid bilayer is deformed around the assembled export apparatus, adopting a locally thinned and outwardly curved conformation[81] that appears to allow the export apparatus to fully traverse the IM thus representing the cytoplasmic facing access point to the

secretory channel. Closed in the isolated flagellar structure, it was suggested that the presence of the needle polymer would confer an open state of the export apparatus necessary for protein translocation[21] but the native injectisome structure is well conserved with the isolated flagellar orthologue in the closed state. Instead, SctU - involved in regulation of secretion hierarchy and likely associated with the SctRST pseudo-helical assembly [73,83], or SctV – the export gate involved in secretion regulation and energization (see below) – likely provide the trigger for the necessary structural remodelling to permit entrance to the export apparatus.

### 1.6.4 The IM rings

Assembled around the export apparatus, IM anchored SctD and SctJ form concentric 24-mer rings (Fig. 1.1b-c, 1.2b). Due to the highly symmetric nature, the IM rings are readily aligned during SPA and their assembled structure was one of the first resolved to high-resolution showing intimate packing of the two ring oligomers[26]. SctD has a small N-terminal cytoplasmic domain (not resolved in SPA) and larger periplasmic domain separated by a single transmembrane helix and SctJ is anchored to the IM via an N-terminal lipidation and variably present C-terminal transmembrane helix. Both proteins exploit a superficially similar small mixed alpha/beta motif termed a ring-building motif[22] present three times in PrgH and once in PrgK, with conserved self-association interfaces mediating oligomerization[25]. This domain and oligomerization interface are also observed in the OM secretin SctC (N1 and N3 domains). In the assembled NC, the pore formed at the centre of the SctJ ring provides a nanodisc-like environment for SctRSTU within[63] forming direct hydrophobic and polar protein-protein interactions. Interestingly, in absence of the export apparatus, the IM rings in *Salmonella* form

unphysiological 23-mers supporting the initial formation of the SctRSTU complex which serves as a template for IM ring oligomerization[81].

### 1.6.5 The export gate and sorting platform

Formation of a T3S competent complex is completed with assembly and recruitment of the membrane embedded export gate SctV and the cytoplasmic sorting platform (Fig. 1.2d). An X-ray crystallographic structure of the cytoplasmic domain of *Shigella* SctV shows a nonomeric ring[79], which is localized to a toroidal region of density below the export apparatus in cryo-ET reconstructions[18,81]. No high-resolution structures are available for the N-terminal membrane embedded region, which is predicted to contain 6-8 transmembrane helices; given the observed nonameric oligomerization of the cytoplasmic domain this would quite remarkably therefore indicate formation of an inner membrane structure of several dozens of helices, potentially the largest bacterial inner membrane structure to date. A recent cryo-ET study shows that the N-terminal membrane region assembles into a pore in the deformed IM patch around the export apparatus, although with few apparent direct interactions, observations that could explain why SctV has been difficult to capture as part of the NC or assemble independently[81]. The sorting platform encompasses the soluble proteins that associate with the NC and are required for secretion. These include the T3S ATPase SctN, associated central (SctO) and peripheral (SctL) stalks, SctQ, with homology to the flagellar C-ring, and SctK. Like SctV, these proteins are lost during purification of the NC and high-resolution details of the assembled complex is lacking. In situ cryo-ET demonstrates that SctK, SctL and SctQ assemble into six “pods” that link the central ATPase to the NC via an interaction with the N-terminal cytoplasmic domains of SctD,

which undergo reorganization from a 24-mer ring into discrete tetrameric bundles[18] (Fig. 1.1a). An SPA structure is available for the EPEC T3SS ATPase EscN in complex with central stalk EscO, which demonstrates structural conservation with the asymmetric rotary ATPase heterohexamers despite the homohexameric nature[28] (Fig. 1.1e). The evolutionary conservation of the ATPase and associated central/peripheral stalks suggests a related rotary catalytic mechanism is involved in T3S. The sorting platform is highly dynamic with data from *Yersinia*[84] and flagellar[85] showing continuous recycling of the components Although still not clearly understood, it is suggested that the sorting platform is involved in the targeting of effectors, typically in complex with chaperones that maintain them in a semi-unfolded state primed for secretion, to the injectisome prior to dissociation and subsequent insertion into the secretory channel. The ATPase and export gate are implicated in the energization of T3S. Secretion depends on the proton motif force (PMF)[86,87], and it is proposed that export gate SctV acts as protein-proton antiporter[88] although the nature of this and how it confers protein translocation is unclear. Although not the predominant energy source, the ATPase has been shown to improve the efficiency of the export gate via an interaction of the central stalk with the export gate cytoplasmic domains[87,89].

Together, the sorting platform, export gate and IM rings create a T3S competent complex capable of secreting the T3S-dependent components required to complete injectisome assembly including the needle filament (SctI and SctF), and downstream tip (SctA) and translocon (pore-forming) complex (SctB and SctE).

### 1.6.6 The OM secretin

The first atomic structure of a secretin pore from any system was the *Salmonella* T3SS secretin InvG (SctC) showing a massive double walled beta-barrel with 15-mer stoichiometry[26] (Fig. 1.1b). Subsequent structures of T2SS secretins demonstrated conservation in stoichiometry and architecture of the common core domains[90,91]. Based on the structure, it was proposed that, after targeting to the OM via the cognate pilotin InvH (a T3SS chaperone), InvG assembles to form a 15-mer pre-pore associated with the inner-leaflet of the OM. An essential aspect of this step is partial insertion (analogous to a cationic peptide) in the inner leaflet of the OM by a externalized highly conserved amphipathic loop motif, providing a proposed perturbing effect believed to support subsequent BAM-independent insertion of the adjacent 45 stranded beta-barrel transmembrane region into the OM[26] (Fig. 1.2c). Interestingly, the collapsed conformation of the predicted membrane spanning beta-sheet agrees with an invagination in the OM around the secretin in both *Shigella* and *Salmonella* injectisomes visualized using in situ cryo-ET[18,92]. Unresolved in earlier SPA structures, however, were the structural details of the N0 and N1 IM "coupling" domains, small modular domains at the secretin N-terminus that are varied between secretion systems depending on the varying and specialized IM components they couple to. Successful whole particle alignment revealed that this region remarkably contains 16 copies of the N0 and N1 domains, in contrast to the 15-mer OM beta-barrel pore, with the extra domain inserted in a restricted specific subset of locations assembled around SpaR and the first SpaP of the internal export apparatus[63] (Fig. 1.1b-c). How the IM and OM components couple has been enigmatic in light of the assumed 15- vs 24- fold symmetry mismatch. The observed

internal stoichiometry variation within the secretin provides an elegant evolutionary answer to this conundrum; changing the symmetry to 16-fold facilitates a repeating 8-fold interface composed of two InvG and three PrgH monomers. Specifically, the C-terminus of two out of every three PrgH monomers forms a three-stranded beta sheet that inserts between the N0 domains forming a circular 80-stranded beta-sheet, while the remaining PrgH adopts a similar secondary structure that instead wraps around the outside and under the beta-sheet, together creating an expansive and cohesive collective interface[63].

### 1.6.7 The needle filament

Within the NC base, the needle polymerizes atop the export apparatus with the pseudo-helical nature of SctRSTU defining the helical parameters. Structures of the *Salmonella* needle polymer PrgI (SctF) have been determined in isolation[80,93] and also from intact NC particles[94] showing an 11-start helical polymer with the N-terminal helix on the outside and the C-terminal helix lining the channel; however, how this was assembled within the injectisome, especially the role of PrgJ (SctI) historically thought to form an inner rod accounting for the span of the filament within the basal body scaffold, were only determined with the recent SPA structure of the NC[63] (Fig. 1.1c). Six PrgJ monomers were observed to assemble on the SpaR:SpaP pseudo-hexamer (Fig. 1.2e), functioning as an adaptor for downstream PrgI polymerization (Fig. 1.2g). Unexpectedly, the N-terminus of five PrgJ subunits (only the C-terminal helix of the first PrgJ is ordered) interacts with the InvG N0-N1 domains via beta-strand complementation further anchoring the filament within the NC lumen (Fig. 1.2f). Four distinct states with varying PrgI stoichiometries were captured in the SPA sample – one with no needle and secretin periplasmic

gate closed, two with short needles trapped in the secretin lumen with partially opened gates, and one with a fully assembled needle and the secretin gate fully opened - revealing not only the basis for needle assembly but also the structural changes associated with opening of the OM secretin periplasmic gate. These structures support a gating mechanism initiated by needle polymerization akin to opening a locked gate[26,63,80]. In the first stage –allosteric opening – needle assembly triggers a rearrangement of the N3 domain ring moving it from a conformation that supports the closed gate conformation, effectively unlocking it. The second stage – steric opening – involves the re-orientation of the inner beta-barrel hairpins upon continued needle polymerization which serves to push open the gate (Fig. 1.1c, 1.2g).

### 1.6.8 The translocators

The final components to assemble are the translocators, which terminate the needle polymer and form a pore in the host cell membrane (Fig. 1.1a, 1.2h). Upon attainment of the correct needle length sensed by ruler protein SctP, secretion switches from the early needle, rod and ruler substrates to the mid translocators[11]. The tip, SctA, forms a homopentadecameric cap on the needle, which acts as a bridge for the pore forming translocon that subsequently inserts into the host cell membrane. The tip has been imaged at lower resolution in *Shigella*[95], which suggests it continues the helical nature of the needle. The translocon pore contains the major and minor components SctB and SctE respectively, which are proposed to form only on insertion in the host cell membrane and the specific lipid environment therein and has been especially recalcitrant to both isolation and recombinant reconstitution. Remarkably, Liu and colleagues have used cryo-ET to image the association between *Salmonella* injectisomes and cultured epithelial cells,

capturing the intact translocon deployed in the host cell plasma membrane[96] (Fig. 1.1a). Although still at low-resolution with undefined stoichiometry, the study reveals a general donut shape approximately 13 nm in diameter with a central pore situated atop the needle tip complex. Interestingly, the study also defines an inward curve in local membrane morphology surrounding the inserted pore. Thus, the three membrane spanning regions of the injectisome are defined by local changes in membrane structure, likely exploited to facilitate membrane insertion and span.

### 1.6.9 The remaining challenges

In conclusion, the last several years have been an exciting time for T3SS structure and function with the advances in cryo-EM technology facilitating tremendous progress in the field including both high-resolution determination of varied assembly states and complimentary imaging of these nanomachines in their native cellular context. As we progress to a complete molecular understanding, several challenges remain: Firstly, some structures of individual components and subcomplexes are yet to be determined. Notable amongst these are the export gate central to secretion energization and regulation, and the associated assembled sorting platform complex. At the opposite end, the nature of the translocon pore embedded in the host cell membrane remains poorly understood. The ultimate aim is a high-resolution structure of the entire injectisome. In order to achieve this either more sensitive means of isolating the injectisome to preserve all components on purification for SPA is required, or, perhaps more feasible, with improvements in data collection and processing for cryo-ET, which, even now is capable of achieving sub-nanometer resolutions.

## 1.7 Secretion hierarchy

To allow all the components to assemble correctly, the T3SS is strictly regulated on the secretion hierarchy. Once the needle complex base and sorting platform are fully assembled, substrates are secreted in an orderly manner. The secreted components are divided into three groups, early substrates, which include inner rod (PrgJ) and needle filaments (PrgI), middle substrates, which are the translocon complex and late substrates, which are various effector proteins[11]. The early inner rod and needle are secreted first with rod polymerizing on the export apparatus, and then the needle filaments build on the short adaptor formed by the rod[63]. There is also evidence to suggest that proper assembly of the inner rod requires the presence of PrgI[97]. The secretion of the needle filament protein subunits is stopped once the polymerizing needle reaches a certain length. The length of the needle is also strictly regulated such that different species are distinguished by various needle lengths in keeping with the specific distance to the host surface.[11] There are two proteins involved in the needle length regulation in *Salmonella*, InvJ and SpaS. InvJ is believed to act as a flexible molecular ruler, measuring the growing needle length and signalling the stop of further polymerization once the required span is achieved [98,99]. SpaS is believed to be involved in the chronological switch of secretion from early to middle substrates, thus terminating the secretion of the PrgI needle[100,101]. The autocleavage of SpaS by an asparagine driven intein like mechanism involving a highly conserved NPTH sequence stops secretion of the inner rod and needle and initiates the secretion of the components of the translocator complex[100]. The autocleavage itself does not begin the secretion of the middle substrates; rather, it is believed the autocleavage allows SpaS to adopt a different

conformation and consequent set of non-covalent interactions with an alternate region of the export apparatus complex, in turn promoting the secretion of middle substrates[100]. Lastly, once the translocator complex is formed, InvE stops the secretion of the middle substrates and starts the secretion of the late substrates, which are the protein effectors essential for pathogenicity[102]. It is thought that InvE binds to the inner rod, thus stopping the secretion of the effector proteins[103,104]. When the translocator complex, a presumed pore forming complex, is properly formed, a signal is transmitted through the needle filament, leading to the release of InvE and thus allows the secretion of effector proteins.

Cryo-EM studies also help to reveal the path of the substrate during secretion. In an earlier study, one of the native substrates, SipA, was extended to three times its native length and labelled with a GFP at its C-terminus as a cytoplasmic anchor. This modified substrate was designed to fully occupy the secretion channel and prevent the escape of the substrate[65,105]. Although the substrate could only be characterized at low local resolution in the final cryo-EM map, essentially a tube-like structure, the secretion path it delineates through the needle is clear if not surprising. It supports the hypothesis of the path of secretion through the center channel of the needle complex with only a small conformational change in the export apparatus to accommodate its passage.

## 1.8 Thesis objective

The threat of antibiotic resistance promotes more and more research on developing new drugs and vaccines and finding new targets. The virulence factor is a good alternative to antibiotics as it poses less selection pressure on the bacteria, leading to less development of resistance to the drugs. Among all the potential virulence factors, the bacteria secretion machinery remains one of the most attractive targets for the new development of drugs and vaccines due to its crucial role in the survival and pathogenicity of bacteria. Understanding function and assembly is key to understanding how those massive secretion systems work and will help identify new targets and improve the existing drugs. With the rapid development of cryo-EM, the structural study of these massive secretion systems becomes possible, and obtaining a fully assembled T3SS will significantly improve our understanding of this nanomachine and aid in the new development of drugs and vaccines. The main topic of this thesis covers the structural and dynamic aspects of assembly of the needle complex of the T3SS.

Chapter 2 details a low resolution view of the overall T3SS needle complex at 7Å resolution and the data processing strategies used to obtain new high resolution structural information for key components of the injectisome as determined by cryo-EM methods[80]. Using partial signal subtraction and focused refinement, the inner membrane (IM) rings PrgH and PrgK reached the final resolution of 3.6Å and the outer membrane ring InvG to 3.9Å, sufficient to unambiguously trace the majority of the structural model including identification of side chains. In addition,

shearing off and isolation of the T3SS assembled needle filament (PrgI), allowed for data collection and subsequent helical reconstruction to 3.3Å. Overall, the filament structure agrees with the previously determined in-vitro reconstituted solid-state NMR (ssNMR) structure[20] with a few subtle but important differences due to the native sample vs. the in-vitro reconstituted sample. The pattern of helical charges within the hollow lumen of the needle in combination with secretion-based assays of point mutants therein hints at the possible mechanism of substrate passage. The IM ring structure in the needle complex determined here is near identical to the previously determined basal body IM structure determined in a needle-less assembly intermediate[26], indicating the inherent stability of this IM base platform during T3SS assembly. Lastly, the structure of the large gated OM portal, the secretin InvG, in an open conformation is the first for this broadly dispersed family (including in other bacterial secretion systems such as the Type II and Type IV pilus systems)[78]. The determined reconstructions directly show the major structural changes needed in the opening of the inner gate to allow passage of the needle substrate with a resulting change in pore size from 15Å (closed) to 75Å (open, needle bound form). Collectively, these structures, in combination with supporting point mutagenesis analysis, delineate the molecular details of the gating mechanism of the giant OM pore and set the foundation for future work on this fascinating nanomachine.

Chapter 3 details multiple high-resolution maps of the T3SS needle complex in multiple assembly stages as determined by single particle cryo-EM[63]. There are four maps, representing particle classes in different assembly stages, ranging from a needle-less complex base to a fully assembled needle complex. The series of maps shows the assembly steps of the needle complex

and shows multiple sub-complex structures at high resolution in the native environment. Importantly, the maps, generated by advanced data processing methods, unveiled the mystery behind the long-standing question regarding the apparent symmetry mismatch between the IM ring complex with 24-fold symmetry and the OM secretin ring complex with 15-fold symmetry (as well as poor local order in that coupling region). By re-evaluating the data processing and refinement strategy, newly generated reconstructions unambiguously showed local introduction of a single monomer in only the inner membrane coupling region of the OM secretin (the N-terminal most ring building motifs); the OM ring complex formed localized 16-fold symmetry at the interface between the IM ring complex while preserving the unambiguous 15-fold symmetry of the central beta region of the OM secretin pore. The coupling region therefore created a repeating 8-fold symmetric set of extensive intimate noncovalent interactions, supported by our mutagenesis studies and in keeping with the expected stable interface between the IM and OM components. In addition, the clearly resolved export apparatus and inner rod density in this work, the first for the T3SS system, answers prior questions as to their role in gating the inner membrane entrance of the system and setting the helicity and anchoring the extensive needle filament respectively. The structural information obtained from these maps, along with supporting data, revealed key steps in assembly and function of the needle complex, and the protocols established to get those maps provide an important foundation for further study of the injectisome's secretion and inhibition.

# Chapter 2: Cryo-EM analysis of the T3S injectisome reveals the structure of the needle and open secretin

## 2.1 Introduction

The injectisome is a specialized bacterial organelle that utilizes a type III secretion system (T3SS) to translocate effector proteins from the bacterial cytosol directly into a eukaryotic host. The injectisome has been the subject of extensive structural analysis from both reductionist and systems standpoints with high-resolution techniques such as X-ray crystallography and NMR used to solve the structures of some of the constituent proteins and traditionally lower resolution methods such as cryo-EM and tomography used to probe the overall architecture[16,17,106,18,23,24,26,71,75,93,100]. Recent advances in single-particle cryo-EM technologies have facilitated a radical improvement in the attainable resolution of the method. We recently published the cryo-EM structure of the prototypical *S.* Typhimurium SPI-1 basal body detailing the molecular architecture of the three oligomeric ring forming proteins that span the bacterial IM and OM in their assembled state, and revealing the massive double walled 60 stranded β-barrel structure of the family of OM secretin portals[26]. In that study, we used a T3S incompetent mutant lacking the N-terminal cytoplasmic domain of PrgH ($PrgH_{130-392}$), which assembled basal body complexes lacking the internal rod/needle filament and with the secretin periplasmic gate in a closed conformation. The atomic details governing the mechanism and specificity of molecular gating in eukaryotic and prokaryotic membrane transport has generated

exceptional interest in the past two decades. From smaller ion channels with pore sizes of 3-4 Å through to systems which transport large macromolecular substrates the underlying question is how these systems provide for passage of only select substrates while preventing deleterious release or uptake of other molecules? In the case of the giant secretin OM portals of bacteria, with inner lumens of more than 75 Å[78], the necessary gating becomes a fascinating and critical functional aspect. To address this and further elucidate the structural basis for injectisome assembly and function, including that of the central needle component, we have conducted here a cryo-EM analysis of the assembled native needle complex resulting in 3D reconstructions of the IM and OM rings at resolutions permitting atomic model building (3.6 Å and 3.9 Å respectively) and the 3D reconstruction of an isolated helical needle filament at 3.3 Å resolution. Collectively, these structures shed significant light on the process of injectisome assembly, needle architecture and span, and importantly capture the precise molecular details of the open gate conformation of the widely dispersed giant outer membrane secretin portal family.

## 2.2 Methods

### 2.2.1 Expression and purification

The needle complex was purified as follows with similar methods as previous work[16]. The ΔflhD knockout *Salmonella* strain SL1344[107,108]was transformed with a plasmid expressing the T3SS transcription activator HilA on a pBAD plasmid. The cells were grown in LB with addition of 0.3M NaCl with low aeration at 37 °C to $OD_{600}$ 0.6. Then 0.12 % L-arabinose was added to induce expression until the early stationary phase. Cells were harvested by low speed centrifugation and lysed by adding 0.35% lauryldimethylamine oxide (LDAO) in lysis buffer

(150 mM phosphate pH 7.4, 0.5 M sucrose, 1.4 mg/ml lysozyme, 12.2 mM EDTA). The needle complex was obtained by centrifugation at 230000g using a 45 Ti rotor (Beckman) for 45 mins and resuspended in buffer A (10 mM phosphate pH 7.4, 0.5 M NaCl, 0.5 % LDAO), before loading on to a 27.5% CsCl gradient for overnight centrifugation. All the fractions from the gradients were examined on SDS PAGE to determine which fractions contained the needle complex. The fractions that contained the needle complex were further purified on Superose 6 10/300 GL column (GE healthcare) with buffer B (10 mM phosphate pH 7.4, 0.5 M NaCl, 0.1 % LDAO).

Isolated needles were purified as follows: non-flagellated ΔPrgI *Salmonella* Typhimurium was complimented with PrgI and hilA on pBAD plasmid. The cells were grown under the same conditions as above. The filaments were removed by vigorously stirring. The intact cells were removed by low-speed centrifugation. Macromolecular contaminations were removed with a 45000g spin using a 45 Ti rotor and needle filaments were collected with a 120000g spin for 1 hour. The pellets were resuspended in TE buffer and loaded on 38.5% CsCl gradient. The fraction that contained needle filaments were used for Cryo-EM study.

### 2.2.2 Cryogenic electron microscopy

For the needle complex, a thin carbon layer was floated onto C-flat™ Holey Carbon (1.2/1.3) grids. Aliquots of 3 µl of the needle complex sample were applied to glow-discharged grids (20s on carbon side). The grids were blotted for 3 s at 100 % humidity with an offset of -20 and plunge-frozen into liquid ethane using a Vitrobot Mark IV (Thermo Fisher, formerly FEI

Company). Grids were imaged on a 300 keV Titan Krios cryogenic electron microscope (Thermo Fisher, formerly FEI Company) equipped with a Falcon 3 direct electron detector. Images were taken on the Falcon 3 in electron counting mode at a calibrated magnification of 47000, corresponding to 1.75 Å per physical pixel. The dose rate on the specimen was set to be 0.6 electrons per $Å^2$ per s and total exposure time was 60 s, resulting in a total dose of 40 electrons per $Å^2$. With dose fractionation set at 1.5 s per frame, each movie series contained 38 frames and each frame received a dose of 1.05 electrons per $Å^2$. Fully automated data collection was carried out using EPU software with a nominal defocus range set from −1.25 to −4 μm[109].

For the isolated needle, aliquots of 2.5 μl of the needle filament sample were applied to glow-discharged (60s on carbon side) Quantifoil grids (Copper, 300 mesh, R1.2/1.3). The grids were blotted for 3 s at 100% humidity and plunge-frozen into liquid ethane using a Vitrobot Mark IV (Thermo Fisher, formerly FEI Company). Grids were imaged on a 300 keV Titan Krios cryogenic electron microscope (Thermo Fisher, formerly FEI Company) equipped with a spherical aberration corrector, an energy filter (Gatan GIF Quantum) and a post-GIF Gatan K2 Summit direct electron detector. Images were taken on the K2 camera in dose-fractionation mode at a calibrated magnification of 37037, corresponding to 1.35 Å per physical pixel (0.675 Å per super-resolution pixel). The dose rate on the specimen was set to be 5.5 electrons per $Å^2$ per s and total exposure time was 10 s, resulting in a total dose of 55 electrons per $Å^2$. With dose fractionation set at 0.25 s per frame, each movie series contained 40 frames and each frame received a dose of 1.37 electrons per $Å^2$. An energy slit with a width of 20 eV was used during

data collection. Fully automated data collection was carried out using SerialEM with a nominal defocus range set from −1.2 to −2.5 μm[110].

### 2.2.3 Image processing

For the needle complex, a total of 1110 movie series were collected at 1.75 Å per pixel. Motion correction was done using MotionCor2[111]. The contrast transfer functions (CTFs) of the summed and dose weighted micrographs were determined using CTFFIND4[112]. Approximately 2,000 particles were manually boxed out from selected micrographs to generate reference-free 2D-class averages by Relion[56]. The representative 2D-class averages were also used as templates for automated particle picking for the entire dataset in Relion. With 88687 particles automatically picked, reference-free 2D classification was performed in Relion and 82678 particles were kept in the good class averages. 3D classification in Relion generated one good 3D map out of four with 58198 particles. 3D refinement was performed with these particles using Relion, without or with imposing 24-fold symmetry resulting in reconstructions to 9.4 Å and 4.6 Å respectively based on Fourier shell correlations (FSCs)[48] calculated from unmasked maps in Relion at 0.143 criterion without post-processing. After post-processing FSCs calculated from masked maps in Relion at 0.143 criterion is 7.4 Å and 3.6 Å respectively. Local resolution estimations were performed using Relion with the unmasked C24 map, showing that certain regions exhibit better resolvability than the overall resolution. To resolve the secretin region, particle subtraction and masked focus alignment refinement were used[67]. The non-secretin region was subtracted from the rest of the micrograph in Relion and the signal subtracted particles were subjected to reference free 2D-class averaging by Relion[56]. 3D classification in

Relion with 15-fold symmetry generated one good 3D map out of four with 26280 particles. 3D refinement with 15-fold symmetry results in the secretin region resolved at 4.4 Å based on FSCs calculated from unmasked maps in Relion at 0.143 criterion without post-processing and 4.1 Å after post-processing. Local resolution estimations were performed using relion with the unmasked C15 map, showing that certain regions exhibit better resolvability than the overall resolution. The secretin and N3 domain was processed in same way with 34096 particles and the final resolution after post-processing was 3.9 Å.

For the isolated needle, a total of 3753 movie series were collected at super resolution (0.675 Å per pixel). Motion correction was done using Unblur[113], with data binned by two (1.35 Å per pixel after binning), and all 40 frames were aligned and summed to a single micrograph with dose-filtering using Sum_movie for further processing[113]. The contrast transfer functions (CTFs) of the summed micrographs were determined using CTFFIND4[112]. Helical processing the helical filament was done following the procedures in Relion[56,114]. Manual picking of the helical filament was done for 30 micrographs and ~6,000 helical segments were extracted that generated initial 2D-class averages as templates for automated particle picking for the entire dataset. With ~1,220,000 particles automatically picked, reference-free 2D classification was performed and ~867,000 particles were kept in the good class averages. Initial helical symmetry parameters were known and used from earlier work[19]. 3D classification with local searches of helical symmetry generated four classes with variations in helical symmetries. Further high-resolution classifications and auto refinements were performed in each helical group and generated final maps after semi-automated post-processing. The final map of Group 1 (41,700

particles included) was solved at 6.8 Å resolution with helical rise of 4.45 Å and turn 63.34°. The final map of Group 2/4 (combined initial Class 2 and 4, 53,000 particles included) was solved at 3.3 Å resolution with helical rise of 4.33 Å and turn 63.34°. The final map of Group 3 (16,500 particles included) was solved at 7.2 Å resolution with helical rise of 4.21 Å and turn 63.44°. All resolutions were determined at a criteria of 0.143 FSC using gold-standard refinement procedures and high-resolution noise substitution to correct soft-mask effects[48].

### 2.2.4 Model building and refinement

For the IM rings, the cryo-EM basal body atomic structure of the PrgH and PrgK rings (PDB 5TCP) was docked into the 24-fold averaged needle complex reconstruction with Chimera[115] and refined using both Phenix real space refine[116] and density-guided iterative local refinement as implemented in Rosetta[117] taking into account the ring symmetry at all steps. The final model has a good fit to the map (Chimera correlation coefficient (CC) score 0.88) and good statistics (Molprobity[118] score 1.92, Ramachandran favored 96.82 % & outliers 0 %, EMRinger[119] score of 2.63). For InvG, the structure of the isolated closed conformation (PDB 5TCQ) was docked into the reconstruction using Chimera and the structurally different regions were rebuilt using COOT[120]. Both Phenix and Rosetta were used for refinement as above. The final model has good fit to the map (Chimera CC 0.87) and good statistics (Molprobity score 2.4, Ramachandran favored 94.76 %, outliers 0 %, EMRinger score of 1.23). For the isolated needle, the density permitted *de novo* model building in COOT[121] and further refinement in both Phenix and Rosetta (using helical symmetry restraints). The final model has a good fit to the map

(Chimera CC 0.92) and good statistics (Molprobity score 2.7, Ramachandran favored 96.05 %, outliers 0 %) and an EMRinger score of 1.83.

### 2.2.5 Liquid chromatography tandem mass spectrometry

In-gel digestion for the basal body sample was carried out as described[122]. Digested proteins were loaded onto a Bruker Impact II Q-ToF mass spectrometer. Peptide separation was carried out on a 50 cm in-house packed 75 µm C18 column by a Proxeon EasynLC UPLC system, using 120 minute water:acetonitrile gradients. Eluted peptides were ionized in positive ion mode, collecting MS/MS spectra for the top 15 peaks >500 counts, with a 30 second dynamic exclusion list.

Resulting data files were loaded into MaxQuant v1.5.1.0 for analysis[123]. Identification used a 0.006 da MS tolerance and 40 ppm MSMS tolerance, and a *Salmonella* specific protein database containing all annotated *Salmonella* proteins extracted from Uniprot.

## 2.3 Results

We purified intact *S.* Typhimurium SPI-1 needle complexes as previously described[16]. Single particle reconstruction resulted in unsymmeterized maps to 7.4 Å resolution as judged by gold-standard FSC[48] revealing the molecular envelope for the IM and OM rings of the basal body and the encompassed substructures with attached needle filament (Fig. 2.1a). Comparison to the previously determined basal body reconstruction[26] reveals the general localized changes within the core secretin domains necessary to accommodate the assembled needle filament;

notably the extended open orientation of the periplasmic gate (Fig. 2.1b). Outside of these regions, the overall structure is remarkably static between assembly states.

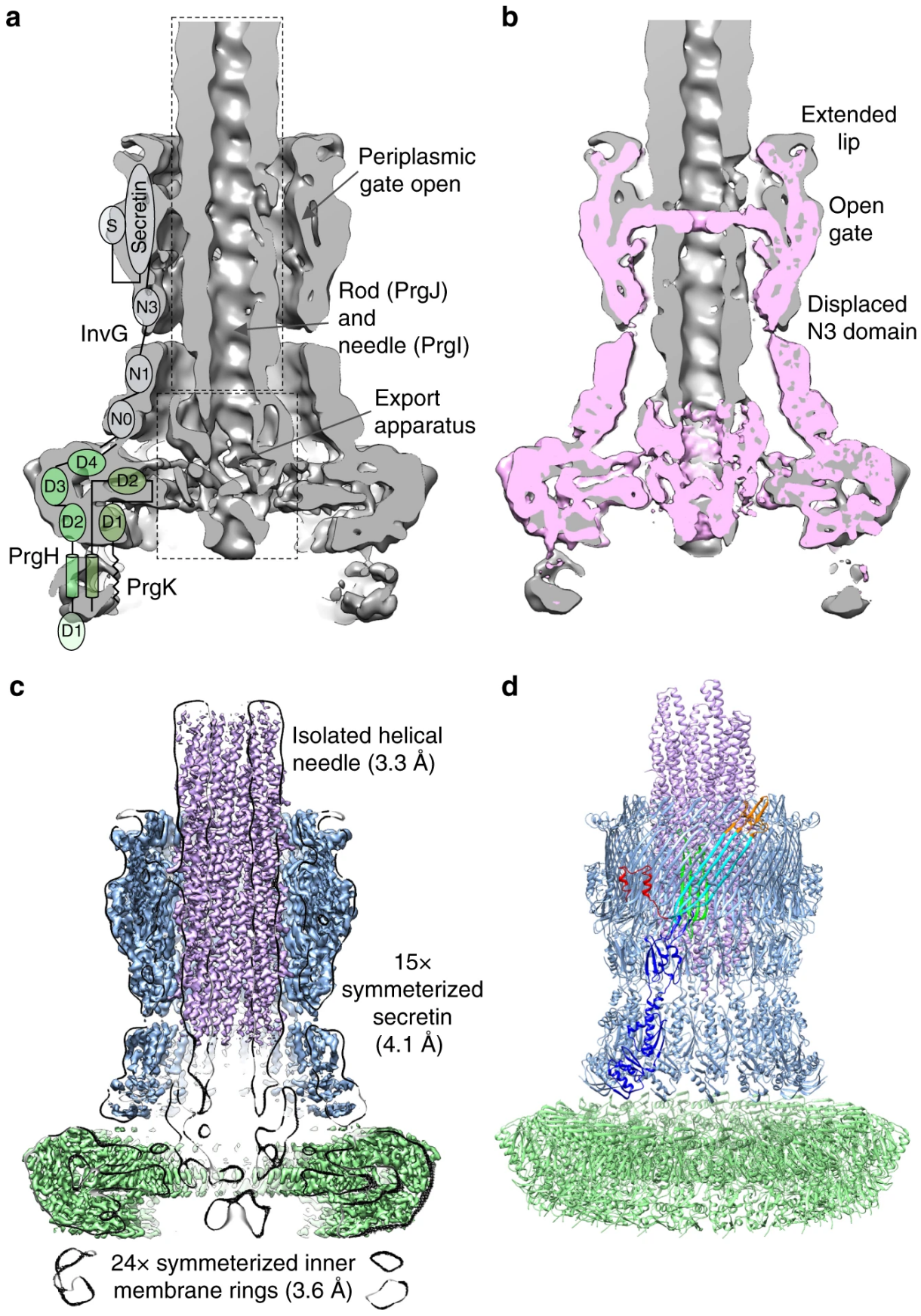

a
Periplasmic gate open
S
Secretin
N3
InvG
Rod (PrgJ) and needle (PrgI)
N1
N0
Export apparatus
D4
D3
D2
D2
D1
PrgH
PrgK
D1
b
Extended lip
Open gate
Displaced N3 domain
c
Isolated helical needle (3.3 Å)
15× symmeterized secretin (4.1 Å)
24× symmeterized inner membrane rings (3.6 Å)
d

**Figure 2.1: Cryo-EM structures of the injectisome needle complex and isolated needle**

**(a) Needle complex C1 reconstruction (low pass filtered from 7.4 Å reconstruction to highlight overall features) cut away at the mid-section. The domain annotation of PrgH, PrgK and InvG is overlaid on the left. Boxed regions indicate the periplasmic region of the export apparatus and the rod/needle filament. (b) Central slice view of needle complex reconstruction (grey) overlaid with the 6.3 Å basal body reconstruction (EMD-8400) (pink). (c) Reconstructions for the 24-fold averaged IM rings (green; 3.6 Å resolution), the 15-fold averaged secretin (blue; 4.1 Å resolution) and the isolated needle (magenta; 3.3 Å resolution). High-resolution reconstructions overlaid on C1 reconstruction shown as central slice (black). The needle was fit into the needle complex C1 map using Chimera and agrees well with the wider part of the needle complex filament. (d) Refined structures for $InvG_{34-557}$ (blue), $PrgH_{171-364}$ (green), $PrgK_{20-203}$ (green) and $PrgI_{3-80}$ (magenta). One monomer encompassing $InvG_{34-557}$ is coloured according to structural domains: N0-N3 domains (blue); outer β-sheet (cyan); inner β-sheet (green); secretin domain lip (orange); S domain (red).**

### 2.3.1 The structure of the 24-mer IM rings

The 24-fold features of the IM rings in the C1 reconstruction were clearly visible and further C24 symmetry averaging resulted in a reconstruction to 3.6 Å resolution (compared to 4.3 Å for the basal body IM rings; EMD-8398) allowing for refinement of the atomic structures of the periplasmic domains of $PrgH_{171-364}$ and $PrgK_{20-203}$ (Fig. 2.1c-d, Fig. 1c, Table A1). Consistent with the conserved dimensions of the IM rings between the needle complex here and that of the previous basal body reconstructions (Fig. 2.1b), the refined PrgH and PrgK ring atomic models determined from each overlap closely (overall Cα RMSD 0.7 Å) suggesting an inherently highly

stable structure with no significant conformational change required in this region upon needle assembly.

