Structural Studies of the Type Three Secretion System Export Apparatus

Lucas Kuhlen

Sir William Dunn School of Pathology

Department of Chemistry

Christ Church

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Trinity term, 2019

Structural Studies of the Type Three Secretion System Export Apparatus

Lucas Kuhlen, Christ Church (Sir William Dunn School of Pathology and Department of Chemistry). Thesis submitted for the degree of Doctor of Philosophy at the University of Oxford, Trinity term 2019.

Abstract

One of the most widespread bacterial secretion systems is the type three secretion system (T3SS), which is critical for the virulence of many pathogens and responsible for assembly of the bacterial flagellum. The virulence-associated T3SS is a molecular syringe that uses a highly conserved export apparatus (EA) to secrete effector proteins through a needle directly into the host cell cytoplasm. The flagellar T3SS secretes flagellar subunits through the EA which then self-assemble into a filament. Despite recent progress in the structural biology of T3SS, the highly conserved membrane domains of the five inner membrane proteins of the EA have evaded structural characterisation.

Here, multiple cryo-EM structures of complexes of the three core EA proteins from both flagella (FliPQR) and injectisomes (SctRST) are presented, revealing a pseudohexameric, helical complex. Remarkably, the complex is made up of a series of long, highly kinked helical hairpins rather than the predicted canonical transmembrane helices. Many of the hydrophobic residues predicted to lie in transmembrane helices are instead buried at the interfaces between subunits, leaving only a small membrane domain. In the assembled T3SS this complex is encased in the basal body, above the plane of the inner membrane, and connects to the helical filament, whose helical parameters it matches. The membrane domain of the fourth export apparatus protein, FlhB, is shown to form a complex with FliPQR, and the complex structure suggests a mechanism of how interactions between FlhB and FliPQR could lead to opening of the complex. Finally, stable rings of full-length FlhA/SctV, the last EA protein, were prepared and cryo-EM confirmed the structure of the nonameric cytoplasmic domain. While little high-resolution information of the membrane domain was obtained, these studies provide a framework for future studies.

Together these results reveal the structure of part of the secretion pathway and suggest that the core of the export apparatus templates the assembly of the helical filament.

Acknowledgements

First I would like to thank my supervisor Susan Lea for the great projects that I had the privilege to work on in her lab. It was great to experience how the resolution revolution in cryo-EM is changing all of structural biology and I especially liked applying these exciting new tools to study the unique proteins of the export apparatus. I also want to thank my second supervisor, Carol Robinson, for the wonderful opportunity to learn about native mass spectrometry of membrane proteins.

I am grateful to Patrizia Abrusci for showing me how everything in the lab works when I joined the lab. This project wouldn't have been possible without her hard work characterising many export apparatus constructs and I want to thank her for her efforts in teaching me how to purify membrane proteins.

I also want to thank Joe Gault, who was a great supervisor in my rotation project and taught me native mass spectrometry. Many thanks to Steven Johnson for all the help with cryo-EM data analysis, illuminating discussions, MALS and thanks to Susan and Steve for reviewing the chapters of this thesis. Thanks also to Justin Deme and COSMIC for help with cryo-EM data collection.

Thanks to the other students in the lab who have become my friends. Thank you to my mum for igniting my passion for host-pathogen interactions and to Sharon for listening to stories about T3SS and cryo-EM.

Finally, thanks to the Wellcome Trust for generous financial support.

Publications resulting from this study

Structure of the core of the type III secretion system export apparatus

Kuhlen L[a], Abrusci P[a], Johnson S[a], Gault J, Deme J, Caesar J, Dietsche T, Mebrhatu MT, Ganief T, Macek B, Wagner S, Robinson CV, Lea SM (2018). Nat Struct Mol Biol **25**: 583-590.

The Structure of an Injectisome Export Gate Demonstrates Conservation of Architecture in the Core Export Gate between Flagellar and Virulence Type III Secretion Systems

Johnson S[b], Kuhlen L[b], Deme JC, Abrusci P, Lea SM, (2019) MBio **10**: e00818-19.

Preprint:

The flagellar substrate specificity switch protein FlhB assembles onto the extra-membrane export gate to regulate type three secretion

Kuhlen L, Johnson S, Zeitler A, Bäurle S, Deme JC, Debo R, Fisher J, Wagner S, Lea SM, (2019), bioRxiv 686782; doi: https://doi.org/10.1101/686782

Author contributions

Structure of the core of the type III secretion system export apparatus

Kuhlen L[a], Abrusci P[a], Johnson S[a], Gault J, Deme J, Caesar J, Dietsche T, Mebrhatu MT, Ganief T, Macek B, Wagner S, Robinson CV, Lea SM

[a]These authors contributed equally to this work

L.K. performed experiments: strain and plasmid construction, complex purification, nMS, cryo-EM grid optimization, cryo-EM data analysis and model building and analysis. P.A. performed experiments: strain and plasmid construction, complex purification, initial cryo-EM grid optimization and cryo-EM data analysis. S.J. performed experiments and designed and supervised experiments for characterization of protein complexes, cryo-EM data analysis, structure determination and analysis, and wrote the first draft of manuscript and figures with S.M.L.

The Structure of an Injectisome Export Gate Demonstrates Conservation of Architecture in the Core Export Gate between Flagellar and Virulence Type III Secretion Systems

Johnson S[b], Kuhlen L[b], Deme JC, Abrusci P, Lea SM

[b]These authors contributed equally to this work

S.J. designed, supervised and performed experiments, did characterization of protein complexes, cryo-EM data analysis, structure determination and analysis, and wrote the manuscript with S.M.L. L.K. performed experiments and did strain and plasmid construction, complex purification, native mass spectrometry, cryo-EM grid optimization, cryo-EM data analysis, and model building and analysis.

Abbreviations

APBS	Adaptive Poisson-Boltzmann solver
ATP	Adenosine triphosphate
CID	Collision induced dissociation
CMC	Critical micelle concentration
CTF	Contrast transfer function
CV	Column volume
DNA	Deoxyribonucleic acid
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen (German Collection of Microorganisms and Cell Cultures)
DTT	Dithiothreitol
EA	Export apparatus
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
EMDB	Electron Microscopy Data Bank
FLP	Flippase
FRT	FLP recombination target
fT3SS	Flagellar type 3 secretion system
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IM	Inner membrane
MALS	Multi angle light scattering
MSA	Multiple sequence alignment
MWCO	Molecular weight cut off
NA	Numberical aperture
nMS	Native mass spectrometry
OD600	Optical density at 600 nm
OM	Outer membrane
PCR	Polymerase chain reaction
PDB	Protein data bank
pmf	Proton motive force
PR	Pathogenesis-related

RBS	Ribosome binding site
rpm	Revolutions per minute
SCV	Salmonella-containing vacuole
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SPI-1	Salmonella pathogenicity island 1
SPI-2	Salmonella pathogenicity island 2
T1SS	Type three secretion system
T2SS	Type three secretion system
T3SS	Type three secretion system
T4SS	Type four secretion system
T5SS	Type five secretion system
T6SS	Type six secretion system
T7SS	Type seven secretion system
T8SS	Type eight secretion system
T9SS	Type nine secretion system
UV	Ultraviolet
vT3SS	Virulence associated type 3 secretion system

Table of contents

List of Figures

List of Tables

1 Introduction

1.1 Bacterial protein secretion in pathogenesis

Cells are bounded by membranes made up of lipid bilayers and protein (Gorter and Grendel 1925; Singer and Nicolson 1972). In order to affect their surroundings, cells must move molecules across the membrane or insert them into the membrane. While some molecules may diffuse across the membrane passively, the most diverse class of biological molecules, polypeptides, generally require dedicated mechanisms to cross the membrane. In bacteria, a large array of secretion systems has evolved which allow secretion of proteinaceous effectors out of the cytoplasm (Green and Mecsas 2016). These secretion systems have been weaponised in the molecular warfare among bacteria and between bacterial pathogens and the immune systems of eukaryotic hosts (Lee and Schneewind 2001) making bacterial secretion a medically relevant phenomenon. In addition to acting alone, for example as toxins, some secreted proteins are assembled into large extracellular appendages such as pili or flagella (Fronzes et al. 2008), which can also be tools for virulence. Most pathogens adhere to host tissue, secrete toxins that disable host immunity or invade cells to hide from the immune system and finally exit the host again in large numbers. All of these can be mediated by secretion systems (Lee and Schneewind 2001).

Bacterial secretion systems can be classified into one step and two step systems to distinguish those systems that rely on the Sec and Tat pathways to export proteins to the periplasm in a first step followed by a second step of secretion across the outer membrane, from the nanomachines that secrete across both membranes in a single step (Green and Mecsas 2016). Type 2, 5, 8 and 9 secretion systems (Figure 1.1) are two step secretion systems, all of which have also been linked to the formation of extracellular structures (Barnhart and Chapman 2006; Korotkov et al. 2012; Leyton et al. 2012; Lasica et al. 2017). In the first step proteins are exported into the periplasm through the Sec translocon or the Tat pathway. Sec secretion substrates are recognised by either SRP (co-translationally) or SecA (co- or post-translationally) and delivered to the SecYEG translocon, which uses energy from ATP hydrolysis and the proton motive force (pmf) to secrete unfolded polypeptides into the periplasm or insert them into the membrane in the case of integral membrane proteins (Papanikou et al. 2007). Chaperones and other

co-factors are associated with the translocon (Sachelaru et al. 2014). The Sec independent Tat pathway (Sargent et al. 1998; Weiner et al. 1998) however secretes folded proteins using energy from the pmf (Palmer and Berks 2012). Proteins are thought to be recognised by the TatBC complex and the membrane is permeabilised through an unknown mechanism by TatA (Bolhuis et al. 2001; Rollauer et al. 2012). In Gram-positive bacteria without an outer membrane, these general secretion systems are sufficient for secretion into the extracellular space.

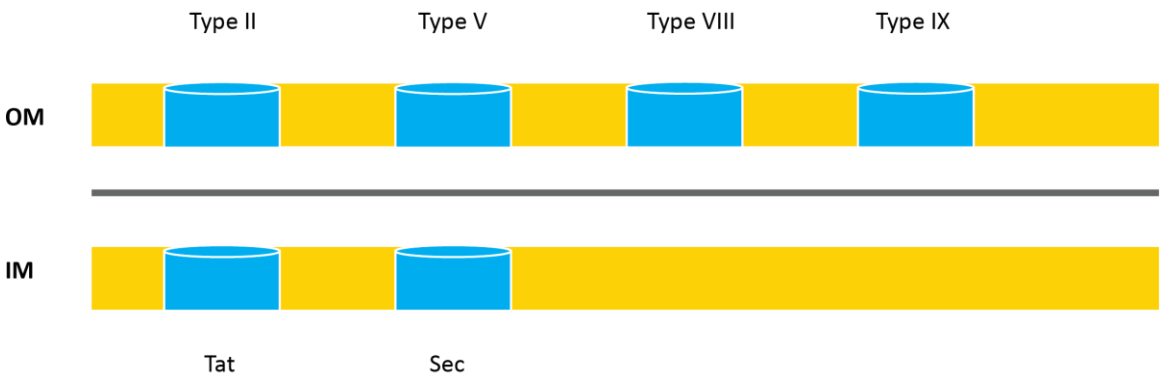


Figure 1.1 Two step secretion systems.
Membrane pores in two step secretion systems.

The type 2 secretion system (T2SS) consists of a secretin in the outer membrane through which folded proteins are pushed by a pseudopilus according to the piston model (Korotkov and Sandkvist 2019). The secretin is a giant 15-fold symmetric beta barrel pore (Majewski et al. 2018). A well-known T2SS substrate is the cholera toxin, which is an AB toxin whose active subunit causes secretion of chloride ions by epithelial cells in the intestine, resulting in watery diarrhoea and bacterial shedding (Sikora 2013). A related secretion machinery is used for the assembly of type IV pili, which are responsible for twitching motility and have been implicated in the pathogenesis of *Pseudomonas aeruginosa* (Mattick 2002).

Type 5 secretion differs from the other secretion systems in that there is no dedicated secretion machinery. Instead, most substrates are autotransporters, consisting of a transport domain that forms the outer membrane beta barrel pore (Oomen et al. 2004) through which the passenger domain is secreted (Leyton et al. 2012). Alternatively, in

some cases a separate protein can form the pore. An example is the *Yersinia* autotransporter YadA, an essential adhesin possessing a collagen-binding head domain and a Factor H binding stalk domain, which confers resistance to complement mediated immunity (Muhlenkamp et al. 2015). A conceptually similar system is the chaperone-usher pathway, in which an outer membrane pore, the usher, and a chaperone are responsible for secretion and assembly of pilus subunits. This type of pilus is of particular importance in uropathogenic *Escherichia coli* infections (Waksman and Hultgren 2009).

The type 8 secretion system pore is a nonameric beta barrel through which unfolded proteins can pass to assemble into bacterial amyloid (Goyal et al. 2014) and the type 9 secretion system pore is a large monomeric beta barrel through which folded substrates are secreted (Lauber et al. 2018). Type 9 secretion is also associated with gliding motility.

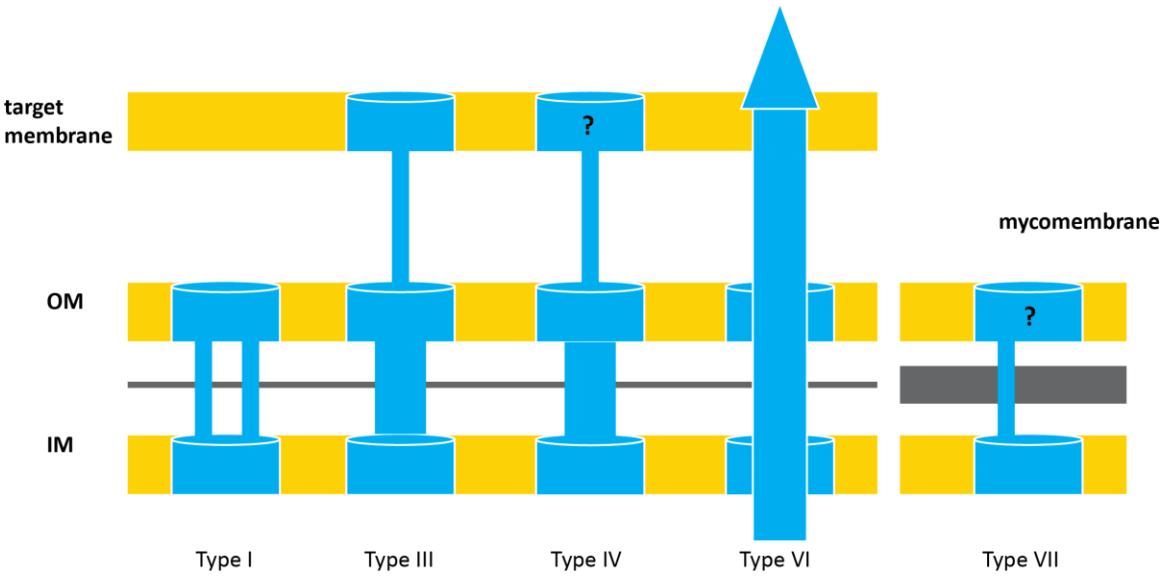


Figure 1.2 One step secretion systems.
Membrane pores in one step secretion systems. The presence of a pore in the mycomembrane in type 7 secretion is not known.

One step secretion systems include type 1, 3, 4, 6 and 7 secretion systems (Figure 1.2) and they secrete substrates across the entire cell envelope in a single step (Thomas et al. 2014; Ates et al. 2016; Galan and Waksman 2018). Type 1 secretion is a common mechanism of secretion using an ABC transporter in the inner membrane (Lin et al. 2015; Morgan et al. 2017) linked to an outer membrane protein such as TolC (Koronakis et al.

2000). Secretion substrates are secreted in a largely unfolded state. The most well characterised substrate is the haemolysin HlyA from uropathogenic *E. coli* which is thought to form pores in host cell membranes (Wiles and Mulvey 2013).

Type 4 secretion systems (T4SS) are ATP hydrolysis powered nanomachines that can secrete proteins and DNA (Chandran Darbari and Waksman 2015). They are found in animal and plant pathogens as well as bacterial conjugative DNA transfer mechanisms. The T4SS consists of a large, envelope embedded complex (Low et al. 2014) and often includes a pilus, such as the T-pilus of *Agrobacterium tumefaciens* through which DNA is injected into host plant cells (Babic et al. 2008). However, T4SS dependent secretion of periplasmic proteins, i.e. two step secretion, has been described for some human pathogens. *Bordetella pertussis* secretes the pertussis toxin, an AB toxin, first through the Sec pathway into the periplasm and then through a T4SS without a pilus (Wallden et al. 2010). It is not known if the large AB toxin is secreted through the small outer membrane pore of the T4SS.

The type 6 secretion system is a molecular “spear gun” in which an effector protein at the tip of the molecular spear is propelled into another cell (Bock et al. 2017). The membranes of the secreting cell are protected by guiding secretion through the membrane core complex whose periplasmic structure is known (Yin et al. 2019), but the inner and outer membrane domains of the membrane core complex have not been solved to high resolution yet. Only the inner membrane pore of the type 7 secretion system of Gram-positive bacteria has been characterised (Beckham et al. 2017). It is not known how substrates pass the mycomembrane in *Mycobacteria*.

1.2 Type three secretion

1.2.1 Discovery of type three secretion

Type 3 secretion (Salmond and Reeves 1993; Deng et al. 2017) is a very widespread mechanism of bacterial protein secretion found in rhizobial symbionts and many pathogens including *Yersinia pestis*, the causative agent of plague (Kitasato 1894; Yersin 1894). The three biovars of *Yersinia pestis* (antiqua, mediaevalis and orientalis) (Devignat

1951) are responsible for the Plague of Justinian (antiqua) (Wagner et al. 2014), the Black Death (mediaevalis) (Bos et al. 2011) and the modern pandemic (Parkhill et al. 2001). Protein secretion was discovered in *Y. pestis* due to growth restriction in the absence of calcium ions at 37 °C (the low calcium response) (Kupferberg and Higuchi 1958). Upon withdrawal of calcium, *Yersinia* secretes a large quantity of plasmid encoded Yop proteins (Michiels et al. 1990) and this is required for virulence (Goguen et al. 1984). However, external administration of Yop proteins to human cells has little effect. Instead, microinjection of Yop proteins revealed that their targets are intracellular (Rosqvist et al. 1991) and Yop translocation by *Yersinia* into host cells is dependent on contact with the target cell (Rosqvist et al. 1994). In addition, Yop proteins do not have a traditional signal sequence, but are recognised by an N-terminal signal (Michiels and Cornelis 1991; Anderson and Schneewind 1997) which is sufficient for protein secretion. A second mechanism of substrate recognition is through specific chaperones that bind their cognate effector protein in the cytoplasm prior to secretion (Wattiau and Cornelis 1993).

As gene sequences became available, a high degree of sequence conservation was observed between type 3 secretion proteins and certain flagellar proteins (Aizawa 2001), especially those thought to be responsible for transport across the inner membrane, for example SctV/FlhA (Plano et al. 1991; Ramakrishnan et al. 1991). Structural analyses revealed further homology even in the absence of identifiable sequence conservation: A needle complex corresponding to the type 3 secretion system (T3SS) nanomachine was isolated from *Salmonella* cells and shown to resemble a syringe consisting of a series of rings embedded in the inner and outer membranes to which a thin needle is attached (Kubori et al. 1998), suggesting that the T3SS is a molecular syringe that injects toxic effector proteins into host cells, which is why it is also known as the injectisome. Cryo-EM later demonstrated that secretion substrates travel through the needle in a largely unfolded state, but alpha helical secondary structure may be maintained (Radics et al. 2014). The structure of the needle complex (Kubori et al. 1998) resembles that of the flagellar basal body (Aizawa et al. 1985), further reinforcing the similarity of the two systems.

Bacteria have been known to be motile since their discovery (Macnab 1999). Although they were initially speculated to possess paws enabling motility, improvements in microscopy led to the first observation of flagella (Ehrenberg 1838). Flagella rotate like a propeller to enable motility, as shown most conclusively by observation of rotation in tethered cells (Berg 1974; Silverman and Simon 1974). The energy for rotation comes from the pmf, not ATP (Larsen et al. 1974a; Manson et al. 1977). Like the injectisome needle (Cordes et al. 2003), the flagellar filament is a hollow tube (Lowy and Hanson 1965) and growth occurs at the distal end of the flagellar filament (Iino 1969; Emerson et al. 1970) or needle (Kimbrough and Miller 2000) from secreted subunits as they emerge from the growing filament/needle and fold. Analogous to secretion in injectisomes, flagellar subunits bind cognate chaperones in the cytoplasm, both to allow recognition as substrates and avoid premature formation of oligomers in the cytoplasm (Auvray et al. 2001).

1.2.2 Flagella and injectisomes in motility and pathogenesis

Flagellar motility has long been known to be related to chemotaxis (Engelmann 1881), the movement towards favourable substances and away from unfavourable chemicals. Importantly, rotation of *E. coli* flagella was seen to change from counterclockwise (CCW) in the presence of attractants to clockwise (CW) in the presence of repellents (Larsen et al. 1974b). While change of direction of the rotation would only cause a change from forward to backward movement if a single flagellum is present, peritrichous flagella, that originate from many different points around the cell envelope, are found in the most well studied organisms, *Salmonella enterica* serovar Typhimurium (*Salmonella* Typhimurium) and *E. coli*. Peritrichous flagella form a helical bundle during CCW rotation and this bundle comes apart when one or more flagella rotate CW, causing the cell to tumble (Macnab and Koshland 1974). This reorients the cell into a random new direction at the end of the tumble, leading the cell to explore its environment in a random walk.

In order to allow the cell to swim towards nutrients and away from harmful substances, the frequency of tumbling/direction changes is controlled, allowing the cell to follow concentration gradients despite being too small to sense a gradient along its cell body (Berg and Brown 1972). In *E. coli* a reduction in concentration of chemoattractant

substances or an increase in concentration of chemorepellent substances causes tumbling frequency to increase, leading to a biased random walk towards attractants and away from repellents.

Different swimming behaviours are found in other flagellated species. For example, a more complex mechanism is found in marine species, like *Vibrio alginolyticus*, with a single polar flagellum. CCW rotation of the flagellum causes forward movement and CW movement reverses the direction of swimming, causing the cell to backtrack. Prior to resumption of forward swimming, the flagellum acts like a rudder and causes a random change in direction (Xie et al. 2011).

Due to the fitness advantages of motility and chemotaxis, flagella are common in bacterial pathogens (Josenhans and Suerbaum 2002). But in addition to motility, flagella can have further biological functions in pathogenicity including adhesion, biofilm formation and invasion (Duan et al. 2013). Furthermore, virulence factors can be secreted through flagella (Young et al. 1999) and in *Vibrio cholerae* flagella are used as a sensor of the mucin layer in the host intestine (Liu et al. 2008). Due to the widespread occurrence of flagella and their importance in virulence, the protein making up the flagellar filament, flagellin/FliC, is a major target of immune receptors (Hayashi et al. 2001).

The injectisomes is a molecular syringe, consisting of a basal body embedded in the cell envelope and an extracellular needle (Kubori et al. 1998). Upon binding to host cells, pathogens can insert a translocon connected to the tip of the needle into the eukaryotic plasma membrane and secrete proteins into the host cytoplasm, passing three membranes in a single step secretion mechanism (Hakansson et al. 1996). The secreted effectors are very diverse and have different functions that allow each pathogen to cause disease. Effectors can manipulate host cell function to promote adherence to target cells, invasion, direct cytotoxicity or subversion of immune system function (Coburn et al. 2007). The T3SS is often essential for virulence. For example, *Salmonella* Typhimurium encodes two injectisomes on its two pathogenicity islands (SPI-1 and SPI-2)

(Galan and Curtiss 1989; Ochman et al. 1996; Shea et al. 1996). The SPI-1 injectisome is expressed early in infection and is responsible for invasion and early stages of intracellular survival (Zhou and Galan 2001). Unusually, *S.* Typhimurium triggers its own endocytosis by endothelial cells. Multiple SPI-1 effectors, including the bacterial actin binding proteins SipA and SipC, co-operate to induce actin polymerisation and uptake of *S.* Typhimurium cells (Hayward and Koronakis 1999; Zhou et al. 1999). The cells are then maintained in a *Salmonella*-containing vacuole (SCV) which does not fuse with endosomes or lysosomes. The SPI-2 injectisome is largely responsible for SCV maturation and prevention of killing by phagocytes (Vazquez-Torres et al. 2000). *S.* Typhimurium also exploits the immunogenicity of flagellin by secreting the protein through the SPI-2 injectisome (Lyons et al. 2004).

In contrast, *Yerinia* does not hide from the immune system inside host cells and instead kills immune cells. *Yersinia pestis* preferentially targets cells of innate immunity for effector injection, resulting in cell death (Marketon et al. 2005). Effectors including YopM (Kerschen et al. 2004) and YopJ (Orth 2002) can cause cell death directly and the effector YopH (Persson et al. 1999) prevents phagocytosis.

Injectisomes are also responsible for establishing symbiotic relationships, in particular between leguminous plants and nitrogen fixing rhizobium bacteria (Viprey et al. 1998). In order to establish symbiosis, effectors such as NopL downregulate the plant immune response and inhibit accumulation of the defensive pathogenesis-related (PR) proteins (Bartsev et al. 2004) which allows the bacteria to colonise the root and induce the formation of nodules.

1.3 Structural biology of type three secretion systems

As sequence information of the subunits that make up injectisomes and flagella has become available, homology between most parts of the two systems has been realised (Figure 1.3). But while a single set of gene names is used for flagella, multiple nomenclatures have traditionally been used for injectisome genes. A universal Sct nomenclature has been proposed (Hueck 1998) and this will be used in this thesis. The

homology between subunits of the two systems can be seen not only at the genetic level, but increasingly also in structures of subunits as well as the intact nanomachines (Diepold and Armitage 2015; Portaliou et al. 2016). Furthermore, both systems employ specific chaperones to recognise secretion substrates (Michiels and Cornelis 1991; Auvray et al. 2001), both depend on the pmf for secretion (Minamino and Namba 2008; Lee and Rietsch 2015) and both nanomachines are assembled hierarchically, partially dependent on the order of gene expression (Minamino 2018; Wagner et al. 2018). Assembly of the helical filaments is precisely controlled through a molecular ruler protein SctP and secretion specificity is switched in injectisome assembly after the helical needle has grown to a certain length (Journet et al. 2003; Hughes 2012). The homologous system in flagella, using the ruler protein FliK, controls the length of the hook (Erhardt et al. 2010) (Figure 1.3). Due to these similarities, flagella can be seen as a type of type three secretion system (fT3SS) that is evolutionarily related to the non-flagellar or virulence-associated type three secretion system (vT3SS).

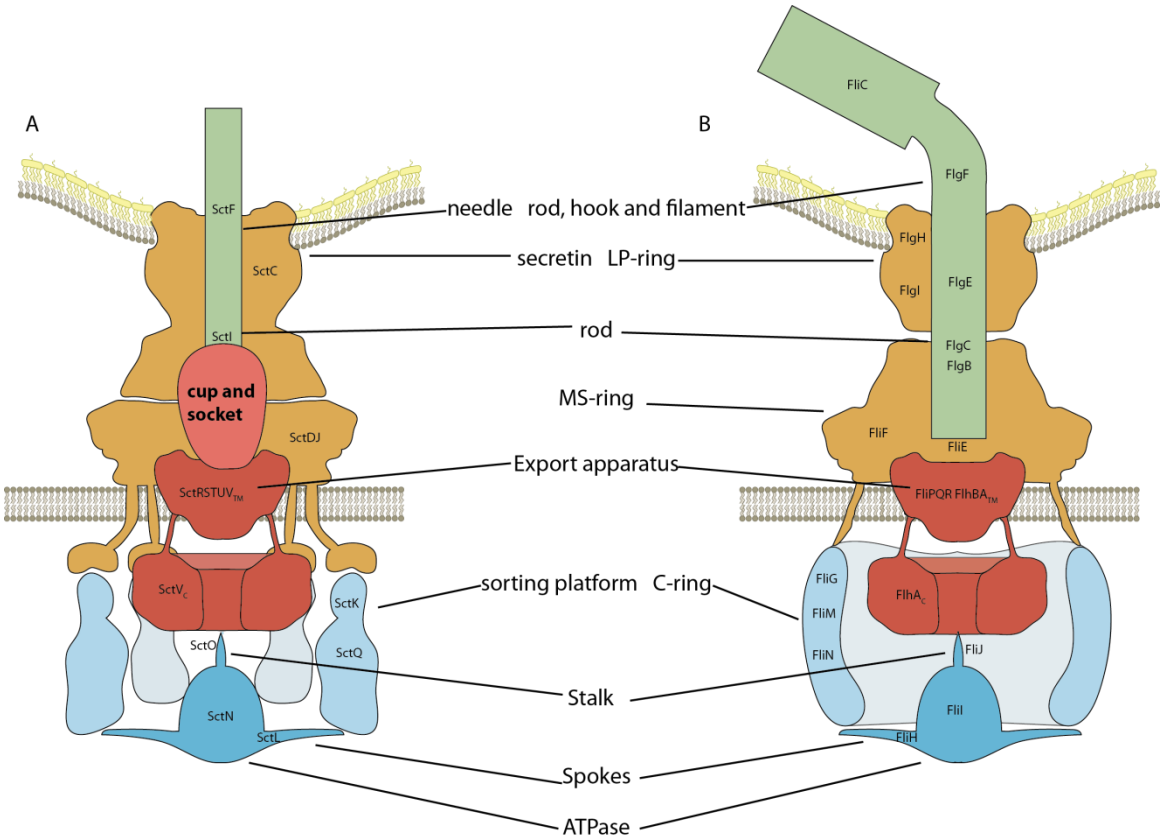


Figure 1.3 Cartoon model of vT3SS and fT3SS.
Cartoon model of the organisation of vT3SS (**A**), and fT3SS (**B**) before the work on this thesis was begun.

1.3.1 Structure of T3SS: the basal body

The main structural component of the T3SS is the basal body; a large complex made up of a series of circularly symmetric rings in each of the bacterial membranes (Akiba et al. 1991; Ueno et al. 1992; Schraidt and Marlovits 2011). Unlike the highly conserved membrane proteins they surround, the basal bodies of injectisomes and flagella have diverged in evolution, possibly related to the mechanical stresses associated with the rotation of the flagellum. The inner membrane ring of the injectisome is composed of two proteins, SctD and SctJ, each of which is built from multiple structurally related domains termed ring building motifs (RBMs) (Spreter et al. 2009). They are each anchored in the membrane by a single transmembrane helix, with SctJ also possessing a covalently attached lipid at the N-terminus in most species (Allaoui et al. 1992). Together they form a 24-fold symmetric ring, with 24 copies of SctD enveloping a tightly packed ring of SctJ. This in turn surrounds a central periplasmic structure known as the cup and socket (Schraidt and Marlovits 2011). Recent near-atomic resolution structures of the vT3SS basal body since this study began have made the stoichiometry and structure of the SctDJ ring unambiguous (Worrall et al. 2016).

In contrast, the equivalent structure in fT3SS is formed by a single protein, FliF, which forms a structure known as the MS-ring (Ueno et al. 1992). FliF has previously been visualised using low-resolution cryo-negative stain electron microscopy which led to an estimated stoichiometry of approximately 26 (Suzuki et al. 2004) and similar estimates of the stoichiometry have been arrived at through the analysis of basal bodies purified from intact flagella (Thomas et al. 2006). Recently, high-resolution cryo-EM of FliF revealed stoichiometries of 33 and 34 and multiple different symmetries in parts of the structure formed by the single chain of FliF (Johnson et al. 2019). Housed in a specialised patch of membrane at the centre of this inner membrane ring is thought to be the highly conserved export apparatus through which secretion substrates are exported (Wagner et al. 2010; Dietsche et al. 2016; Minamino 2018).

In the periplasm the inner membrane ring contacts the outer membrane pore (Schraidt and Marlovits 2011). In vT3SSs this is the secretin, SctC. Earlier low resolution cryo-EM structures of related T2SS secretins had suggested variable stoichiometry, with

assemblies ranging from 12 to 15-mers proposed. However, recent advances in cryo-EM have led to the determination of structures of the *S.* Typhimurium vT3SS basal body (Worrall et al. 2016), including the secretin, and multiple T2SS secretins (Hay et al. 2017; Yan et al. 2017), revealing a conserved architecture and stoichiometry. The core secretin is a giant pentadecameric beta-barrel penetrating the outer membrane at its upper surface. Unusually for outer membrane proteins, the beta-barrel extends far into the periplasm and is double-walled. N-terminal to the barrel structure is a series of globular domains that display structural homology to the RBM motifs of the inner membrane components. In vT3SS these RBM domains span the periplasm and connect the secretin to SctDJ (Worrall et al. 2016). Although the periplasmic RBMs were assumed to be 15-fold symmetric, following the symmetry of the membrane embedded domain, recent results suggest that an additional fragment of SctC is part of the periplasmic complex, making it 16-fold symmetric at the periplasmic ring that contacts SctDJ (Goessweiner-Mohr et al. 2019; Hu et al. 2019). Interestingly, although unrelated in sequence, the periplasmic collar of the inner membrane FliF ring seems to occupy the same space in the basal body as the N-terminal secretin RBM domain (Suzuki et al. 2004).

Secretins are thought to assemble independently of the rest of the secretion system (Diepold et al. 2010). Secretin assembly is facilitated by pilotins. The pilotin promotes proper assembly of the secretin and targets insertion of the secretin monomers into the outer membrane rather than the inner membrane (Majewski et al. 2018). Leakage through the secretin prior to the completion of the secretion system is prevented by a periplasmic gate made up of two sites called GATE1 and GATE2 (Spagnuolo et al. 2010). A pair of structures of the *S.* Typhimurium vT3SS basal body with an assembled needle, termed needle complex, and without the needle have recently elucidated the molecular basis of this gating mechanism (Worrall et al. 2016; Hu et al. 2018). It is thought that in T3SS secretins assembly of the needle filament causes the gate to flip from a position parallel to the membrane to one that is perpendicular, where it becomes part of the secretin beta-barrel.

The outer membrane components of the fT3SS are less well understood. Flagella of most Gram-negative species are thought to pass the peptidoglycan and outer membrane via

the P and L-rings, formed by FlgI and FlgH respectively (Jones et al. 1987). Neither of these proteins is predicted to be related to secretins. However, there are currently no high resolution structures of either ring and FlhH and FlgI do not resemble any protein of known structure at the sequence level. Interestingly, it has recently been shown that when flagella are ejected under starvation conditions a plug appears in the outer membrane ring (Ferreira et al. 2019), but it is not known whether the plug is made up of a separate protein or formed by FlgH or FlgI. Both secretins and the inner membrane base of vT3SS have the ability to assemble independently and subsequently associate with each other (Wagner et al. 2018), however assembly of the secretin may be promoted by the clearing of peptidoglycan near the growing inner membrane ring (Burkinshaw et al. 2015). In contrast, assembly of L- and P-rings were thought to be dependent on the presence of the rod and inner membrane base (Kubori et al. 1992; Minamino and Macnab 1999; Minamino and Namba 2004). Recently, assembly intermediates of both individual inner membrane rings and L-P rings could be observed by cryo-ET in multiple species (Kaplan et al. 2018). Separate assembly of outer and inner membrane sub-complexes followed by association of the subcomplexes may therefore be possible in both systems.

Running through the centre of the basal body are the inner rod/needle of vT3SSs and the proximal/distal rod of fT3SSs (Radics et al. 2014; Fujii et al. 2017). These subunits are secreted by the export apparatus and assemble onto the distal end of a growing helical structure. In so doing the axial components of the T3SS are built by the secretion system itself.

1.3.2 Structure of T3SS: the C-ring/sorting platform and the ATPase

At its cytoplasmic face, the basal body contacts another large assembly. In fT3SS, FliF binds a continuous cytoplasmic ring structure termed the C-ring made up of FliG, FliM and FliN (Lynch et al. 2017; Xue et al. 2018). Together, FliF and the C-ring form the rotor of the flagellar motor, with the C-ring also playing roles in flagellar rotational switching in response to chemotactic stimuli and in the recruitment of substrates for secretion. The vT3SS counterpart of the C-ring, known as the sorting platform, on the other hand has been shown consist of six “pods” (Hu et al. 2015) rather than a continuous ring.

The cytoplasmic components of the T3SS are thought to recruit secretion substrates before they pass through the export apparatus. In vT3SSs, a complex of SctQ, $SctQ_C$, SctK and SctL forms the sorting platform (Lara-Tejero et al. 2011) that binds substrates sequentially and delivers them to the export apparatus. The vT3SS protein SctQ and $SctQ_C$, the product of an internal translation start site within the SctQ gene that produces a C-terminal fragment, can form a 1 to 2 complex. This is equivalent to the flagellar complex formed by 3 FliM and 1 FliN as both are made up of 3 SpoA2 and 1 SpoA1 domain (Notti et al. 2015; McDowell et al. 2016). Although it was originally thought that SctQ-Q_C/FliMN both form a ring in their respective systems, it is now clear that in vT3SSs the sorting platform consists of six pods (Hu et al. 2015) (Figure 1.4A) attached to a cytoplasmic extension of SctD – though how the 24 copies of this extension are coordinated to generate only 6 attachment sites is not currently understood. The highest resolution *in situ* reconstruction of an injectisome has recently been achieved in *Salmonella* (Hu et al. 2017). This revealed the location of a domain unique to vT3SS. The N-terminal domain of the basal body subunit SctD could be shown to connect the inner membrane rings with the sorting platform, a role carried out by the unrelated protein FliG in flagella (Lynch et al. 2017; Xue et al. 2018). Interestingly, it does not form a continuous ring, but six clusters, each connected to one of the pods of the sorting platform (Figure 1.4B-C). In the absence of the sorting platform, it does not appear to form clusters, suggesting that the basal body may co-ordinate assembly of the cytoplasmic sorting platform.

Despite the similarity of the FliMN and SctQ-Q_C complexes, it is still controversial whether both SctQ and $SctQ_C$ are structural components of the sorting platform. $SctQ_C$ has been reported to be essential in *Shigella* (McDowell et al. 2016) and *Yersinia* (Bzymek et al. 2012) but appears dispensable in both *Salmonella* SPI-1 (Lara-Tejero et al. 2019) and SPI-2 (Yu et al. 2011) vT3SS. In the *Salmonella* SPI-1 injectisome $SctQ_C$ has been reported to contribute to assembly of the sorting platform but is absent in the final assembly (Lara-Tejero et al. 2019).

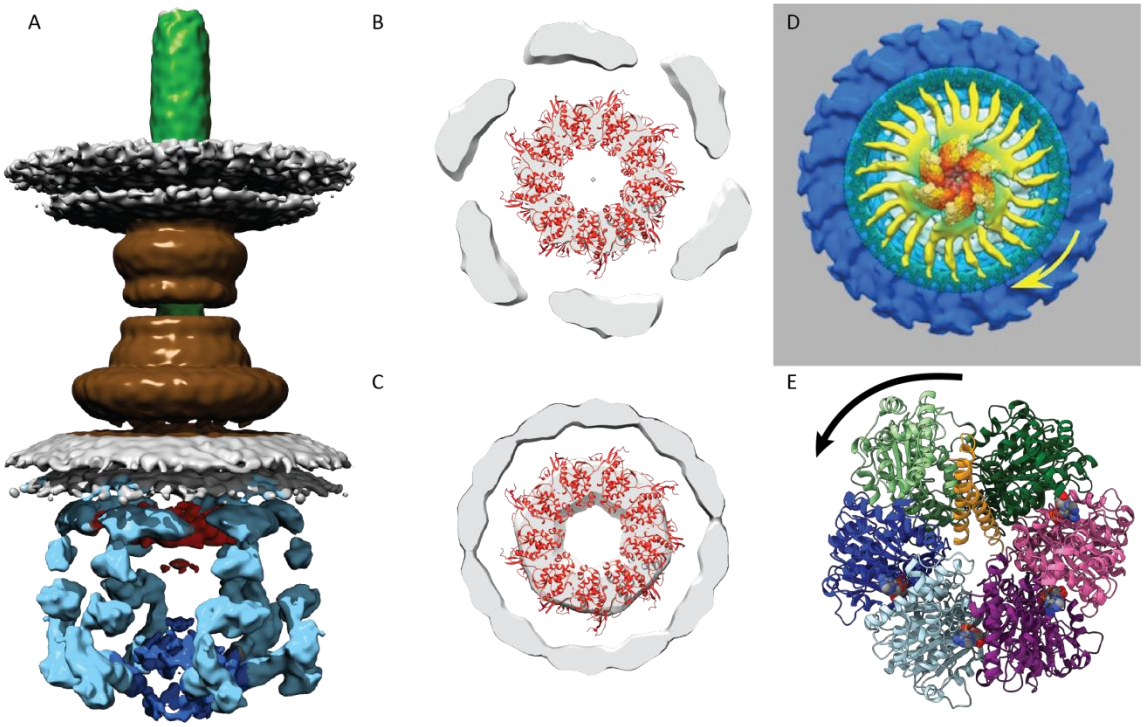


Figure 1.4 The cytoplasmic assemblies.
A, Tomographic reconstruction of the vT3SS from *S.* Typhimurium SPI-1 with C6 symmetry applied (EMD-8544). **B**, Slice of the tomogram showing the six clusters of density corresponding to the cytoplasmic domain of SctD surrounding the toroidal density corresponding to the cytoplasmic domain of the export apparatus protein SctV whose structure is shown in red (PDB: 4A5P). **C**, Equivalent slice as in (**B**) of the tomogram of a *sctQ* mutant (EMD-8445) after applying C6 symmetry. **D**, View from the cytoplasm of the subtomogram average of the cytoplasmic parts of the fT3SS of *Borrelia* (EMD-9122) with an arrow indicating the rotation of the C-ring (reproduced from (Qin et al. 2018)). **E**, Structure of the vT3SS ATPase viewed from the inner membrane with an arrow indicating the direction of the rotation of the stalk (PDB: 6NJP).

In fT3SS the C-ring interacts with the basal body protein FliF through a strong interaction of the cytoplasmic C-terminal helix of FliF with the C-ring protein FliG (Lynch et al. 2017; Xue et al. 2018). $FliF_C$ and $FliG_N$ are parts of a split domain and form a 1:1 heterodimer, and FliG can be fused to FliF in some organisms (Francis et al. 1992). Modelling of the assembled FliF ring suggested a stoichiometry of approximately 25 or 26 (Suzuki et al. 2004; Bergeron 2016). However, the C-ring is thought to have a stoichiometry of approximately 34 (Thomas et al. 1999). Different models of how this symmetry mismatch is bridged have been proposed (Stock et al. 2012), including an interaction between FliG and FliMN that is not 1:1 due to gaps in the FliG ring compared to the FliMN ring, two modes of interaction between FliM and FliG or dynamic exchange of

FliMN in the flagellum with cytoplasmic pools. The recent cryo-EM structure of FliF showed that a majority of recombinant MS-rings were 33-fold, but some were 34-fold symmetric (Johnson et al. 2019), suggesting that the symmetries of C-ring and MS-ring are matched, although whether the stoichiometry in native flagella is 33 or 34 or a mixture of stoichiometries is not clear yet. This is consistent with the existence of a natural FliF-FliG fusion in *Chlamydia* (Bergeron 2016), although flagella are not known to occur in *Chlamydia*. As the diameter of the C-ring is not conserved in evolution, there may be other stoichiometries in organisms with unusually large or small C-rings (Abrusci et al. 2013; Qin et al. 2018). A cryo-EM study of the C-ring in a deletion fusion mutant of FliF and FliG in *Salmonella* missing the FliF C-terminus and FliG N-terminus found that the density previously proposed to correspond to these two termini disappeared, confirming the assignment of this most membrane-proximal part of the C-ring (Sakai et al. 2019). Interestingly, the diameter of the C-ring was smaller in this mutant because the majority of the assemblies were 31-fold symmetric rather than 34-fold as found in the wild type.