### 2.3.2 The export apparatus and needle filament

The low-resolution substructure density localized within the central lumen of the stacked IM PrgH and PrgK rings at the base of the needle filament has historically been referred to as the socket and cup. It is present in native and T3S incompetent mutants alike[25,26], the latter of which consequently lack the rod/needle filament (themselves T3S substrates). Mutational, mass spectrometry and low resolution EM analysis have attributed this substructure density to regions of the IM associated export apparatus components SpaP, SpaQ, SpaR and SpaS[72,73,100], with knockout of one or more removing the substructure in vitro and in situ[18,72]. Remarkably, a recent cryo-EM structure of the isolated recombinant export apparatus complex from the bacterial flagellar system (long recognized to encode several analogous IM components to the injectisome) shows that this substructure is composed of the flagellar orthologues FliP/Q/R (25-30% sequence identity with *Salmonella* SPI-1 SpaP/Q/R) in their entirety despite predictions they were integral membrane proteins with 4, 2 and 6 predicted transmembrane domains respectively. The complex adopts a helical assembly of dimensions ~90 x ~60 x ~60 Å with 5:4:1 stoichiometry (FliP/Q/R) with the single FliR bridging the IM facing FliQ tetramer and the more periplasmic facing FliP pentamer and elegantly explains previous cross-linking and mass spectrometry studies[21,72–74]. Complicated by the local symmetry mismatch (e.g. with the 24mer IM rings which encompass the export apparatus) and potential relative conformational variation, we have been unable to resolve the high resolution structural features of this

substructure in the current needle complex data or in the prior basal body reconstruction. However, whereas the latter presented a rather flat, symmetrical pore due to averaging out of the local structure (see Fig. 2.2b), the needle complex reconstruction here provides an inherently more detailed, asymmetric substructure than previously observed (Fig. 2.2a,c). Overlay of the flagellar FliP/Q/R structure matches well throughout suggesting the structure of the flagellar and non-flagellar T3SS export apparatus will be conserved, providing evidence for the precise span of the export apparatus within the assembled T3SS (Fig. 2.2d).

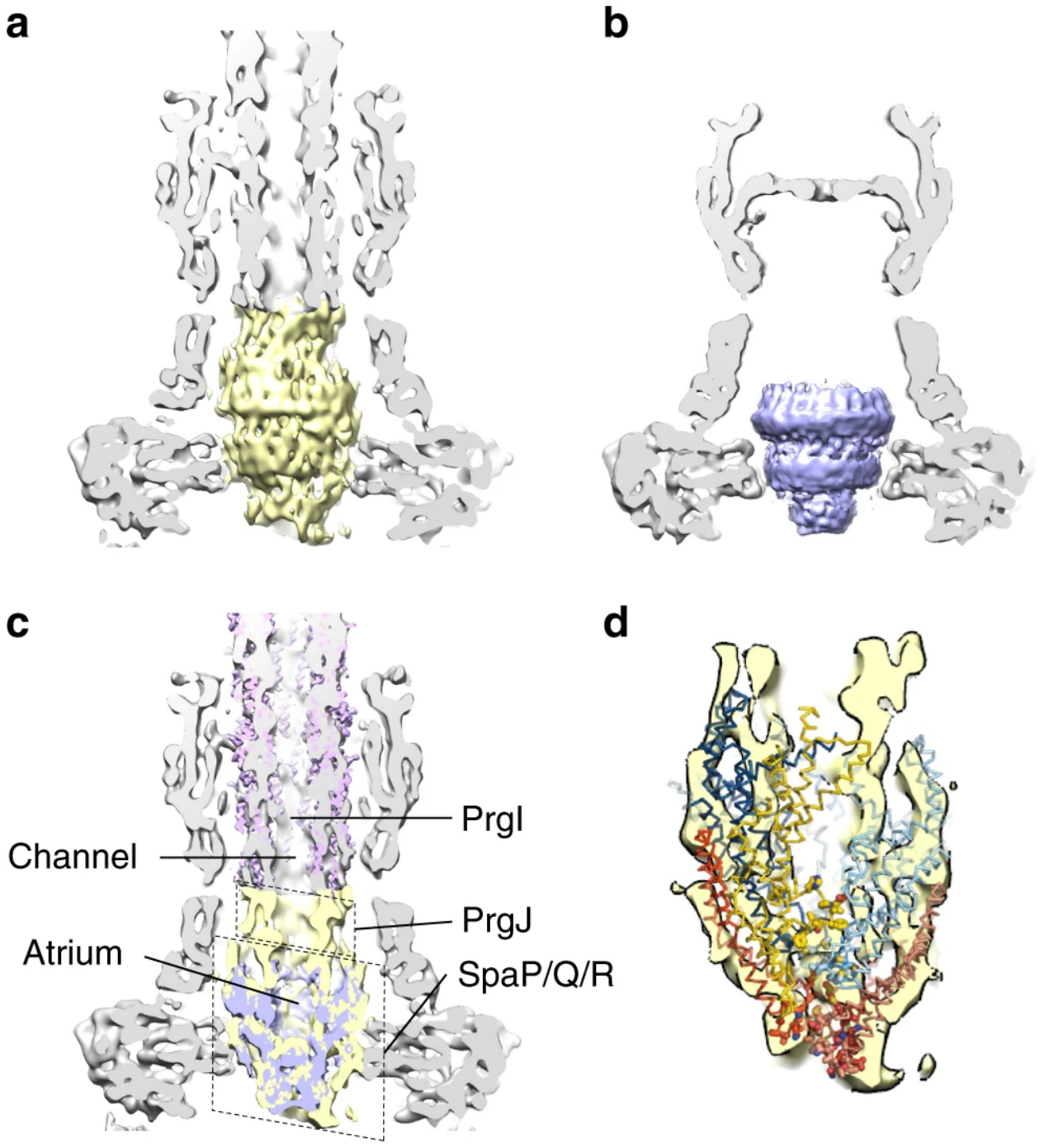



**Figure 2.2: The export apparatus forms an asymmetric substructure composed of SpaP and SpaQ and SpaR**

**(a) The region of the needle complex C1 reconstruction corresponding to the export apparatus and putative rod colored yellow. Slabbed remainder of reconstruction colored grey. (b) The region of previous C1 basal body reconstruction**[26] **corresponding to the export apparatus colored purple. The map features a flat, symmetrical pore, averaged out by the signal of the 24mer IM rings (24mer repeating features evident on upper surface). Slabbed remainder of reconstruction colored grey. (c) Slabbed view of overlay of (a) and (b) showing export apparatus region from needle complex map (yellow) is asymmetric compared to the same region from the basal body map (purple) (lower box). The 3.3 Å needle map (magenta) fits well (correlation coefficient = 0.95) into the wider part of the filament density. We propose the boxed narrower region between the base of the filament and the export apparatus represents the PrgJ rod, which has been proposed to form a short adaptor between export apparatus and needle**[74]**. (d) The recent structure of the *Salmonella* flagellar export apparatus**[124] **FliP (blue)/Q (red)/R (yellow)) forms a complex with stoichiometry of 5:4:1 and overlays very well with the needle complex export apparatus substructure. The subunits form a helical assembly ideal as a structural foundation for the assembly of the rod and needle filament. Atomic model figure reproduced under the CC-BY 4.0 International license.**

The needle filament observed in our reconstruction forms a continuation of the export apparatus with the internal channel diameter the same as the export apparatus atrium (Fig. 2.1b, Fig. 2.2c). The needle (PrgI) is proposed to be initiated by a homologous protein historically termed the rod (PrgJ) - both detected by MS analysis here (Table A2) but absent, as designed, from the purified secretion incompetent basal body[26]. Although the general dimensions and span of the filament are clearly observed, like with the export apparatus and the issues with localized symmetry mismatch with surrounding components, we were unable to sufficiently resolve the high-resolution structural features within the assembled needle complex to allow model building and

refinement. To further study the atomic details of the needle therefore, we isolated native filaments from assembled needle complexes and subjected them to cryo-EM analysis resulting in a reconstruction with helical averaging to 3.3 Å resolution with excellent density for side chains throughout (Fig. A2, Fig. 2.3c, Table A1) allowing us to build a de novo model of PrgI lacking only the first two N-terminal residues (Fig. 2.1c-d, Fig. 2.3). The PrgI monomer adopts a helix-turn-helix motif that polymerizes via an extensive network of interactions along both helices with surrounding subunits (Fig. 2.3a,b). The structure supports the solid-state NMR model of PrgI presented by Loquet et al.[93] (Cα RMSD 1.5 Å over 78 residues), which was fundamentally different from prior low resolution cryo-EM derived models[19,95]; however, some notable differences in the monomeric structure and helical packing as well as the definition of the majority of side chain positions extend our understanding of the intra-subunit interactions and the internal channel passaged by effectors. The kinked loop in the N-terminal helix (residues 20-23) is more extended and forms a tighter interface with the N-terminus of monomer i + 6 while the C-terminal Arg80 has its side chain and carboxylate flipped, altering the interactions with monomers i - 1, i – 5 and i – 6 (Fig. 2.3d,e). The refined polymer forms a helical assembly with approximately 11 subunits per two turns but a minor difference in the helical rise compared to the solid-state NMR model (helical rise 4.33 Å vs ~4.2 Å respectively) creates a small but significantly accumulating shift in subunit packing (Fig. 2.3a,b). Along with the specific structural changes at the C-terminus, this creates a right handed spiralled groove with polar and hydrophobic character extending along the needle lumen which was not observed in the NMR model, with a cluster of conserved residues around the C-terminus defining the raised edges of the groove (Fig. 2.3f-h). Taking into account the resolved side-chains here (Fig. 2.3b,c), the axial

diameter is considerably less than previously measured in lower resolution reconstructions[19,125,126] (~15 Å vs ~25 Å, only very closely accommodating a single helix for example) suggesting unfolded or partially unfolded secreted effectors would necessarily track along the deeper and longer helical path during passage (Fig. 2.3f-h). Indeed, previous scanning alanine mutations in the Shigella orthologue MxiH showed that mutation of conserved charged residues that here define the raised groove (equivalent to PrgI Lys66, Lys69, Asp70, Arg80; Fig. 3e-g) had differential effects on needle polymerization and length, secretion, and regulation[127] suggesting a specific contribution of the interior needle channel surface in the hierarchical secretion of T3S effectors as opposed to merely providing a channel for passive diffusion. The helical nature of the secretion channel also perhaps raises the intriguing possibility that an as yet undefined rotational force, perhaps linked to energy from the proton motive force required for secretion[86,87], might contribute to effector secretion. Fitting of the cryo-EM derived model into the C1 needle complex map shows a good agreement (Chimera correlation coefficient = 0.95) between the wider part of the filament density (Fig. 2.1c, Fig. 2.2c). This needle placement suggested by our data here, anchored deep into the secretin lumen in keeping with its needed stability within the relatively severe mechanical environment of the infected cell, is consistent with a recently proposed role of the protein PrgJ as a relatively small bridge or adaptor between the IM export apparatus and the needle filament[74] rather than an extended rod as its historical name unfortunately had predicted (here accounting for the narrowing region of density at the base of the filament (Fig. 2.2c)). Taken together, these structures now demonstrate that the SpaP/Q/R export apparatus complex is positioned to provide a structural foundation for the assembly of the needle filament with the helical, pseudo-hexameric arrangement of the central 5x

SpaP and 1x SpaR subunits an ideal template for the downstream T3S dependent polymerization of the helical filament, with several (likely hexameric) copies of PrgJ (predicted stoichiometry 3-6 by mass spectrometric analysis[74]) putatively assembling onto the SpaR/P platform (cross-linking demonstrates close interaction[73]) and in turn providing a helical template to initiate the helical needle polymerization.

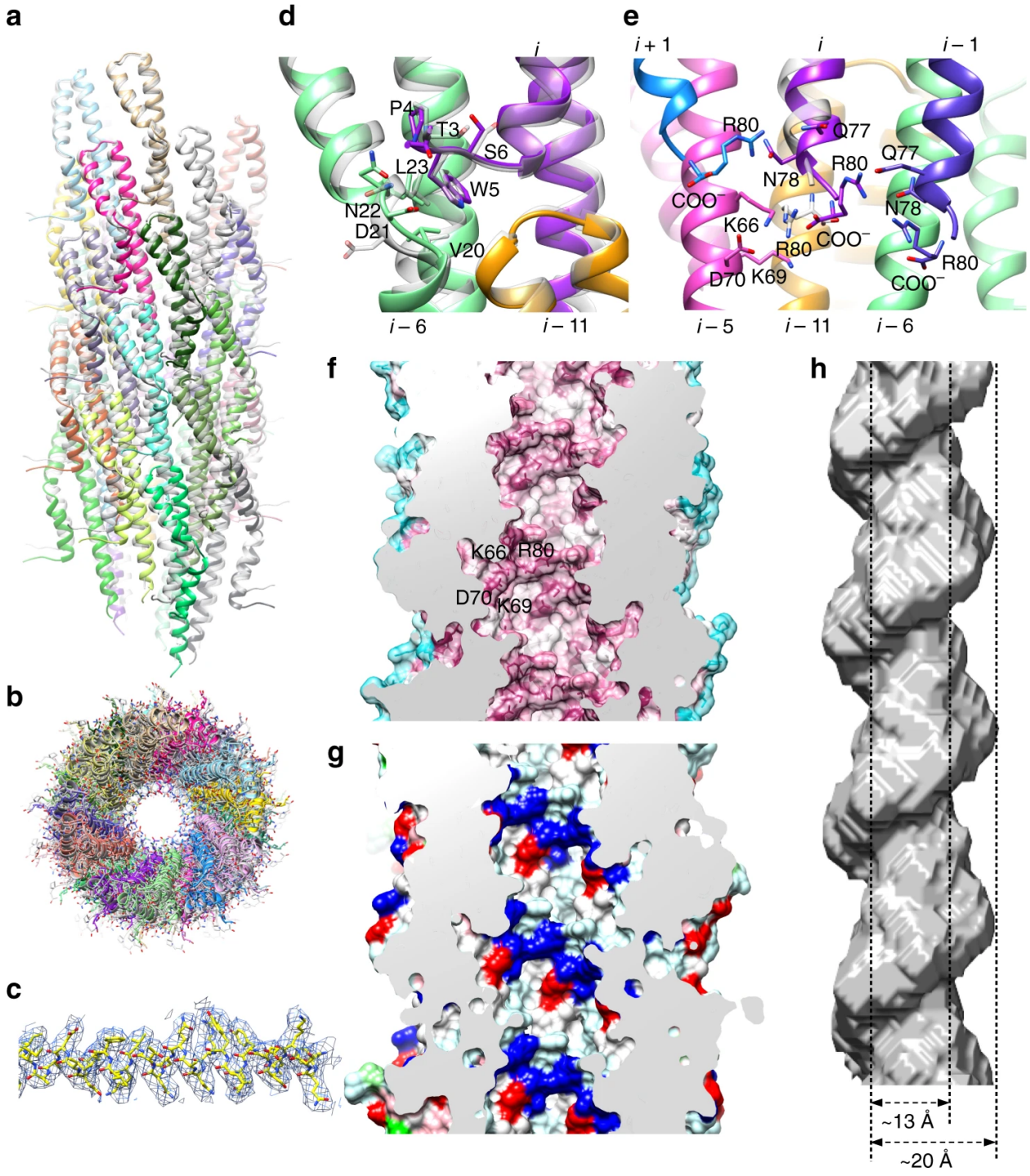



**Figure 2.3: Structure of the PrgI needle filament at 3.3 Å defines a spiral groove for effector secretion**

**(a) and (b) Side and top views of the helical packing of the PrgI needle with 5.7 monomers per turn and a helical rise of 4.33 Å. The ssNMR structure of PrgI (PDB 2LPZ; grey transparent) is overlaid showing the subtle differences in helical parameters create a small but accumulating change in subunit packing. Side chains, resolved in the reconstruction here, are shown in b highlighting the constriction of the interior channel. (c) Representative density (3.3 Å resolution). (d) Zoomed in view showing interface of *i* , *i* - 6 and *i* - 11 monomers formed around the N-terminal loop of *i*. The ssNMR structure of PrgI (grey transparent) is overlaid. The variable N-terminal loop here is packed closer to the kinked helix of monomer *i* - 6. The kink (residues 20-23) adopts a different conformation compared to the ssNMR structure. (e) Zoomed view of the interaction network around the C-terminal $Arg_{80}$. The ssNMR structure is overlaid on monomer *i* showing the $Arg_{80}$ side chain and carboxylate in a flipped orientation. Here, the $Arg_{80}$ side chain guanidinium group interacts with $Gln_{77}$ and $Asn_{78}$ on monomer *i* – 1 while the carboxylate interacts with $Lys_{66}$ on *i* – 5. (f) The interior channel is significantly more conserved than the outer surface (conservation colored from cyan (low) to maroon (high)). The cluster of conserved residues around the C-terminus as in e define the raised edge of a right handed spiral that extends the length of the lumen. The groove is lined by repeating deeper pockets defining the path of effector secretion. (g) Same view as in f colored according to residue type: hydrophobic – grey, aromatic – light pink, polar – light cyan, positive – blue, negative – red, cysteine – light yellow, proline – light green, glycine – green. The raised edge of the groove is demarcated by charged residues while the groove itself is predominantly polar and hydrophobic. (h) Surface corresponding to the interior needle lumen highlighting the right handed spiral and dimensions of the channel.**

### 2.3.3 The gate open state of the OM secretin

In these needle complex samples, the strong 24mer symmetry of the IM rings dominates the 2D class averages, a consequence of which is that the region of the C1 needle complex

reconstruction corresponding to the ~1 MDa OM secretin component is somewhat smeared out and lacked obvious symmetric or high-resolution features. We therefore employed a partial signal subtraction and masked focussed local refinement procedure[67] with the obvious 15-fold symmetry resulting used subsequently in averaging to produce reconstructions to 4.1 Å resolution for the entire secretin and 3.9 Å resolution for the core secretin structure encompassing the N3, secretin and S domains (Fig. 2.1c-d, Fig. A3). The density in the core region ($InvG_{176-557}$) showed defined secondary structure with the majority of side chains clearly visible (Fig. A3h) allowing for unambiguous building and refinement of the structure including significant additionally resolved and structurally changed regions from the previously determined closed form[26]. Although the more peripheral N-terminal N0 and N1 domains were of lower local resolution (Fig. A3g), the map was of sufficient quality to reveal secondary structure elements for the N1 domain in particular allowing us to position 15 copies of our prior $InvG_{34-173}$ crystal structure for that domain (PDB 4G08[25]) (Fig. A5e) further expanding the defined model of the injectisome. The final secretin model spanning all domains (residues 34-557, N3 domain loop 228-251 disordered; Fig. 2.4, Table A1) was refined with density restrained symmetry refinement in Rosetta and notably reveals the dramatic conformational changes involved in secretin gate opening during needle assembly, a fundamental and functionally critical aspect of this giant gated pore family.

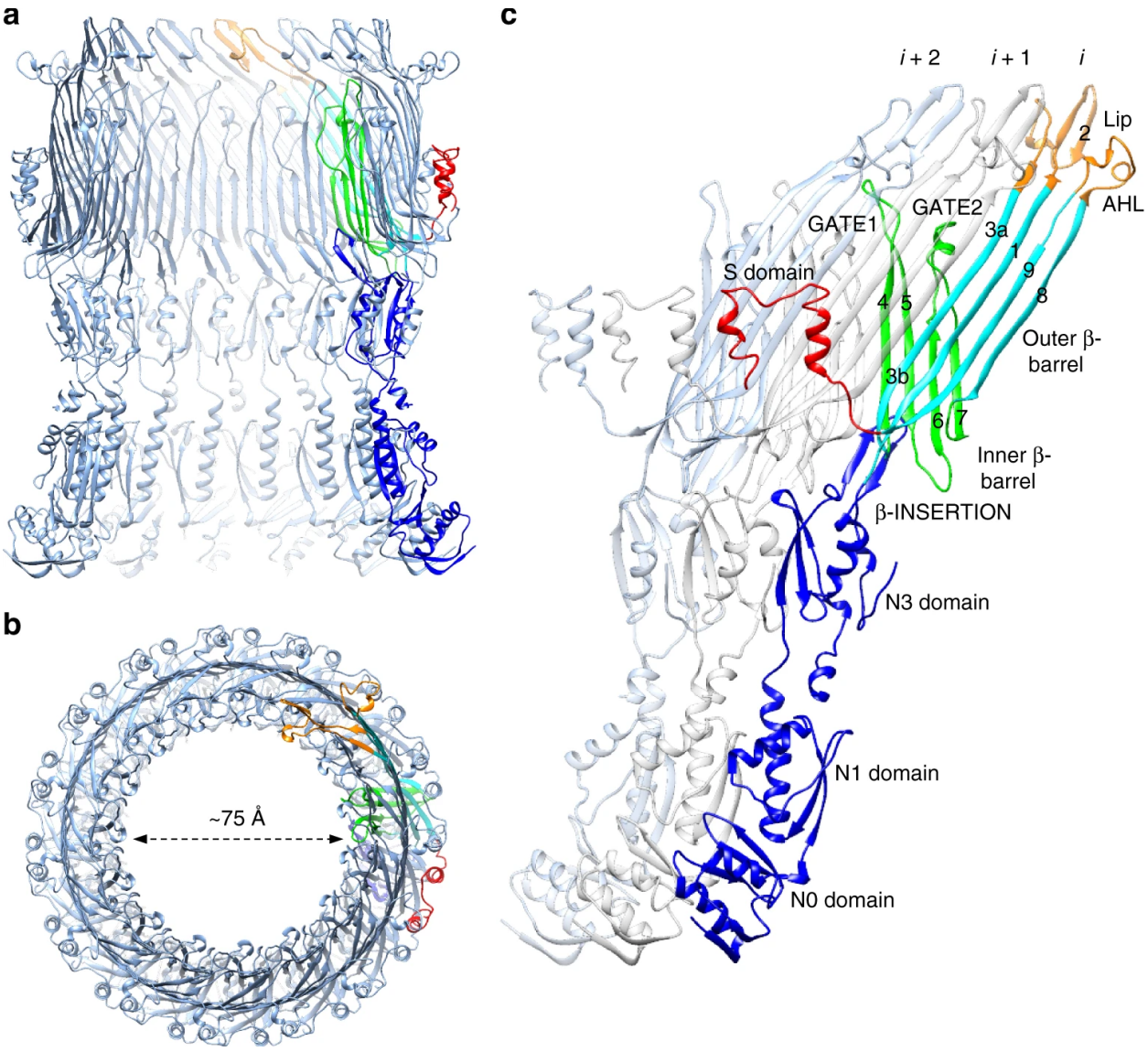



**Figure 2.4: Structure of the InvG secretin pore in the open state**

**InvG$_{34-557}$ secretin pore structure viewed from (a) side slab, (b) top (OM perspective) and (c) three monomers highlighting inter- and intra- domain packing of monomers *i*, *i* + 1 and *i* + 2. One monomer coloured according to structural features as per Figure 1. Secretin domain β-strand numbers as per our previous closed InvG structure**[26] **indicated in c.**

Overlay of the needle complex and prior closed basal body reconstructions illustrates that the primary sites of conformational change upon needle assembly are in the conserved core secretin region: the N3 ring domain - the tightly oligomerized foundation upon which the secretin assembles[26] - and the inner β-barrel and membrane associating lip of the central secretin domain, all of which have reoriented to accommodate the presence of the needle filament (Fig. 2.1b). The N-terminal N0 and N1 domains, outer β-barrel and C-terminal surface exposed S domain conversely show very little structural difference between assembly states; as with the static IM ring components they appear to provide a highly stable scaffold or shell to buttress the inner conformational changes required for needle passage. The most dramatic conformational change upon needle assembly is in the inner β-barrel. In the previously observed closed state[26], the extended β-hairpin formed by strands 4 and 5 (numbering consistent with the closed InvG structure[26], here termed GATE1 based on mutational analysis in secretin pIV[128]) is kinked and twisted around residues $Asn_{386}$ and conserved $Gly_{407}$, with the spoke like radial projections afforded by the pentadecameric symmetry forming a collective barrier across the ~75 Å secretin lumen that is permeable only to small molecules[128–130]. In the T2SS, mutation of the $Asn_{386}$ equivalent in PulD ($Gly_{458}$) increased antibiotic sensitivity[131] and accordingly resulted in a partial opening of the gate in the cryo-EM structure of GspD ($Gly_{453}$)[91]. In the fully open state defined here, the GATE1 hairpin is now unkinked, with the residues at the tip having moved position by a remarkable ~40 Å (measured from the $Gly_{395}$ Cα) (Fig. 2.5). The open position is stabilized by intimate packing against the outer β-barrel of itself and neighbouring monomer $i + 1$ (monomers defined clockwise looking down from the OM),

and extending vertically toward the lip region of monomer *i* + 2 (Fig. 2.4c, 2.5b,d) with a dramatic increase in buried surface area between the inner and outer β-barrels (2111 $Å^2$ per monomer compared to 1153 $Å^2$ in the closed state) (Fig. 2.5b-e). Side chain density is defined for the large majority of residues along the length of both GATE1 strands with the exception of 5 residues ($Glu_{396}$-$Ala_{400}$) at the tip. Although no longer kinked, the twist in GATE1 is maintained in the extended state such that strand 4 is oriented closer to the outer β-barrel forming multiple specific interactions. Of note, $Lys_{392}$ and $Ile_{394}$ at the C-terminus interact with residues on the outer β-barrel and lip of monomers *i* + 1 and *i* + 2 (Fig. 2.5b,d). In our previous secretion assays, mutation of either $Lys_{392}$ or $Ile_{394}$, as well as $Gly_{395}$, which forms a tight turn in both closed and open gate conformations, significantly impacted secretion[26]. The inner β-barrel strands 6 and 7 (termed GATE2 here) are straightened around $Gly_{430}$ and $Gly_{451}$ compared to the closed form. More significantly, the GATE2 loop (residues $Asp_{433}$-$Gly_{451}$) undergoes a 180° rotation, effectively repositioning the regions oriented above the GATE1 hairpin in the closed state against the outer β-barrel in the open state (Fig. 2.5). These changes in GATE1 and GATE2 of the inner β-barrel result in an increase in the secretin pore diameter from ~15 Å in the closed state to a remarkable ~75 Å in the open state (Fig. 2.4b, 2.5a).

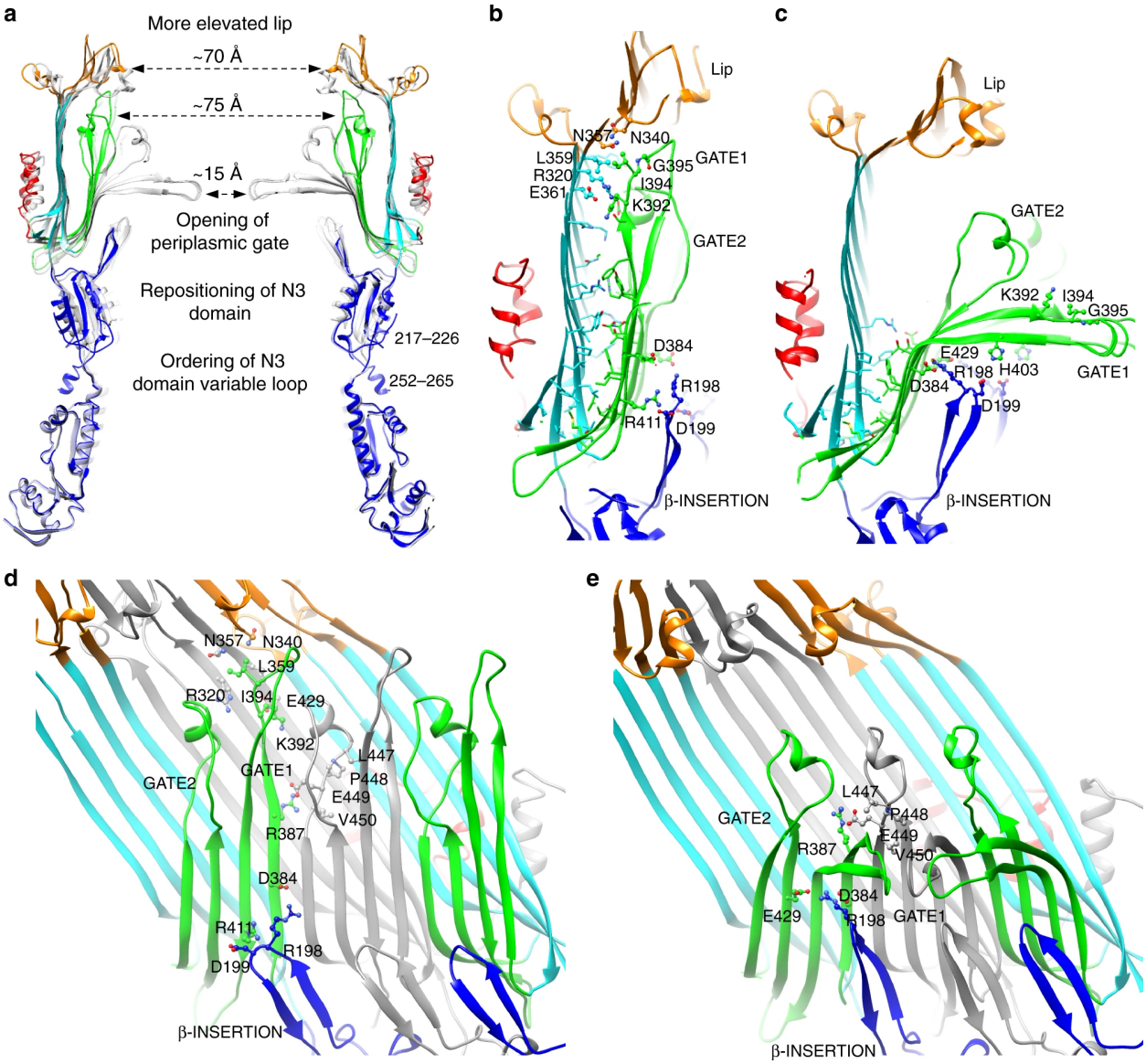



**Figure 2.5: Structural changes in involved in InvG gate opening**

**(a) Comparison of the InvG secretin gate open state (coloured as per Figure 2.1) and the closed state (grey transparent). The major structural changes are the ordering of the N3 domain variable loop (residues 217-226 and 252-265, disordered in other secretin structures) and accompanying change in N3 domain position, the opening of the periplasmic gate involving the repositioning of GATE1 and GATE2, and the more elevated lip region caused by interactions with GATE1 and the insertion of the assembled needle filament. (b) and (c)**

**– side view - and (d) and (e) – inside view - compare the interface between outer and inner β-barrels in the open and closed states respectively with accompanying interface areas of 2111 Å$^2$ and 1153 Å$^2$, residues forming the open and closed interfaces shown as sticks in b and c. The core interface at the base of the β-sandwich formed by the outer and inner β-barrels is predominantly hydrophobic while the region formed by the extension of the GATE1 and GATE2 hairpins is mostly polar in nature. Key interactions defining the open and closed conformations between the N3 domain β-INSERTION, the inner β-barrel GATE1 and GATE2 motifs, and the upper outer β-barrel and lip are labelled and shown as ball and stick. The middle subunit is coloured grey in d and e to define the outer and inner β-barrels. The closed gate conformation is supported by interactions of the N3 β-INSERTION, specifically Arg$_{198}$, with the kinked regions of the GATE1 (Asp$_{384}$) and GATE2 (Glu$_{429}$) hairpins and further supported by the surrounding network of interactions. In the open gate conformation, the N3 β-INSERTION interface is disrupted and the gate forming GATE1 hairpin is now extended toward the lip and packed against the outer β-barrel with Ile$_{394}$ at the tip packing in a hydrophobic notch formed by the side chains of Arg$_{320}$, Asn$_{340}$, Asn$_{357}$ and Leu$_{359}$. The GATE2 hairpin undergoes a significant rotation with residues stabilizing the closed gate - Leu$_{447}$, Pro$_{448}$, Glu$_{449}$ and Val$_{450}$ – packing against the outer β-barrel wall. A salt bridge between GATE2 Glu$_{449}$ and GATE1 Arg$_{387}$ is maintained between closed and open conformations.**

The secretin outermost lip constitutes a kinked 45-stranded β-barrel, capping the top of the 60-stranded outer β-barrel wall[26] (Fig. 2.4). With one strand less per monomer, the well-ordered lip was observed in the basal body closed reconstruction to collapse inwards to structurally accommodate the reduced strand count. Decorated on the exterior by an amphipathic helical loop (the AHL) - the most conserved sequence across the secretin family[26] – this region was proposed on structural features to be central to membrane association, BAM-independent

insertion and OM span[26]. In the needle assembled open state here, the lip has been pushed up and out to accommodate the passage of the filament, increasing the outer pore diameter from ~65 Å in the closed state to ~70 Å in the open state (Fig. 2.5a). The presence of the needle supports a more upright lip orientation consistent with the recent in situ cryo-ET structure of the *S.* Typhimurium T3SS which demonstrated the membrane invagination we had noted previously[18,92,132] to be more pronounced in a mutant lacking the needle compared to the assembled needle complex[18] (Fig. A4). Interestingly, recent structures of several T2SS secretins have revealed their lip region consists of 4 rather than 3 β-strands per monomer and as such present a non-kinked continuation of the outer β-barrel wall [90,91] (Fig. A5a,b). Sequence analysis suggests the three-stranded lip architecture observed for InvG will be common in other T3SS secretins and may reflect the necessity for a closer fitting and more stable interface with the needle, which is more statically attached in the injectisome than, by comparison, the dynamically attached (pseudo)pilus of the T2SS or T4PS which are in a fluctuating state of extension and retraction.