The cytoplasmic components of the T3SS have recently been realised to be far more dynamic than the membrane embedded proteins. Fluorescence imaging has revealed that many of the cytoplasmic proteins exist in a cytoplasmic pool as well as the membrane-localised complete assembly and that the two populations are in constant exchange. Similarly, the flagellar C-ring components FliM and FliN can also exchange with their cytoplasmic pools (Kihara et al. 1996; Delalez et al. 2010; Fukuoka et al. 2010; Kudryashev et al. 2010; Delalez et al. 2014). A recent study using super-resolution microscopy to track single molecules defined distinct complexes of SctQ and SctL oligomers (Rocha et al. 2018). Interestingly, the distribution of oligomers shifted to smaller species when secretion was turned off. However, it is not yet clear through what mechanism the cytoplasmic complexes regulate secretion.

SctN/FliI, the ATPase, is at the centre of the sorting platform/C-ring and has been visualised as a central 'hub' in tomograms of the *Shigella* vT3SS (Hu et al. 2015). The homology between SctN/FliI and F-ATPases has lead to the proposal that in the fully assembled complex SctN/FliI is a hexamer (Imada et al. 2007). Although ATP hydrolysis is thought to contribute to the energy requirement of secretion, it is dispensable under

some circumstances, while the pmf is thought to be essential for secretion (Minamino and Namba 2008). The sorting platform is connected to the ATPase by spokes that vary in length depending on C-ring size (Abrusci et al. 2013). SctN/FliI is known to interact with SctL/FliH in a one to two complex *in vitro* in both systems (Imada et al. 2016) predicting a model in which 6 SctL/FliH dimers form 6 spokes. However, a recent cryo-ET study of the periplasmic flagella of *Borrelia* (Qin et al. 2018) has overturned this hypothesis about the stoichiometry in the native complex. The reconstruction revealed a 46-fold symmetric C-ring complex that is connected to the central hexameric ATPase through 23 spokes, thought to consist of two FliH molecules each. A subpopulation of cells appears to have a different number of spokes, ranging from 21 to 24. Computational classification revealed rotation of the ATPase complex relative to the stators of the flagellar motor, the MotAB complex. As the ATPase is attached to the C-ring via the spokes, it has been proposed that the C-ring and the ATPase rotate as a single rigid body (Figure 1.4D). More high-resolution tomographic reconstructions of flagella of other organisms are required to clarify the conservation of the *in vivo* stoichiometry of the SctN/FliI-SctL/FliH interaction.

The monomer of the ATPase has been extensively characterised and is homologous to the F/V-ATPases (Imada et al. 2007; Zarivach et al. 2007). However, as the ATPase complex is a homohexamer of SctN/FliI, unlike the trimer of heterodimers complex of F/V-ATPases, it was not clear how rotation of the stalk, SctO/FliJ, is achieved. The cryo-EM structure of the assembled hexamer revealed that importantly the stalk is in the centre of the ATPase and the SctN homohexamer is asymmetric (Majewski et al. 2019): two subunits each are in the empty state and the other four are bound to a transition state analogue of ATP. When looking onto the membrane proximal face of the complex the transition state analogue is gradually bound more tightly going clockwise. Based on this structure a rotation mechanism was proposed in which two subunits each are in the ATP bound state, the ADP bound state or the empty state, similar to F/V-ATPases. The stalk contacts the empty subunits through electrostatic interactions and rotates counter-clockwise (Figure 1.4E) as ATP is hydrolysed by each subunit in turn. Conformational changes in the ATPase subunits expose a negatively charged surface in the empty state which a positively charged region of the stalk binds, resulting in rotation. The direction

of this rotation is the same as that of the flagellum of *E. coli* and *Salmonella* when the cell is moving forwards. The function of rotation is not clear but the interaction of SctO/FliJ with the export apparatus protein SctV/FlhA may cause conformational changes in its cytoplasmic domain that facilitate dissociation of chaperone-substrate complexes which are known to dock onto SctV/FlhA (Bange et al. 2010). Another hypothesised function of the ATPase is to create a local proton gradient that could power the export apparatus (Morimoto et al. 2016; Terashima et al. 2018). The structural similarity of the SctN hexamer to F/V-ATPases supports this, however, it is not clear which protein would function as an equivalent of the F-ATPase c ring to facilitate proton pumping. One candidate protein that has been observed to conduct protons and sodium ions (Minamino et al. 2016) is the membrane domain of the export apparatus protein FlhA/SctV (see below).

1.3.3 Structure of T3SS: needle and flagellar filament

At its extracellular face, a helical filament grows out of the basal body. The flagellar filament is made up of flagellin (FliC). However, contrary to expectations of a helical filament and unlike the vT3SS needle, the flagellum is not a straight filament. Instead, it is shaped like a wave. It was proposed (Asakura 1970) that this may be explained by considering the FliC helix, which is made up of 5.5 subunits per turn (11 subunits in two turns), to be made up of 11 protofilaments that together form a tube. The protofilaments can be in two different conformations (L and R). Small differences in the length of these protofilaments would produce twisted structures. The structures of the L-type (Yonekura et al. 2003) and R-type (Maki-Yonekura et al. 2010) protofilaments confirm this model. Switching from CCW to CW rotation of the flagellum is though to change the left handed flagellar supercoil (9 L-type and 2 R-type protofilaments) into a right-handed supercoil (between 4 and 6 R-type protofilaments) which is unable to form a bundle and so the filaments of each flagellum detach from each other, resulting in tumbling rather than forward swimming (Maki-Yonekura et al. 2010). In contrast, the vT3SS needle is a straight helical filament with similar helical parameters as the flagellar filament (Cordes et al. 2003; Loquet et al. 2012).

The filament and needle are attached to their respective basal bodies via a helical inner rod (vT3SS) or the rod and hook (fT3SS). In the past few years, cryo-EM has produced near-atomic structures for the major periplasmic filamentous components of both systems. In fT3SS, the proximal rod is formed sequentially by single helical turns of FlgB, FlgC and FlgF (Zhao et al. 2013), followed by approximately 5 helical turns of FlgG (Jones et al. 1990; Fujii et al. 2017). The helical rod is proposed to be anchored onto the circularly symmetric motor structure via an adapter protein called FliE, which shares features with the more distal rod proteins (Kubori et al. 1992; Minamino et al. 2000). A 7 Å cryo-EM map of a filament formed from by a mutated FliG, called polyrod, (Fujii et al. 2017) has confirmed its structural homology with the more distal hook filament (Samatey et al. 2004), with subtle differences in domain orientations proposed to impart rigidity to the rod structure.

In vT3SS, cryo-EM of the fully assembled basal body has determined that the bulk of the periplasmic lumenal filament is composed of the needle that continues into the extracellular space (Hu et al. 2018). The protein previously identified as the “inner rod”, termed SctI, is proposed instead to function as an adaptor protein linking the needle to the “cup and socket” structure (Marlovits et al. 2006) through just a single helical turn made up of 5 or 6 subunits, consistent with earlier mass spectrometry data (Hu et al. 2018; Goessweiner-Mohr et al. 2019; Torres-Vargas et al. 2019). Similar to FliE, the cup and socket had been suggested to be a symmetry adapter, but the identities of the proteins making up this structure were not known.

At its distal end the needle is capped by a tip complex. This complex is thought to be made up of four subunits of the tip protein IpaD (*S. flexneri* nomenclature) and one copy of IpaB, but can be formed by five IpaD molecules when IpaB is deleted (Cheung et al. 2015). Following activation of secretion, the translocators, IpaB and IpaC, are secreted and inserted into the membrane to form the translocon. The structure and activity of the translocon is the least well understood part of type three secretion. Although it has long been assumed that the translocon is continuous with the needle, it has only recently become possible to image the point of contact between the needle and the host cell. Cryo-ET reconstructions of *Chlamydia* (Nans et al. 2015) and *Salmonella* (Park

et al. 2018) cells in contact with host cells revealed the presence of a pore-like object in the host cell membrane that is connected to the needle tip. This object could be probed in detail in *Salmonella* to show that this presumptive pore *in situ* is much smaller than what had previously been observed with purified translocon components (Ide et al. 2001), suggesting that the tip complex may have a role in proper assembly of the translocon. The *in* situ structure of the translocon also unexpectedly protrudes into the host cytoplasm, however it has not yet been possible to unambiguously determine that this structure consists of the translocators and no other components. Future studies will be needed to elucidate the structure of the translocon and whether it plays a role in the process of secretion beyond forming an open channel.

1.4 The export apparatus

When searching for genes responsible for secretion, five inner membrane proteins (Figure 1.5) were identified as potential components of an inner membrane export apparatus (EA) through which the secretion substrates could be secreted from the cytoplasm into the growing flagellum (Minamino and Macnab 1999). The EA was hypothesised to reside within a patch of membrane at the centre of the basal body. The high conservation of these and other proteins originally led to the idea that both injectisomes and flagella utilise a similar type 3 secretion machinery centred around the shared export apparatus (Macnab 1999).

Despite today's very detailed molecular understanding of the structure and function of T3SS (Portaliou et al. 2016) the critical step of how the EA allows secretion substrates to cross the inner membrane is still not understood. Revealing the structural basis for permeabilising the inner membrane in T3SS is the aim of this thesis.

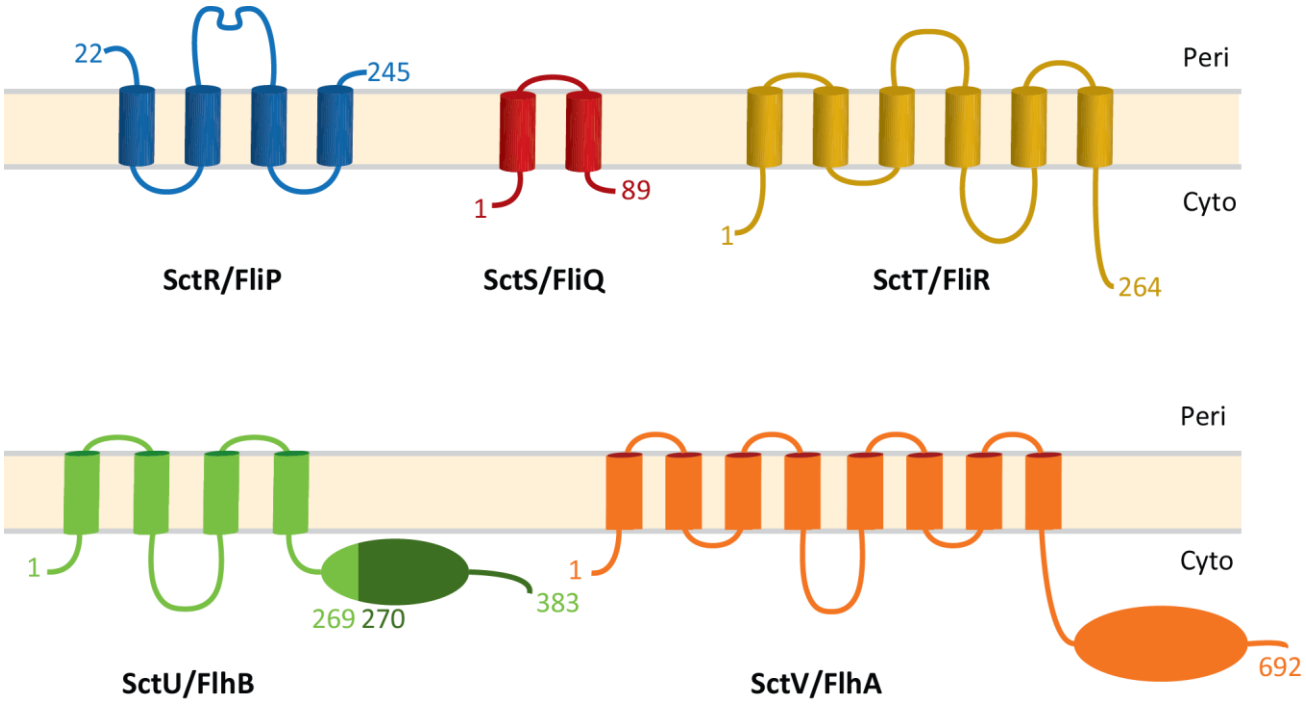


Figure 1.5 Topology diagrams of the five EA proteins.
Predicted topologies of the EA proteins following (Erhardt et al. 2017). The structures of the cytoplasmic domains of SctU/FlhB and SctV/FlhA (Zarivach et al. 2008; Worrall et al. 2010) and the periplasmic domain of FliP have been studied crystallographically (Portaliou et al. 2016; Fukumura et al. 2017) but the structures of the membrane domains were not known. FliP contains a signal sequence (Malakooti et al. 1994).

1.4.1 The major component of the apparatus

SctV/FlhA is the largest EA subunit and is made up of a large transmembrane domain ($SctV_{TM}$/$FlhA_{TM}$) and a large cytoplasmic domain($SctV_C$/$FlhA_C$) (Figure 1.5). Although different oligomeric states of SctV/FlhA had been proposed (Lilic et al. 2010; Diepold et al. 2011; Li and Sourjik 2011), the nature of these oligomers was only revealed in the crystal structure of *S. flexneri* $SctV_C$. The protein formed a nonameric ring (Figure 1.6A) in the crystal and tomography supported the *in vivo* relevance of the crystallographic symmetry (Abrusci et al. 2013). The structure implied that the transmembrane domain of SctV/FlhA also forms a nonameric ring and that the remaining EA proteins form the secretion channel inside of this transmembrane ring.

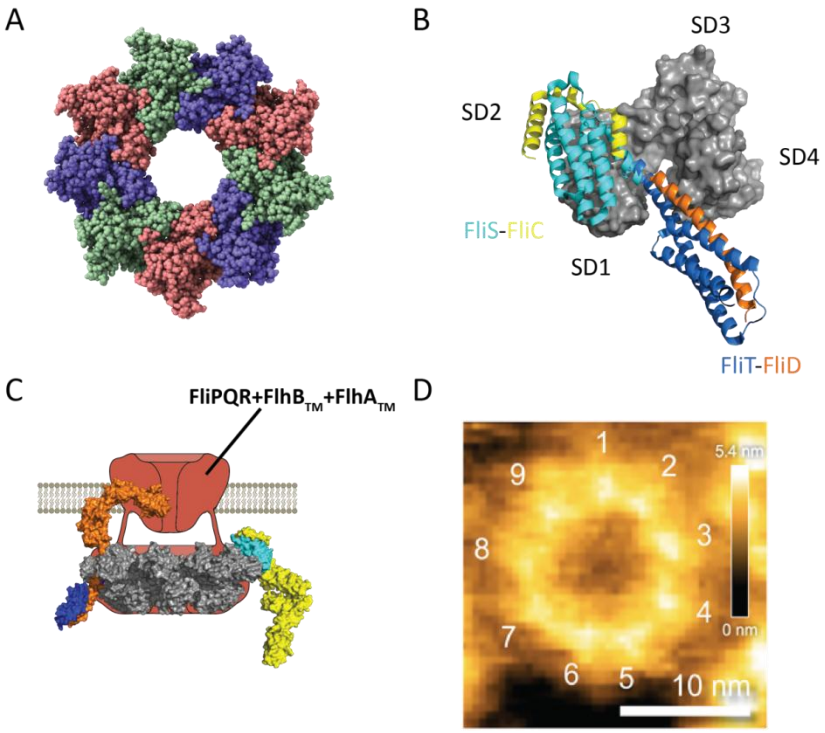


Figure 1.6 The major export apparatus subunit SctV binds to secretion substrates in the cytoplasm.
A, Crystal structure of SctV$_C$ from Shigella, which crystallised as a nonameric ring (PDB: 4A5P). **B**, Superposition of the structures of the SctV cytoplasmic domain (grey, surface representation) in complex with the chaperone FliS and FliC (PDB: 6CH3) and the chaperone FliT and FliD (PDB: 6CH2). The four subdomains (SD1-SD4) are annotated. **C**, Nonameric model of FlhA in complex with a model of full-length FliS and FliC (right, modelled using 6CH3, 1UCU and 5MAW) and FliT and FliD (left, modelled using 6CH2, 5GNA, 5H5V). **D**, AFM image of FlhA$_C$ on mica (reproduced from (Terahara et al. 2018)).

FlhA$_C$ is made up of four subdomains (SD1-4) (Figure 1.6B) and has been proposed to act as a docking site for flagellar filament proteins in complex with their cognate chaperones (Bange et al. 2010), prior to the secretion of the substrate. Recent crystal structures of FlhA$_C$ bound to fragments of the flagellar filament-capping protein (FliD)/chaperone (FliT) complex and the flagellar filament protein (FliC)/chaperone (FliS) complex have shed some light on the precise mode of interaction with substrates (Xing et al. 2018). The structures reveal that FlhA recognises these substrates only through the cognate chaperone (Figure 1.6B). The chaperones FliT and FliS, but not FlgN, require binding of the substrate to induce a conformation compatible with binding to FlhA$_C$. In both the crystal structures of the ternary complexes of FliD-FliT- SctV$_C$ and FliT-FliS- SctV$_C$ a highly conserved tyrosine residue in the chaperone makes extensive hydrophobic contacts with SctV$_C$. Both chaperones can be seen to bind the same site on FlhA$_C$, a cleft between

subdomains SD1 and SD2, and NMR measurements indicate that the third flagellar chaperone, FlgN, also binds this site via a tyrosine residue. Docking of models of full length substrate/chaperone complexes onto a model of the fT3SS SctV positioned the substrates adjacent to the inner membrane, presumably primed for entry into the secretory pathway (Figure 1.6C). However, it is unclear from the structures what the mechanism of release of the secretion substrate from the chaperone would be, or what the nature of the factors determining secretion order is. It has been noted that chaperone binding site identified in the crystal structures is not conserved in vT3SS SctV homologues (Xing et al. 2018), although this may reflect the relative lack of conservation of the chaperones between systems.

Some insight into the mechanism of $FlhA_C$ was provided recently by combining mutagenesis of FlhA with high-speed atomic force microscopy (AFM) imaging (Terahara et al. 2018) (Figure 1.6D). $FlhA_C$ was found to assemble into nonameric rings on mica. This AFM assay was used to verify the importance of a flexible linker at the N-terminal end of the cytoplasmic domain: Alanine substitutions that disrupted the interaction between the linker of one subunit with the neighbouring subunit in the ring disrupted nonamerisation. These mutations also reduced filament assembly and decreased the affinity of $FlhA_C$ for the chaperone FlgN, leading to the hypothesis that FlgN and its cognate secretion substrates may bind across more than one SctV/FlhA subunit, but this is inconsistent with the crystal structures of FlhA-substrate crystal structures (Xing et al. 2018). However, because most of these mutants were still able to form hook-basal bodies (HBBs), requiring secretion of early substrates, it was also proposed that the $SctV_C$/$FlhA_C$ linker is responsible for triggering a conformational switch in the nonamer ring that alters secretion from early to late substrates. Conformational changes in the SctV/FlhA family is lent credence by the large range of conformations observed in the various crystal structures of SctV/FlhA homologues, manifested as a hinging movement around the central sub-domain to form an open or a closed state (Bange et al. 2010; Saijo-Hamano et al. 2010; Worrall et al. 2010; Abrusci et al. 2013). Recent crystallographic data suggest that the closed state is responsible for binding early substrates and that following specificity switching a conformational change is triggered from the closed to the open state (Inoue et al. 2019b), however, so far only the closed

state has been observed in the assembled nonameric form (Abrusci et al. 2013) and it is not clear how the observed conformational changes in the monomer relate to the nonamer. It is currently unclear how this proposed switching mechanism would interface with the proposed switching mechanisms attributed to SctU/FlhB-SctP/FliK (section 1.4.2). Also unclear is how these chaperone binding and secretion switch models would apply to vT3SS SctV as the chaperone binding site is not conserved (Xing et al. 2018) and a different site has previously been suggested as important for secretion (Abrusci et al. 2013).

In addition, vT3SS utilise a second specificity switch. After secretion of the translocators, the final category of substrates, the effector proteins, are secreted. One of the proteins regulating this switch is the gatekeeper, SctW, which is thought to hold the secretion system in an inactive state until a target host cell has been recognised (Diepold and Wagner 2014). Recently, it has been shown in multiple systems (Shen and Blocker 2016; Portaliou et al. 2017; Yu et al. 2018) that mutations in the C-terminal sub-domain of $SctV_C$ (termed SD4) result in increased secretion of late effectors and reduced secretion of translocators, likely due to a weakened interaction of SctW with SctV. Although the structural details of the SctV-SctW interaction and its effects on the conformation of $SctV_C$ have not yet been clarified, it is clear that SctV has a role in the regulation of both secretion specificity switches of vT3SSs.

One of the least well understood components of the export apparatus is the transmembrane domain of SctV/FlhA ($SctV_{TM}$/$FlhA_{TM}$). It is thought to play a role conducting ions across the membrane to power secretion (Minamino et al, 2016). Mutation of every conserved charged residue in the predicted transmembrane regions of the fT3SS export apparatus proteins demonstrated that SctV is most likely the only candidate for coupling the energy from the pmf with protein export (Erhardt et al. 2017). In particular, essential charged residues that could not be mutated while maintaining function were located in the FHIPEP motif, a short, conserved motif between the predicted transmembrane helices 4 and 5 of the SctV membrane domain. Some of these were found to facilitate self-interaction of the FHIPEP motif using suppressor mutations and double cysteine mutants, implying that the FHIPEP motif forms a membrane-

proximal ring. Erhardt et al also propose that the protonation state of the cluster of charged residues in this ring is sensed by the cytoplasmic domain and that proton conductance is coupled to a cyclical movement of $SctV_C$ to and from the membrane through retraction and extension of the linker between the two domains. However, structural evidence for such a mechanism is currently lacking and in the absence of structural information about the transmembrane domain, mutational data has to be interpreted with care. Proposed proton conducting charged residues could be difficult to distinguish from residues merely important for the structural integrity of the protein.

1.4.2 The switch protein

SctU/FlhB is a protein with a membrane and a cytoplasmic domain (Figure 1.5). Its classification as an EA subunit implies that the cytoplasmic domain ($SctU_C$/$FlhB_C$) is positioned very close to the secretion pore in the inner membrane. $SctU_C$/$FlhB_C$ has been genetically linked to SctP/FliK and hook length control as deletions in FliK led to disrupted hook length and suppressor mutations were isolated in $FlhB_C$ (Williams et al. 1996). SctP/FliK is a secretion substrate and secretion levels are highest before hook completion (Minamino et al. 1999). This led to the proposal that FliK measures hook length and sends a signal to $FlhB_C$ when the hook structure is completed. $FlhB_C$ would then change the substrate specificity of the export apparatus from early substrates (subunits of the rod and hook) to late substrates (subunits of the filament). $FlhB_C$ was found to be cleaved specifically between residues Asn269 and Pro270 (*S.* Typhimurium) in a conserved NPTH motif (Minamino and Macnab 2000) which splits $FlhB_C$ into $FlhB_{CN}$ and $FlhB_{CC}$, and this cleavage is autocatalytic (Ferris et al. 2005). Interestingly, when cleavage was prevented by mutating the cleavage site secretion specificity was locked in the early substrates state (Fraser et al. 2003). Similar results were obtained with SctU and SctP and changes in SctP length resulted in changes in needle length (Sorg et al. 2007). Despite long debate about the mechanism through which SctP/FliK signals switching to SctU/FlhB, it is now accepted that SctP/FliK is a molecular ruler (Journet et al. 2003; Hughes 2012) whose N-terminal domain is unfolded as it is secreted and thereby fills the hook/needle. Only if the N-terminal domain is extended inside the hook/needle of the appropriate size then the N-terminus of the ruler reaches the end of the growing filament at the same time as a conserved motif of the C-terminal domain of

the ruler will be positioned to allow binding to $SctU_C$/$FlhB_C$ (Bergeron et al. 2016) in the cytoplasm, leading to a potential conformational change in $SctU_C$/$FlhB_C$ that results in specificity switching. As there is no evidence for this conformational change, another model proposes a physical obstruction of the EA by the C-terminal fragment, $SctU_{CC}$/$FlhB_{CC}$, which is released upon secretion of the fragment (Bjornfot et al. 2009; Frost et al. 2012). Interestingly, the C-terminal fragment $SctU_{CC}$/$FlhB_{CC}$ is additionally expressed from a separate gene in some organisms and this second $SctU_{CC}$/$FlhB_{CC}$ has been proposed to function as a spare part (Wand et al. 2006). It is not known why a spare part is required.

The crystal structure of $SctU_C$ showed that the mechanism of auto-cleavage is through a cyclisation reaction of the asparagine which breaks the peptide bond to the proline in the NPTH motif (Zarivach et al. 2008). The structure is highly conserved in both types of T3SS (Meshcheryakov et al. 2013). Despite this extensive characterisation of $SctU_C$/$FlhB_C$, little was known about the putative membrane domain, or how the interaction between SctP/FliK and SctU/FlhB causes switching. Even the stoichiometry of SctU/FlhB was still unclear, as detergent purified FlhB is dimeric (McMurry et al. 2015), but estimates from extracted basal bodies suggested presence of only one copy (Zilkenat et al. 2016).

1.4.3 The core of the apparatus: the export gate

Not much was known about the location and function of the SctRST/FliPQR proteins except that they were known to be associated with the basal body in both fT3SS (Fan et al. 1997) and vT3SS (Wagner et al. 2010). The proteins FliP and FliR are a possible nucleation point for assembly of the basal body as through an elegant use of temperature-sensitive mutants the order of incorporation of flagellar subunits could be monitored using autoradiography which revealed two previously uncharacterised proteins, now known to be FliP and FliR, to form “pools” (complexes) at an early stage of assembly (Jones and Macnab 1990). FliP and FliR synthesised early were preferentially incorporated into flagella, except in FliF mutants, indicating an interaction between FliPR and FliF and suggesting that a complex of FliP and FliR could nucleate FliF assembly. The pattern of incorporation also suggests that FliPR synthesised after basal body assembly cannot be incorporated, but this could not be shown at the time. More recently, this

inability to assemble SctRST into pre-assembled basal bodies was was shown in vT3SS (Wagner et al. 2010). Although the vT3SS basal body, like FliF (Ueno et al. 1992), could assemble in the absence of the entire EA, only SctV could be incorporated into pre-assembled basal bodies. These experiments illustrate the central position of SctRST/FliPQR not just in secretion but also in assembly of the secretion system. Furthermore, the stoichiometry of SctV/FlhA (Abrusci et al. 2013) suggested that the membrane domain of this protein forms a large ring which would surround the other EA proteins, leaving SctRST/FliPQR to form the secretion channel or export gate in the inner membrane.

In addition, the socket and cup feature at the centre of the basal body was observed to disappear in SctRST mutants (Wagner et al. 2010) although it was not clear if the density disappeared because it corresponds at least in part to the export gate or because its components require secretion through the export gate. At the time it was proposed that the cup corresponds to a periplasmic part of the export gate proteins and the socket to the inner rod, as the socket density, but not the cup, also disappeared in other mutants that destabilised the inner rod (Marlovits et al. 2006). Excitingly, recent results suggest that in the absence of SctRST the inner membrane ring of the basal body assembles into a non-physiological dead end complex with 23-fold rather than 24-fold symmetry (Butan et al. 2019; Goessweiner-Mohr et al. 2019), illustrating the importance of SctRST for the correct nucleation of SctDJ assembly.

One of the reasons for the few studies directly characterising SctRST/FliPQR is the extreme hydrophobicity of their sequences. Apart from a small stretch of residues in SctR/FliP, which have been proposed to form a periplasmic domain (Fukumura et al. 2014), the polypeptides are generally thought to be almost completely membrane embedded (Figure 1.5) based on topology prediction algorithms and experimental topology mapping (Dietsche et al. 2016; Taylor et al. 2016; Erhardt et al. 2017; Ward et al. 2018). However, the topology is not predicted consistently and especially the topology of SctT/FliR has been uncertain (Dietsche et al. 2016; Taylor et al. 2016) and prediction of the energetics of SctT transmembrane helices in the membrane suggested that a transmembrane location for two predicted transmembrane helices at the N-

terminus of FliR is energetically unfavourable (Dietsche et al. 2016). Further confusion comes from a study of the EA proteins in the vT3SS of *Xanthomonas*, which used a PhoA-LacZ fusion strategy and found that only the N-terminus of SctT was membrane embedded and the majority of the structure was periplasmic (Berger et al. 2010). Despite this uncertainty, the complex is thought to reside in the specialised area of inner membrane surrounded by the basal body (Macnab 2004). The location of N- and C-terminus in cytoplasm or periplasm is also not clear, only the N-terminus of FliP is known to be periplasmic due to the presence of a cleaved signal sequence (Ohnishi et al. 1997). The termini of SctS/FliQ and SctT/FliR are thought to be cytoplasmic (Erhardt et al. 2017), but this is not certain. A fourth protein specific to fT3SS, FliO, is now known to be a chaperone for the assembly of a FliPQR complex rather than a component of the EA (Barker et al. 2010; Fabiani et al. 2017; Fukumura et al. 2017). The difficulty in determining the exact topology indicates the export gate complex to be an interesting structural target with potential general implications for the entire field of membrane protein biology.

Recently, two groups reported preliminary results on purification and characterisation of a subcomplex of the EA (Dietsche et al. 2016; Fukumura et al. 2017). While the Namba group reported that a hexamer of FliP forms a ring that they proposed to be the core export gate (Fukumura et al. 2017), the Wagner group in collaboration with the Lea and Robinson groups in Oxford used native mass spectrometry (nMS) and negative stain EM to show that a pentameric ring of SctR formed the export gate (Dietsche et al. 2016). The latter study formed the basis of the experiments described in chapter 3, an attempt to reconcile these two models. The observation of a ring shaped complex in both studies is supportive of the model that an oligomer of SctR/FliP is responsible for making the inner membrane permeable to secretion substrates. However, the high degree of sequence conservation between FliP and SctR suggests that the stoichiometry of this complex is conserved. An additional complication of the proposed models involving perfectly symmetric FliP/SctR oligomers is that such a complex would be expected to have multiple binding sites for other EA subunits such as SctT/FliR, but just a single SctT/FliR subunit is thought to be part of T3SS (Zilkenat et al. 2016).

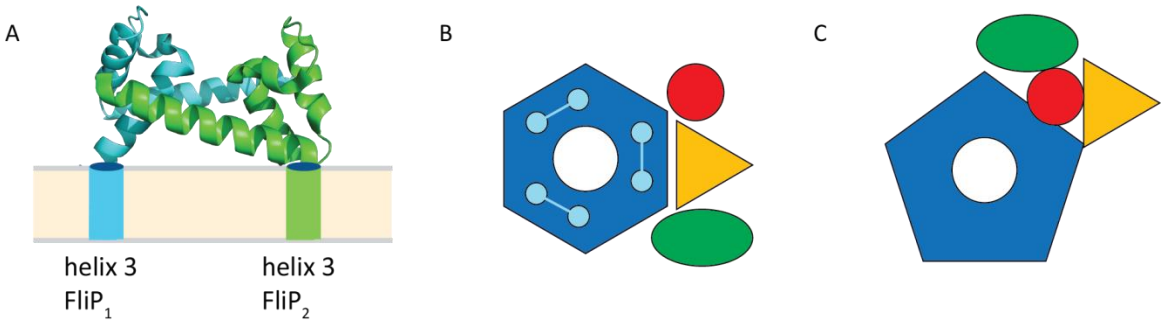


Figure 1.7 The export apparatus subunits FliP, FliQ and FliR.
A, Model for dimerization of FliP based on the crystal structure of a small, soluble fragment of the protein (Fukumura et al. 2017). **B**, Model of the FliPQR-FlhB complex as a hexamer of FliP (blue) stabilised by dimerization of the FliP periplasmic domains (cyan). FliR (yellow), FliQ (red) and FlhB (green) bind to the outside the FliP ring. Based on (Fukumura et al. 2017). **C**, Model of the SctRSTU complex as a pentamer of SctR (blue). SctT (yellow), SctS (red) and SctU (green) bind to the outside the SctR ring. Based on (Dietsche et al. 2016).

Additionally, the organisation of the FliP ring was studied. The putative periplasmic domain of FliP from *T. maritia* was found to form a tetramer in solution and in the crystal (Fukumura et al. 2014) opening the possibility that one of the interfaces between FliP periplasmic domains in the crystal is responsible for formation of a FliP dimer (Figure 1.7A). The dimer interactions between FliP subunits were further confirmed by *in vivo* photo-crosslinking (Fukumura et al. 2017). Based on this potential dimerization via the periplasmic domain, the hexameric ring was proposed to be a trimer of dimers (Figure 1.7B).

The trimer of dimers model is not compatible with the pentameric SctR complex (Dietsche et al. 2016). In this study, the stoichiometry of the SctRT complex was shown to be 5 SctR and 1 SctT by native MS. SctS was not observed despite being co-expressed, possibly due to dissociation during purification. Negative stain EM revealed two types of objects: the majority of particles were classified as a smooth ring while a minority were rings with a small protrusion. This was interpreted as a mixture of circular $SctR_5$ complexes and $SctR_5$ rings decorated with a single SctT, leading to a model centred on a pentameric SctR ring (Figure 1.7C). In both models the switch protein SctU/FlhB was suggested to bind to the SctR/FliP complex.

The two models of the complex of SctRSTU and FliPQR-FlhB are incompatible, but these highly conserved EA proteins are expected to form a similar complex in both native T3SS as even many of the proteins less conserved at the sequence level assemble into highly similar structure (Figure 1.3). The study of the stoichiometry and structure of this complex were a high priority in understanding this most central part of type three secretion and putative pore in the inner membrane.

1.5 Aims of this thesis

The aim of this thesis is the establishment of structural understanding of the membrane proteins of the export apparatus of T3SS.

1. Reconcile the two different models of the stoichiometry of the SctRST/FliPQR complex (chapter 3)

2. Determine the structure of FliPQR and examine whether this complex forms a pore in the inner membrane (chapter 4)

3. Test the structural conservation of the FliPQR complex across different systems (chapter 5)

4. Study the structure and function of the membrane domain of SctU/FlhB (chapter 6)

5. Determine the structure of the membrane domain of SctV/FlhA (chapter 7)

2 Materials and methods

2.1 Chemicals

All chemicals were from Sigma Aldrich unless indicated otherwise.

2.2 Common media and buffers

Media name	Composition
LB	10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl
2YT	16 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl
TB (Fisher Scientific)	12 g/L peptone, 24 g/L yeast extract, 2.2 g/L Potassium Phosphate, Monobasic, 9.4 g/L Potassium Phosphate, Dibasic, 0.4% (v/v) glycerol

Table 2.1 Common bacterial growth media.

Buffer name	Composition
TBS	150 mM NaCl, 100 mM Tris, 1 mM EDTA, pH 8
HBS	50 mM HEPES, 300 mM NaCl, pH 8
PBS	137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4
RF1	15% (v/v) glycerol, 100 mM RbCl, 50 mM $MnCl_2$, 30 mM potassium acetate, 10 mM $CaCl_2$, pH 5.8
RF2	15% (v/v) glycerol, 10 mM RbCl, 10 mM MOPS, 75 mM $CaCl_2$, pH 6.8

Table 2.2 Common buffers.

2.3 Molecular biology

2.3.1 Strains and plasmids

Plasmid	Vector	Insert/Reference	Species	Chapter
pMIB5689	pT12	$SctRST^{TEV2xStrepII}$ SctT 1-251	*S.* Typhimurium SPI-1	3
pT12_FliPQR	pT12	$FliPQR^{TEV2xStrepII}$	*S.* Typhimurium	3
pT12_FliOPQR	pT12	$FliOPQR^{TEV2xStrepII}$	*S.* Typhimurium	3
pT12_Spa24-9-29	pT12	$SctRST^{TEV2xStrepII}$	*S. flexneri*	3
pT12_FliOP	pT12	$FliO^{his}P$	*S.* Typhimurium	3
pT12_SpaPQR_fl	pT12	$SctRST^{TEV2xStrepII}$	*S.* Typhimurium SPI-1	3
pT12_Spa24-9-29t	pT12	$SctRST^{TEV2xStrepII}$ SctT 1-239	*S. flexneri*	3

pT12_FliOPQRt	pT12	FliOPQR^{TEV2xStrepII} FliR 1-238	*S.* Typhimurium	3
pT12_YscRST	pT12	SctRST^{TEV2xStrepII}	*Y. enterocolitica*	3
pBAD_FlhBA	pBAD	FlhBA^{his}	*S.* Typhimurium	3
pT12_FliOPQR_PSESH	pT12	FliOPQR^{TEV2xStrepII}	*P. savastanoi*	3
pT12_FliOPQR_VIBMI	pT12	FliOPQR^{TEV2xStrepII}	*V. mimicus*	3
pT12_FliOPQR_Q_E46A	pT12	FliOPQR^{TEV2xStrepII} FliQ E46A	*S.* Typhimurium	4
pT12_FliLMNOPQR	pT12	FliLMNOPQR^{TEV2xStrepII}	*S.* Typhimurium	4
pT12_HrcRST	pT12	SctRST^{TEV2xStrepII}	*P. savastanoi*	5
pT12_SsaRST	pT12	SctRST^{TEV2xStrepII}	*S.* Typhimurium SPI-2	5
pT12_FliOPQR_Q_Δ6-84	pT12	FliOPR^{TEV2xStrepII} FliQ Δ6-84	*S.* Typhimurium	5
pT12_FliOPQRtt_PSESH	pT12	FliOPQR^{TEV2xStrepII} FliR Δ165-261	*P. savastanoi*	5
pBAD_FlhBA	pBAD	^{his}FlhB	*S.* Typhimurium	6
pBAD_FlhBA	pBAD	FlhB^{his}	*S.* Typhimurium	6
pBAD_FlhBA	pBAD	^{tst}FlhB	*S.* Typhimurium	6
pT12_FliOPQR_FlhB_VIBMI	pT12	FliOPQR FlhB^{TEV2xStrepII}	*V. mimicus*	6
pT12_FliOPQR_FlhB_PSESH	pT12	FliOPQR FlhB^{TEV2xStrepII}	*P. savastanoi*	6
pT12_Spa24-9-29-40	pT12	SctRSTU^{TEV2xStrepII}	*S. flexneri*	6
pT12_SsaRSTU	pT12	SctRSTU^{TEV2xStrepII}	*S.* Typhimurium	6
pT12_FliOPQR_FlhB_CLOSG	pT12	FliOPQR FlhB^{TEV2xStrepII}	*C. sporogenes*	6
pT12_FliO_SALTY_PQR_FlhB_CLOSG	pT12	$FliO_{ST}$ FliPQR $FlhB_{Cs}$^{TEV2xStrepII}	*C. sporogenes* and *S.* Typhimurium	6
pKD3		(Datsenko and Wanner 2000)		6
pKD46		(Datsenko and Wanner 2000)		6
pCP20		(Datsenko and Wanner 2000)		6
pT12_W1	pT12	FliOPQR^{TEV2xStrepII}	*E. coli* W	6
pT12_W2	pT12	FliOPQR^{TEV2xStrepII} FliQ A41W	*E. coli* W	6
pT12_W3	pT12	FliOPQR^{TEV2xStrepII} FliQ A41D	*E. coli* W	6
pT12_W4	pT12	FliOPQR^{TEV2xStrepII} FliQ A41R	*E. coli* W	6
pT12_W5	pT12	FliOPQR^{TEV2xStrepII} FliQ ΔA41	*E. coli* W	6
pT12_W6	pT12	FliOPQR^{TEV2xStrepII} FliQ T42W	*E. coli* W	6
pT12_W7	pT12	FliOPQR^{TEV2xStrepII} FliQ T42D	*E. coli* W	6

pT12_W8	pT12	FlhB^{TEV2xStrepII}	*E. coli* W	6
pBAD_W1	pBAD	FlhB^{TEV2xStrepII}	*E. coli* W	6
pBAD_W2	pBAD	FlhB^{TEV2xStrepII} G132A	*E. coli* W	6
pBAD_W3	pBAD	FlhB^{TEV2xStrepII} G132D	*E. coli* W	6
pBAD_W4	pBAD	FlhB^{TEV2xStrepII} Δ133-135	*E. coli* W	6
pBAD_W5	pBAD	FlhB^{TEV2xStrepII} Δ118-120	*E. coli* W	6
pBAD_W6	pBAD	FlhB^{TEV2xStrepII} Δ115-120	*E. coli* W	6
pBAD_W7	pBAD	FlhB^{TEV2xStrepII} Δ113-137	*E. coli* W	6
pBAD_W8	pBAD	FlhB^{TEV2xStrepII} R135A	*E. coli* W	6
pBAD_W9	pBAD	FlhB^{TEV2xStrepII} L127A	*E. coli* W	6
pBAD_W10	pBAD	FlhB^{TEV2xStrepII} L127D	*E. coli* W	6
pBAD_W11	pBAD	FlhB^{TEV2xStrepII} L127A M136A	*E. coli* W	6
pBAD_W12	pBAD	FlhB^{TEV2xStrepII} L127D M136D	*E. coli* W	6
pT12_V1	pT12	FliOPQR^{TEV2xStrepII} FliQ A41W	*V. mimicus*	6
pT12_V2	pT12	FliOPQR^{TEV2xStrepII} FliQ A41D	*V. mimicus*	6
pET28_FlhA	pET28	FlhA^{TEV2xStrepII}	*S.* Typhimurium	7
pET28_FlhA_4C	pET28	FlhA^{TEV2xStrepII} D170C, R175C, K423C, D553C	*S.* Typhimurium	7
pET28_FlhA_takaii	pET28	FlhA^{TEV2xStrepII}	*T. takaii*	7
pET28_FlhA_takaii_2C	pET28	FlhA^{TEV2xStrepII} D162C, R178C	*T. takaii*	7
pET28_FlhA_2C	pET28	FlhA^{TEV2xStrepII} D170C, R175C	*S.* Typhimurium	7
pT12_FlhA$_{ST}$-GFP	pT12	FlhA-GFP^{TEV2xStrepII}	*S.* Typhimurium	7
pT12_FlhA$_{Ct}$-GFP	pT12	FlhA-GFP^{TEV2xStrepII}	*C. trachomatis*	7
pT12_FlhA$_{Tt}$-GFP	pT12	FlhA-GFP^{TEV2xStrepII}	*T. takaii*	7
pT12_FlhA$_{Mx}$-GFP	pT12	FlhA-GFP^{TEV2xStrepII}	*M. xanthus*	7
pT12_FlhA$_{Vm}$-GFP	pT12	FlhA-GFP^{TEV2xStrepII}	*V. mimicus*	7
pT12_FlhA$_{Vp}$-GFP	pT12	FlhA-GFP^{TEV2xStrepII}	*V. parahaemolyticus*	7
pT12_SctV$_{Ye}$-GFP	pT12	SctV-GFP^{TEV2xStrepII}	*Y. enterocolitica*	7
pT12_SctV$_{Sf}$-GFP	pT12	SctV-GFP^{TEV2xStrepII}	*S. flexneri*	7
pT12_SctV$_{Ps}$-GFP	pT12	SctV-GFP^{TEV2xStrepII}	*P. savastanoi*	7
pT12_SctV$_{Ct}$-GFP	pT12	SctV-GFP^{TEV2xStrepII}	*C. trachomatis*	7
pT12_SctV$_{Mx}$-GFP	pT12	SctV-GFP^{TEV2xStrepII}	*M. xanthus*	7
pT12_SctV$_{ST1}$-GFP	pT12	SctV-GFP^{TEV2xStrepII}	*S.* Typhimurium	7
pT12_SctV$_{ST2}$-GFP	pT12	SctV-GFP^{TEV2xStrepII}	*S.* Typhimurium	7
pT12_SctV$_{Sf}$-GFP M5	pT12	SctV-GFP^{TEV2xStrepII} E502A, K504A, R545A, R560A, R577A	*S. flexneri*	7
pT12_FlhA$_{ST}$TM-GFP	pT12	FlhA-GFP^{TEV2xStrepII} 1-342	*S.* Typhimurium	7
pT12_FliF-GFP	pT12	FliF-GFP^{TEV2xStrepII}	*S.* Typhimurium	7

pT12_FlhA$_{Ct}$-tevGFP	pT12	FlhA-^{TEV}GFP^{TEV2xStrepII}	*C. trachomatis*	7
pT12_FlhA$_{Ct}$	pT12	FlhA^{TEV2xStrepII}	*C. trachomatis*	7
pT12_FlhA$_{Vp}$-tevGFP	pT12	FlhA-^{TEV}GFP^{TEV2xStrepII}	*V. parahaemolyticus*	7
pT12_FlhA$_{Vp}$	pT12	FlhA^{TEV2xStrepII}	*V. parahaemolyticus*	7
pT12_SctV$_{Ye}$-tevGFP	pT12	SctV-^{TEV}GFP^{TEV2xStrepII}	*Y. enterocolitica*	7
pT12_SctV$_{Ye}$	pT12	SctV^{TEV2xStrepII}	*Y. enterocolitica*	7
pT12_SctXYV$_{Ye}$-tevGFP	pT12	SctXYV-^{TEV}GFP^{TEV2xStrepII}	*Y. enterocolitica*	7
pT12_FliPQR-FlhBA$_{Vp}$	pT12	FliPQR-FlhBA-^{TEV}GFP^{TEV2xStrepII}	*V. parahaemolyticus*	7
pT12_FlgD	pT12	FlgD	*V. parahaemolyticus*	7
pT12_FlhAc$_{Vp}$-tevGFP	pT12	FlhA$_C$-^{TEV}GFP^{TEV2xStrepII} 337-696	*V. parahaemolyticus*	7

Table 2.3 Table of plasmids.