The N3 domain abuts the periplasmic base of the inner β-barrel formed by GATE1 and GATE2 (Fig. 2.4, 2.5) and has been demonstrated through mutagenesis to be essential for secretin oligomerization[131,133,134]. N3 belongs to the family of small mixed α/β domains we previously termed ring building motifs[22] (RBMs) which have been observed in numerous ring forming components from different secretion systems. Although often sharing a common general oligomerization interface[22,25,128,129] the variation of non-covalent interactions and surface areas observed suggest inherent degrees of stability in RBM oligomerization that may reflect

their role in assembly. The structure of InvG[26] and subsequent T2SS secretins[90,91] demonstrated that the N3 domain forms a ring oligomer with a consistently hydrophobic RBM interface that is also intimately associated with the underside of the inner β-barrel of the neighbouring *i* + 1 monomer (Fig. 2.4c, Fig. A5a,b), providing a stable foundation for secretin oligomerization. Here, in the open state, the N3 domain is rotated outwards slightly (Fig. 2.5a). The β-hairpin insertion specific to T3SS secretin orthologues (referred to here as β-INSERTION), observed to form a strut supporting the conformation of GATE1 and GATE2 in the closed state[26], is pulled away and twisted, disrupting the specific interactions of conserved $Arg_{198}$ at the β-INSERTION turn with $Asp_{384}$ and $Glu_{429}$ at the kinked regions of GATE1 and GATE2 respectively (Fig. 2.5b-e). The neighbouring conserved $Asp_{199}$ also shifts to interact with $Arg_{411}$ at the base of GATE1 in the open state (Fig. 2.5b-e). Mutation of $Asp_{199}$ completely abrogated secretion with $Arg_{198}$ having a milder phenotype in our previous assays[26] supporting the importance of these observed interactions in the two functional open/closed conformational states of InvG.

The quality of the needle complex local reconstruction in the OM secretin region has also permitted the observation of further structural features previously not resolved: First, we have been able to model additional sequences of the variable insertion (residues 217-267) between the first helix and second strand of the N3 domain. The now resolved 217-227 and 252-267 residues demonstrate an unexpected contribution to the structural stabilization of the secretin N domains (Fig. 2.5a, Fig. A5c, d). In the prior basal body reconstruction, we suggested this unresolved loop might account for the map features forming a periplasmic constriction evident at lower contour

levels[26]. However, with the needle filament occupying the secretin lumen in the assembled state here, the now clear atomic details of the loop show it threads out between the N3 and N1 domains and participates in both intra and inter domain interactions. The N-terminal region (residues 217-227) contributes to the N3 domain RBM interface, effectively chelating between the first strand of monomer *i* – 1 and the second helix of monomer *i* (Fig. A5c, d) while residues $Met_{252}$-$Gln_{260}$ form an α-helix which packs between the N1-N3 domain loops of monomer *i* and *i* - 1 and interacts with the N1 domain of monomer *i* (Fig. A5c,d). Collectively these additional interfaces serve to further anchor and stabilize the needle bound form of the injectisome at the point of greatest structural change. Second, we have been able to position the peripheral N0 and N1 domains, which bridge the periplasmic space to the IM rings with the N-terminal residues of the N0 domain interacting with the C-terminus of PrgH[17,26,75] (Fig. 2.1a, Fig. A5e; see above). Interestingly, while the structure validates our previous Rosetta density guided symmetry models of the $InvG_{34-172}$ crystal structure[25,26], the data here confirms the direction of the N1 domain RBM interface is indeed opposite to the N3 domain (Fig. A5a). This is in contrast to the recent cryo-EM structures of several T2SS secretins[90,91] where the N1 and N2 (a distinct RBM motif not present in T3SS secretins) domains are shown to have the same orientation with the relative inter N domain packing more staggered compared to InvG (Fig. A5a,b). These rather unexpected distinctions again clearly underlie the plasticity of the RBM in modulating periplasmic span and inner membrane coupling in these functionally diverse secretion systems.

### 2.3.4 A mechanism for substrate mediated secretin gate opening

We previously envisaged a substrate induced gate opening mechanism for the OM secretin during needle assembly based on comparison of the basal body structure to previous low resolution (10-20 Å) studies of the needle complex[17,105]. The near-atomic resolution structure of the needle complex here provides details of the discrete structural changes occurring at the molecular level and permits further elaboration. We hypothesize a gate opening mechanism broadly divided into two stages: an initial allosteric transition to unlock the gate and a secondary steric phase to push the gate open. Rod/needle polymerization within the secretin lumen would bring the filament into contact with the N3 domain, including the regions of the N3 domain loop that account for the periplasmic constriction visible in the isolated cryo-EM basal body[26] and in situ cryo-ET reconstruction[18], triggering a reorientation/ordering to adopt the conformation we observe here and contributing lateral stability to withstand the continued force of the assembling needle (Fig. 2.6a-b). The accompanying change in N3 domain orientation we observe alters the interface with the inner β-barrel, which we propose triggers an allosteric transition that effectively unlocks the gate (Fig. 2.6b). Central to this appears to be the disruption of interactions of the β-INSERTION with GATE1 and GATE2: $Arg_{198}$ at the β-INSERTION turn is at the centre of a network of polar and hydrophobic interactions that serve to stabilize the closed gate conformation of GATE1 and GATE2 (Fig. 2.5b-e). The reorientation of β-INSERTION and disruption of this interface could enable the downstream conformational changes associated with gate opening. Supporting an important functional role, mutation of $Arg_{198}$ and neighbouring $Asp_{199}$ impacted secretion in our previous assays[26]. GATE2, although not contributing directly

to the physical barrier per se (the role of GATE1), appears to play rather a critical buttressing role for GATE1, undergoing a significant displacement between closed and open states (Fig. 2.5b-e). In the closed state, residues at the N-terminus of strand 7 in GATE2 pack between kinked GATE1 β-hairpins of monomers *i* and *i* – 1 holding them in a closed position (Fig. 2.5e), whereas in the open state this same region is rotated 180° and packed against the outer β-barrel and the extended GATE1 of monomer *i* (Fig. 2.5d). A specific interaction between GATE2 $Glu_{449}$ and $Arg_{387}$ of GATE1 of monomer *i* – 1 is maintained between the two states suggesting the movement of these two hairpins occurs in concert (Fig. 2.5d, e), with the rotation of GATE2 potentially initiating the lifting of the GATE1 and/or providing structural support to guide the extended β-strands during opening. Mutants found to increase the permeability of the filamentous phage secretin pIV map to GATE2 in addition to GATE1 (Fig. A6) highlighting the influence its conformation has on the extent of gate opening. Continued needle polymerization within the secretin lumen could provide additional energy to sterically push the gate open with the resulting reorientation of the entire inner β-barrel we observe (Fig. 2.6c-d). The resulting complementary fit, notably at the T3SS specific adaptations of the N3 domain β-INSERTION and three-stranded collapsed β-barrel outer lip, and complementary electrostatics of the N3 and β-INSERTION (Fig. A7) would contribute to the stable anchoring of the needle through the OM, essential for the assembled needle complex that must endure considerable extracellular forces in the infected host.

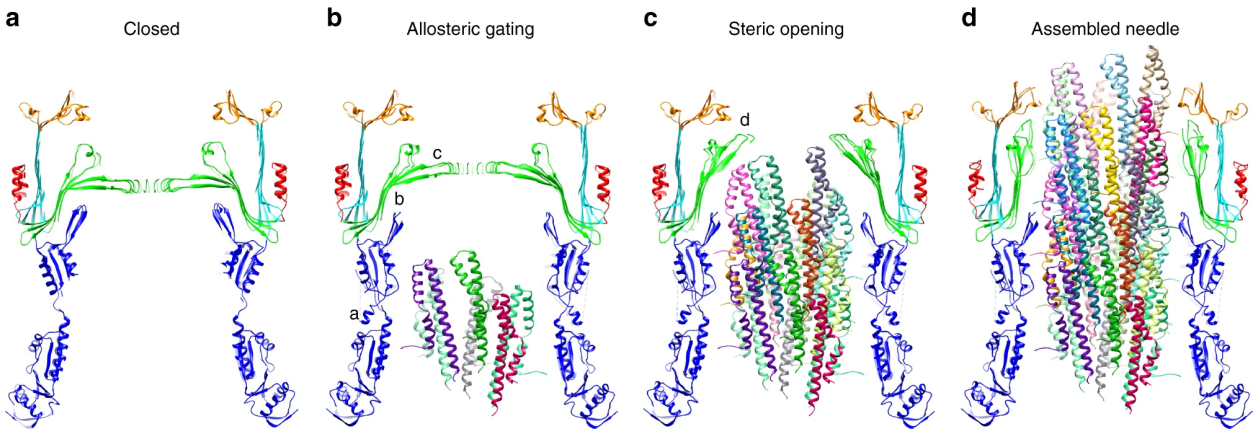



**Figure 2.6: Proposed secretin gating mechanism**

**(a) Closed secretin. (b) Initial rod/needle polymerization within the secretin lumen contacts the N3 domain triggering the ordering of the variable N3 loop[a] providing lateral stability and altering the N3 domain-inner β-barrel interface[b]. This in turn disrupts the interactions of the N3 β-INSERTION with the GATE1 and GATE2 kinks and unlocks the gate[c]. (c) Continued needle polymerization sterically pushes the gate open[d]. (d) Final assembled needle with fully open gate.**

## 2.4 Discussion

In this work, we have solved the cryo-EM structures of the T3SS needle complex and isolated needle filament at resolutions permitting atomic modelling. These structures provide essential understanding of injectisome structure and importantly provide the view and molecular understanding of the pivotal elements involved in substrate specific gating across one of the largest known membrane spanning portals. This work and the resolution it afford in turn paves the way for potential structure-based design of antimicrobials (compounds that block gate

opening or needle polymerization for example) against the many important clinical and community pathogens that rely on a T3SS for their virulent effects

# Chapter 3: Cryo-EM snapshots of T3S injectisome needle complex membrane coupling and assembly

## 3.1 Introduction

Assembly of the *Salmonella* SPI-1 injectisome proceeds via two steps: First, eight *sec*-dependent membrane proteins - SpaPQRS, PrgHK, InvA, InvG (unified nomenclature SctRSTU, SctDJ, SctV, SctC) - assemble to form a complex spanning the bacterial IM, periplasmic space, and OM[17,26,72,79]. Second, association of cytoplasmic proteins including the ATPase complex[71] results in T3S of the internalized hollow needle (PrgIJ/SctFI), and terminating tip/translocon (SipBCD). When isolated from native membranes, weakly associated components including InvA and SpaS (verified by mass spectrometry[74]) typically dissociate. The resulting NC-base (*sec*-dependent components) and NC (with addition of T3S dependent needle) (Fig. 3.1a), have been crucial *in vitro* tools towards understanding function of this complex nanomachine; comparison of their isolated structures to the corresponding *in situ* cryo-ET reconstructions demonstrates near native arrangement[17,18,26].

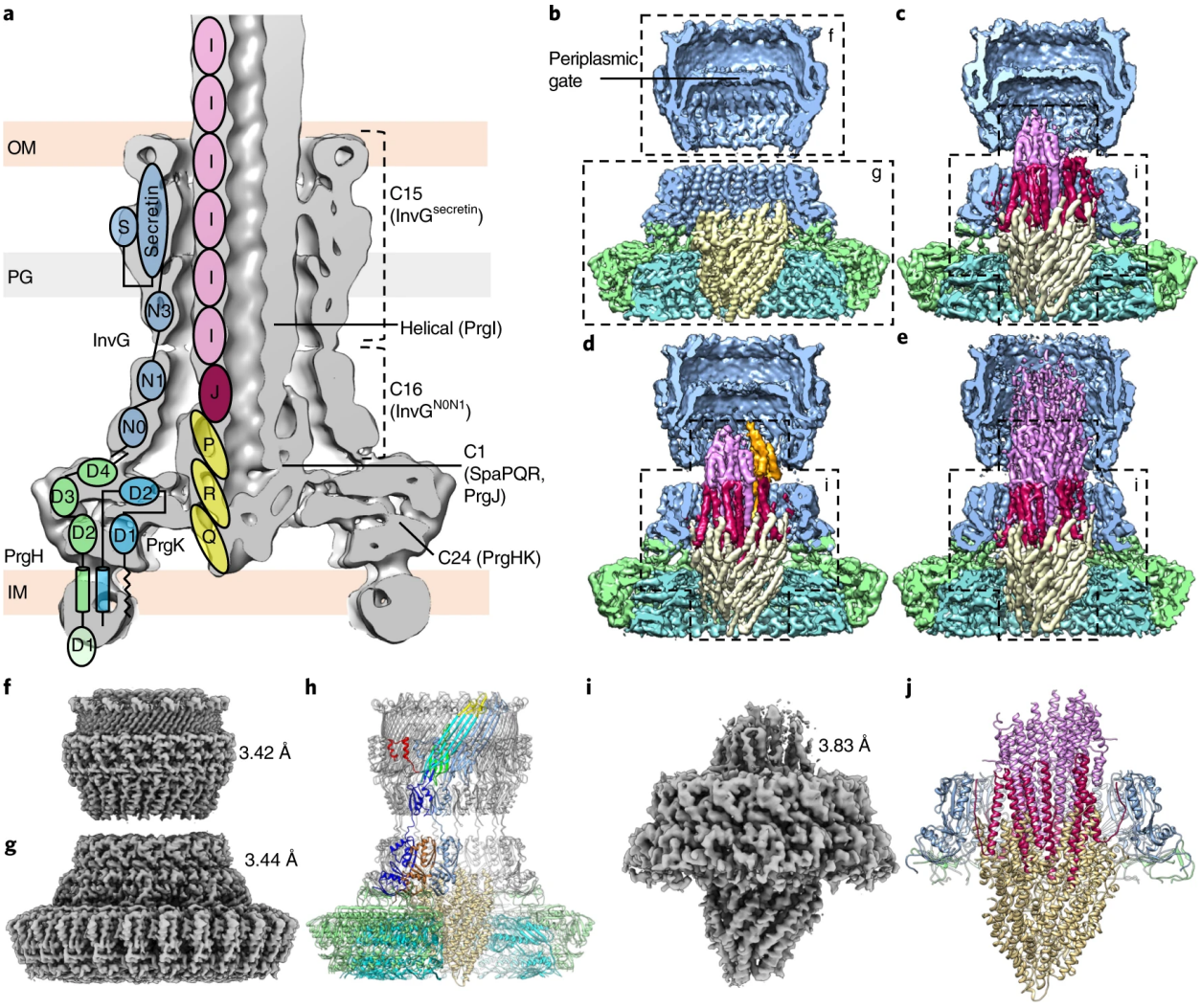



**Fig. 3.1: Cryo-EM snapshots capture the molecular basis of needle complex assembly.**

**a, Overview of the T3S NC. Low-pass filtered NC reconstruction15 showing general domain architectures (left) and varying symmetries of the assembled components (right). PrgH, green; PrgK, cyan; SpaPQR, yellow; PrgJ, cherry; PrgI, pink; InvG, blue. IM, inner membrane; PG, peptidoglycan; OM, outer membrane. b–e, Reconstructions of the entire NC (C1) showing different states of needle assembly from the NC base with no needle and the secretin periplasmic gate closed (b; 3.8-Å resolution, coloured as in a) and partially unlocked (c; 6.0-Å resolution), and unlocked (d; 4.75-Å resolution, additional PrgI subunits in**

**orange) states with the secretin periplasmic gate partially opened and initial PrgJ and PrgI subunits trapped within, and a gate open state (e; 5.15-Å resolution) with the needle assembled throughout and the fully open gate. The boxed regions correspond to the focused refinements shown in f,g and i, as labelled. f, Focused refinement of the InvGsecretin pore (3.42-Å resolution, 15-fold symmetry averaged). g, Focused refinement of the IM complex composed of PrgHK, SpaPQR and InvGN0N1 (3.44-Å resolution, no symmetry applied). h, Atomic structure of the NC base. To highlight the stoichiometry and domain arrangement in the InvG secretin, one monomer is coloured according to structural features with the N-terminal, modular N0, N1, N3 domains coloured in cobalt blue, the secretin outer β-barrel in cyan, the inner β-barrel in green, the lip β-barrel in yellow and the S domain in red. The adjacent subunit is coloured light blue and the remainder are coloured grey. The additional InvGN0N1 domain is coloured orange. PrgHK and InvGN0N1 are slabbed to better show the internal SpaPQR complex. i, Focused refinement of the core SpaPQR complex with attached PrgIJ needle and surrounding InvGN0N1 ring at a resolution of 3.83 Å. j, Atomic structure of the SpaPQR complex with attached PrgIJ needle and surrounding InvGN0N1 ring.**

Cryo-EM studies of the NC-base and NC have elucidated atomic structures of the IM rings[17,26,80], and OM secretin in closed[26] and open states[80]. However, varied symmetries throughout (Fig. 3.1a) have confounded further analyses, with many features of the assembled complex remaining unresolved, notably the export apparatus, attached PrgIJ needle, and specific intermolecular interactions that mediate IM/OM coupling. Here we have captured four structural snapshots resolving the entire NC in different stages of assembly, remarkably revealing atomic details of each component of the NC including those that elucidate function of the export apparatus and PrgJ in initiating needle assembly and answering the enigmatic question of how the IM and OM rings couple given the apparent incompatible symmetries.

## 3.2 Methods

### 3.2.1 Sample preparation

A non-flagellated *ΔprgH Salmonella* strain[135] was used, complemented with a pWSK29 plasmid harboring the *prg* operon modified to only express a thrombin-cleavable N-terminal 10XHis–tagged PrgH with the *lac* promotor deleted so expression is driven by the native *prgH* promotor with no induction[135]. The *prgH* complemented *Salmonella* strain was transformed with a pBAD plasmid expressing the T3SS transcription activator *hilA*[136]. Bacteria were grown with mild aeration in LB broth supplemented with 0.3 M NaCl to an $A_{600}$ of approximately 0.5, arabinose was added to a final concentration of 0.015% and grown overnight at 30 °C. Cells were resuspended in 0.1 M Tris (pH 8.0), 0.75 M sucrose in the presence of lysozyme (0.4 mg/ml) and EDTA (1 mM), and incubated with stirring. Cells were lysed with 1% lauryldimethylamine oxide (LDAO) in the presence of protease inhibitor tablets (Roche) before adding 10 mM $MgCl_2$ and 500 mM NaCl to the lysate. Cell debris was removed by low-speed centrifugation and needle complexes were pelleted by high-speed centrifugation (Beckman, 45Ti rotor, 40,000 r.p.m., 45 min, 4 °C). The pellet was resuspended in 0.5% LDAO in 10 mM Tris-HCl pH 8, 0.5 M NaCl, 5 mM EDTA and adjusted to a final concentration of 30% (w/v) of CsCl. Following this, 5-ml samples were centrifuged for 16–20 h at 25,000 r.p.m. in a Beckman SW41 rotor. Fractions containing needle complex were combined and desalted into 10 mM Tris (pH 8.0), 0.5 M NaCl, 0.2% LDAO before purification using nickel-charged Chelating Sepharose Fast Flow (GE Healthcare). The sample was further concentrated before gel filtration on a

Superose 6 column (GE Healthcare) equilibrated in 10 mM Tris (pH 8.0), 0.5 M NaCl, 0.2% LDAO.

### 3.2.2 Cryogenic electron microscopy

2.5 µl of the purified needle complex was applied to glow-discharged (30 s on each side) Quantifoil grids (Gold, 400 mesh, R1.2/1.3). The grids were blotted for 3 s at 100% humidity and plunge-frozen into liquid ethane using a Vitrobot Mark IV (FEI Company). Grids were imaged on a 300 keV Titan Krios cryogenic electron microscope (Thermo Fisher Scientific) equipped with a spherical aberration corrector, an energy filter (Gatan GIF Quantum) and a post-GIF Gatan K2 Summit direct electron detector. Images were taken on the K2 camera in dose-fractionation and super-resolution mode at a calibrated magnification of 29,240, corresponding to 0.855 Å per super-resolution pixel (1.71 Å per physical pixel). The dose rate on the specimen was set to be 3.42 electrons per $Å^2$ per s and total exposure time was 15 s, resulting in a total dose of 51.3 electrons per $Å^2$. With dose fractionation set at 0.3 s per frame, each movie series contained 50 frames and each frame received a dose of 1.03 electrons per $Å^2$. An energy slit with a width of 20 eV was used during data collection. Fully automated data collection was carried out using SerialEM with a nominal defocus range set from − 1.5 to − 3 µm.

### 3.2.3 Image processing

A total of 2575 movie series were collected at super resolution at 0.855 Å per pixel and binned to 1.71 Å per pixel for further processing. Motion correction was done using MotionCor2[111]. The contrast transfer functions (CTFs) of the summed micrographs were determined using

CTFFIND4[112]. Approximately 2,000 particles were manually boxed out from selected micrographs to generate an initial model using reference-free 2D-class averages by Relion 3.0[137]. The representative 2D-class averages were also used as templates for automated particle picking for the entire dataset in Relion 3.0. With 459661 particles automatically picked, reference-free 2D classification was performed in Relion and 235618 particles were kept in the good class averages. Two good initial maps were generated in Relion, with most of particles (~85 %) contributing to a model with no needle (NC-base), and rest (~15 %) contributing to a model with an internal needle substructure; however, 3D classification with either model was not able to obtain a class with the internal needle. Thus, all the particles were combined and intended to be separated at later stage of processing. One round of 3D classification and refinement was done to generate a map for Bayesian polishing and CTF refinement[137]. Another round of refinement was done on the "shiny" particles to generate a new map that was subsequently used for partial signal subtraction. The density of $InvG^{secretin}$ and PrgHK ring was subtracted, and the subtracted particles were subjected to several rounds of *ab initio*, heterogenous and homogeneous refinement in cryoSPARC[53]. The resulting two reconstructions reproduced the features of the initial models and refined to ~3.5 Å ("no needle") and ~3.8 Å ("with needle"). The angular information determined in cryoSPARC was used in Relion on the unsubtracted particles to generate two C1 maps that contained all the components of the NC without or with the internal needle. The "no needle" NC-base reconstruction showed lower local resolution at the $InvG^{secretin}$ pore; thus we further classified the particles focusing at the relative position between the $InvG^{secretin}$ ring and the export apparatus. Two classes with much improved local density of the $InvG^{secretin}$ ring were obtained and showed the alternate position of extra $InvG^{N0N1}$. To

achieve the best local resolution, partial signal subtraction and focused refinement was done for both InvG$^{secretin}$ with C15 symmetry imposed and rest of the complex with no symmetry imposed and obtained reconstructions at 3.42 Å and 3.44 Å resolution respectively. The unsubtracted particle dataset contributing to the "with needle" model was treated in a similar fashion although the signal was not sufficient to distinguish different classes corresponding to the alternate InvG$^{N0N1}$ location. Thus, a symmetry imposed (C15) 3D classification of the InvG$^{secretin}$ was carried out, and three distinctive classes were obtained with different degrees of periplasmic gate opening. Particles associated with these classes were selected and refined in Relion using the angles determined previously from the subtracted C1 SpaPQR, InvG$^{N0N1}$, PrgIJ complex. Three reconstructions were obtained (4.75 – 6 Å resolution) showing different lengths of the internal needle and degree of secretin gate opening. Resolution was determined at a criteria of 0.143 FSC[48] gold-standard refinement procedure and high-resolution noise substitution to correct soft-mask effects.

### 3.2.4 Model building and analysis

For the focused refinement reconstruction of the InvG$^{secretin}$ pore in the NC-base at 3.42 Å (C15 symmetry applied), the previous structure (PDB 5TCQ) was docked into the map using Chimera[115] and refined using both Phenix real space refine[116] and density-guided iterative local refinement as implemented in Rosetta[117], taking into account the 15x symmetry. Further model building was carried out in COOT[120]. The final model has a good fit to the map (EMRinger[119] score 2.6) and good statistics (Molprobity[118] score 1.50, Ramachandran favored 94.21 % & outliers 0 %) (Table B1). For the focused refinement reconstruction of the IM

complex in the NC-base (corresponding to PrgHK, SpaPQR, and InvG$^{N0N1}$) at 3.44 Å resolution (no symmetry applied), the previous structures of the 24-mer PrgHK rings (PDB 5TCP) and 16 copies of the InvG$^{N0N1}$ domain (PDB 4G08) were docked into the map using Chimera. *De novo* model building of the missing regions of PrgH, InvG, and the entire SpaPQR complex was carried out in COOT. The resulting model was refined in Phenix real space refine. The final model has a good fit to the map (EMRinger score 2.0) and good statistics (Molprobity score 1.28, Ramachandran favored 96.40 % & outliers 0.02 %) (Table B1). For the focused refinement reconstruction of the SpaPQR complex with attached PrgIJ needle and surrounding InvG$^{N0N1}$ ring at 3.83 Å resolution (no symmetry applied), the equivalent coordinates from the NC-base structure were docked into the map using Chimera[115]. Rebuilding of the structurally different and missing regions of SpaPQR, and de novo model building of PrgJ were carried out in COOT. PrgI monomers were subsequently docked in Chimera and rebuilt in COOT where necessary. The resulting model was refined in Phenix real space refine. The final model has a good fit to the map (EMRinger score 1.0) and good statistics (Molprobity score 1.72, Ramachandran favored 95.72 % & outliers 0.08 %) (Table B1).

Structural analysis and figure generation were carried out in PyMOL[138] and Chimera. Remote homology detection was carried out with HHPred[139].

### 3.2.5 Generation of *prgH and prgJ* mutants in *Salmonella* SL1344

We used the *sacB* gene-based allelic exchange method and the suicide vector pRE112[140] to generate in-frame, non-polar deletion mutants of *prgH* and *prgJ* in *Salmonella enterica* serovar

Typhimurium strain SL1344 (streptomycin-resistant) and a non-flagellated derivative SL1344 *flhD::kan* (streptomycin- and kanamycin- resistant).

Deletion mutants of *prgH* (Δ*prgH*): Two DNA fragments flanking upstream (~0.8 kb) and downstream (~1.0 kb) of the coding region of *prgH* were generated by PCR using primer pairs PrgH-1 (*Kpn*I underlined) (5'-GCGGTACCGCCATTTTGGTTTGCTGC-3') and DprgH-R (*Nhe*I underlined) (5'-GCGCTAGCGCCTGTCAGCAATGGAAACTC-3'), or DprgH-F (*Nhe*I underlined) (5'-GCGCTAGCAAAATGAGCCCAGGCCATTGG-3') and PrgH-2 (*Sac*I underlined) (5'-GCGAGCTCGAGAGCACGCCCTCCATCGTC-3'). The two PCR products were first cloned into pCR2.1-TOPO (Invitrogen) and verified by DNA sequencing. After digested with *Kpn*I/*Nhe*I and *Nhe*I/*Sac*I, respectively, the two DNA fragments were gel-purified and cloned into *Kpn*I/*Sac*I-digested pRE112 in a 3-way ligation, generating pRE112-ΔprgH which contains 0.8 kb upstream and 1.0 kb downstream flanking sequences of *prgH* and the *prgH* gene with an internal in-frame deletion from nucleotides 103 to 1137 (about 88% of the coding region). The coding regions of the N-terminal 34 and the C-terminal 13 amino acid residues of *prgH* are retained to avoid any polar effects on the expression of the downstream open reading frames. An *Nhe*I site was introduced into the deletion site. The suicide vector pRE112-ΔprgH was transformed into *E. coli* strain MFD*pir*[141] by electroporation, and introduced into *Salmonella* strain SL1344 or its non-motile derivative SL1344 *flhD::kan* by conjugation. After sucrose selection, *Salmonella* colonies resistant to sucrose and streptomycin and sensitive to chloramphenicol were screened for deletion of *prgH* by PCR. The *prgH* mutants were further verified by multiple PCRs. The *prgH* deletion mutant thus created showed abolished

SPI-1 T3SS mediated protein secretion, and this secretion defect can be fully complemented by expressing PrgH from a plasmid.

Deletion mutants of *prgJ* (Δ*prgJ*): Primer pairs PrgIJ-1 (*Kpn*I underlined) (5'-GCGGTACCCCAATACAGGTCGGTGAATTGC-3') and DprgJ-R (*Nhe*I underlined) (5'-GCGCTAGCAGGGACAATAGTTGCAATCGAC-3'), and DprgJ-F (*Nhe*I underlined) (5'-GCGCTAGCAAAGGAGTCGGGGCTGTTGAAAC-3') and PrgIJ-2 (*Sac*I underlined) (5'-GCGAGCTCCTCATTTGCCGCCGCCCTGGC-3'), were used to construct pRE112-ΔprgJ, which contains 1.2 kb upstream and 0.9 kb downstream flanking sequences of *prgJ* and the *prgJ* gene with an internal in-frame deletion from nucleotides 25 to 267 (about 88% of the coding region), while keeping the coding regions of the N-terminal 8 and the C-terminal 12 amino acid residues of *prgJ*. The pRE112-ΔprgJ plasmid was then used to generate Δ*prgJ* mutants in *Salmonella* strains.

### 3.2.6 SPI-1 secretion assays

Secretion assays were carried out using wild-type or Δ*prgH* or Δ*prgJ* deletion strains (see above) complemented with modified pWSK29 plasmids expressing *prgH* or *prgJ* respectively from the native *prg* operon promotor[135] with no induction. *Salmonella* strains were grown overnight in LB broth (supplemented with antibiotics if required) at 37 °C with shaking at 225 r.p.m. The cultures were then diluted 1:100 into 4 ml of fresh LB, with appropriate antibiotics, and grown under the same conditions for 6 h to induce SPI-1 type III secretion without induction. The cultures were then centrifuged at 16,100g for 10 min to pellet the bacteria and the

bacterial pellet was re -suspended in SDS–PAGE sample buffer to generate whole cell lysates. The growth medium supernatant was collected and passed through a Millex-GV 0.22 μm filter unit (Millipore) to remove any remaining bacteria and the secreted proteins were precipitated with a final concentration of 10% (v/v) trichloroacetic acid. The secreted proteins were then collected by centrifugation at 16,100g for 30 min, the protein pellet was dried in air and dissolved in SDS–PAGE sample buffer, with residual trichloroacetic acid neutralized using 0.5 μl of saturated Tris. The amount of the sample buffer used to re-suspend the bacterial pellet or dissolve the precipitated proteins was normalized according to the $A_{600}$ values of the cultures to ensure equal loading of the samples. The secreted proteins were run on 12% SDS–PAGE gels and stained with Coomassie blue G250.

## 3.3 Results

### 3.3.1 Cryo-EM snapshots capture the molecular basis of needle complex assembly

Using polyhistidine-tagged PrgH-expressing plasmids that restored secretion in a Δ*prgH* background (Fig. B1b), we observed time dependent distribution of NC and NC-base-like complexes (Fig. B1a). Overnight cultures resulted in purified samples enriched with particles of varying needle lengths, we propose a result of both needle shearing/depolymerization and increased levels of immature NC bases due to insufficient membrane and cytoplasmic components. To further investigate these assembly states, single particle cryo-EM analysis was carried out (Table B1). Partial signal subtraction and focused refinement[67] of SpaPQR and surrounding periplasmic InvG N0 and N1 domains ($InvG^{N0N1}$) followed by 3D classification and refinement of unsubtracted particles using determined orientations (Fig. B1, B2) resulted in

reconstructions resolving the entire NC in four distinct assembly snapshots (Fig. 3.1b-e, Fig. B1). The predominant state is the NC-base (3.80 Å resolution), lacking needle and the secretin gate closed (Fig. 3.1b). Due to small variations in relative domain orientation (centered at ~f in Fig 3.2a), the region corresponding to the InvG secretin domain ($InvG^{secretin}$) has lower local resolution in context of the whole particle maps; focused refinements however result in two reconstructions at ~3.4 Å that collectively account for the entire assembly ($InvG^{secretin}$ and $InvG^{N0N1}$+SpaPQR+PrgHK) (Fig. 3.1f-g, Fig. B1f, B2). Three less populated NC states capture an internal PrgIJ needle of increasing length with concomitant degrees of secretin gate opening: two intermediate structures with a short needle and partially opened gate (6.00 Å and 4.75 Å resolution), and an open state with the needle occupying the NC lumen and the gate fully opened (5.15 Å resolution) (Fig. 3.1c-e, Fig. B1e). Focused refinement of the SpaPQR complex with attached PrgIJ needle and surrounding $InvG^{N0N1}$ ring common to these states produced a reconstruction to 3.83 Å resolution (Fig. 3.1i, Fig. B1f, B2) showing the molecular details of PrgJ mediated initiation, interaction, and anchoring of the PrgI needle (Fig. 3.1j). Remarkably, 16 copies of the terminal $InvG^{N0N1}$ domain are observed in all states, in contrast to the firmly established 15 copies of the OM spanning $InvG^{secretin}$ β-barrel[16,17,26] (Fig. 3.1h).

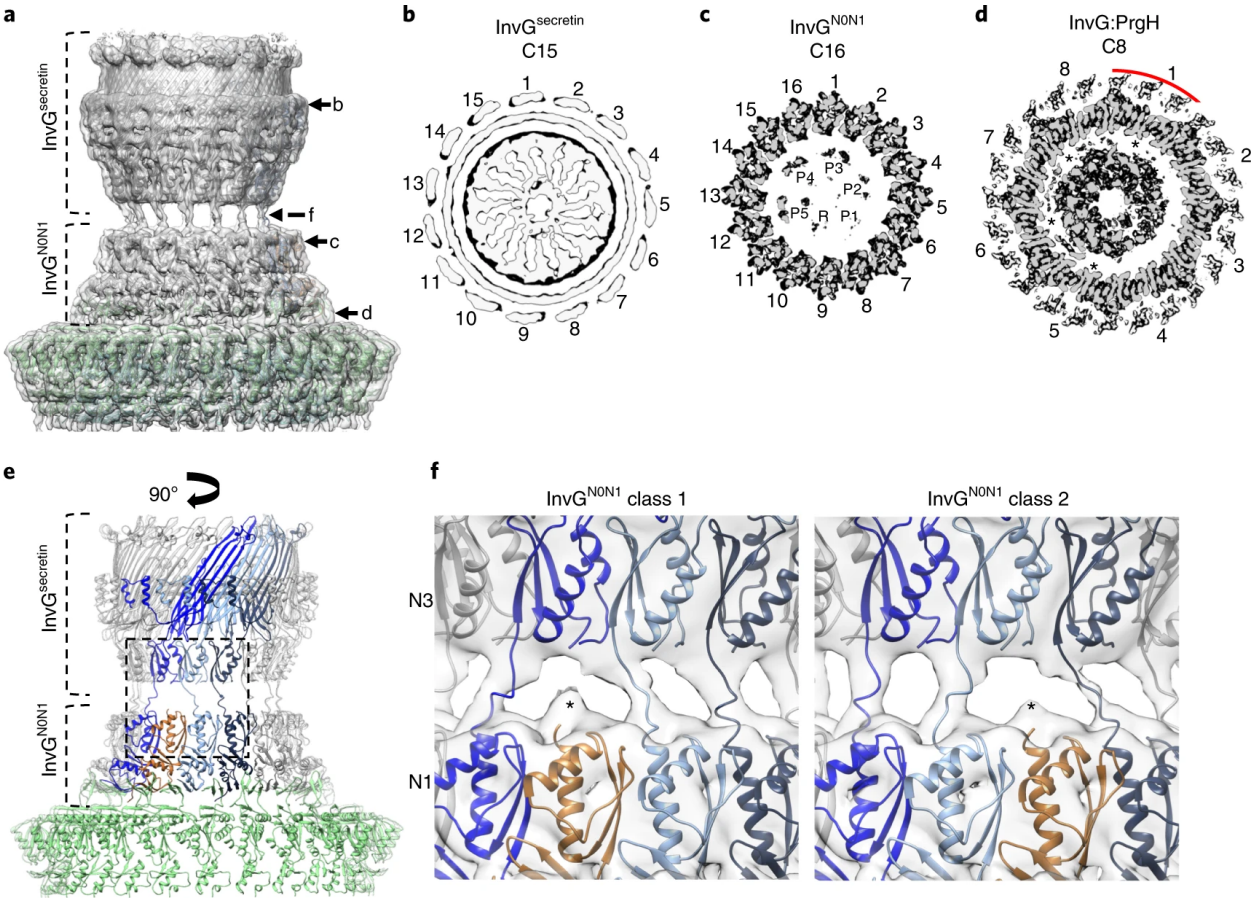



**Fig. 3.2: Recruitment of an extra secretin monomer in the periplasm locally alters ring stoichiometry and facilitates multivalent symmetrical coupling of the OM and IM T3S components.**

**a, C1 reconstruction of NC base showing displaced and tilted InvG$^{secretin}$ relative to InvG$^{N0N1}$ and the IM complex. Slab planes in b–d and view perspective in f are indicated. b–d, Slab views at planes marked in a showing 15-fold symmetry of the InvG$^{secretin}$ (b), 16-fold symmetry of the InvG$^{N0N1}$ domains (c) and 8-fold symmetry of the PrgH$_{24}$:InvG$^{N0N1}_{16}$ interface (d) facilitated by the extra InvG$^{N0N1}$. The repeating PrgH$_3$:InvG$^{N0N1}_2$ asymmetric unit is marked in red. The level-dependent ordering of the SpaP$^{insertion}$ loop at the interface with the InvG N0 domains is also evident (marked with an asterisk) with SpaP$^5$ (P5 in c) the most well resolved. e, Atomic model showing InvG (blue, orange and grey) and PrgH (green). Three neighbouring InvG monomers are coloured blue, light blue and dark blue, with the extra InvG$^{N0N1}$ coloured orange. f,**

**Zoom of boxed region in e showing two NC base classes with alternative positions for the additional InvG$^{N0N1}$ (orange, marked with an asterisk). The loops connecting the InvG$^{N0N1}$ and InvG$^{secretin}$ domains are best observed in the lower resolution (non-postprocessed) reconstructions. Coloured as in e. Residues beyond InvG$^{N0N1}$ in the additional monomer are not observed.**

### 3.3.2 Local OM secretin stoichiometry variation facilitates interaction with the IM components

A significant unanswered puzzle of injectisome assembly is how the IM and OM components couple to create a stable scaffold. The conundrum arises from a symmetry mismatch, with the OM secretin having an assumed 15-fold symmetry throughout[17], supported by high resolution structures of the InvG$^{secretin}$ pore[26,80], and the IM components established as 24-mer nested rings. Our structures resolve this, revealing the periplasmic most InvG$^{N0N1}$ ring locally incorporates an additional InvG monomer, which is not continuous to the strictly 15-mer OM spanning pore; residues beyond InvG$^{N0N1}$ in the additional monomer are not observed (Fig. 3.2). The result is a central focal point of IM/OM component coupling involving the 24-mer PrgH ring and 16 copies of InvG$^{N0N1}$ with a repeating interface of 3:2 stoichiometry (Fig. 3.2c-e, Fig. 3.3b-c). This unprecedented stoichiometry variation within the secretin is an unambiguous feature in the dataset; we found no evidence of different compositions. We extended analyses to our previous mutant NC-base data[26] and a native NC sample with intact needles with analogous results (Fig. B4a-b). Further, this region superposes near perfectly with the *in situ* cryo-ET structure of the SPI-1 injectisome[18] (Fig. B4c). Thus, the collective structural evidence

supports that the mixed secretin stoichiometry we observe in three distinct samples is a requisite feature of the functional *Salmonella* SPI-1 injectisome, demonstrating a remarkable evolutionary adaptation to couple the IM rings and OM secretin via a multivalent, reinforced interface.

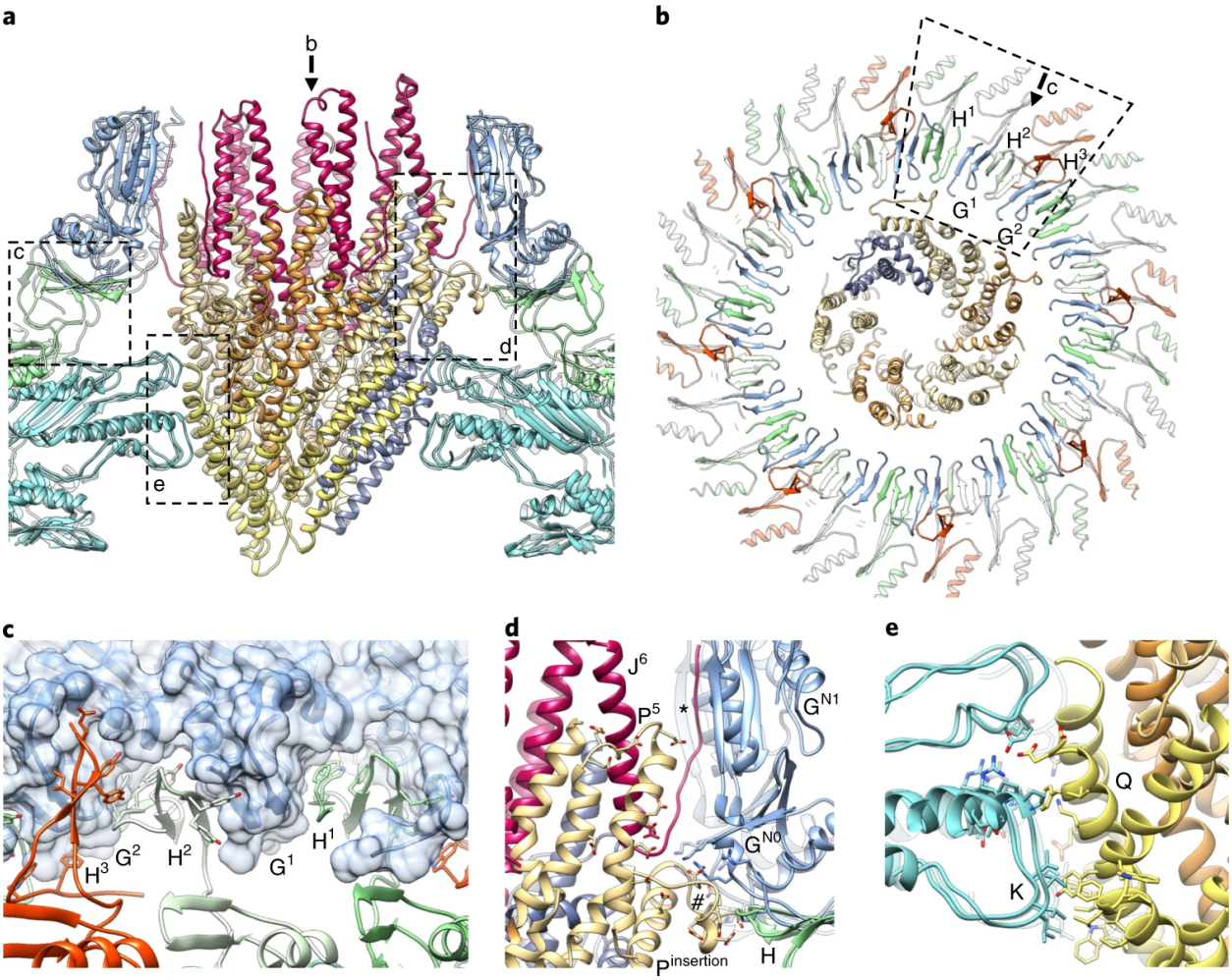



**Fig. 3.3: Molecular interactions that govern NC assembly.**

**a, Central view of the IM complex showing multiple intercomponent interactions that define the NC. PrgH, green; PrgK, cyan; InvG, light blue; SpaP$^{1–5}$ alternating light and dark orange; SpaR, dark blue; SpaQ, yellow; PrgJ, cherry; PrgI is omitted from the figure for clarity. Labels correspond to regions shown in b–**

**e. b, Top-down view (from OM perspective) of the PrgH–InvG coupling interface. InvG coloured light blue, two PrgH monomers per repeating asymmetric unit (boxed) that interchelate with InvG to form the circular, 80-stranded β-sheet, coloured green and light green, respectively, and the third, which forms a similar secondary structure, but packs around the exterior and underside of the interface, is coloured red. c, Side view of the $PrgH_3$:$InvG^{N0N1}{}_2$ asymmetric unit boxed in b. d, The collective interface between SpaP, PrgJ, InvG and PrgH. In the presence of PrgJ, $SpaP^{insertion}$ is more ordered and, together, they form a cohesive interface with InvG mediated by the interaction of the ordered $SpaP^{insertion}$ loop with the InvG N0 domain (indicated by a hashtag), and the PrgJ N-terminal loop at the InvG N1 domain oligomerization interface via a β-strand complementation-like interaction (indicated by an asterisk). Enrichment of negatively and positively charged residues (shown as sticks), respectively, creates complimentary electrostatic surfaces. e, The interface between PrgK and the SpaPQR complex. PrgK creates a membrane-like environment that houses SpaPQR with the two centrally projecting loops providing both hydrophobic and polar interactions with the corresponding exterior surface of SpaPQR.**

Clearly defined in these new reconstructions, the $PrgH_{24}$:$InvG_{16}$ interface has local C8 symmetry (Fig. 3.2d, 3.3b-c). The $PrgH_3$:$InvG^{N0N1}{}_2$ repeating interface is mediated by the PrgH C-terminus (unresolved in all previous cryo-EM and X-ray models[23–26]) and the InvG N0 domain (also less resolved in prior EM reconstructions that assumed 15-fold symmetry[17,26,80]). Sixteen PrgH C-termini bind between the 16 $InvG^{N0N1}$ domains, creating a circular, 80 stranded β-sheet, while the remaining eight pack around and under the β-sheet to further reinforce the interface (Fig. 3.3b-c), explaining cross-links from several studies (Fig. B5a)[16,142]. Collectively, these interactions form an extensive, hydrophobic interface (area = 3141.3 $Å^2$, $\Delta^iG$ = -38.7 kcal/mol per repeat)[143] that underrides the stability of the assembled NC. Accordingly, deletion of the

last six residues of PrgH mapping to the terminal β-sheet (Fig. B5a-b) was found to abolish secretion and severely impact NC assembly[142]. Mutation of this β-strand here shows conserved aromatic residues $Trp^{386}$ and $Phe^{388}$ at the interface are critical determinants of IM/OM coupling (Fig. B5a-b).

The extra $InvG^{N0N1}$ domain is observed in two neighboring positions. In the NC-base structure, the $InvG^{secretin}$ pore is laterally displaced and oriented at a tilt with respect to the IM complex (Fig. 3.2a), a result of the stretching of loops connecting the $InvG^{N0N1}$ and $InvG^{secretin}$ domains to accommodate the extra $InvG^{N0N1}$ subunit in the peripheral ring (Fig. 3.2e-f). Two classes were observed in the NC-base dataset differing by relative orientation of the displaced $InvG^{secretin}$ pore with respect to the SpaPQR complex (Fig. 3.2f, Fig. B2). This suggests that incorporation of the extra $InvG^{N0N1}$ is not confined to a specific position but happens stochastically within a restricted subset of locations; the implications for assembly are discussed below.

### 3.3.3 Export apparatus SpaPQR is the structural focal point of NC assembly

The $SpaP_5Q_4R_1$ complex accounts, in entirety, for central density at the heart of the NC, abutting the IM in contrast to previous predictions as multi-span membrane proteins[72,73]. It is a remarkable focal point of NC assembly with myriad interactions to NC-base and needle components as observed for the first time here, and explaining drastic decrease in NC-base assembly efficiency in the absence of either SpaP, SpaQ or SpaR[72]. The stoichiometry and architecture – a right handed pseudo-helical assembly closed at the base, with a central atrium 15-20 Å wide – is well conserved with the 4.2 Å resolution structure of the recombinant, isolated

flagellar FliPQR complex[21] (~20-30 % sequence identity; Cα RMSD ~1.2 – 3.2 Å). The pseudo-helical polymer is built up of a repeating helix-hairpin-helix motif present once in SpaQ, twice in SpaP, and three times in SpaR (Fig. B6d-e), which share a kinked structure that narrows the assembly toward the IM. The IM facing hairpin loops create a constriction that occludes the channel, previously termed in the isolated FliPQR structure[21] as R-plug and P/R-gasket (Fig. B6b). Consistent with the complex representing the entry portal for secreted proteins into the needle channel, mutation of these in flagellar orthologues impacts membrane permeability[144] or rescues secretion deficient mutants[145]. Substrate passage would clearly necessitate opening of SpaPQR to the IM; an iris-like opening model of the FliPQR structure involving straightening of the kinked helices was proposed[21] and, using a prior low-resolution (~10 Å) cryo-EM NC reconstruction[17] as a reference, it was suggested that the presence of the attached PrgIJ needle would confer the open state. However, we do not observe this here, with SpaPQR in a closed conformation in both absence and presence of PrgIJ and the major sites of conformational change restricted to the periplasmic-most helical repeats of SpaPR that bind PrgJ (Fig. B6c) (see below). This observation suggests a new hypothesis that interaction of the export apparatus with the IM embedded export gate InvA and/or SpaS, both absent from our structures, will be largely responsible for conformational control of gating/opening to regulate access to the needle lumen during secretion.

SpaPQR presents hydrophobic and hydrophilic surfaces toward the IM and periplasm respectively (Fig. B6f); the PrgK ring is assembled around, providing a membrane-like amphipathic environment (a built in nanodisc) (Fig. 3.3d, Fig. B6f) forming specific interactions

with SpaQ and the C-terminal most helical motif of SpaR (the packing around the surface exposed first SpaP (SpaP$^{1}$) is by contrast looser and discussed further below). A band enriched in aromatic and positively charged residues, more typical of membrane insertion, is shielded within the PrgK nanodisc (Fig. 3.3d). Remarkably, the resolution affords observation of the varied interactions formed by PrgK, details lost in previous 24-fold symmetry averaged reconstructions[17,26]. The detergent micelle is evident around the exposed hydrophobic surface of SpaPQR consistent with this region penetrating the outer leaflet of the IM (Fig. B6f).

The exterior lip of SpaPQR is defined by an insertion between the first two helical motifs of SpaP (SpaP$^{insertion}$) (Fig. 3.3a, d, Fig. B6d-e). This region is the most sequence variable between injectisome and flagellum (Fig. B7a) reflecting the specialized role observed here. The SpaP$^{insertion}$ forms a negatively charged surface, complimentary to the interior electropositive surface of the surrounding InvG N0 domains (Fig. B6g). Notably, the presence of PrgJ, which assembles around SpaP$^{insertion}$ (see below), results in a more ordered and cohesive collective interface. The helical rise of SpaPQR presents SpaP at increasing levels relative to InvG with an associated correlation of ordering of the C-terminal region of SpaP$^{insertion}$ (Fig. 3.2d). We are only able to fully trace the corresponding loop in SpaP$^{5}$ (albeit at a lower local resolution compared to the rest of the subunit) (Fig. B6b) where it is observed to directly interact with the InvG N0 domain and the PrgH C-terminus (Fig. 3.3a,d), an interface validated by identified but not previously understood cross-links between InvG and SpaP[21] (Fig. B5a). The location, increased system-specific sequence variation, mutually complimentary interfaces, and specific interactions within the NC observed here suggest that this insertion in SpaP orthologues provides

a tailored adaptation to couple the T3S export apparatus complex to the respective and distinct OM scaffolds of the injectisome or flagellar systems.

### 3.3.4 PrgJ anchors needle to SpaPQR and surrounding secretin

Our structure of SpaPQR with a short PrgIJ polymer (Fig. 3.1i-j) defines an initiation and anchoring mechanism of needle polymerization. Structures of the isolated helical needle PrgI are known[80,93]; however, how it is nucleated within the NC, including proper stoichiometry, and structure/function roles of PrgJ, has remained unclear[17,80,105]. Six PrgJ subunits bind $SpaP_5R_1$ (Fig. 3.4), consistent with previous and at the time puzzling proteomic analysis[74] given the historical interpretation of PrgJ as an "inner-rod" accounting for the filament within the span of the basal body[17,75]. The first PrgJ ($PrgJ^1$) is partially ordered, its C-terminal helix bound in an extensive groove at the interface of SpaR and $SpaP^1$ (1328 $Å^2$, $\Delta^iG$ = -34.0 kcal/mol)[143] (Fig. 3.4a-c), the N-terminal helices of which undergo binding-induced conformational reordering to accommodate PrgJ (Fig. B6c). The subsequent four PrgJ subunits ($PrgJ^{2-5}$) bind at SpaP-SpaP interfaces; however, they adopt helical hairpins similar to PrgI (Fig. 3.4). Unexpectedly, the N-terminal region of PrgJ interacts with a surrounding $InvG^{N0N1}$ oligomerization interface via β-strand complementation with the InvG N1 domain β-sheet (Fig. 3.4a-b, Fig. B7b), consistent with earlier detected pulldowns of EPEC rod and secretin orthologues[146]. In $PrgJ^2$, this is observed in alternate locations (Fig. 3.4b, d), we propose a result of recruitment of the additional $InvG^{N0N1}$ domain (discussed below). The final $PrgJ^6$ binds at the $SpaP^5$-SpaR interface (Fig. 3.4b, d). The structure and assembly of PrgJ, tailored to the varied binding surfaces presented by SpaPR and anchored to InvG, explain its necessity for

needle assembly (as opposed to direct binding of PrgI), in turn providing a more homogenous platform to initiate polymerization of the downstream PrgI needle, and explaining prior cross-linking data[73] and mutants[147–149] (clustered in the C-terminal helix) that impact needle assembly and secretion. Notably, mutation of PrgJ $Asp^{31}$ (N-terminal helix interacting with SpaP) or $Tyr^{77}$ (C-terminal helix at the interface with SpaR or SpaP) completely abrogated secretion[147,148], with the latter also abolishing both PrgJ incorporation and needle assembly. $Tyr^{77}$ forms key interactions with SpaR, SpaP, and the PrgJ N-terminal helix, that contribute to the direct anchoring and correct tertiary structure of PrgJ (Fig. B5d). We note these mutants were identified using plasmid-complementation, which has been suggested to result in artifacts when specifically assaying for needle length (due to secretion competition between early substrates[150]); however, the especially deleterious effects of these mutants in agreement with the structure we observe support a critical role in PrgJ assembly. Interestingly, we identified mutations of conserved positively charged residues lining the interior channel of PrgJ that allow secretion of early and middle, but not late, substrates (Fig. B5c-f) suggesting a specific role of the channels internal surface in regulating secretion hierarchy (discussed below).

In this shortened needle particle reconstruction, five PrgI needle protomers subsequently bind both PrgJ and SpaP, completing the first two turns of the needle and providing the base for further polymerization (Fig. 3.4a-b, d). The N-terminal regions of these initial PrgI subunits are variably ordered without the high-affinity PrgI-PrgI binding sites normally present downstream in the homo-oligomeric helical assembly. Significantly, the direct interactions that serve to anchor PrgJ and PrgI to the $SpaP_5R_1$ ring mimic those observed in the assembled PrgI needle,

where individual subunits are incorporated via homotypic interactions of the inner-facing C-terminal helix with the lateral and axial neighboring PrgI protomers (Fig. B6h). Thus, the disposition of the $SpaP_5R_1$ N-terminal helices specifically defines the 11-start helical symmetry parameters of the polymerized PrgI needle. Continued needle polymerization is mediated via lateral interactions with neighboring PrgI subunits and axial interactions with the PrgJ hairpin loop, accounting for the weakly attached needle phenotype observed for a PrgJ $Asp^{63}$ mutant contained therein[148] (Fig. B5d).

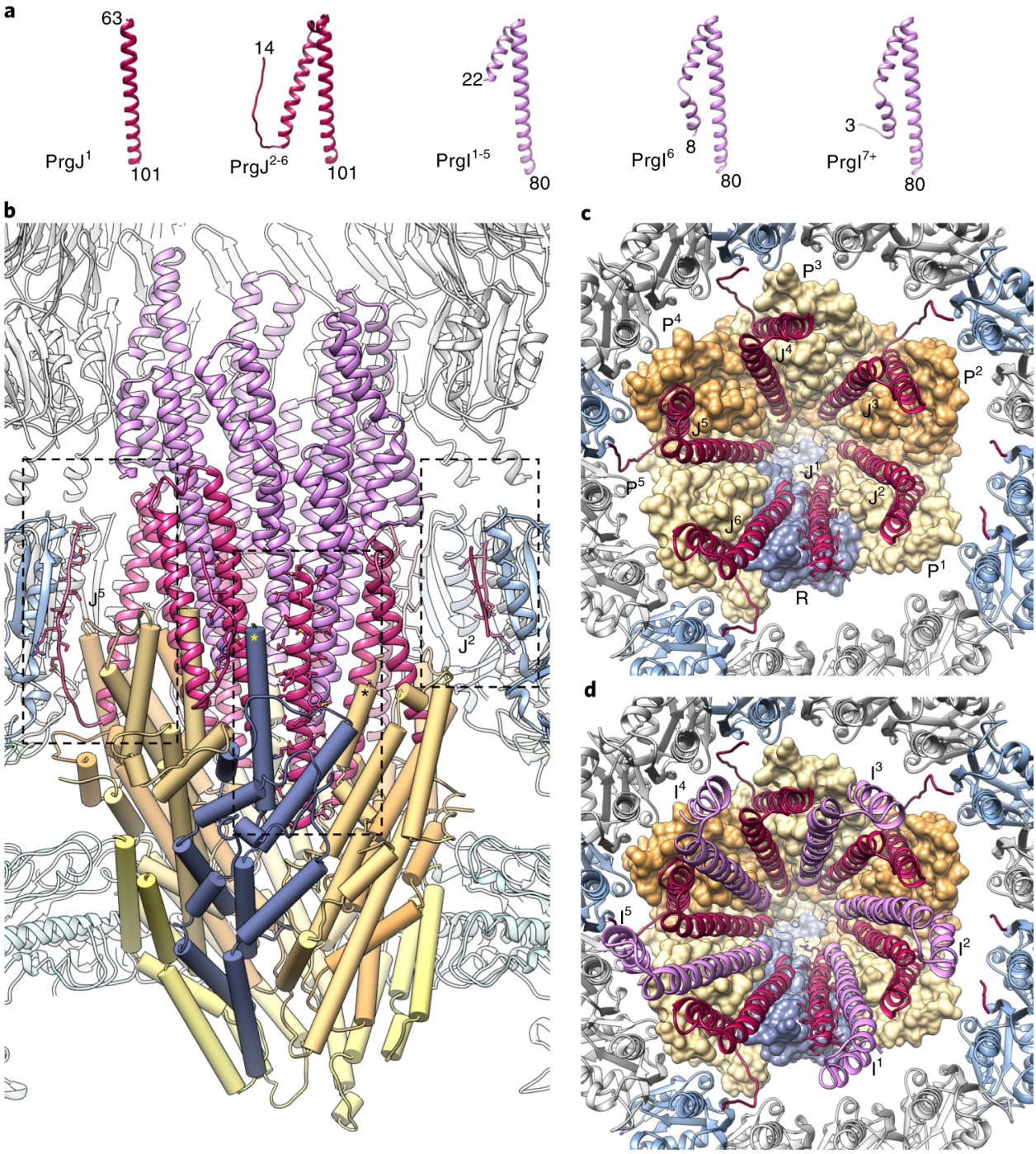



**Fig. 3.4: The SpaPQR complex forms an assembly platform for PrgJ and PrgI needle polymerization.**

**a, Atomic models of PrgJ (cherry) and PrgI (pink) showing the varied ordered states observed. Only the C-terminal helix of $PrgJ^{1}$ is ordered while $PrgJ^{2-6}$ are ordered throughout. The first turn of PrgI subunits $PrgI^{1-6}$ show different degrees of ordering in the N-terminus without the typical neighboring interactions of the assembled needle. b, Central view of the SpaPQR complex (tubes colored as in Fig. 3) with PrgJ and PrgI attached. The central box shows the location of the ordered C-terminal helix of $PrgJ^{1}$, bound between the N-terminal helices of SpaR and SpaP, denoted by yellow and black asterisks respectively. Left and right boxes show the interaction of the PrgJ N-terminal loop (sticks) with the InvG N1 domains via a beta-strand complementation interaction. In $PrgJ^{2}$ ($J^{2}$; right), situated at the site of the differential location of the extra $InvG^{N0N1}$ domain, the loop is observed in alternate positions. c, Top view of SpaPQR in surface representation showing the binding sites for $PrgJ^{1-6}$. The ordered C-terminal helices bind at the interface of SpaR and $SpaP^{1}$ ($PrgJ^{1}$), SpaP and SpaP ($PrgJ^{2-5}$), and SpaP and SpaR ($PrgJ^{6}$). The N-terminal helices of $PrgJ^{2-6}$ bind the $SpaP^{insertion}$, and the N-terminal loop binds the InvG N1 domain interface (The surrounding $InvG^{N0N1}$ domains are colored blue if bound to the N-terminal loops of PrgJ). d, As c with further $PrgI^{1-5}$ subunits shown. These make up the first full turn of the 11-start helical needle polymer.**

### 3.3.5 Structural snapshots of needle assembly and secretin gating

Needle polymerization necessitates opening of the secretin periplasmic gate, formed by the pentadecameric radial projections of an inner β-barrel hairpin that restricts the pore in the closed state[26]. Remarkably the four distinct assembly snapshots captured here (Fig. 3.1b-e) reveal the stepwise conformational changes of the $InvG^{secretin}$ pore. The $InvG^{N0N1}$ domains and IM rings by contrast remain static throughout. The closed and open secretin gate states reflect our previous structures of the isolated $InvG^{secretin}$ oligomer[26] and a focused refinement of the NC secretin[80] respectively, which revealed the structural changes associated with the presence of

the needle (although itself poorly resolved due to symmetry mismatch in that reconstruction). We hypothesized a two-stage gating mechanism[80]: an allosteric step where the assembling needle would trigger the observed changes in the N3 ring domain that would unlock the gate (disruption of the interactions between the T3S specific β-hairpin insertion[26] of N3 and the secretin inner β-barrel that stabilizes the closed gate[80]), and a steric step where continued polymerization would provide the force to straighten and reorient the inner β-barrel to create a ~75 Å pore that houses the internal needle. We clearly observe these features in the NC states here, with notably PrgI needle subunits additionally individually resolved throughout (Fig. 3.5a, c). We also capture the overall straightening of the InvG$^{secretin}$ pore with respect to the IM complex and InvG$^{N0N1}$ domains on initial needle insertion (Fig. 3.5b). This shift in InvG$^{secretin}$ pore orientation and potential strain on the connecting loops with the InvG$^{N0N1}$ domains may explain the associated ordering of the N3 domain variable loop[80] (Fig. 3.5b) (disordered in the closed state; Fig. 3.5a), which fortifies the NC scaffold via interactions at the N3 domain oligomerization interface and buttressing against the N1 domain ring below.

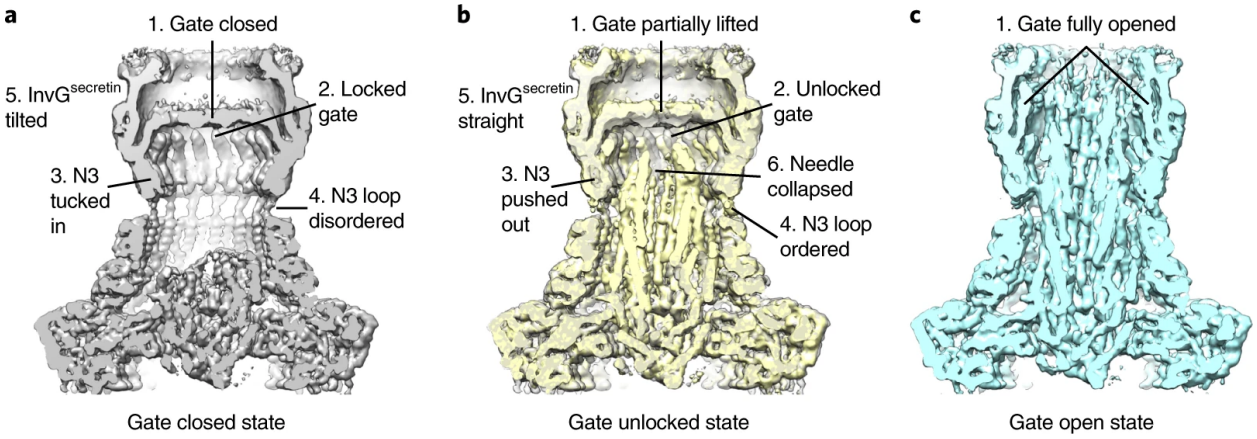


**Fig. 3.5: The structural snapshots capture the transitions associated with needle assembly and secretin gating.**

**Reconstructions have been low pass filtered to highlight the general structural features. a, In the gate closed state of the NC-base (grey), the InvG$^{secretin}$ perisplasmic gate (1) is held closed via stabilizing interactions of the N3 domain β-hairpin insertion specific to T3S secretin homologues (2). The bottom of the N3 domains is tilted into the lumen (3) and the N3 variable loop is disordered (4). The InvG$^{secretin}$ pore is tilted with respect to the rest of the complex (5). b, Assembly of the initial PrgJ and PrgI subunits atop the SpaPQR complex (yellow, gate closed state overlaid in transparent grey) collectively pushes the N3 domains out (3) with re-orientation of the N3 domain β-hairpin (2); without the stabilizing interactions the gate is partially opened (1). The N3 variable loop is ordered, reinforcing both the N3 domain oligomerization interface and interfacing with the N1 domain ring (4). The InvG$^{secretin}$ orientation has straightened due to needle insertion (5). The subunits at the tip of the short needle are collapsed into the channel without the lateral packing interactions of the assembled needle (6). c, Further needle polymerization results in the full extension of the inner β-barrel (1) to accommodate the central channel (cyan).**

Significantly, the snapshots with short needles have captured intermediate states between open and closed OM secretin. In the unlocked state (Fig. 3.1d, 3.5b), the needle is composed of 11 PrgI subunits. The terminal two immediately under the InvG$^{secretin}$ gate are partially ordered while the remaining four that cap the needle are collapsed into the channel without the lateral packing present in the assembled helical polymer (Fig. 3.5b). Around this, the InvG$^{secretin}$ N3 ring domain conformation is similar to fully assembled needle described above: collectively pushed out with the T3S specific β-hairpin in the "unlocked" position, the relative orientation of the InvG$^{secretin}$ pore straightened, and the N3 domain variable loop ordered (Fig. 3.5b). The secretin

periplasmic gate is partially raised compared to the fully closed conformation. The partially unlocked state lacks a neighboring PrgI triplet (Fig. 3.1c). The N3 domains surrounding this consequently adopt a more closed conformation: tilted in, the β-hairpin extending toward the gate (locked), and the variable loop disordered. These snapshots demonstrate directly that the changes in the N3 gating domain are specifically induced by interactions with PrgI "substrate" and is consistent with our proposed allosteric gating. The degree of gate opening correlates with PrgI stoichiometry (Fig. B8i-k), consistent with the proposed steric opening phase, and continued PrgI polymerization, as captured in the open state (Fig. 3.1d, 3.5c), would provide the additional force to support the fully open gate conformation. We note these snapshots with attached PrgIJ needle may represent needle depolymerization during purification and thus secretin gate closing; however, we envisage that the general structural transitions will be equally applicable in reverse.