Strain	Reference
BL21 (DE3)	(Wood 1966)
Mt56	(Baumgarten et al. 2017)
E. coli W	obtained from DSMZ (DSM 1116)
E. coli W ΔFliOPQR	this study
E. coli W ΔFlhB	this study

Table 2.4 Table of strains.

2.3.2 DNA amplification

Polymerase chain reaction (PCR) was performed using the Q5 kit (New England Biolabs (NEB)) following the manufacturer's recommendations. Purified plasmid or genomic DNA template was mixed with forward and reverse primers, water and PCR master mix (Table 2.5). The reaction mixture was then transferred to a T100 thermal cycler (Bio-Rad) and heated (Table 2.6). When using plasmid DNA as the template, 1 µl DpnI was subsequently added to the reaction mixture and the solution was incubated overnight at 37 °C. Primers for amplification of plasmids had no overhangs and primers for amplification of inserts had 15 to 20 nucleotide 5' overhangs complementary to the

cloning site of the plasmid of interest. Primers for site-directed mutagenesis were complementary to the sequence immediately upstream or downstream of the mutation site and had a 10 to 20 nucleotide 5' overhang complementary to the desired mutant sequence.

Solution	Volume (µl)	Final concentration
Q5 high-fidelity 2X master mix (NEB)	25	1X
10 µM forward primer	2.5	0.5 µM
10 µM reverse primer	2.5	0.5 µM
template DNA (5-20 ng/µl)	1	0.1-0.4 ng/µl
water	19	

Table 2.5 PCR reaction composition.

Stage	Temperature (°C)	Time (seconds)	Cycles
Initial denaturation	98	30	1
Denaturation	98	8	25
Annealing	65	25	
Elongation	72	25 per kilobase	
Final elongation	72	420	1

Table 2.6 PCR protocol.

2.3.3 Agarose gel electrophoresis

50 µl of PCR product were mixed with 10 µl of 6X purple DNA loading dye (NEB) and loaded on a 50 ml 0.7 to 1% (w/v) agarose gel in TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.5) containing 1X SYBR Safe (Invitrogen). Gels were run for 20 to 30 minutes at 100 V in a Sub-Cell GT agarose gel electrophoresis system (Bio-Rad) using TAE as the running buffer. Bands were visualised on a Safe Imager 2.0 blue light transilluminator (Fisher Scientific) and cut out of the gel. When the DNA was not used downstream, gels were imaged on a Gel Doc (Bio-Rad). DNA was extracted from cut bands using the QIAquick gel extraction kit (Qiagen) according to the manufacturer's recommendations.

2.3.4 Cloning: Gibson assembly and in vivo assembly (IVA)

New constructs were generated by assembly of PCR products using overlapping regions of 15 to 20 nucleotides at both ends of the PCR products. Gibson assembly was performed by mixing 1 µl of PCR product of both plasmid and insert purified by agarose gel electrophoresis with 3 µl of water and 5 µl of NEBuilder HiFi DNA assembly master mix (NEB) and incubating the reaction mixture at 50 °C for 1 hour. 5 µl of the product was then introduced into chemically competent cells by transformation (section 2.3.6). Alternatively, 1 µl of unpurified PCR product of both plasmid and insert were mixed and used for transformation of competent cells, in order to allow for assembly of the plasmid *in vivo*.

Constructs for expression of strep tagged proteins were cloned into a pT12 vector and constructs for co-expression of proteins together with the pT12 vector were cloned into a pBAD vector (Figure 2.1). The pT12 vector was provided by collaborators in Samuel Wagner's laboratory (Dietsche et al. 2016).

2.3.5 Cloning: Source of genomic DNA

Genomic DNA from *S.* Typhimurium (Le Minor and Popoff 1987) was obtained from the Tang laboratory at the Sir William Dunn School and genomic DNA from *Vibrio mimicus* (strain DSM 19130) (Davis et al. 1981), *Pseudomonas savastanoi* (DSM 21482) (Hayward and Waterston 1965), *Thermosulfidibacter takaii* (DSM 17441) (Nunoura et al. 2008), *Clostridium sporogenes* (DSM 795) (Bergey et al. 1923), *Myxococcus xanthus* (DSM 16526) (Beebe 1941), *Vibrio parahaemolyticus* (DSM 10027) (Sakazaki et al. 1963) and *Chlamydia trachomatis* (DSM 19411) (Everett et al. 1999) were obtained from DSMZ. The virulence plasmids of *Shigella flexneri* (Castellani and Chalmers 1919) and *Yersinia enterocolitica* (Frederiksen 1964) were obtained from Dr Patrizia Abrusci.

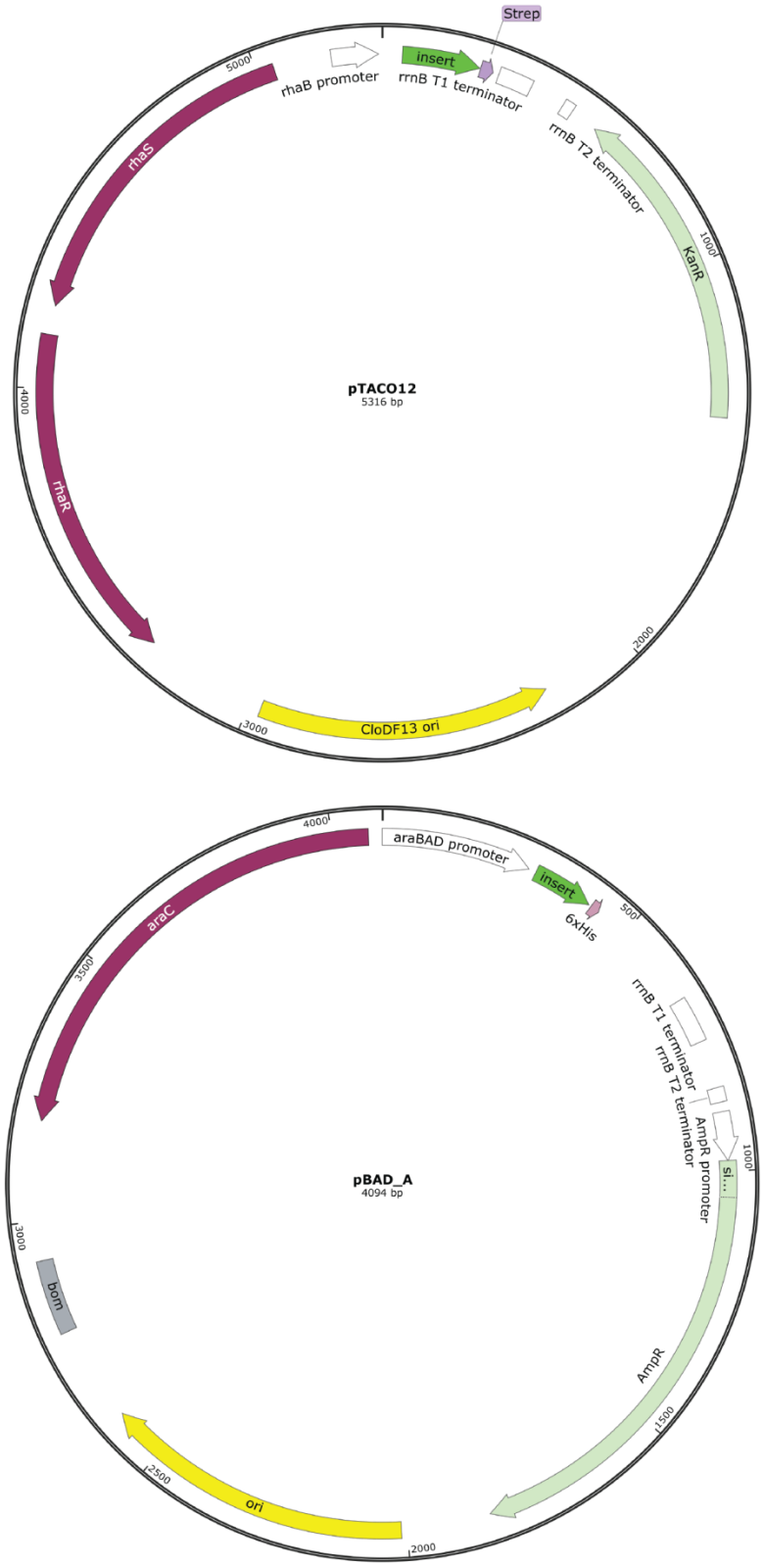


Figure 2.1 plasmid maps of the pT12 (pTACO12) and pBAD vectors.

2.3.6 Bacterial cell transformation

E. coli cells were made chemically competent using calcium/rubidium chloride (Table 2.7) and transformation was carried out using heat shock (Table 2.8).

1.	Grow a 5 ml overnight culture of the strain of interest in 2YT media
2.	Inoculate 60 ml of 2YT media with 0.3 ml of the overnight culture in a 250 ml flask until the OD600 reaches 0.4 to 0.6
3.	Incubate the culture on ice for 10 minutes
4.	Spin the culture (5,000x g, 10 minutes, 4 °C) and discard the supernatant
5.	Resuspend the pelleted cells in 20 ml RF1 and incubate on ice for 15 to 60 minutes
6.	Spin the culture (5,000x g, 10 minutes, 4 °C) and discard the supernatant
7.	Resuspend the pelleted cells in 4.8 ml RF2 and incubate on ice for 15 minutes
8.	Aliquot the cell suspension in 50 µl aliquots into tubes on dry ice and store at -80 °C

Table 2.7 Preparation of chemically competent cells.

Step		Time (minutes)	
		Cloning or multiple plasmids	Single plasmid
1.	Add 1 µl of plasmid (10 to 500 ng/µl) or 2 µl of IVA mixture or 5 µl of Gibson assembly product to one aliquot of competent cells		
2.	Incubate on ice	30	5
3.	Heat shock in a water bath at 42 °C	1	1
4.	Incubate on ice	2	2
5.	Add 1 ml (cloning) or 100 µl (purified plasmid) of 2YT media to the mixture		
6.	Incubate at 37 °C in a shaking incubator	60	0
7.	Spread on an LB agar plate containing the appropriate antibiotic		
8.	Incubate overnight at 37 °C		

Table 2.8 Transformation of chemically competent cells.

2.3.7 Generation of bacterial chromosomal deletion mutants

Chromosomal deletion mutants were made using the λ Red system (Datsenko and Wanner 2000). The principle of this system is the recombination between a PCR product with 50 bp flanking regions homologous to the sequences upstream and downstream of the desired deletion, which is catalysed by the Red recombinase expressed from plasmid pKD46 (Datsenko and Wanner 2000). A selectable marker is amplified with primers containing overhangs to introduce the desired flanking homology regions, allowing its insertion in the genome in place of the target sequence. Following successful gene replacement, the inserted marker, which is flanked by FRT sites, is removed by expressing the FLP recombinase from a helper plasmid.

E. coli strain W (DSM 1116) (Castellani and Chalmers 1919) carrying plasmid pKD46 was made electrocompetent (Table 2.9) and transformed with a PCR product containing the chloramphenicol resistance cassette of plasmid pKD3 (Datsenko and Wanner 2000) flanked by sequences homologous to the sequences upstream and downstream of the desired target sequence for deletion. Transformation was carried out using an E. coli Pulser transformation apparatus (Bio-Rad) according the manufacturer's recommendations. Cells resulting from electroporations with a time constant between 4 and 6 seconds were grown in 1 ml 2YT media for 2 hours and plated on LB agar plates containing 20 μg/ml chloramphenicol. Selected colonies were made chemically competent (Table 2.7), transformed with plasmid pCP20 (Datsenko and Wanner 2000), and plated on LB agar plates containing ampicillin (100 μg/ml). The plates were incubated at 30 °C and picked colonies grown overnight in 2YT media at 37 °C. The origin of replication of pCP20 and pKD46 is temperature sensitive, allowing the plasmids to be cured by growth at 37 °C.

1.	Grow a 5 ml overnight culture of *E. coli* W carrying pKD46 in 2YT media supplemented with carbenicillin (50 µg/ml) at 30 °C
2.	Inoculate 25 ml 2YT media supplemented with ampicillin (100 µg/ml) in a 125 ml flask with the overnight culture to a final OD600 of 0.05 and grow at 30 °C in a shaking incubator
3.	When the cells have reached an OD600 of 0.15, add arabinose to a final concentration of 0.3%
4.	When the cells have reached an OD600 of 0.4 to 0.6, chill the culture on ice for 15 minutes
5.	Spin the culture (5,000x g, 10 minutes, 4 °C) and discard the supernatant
6.	Resuspend the pelleted cells in 10 ml ice-cold sterile 10% (v/v) glycerol
7.	Repeat step 5 and 6 two more times
8.	Spin the suspension (5,000x g, 10 minutes, 4 °C) and discard the supernatant
9.	Resuspend the pelleted cells in 1 ml ice-cold sterile 10% (v/v) glycerol
10.	Aliquot the cell suspension in 50 µl aliquots into tubes on dry ice and store at -80 °C

Table 2.9 Preparation of electrocompetent cells for lambda Red recombineering.

2.4 Cell biology

2.4.1 Motility assays

E. coli W mutant strains generated by deleting FlhB or FliOPQR (section 2.3.7) were transformed (section 2.3.6) with plasmids containing wild type or mutant sequences of the deleted proteins. FlhB constructs were in the vector pBAD and FliOPQR constructs were in the vector pT12. 5 ml cultures of all constructs were grown overnight in 2YT media containing the appropriate antibiotic. On the following day, a soft agar plate was prepared using 2YT media containing 0.28% (w/v) agar, the appropriate antibiotic and 0.1% (w/v) arabinose for pBAD constructs or 0.5% (w/v) rhamnose monohydrate for pT12 constructs. 3 µl of each overnight culture were injected in separate spots and the plates were incubated overnight at room temperature. Experiments were performed in triplicate. The diameter of the circle surrounding each inoculation spot of the mutant constructs was compared to the wild type on the same plate.

2.4.2 Fluorescence microscopy

A 5 ml culture of BL21 transformed with the GFP fusion construct of interest in the pT12 vector was grown overnight in 2YT media containing 0.01% (w/v) rhamnose monohydrate and 30 µg/ml kanamycin. The next morning, 8 µl of the culture were applied to a glass slide and covered with a coverslip. The cells were imaged on a Zeiss 880 inverted microscope equipped with a plan-apochromat 63x/1.4 NA objective using an Airyscan detector. GFP was excited using a laser (488 nm).

2.5 Biochemistry

2.5.1 Expression and purification of export gate complexes for native mass spectrometry

The following protocol was used for preparation of protein samples for native mass spectrometry (nMS) in chapter 3: 100 ml of TB supplemented with kanamycin (30 µg/ml) were inoculated with *E. coli* Mt56 (Baumgarten et al. 2017) carrying the plasmid of interest in the morning. In the evening, up to 16 2.5 L Tunair flasks (Sigma) were filled with 1 L of TB supplemented with kanamycin (30 µg/ml) each. The media was inoculated by diluting the preculture to a final OD600 of 0.02. The cells were grown at 37 °C and 180 rpm overnight without inducing expression, as collaborators in Samuel Wagner's laboratory reported that expression of export apparatus proteins was toxic. The cells were harvested by spinning the culture (5000x g, 10 minutes, 4 °C). The supernatant was discarded and the pellet was washed in PBS. TBS (17.5 ml per litre of the original culture) containing a protease inhibitor cocktail, lysozyme (1 mg/ml) and Dnase I (10 µg/ml) was added to the cell pellet and the pellet was resuspended by shaking and pipetting up and down. The cell suspension was lysed in three passes through an Emulsiflex C5 (Avestin) homogenizer at 15,000 psi. The lysate was spun (24,000x g, 20 minutes, 4 °C) and membranes were pelleted by spinning the resulting supernatant in an ultracentrifuge (200,000x g, 60 minutes, 4 °C) using a 45Ti rotor (Beckman). The supernatant was discarded and 5 ml TBS was added to each pellet. The pellets were disrupted using a spatula and transferred to a dounce homogenizer using a positive displacement pipette. The membranes were homogenized with the loose pestle and then with 10 strokes of the tight pestle. The protein concentration in the homogenized membranes was

measured using a BCA assay kit (ThermoFisher Scientific) and the suspension was diluted to a final protein concentration of 6 mg/ml in a glass bottle.

DDM was added from a 10% (w/v) stock solution to a final concentration of 1% (w/v) and the solution was stirred for one hour at 4 °C. The dissolved membranes were spun again (200,000x g, 60 minutes, 4 °C) and a 5 ml StrepTactin Superflow high capacity column (IBA), binding the strep tag (Schmidt and Skerra 2007), was equilibrated with TBS containing 0.01% (w/v) DDM. The supernatant of the spun membranes was transferred to a glass bottle and recirculated through the StrepTactin column for two to three hours. The column was washed until it appeared white using TBS containing 0.1% (w/v) DDM and then the same buffer was recirculated through the column overnight. The next morning, the column was transferred to an AKTA purifier. The running buffer was TBS containing 0.016% (w/v) DDM using a flow rate of 1 ml/min. The column was washed until the UV absorbance at 280 nm in the flow through stabilized and then the protein was eluted using 10 ml TBS containing 0.016% DDM and 10 mM d-desthiobiotin. The peak fractions were concentrated to 500 µl in a concentrator with a 100 kDa MWCO (Vivaspin). Before using the concentrator, it was first washed with milli-Q water and then TBS containing 0.016% DDM. The eluate was concentrated by spinning the concentrator at 750x g for 2 minutes at 4 °C repeatedly until the volume reached 0.5 ml. After each spin the solution was mixed by pipetting up and down.

A Superdex 200 10/300 column (GE) was equilibrated in HBS containing 0.016% (w/v) DDM on an AKTA purifier and the concentrated eluate was injected into a washed 0.5 ml loop. The sample was injected into the column and eluted with HBS at a flow rate of 0.5 ml/min.

2.5.2 Purification of FliP

FliP was expressed with an N-terminal his-tag between the signal peptide and the N-terminus of the mature protein. The gene was part of an operon with FliO in a pT12 plasmid. The protein was expressed and the membranes were prepared following the same protocol as was used for other proteins (section 2.5.1). The membranes (7.9 grams) were homogenised in 72 ml binding buffer (50 mM Tris, pH 8, 300 mM NaCl, 20 mM

imidazole, 5% (v/v) glycerol) and 8 ml of 10% (w/v) DDM was added. After stirring for 30 minutes at 4 °C, unsolubilised material was removed by centrifugation (75,000x g, 20 minutes, 4 °C) and the supernatant was applied to a 5 ml Ni-NTA column (Qiagen). The column was connected to an AKTA purifier (GE) and washed with binding buffer containing 0.1% (w/v) DDM until the UV absorbance of the flow through stabilized. Protein was eluted in a 50 to 400 mM imidazole gradient in binding buffer containing 0.1% (w/v) DDM. Fractions were analysed by SDS-PAGE and the fractions containing FliP were pooled and concentrated. The protein was then subjected to SEC on a Superdex 200 increase column using 20 mM Tris, pH 8, 300 mM NaCl, 2 mM EDTA and 0.1% (w/v) DDM as the running buffer. Fractions containing FliP but not FliO by SDS-PAGE were pooled and concentrated.

2.5.3 Modified protocol for expression and purification of Strep-tagged membrane protein complexes

The following protocol was used for preparation of protein samples in chapters 4 to 6: The earlier protocol for purification of membrane proteins for nMS (section 2.5.1) was modified and simplified: BL21 cells were used instead of Mt56 cells, as the genome sequence of Mt56 showed that the difference between BL21 (DE3) (Wood 1966) and Mt56 was a mutation in the T7 polymerase, which is not relevant for expression using the pT12 plasmid (Baumgarten et al. 2017). Cells were grown at 200 rpm rather than 180 rpm shaking in order to improve oxygenation and achieve higher growth rates and final cell densities. Expression was induced with rhamnose as cells continued growing when induced (Figure 4.20). The ultracentrifugation time and speed for isolating membranes from lysate were increased in order to collect more membrane and achieve higher yields. Detergent was added in a specific ratio to the wet weight of the membranes instead of measuring membrane protein concentration in the homogenised membranes in order to simplify and speed up the protocol. The solubilised membranes were clarified by centrifugation at a lower speed and then filtered in order to avoid using an ultracentrifuge, which would require a longer preparation time. A StrepTactin resin from GE rather than IBA was used.