## 3.4 Discussion

The presented structures provide a significant leap forward in understanding of the architecture, assembly and function of the T3S injectisome. Internal variation in OM secretin stoichiometry and coupling interface with IM PrgH, as well as capture of the SpaPQR export apparatus in its NC environs, the role of the PrgJ rod as a connector that anchors needle to export apparatus and NC-base, combined with stepwise needle induced conformational changes within the $InvG^{secretin}$ pore are each milestone contributions in understanding of this complex dual membrane spanning nanomachine.

Alternative pathways of NC assembly in different systems have been proposed[72,82] (see ) Fig. B8). The *Salmonella* SPI-1 injectisome NC is proposed to follow an inside-out pathway of assembly with the IM rings nucleated by export apparatus components prior to recruitment of the OM secretin[72,151] (supported by formation of otherwise intact NC in a Δ*invG* mutant[16]) and is discussed here. Efficient assembly of IM components depends on SpaPQR, which nucleates oligomerization of the encapsulating PrgHK nested rings[72], captured here with $SpaP_5Q_4R_1$:$PrgK_{24}PrgH_{24}$ stoichiometry (Fig. B8a,c). The export apparatus proteins SpaPQRS have historically been predicted as multi-span IM proteins based on their hydrophobic helical character[73,152]; however, the structure of isolated flagellar $FliP_5Q_4R_1$ predicted instead location on the outer leaflet of the IM[21]. The $SpaP_5Q_4R_1$ structure within the assembled NC here demonstrates this directly, further revealing that the PrgK ring specifically provides a membrane-like environment to surround the export apparatus complex (Fig. 3d). This natural nanodisc provides hydrophobic and polar interactions to the amphipathic SpaPQR surface (Fig. B6f) that could assist in incremental removal from the lipid bilayer as proposed for FliPQR assembly[21], with mutually stabilizing interactions between PrgK and SpaPQR providing an explanation as to why NC-base assembly is inefficient in the absence of export apparatus[72]. This partitioning could also explain the distorted local bilayer structure observed in lower resolution *in situ* cryo-ET analyses[18]. We observe space for additional components within the PrgK nanodisc around the C-terminal repeat of SpaR and the neighboring $SpaP^1$ (lacking an interacting SpaQ) (Fig. B9a). A less stable $FliP_5Q_5R_1$ complex was detected[21], and we speculate this represents the native stoichiometry with the peripheral location of the first SpaQ (FliQ) leading to more ready dissociation. SpaS, also typically lost in purification[17,74,80], is

predicted as a single copy[74] with gene-fusion[83] and cross-linking[73] localizing it near the C-terminus of SpaR and SpaP, here mapping to this same loosely packed region. SpaS is critical for regulation of secretion hierarchy via its C-terminal autoprotease domain[100]. It is predicted to have four N-terminal transmembrane domains; however, we detect similarity to helical repeats of FliPQR (Fig. B9b) suggesting again participation in the pseudo-helical export apparatus assembly on the membrane surface. This could place the C-terminal substrate-switching domain of SpaS near the entrance to the export apparatus, well positioned to regulate secretion via coordinating opening/access to the secretory channel possibly in combination with the export gate component InvA.

OM secretins assemble independently from the IM components they couple to[78,82]. Recent structures from the T2SS[90,91] and T3SS[26,80] show the β-barrel domain forms a pentadecameric pore. We previously proposed roles of the highly conserved amphipathic helical loop (AHL) in OM association and the 45-stranded β-barrel "lip" in subsequent BAM-independent OM insertion and span[26] (Fig. B8b,d). Extending this, we observe ordered detergent in the closed state InvG$^{secretin}$ reconstruction (Fig. B3), with the bound orientation consistent with a role of the AHL in penetrating the lipid bilayer and the local OM distortion observed *in situ*[18,92]. The peripheral N-terminal domains (N0 and N1) of isolated T3SS secretins studied by EM have been uniformly poorly ordered[22,26,153] suggesting destabilization without the IM rings they pack against; indeed we suggested the inherent weaker self-association interfaces may provide needed plasticity for assembly[80]. Our structure here reveals the extensive interface between OM and IM components, completed via incorporation of

an additional InvG$^{N0N1}$ with no downstream ordered counterpart in the InvG$^{secretin}$ OM pore. The specific ring position, restricted to two neighboring locations (Fig. 3.2f), points to ordered yet stochastic assembly; likely either initiating or terminating IM/OM component coupling. We favor a model where sequential assembly of the 15 InvG$^{N0N1}$ domains belonging to the InvG$^{secretin}$ pore is initiated first at the site of the interaction with SpaP$^{5}$, with its favorably oriented complimentary electrostatic interface (Fig. 3.3b, Fig. B6g). The InvG$^{N0N1}$ self-association, multivalent interface with PrgH, and interactions with SpaP would drive further coupling. In the case of SpaP, the helical rise of the SpaPQR complex could contribute to more rapid assembly from SpaP$^{5}$ around to SpaP$^{1}$ (clockwise when viewed from the OM perspective) (Fig. B9c). To complete the symmetric interface with PrgH, we envisage that the additional InvG$^{N0N1}$ subunit, located around SpaP$^{1}$, would be subsequently recruited from a pool of membrane associated secretin monomers, which are targeted to the OM in a partially folded state in complex with cognate pilotin InvH[154] (Fig. B8h). Further experiments are required to investigate this hypothesis. This capacity at the N0-N1 domains appears to be specific to the T3SS; a recent structure of an assembled T2SS shows 15-fold symmetry of the secretin throughout[155]. Interestingly, recombinantly expressed T2SS secretin GspD was observed to form a minor population of hexadecameric rings[91] suggesting the stoichiometry does not preclude oligomerization and membrane association. Instead, it could be the necessary tight interface with the contained needle that provides a selective pressure to maintain the 15-fold symmetry within the T3S InvG$^{secretin}$ pore. The role/fate of the unresolved C-terminus of the additional InvG monomer is unclear. We see no evidence of it and presume it has been degraded; however, if this is physiological and mediated by a specific protease remains to be determined.

In the monomeric and prepore states, the beta-barrel secretin domains are largely unstructured[156], associated with the OM via cognate pilotins and/or OM associating motifs[26] and it is possible that the additional unstructured InvG$^{secretin}$ domain maintains these interactions in the assembled NC.

The N-terminal loops of PrgJ$^{2-6}$ additionally bind at the internal InvG$^{N0N1}$ oligomerization interfaces, further reinforcing the assembly and anchoring the PrgIJ needle (Fig. 3.3b, 3.4b). Although the interacting PrgJ motif is not well conserved (Fig. B7b), the extended N-terminal sequence is a common feature and association between rod and secretin has been detected in EPEC[146], suggesting conservation of this interaction in other systems. Unlike PrgJ$^{3-6}$, which preferentially bind the equivalent InvG$^{N0N1}$, the PrgJ$^{2}$ N-terminus is observed in alternate locations at the site of the extra InvG$^{N0N1}$ (Fig. B9c). This could be a result of the helical presentation of PrgJ relative to InvG$^{N0N1}$, or could suggest that PrgJ is secreted at some point prior to recruitment of the final InvG$^{N0N1}$, with the N-terminus of PrgJ$^{2}$ binding the alternate location (orange box in Fig. B9c) when its preferred binding site (blue box in Fig. B9c) is unavailable. Although this remains to be validated, T3S of early substrates is not dependent on secretin coupling with needle polymerization still occurring, albeit inefficiently, in a Δ*invG* strain[16]. PrgJ rod orthologues are proposed to regulate localization and activity of T3SS-dedicated lytic enzymes to locally clear peptidoglycan[157] during assembly, which would require them to be transiently accessible before encapsulation within the NC lumen. Thus, we propose that, under native conditions, PrgJ is assembled prior to IM/OM coupling, instigating local peptidoglycan clearing and assisting inefficient coupling of IM and OM complexes (Fig.

B8e-h). We note formation of NC bases in T3S deficient mutants[80,158] suggests PrgJ is not essential in this, rather we envisage it contributes to efficient assembly under native expression conditions as suggested by a reduction in T3S on deletion of the T3SS lytic transglycosylase in *Citrobacter rodentium*[157].

The structure of PrgJ has implications for models of needle length control and substrate switching from early (needle and rod) to middle (needle tip and translocases required for pore formation in host membrane) and late substrates (effectors destined for the host). In the *Salmonella* SPI-1 injectisome, needle length is controlled by InvJ and two models have been proposed. In the first, InvJ acts as a molecular ruler to directly detect needle length and trigger the switch via an interaction with export apparatus protein SpaS[158], similar to models proposed for other T3SSs[159] and flagellar hook length control[98]. In the second model, based on interpretation that the PrgJ inner rod accounted for the span within the basal body, and deletion of *invJ* resulted in long, weakly attached needles and a reported apparent absence of the inner rod structure[148,158], InvJ instead functions to mediate assembly of a PrgJ inner rod in parallel with the PrgI needle, the timing and termination of which provides the trigger for substrate switching[148,158]. The PrgJ structure here, which does not self-associate and forms an intimate and seamless polymerization with PrgI on $SpaP_5R_1$, is inconsistent with this later model. PrgJ mutants have also been identified that selectively abolish secretion of middle substrates, assembling wild-type length NCs but with reduced level of cellular invasion[148]. These predominantly map to the base of the N-terminal helix, ordered in $PrgJ^{2-6}$ and involved in the interaction with SpaP (Fig. B5d). Thus, although the independent assembly of the inner rod

as a trigger for substrate switching seems unlikely given our structure, the correct anchoring of PrgJ to the export apparatus could be a requirement for efficient switching from early to middle substrates to occur.

We have also identified a PrgJ mutant that abrogates secretion of late but not early/middle substrates (Fig. B5c); mapped to the structure, these residues define an interior electropositive path which mutation to alanine significantly alters (Fig. B5d-f). This is consistent with our earlier observations of the PrgI needle interior[80] and mutation of equivalent residues in *Shigella* needle MxiH similarly abolishes ability to induce secretion of late effectors without impacting tip/translocon components[127]. This positive charge distribution is also shared by the interior atrium of the SpaPQR export apparatus here (Fig. B5e-f), and the export apparatus and analogous filament components of the flagellum[21], raising the intriguing possibility that local electrostatic forces are critical for protein translocation, either directly via interactions with different substrate classes or indirectly via altered conductance of ions or other small chemical signals that may regulate secretion hierarchy[11].

In summary, we have determined structures of the *Salmonella* SPI-1 injectisome NC in four states, providing insight into NC assembly. Notably, our model describes a stoichiometric mismatch within the OM secretin, an elegant evolutionary adaptation to couple the incompatible symmetries of the OM and IM complexes. In addition to complete visualization of NC architecture, these structures provide the first detailed atomic framework to understand the wealth of genetic, biochemical, and biophysical studies of this intensely studied molecular

nanomachine that drives pathogenicity in many notorious and drug resistant Gram-negative bacterial pathogens.

# Chapter 4: Conclusion

The structural characterization of the injectisome has advanced over the past 30 years. A significant portion of the complex has been determined using X-ray crystallography and NMR methods of isolated components in a non-native environment, with the inherent technical limitations in structure determination therein limiting functional understanding[23–25,100,157]. With the recent advances in infrastructure and ultimate resolution of single particle cryo-EM studies[160], as well as more and more ways to stabilize membrane protein complexes, the structure determination of the largely intact injectisome may not be far off in the future. Cryo-ET analysis, on the other front, has also benefited from the EM infrastructure advances and has made huge progress in elucidating the more loosely attached components in a truly native cellular environment[18,81,96].

## 4.1 Results of this work

In this work, the complete structure of the needle complex has been determined with the advance of gentle sample growth and preparation and the cryo-EM technique[63,80]. Many newly determined sub-complexes have answered many long-standing questions in the T3SS needle complex. The needle filament structure largely confirmed the previous ssNMR model[20] but also highlighted the crucial differences in terms of needle filament helicity and assembly in vivo vs. in vitro, with and the repeating electrostatic surfaces suggesting a new potential mechanism of virulence effector substrate passage through the needle length to the host cytosol. The InvG secretin structure is the first one in the family captured in an open conformation and revealed the

molecular details of the gating mechanism of this giant outer membrane portal. The symmetry mismatch within the InvG secretin itself, with localized recruitment of an additional monomer to the coupling domain only, explained the long-standing symmetry mismatch mystery with an elegant evolutionary adaptation. The structure of the export apparatus is the first in a T3SS injectisome native environment and confirmed the previous hypothesis that these predicted membrane proteins[124] actually reside atop of the membrane within the "nanodisc" like environment made by the surrounding IM nested rings. The structures show as well that the export apparatus forms the basic helical repeats for the inner rod facilitated correct polymerization of the needle filament. The inner rod structure showed the location and unexpected lack of self-polymerization by the rod subunits themselves, instead playing a crucial role as a structural adaptor and mediator of proper helical parameters between the export apparatus and filament. This feature is likely at the heart of why reconstituted needle filaments in earlier studies had altered helical parameters. Further, the snapshot of multiple assembly stages pictured in our work provided a clear view of the assembly of the needle complex including the growing needle filament polymerization and consequent secretin gate opening at various stages, all highlighting the amazing dynamic assembly of this giant nanomachine. Most importantly, having all the components together in one map with multiple snapshots in different assembly steps allowed us to visualize the complex interaction between all components and dependent conformational changes that are needed for the assembly and secretion process.

The structural information obtained from the series maps greatly advanced our understanding of the structure, assembly, and dynamic of the needle complex and set a strong foundation for the

following up research and future development of therapeutics. The techniques involved in this work, particularly the strain engineering and native expression and isolation of the intact (as possible) massive complex and the complicated cryo-EM data processing strategies including those needed to overcome the local symmetry mismatch amongst the many components of the T3SS system, can be valuable for further similar works in the T3SS and other complex membrane localized transport systems.

## 4.2 The limitations and further experiments

### 4.2.1 The missing components: InvA

Although the complete needle complex structure has been obtained through cryo-EM, there are a few more weakly associated but functionally critical components of the complete injectisome that are typically lost in even the most gentle of purification procedures[17,63,74,81]. For example several components of the inner membrane anchored “cytosolic sorting platform” and the essential ATPase[28] are missing in these preparations. In particular, the expected gate protein, InvA, is not resolved in the needle complex maps. The structure of the c-terminal domain of InvA, which is a soluble region protruding into the cytosol, has been determined by x-ray crystallography[79] and, more recently, by cryo-EM[29–31] highlighting a clear nonomeric stoichiometry. The N-terminal multi-span transmembrane region, which is presumed to interact with the export apparatus components determined in our needle complex reconstructions and thought to drive secretion via an unknown PMF fuelled process, is not resolved in the cryo-EM density in these studies[30,31]. Thus, the native expression of the InvA in the context of the fully assembled T3SS is an appealing next hurdle for stabilization and structural studies. To date, the

highest resolution of the cytosolic sorting platform has been accomplished by sub-tomogram averaging in cryo-ET studies, with focused refinement in that region to again overcome symmetry mismatch and signal domination by other components[18,81]. Although of relatively low resolution, an overall seahorse shape can be seen in the transmembrane region, and many dynamic interactions between the c-terminal cytosolic domain and the sorting platform are also observed, all indicating the protein can form a stable conformation in an assembled form if appropriately gentle extraction methods can be devised. Another cryo-ET sub-tomogram averaging study shows that the transmembrane region of InvA itself can also form stable rings without the c-terminus[18]. Together, it provides a promising path to isolate the transmembrane region of the protein in its native environment. Recently, another effort to capture InvA by extracting the complex in a milder detergent has been published. Although InvA was not captured in the resulting cryo-EM reconstruction, the membrane that InvA is embedded in can be seen in the 2D class averages as well as the final 3D volume[81]. The choice of mild DDM detergent may still be too harsh for this loosely associated component, and I predict that not removing the potentially dynamic C-terminal domain from this construct could also lead to interference and disassociation from the export apparatus during the purification. The combination of an even more gentle detergent such as GDN and removal of the c-terminal cytosolic domain “cross talk”, could lead to the isolation of a stable complex that is suitable for high-resolution cryo-EM study.

### 4.2.2 The missing components: SpaS

Another protein similar to InvA that is lost during the purification of the needle complex is the inner membrane anchored protein SpaS. The autocleavage of SpaS is responsible for switching secretion chronology from the middle to the late substrates[100]. The C-terminal domain that is responsible for autocleavage was solved by x-ray crystallography, elucidating the residues involved and catalytic mechanism[100]. More recently a cryo-EM study captured part of the N-terminal region of SpaS interacting with the export apparatus[161]. However, presently, being unable to accurately place the C-terminal domains in the context of a fully assembled needle complex the exact mechanism of the substrate switch autocleavage remains unknown. Together, the complex purification can benefit from a further optimized purification scheme. Combined with the power of SPA of cryo-EM, it is likely a more complete injectisome structure can be obtained and providing more insight into the function and regulation of the secretion process.

### 4.2.3 The path of secretion

Another prospect that can benefit from a more complete injectisome map is to identify the path of secretion as well as the mechanism of secretion. A recent cryo-EM needle complex map in complex with a specifically designed substrate trapped in the center lumen provides some insight about the path of the translocation as well as subtle conformational change that happen along with the delivery of the substrate[65]. But due to the limitation of local resolution as well as missing key components of injectisome (mentioned above), the secretion mechanism is still not understood. By maintaining the gate protein InvA and the switch SpaS in future characterized

complexes, the molecular details underlying the path of secretion will become complete. Although the retention of the loosely attached components may be achieved by a more optimized purification scheme, the resolve of the lower local resolution of the substrate may only be overcome by the further improved cryo-EM technology. As the substrate occupies a long-length of a symmetry diverse channel, it is obvious the substrate will not be in the same conformation in the secretion channel of every protein complex. To resolve the substrate in a better resolution, a lot more data needs to be collected to obtain enough particles that have a similar conformation of the substrate. The 3D reconstitution from those relatively homogeneous particles may achieve high enough resolution to allow the *de novo* building of the substrate. The high-resolution structural information can help understanding of the secretion mechanism and will set a solid foundation for the development of drugs to inhibit the T3SS. So far, there are a few good candidates of the lead structure of the inhibitors[4], but without knowing where these specifically bind to the complex or the mechanism of inhibition, further optimization is challenging. The studies here already significantly pave the way for such understanding going forward. With the rapid development of the direct electron detector with faster and faster readout speed and the advances higher throughput data collection strategies, it is possible to collect enough particles and fully resolve the substrate and inhibitors in the near future.

### 4.2.4 A complete assembly snapshots

Lastly, although the T3SS needle complex has been captured in four different assembly stages in our studies, there are still many steps that we do not have a direct structure to support our assembly hypothesis. As powerful as SPA can be, if the subcomplex cannot be isolated, it would

not work. To overcome the difficulty of complex isolation, the sub-tomogram averaging approach is a good alternative pathway, particularly when the components in question have a structure at high resolution by single particle cryo-EM or X-ray crystallography so they can be docked into the lower resolution sub-tomogram maps. This method has already proven to work well in terms of identifying the location of several of the components in the cytosolic sorting platform[18]. Due to the weaker association between the proteins in the sorting platform, the sorting platform has not yet been purified as one stable complex. Thus, the architecture of that region was a mystery. The determination of the 17Å sub-tomogram average of the entire injectisome combined with GPF attached to each component unambiguously determined the location of all the proteins that make up the sorting platform. This approach could be further used to determine the assembly sequence of the injectisome in situ as well, provided enough data at the different assembly stages can be collected.

### 4.2.5 Summary

To summarize, the goal of this work was to further our understanding of the T3SS using the cryo-EM technique. The new structures that were determined in this work significantly advance our understanding of the T3SS needle complex in terms of structure, function, and assembly. The methods that developed and improved in this work also set a solid foundation for future work on the T3SS or other complex nanomachines. Overall, the information gained from this work also provides a strong step forward for the further study of the T3SS in terms of structure-guided development of therapeutics that can help combat antibiotic resistance.

## Bibliography


1. WHO: **Antimicrobial Resistance Global Report on Surveillance**. 2014, doi:10.1016/j.giec.2020.06.004.

2. JIM O'NEILL: **TACKLING DRUG-RESISTANT INFECTIONS GLOBALLY: FINAL REPORT AND RECOMMENDATIONS**. 2016,

3. Wang W, Arshad MI, Khurshid M, Rasool MH, Nisar MA, Aslam MA, Qamar MU: **Antibiotic resistance : a rundown of a global crisis**. *Infect Drug Resist* 2018,

4. Lyons BJE, Strynadka NCJ: **On the road to structure-based development of anti-virulence therapeutics targeting the type III secretion system injectisome**. *Medchemcomm* 2019, **10**:1273–1289.

5. Beceiro A, Tomás M, Bou G: **Antimicrobial resistance and virulence: A successful or deleterious association in the bacterial world?** *Clin Microbiol Rev* 2013, **26**:185–230.

6. Costa TRD, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, Waksman G: **Secretion systems in Gram-negative bacteria: Structural and mechanistic insights**. *Nat Rev Microbiol* 2015, **13**:343–359.

7. Lycklama a Nijeholt JA, Driessen AJM: **The bacterial Sec-translocase: Structure and mechanism**. *Philos Trans R Soc B Biol Sci* 2012, **367**:1016–1028.

8. Palmer T, Berks BC: **The twin-arginine translocation (Tat) protein export pathway**.

*Nat Rev Microbiol* 2012, **10**:483–496.

9. Kanonenberg K, Schwarz CKW, Schmitt L: **Type I secretion systems - a story of appendices**. *Res Microbiol* 2013, **164**:596–604.

10. Nivaskumar M, Francetic O: **Type II secretion system: A magic beanstalk or a protein escalator**. *Biochim Biophys Acta - Mol Cell Res* 2014, **1843**:1568–1577.

11. Deng W, Marshall NC, Rowland JL, McCoy JM, Worrall LJ, Santos AS, Strynadka NCJ, Finlay BB: **Assembly, structure, function and regulation of type III secretion systems**. *Nat Rev Microbiol* 2017, **15**:323–337.

12. Hueck CJ: **Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants**. *Microbiol Mol Biol Rev* 1998, **62**:379–433.

13. Diepold A, Armitage JP: **Type III secretion systems: the bacterial flagellum and the injectisome**. *Philos Trans R Soc B Biol Sci* 2015, **370**.

14. Kubori T, Matsushima Y, Nakamura D, Uralil J, Lara-Tejero M, Sukhan A, Galán JE, Aizawa SI: **Supramolecular structure of the salmonella typhimurium type III protein secretion system**. *Science (80- )* 1998, **280**:602–605.

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

16. Schraidt O, Lefebre MD, Brunner MJ, Schmied WH, Schmidt A, Radics J, Mechtler K, Galán JE, Marlovits TC: **Topology and Organization of the Salmonella typhimurium Type III Secretion Needle Complex Components**. *PLoS Pathog* 2010, **6**:e1000824.

17. Schraidt O, Marlovits TC: **Three-dimensional model of Salmonella's needle complex at subnanometer resolution**. *Science (80- )* 2011, **331**:1192–1195.

18. Hu B, Lara-Tejero M, Kong Q, Galán JE, Liu J: **In Situ Molecular Architecture of the Salmonella Type III Secretion Machine**. *Cell* 2017, **168**:1065–1074.

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

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

21. Kuhlen L, Abrusci P, Johnson S, Gault J, Deme J, Caesar J, Dietsche T, Mebrhatu MT, Ganief T, Macek B, et al.: **Structure of the core of the type iii secretion system export apparatus**. *Nat Struct Mol Biol* 2018, **25**:1–8.

22. Spreter T, Yip CK, Sanowar S, André I, Kimbrough TG, Vuckovic M, Pfuetzner RA, Deng W, Yu AC, Finlay BB, et al.: **A conserved structural motif mediates formation of the periplasmic rings in the type III secretion system**. *Nat Struct Mol Biol* 2009,

**16**:468–476.

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

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

25. Bergeron JRC, Worrall LJ, Sgourakis NG, DiMaio F, Pfuetzner RA, Felise HB, Vuckovic M, Yu AC, Miller SI, Baker D, et al.: **A Refined Model of the Prototypical Salmonella SPI-1 T3SS Basal Body Reveals the Molecular Basis for Its Assembly**. *PLoS Pathog* 2013, **9**.

26. Worrall LJ, Hong C, Vuckovic M, Deng W, Bergeron JRC, Majewski DD, Huang RK, Spreter T, Finlay BB, Yu Z, et al.: **Near-atomic-resolution cryo-EM analysis of the Salmonella T3S injectisome basal body**. *Nature* 2016, **540**:597–601.

27. Korotkov K V., Delarosa JR, Hol WGJ: **A dodecameric ring-like structure of the N0 domain of the type II secretin from enterotoxigenic Escherichia coli**. *J Struct Biol* 2013, **183**:354–362.

28. Majewski DD, Worrall LJ, Hong C, Atkinson CE, Vuckovic M, Watanabe N, Yu Z,

Strynadka NCJ: **Cryo-EM structure of the homohexameric T3SS ATPase-central stalk complex reveals rotary ATPase-like asymmetry**. *Nat Commun* 2019, **10**:1–12.

29. Majewski DD, Lyons BJE, Atkinson CE, Strynadka NCJ: **Cryo-EM analysis of the SctV cytosolic domain from the enteropathogenic E. coli T3SS injectisome**. *J Struct Biol* 2020, **212**:107660.

30. Matthews-Palmer TRS, Gonzalez-Rodriguez N, Calcraft T, Lagercrantz S, Zachs T, Yu XJ, Grabe GJ, Holden DW, Nans A, Rosenthal PB, et al.: **Structure of the cytoplasmic domain of SctV (SsaV) from the Salmonella SPI-2 injectisome and implications for a pH sensing mechanism**. *J Struct Biol* 2021, **213**:107729.

31. Kuhlen L, Johnson S, Cao J, Deme JC, Lea SM: **Nonameric structures of the cytoplasmic domain of FlhA and SctV in the context of the full-length protein**. *PLoS One* 2021, **16**:1–10.

32. Luo Y, Bertero MG, Frey EA, Pfuetzner RA, Wenk MR, Creagh L, Marcus SL, Lim D, Sicheri F, Kay C, et al.: **Structural and biochemical characterization of the type III secretion chaperones CesT and SigE**. *Nat Struct Biol 2001 812* 2001, **8**:1031–1036.

33. Yip CK, Finlay BB, Strynadka NCJ: **Structural characterization of a type III secretion system filament protein in complex with its chaperone**. *Nat Struct Mol Biol 2004 121* 2004, **12**:75–81.

34. Allen KJ, Rogan D, Brett Finlay B, Potter AA, Asper DJ: **Vaccination with type III**

**secreted proteins leads to decreased sheddingin calves after experimental infection with Escherichia coli o157**. *Can J Vet Res* 2011, **75**:98–105.

35. Raymond B, Young JC, Pallett M, Endres RG, Clements A, Frankel G: **Subversion of trafficking, apoptosis, and innate immunity by type III secretion system effectors**. *Trends Microbiol* 2013, **21**:430–441.

36. Van Der Heijden J, Finlay BB: **Type III effector-mediated processes in Salmonella infection**. *Future Microbiol* 2012, **7**:685–703.

37. Beeckman DSA, Vanrompay DCG: **Bacterial secretion systems with an emphasis on the chlamydial Type III secretion system.** *Curr Issues Mol Biol* 2010, **12**:17–41.

38. Santos AS, Finlay BB: **Bringing down the host: Enteropathogenic and enterohaemorrhagic Escherichia coli effector-mediated subversion of host innate immune pathways**. *Cell Microbiol* 2015, **17**:318–332.

39. B. BJ, Wang X, Viboud GI, Brodsky I: **Modulation of innate immune responses by Yersinia type III secretion system translocators and effectors**. *Cell Microbiol* 2013, **15**:1622–1631.

40. Galán JE: **Salmonella Typhimurium and inflammation: a pathogen-centric affair**. *Nat Rev Microbiol* 2021, **0123456789**.

41. Frank J: **Three-Dimensional Electron Microscopy of Macromolecular Assemblies**.

*Three-Dimensional Electron Microsc Macromol Assem* 2010, doi:10.1093/acprof:oso/9780195182187.001.0001.

42. Bhella D: **Cryo-electron microscopy: an introduction to the technique, and considerations when working to establish a national facility**. *Biophys Rev 2019 114* 2019, **11**:515–519.

43. Adrian M, Dubochet J, Lepault J, McDowall A: **Cryo-electron microscopy of viruses**. *Nature* 1984, **308**:32–33.

44. Cheng Y: **Single-particle cryo-EM-How did it get here and where will it go**. *Science (80- )* 2018, **361**:876–880.

45. Wade RH: **A brief look at imaging and contrast transfer**. *Ultramicroscopy* 1992, **46**:145–156.

46. De Rosier DJ, Klug A: **Reconstruction of three dimensional fiber structures from orthogonal projections**. *Nature* 1968, **217**:130–134.

47. Frank J: **Averaging of low exposure electron micrographs of non-periodic objects**. *Ultramicroscopy* 1975, **1**:159–162.

48. Chen S, McMullan G, Faruqi AR, Murshudov GN, Short JM, Scheres SHW, Henderson R: **High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy**.

*Ultramicroscopy* 2013, **135**:24–35.

49. Mooney P: **Optimization of Image Collection for Cellular Electron Microscopy**. *Methods Cell Biol* 2007, **2007**:661–719.

50. McMullan G, Faruqi AR, Henderson R: *Direct Electron Detectors*. Elsevier Inc.; 2016.

51. Li X, Mooney P, Zheng S, Booth CR, Braunfeld MB, Gubbens S, Agard DA, Cheng Y: **Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM**. *Nat Methods 2013 106* 2013, **10**:584–590.

52. Bai XC, Fernandez IS, McMullan G, Scheres SHW: **Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles**. *Elife* 2013, **2013**.

53. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA: **CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination**. *Nat Methods* 2017, **14**:290–296.

54. Sigworth FJ: **A maximum-likelihood approach to single-particle image refinement**. *J Struct Biol* 1998, **122**:328–339.

55. Scheres SHW: **A bayesian view on cryo-EM structure determination**. *J Mol Biol* 2012, **415**:406–418.

56. Scheres SHW: **RELION: Implementation of a Bayesian approach to cryo-EM structure determination**. *J Struct Biol* 2012, **180**:519–530.

57. Huiskonen JT: **Image processing for cryogenic transmission electron microscopy of symmetry-mismatched complexes**. *Biosci Rep* 2018, **38**:1–13.

58. Roh SH, Hryc CF, Jeong HH, Fei X, Jakana J, Lorimer GH, Chiu W: **Subunit conformational variation within individual GroEL oligomers resolved by Cryo-EM**. *Proc Natl Acad Sci U S A* 2017, **114**:8259–8264.

59. Huang G, Zhang Y, Zhu X, Zeng C, Wang Q, Zhou Q, Tao Q, Liu M, Lei J, Yan C, et al.: **Structure of the cytoplasmic ring of the Xenopus laevis nuclear pore complex by cryo-electron microscopy single particle analysis**. *Cell Res* 2020, **30**:520–531.

60. Liu H, Cheng L: **Cryo-EM shows the polymerase structures and a nonspooled genome within a dsRNA virus**. *Science (80- )* 2015, **349**:1347–1350.

61. Kotecha A, Wang Q, Dong X, Ilca SL, Ondiviela M, Zihe R, Seago J, Charleston B, Fry EE, Abrescia NGA, et al.: **Rules of engagement between αvβ6 integrin and foot-and-mouth disease virus**. *Nat Commun* 2017, **8**.

62. Morais MC, Tao Y, Olson NH, Grimes S, Jardine PJ, Anderson DL, Baker TS, Rossmann MG: **Cryoelectron-microscopy image reconstruction of symmetry mismatches in bacteriophage φ29**. *J Struct Biol* 2001, **135**:38–46.

63. Hu J, Worrall LJ, Vuckovic M, Hong C, Deng W, Atkinson CE, Brett Finlay B, Yu Z, Strynadka NCJ: **T3S injectisome needle complex structures in four distinct states reveal the basis of membrane coupling and assembly**. *Nat Microbiol* 2019, **4**:2010–

2019.

64. Guo F, Liu Z, Vago F, Ren Y, Wu W, Wright ET, Serwer P, Jiang W: **Visualization of uncorrelated, tandem symmetry mismatches in the internal genome packaging apparatus of bacteriophage T7**. *Proc Natl Acad Sci U S A* 2013, **110**:6811–6816.

65. Miletic S, Fahrenkamp D, Goessweiner-Mohr N, Wald J, Pantel M, Vesper O, Kotov V, Marlovits TC: **Substrate-engaged type III secretion system structures reveal gating mechanism for unfolded protein translocation**. *Nat Commun* 2021, **12**:1–14.

66. Koning RI, Gomez-Blanco J, Akopjana I, Vargas J, Kazaks A, Tars K, Carazo JM, Koster AJ: **Asymmetric cryo-EM reconstruction of phage MS2 reveals genome structure in situ**. *Nat Commun* 2016, **7**:1–6.

67. Bai X, Rajendra E, Yang G, Shi Y, Scheres SHW: **Sampling the conformational space of the catalytic subunit of human g-secretase**. *Elife* 2015,

68. Zhou Q, Huang X, Sun S, Li X, Wang HW, Sui SF: **Cryo-EM structure of SNAP-SNARE assembly in 20S particle**. *Cell Res* 2015, **25**:551–560.

69. Ilca SL, Kotecha A, Sun X, Poranen MM, Stuart DI, Huiskonen JT: **Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes**. *Nat Commun* 2015, **6**:4–11.

70. Sun Z, El Omari K, Sun X, Ilca SL, Kotecha A, Stuart DI, Poranen MM, Huiskonen JT:

**Double-stranded RNA virus outer shell assembly by bona fide domain-swapping**. *Nat Commun* 2017, **8**.

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

72. Wagner S, Königsmaier L, Lara-Tejero M, Lefebre M, Marlovits TC, Galán JE: **Organization and coordinated assembly of the type III secretion export apparatus**. *Proc Natl Acad Sci* 2010, **107**:17745–17750.

73. Dietsche T, Tesfazgi Mebrhatu M, Brunner MJ, Abrusci P, Yan J, Franz-Wachtel M, Schärfe C, Zilkenat S, Grin I, Galán JE, et al.: **Structural and Functional Characterization of the Bacterial Type III Secretion Export Apparatus**. *PLoS Pathog* 2016, **12**:1–25.

74. Zilkenat S, Franz-Wachtel M, Stierhof Y-D, Galán JE, Macek B, Wagner S: **Determination of the Stoichiometry of the Complete Bacterial Type III Secretion Needle Complex Using a Combined Quantitative Proteomic Approach**. *Mol Cell Proteomics* 2016, **15**:1598–1609.

75. Marlovits TC, Kubori T, Sukhan A, Thomas DR, Galán JE, Unger VM: **Structural insights into the assembly of the type III secretion needle complex**. *Science (80- )* 2004, **306**:1040–1042.

76. Beck M, Baumeister W: **Cryo-Electron Tomography: Can it Reveal the Molecular Sociology of Cells in Atomic Detail?** *Trends Cell Biol* 2016, **26**:825–837.

77. Romano FB, Tang Y, Rossi KC, Monopoli KR, Ross JL, Heuck AP: **Type 3 secretion translocators spontaneously assemble a hexadecameric transmembrane complex**. *J Biol Chem* 2016, **291**:6304–6315.

78. Majewski DD, Worrall LJ, Strynadka NC: **Secretins revealed: structural insights into the giant gated outer membrane portals of bacteria**. *Curr Opin Struct Biol* 2018, **51**:61–72.

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

80. Hu J, Worrall LJ, Hong C, Vuckovic M, Atkinson CE, Caveney N, Yu Z, Strynadka NCJ: **Cryo-EM analysis of the T3S injectisome reveals the structure of the needle and open secretin**. *Nat Commun* 2018, **9**:1–11.

81. Butan C, Lara-Tejero M, Li W, Liu J, Galan J: **High-resolution view of the type III secretion export apparatus in situ reveals membrane remodeling and a secretion pathway**. *bioRxiv* 2019, doi:10.1101/709592.

82. Diepold A, Amstutz M, Sorg I, Jenal U, Cornelis GR: **Deciphering the assembly of the Yersinia type III secretion injectisome**. 2010, **29**:1928–1940.

83. Arnam JS Van, Mcmurry JL, Kihara M, Macnab RM: **Analysis of an Engineered Salmonella Flagellar Fusion**. *J Bacteriol* 2004, **186**:2495–2498.

84. Lara-Tejero M, Qin Z, Hu B, Butan C, Liu J, Galán JE: **Role of SpaO in the assembly of the sorting platform of a Salmonella type III secretion system**. *PLoS Pathog* 2019, **15**:1–20.

85. Bai F, Morimoto Y V., Yoshimura SDJ, Hara N, Kami-Ike N, Namba K, Minamino T: **Assembly dynamics and the roles of FliI ATPase of the bacterial flagellar export apparatus**. *Sci Rep* 2014, **4**:1–7.

86. Wilharm G, Lehmann V, Krauss K, Lehnert B, Richter S, Ruckdeschel K, Heesemann J, Trülzsch K, Tru K: **Yersinia enterocolitica Type III Secretion Depends on the Proton Motive Force but Not on the Flagellar Motor Components MotA and MotB Yersinia enterocolitica Type III Secretion Depends on the Proton Motive Force but Not on the Flagellar Motor Components M**. *Infect Immun* 2004, **72**:4004–4009.

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

88. Minamino T, Morimoto Y V., Hara N, Aldridge PD, Namba K: **The Bacterial Flagellar Type III Export Gate Complex Is a Dual Fuel Engine That Can Use Both H+ and Na+ for Flagellar Protein Export**. *PLoS Pathog* 2016, **12**:1–18.

89. Minamino T, Morimoto Y V., Hara N, Namba K: **An energy transduction mechanism**

**used in bacterial flagellar type III protein export**. *Nat Commun* 2011, **2**.

90. Hay ID, Belousoff MJ, Lithgow T: **Structural basis of type 2 secretion system engagement between the inner and outer bacterial membranes**. *MBio* 2017, **8**:1–6.

91. Yan Z, Yin M, Xu D, Zhu Y, Li X: **Structural insights into the secretin translocation channel in the type II secretion system**. *Nat Struct Mol Biol* 2017, **24**:177–183.

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

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

94. Guo EZ, Desrosiers DC, Zalesak J, Tolchard J, Berbon M, Habenstein B, Marlovits T, Loquet A, Galán JE: *A polymorphic helix of a Salmonella needle protein relays signals defining distinct steps in type III secretion*. 2019.

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

96. Park D, Lara-Tejero M, Waxham MN, Li W, Hu B, Galán JE, Liu J: **Visualization of the**

**type III secretion mediated salmonella–host cell interface using cryo-electron tomography**. *Elife* 2018, **7**:1–15.

97. Goessweiner-Mohr N, Kotov V, Brunner MJ, Mayr J, Wald J, Kuhlen L, Miletic S, Vesper O, Lugmayr W, Wagner S, et al.: **Structural control for the coordinated assembly into functional pathogenic type-3 secretion systems**. *bioRxiv* 2019, doi:10.1101/714097.

98. Erhardt M, Singer HM, Wee DH, Keener JP, Hughes KT: **An infrequent molecular ruler controls flagellar hook length in Salmonella enterica**. *EMBO J* 2011, **30**:2948–2961.

99. Journet L, Agrain C, Broz P, Cornelis GR: **The Needle Length of Bacterial Injectisomes Is Determined by a Molecular Ruler**. *Science (80- )* 2003, **302**:1757–1760.

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

101. Wiesand U, Sorg I, Amstutz M, Wagner S, van den Heuvel J, Lührs T, Cornelis GR, Heinz DW: **Structure of the Type III Secretion Recognition Protein YscU from Yersinia enterocolitica**. *J Mol Biol* 2009, **385**:854–866.

102. Büttner D: **Protein Export According to Schedule: Architecture, Assembly, and Regulation of Type III Secretion Systems from Plant- and Animal-Pathogenic**

**Bacteria**. *Microbiol Mol Biol Rev* 2012, **76**:262–310.

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

104. Frost S, Ho O, Login FH, Weise CF, Wolf-Watz H, Wolf-Watz M: **Autoproteolysis and Intramolecular Dissociation of Yersinia YscU Precedes Secretion of Its C-Terminal Polypeptide YscUCC**. *PLoS One* 2012, **7**.

105. Radics J, Königsmaier L, Marlovits TC: **Structure of a pathogenic type 3 secretion system in action**. *Nat Struct Mol Biol* 2014, **21**:82–87.

106. Yip CK, Strynadka NCJ: **New structural insights into the bacterial type III secretion system**. *Trends Biochem Sci* 2006, **31**:223–230.

107. Hoiseth SK, Stocker BAD: **Aromatic-dependent Salmonella typhimurium are non-virulent and effective as live vaccines**. *Nature* 1981, **291**:238–239.

108. Kroger C, Dillon SC, Cameron ADS, Papenfort K, Sivasankaran SK, Hokamp K, Chao Y, Sittka A, Hebrard M, Handler K, et al.: **The transcriptional landscape and small RNAs of Salmonella enterica serovar Typhimurium**. *Proc Natl Acad Sci* 2012, **109**:E1277–E1286.

109. Tan YZ, Cheng A, Potter CS, Carragher B: **Automated data collection in single particle electron microscopy**. *Reprod Syst Sex Disord* 2016, **65**:43–56.

110. Mastronarde DN: **Automated electron microscope tomography using robust prediction of specimen movements**. *J Struct Biol* 2005, **152**:36–51.

111. Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA: **MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy**. *Nat Methods* 2017, **14**:331–332.

112. Rohou A, Grigorieff N: **CTFFIND4: Fast and accurate defocus estimation from electron micrographs**. *J Struct Biol* 2015, **192**:216–221.

113. Grant T, Grigorieff N: **Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6**. *Elife* 2015, **4**:1–19.

114. He S, Scheres SHW: **Helical reconstruction in RELION**. *J Struct Biol* 2017, **198**:163–176.

115. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE: **UCSF Chimera - A visualization system for exploratory research and analysis**. *J Comput Chem* 2004, **25**:1605–1612.

116. Adams PD, Afonine P V., Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, et al.: **PHENIX: A comprehensive Python-**

**based system for macromolecular structure solution**. *Acta Crystallogr Sect D Biol Crystallogr* 2010, **66**:213–221.

117. Wang RYR, Song Y, Barad BA, Cheng Y, Fraser JS, DiMaio F: **Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta**. *Elife* 2016, **5**:1–22.

118. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC: **MolProbity: All-atom structure validation for macromolecular crystallography**. *Acta Crystallogr Sect D Biol Crystallogr* 2010, **66**:12–21.

119. Barad BA, Echols N, Wang RYR, Cheng Y, Dimaio F, Adams PD, Fraser JS: **EMRinger: Side chain-directed model and map validation for 3D cryo-electron microscopy**. *Nat Methods* 2015, **12**:943–946.

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

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

122. Chan QWT, Howes CG, Foster LJ: **Quantitative Comparison of Caste Differences in Honeybee Hemolymph**. *Mol Cell Proteomics* 2006, **5**:2252–2262.

123. Cox J, Mann M: **MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification**. *Nat Biotechnol* 2008, **26**:1367–1372.

124. Kuhlen L, Abrusci P, Johnson S, Gault J, Deme J, Caesar J, Dietsche T, Mebrhatu MT, Ganief T, Macek B, et al.: **Structure of the Core of the Type Three Secretion System Export Apparatus**. *bioRxiv* 2018, doi:10.1101/249128.

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

126. Vitold G, Wolfgang S, Oliver S, Thomas M, Edward E: **The Structure of the Salmonella typhimurium Type III Secretion System Needle Shows Divergence from the Flagellar System**. *J Mol Biol* 2010, **396**:1392–1397.

127. Kenjale R, Wilson J, Zenk SF, Saurya S, Picking WL, Picking WD, Blocker A: **The needle component of the type III secreton of Shigella regulates the activity of the secretion apparatus**. *J Biol Chem* 2005, **280**:42929–42937.

128. Spagnuolo J, Opalka N, Wen WX, Gagic D, Chabaud E, Bellini P, Bennett MD, Norris GE, Darst SA, Russel M, et al.: **Identification of the gate regions in the primary structure of the secretin pIV**. *Mol Microbiol* 2010, **76**:133–150.

129. Burghout P, van Boxtel R, van Gelder P, Ringler P, Müller SA, Tommassen J, Koster M:

**Structure and electrophysiological properties of the YscC secretin \rfrom the type III secretion system of Yersinia enterocolitica**. *J Bacteriol* 2004, **186**:4645–4654.

130. Disconzi E, Guilvout I, Chami M, Masi M, Huysmans GHM, Pugsley AP, Bayan N: **Bacterial secretins form constitutively open pores akin to general porins**. *J Bacteriol* 2014, **196**:121–128.

131. Guilvout I, Chami M, Berrier C, Ghazi A, Engel A, Pugsley AP, Bayan N: **In Vitro Multimerization and Membrane Insertion of Bacterial Outer Membrane Secretin PulD**. *J Mol Biol* 2008, **382**:13–23.

132. Nans A, Kudryashev M, Saibil HR, Hayward RD: **Structure of a bacterial type III secretion system in contact with a host membrane in situ**. *Nat Commun* 2015, **6**:1–8.

133. Guilvout I, Nickerson NN, Chami M, Pugsley AP: **Multimerization-defective variants of dodecameric secretin PulD**. *Res Microbiol* 2011, **162**:180–190.

134. Salzer R, D'Imprima E, Gold VAM, Rose I, Drechsler M, Vonck J, Averhoff B: **Topology and structure/function correlation of ring-and gate-forming domains in the dynamic secretin complex of thermus thermophilus**. *J Biol Chem* 2016, **291**:14448–14456.

135. Kimbrough TG, Miller SI: **Contribution of Salmonella typhimurium type III secretion components to needle complex formation**. *Proc Natl Acad Sci* 2000, **97**:11008–11013.

136. Eichelberg K, Galán JE: **Differential regulation of Salmonella typhimurium type III secreted proteins by pathogenicity island 1 (SPI-1)-encoded transcriptional activators InvF and HilA**. *Infect Immun* 1999, **67**:4099–4105.

137. Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, Scheres SH: **New tools for automated high-resolution cryo-EM structure determination in RELION-3**. *Elife* 2018, **7**:1–22.

138. **The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.** [date unknown],

139. Zimmermann L, Stephens A, Nam S, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V: **A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core**. *J Mol Biol* 2018, **430**:2237–2243.

140. Edwards RA, Keller LH, Schi DM: **Improved allelic exchange vectors and their use to analyze 987P fimbria gene expression**. 1998, **207**:149–157.

141. Ferrie L, Hemery G, Nham T, Gue A, Mazel D, Beloin C, Ghigo J: **Silent Mischief : Bacteriophage Mu Insertions Contaminate Products of Escherichia coli Random Mutagenesis Performed Using Suicidal Transposon Delivery Plasmids Mobilized by Broad-Host-Range**. *J Bacteriol* 2010, **192**:6418–6427.

142. Sanowar S, Singh P, Pfuetzner RA, André I, Zheng H, Spreter T, Strynadka NCJ, Gonen T, Baker D, Goodlett DR, et al.: **Interactions of the transmembrane polymeric rings of**

**the Salmonella enterica serovar Typhimurium type III secretion system**. *MBio* 2010, **1**:1–8.

143. Krissinel E, Henrick K: **Inference of Macromolecular Assemblies from Crystalline State**. 2007, doi:10.1016/j.jmb.2007.05.022.

144. Erhardt M, Wheatley P, Kim EA, Hirano T, Zhang Y, Sarkar MK, Hughes KT, Blair DF, Ak EW, Author MM: **Mechanism of type-III protein secretion: Regulation of FlhA conformation by a functionally critical charged-residue cluster contributed to conception, design and analysis of experiments, and drafting the paper; P HHS Public Access Author manuscript**. *Mol Microbiol* 2017, **104**:234–249.

145. Ward E, Renault TT, Kim EA, Erhardt M, Hughes KT, Blair DF: **Type-III secretion pore formed by flagellar protein FliP**. *Mol Microbiol* 2018, **107**:94–103.

146. Sal-man N, Deng W, Finlay BB: **EscI : a crucial component of the type III secretion system forms the inner rod structure in enteropathogenic Escherichia coli**. 2012, **125**:119–125.

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

148. Lefebre MD, Galan JE: **The inner rod protein controls substrate switching and needle length in a Salmonella type III secretion system**. *Proc Natl Acad Sci* 2014, **111**:817–

822.

149. El Hajjami N, Moussa S, Houssa J, Monteyne D, Perez-Morga D, Botteaux A: **The inner-rod component of Shigella flexneri type 3 secretion system, MxiI, is involved in the transmission of the secretion activation signal by its interaction with MxiC**. *Microbiologyopen* 2018, **7**:1–11.

150. Wee DH, Hughes KT: **Molecular ruler determines needle length for the Salmonella Spi-1 injectisome**. *Proc Natl Acad Sci* 2015, **2015**:1–6.

151. Diepold A, Wagner S: **Assembly of the bacterial type III secretion machinery**. 2014, **38**:802–822.

152. Berger C, Robin GP, Bonas U, Koebnik R: **Membrane topology of conserved components of the type III secretion system from the plant pathogen Xanthomonas campestris pv . vesicatoria**. *Microbiology* 2010, doi:10.1099/mic.0.039248-0.

153. Kowal J, Chami M, Ringler P, Müller SA, Kudryashev M, Castaño-Díez D, Amstutz M, Cornelis GR, Stahlberg H, Engel A: **Structure of the dodecameric Yersinia enterocolitica secretin YscC and its trypsin-resistant core**. *Structure* 2013, **21**:2152–2161.

154. Koo J, Burrows LL, Lynne Howell P: **Decoding the roles of pilotins and accessory proteins in secretin escort services**. *FEMS Microbiol Lett* 2012, **328**:1–12.

155. Chernyatina A, Harry L: **Architecture of a bacterial type II secretion system**. *bioRxiv* 2018,

156. Guilvout I, Chami M, Disconzi E, Bayan N, Pugsley AP, Huysmans GHM: **Independent Domain Assembly in a Trapped Folding Intermediate of Multimeric Outer Membrane Secretins**. *Struct Des* 2014, **22**:582–589.

157. Burkinshaw BJ, Deng W, Lameignère E, Wasney GA, Zhu H, Worrall LJ, Finlay BB, Strynadka NCJ: **Structural analysis of a specialized type III secretion system peptidoglycan-cleaving enzyme**. *J Biol Chem* 2015, **290**:10406–10417.

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

159. Wagner S, Stenta M, Metzger LC, Dal M, Cornelis GR: **Length control of the injectisome needle requires only one molecule of Yop secretion protein P ( YscP )**. *Proc Natl Acad Sci* 2010, **107**.

160. Kühlbrandt W: **The resolution revolution**. *Science (80- )* 2014, **343**:1443–1445.

161. Kuhlen L, Johnson S, Andreas Zeitler, Sandra Bäurle, Deme JC, Rebecca Debo, Joseph Fisher, Samuel Wagner, Lea SM: **The flagellar substrate specificity switch protein FlhB assembles onto the extra-membrane export gate to regulate type three secretion**. *bioRxiv* 2019,

162. Kudryashev M, Stenta M, Schmelz S, Amstutz M, Wiesand U, Castaño-díez D, Degiacomi MT, Münnich S, Bleck CKE, Kowal J, et al.: **In situ structural analysis of the Yersinia enterocolitica injectisome**. 2013, doi:10.7554/eLife.00792.

# Appendices

## Appendix A Chapter 2 supplementary information

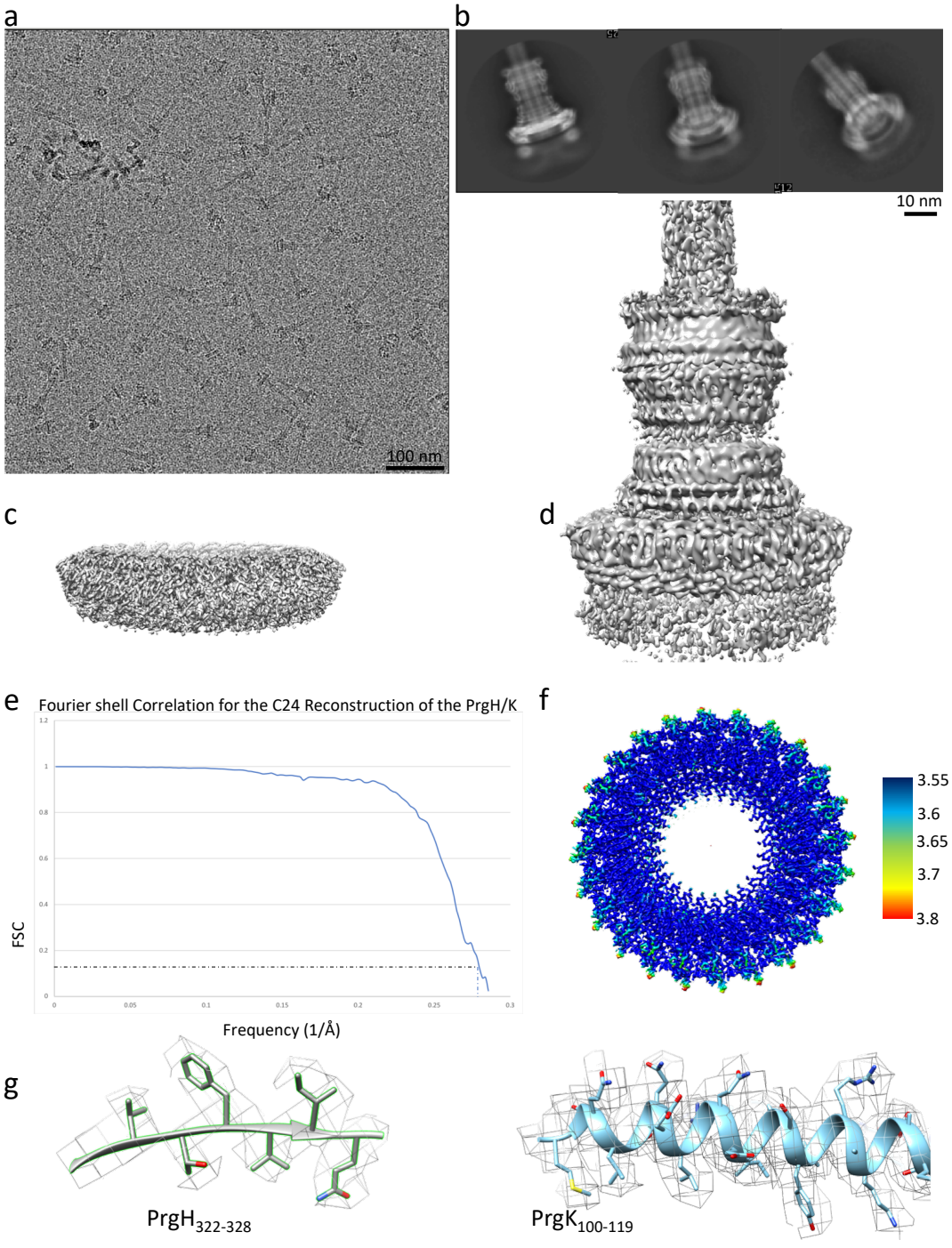

a
b
100 nm
10 nm
c
d
e
Fourier shell Correlation for the C24 Reconstruction of the PrgH/K
FSC
Frequency (1/Å)
f
3.55
3.6
3.65
3.7
3.8
g
PrgH322-328
PrgK100-119

**Fig. A1: 3D Reconstructions of the needle complex and 24-fold symmetry averaged inner membrane rings. a) Representative micrograph of the needle complex. b) Selected reference-free 2D class averages. c) C24 symmetry averaged inner membrane region at 3.6 Å resolution d) C1 (no symmetry imposed) reconstructed map of the needle complex. e) FSC of the C24 reconstruction calculated in Relion. f) Local resolution estimations of the C24 map from ResMap. Bottom slabbed view of the C24 reconstructed map of the basal body complex. g) Representative density for the inner membrane rings (3.6 Å resolution).**

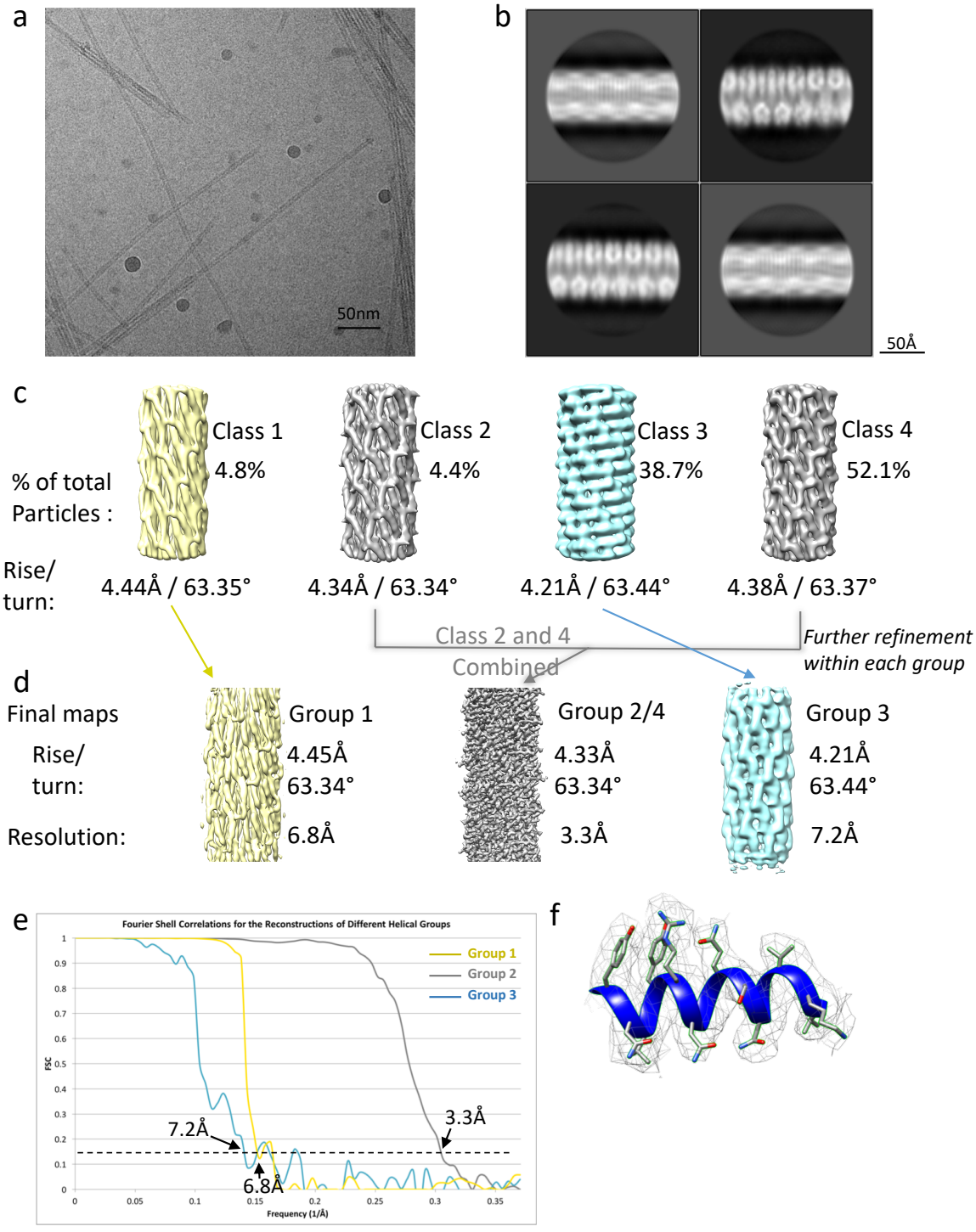



**Fig. A2: 3D Reconstruction of the isolated needle at 3.3 Å resolution. a) Representative micrograph of the isolated**

**needle. b) Selected reference-free 2D class averages. c) 3D classes, class 2 and 4 combined for final high-resolution refinement. d) Final maps for different classes revealing variation in helical rise. Group2/4 refined to high resolution. e) FSC of the different helical reconstructions. f) Representative density (3.3 Å resolution).**

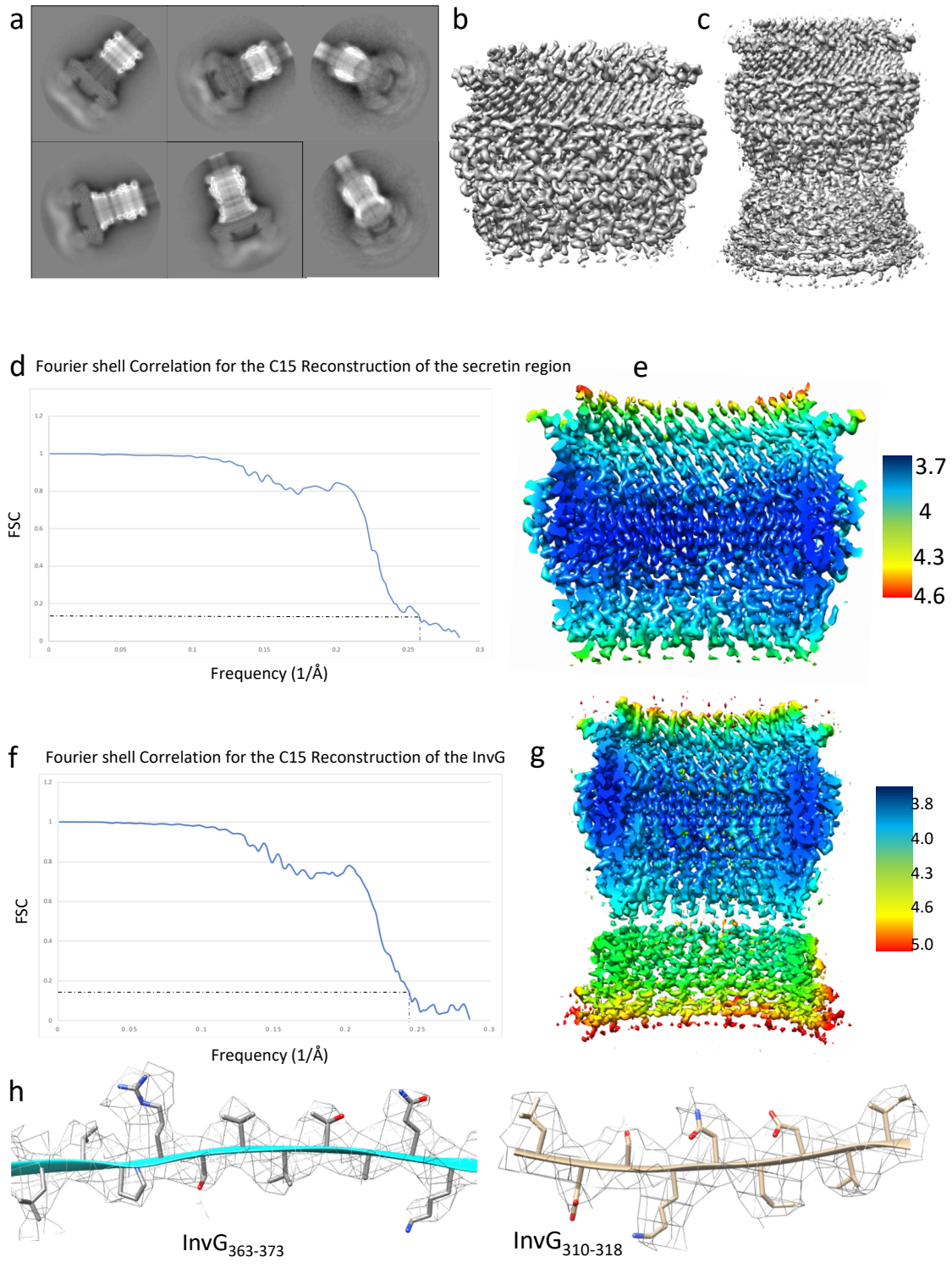



**Fig. A3: 3D Reconstructions of the 15-fold symmetry averaged outer membrane secretin.**

**a) Selected local reference-free 2D class averages of the core N3, secretin and S domains (top row) or full-length secretin (bottom row). b) C15 symmetry averaged core secretin at 3.9 Å resolution c) C15 symmetry averaged full-length secretin at 4.1 Å resolution. d) FSC for core secretin. e) Local resolution estimations of core secretin from ResMap. Slabbed view. f) FSC for full-length secretin. g) Local resolution estimations of full-length secretin from ResMap. Slabbed view. h) Representative density for the inner membrane rings (3.9 Å resolution).**

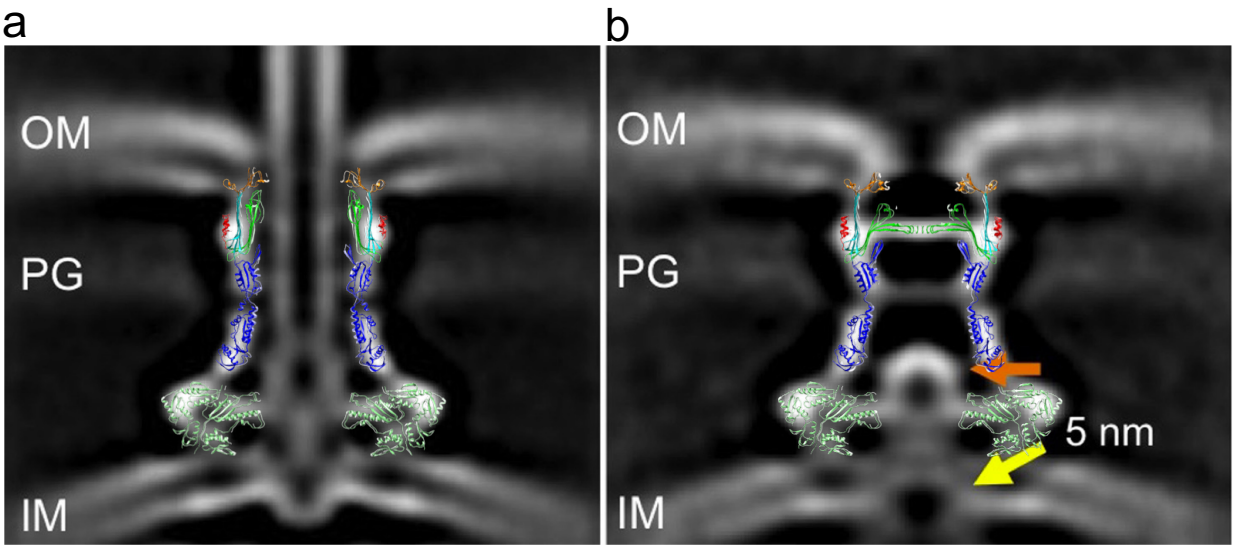



**Fig. A4: Comparison of isolated and *in situ* needle complex and basal body. Overlay of needle complex (a) and basal body (b) atomic models on the *in situ* cryo-ET image of the *S.* Typhimurium T3SS (Figure reproduced from REF). The relative structural span is conserved suggesting extraction from the membrane does not alter the overall structure significantly. Notably, the degree of membrane invagination caused by insertion of the secretin is correlated to orientation of the secretin lip beta-barrel.**