E. coli BL21 cells carrying the plasmid of interest were streaked on an LB agar plate containing the appropriate antibiotic(s) and grown overnight. The cells were resuspended using 2YT media and transferred to 250 ml flask containing 100 ml of 2YT media supplemented with antibiotic. In the evening, up to 16 2.5 L Tunair flasks were filled with 1 L of TB media supplemented with kanamycin (60 µg/ml) and/or ampicillin (100 µg/ml) as appropriate. The media was inoculated by adding the preculture to a final OD600 of 0.02. After 12 hours of growth at 37 °C, 200 rpm, the cells were harvested (5000x g, 15 minutes, 4 °C). TBS containing a protease inhibitor cocktail was added to the cell pellet and the cells were resuspended using a dounce homogeniser pestle. The cell suspension was passed through an Emulsiflex C5 (Avestin) homogeniser at 15,000 psi three times and the lysate was clarified by centrifugation (24,000x g, 20 minutes, 4 °C). The supernatant was spun in an ultracentrifuge using a 45Ti rotor (Beckman) (235,000x g, 60 to 180 minutes, 4 °C). The membrane pellet was transferred to a dounce homogeniser and homogenised using 8 ml of TBS per gram of wet membrane. 1 ml of a 10% (w/v) stock solution of detergent (DDM or LMNG) was added per gram of membrane, resulting in a final concentration of 1% (w/v). After gentle stirring for one hour at 4 °C the solution was clarified by centrifugation (75,000x g, 20 minutes, 4 °C) and the supernatant was filtered through a 0.45 µm filter. A 5 ml StrepTrap HP column (GE) was equilibrated with TBS containing 0.02% (w/v) DDM or 0.01% (w/v) LMNG. The filtered supernatant was applied to the column at a flow rate of 2 ml/min and the column was washed with at least 50 ml of TBS containing 0.02% (w/v) DDM or 0.01% (w/v) LMNG. The protein was eluted with 10 ml TBS containing 0.02% (w/v) DDM or 0.01% (w/v) LMNG and 10 mM d-desthiobiotin. The eluate was concentrated to 500 µl in a concentrator with a 100 kDa MWCO concentrator (Amicon). A Superdex 200 increase 10/300 (GE) or Superose 6 increase 10/300 column (GE) was equilibrated in TBS containing 0.02% (w/v) DDM or 0.01% (w/v) LMNG on an AKTA purifier (GE) and the sample was injected from a pre-washed 0.5 ml loop. The protein was eluted with 30 ml TBS buffer containing 0.02% (w/v) DDM or 0.01% (w/v) LMNG at a flow rate of 0.7 ml/min. Where small alterations to this protocol were made, these are described in the relevant results chapters.

2.5.4 Detergent screen for co-purification of FlhB with FliPQR

100 ml TB supplemented with kanamycin (30 μg/ml) were inoculated with *E. coli* Mt56 carrying a pT12 plasmid expressing the *S.* Typhimurium FliOPQR operon with a C-terminal strep tag on FliR and a pBAD plasmid expressing the *S.* Typhimurium FlhBA operon with a C-terminal his tag on FlhA in the morning. In the evening, 4 L TB media in 2.5 L Tunair flasks were inoculated with the preculture to a final OD600 of 0.02 and the cultures were grown at 37 °C, 180 rpm. After 30 minutes rhamnose monohydrate and arabinose were each added to a final concentration of 0.02%. After 12 hours, additional rhamnose monohydrate and arabinose were added to a final concentration of 0.06%. After an additional 3 hours cells were harvested by centrifugation (5000x g, 10 minutes, 4 °C). 100 ml of TBS containing a protease inhibitor cocktail, lysozyme (1 mg/ml) and Dnase I (10 μg/ml) were added and the cells were resuspended by pipetting. The cells were lysed by passing the suspension through an Emulsiflex C5 (Avestin) homogeniser at 15,000 psi three times. The lysate was clarified by centrifugation (24,000x g, 20 minutes, 4 °C) and the supernatant was spun in an ultracentrifuge using a 45Ti rotor (Beckman) (200,000x g, 180 minutes, 4 °C). The pelleted membrane fraction was transferred to a dounce homogenizer and homogenized in 80 ml TBS. The homogenate was split into eight tubes and 1 ml of detergent was added to each tube from a 10% (w/v) stock solution. The detergents used were DDM, GDN, LMNG, thesit, OG, CHAPS, a mixture of 0.5 ml each of 10% (w/v) DDM and GDN and a mixture of 0.5 ml each of 10% (w/v) DDM and GDN plus 0.293 ml 7.5 M ammonium acetate. 1 ml of 50% (v/v) streptactin XT beads (IBA) were added to each tube and all tubes were rotated at 4 °C overnight. Beads were pelleted (3,200x g, 10 minutes, 4 °C), the supernatant was discarded and the beads were resuspended in TBS containing 0.01% (w/v) detergent (LMNG, GDN, thesit) or 0.02% (w/v) detergent (DDM) or 1% (w/v) detergent (OG, CHAPS). Beads resulting from solubilisation in DDM GDN mixtures were washed only with GDN in TBS. This step was repeated two times. Protein was eluted by resuspending the beads in 0.75 ml 50 mM biotin in TBS. Beads were pelleted (3,200x g, 10 minutes, 4 °C), and the supernatant was kept. The step was repeated and the two supernatants were pooled. After SDS-PAGE analysis the DDM, GDN, LMNG, thesit and OG samples were submitted for proteomics in the Dunn School proteomics facility.

2.5.5 Co-purification of *Salmonella* FliPQR and FlhB

C-terminally Strep tagged FliR as part of the FliLMNOPQR operon in the pT12 vector was co-expressed with N-terminally His tagged FlhB in a pBAD plasmid. The two plasmids were co-transformed into BL21 cells. In order to avoid loss of the pBAD plasmid during overnight expression, carbenicillin (100 µg/ml) was used for maintaining the plasmid. In addition, the TB media was supplemented with 60 µg/ml kanamycin, 0.1% (w/v) rhamnose monohydrate, 0.1% (w/v) arabinose and 0.05% (w/v) glucose. Glucose represses induction by rhamnose and arabinose due to catabolite repression (Guzman et al. 1995), resulting in autoinduction once the glucose has been fully metabolised. Protein was prepared using a slightly modified version of the standard protocol (section 2.5.3). The cells were grown overnight at 37 °C, 200 rpm and lysed in an emulsiflex C5 homogeniser (Avestin). Membranes were prepared by spinning the clarified lysate (235,000x g, 3 hours, 4 °C). The resulting 14 grams of membranes were homogenised in TBS and 11 ml of a 10% (w/v) LMNG stock solution were added to a final concentration of 1% (w/v). This is a lower ratio of detergent to membrane than used previously. The amount of detergent was reduced compared to the standard protocol in order to avoid "washing off" the FlhB subunit with excess detergent, as it may be stabilised by essential lipids that could be removed by detergent. The solubilised membrane proteins were applied to a 5 ml StrepTrap column (GE) which was washed with TBS containing 0.01% (w/v) LMNG and the purified FliPQR-FlhB complex was eluted using TBS containing 0.01% (w/v) LMNG and 10 mM desthiobiotin. In all subsequent steps, TBS was made without EDTA and supplemented with 0.01% (w/v) LMNG. The protein was concentrated and further purified by SEC on a Superdex 200 increase 10/300 column (GE)

2.5.6 SDS-PAGE

15% (v/v) polycrylamide gels were made as a 15% (v/v) resolving gel (15% (v/v) Acrylamide/Bis-acrylamide (ratio 19:1), 375 mM Tris, pH 8.8, 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulfate, 0.1% (v/v) tetramethylethylenediamine) underneath a 5.1% (v/v) stacking gel (5.1% (v/v) Acrylamide/Bis-acrylamide (ratio 19:1), 125 mM Tris, pH 6.8, 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulfate, 0.1% (v/v) tetramethylethylenediamine) and 4-20% gradient Mini-PROTEAN precast gels were purchased from Bio-Rad. Loading buffer was added to samples to reach a final concentration of 1X (5X loading buffer: 250

mM Tris, pH 6.8, 10% (w/v) SDS, 50% (v/v) glycerol, 0.5% (w/v) Bromophenol Blue, 100 mM DTT). Samples were not heated before running the gel. The gels were run in the Mini-PROTEAN tetra cell (Bio-Rad) in SDS-PAGE running buffer (25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS) at 180 V and stained in Instant*Blue*.

2.5.7 Concentration of proteins

Protein concentration was carried out using 100 kDa MWCO concentrators (Amicon) unless indicated otherwise.

2.5.8 SEC-MALS

SEC was performed using a Superdex 200 10/300 column (GE) or a Superose 6 increase 10/300 column (GE) equilibrated in the relevant buffer. 100 µl of purified protein at 1 mg/ml was injected into the column and the protein was eluted at a flow rate of 0.4 ml/min. A Dawn Heleos-II light-scattering detector (Wyatt Technology) and an Optilab-TrEX refractive index monitor (Wyatt Technology) followed the column in-line. ASTRA 6.1.1.17 was used for analysis assuming a dn/dc of 0.186 ml/g for protein and 0.132 ml/g for detergent. SEC-MALS was performed by Dr Steven Johnson.

2.5.9 Exchange of samples into mass spectrometry compatible solvents

Membrane protein samples purified using the detergent DDM or LMNG were subjected to SEC using TBS containing an nMS compatible detergent as a running buffer. The detergent was used at a concentration twice its CMC (0.5% (w/v) for C8E5, 0.5% (w/v) for C8E4, 0.8% (w/v) for C8E6, 0.08% (w/v) for C10E6, 0.15% (w/v) for C10E9 and 0.006% (w/v) for C12E9). 50 µl of purified membrane protein at a concentration of at least 0.7 mg/ml were injected into a Superdex 200 increase 5/150 column (GE) on an AKTA purifier (GE) equilibrated with running buffer and the protein was eluted using a flow rate of 0.1 ml/minute. Peak fractions were pooled and concentrated to below 100 µl using a 100 kDa MWCO concentrator (Amicon). This step was omitted when the detergent for purification and analysis by nMS was identical.

At the same time, a Bio-Spin 6 column (Bio-Rad) was equilibrated by first spinning the column (1,000x g, 2 minutes, 4 °C). Subsequently, 500 µl of 200 mM ammonium acetate containing the detergent of choice at a concentration of twice its CMC were applied to the top of the column and it was spun again (1,000x g, 1 minute, 4 °C). This was repeated three more times. The concentrated protein sample was then applied to the top of the Bio-Spin 6 column and the column was spun in a clean tube (1,000x g, 4 minutes, 4 °C). The eluted sample was stored on ice and used for nMS analysis as soon as possible.

2.5.10 Native mass spectrometry of membrane proteins

Samples were analysed using the Q-exactive hybrid quadrupole-Orbitrap mass spectrometer (ThermoFisher Scientific) modified for transmission and detection of high m/z ions and set up as published (Gault et al. 2016). The sample was sprayed through a glass capillary with a 1.2 mm outer diameter (Harvard Apparatus) made in-house using a P-97 micropipette puller (Sutter Instrument) and gold-coated with a Polaron SC7680 sputter coater (Quorum Technologies) (Hernandez and Robinson 2007). A voltage of 1.2 kV was applied to the needle and the temperature at the ion source was 40 to 60 °C unless otherwise indicated. The collision gas was Argon at a pressure around 10^{-9} mbar and the collision voltage was between 100 and 200 V. When required, additional collisional energy was used by applying a voltage in the source of the instrument.

Mass spectrometry data were analysed using Xcalibur (ThermoFisher Scientific) and peaks were assigned using UniDec (Marty et al. 2015) or manually by calculating the mass using an Excel (Microsoft) spreadsheet obtained from the Carol Robinson group. When ionised, a protein of mass m will produce a series of peaks corresponding to different charge states whose mass to charge ratio is measured by the mass spectrometer. A charge of +z results from z protons binding to the protein, which then has a mass of m+z and a mass to charge ratio of (m+z)/z. Therefore, the values of both m and z can be deduced by measuring the mass to charge ratio of at least two different charge states from the same series of peaks.

2.5.11 Negative stain EM

Carbon-coated grids (Quantifoil) were glow discharged (15 mA, 20 seconds) in an EM ACE200 vacuum coater (Leica). Subsequently, a 10 µl drop of sample at a protein concentration around 0.01 mg/ml unless otherwise specified was applied to a piece of parafilm and the grid was placed on top of the drop with the carbon side facing down. After two minutes, the grid was blotted and placed on a 20 µl drop of 2% (w/v) uranyl acetate. After 10 seconds the grid was blotted again and left on filter paper to dry with the carbon side facing up. Grids were examined on a T12 microscope (FEI) operated at 120 kV equipped with a OneView CMOS camera (Gatan). Images were acquired at a magnification of 68,000 and a pixel size of 1.73 Å.

2.5.12 Preparation of grids for cryogenic EM

Holey carbon-coated grids (Quantifoil) were glow discharged (15 mA, 20 seconds) in an EM ACE200 vacuum coater (Leica). Subsequently, grids were picked up with tweezers and inserted into a Vitrobot Mark IV (FEI) at 100% humidity at room temperature or 4 °C. A 3 µl drop of sample was applied to the grid; the grid was blotted with filter paper and plunged into liquid ethane cooled by liquid nitrogen. For the exact parameters see Table 2.10 and the relevant results sections. All grids were Au quantifoil grids, except FlhA grids, which were copper grids, sample volume was always 3 µl, blot time was always 3 seconds, wait time was always 5 to 10 seconds except for graphene oxide grids where it was 60 seconds, drain time was always 0 seconds and humidity was always 100%.

Sample	Chapter	Concentration (mg/ml)	Mesh size	Spacing	Temperature
FliPQR + FlhB in LMNG, *S.* Typhimurium	4	0.9	400	1.2/1.3	22 °C
FliPQR in LMNG, *P. savastanoi*	5	1	300	1.2/1.3	22 °C
FliPQR in LMNG, *V. mimicus*	5	1	300	1.2/1.3	4 °C
FliPQR in LMNG, *V. mimicus* + fluorinated FC-8	5	2.7	300	1.2/1.3	4 °C
SctRST in LMNG, *S. flexneri*	5	8.4	300	1.2/1.3	22 °C
FliPQR + FlhB in LMNG, *V. mimicus*	6	2.5	300	1.2/1.3	22 °C
SctV in LMNG, *Y. enterocolitica*	7	1	300	1.2/1.3	4 °C
FlhA in LMNG, *V. parahaemolyticus*	7	0.9	300	2/2	4 °C
SctXYV-GFP in GDN, *Y. enterocolitica*	7	0.1 (graphene oxide coated grid)	300	2/2	4 °C

Table 2.10 Grid preparation parameters of selected samples.

2.5.13 Acquisition and analysis of cryo-EM data

Grids were imaged using a Talos Arctica (FEI) or Titan Krios (FEI) microscope equipped with a Falcon 3 (Gatan) or K2 (Gatan) detector respectively. The pixel size for data acquisition on the Titan Krios microscope was 0.85 Å. Data was acquired automatically with EPU (FEI) using a defocus range of approximately 0.5 to 4 µm. In general, pre-processing of the movies, including motion correction and CTF estimation, and particle picking were carried out in SIMPLE (Reboul et al. 2018). The particles were then subjected to 2D classification in RELION-2.0 (Kimanius et al. 2016) and 2D classes with high resolution features were selected. The particles were then further classified in 3D in RELION-2.0 (Kimanius et al. 2016) and RELION-3.0 (Zivanov et al. 2018) and classes with detailed features, little noise and accurate particle alignment as estimated within RELION were selected. Finally a high resolution volume was calculated using refinement and post-processing in RELION. The exact strategy for data processing varied as technology improved throughout the time in which data were acquired for this study. Where strategies for data processing differed from this general outline, the changes that were made are described in the relevant results chapters. The final high resolution map

was used for building an atomic model in Coot (Emsley et al. 2010). Models were built manually using predicted models of FliP, FliQ and FliR (Ovchinnikov et al. 2015) as a guide. The models were then iteratively improved using phenix.real_space_refine (Afonine et al. 2018) and manual editing in Coot.

3 Understanding the architecture of the core of the export apparatus

3.1 Introduction

As described in chapter 1, the proteins FliPQR (fT3SS) or SctRST (vT3SS) are thought to be components of the export apparatus and are also known as the export gate. FliP and FliR are among the first proteins in the nanomachine to assemble (Jones and Macnab 1990). In the case of the vT3SS SctRST it has been shown that basal bodies assembled without SctRST are not functional and cannot be repaired by expression of SctRST post assembly (Wagner et al. 2010). They have been proposed to be located in a specialised patch of inner membrane at the centre of the basal body (Macnab 2004), but the highly hydrophobic nature of these putative integral membrane proteins has complicated biochemical characterisation (Figure 1.5). In the case of the fT3SS a specific chaperone, FliO, has been described, but when this work was begun it was not known whether FliO is necessary for assembly of FliPQR (Barker et al. 2010). Recently, different stoichiometries have been proposed for the SctRST and FliPQR complexes despite the high sequence homology (Figure 1.7). The aim of this chapter is to establish whether the stoichiometries of the two complexes are truly different and which, if any, model of the complex is correct. In order to minimise artefacts from the recombinant expression system, all constructs were cloned as operons including the genes for all three subunits. The export gate genes are found next to each other in most organisms (Macnab 2004).

3.2 Purification of *S.* Typhimurium FliPQR and *S. flexneri* SctRST

Two export gate homologues were initially chosen for purification, one from the fT3SS (*S.* Typhimurium FliPQR) and one from the vT3SS (*S. flexneri* SctRST). The genes were cloned as an operon in the pT12 vector, which has previously been used to produce the *S.* Typhimurium SPI-1 complex SctRT (Dietsche et al. 2016). In the fT3SS gene cluster, the FliO gene is found immediately upstream of FliP and has been proposed to promote stability of FliP (Barker et al. 2010), creating the possibility that FliO is a chaperone for assembly of the FliPQR/SctRST complex. Therefore, an additional construct including FliO in the operon was cloned (Figure 3.1).

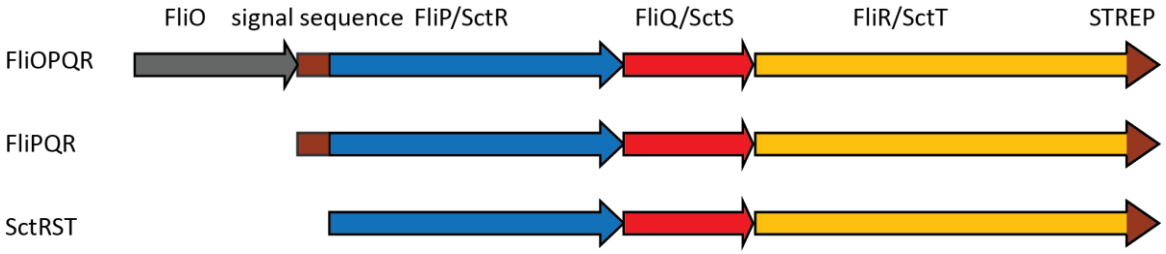


Figure 3.1 Export gate constructs.
Diagram of the operon structure of the export gate constructs in the pT12 vector.

All constructs were expressed and purified following the protocol in section 2.5.1 using the detergent DDM. The export gate complex could be purified using all three constructs, but the sample eluted from the Strep column using the FliPQR construct was lost during concentration before the size exclusion step, likely due to aggregation, and this construct was excluded from further characterisation. The FliOPQR and *S. flexneri* SctRST samples were then further purified by SEC (Figure 3.2A). Interestingly, two peaks were observed in the SEC profile of FliOPQR. SDS-PAGE (Figure 3.2B) showed that the first peak contained two bands corresponding to FliO, consistent with previous analysis of FliO (Ohnishi et al. 1997), in addition to bands for FliPQR, while the second peak contained only FliPQR. The elution volume of the second peak was very similar to the SctRST elution volume and the SEC elution volume of the previously published *S.* Typhimurium SPI-1 SctRT complex (Dietsche et al. 2016), consistent with the second peak corresponding to the assembled FliPQR complex and the first peak corresponding to an assembly intermediate that is still bound to the chaperone FliO.

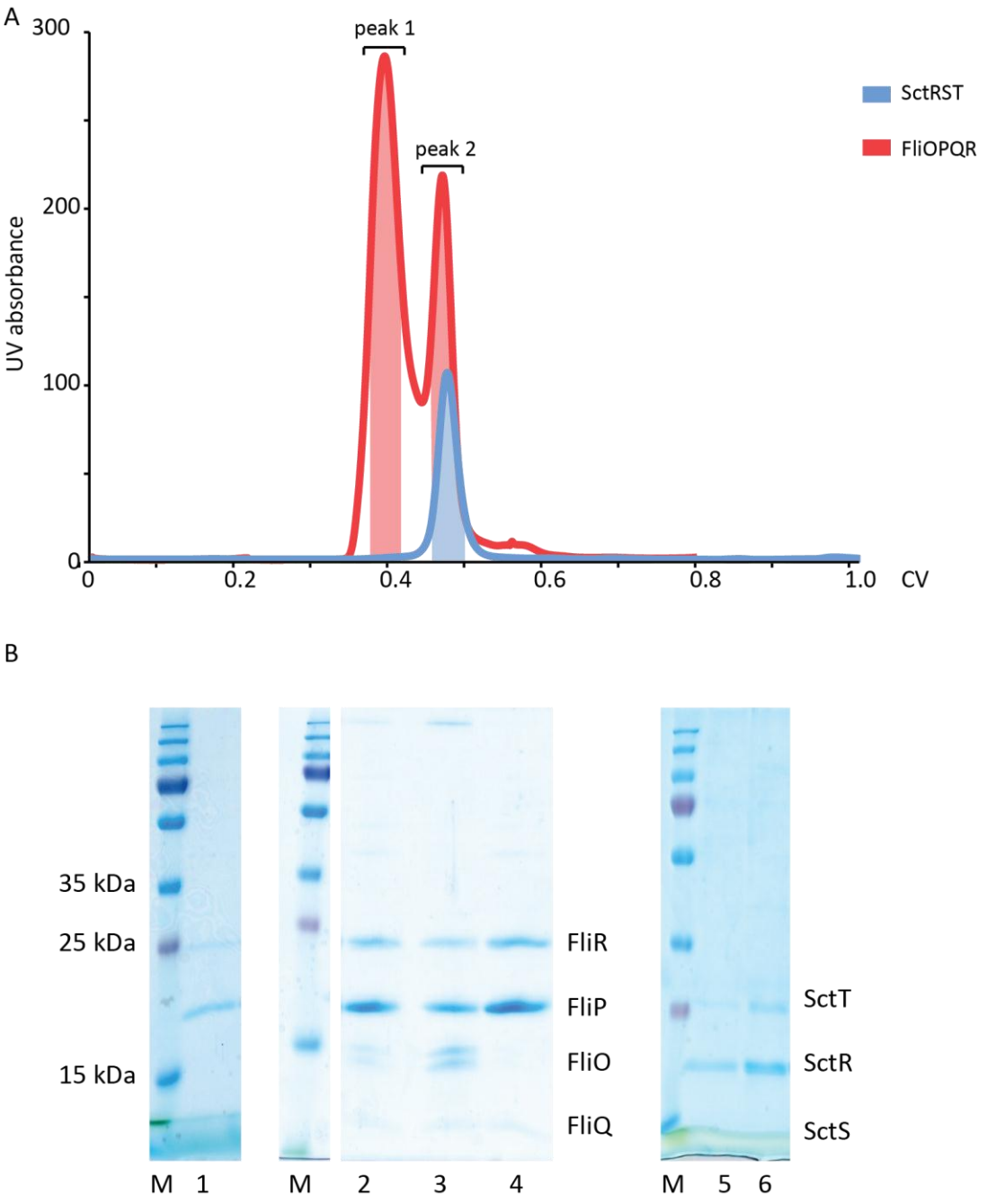


Figure 3.2 Purification of FliPQR and SctRST.
A, SEC of FliOPQR and SctRST. **B**, Representative 15% SDS-PAGE analysis of the result of the purification of a FliPQR, FliOPQR and SctRST homologue. Strep eluate and SEC pools of FliPQR (left), FliOPQR (middle) and SctRST (right). The identities of the proteins were confirmed by proteomics carried out by the Dunn School Advanced Proteomics Facility.