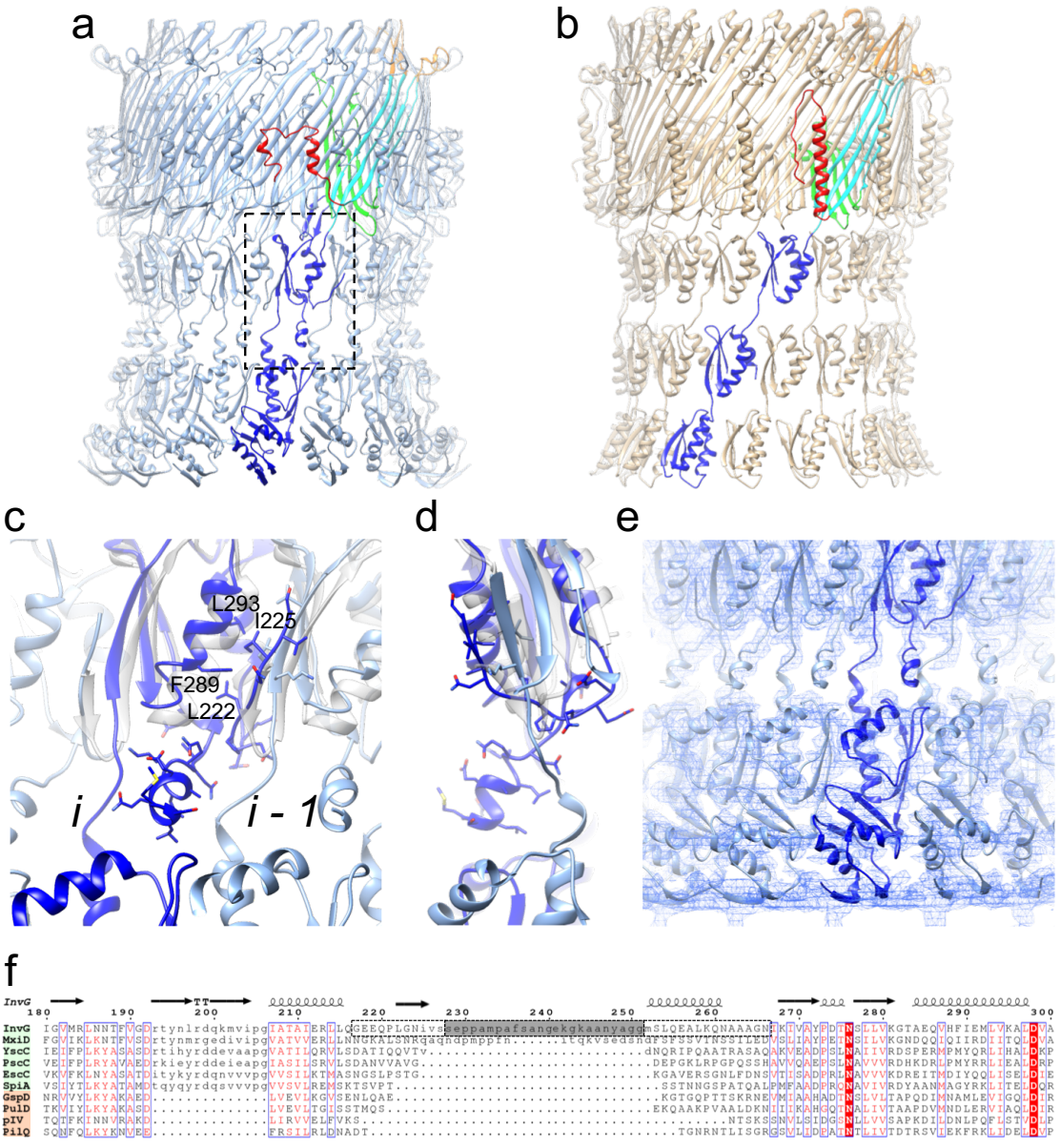



**Fig. A5: Packing of the peripheral N-terminal domains. (a) and (b) Secretin structures from the needle complex here and the T2SS secretin GspD (PDB 5WQ8). One monomer colored according to Fig. 1. (c) and**

**(d) Zoomed view of N3 domain (boxed in a) showing packing of ordered regions of N3 domain loop (shown as sticks). The isolated secretin (closed state) structure is overlaid in transparent grey illustrating the packing of the N-terminal most ordered region, which is folded back against the N3 domain. (e) Density for the N0, N1 and N3 domains, the N1 and N0 domains are less well resolved (also see Fig. 4g) although clear secondary structure density allowed positioning of the InvG$_{33-172}$ crystal structure and further refinement. (f) Multiple sequence alignment for the N3 domain from T3SS (green) and non-T3SS (orange) secretins. The variable N3 domain loop further resolved here is boxed with the region still disordered shaded grey.**

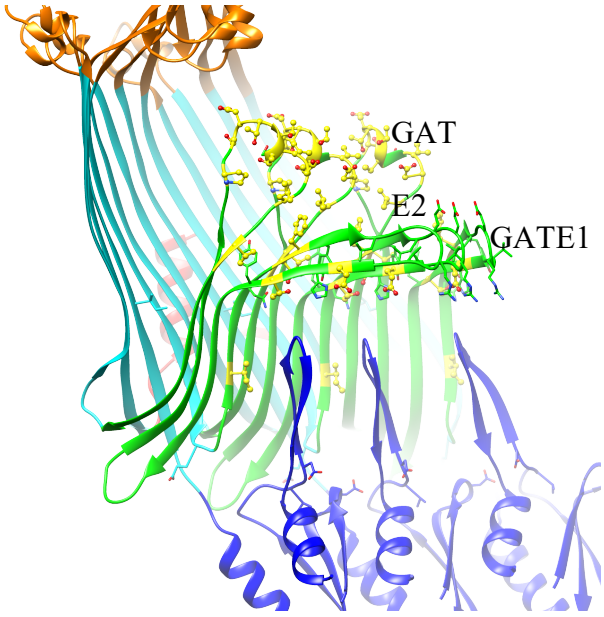



**Fig. A6: Mutations found to affect filamentous phage secretin pIV permeability mapped onto the InvG structure. Mutants map to the N3-secretin interface, the outer β-barrel wall and mostly to the GATE1 and GATE2 hairpins (shown as stick or ball and stick). The mutants found to increase sensitivity to antibiotics, suggesting a bigger impact on the degree of gate opening, are coloured yellow and show as ball and stick with a high concentration mapping to GATE2.**

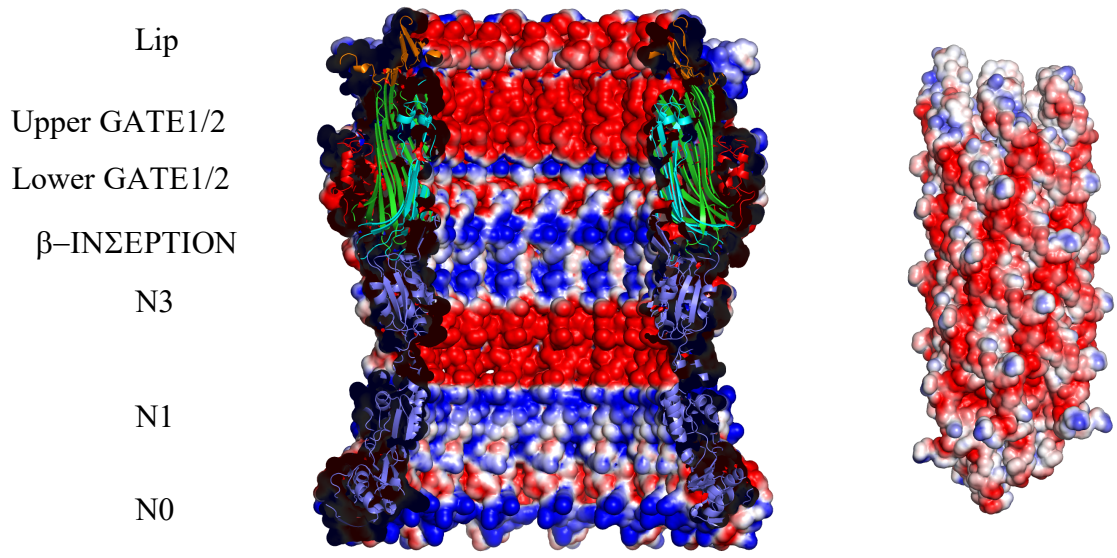



**Fig. A7: Electrostatic surface for InvG in the open state and the PrgI needle. Surface electrostatics calculated with APBS60 for the interior of the open InvG pore (a) and the PrgI needle (b). The InvG N1 and upper N3 domain including the β-INSERTION are mainly positively charged while the lower N3 domain and upper regions of the extended GATE1 and GATE hairpins and lip are electronegative. The β-INSERTION packs closely with the assembled needle filament within the secretin lumen and the charge complementarity with the predominantly negatively charged needle exterior suggest a possible means of further anchoring the needle within the basal body.**

| | #1 Needle complex C1 map | #2 $InvG_{34-557}$ C15 map (EMDB-8915) (PDB 6DV6) | #3 InvG176-557 C15 map (EMDB-8914) (PDB 6DV3) | #4 PrgH/K ring C24 map (EMDB-8913) (PDB 6DUZ) | #5 PrgI Needle Filaments (EMDB-8924) (PDB 6DWB) |
|---|---|---|---|---|---|
| Data collection and processing | | | | | |
| Magnification | 47000 | 47000 | 47000 | 47000 | 37037 |
| Voltage (kV) | 300 | 300 | 300 | 300 | 300 |
| Electron exposure ($e–/Å^2$) | 40 | 40 | 40 | 40 | 55 |
| Defocus range (μm) | 1.25-4 | 1.25-4 | 1.25-4 | 1.25-4 | 1-2.5 |
| Pixel size (Å) | 1.75 | 1.75 | 1.75 | 1.75 | 0.625 |
| Symmetry imposed | C1 | C15 | C15 | C24 | helical |
| Initial particle images (no.) | 80000 | 80000 | 80000 | 80000 | 867000 |
| Final particle images (no.) | 42647 | 26280 | 26000 | 58198 | 53000 |

| | | | | | |
|---|---|---|---|---|---|
| Map resolution (Å)<br>FSC threshold | 0.143 | 0.143 | 0.143 | 0.143 | 0.143 |
| | | | | | |
| Refinement | | | | | |
| Initial model used<br>(PDB code) | | 5TCQ | 5TCQ | 5TCP | |
| Model resolution<br>range (Å) | 7.4 | 4.1 | 3.9 | 3.6 | 3.3 |
| Map sharpening *B*<br>factor (Å$^2$) | -178 | -141 | -144 | -118 | -145 |
| Model composition<br>Non-hydrogen<br>atoms<br>Protein residues<br>Ligands | | 57975<br>7500 | 40950<br>5372 | 72768<br>9072 | 18360<br>2340 |
| *B* factors (Å$^2$)<br>Protein Ligand | | 9.9 | 13.87 | 11 | 19.6 |
| R.m.s. deviations | | | | | |
| Bond lengths (Å) | | 0.009 | 0.010 | 0.007 | 0.016 |
| Bond angles (°) | | 1.37 | 1.49 | 1.25 | 1.3 |
| Validation | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| MolProbity score | | 2.4 | 2.54 | 1.92 | 2.7 |
| Clashscore | | 8.63 | 8.13 | 4.63 | 6.12 |
| Poor rotamers (%) | | 8.59 | 7.64 | 4.3 | 0 |
| Ramachandran plot | | | | | |
| Favored (%) | | 5.24 | 6.5 | 3.18 | 3.95 |
| Allowed (%) | | 94.76 | 93.5 | 96.82 | 96.05 |
| Disallowed (%) | | 0 | 0 | 0 | 0 |

**Table A1. Cryo-EM data collection, refinement and validation statistics.**

| Protein ID | MW (kDa) | Peptides | Relative intensity |
|---|---|---|---|
| PRGK_SALTY | 30 | 32 | 1 |
| PRGI_SALTY | 8.9 | 16 | 0.92 |
| PRGH_SALTY | 44.5 | 48 | 0.74 |
| INVG_SALTY | 61.7 | 64 | 0.30 |
| PRGJ_SALTY | 10.9 | 10 | 0.16 |
| SPAP_SALTY | 25.2 | 9 | 0.06 |
| SIPD_SALTY | 37.1 | 19 | 0.01 |
| SIPA_SALTY | 73.9 | 18 | 0.005 |
| HILA_SALTY | 63 | 8 | 0.002 |
| INVA_SALTY | 76.1 | 10 | 0.002 |
| SPAR_SALTY | 28.5 | 1 | 0.001 |

| | | | |
|---|---|---|---|
| INVH_SALTY | 16.5 | 2 | 0.0004 |
| SIPB_SALTY | 62.5 | 2 | 0.0003 |

**Table A2. Mass spectrometry analysis of the purified needle complex.**

## Appendix B  Chapter 3 supplementary information

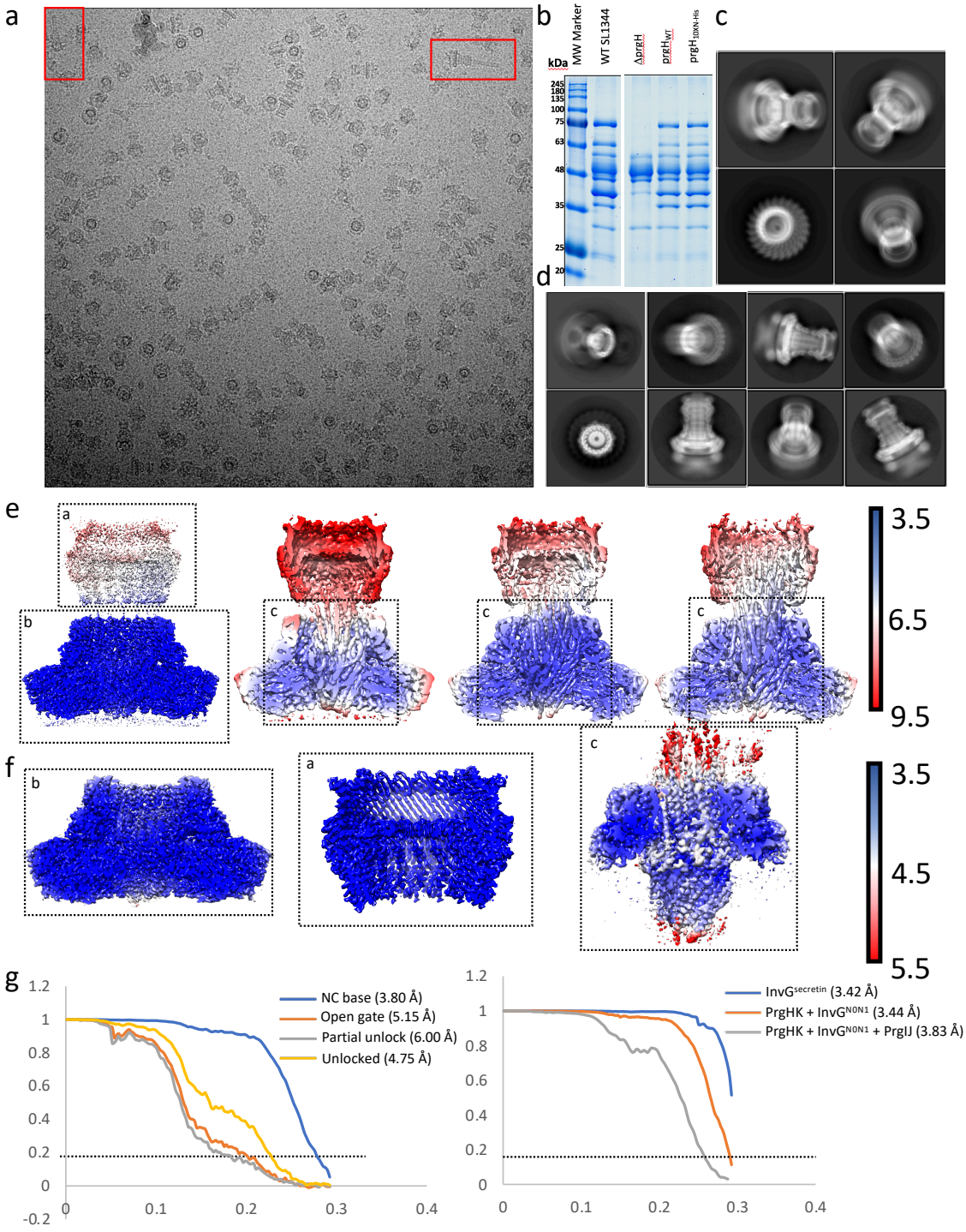

a
b
MW Marker
WT SL1344
ΔprgH
prgH_WT
prgH_10XN-His
kDa
245
180
135
100
75
63
48
35
25
20
c
d
e
a
b
c
3.5
6.5
9.5
f
3.5
4.5
5.5
g
NC base (3.80 Å)
Open gate (5.15 Å)
Partial unlock (6.00 Å)
Unlocked (4.75 Å)
InvG^secretin (3.42 Å)
PrgHK + InvG^N0N1 (3.44 Å)
PrgHK + InvG^N0N1 + PrgIJ (3.83 Å)

**Figure B1: Single particle cryo-EM data processing. a, Representative micrograph of the sample from overnight culture. Fully assembled NCs boxed. Inset is a negative stain EM micrograph of the purified sample after 4-hour expression showing increased proportion of NCs with assembled needles. We propose the variation in needle length is due to shearing/depolymerization and increased levels of immature NC-bases due to insufficient membrane and cytosolic components. Similar assembly distributions were noted in two separate sample preparations. b, Complementation of a prgH knockout with an N-terminally His-tagged PrgH plasmid (prgH 10xNhis) restores secretion as effectively as the full-length prgH plasmid (prgHWT) compared to the wild-type strain (WT SL1344). SDS-PAGE of secreted proteins, with the most abundant proteins labelled on right. Assays were repeated three times with the same results. c, Selected 2D class averages of the data. d, Selected 2D class averages of (from left to right) the subtracted SpaPQR and InvGN0N1 region, open state, unlocked state, and partially unlocked state. e-h, Local resolution of the four NC assembly snapshots as in Fig. 1b-e. The boxed regions correspond to the focused refinements in i-k as labelled. Resolution color indicated on right. i-k, Focused refinement reconstructions of the different regions boxed in e-h. Resolution color indicated on right. l, FSC curves for reconstructions in e-k.**

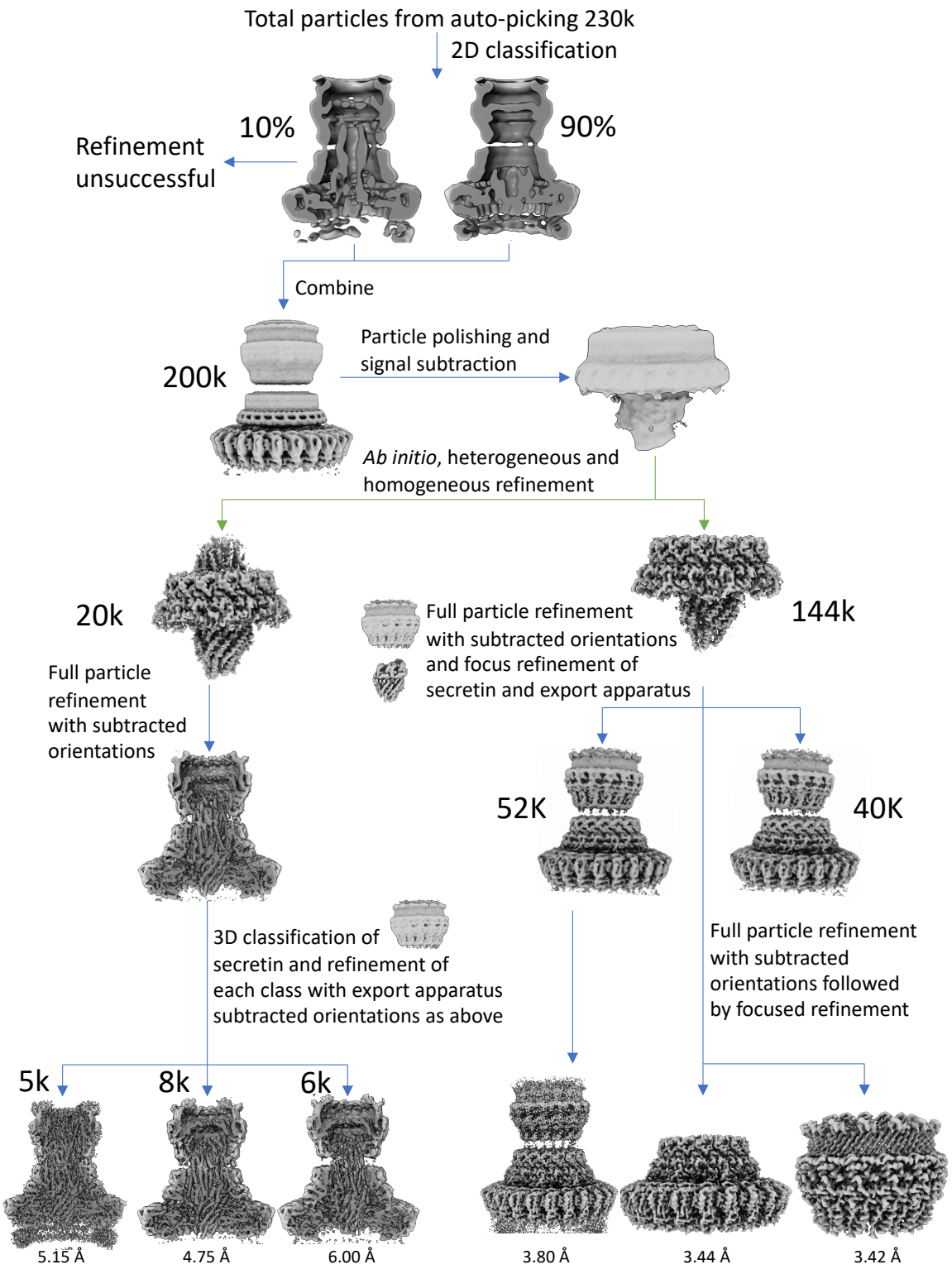

Total particles from auto-picking 230k
2D classification
10%
90%
Refinement unsuccessful
Combine
200k
Particle polishing and signal subtraction
*Ab initio*, heterogeneous and homogeneous refinement
20k
144k
Full particle refinement with subtracted orientations and focus refinement of secretin and export apparatus
Full particle refinement with subtracted orientations
52K
40K
3D classification of secretin and refinement of each class with export apparatus subtracted orientations as above
Full particle refinement with subtracted orientations followed by focused refinement
5k
8k
6k
5.15 Å
4.75 Å
6.00 Å
3.80 Å
3.44 Å
3.42 Å

**Figure B2: Single particle cryo-EM data processing strategy. Blue arrows correspond to processing in Relion, green cryoSPARC.**

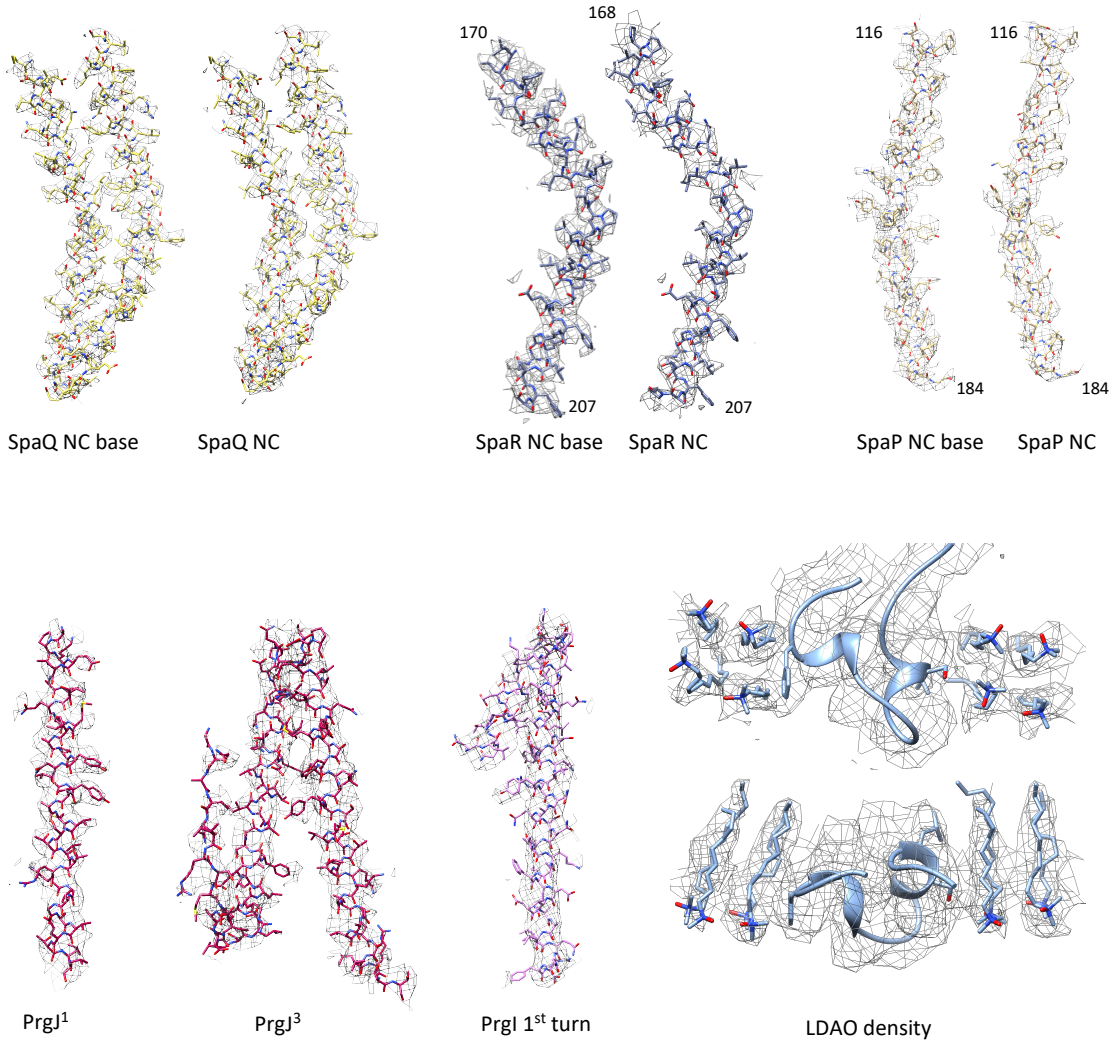


**Figure B3: Sample cryo-EM density. The SpaPQR and PrgIJ components from the focused refinement reconstructions in the absence (3.44 Å resolution) and presence (3.83 Å resolution) of PrgIJ. 16-mer InvGN0N1 region with no symmetry applied at 3.44 Å resolution. C15 InvGsecretin reconstruction at 3.42 Å resolution showing AHL with bound LDAO.**

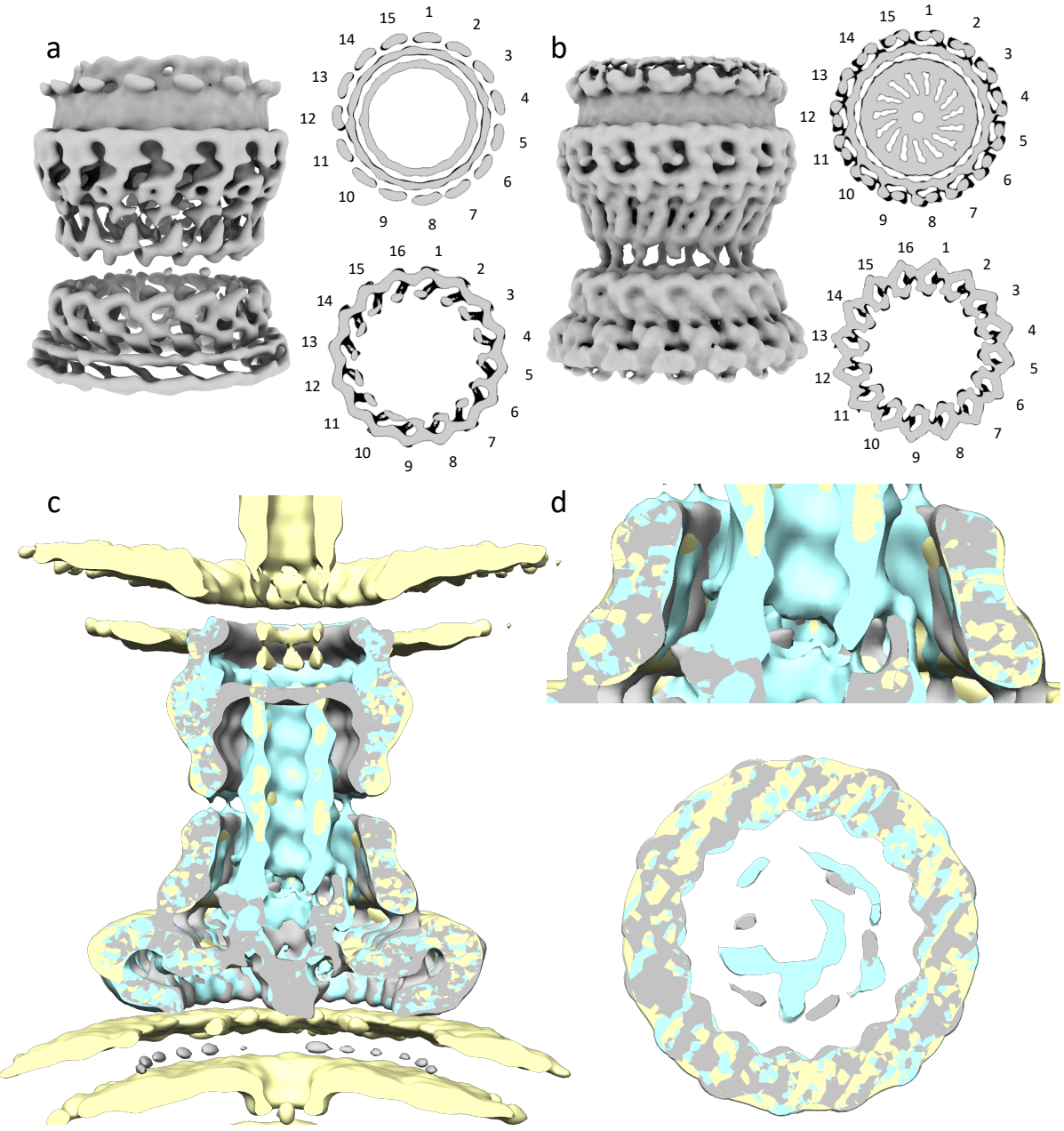



**Figure B4: Mixed secretin stoichiometry in other purified NC samples and comparison to the in situ injectisome. C1 InvG focused refinement from intact NC sample data (left), and top slab views (marked with dashed arrows) at the InvG$^{secretin}$ showing 15 copies (top right) and the InvG$^{N0N1}$ region showing 16 copies**

**(bottom right). b, C1 InvG focused refinement from the PrgH$_{130-392}$ mutant NC-base data**[26] **(left), and top slab views (marked with dashed arrows) at the InvG$^{secretin}$ showing 15 copies (top right) and the InvG$^{N0N1}$ region showing 16 copies (bottom right). c, Superposition of the *in situ* cryo-ET reconstruction**[18] **(yellow), and the low-pass filtered NC-base (grey) and NC reconstructions (cyan). d, Side and top zoomed views of the 16-mer InvG$^{N0N1}$ region from c showing near perfect agreement between high resolution cryo-EM and lower resolution cryo-ET reconstructions.**

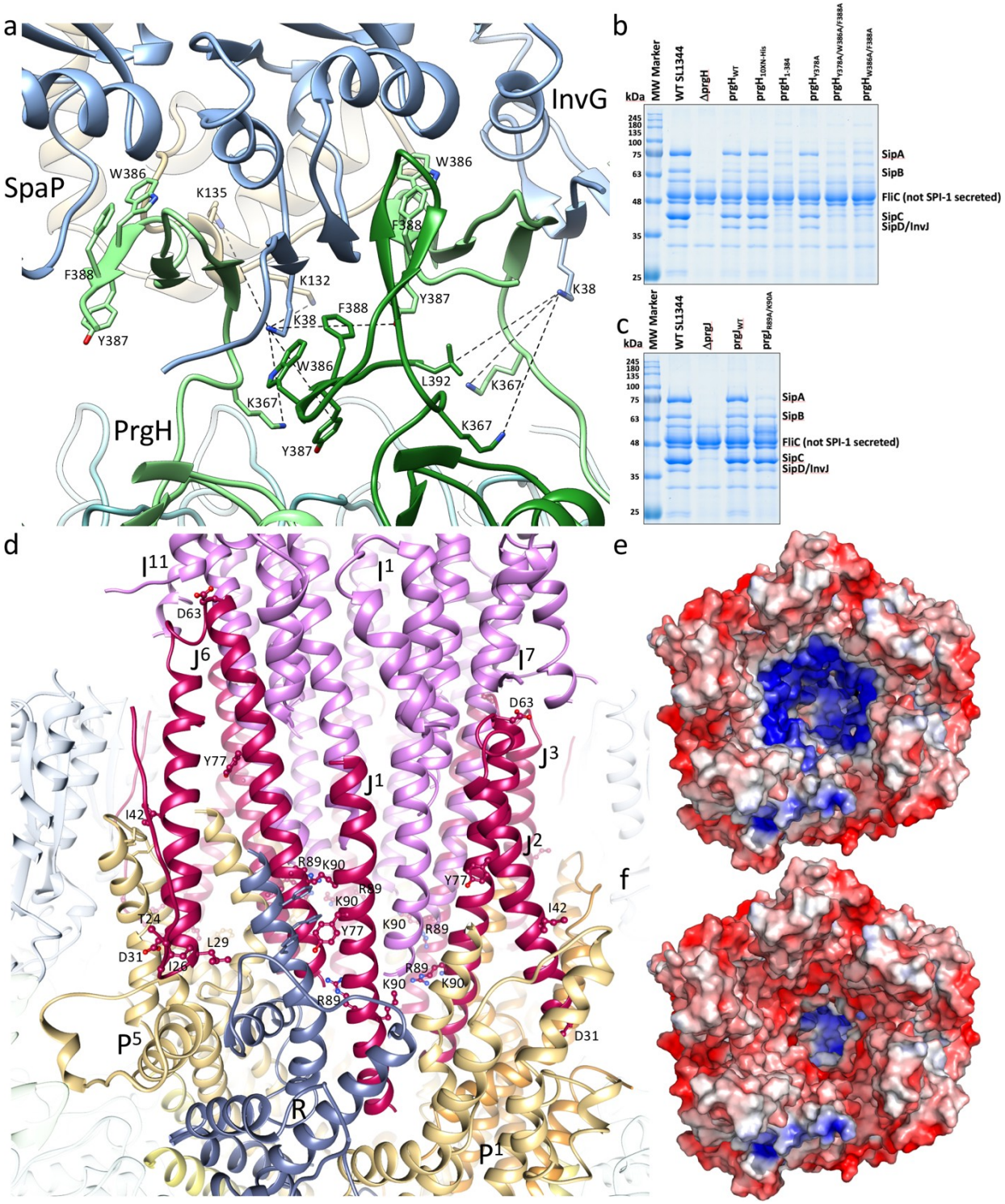

a
InvG
SpaP
PrgH
W386
K135
F388
K132
K38
Y387
L392
K367
b
MW Marker
WT SL1344
ΔprgH
prgH_WT
prgH_1-370N-His
prgH_1-384
prgH_Y379A
prgH_Y379A/W386A/F388A
prgH_W386A/F388A
kDa
245
180
135
100
75
63
48
35
25
SipA
SipB
FliC (not SPI-1 secreted)
SipC
SipD/InvJ
c
MW Marker
WT SL1344
ΔprgJ
prgJ_WT
prgJ_R89A/K90A
d
I11
I1
I7
J6
J3
J1
J2
D63
Y77
I42
R89
K90
T24
L29
D31
I26
P5
R
P1
e
f