SDS-PAGE also revealed the presence of a weak band around the expected size of FliQ/SctS. SctS had not been observed in the *S.* Typhimurium SPI-1 export gate complex even though the entire SctRST operon was expressed (Dietsche et al. 2016). This

indicated that FliQ may be part of the export gate but binding is weaker in the *S.* Typhimurium SPI-1 homologue. Next, the protein from both the FliOPQR peaks and the SctRST peak was pooled separately and analysed by negative stain EM (Figure 3.3).

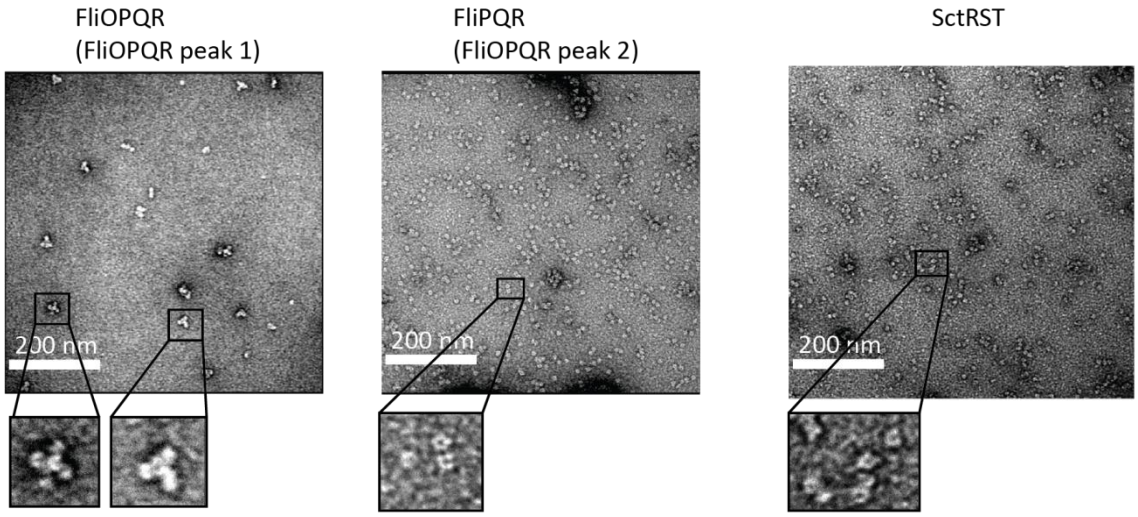


Figure 3.3 Negative stain electron microscopy of the export gate.
Negative stain grids were prepared using samples at 0.003 mg/ml (left), 0.04 mg/ml (middle) and 0.05 mg/ml (right).

The FliPQR and SctRST samples both contained circular particles with a diameter of around 10 nm, similar to the previously published analysis of the *S.* typhimurium SPI-1 SctRT complex (Dietsche et al. 2016). Interestingly, the FliOPQR complex was a heterogeneous mixture of assemblies of three or four rings surrounding a central ring. The similar size of the outer rings and the FliPQR complex suggests that they correspond to each other. The central ring could be formed by the FliO chaperone, but the exact architecture of the FliOPQR complex was not further investigated. After this experiment was performed a study was published suggesting that the central ring in the FliOPQR complex is indeed a FliO ring while the outer rings are FliPQR complexes (Fukumura et al. 2017).

3.3 Stoichiometry of FliPQR

As SEC, negative stain EM and SDS-PAGE analysis all supported successful purification of the *S.* Typhimurium fT3SS export gate FliPQR and the *S. flexneri* vT3SS export gate

SctRST the samples were used for stoichiometric analysis by native mass spectrometry (nMS). The sample buffer was exchanged to 200 mM ammonium acetate (section 2.5.9) and the sample was then analysed using a Q exactive mass spectrometer (section 2.5.10). The sample was sprayed through a glass capillary to which a voltage of 1.2 kV was applied. The source temperature was 40 °C and the collision gas was Argon at a pressure around 10^{-9} mbar. In nMS of membrane proteins collision of the protein-detergent complex with gas molecules is used to dissociate the relatively weakly bound detergent molecules from the protein by collision-induced dissociation (CID) without using so much collisional energy that the protein complex would dissociate.

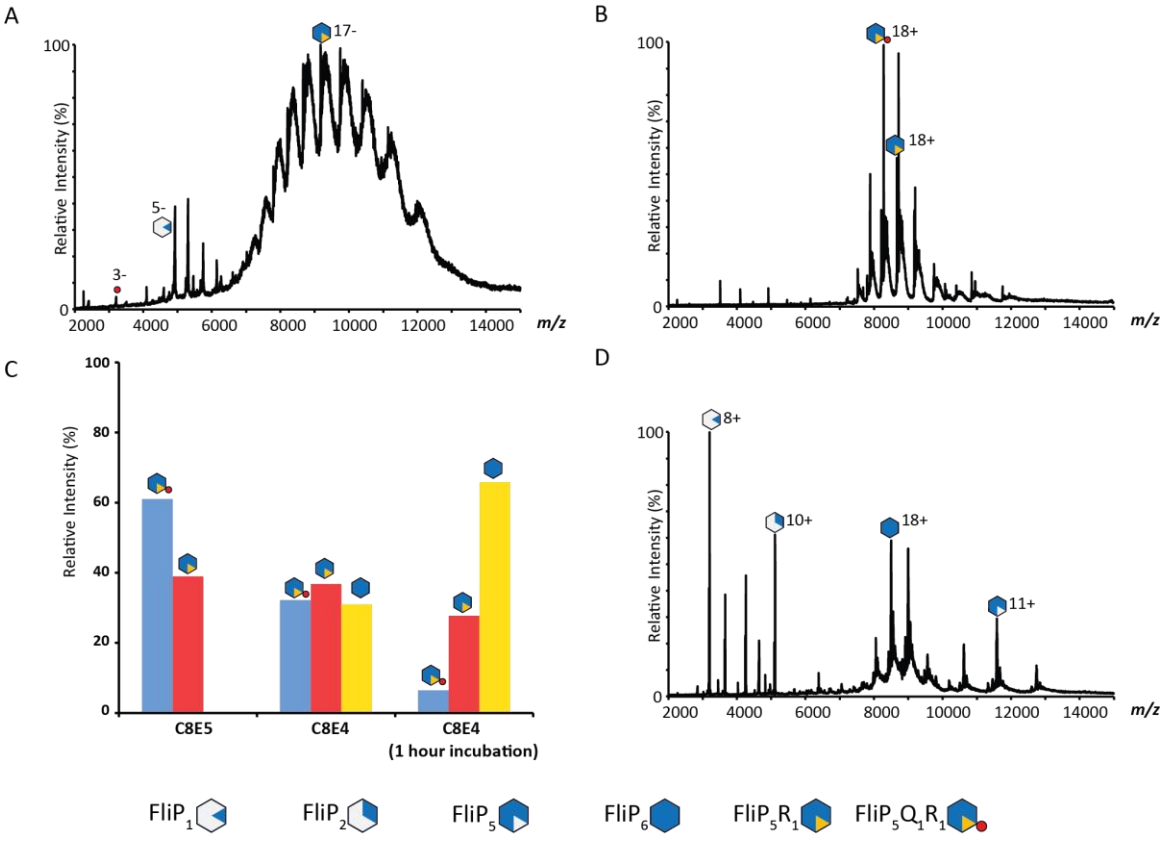


Figure 3.4 nMS of *S.* Typhimurium FliPQR.
A, Mass spectrum of FliPQR in DDM. The sample was analysed in negative mode (Liko et al. 2016) by reversing the polarity on the electrospray needle. The collision voltage was 200 V. **B**, Mass spectrum of FliPQR purified in DDM and exchanged into C8E5. The collision voltage was 150 V. **C**, Relative abundance of different complexes in samples of FliPQR purified in DDM and exchanged into C8E5, C8E4 and a sample exchanged into C8E4 which was left for one hour at 4 °C in between detergent exchange by SEC and buffer exchange using a Bio-Spin column (section 2.5.9). The collision voltage was 150 V. **D**, Mass spectrum of FliP purified in DDM and exchanged into C8E5. The collision voltage was 150 V.

When FliPQR was analysed in DDM, the detergent used for purification, a high voltage (200V) had to be applied in the collision cell to achieve a sufficiently high collisional energy to strip the DDM micelle from the protein. It was possible to observe peaks corresponding to the mass of a $FliP_5R_1$ complex and monomeric FliP and FliQ, which most likely dissociated in the mass spectrometer by CID (Figure 3.4A). Additionally a large part of the spectrum remained unresolved as can be seen in the large, broad peaks to the right of each $FliP_5R_1$ peak indicating the presence of larger species that could potentially include a FliPQR complex.

One reason for why the complex could not be fully resolved could be that DDM binds the complex too tightly, preventing complete stripping of the micelle from the complex in the mass spectrometer. A range of detergents was screened with the aim of finding a detergent in which the intact FliPQR complex could be resolved at a low collision voltage (summarised in Table 3.1). The most stabilising detergent, in which a species including the more fragile FliQ subunit could be observed and in which very little dissociation of monomeric subunits was observed, was C8E5 (Figure 3.4B). This revealed that the complex in this sample had a stoichiometry of five FliP, one FliQ and one FliR. Additionally, FliP hexamers were observed in two detergents, C8E4 and C10E9, both of which also caused partial dissociation of the complex as indicated by the presence of FliP monomers (Table 3.1).

Interestingly, when the FliP monomer peaks were inspected more closely, two peaks were observed corresponding to FliP with different mature sequences (Figure 3.5A). FliP contains a cleaved signal sequence and the cleavage site has been studied previously using N-terminal sequencing (Ohnishi et al. 1997) and was found to be between residues Ala21 and Gln22. However, the majority species observed by mass spectrometry was around 100 Da heavier than expected, consistent with a formylated FliP (21-245). Formylated peptides are not reactive to N-terminal sequencing. It is possible that the smaller FliP (22-245) is a product of degradation or alternative cleavage by the signal peptidase due to the ambiguous signal peptide cleavage site (Figure 3.5B).

Sample	Detergent	Major observed species	Voltage (V)
FliPQR	DDM	$FliP_5R_1$	200
FliPQR	C12E9	not resolved	200
FliPQR	C10E9	$FliP_5R_1$, $FliP_4R_1$, $FliP_6$ and $FliP_1$	200
FliPQR	C10E6	sample aggregated	n/a
FliPQR	C8E4	$FliP_5Q_1R_1$, $FliP_5R_1$, $FliP_6$ and $FliP_1$	150
FliPQR	C8E5	$FliP_5Q_1R_1$, $FliP_5R_1$	150
FliPQR	C8E6	$FliP_5Q_1R_1$, $FliP_5R_1$	150
FliP	C8E5	$FliP_6$, $FliP_2$, and $FliP_1$	150
FliPQR (his tag on FliP)	C8E5	$FliP_5Q_1R_1$, $FliP_5R_1$	150
FliPQR (from FliLMNOPQR construct)	C8E5	$FliP_5Q_1R_1$, $FliP_5R_1$	150

Table 3.1 Summary of the main species observed by nMS of FliPQR in a range of detergents.

The sample in the detergent C10E9 was analysed using xenon rather than argon as the collision gas. Xenon atoms are larger than argon atoms resulting in higher energy collisions between collision gas and protein (Lorenzen et al. 2007). A resolved spectrum could not be obtained using the C10E9 sample with the lower energy argon collision gas.

Figure 3.5 FliP signal peptide cleavage site.

A, Close-up view of the 7+ charge state of the FliP monomer. Two peaks could be observed, one which corresponds to the mass of FliP residues 21 to 245 plus formyl (29 Da) and a second, smaller peak which corresponds to FliP residues 22 to 245. **B**, Sequence of the first 40 residues of FliP in the preprotein and the mature protein following signal peptide removal.

In order to further investigate whether FliP hexamers were present in the original sample despite not being observed in the spectra obtained using samples in C8E5 and DDM or if the hexamer only forms after detergent exchange, FliPQR was analysed after extended incubation in C8E4 (Figure 3.4C). This showed that incubation in a destabilising detergent caused a decrease in $FliP_5Q_1R_1$ and an increase in FliP hexamer abundance, suggesting that one can be converted into the other and that FliQ dissociates from the complex in harsher detergents. In no spectrum could $FliP_6$ be observed bound to either FliQ or FliR. As there was no $FliP_5$ intermediate, it was not clear if the $FliP_6$ complex formed from FliP monomers or from a larger subcomplex of $FliP_5Q_1R_1$.

In order to further validate the use of the detergent C8E5 and confirm that FliP in isolation forms hexamers, as reported by the Namba group (Fukumura et al. 2017), FliP was expressed from a pT12 plasmid containing a FliOP operon with an N-terminal his tag between the FliP signal sequence and the N-terminus of the mature protein. FliP was purified as reported previously (Fukumura et al. 2017) (section 2.5.2) and was then analysed by nMS after exchange into C8E5 and observed to be a pure hexamer (Figure 3.4D). Additionally, FliP monomers and pentamers could be observed. The pentamer peak had a lower charge state than the hexamer, indicating that it was the result of CID. In CID of protein complexes a monomer is ejected from the complex in an unfolded state, taking with it a large proportion of the charge of the intact complex (Pagel et al. 2010). A charge reduced complex is left behind, in this case the FliP pentamer. Thus the charge state indicates that FliP pentamers were not present in the sample before entering the mass spectrometer.

These experiments suggested that $FliP_6$ was an artefact that arose due to the tendency of FliP to assemble into hexamers in the absence of FliQ and FliR either because they are not co-expressed or because they dissociate in a harsh detergent like C8E4. As a control for testing if the his tag affects assembly, a construct of his-tagged FliP including FliQ and C-terminally strep tagged FliR downstream of FliP in the operon was made and the purified protein (section 2.5.1) was analysed after detergent exchange into C8E5. The $FliP_5Q_1R_1$ complex was again observed without evidence of a $FliP_6$ species (Table 3.1).

3.4 Conservation of the core stoichiometry

A previous nMS study suggested that the vT3SS export gate from *S.* Typhimurium SPI-1 is an $SctR_5T_1$ complex (Dietsche et al. 2016), consistent with the $FliP_5Q_1R_1$ complex. SctS was not co-purified, possibly due to the fragility of its association with the SctRT complex. In order to explore the stoichiometry of SctS, *S. flexneri* SctRST was analysed by nMS after exchange into the detergent C8E5, the detergent that was most suitable for analysis of FliPQR. The complex was found to be a mixture of an $SctR_5S_1T_1$ and an $SctR_5S_2T_1$ complex (Figure 3.6A), indicating that the core stoichiometry of the complex is conserved but that the FliQ/SctS subunit is more variable and potentially sensitive to detergent, consistent with the loss of FliQ after incubation in C8E4 (Figure 3.4C).

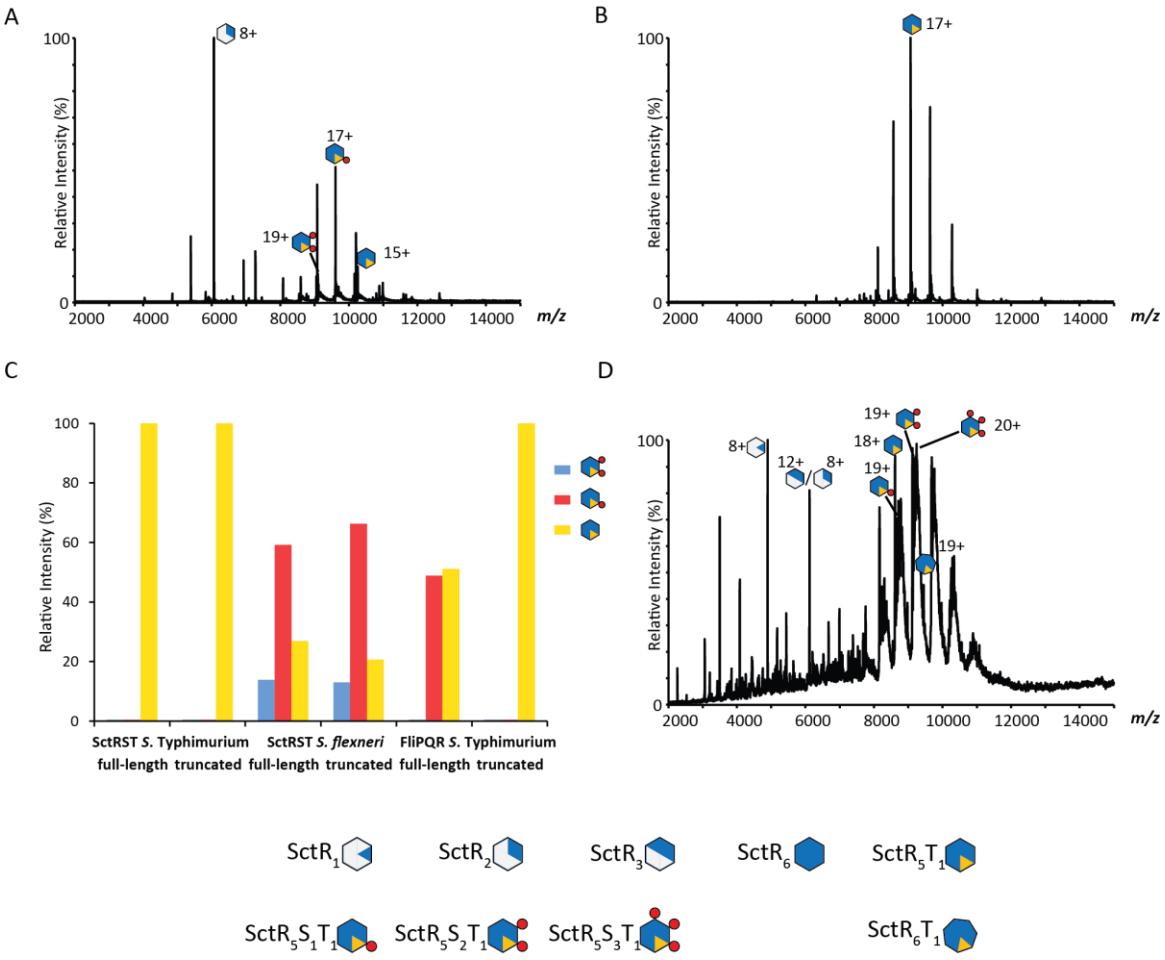


Figure 3.6 nMS of SctRST.
A, Mass spectrum of *S. flexneri* SctRST purified in DDM and exchanged into C8E5. The collision voltage was 200 V. **B**, Mass spectrum of *S.* Typhimurium SPI-1 SctRST. The collision voltage was 150 V. **C**, Relative abundance of $SctR_5T_1/FliP_5R_1$, $SctR_5S_1T_1/FliP_5Q_1R_1$ and $SctR_5S_2T_1/FliP_5Q_2R_1$ in full length and C-terminally truncated constructs in mass spectra obtained under identical conditions. The collision voltage was 150 V in all experiments. **D**, Mass spectrum of *Y. enterocolitica* SctRST. The collision voltage was 150 V.

S. Typhimurium SctRST was previously analysed in the detergent C12E9 and no SctS was observed (Dietsche et al. 2016). In order to test if *S.* Typhimurium SctS can be observed when the sample is prepared in the same way as *S. flexneri* SctRST, the *S.* Typhimurium complex was purified and exchanged into C8E5 following the same protocol as for *S. flexneri* SctRST (section 2.5.1). Although the peaks in the spectrum of *S.* Typhimurium obtained here are sharper than what was observed before (Dietsche et al. 2016), the stoichiometry is the same, $SctR_5T_1$ (Figure 3.6B).

In the *S.* Typhimurium SPI-1 SctRST construct the C-terminus of SctT (residues 252 to 263) was truncated. In order to exclude the possibility that SctS is not bound to the complex due to the mutation of SctT, a full length construct was cloned and analysed in the same way, but the stoichiometry was unaffected (Figure 3.6C). In addition, truncated *S.* Typhimurium FliOPQR (FliR residues 239 to 264) and *S. flexneri* SctRST (SctT residues 240 to 256) constructs were made and the complexes were purified and analysed as before. The results were ambiguous, showing loss of FliQ binding after truncation only in *S.* Typhimurium FliPQR (Figure 3.6C), suggesting that the C-terminus of FliR is part of the FliQ binding site on the FliPQR complex, but other interactions keep SctS associated with the complex, possibly with one of the SctR subunits.

In order to further investigate the stoichiometry of FliQ/SctS, two further vT3SS homologues were prepared of which one yielded interpretable mass spectra: *Y. enterocolitica* SctRST was purified in DDM (section 2.5.1). After screening a small number of detergents, most of which led to aggregation of the protein, the sample was kept in DDM and the buffer was exchanged to the nMS compatible 200 mM ammonium acetate containing 0.016% (w/v) DDM using a Bio-Spin P6 column (section 2.5.9). Then, the sample was diluted 3:1 in 0.8% (w/v) C8E6 in 200 mM ammonium acetate and analysed immediately. Due to the instability of this sample, data could only be recorded for a few minutes before the signal disappeared, suggesting that *Y. enterocolitica* SctRST is highly sensitive to detergents. The spectrum obtained showed a mixture of species (Figure 3.6D), including a core $SctR_5T_1$ complex with up to three SctS subunits as well as an $SctR_6T_1$ complex. Due to the instability of this sample, it is possible that at least some of the observed species are the result of formation of non-physiological complexes,

similar to the *S.* Typhimurium FliP hexamer (Figure 3.4). As the sequences of FliPQR/SctRST are highly conserved, it is likely that the stoichiometry is widely conserved making the consistently observed core $FliP_5R_1$/$SctR_5T_1$ complex bound by a variable number of FliQ/SctS subunits the most likely physiological stoichiometry.