**Figure B5: Validation of the interfaces defining the NC assembly. a, The PrgH:InvG:SpaP interface is validated by cross-links between InvG (blue) Lys$^{38}$ and PrgH (green) Lys$^{367}$, Tyr$^{387}$, and Leu$^{392}$; and SpaP (light orange) Lys$^{132}$ and Lys$^{135}$. These cross-links are consistent with both PrgH C-terminal loop conformations observed here (light vs dark green). The intimate hydrophobic interface between InvG and PrgH also explains the secretion and assembly deficient phenotype associated with deletion of the C-terminal 6 residues from PrgH**[142]**. We further probed this interface, demonstrating Trp$^{386}$ and Phe$^{388}$ are critical determinants of the interaction (see b). b-c, Secretion assays for PrgH (b) and PrgJ mutants (c), carried out in a *prgH* or *prgJ* deletion SL1344 strain complemented with plasmids containing the wild-type *prgH* or *prgJ* sequences (prgH$_{WT}$, prgJ$_{WT}$) or the labelled mutants. SDS-PAGE showing secreted effector SipA (late substrate), translocators SipBCD (middle substrates), and InvJ (early substrate). d, The structure of PrgJ initiating needle polymerization atop SpaPR locates mutants shown to abrogate secretion and PrgJ binding (Asp$^{31}$ and Tyr$^{77}$ on the N-terminal and C-terminal helices respectively), result in weakly attached needles (Asp$^{63}$ at the interface with the N-terminus of a PrgI monomer one full helical turn downstream), results in the specific abrogation of translocator secretion (Thr$^{24}$, Ile$^{26}$, Leu$^{29}$ and Ile$^{42}$, at the interface of the N-terminal loop/helix with the SpaP$^{insertion}$), and selectively abolish late effector secretion (Arg$^{89}$ and Lys$^{90}$ on interior surface, see c, e and f). PrgJ mutants shown as ball-and-stick, key interacting residues on SpaR or SpaP shown as stick. We note these *prgJ* mutants were identified using plasmid-complementation**[147,148] **with further experiments required to fully probe the role of PrgJ assembly in assembly and secretion regulation. e-f, Top view of the SpaPQR-PrgJ complex colored according to calculated electrostatic surface for the native structure (e) and the Arg89Ala, Lys90Ala double mutant (f), which abolishes secretion of late substrates (effector SipA) but not early (InvJ) or middle substrates (SipBCD), see c.**

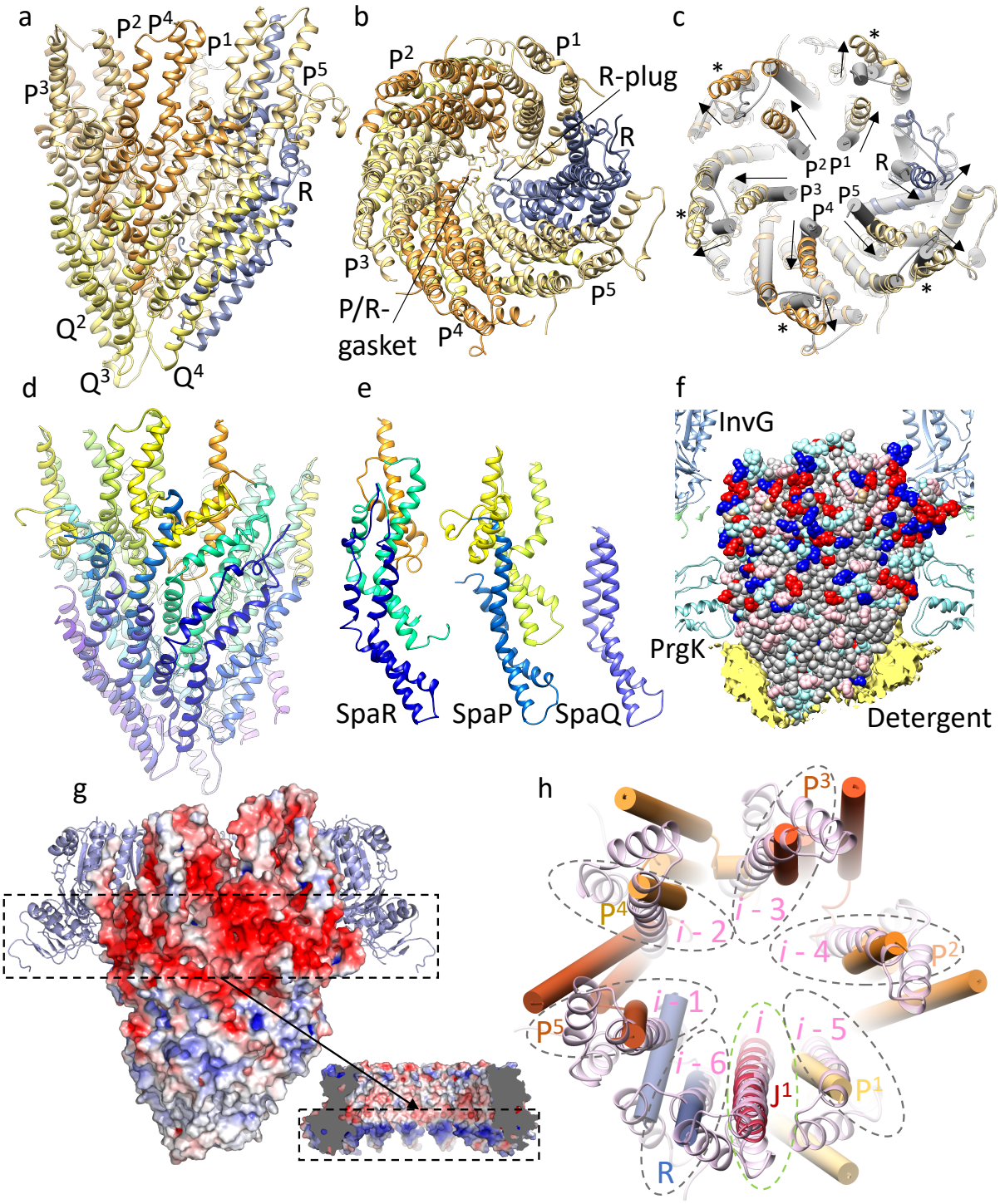

a
P2 P4
P1
P5
P3
R
Q2
Q3
Q4
b
P2
P1
R-plug
R
P3
P/R-
gasket
P4
P5
c
*
P2 P1
R
P3
P5
P4
d
e
SpaR
SpaP
SpaQ
f
InvG
PrgK
Detergent
g
h
P3
P4
i - 3
i - 2
i - 4
P2
i - 1
P5
i
i - 5
i - 6
J1
P1
R

**Figure B6: SpaPQR complex and bound PrgIJ. a, Side view of atomic model of the SpaPQR complex, colored as in Fig. 3a. b, top view of a. c, Comparison of SpaPQR in the presence (cartoon) and absence (tubes) of bound PrgJ and PrgI. PrgJ binding is coupled with significant change to the N-terminal helices of SpaP and SpaR, and the SpaPinsertion (*) indicated with arrows. d, The individual repeating helix-turn-helix motifs that make up SpaPQR and generate the pseudo-helical assembly are individually colored from blue-green-orange (SpaPinsertion colored yellow). The kinked nature of the repeats narrows the assembly towards the IM (bottom). e, Individual repeat motifs in SpaP, SpaQ, and SpaR monomers. f, Space filling representation of SpaPQR complex (positive charge: blue, negative charge: red, hydrophobic: grey, aromatic: pink, polar: cyan) and its surrounding detergent (map in yellow), PrgK (cyan), and InvGN0N1 (light blue). g, Electrostatic surface of SpaPQR-PrgJ complex. Surrounding InvGN0N1 (blue) shown and corresponding, complimentary electrostatic surface of the interacting N0 domain boxed. h, The pseudo-hexameric SpaR and SpaP N-terminal helices, which bind PrgJ, superpose with the 6 subunit turn of the PrgI needle (labelled i – 1 to i – 6 and circled, the green circle corresponds to the location of PrgJ1 or superposed PrgIi). The SpaPR scaffold thus mimics the homotypic polymerization interfaces of the PrgI needle and explicitly defines its helical parameters.**

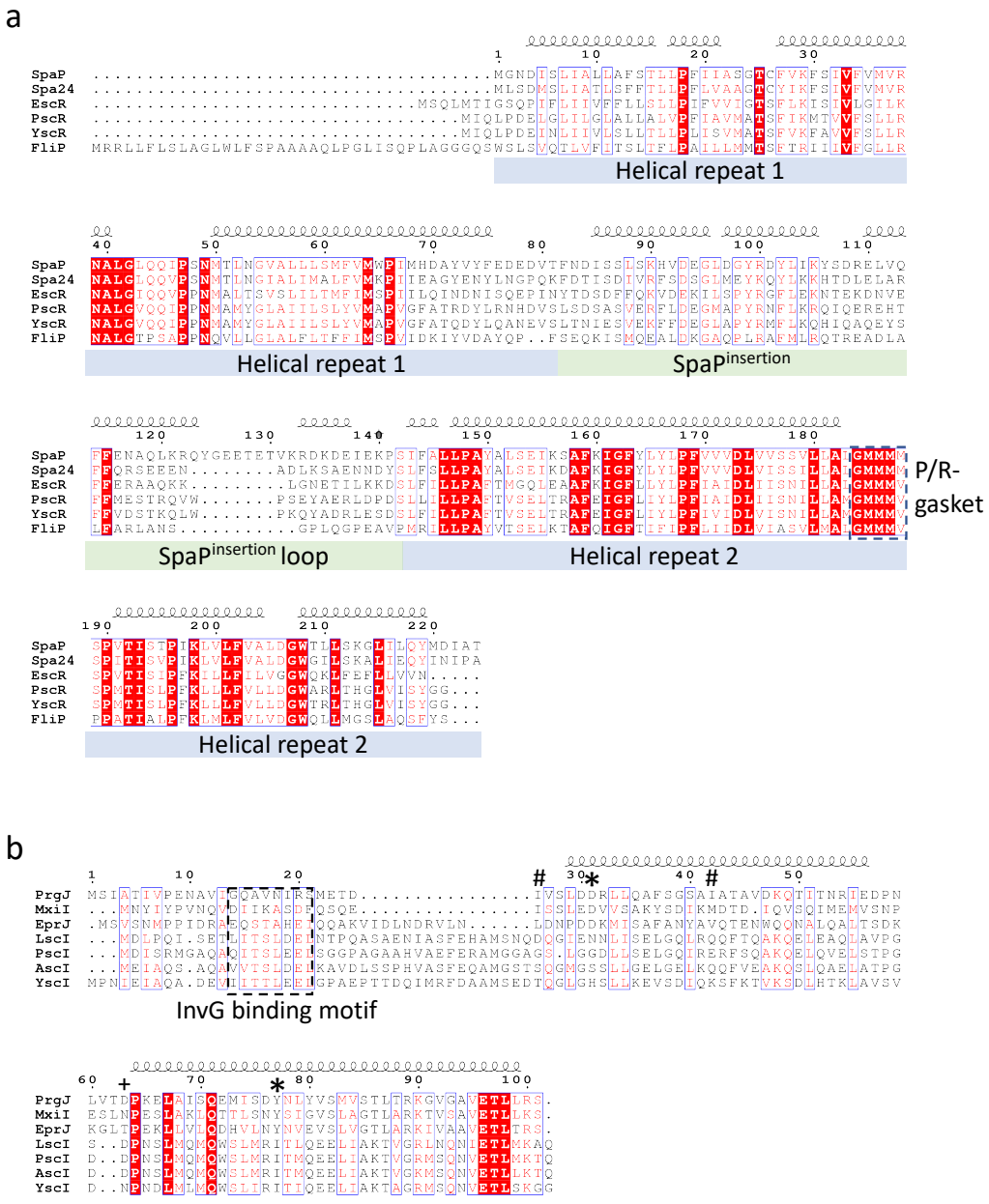
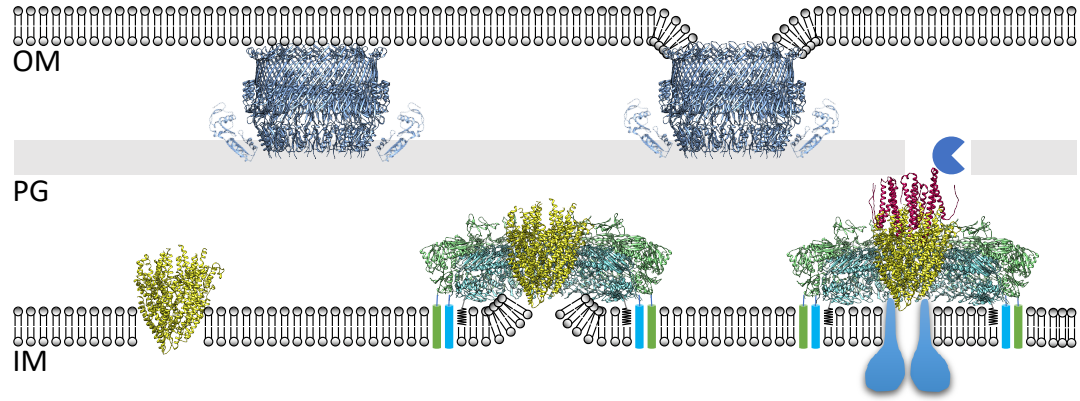
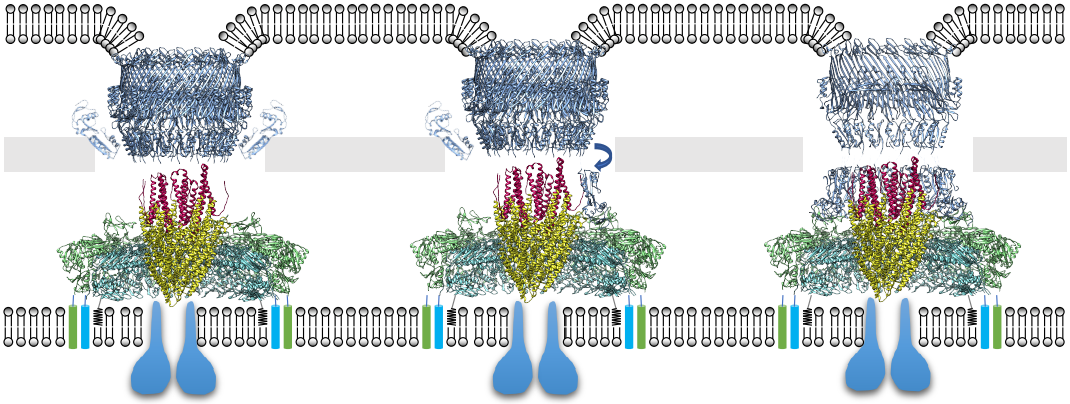



**Figure B7: Sequence alignments of SpaP and PrgJ. Sequence alignment of SpaP homologues. Secondary structure from SpaP annotated above. Spa24, *S. flexneri*; EscR, EPEC; PscR, *P. aeruginosa*; YscR, *Y. pestis*, FliP, *S.* Typhimurium. The position of the helical repeats and SpaP$^{insertion}$ indicated. b, Sequence alignment of**

**PrgJ homologues. PrgJ secondary structure annotated above. MxiI, *S. flexneri*; EprJ, EHEC; LscI, *P. luminescens*; PscI, *P. aeruginosa*; AscI, *A. salmonicida*; YscI, *Y. pestis*. Key mutants labelled: Asp$^{31}$ and Tyr$^{77}$ which abolish PrgJ function (*)**[147,148]**; Thr$^{24}$, Ile$^{26}$, Leu$^{29}$, Ile$^{42}$, which selectively abolish middle substrate (translocases) secretion (#)**[148]**; and Arg$^{89}$ and Lys$^{90}$, which selectively abolish late substrate (effectors) secretion (^), see Fig B6c-e). Blue boxes indicate global similarity determined according to physio-chemical property with a score of 0.7.**


OM
PG
IM



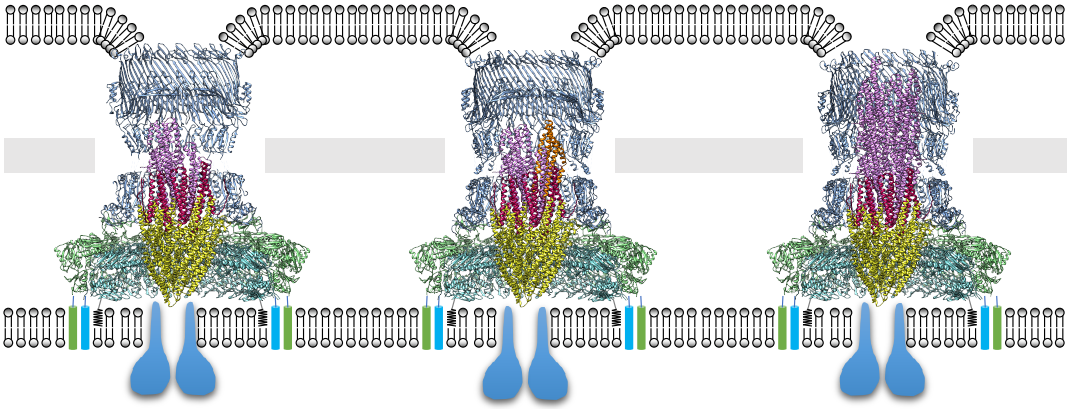


**Figure B8: Assembly of the needle complex. Different assembly pathways have been proposed for the NC with an inside-out model in the *Salmonella* SPI-1 system**[72]**, where assembly starts with the IM export apparatus, and an outside-in model alternatively suggested in *Yersinia***[82]**, where assembly is dependent instead on the OM secretin YscC (InvG) and interaction with PrgH orthologue YscD. This possibly reflects differences in the architecture of the two systems: the *in situ* structure of the *Yersinia* injectisome**[162] **appears elongated compared to *Salmonella***[18] **and *Shigella***[92]**, with YscD correspondingly showing an extended structure with a predicted less intimate/stable interface with YscJ compared to PrgHK**[162]**. Together with published data, our novel structures are more consistent with the inside-out pathway and we propose the following model for NC assembly. Assembly is initiated at the IM with formation of the SpaPQR(S) complex (a) and at the OM by the 15-mer InvG$^{secretin}$ pore (b). Export gate InvA (blue) is recruited and the PrgHK rings assemble around the SpaPQR(S) complex, facilitating its extraction from the IM (c) resulting in the membrane deformation observed in *in situ* cryo-ET**[18]**. The OM secretin spontaneously inserts into the OM (d). Assembly of the cytoplasmic components results in a T3S competent apparatus, secretion and assembly of PrgJ serves to recruit a T3S lytic transglycosylate to locally clear peptidoglycan (e). The OM InvG$^{secretin}$ localizes to the IM complex (f) and coupling of the InvG$^{N0N1}$ domains to PrgH is initiated via interaction with SpaP$^{5}$ (g). InvG$^{N0N1}$ coupling proceeds and is terminated by recruitment of a 16$^{th}$ Invg$^{N0N1}$ domain to complete the symmetric interface with PrgH (h). PrgI secretion and assembly atop PrgJ causes the allosteric gating changes in the InvG$^{N0N1}$ to unlock the gate (i and j) with subsequent steric gate opening via continued PrgI polymerization (k).**

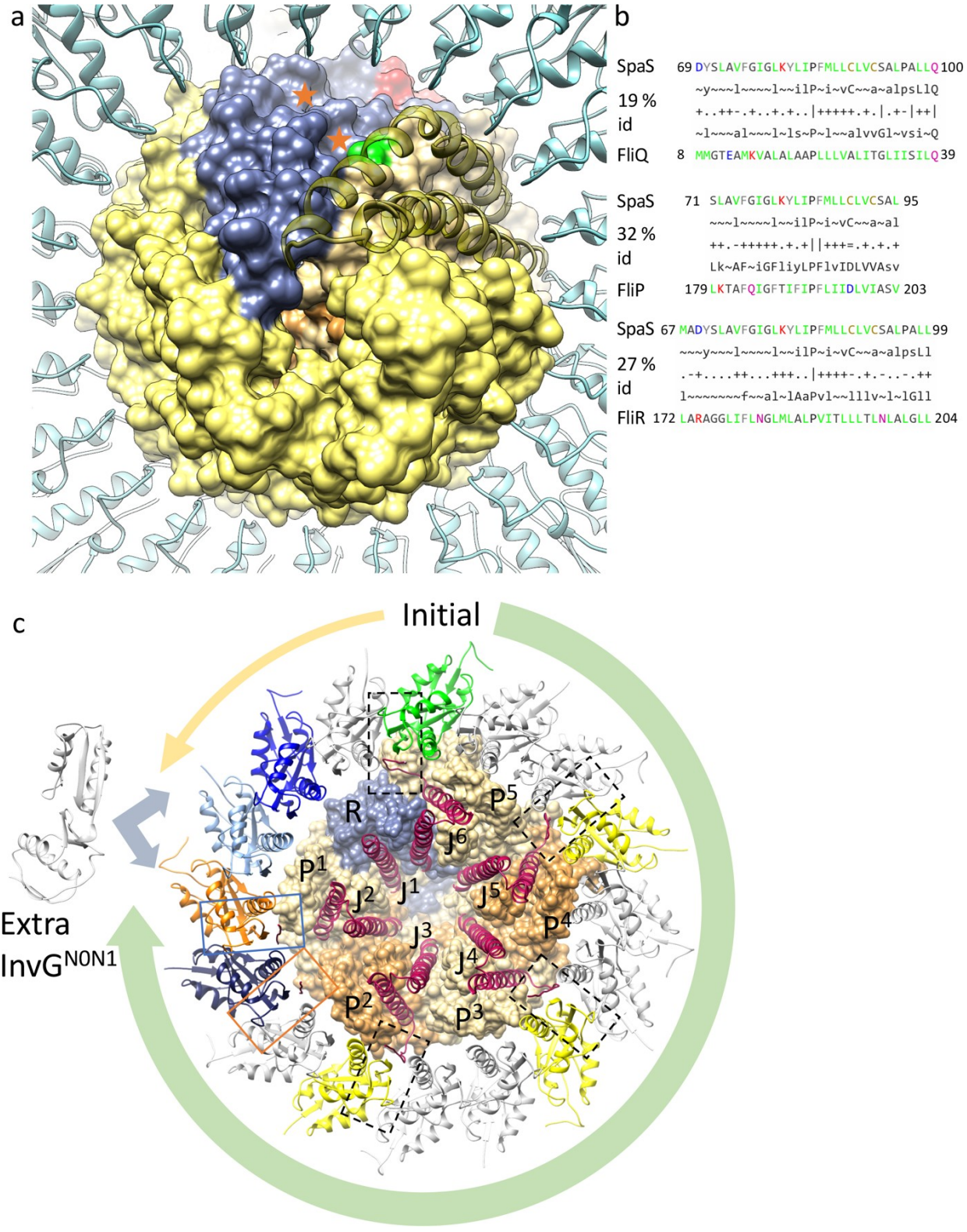
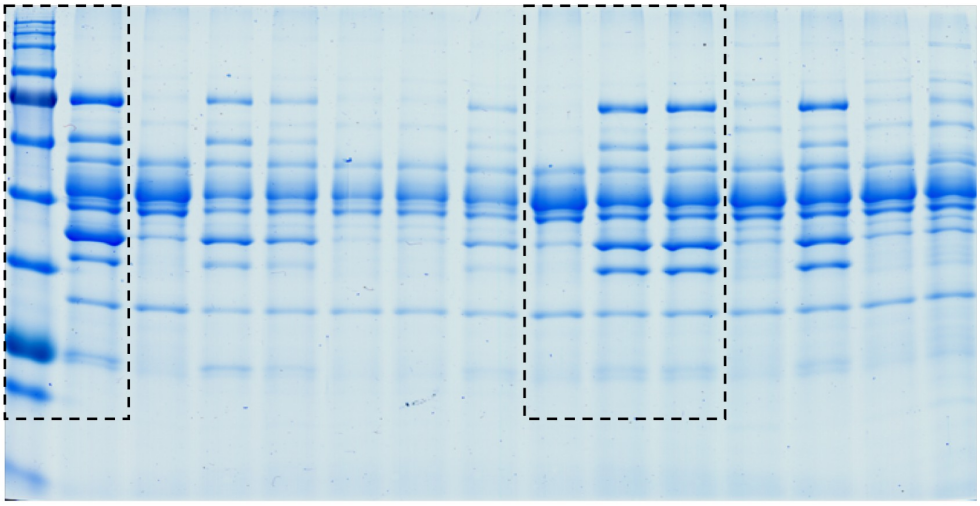

a
b
SpaS 69 DYSLAVFGIGLKYLIPFMLLCLVCSALPALLQ 100
19 % id
FliQ 8 MMGTEAMKVALALAAPLLLVALITGLIISILQ 39
SpaS 71 SLAVFGIGLKYLIPFMLLCLVCSAL 95
32 % id
FliP 179 LKTAFQIGFTIFIPFLIIDLVIASV 203
SpaS 67 MADYSLAVFGIGLKYLIPFMLLCLVCSALPALL 99
27 % id
FliR 172 LARAGGLIFLNGLMLALPVITLLLTLNLALGLL 204
c
Initial
Extra
InvG^NON1
R
P1
P2
P3
P4
P5
J1
J2
J3
J4
J5
J6

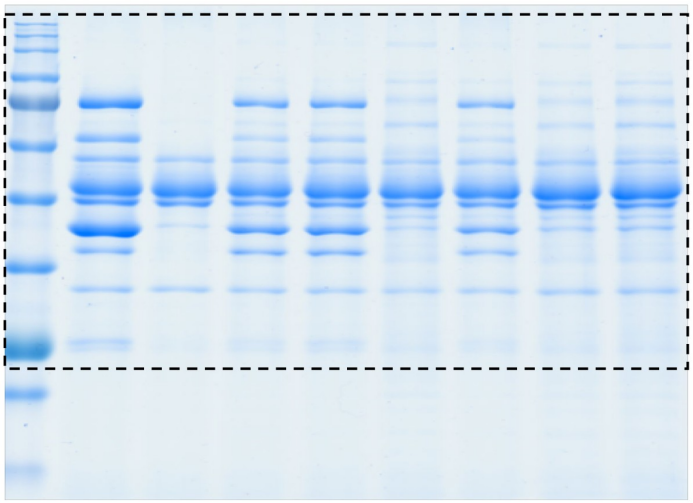
**Figure B9: Composition of the SpaPQR complex and coupling of the IM/OM rings. a, Space sufficient for additional components binding to the SpaPQR complex is observed around the exterior surface of SpaR (blue) and SpaP$^{1}$ (light orange). This is consistent with a FliP$_5$Q$_5$R$_1$ species observed using native mass spectrometry**[21]**, with the additional SpaQ (transparent yellow ribbon) binding the exterior of SpaP$^{1}$ as the other SpaP:SpaQ interactions. The peripheral location and reduced interaction with PrgK would explain its ready dissociation. We also propose that the exterior surface of SpaR (orange stars) represents the binding site for SpaS, with cross-linking**[73] **locating it near the C-terminal helix of SpaP (green) and gene-fusion**[83] **experiments showing it can be linked to the C-terminus of SpaR (red). This would locate the SpaS C-terminal domain, implicated in regulation of secretion hierarchy, near the cytoplasmic facing entrance to the SpaPQR entry portal to the PrgIJ needle channel. b, Similarity of the SpaS predicted N-terminal transmembrane region to the helical repeat motifs of FliP, FliQ, and FliR, supporting a similar helical repeat structure. c, We propose coupling of the 15-mer OM InvG$^{secretin}$ pore and 24-mer IM PrgH is initiated by interaction of an InvG$^{N0N1}$ domain (green) with SpaP$^{5}$, the most extensive interface observed here. Subsequent assembly proceeds primarily in a clockwise direction (viewed from OM perspective; green arrow) driven by interaction with SpaP, PrgH, and PrgJ, with secondary counter-clockwise assembly (orange arrow). Accordingly, incorporation of the 15 InvG$^{N0N1}$ domains belonging to the InvG$^{secretin}$ pore with the observed PrgH$_3$:InvG$_2$ symmetrical interface results in a vacant position around SpaP$^{1}$. To complete IM/OM ring coupling, an extra InvG$^{N0N1}$ domain is recruited (blue arrow), and we observe this here in neighboring, alternate locations (light blue or orange InvG$^{N0N1}$, colored as Fig. 2). The interaction of the PrgJ N-terminal loop with InvG$^{N0N1}$ is boxed; PrgJ$^{3-6}$ interact with evenly distributed (every third) InvG$^{N0N1}$ domains (green/yellow). PrgJ$^{2}$, by contrast, is observed to bind in alternate locations: the expected, evenly spaced InvG$^{N0N1}$ domain (orange) and its neighbor (dark blue). We interpret this to suggest that PrgJ is secreted and assembled prior to IM/OM coupling, with the PrgJ N-terminal loops binding to the InvG$^{N0N1}$ domains of the InvG$^{secretin}$ pore on coupling, prior to recruitment of the extra 16$^{th}$ InvG$^{N0N1}$ domain to complete the coupling interface: when the light blue InvG$^{N0N1}$ is the extra 16$^{th}$ recruited domain, PrgJ$^{2}$ binds its preferred site (blue box), however, when this**

**preferred InvG$^{N0N1}$ is instead the extra 16$^{th}$ (see Fig. 2), PrgJ$^{2}$ instead binds the immediate neighbor (orange box).**

a



b



c

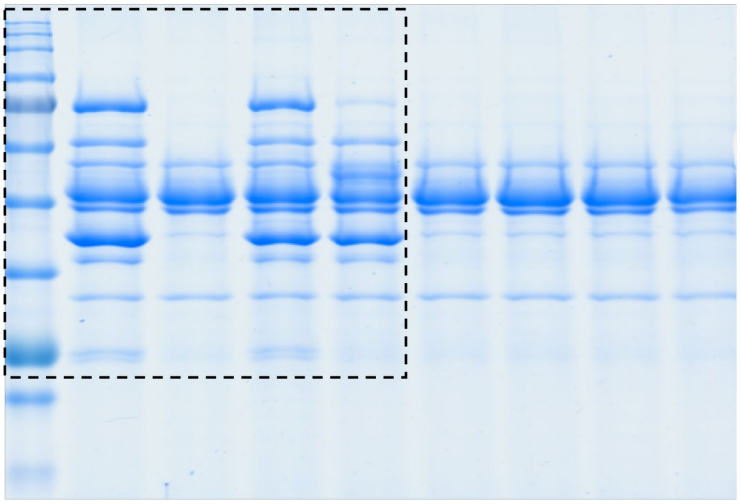


**Figure B10: Full length blots for gels shown in Figure B1b (a), Figure B5b (b) and c (c).**

| | #1 needle complex base (1) C1map (EMDB-8915) (PDB 6DV6) | #2 needle complex base partial unlock (EMDB-8915) (PDB 6DV6) | #3 needle complex base full unlock (EMDB-8914) (PDB 6DV3) | #4 needle complex base open (EMDB-8913) (PDB 6DUZ) | #5 InvG$^{secretin}$ (EMDB-8924) (PDB 6DWB) | #6 Inner complex (EMDB-8924) (PDB 6DWB) | #7 Inner complex with rod (EMDB-8924) (PDB 6DWB) |
|---|---|---|---|---|---|---|---|
| **Data collection and processing** | | | | | | | |
| Magnification | 29240 | 29240 | 29240 | 29240 | 29240 | 29240 | 29240 |
| Voltage (kV) | 300 | 300 | 300 | 300 | 300 | 300 | 300 |
| Electron exposure (e–/Å$^2$) | 15 | 15 | 15 | 15 | 15 | 15 | 15 |
| Defocus range (μm) | 1.5-3 | 1.5-3 | 1.5-3 | 1.5-3 | 1.5-3 | 1.5-3 | 1.5-3 |
| Pixel size (Å) | 1.71 | 1.71 | 1.71 | 1.71 | 1.71 | 1.71 | 1.71 |
| Symmetry imposed | C1 | C1 | C1 | C1 | C15 | C1 | C1 |
| Initial particle images (no.) | 235618 | 235618 | 235618 | 235618 | 235618 | 235618 | 235618 |
| Final particle images (no.) | 51816 | 5319 | 7249 | 5018 | 185855 | 144097 | 25666 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Map resolution (Å)<br>FSC threshold | 0.143 | 0.143 | 0.143 | 0.143 | 0.143 | 0.143 | 0.143 |
| | | | | | | | |
| **Refinement** | | | | | | | |
| Initial model used (PDB code) | | 5TCQ | 5TCQ | 5TCP | | | |
| Model resolution range (Å) | 7.4 | 4.1 | 3.9 | 3.6 | 3.3 | 3.3 | 3.3 |
| Map sharpening *B* factor (Å$^2$) | -178 | -141 | -144 | -118 | -145 | -145 | -145 |
| Model composition<br>Non-hydrogen atoms<br>Protein residues<br>Ligands | | 57975<br>7500 | 40950<br>5372 | 72768<br>9072 | 18360<br>2340 | 18360<br>2340 | 18360<br>2340 |
| *B* factors (Å$^2$)<br>Protein<br>Ligand | | 9.9 | 13.87 | 11 | 19.6 | 19.6 | 19.6 |
| R.m.s. deviations | | | | | | | |
| Bond lengths (Å) | | 0.009 | 0.010 | 0.007 | 0.016 | 0.016 | 0.016 |
| | | 1.37 | 1.49 | 1.25 | 1.3 | 1.3 | 1.3 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Bond angles (°) | | | | | | | |
| Validation | | | | | | | |
| MolProbity score | | 2.4 | 2.54 | 1.92 | 2.7 | 2.7 | 2.7 |
| Clashscore | | 8.63 | 8.13 | 4.63 | 6.12 | 6.12 | 6.12 |
| Poor rotamers (%) | | 8.59 | 7.64 | 4.3 | 0 | 0 | 0 |
| Ramachandran plot | | | | | | | |
| Favored (%) | | 5.24 | 6.5 | 3.18 | 3.95 | 3.95 | 3.95 |
| Allowed (%) | | 94.76 | 93.5 | 96.82 | 96.05 | 96.05 | 96.05 |
| Disallowed (%) | | 0 | 0 | 0 | 0 | 0 | 0 |

**Table B1. Cryo-EM data collection, refinement and validation statistics.**