3.5 Probing the substructure of the export gate complex using atypical CID

When FliPQR was analysed in DDM (Figure 3.4A), negative mode (Liko et al. 2016) was used to generate negatively charged ions which allowed analysis of the intact complex. When nMS was performed with positive charge, as used for all other samples, the complex was observed to break apart (Figure 3.7A), complicating the stoichiometric analysis. Dissociation of a complex in the mass spectrometer into intact subcomplexes is called atypical CID (Hall et al. 2013), and enables the study of the architecture of the complex through analysis of the subcomplexes. No subcomplexes including FliQ were observed, suggesting that at least in *S.* Typhimurium FliPQR this subunit is bound weakly to a more robust FliPR core complex when the complex is extracted in DDM and dissociates quickly in the gas phase inside the mass spectrometer. Subcomplexes of the FliPR complex indicated that FliR is not a peripheral subunit: If FliPR is a FliP pentamer bound by a FliR subunit on the outside (Figure 1.7D), the complex would be expected to first fall apart into a FliR monomer and a closed $FliP_5$ ring, i.e. the dissociation behaviour that was observed for FliQ. While $FliP_5$ was observed, other species, including $FliP_4R_1$, $FliP_3R_1$, $FliP_2R_1$ and $FliP_1R_1$ were also observed. This suggested that $FliP_5R_1$ is a heterohexameric ring rather than a pentamer plus a peripheral FliR subunit. As the structure of the export gate was expected to be ring shaped like a “donut” (Dietsche et al. 2016), FliR is most likely a component of that ring.

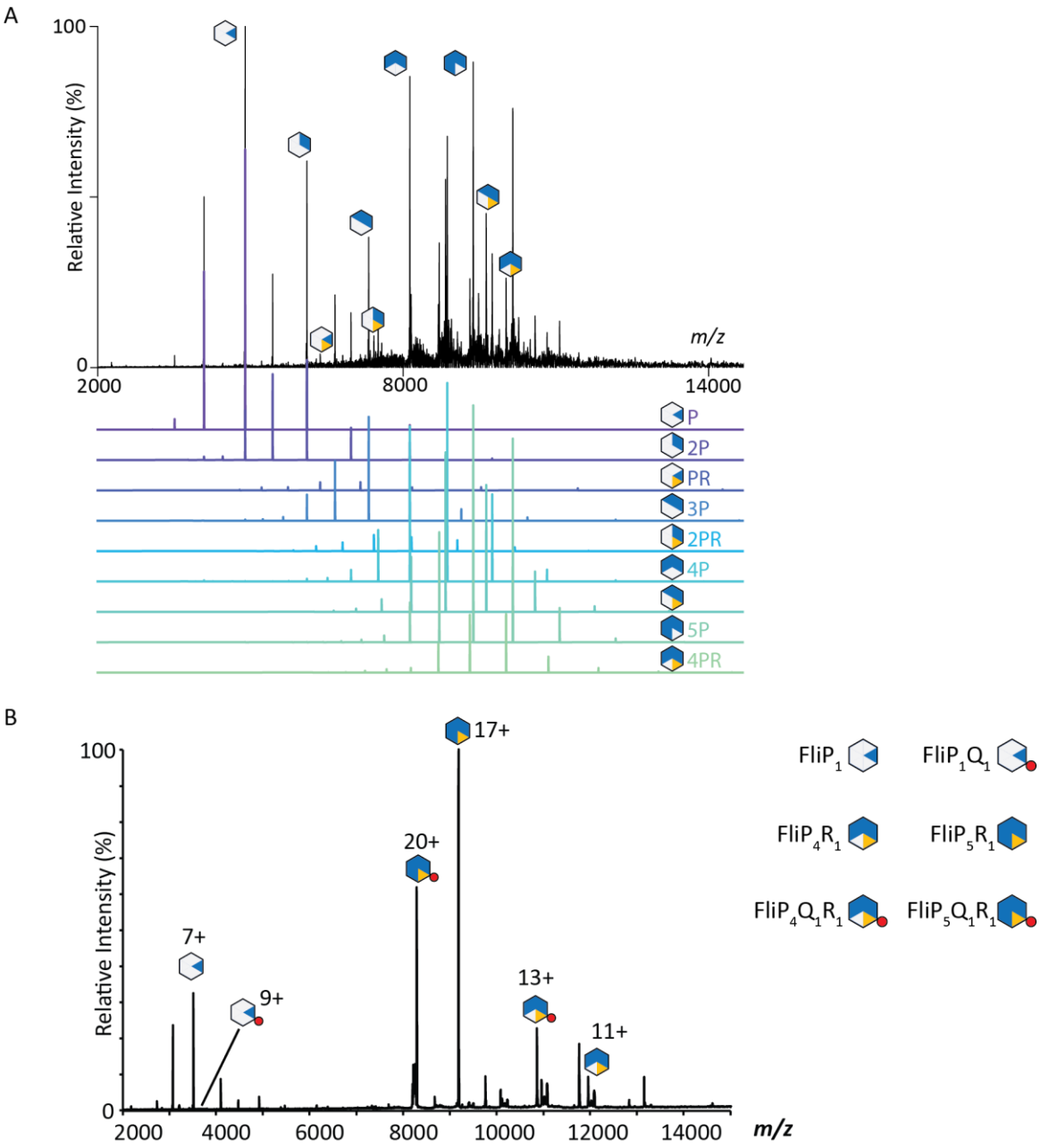


Figure 3.7 atypical CID of FliPQR.
A, Atypical CID of *S.* Typhimurium FliPQR purified in DDM. The collision voltage was 200 V with an addition 150 V of in source activation (Gault et al. 2016), causing collisional activation in both the collision cell and the ion funnel at the entrance of the instrument. **B**, Dissociation of the 20+ charge state of *S.* Typhimurium $FliP_5Q_1R_1$ purified in DDM and exchanged into C8E5. The sample sample was activated in source (100 V) and the 8279 m/z ions were isolated. The ions were dissociated by applying 250 V in the collision cell. This experiment was performed together with Dr Joe Gault.

In order to further investigate how FliQ binds to the complex, FliPQR, exchanged into C8E5, was analysed by selecting an ion corresponding to the 20+ charge state of $FliP_5Q_1R_1$ inside the mass spectrometer followed by CID of just this ion (Figure 3.7B).

Interestingly, one dissociation product was a $FliP_1Q_1$ complex. This showed that FliQ not only contacts FliR (Figure 3.6C) but also FliP, suggesting that this FliQ subunit is bound to the interface of FliR and a FliP subunit in the $FliP_5R_1$ ring. Furthermore, no $FliP_5$ complex was observed, which would be the result of dissociation of FliR, further strengthening the idea that FliR is not peripheral to a central FliP pentamer.

3.6 Stoichoimetry of the export gate purified in the detergent LMNG

The export apparatus subunit FlhB/SpaS is thought to interact with the export gate complex (Van Arnam et al. 2004), but the stoichiometry of this interaction was unclear (McMurry et al. 2015), making a FliPQR-FlhB complex an attractive target for characterisation by nMS. A detergent screen was designed with the aim of finding a detergent that would allow co-purification of *S.* Typhimurium FlhB with FliPQR (section 2.5.4). Purified samples were submitted for proteomic analysis in the Dunn School proteomics facility and co-purification of FlhB was quantified by measuring the relative intensity of FlhB to FliP peptides (Figure 3.8A). This revealed that FlhB was enriched more than 24-fold in LMNG compared to DDM, although co-purification of FlhB was not initially successful to levels stoichiometric with FliPQR (see chapters 4 and 6).

Moreover, increased amounts of FliQ were seen to co-purify with FliPR when extracting and purifying the complex with LMNG (Figure 3.8B). Therefore, the nMS analysis was repeated with FliPQR/SctRST extracted in LMNG and exchanged into C8E5 (section 2.5.9), but the *S.* Typhimurium FliPQR complex could not be fully resolved (Figure 3.8C). The broad nature of these peaks, between 8,000 and 12,000 m/z, suggests binding of molecules of heterogeneous mass to the complex, most likely tightly bound lipids (Bechara et al. 2015). In addition, a FliQ dimer was observed, indicating that the FliQ subunits in the assembled FliPQR complex are in contact with each other.

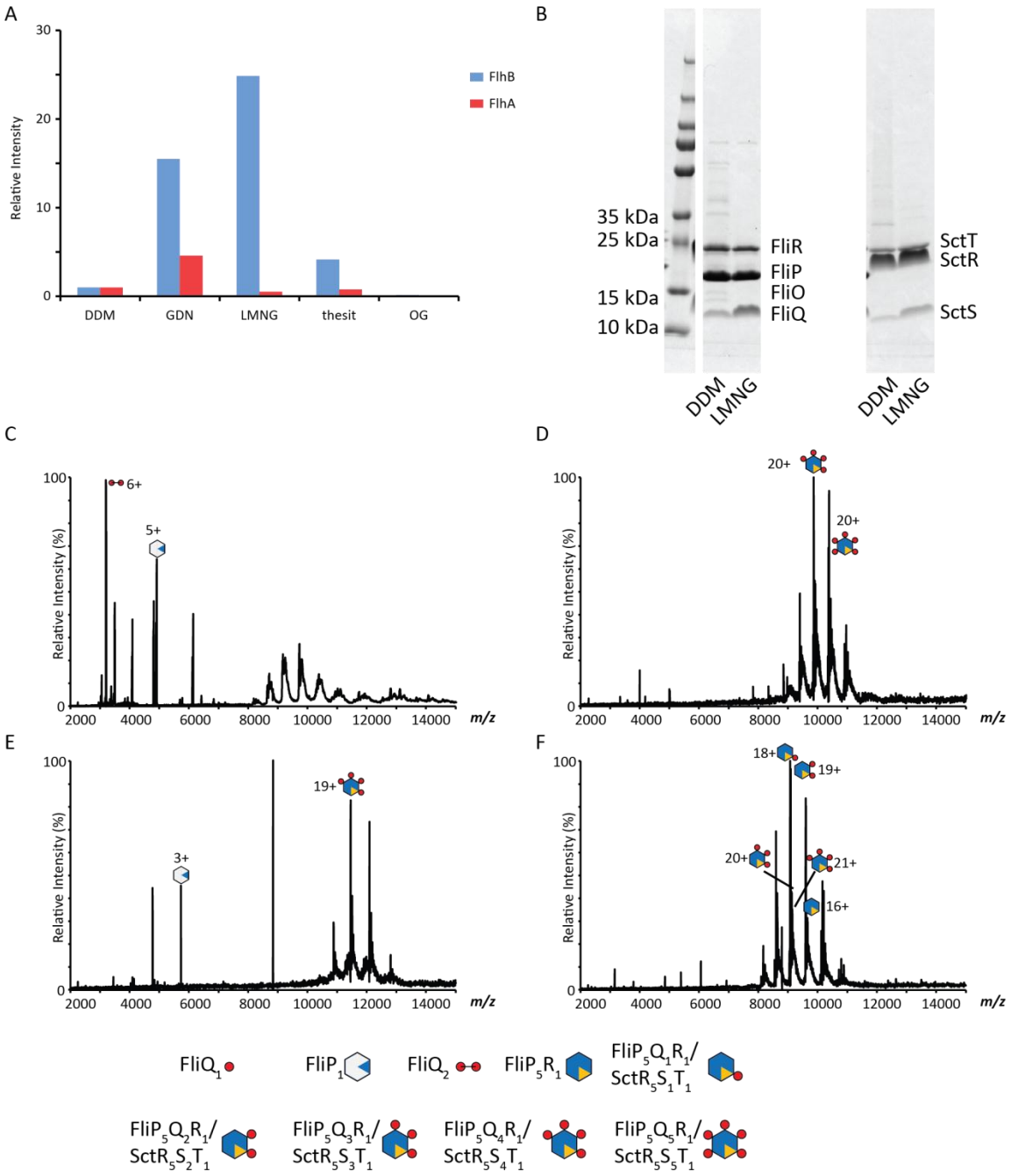


Figure 3.8 Analysis of the export gate in LMNG.
A, Relative intensity of FlhB and FlhA peptides compared to FliP peptides by proteomics in samples of FliPQR co-expressed with FlhBA and purified in a range of detergents. **B**, Representative 4-20% SDS-PAGE gradient gel comparing FliOPQR and SctRST purified in DDM and LMNG. **C**, Mass spectrum of *S.* Typhimurium FliPQR purified in LMNG and exchanged into C8E5. The collision voltage was 200V. **D**, Mass spectrum of *P. savastanoi* FliPQR purified in LMNG and exchanged into C8E5. The collision voltage was 200V. **E**, Mass spectrum *V. mimicus* FliPQR purified in LMNG and exchanged into C8E5. The collision voltage was 200V. **F**, Mass spectrum of *S.* Typhimurium SctRST purified in LMNG and exchanged into C8E5. The collision voltage was 200V.

As the stoichiometry of *S.* Typhimurium FliPQR in LMNG could not clearly be resolved, two further FliPQR complexes, from *Pseudomonas savastanoi* pathovar phaseolicola strain 1448A (DSM 21482) and *Vibrio mimicus* (DSM 19130), were cloned as FliOPQR operons in the same way as the *S.* Typhimurium FliOPQR construct and were expressed and purified in LMNG (section 2.5.3). While *P. savastanoi* FliPQR was observed to be a mixture of $FliP_5Q_4R_1$ and $FliP_5Q_5R_1$ complexes by nMS (Figure 3.8D), *V. mimicus* FliPQR was a homogeneous $FliP_5Q_4R_1$ complex (Figure 3.8E). *S. flexneri* SctRST purified in LMNG (section 2.5.3) was observed to be a complex mixture of $SctR_5T_1$, $SctR_5S_1T_1$, $SctR_5S_2T_1$, $SctR_5S_3T_1$ and $SctR_5S_4T_1$ (Figure 3.8F). This suggested that even though LMNG allowed co-purification of a higher number of FliQ/SctS subunits per complex compared to DDM, the additional subunits may be more sensitive to detergent or more weakly bound than the single FliQ subunit that was previously predicted to be located at a FliP/FliR interface (section 3.5). The average sequence identity between *S*. Typhimurium FliPQR and *V. mimicus* FliPQR is 39%, the identity between *S.* Typhimurium and *P. savastanoi* is 44% and the identity between *S.* Typhimurium FliPQR and *S. flexneri* SctRST is 26%.

3.7 Modelling of the architecture of the export gate

The interaction between FliP, FliQ and FliR was additionally analysed using bioinformatics. The evolutionary co-variation was calculated using the GREMLIN server (http://openseq.org/index.php) (Ovchinnikov et al. 2014). Residues of both predicted transmembrane helices of FliQ co-vary with residues in the C-terminal half of FliP (Figure 3.9A) indicating a large interface between the two proteins that could potentially occur at each FliP subunit. This would support a FliQ stoichiometry of five per complex. Extensive contacts between FliP and FliR throughout the entire sequence (Figure 3.9B) support the position of FliR as part of a $FliP_5R_1$ ring as predicted above (Figure 3.7). In contrast, only helix one of FliQ co-varies with FliR (Figure 3.9C). This was consistent with the most tightly bound FliQ binding at the FliP-FliR interface in the FliPR ring as suggested by the analysis of truncated constructs (Figure 3.6). The sensitivity of FliQ/SctS to detergents and dissociation in the mass spectrometer suggests that it peripherally surrounds the FliPR core, most likely in a continuous layer as suggested by the FliQ interaction with itself (Figure 3.8C). The stoichiometry of FliQ cannot be determined precisely due to the possibility of dissociation of this fragile subunit during the

purification, but the highest observed number per complex, five, is consistent with the predicted interaction between FliP and FliQ, making this the most likely correct stoichiometry.

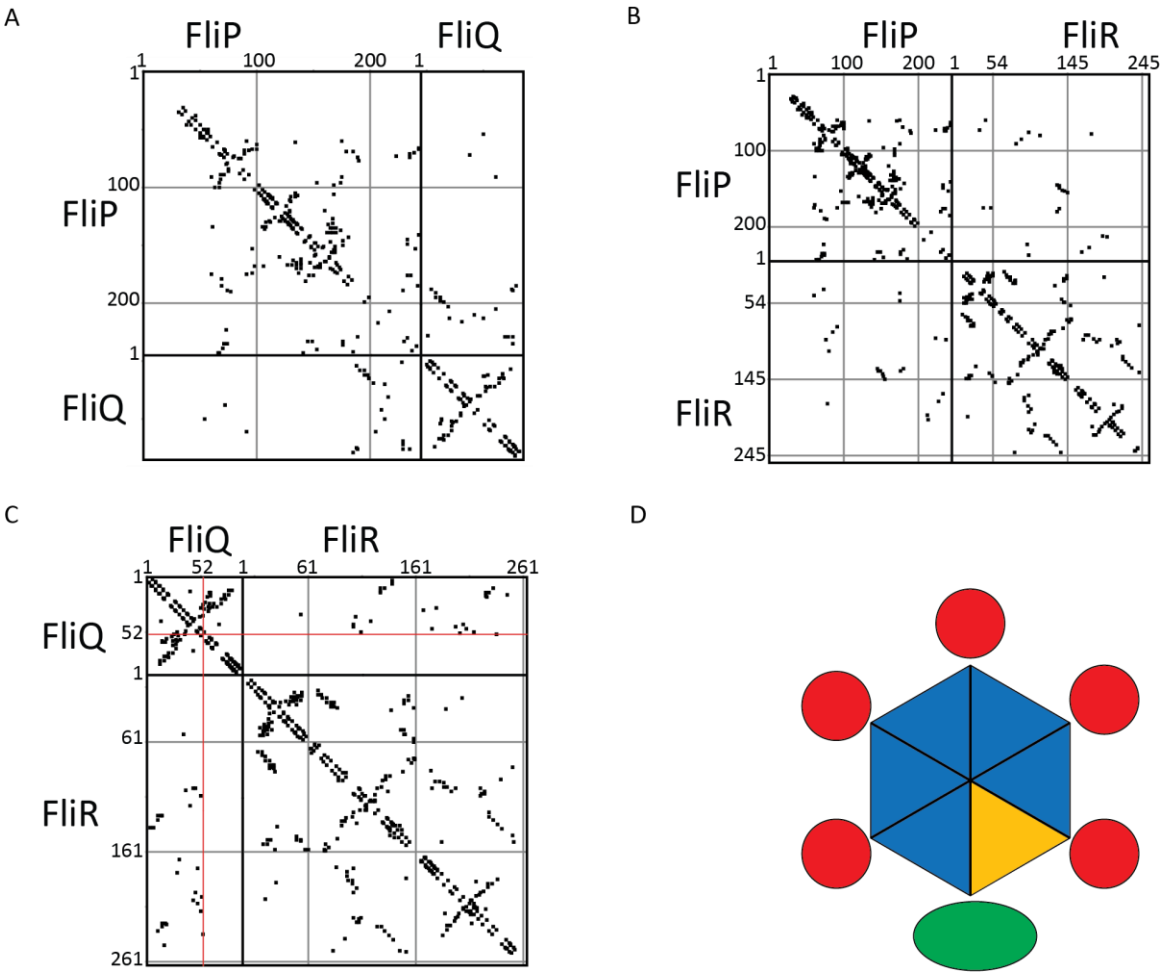


Figure 3.9 Evolutionary covariation.
A-C, Covariation within FliPQR calculated using GREMLIN (Ovchinnikov et al. 2014). Only contacts with a probability score greater than 0.75 are displayed. Contacts were calculated from pairs of sequences. Contacts within a single sequence are in the squares at the top left and bottom right while contacts between the proteins are in the rectangles at the top right and bottom left. The FliP C-terminal 100 residues co-vary with residues spread throughout the entire FliQ sequence (**A**), FliP and FliR co-vary throughout both of their sequences (**B**) and the N-terminal 52 residues of FliQ co-vary with residues spread throughout the entire FliR sequence (**C**). The MSAs contained 2,745 (**A**), 3,755 (**B**) and 3,594 aligned sequences. **D**, Cartoon model of the architecture of the FliPQR-FlhB complex. FliP=blue, FliR=yellow, FliQ=red, FlhB=green.

Additionally, FlhB is thought to contact FliR as a fusion of the two proteins is partially functional and exists naturally in *Clostridium* (Van Arnam et al. 2004), which is supported

by interactions between SctRST and SctU (FliPQR and FlhB) shown by *in vivo* photo-crosslinking (Dietsche et al. 2016). This would place FlhB in the peripheral layer made of FliQ that surrounds FliPR. Summarising all the data, a new model for the architecture of the export gate complex could be proposed (Figure 3.9D). The core complex is proposed to be a heterohexameric ring made up of 5 FliP subunits and 1 FliR subunit. This is surrounded by a ring of 5 FliQ subunits and 1 FlhB subunit.

3.8 Discussion

In this chapter the stoichiometry of the export gate complex FliPQR/SctRST was shown to be conserved in flagellar and virulence T3SS. The core is formed by a heterohexameric ring made up of five FliP/SctR subunits and one FliR/SctT subunit. The exact stoichiometry of FliQ/SctS could not be determined due to its sensitivity to detergent, but it was clear that there were more than one FliQ/SctS subunit with a maximum of five observed in complex with FliPR (Figure 3.8). Combining all data with bioinformatic analysis allowed the construction of a new model for export gate architecture of both vT3SS and fT3SS (Figure 3.9D). This model contradicts both older models (Dietsche et al. 2016; Fukumura et al. 2017).

3.8.1 Comparison with older models

The disagreement between the data presented here and the SctRT pentamer model (Dietsche et al. 2016) can be explained by misinterpretation of negative stain data. Previously, *S.* Typhimurium $SctR_5T_1$ was analysed by negative stain EM (Figure 3.10). A mixture of particles were observed, the majority were rings while some appeared to be rings with an additional subunit outside the ring. The latter were proposed to be SctR pentamers with an SctT subunit decorating the outside of the SctR ring and the former were proposed to correspond to SctR rings from which the SctT subunit had dissociated. However, a different possibility is that these are simply different views of the same object. As shown above (Figure 3.4), dissociation of FliR leads to formation of FliP hexamers, which are also thought to be rings (Fukumura et al. 2017). The tendency of FliP to form hexameric ring structures when other export gate subunits are not present, either due to absence in the recombinant expression system or due to instability in a

relatively harsh detergent, suggests that the FliP homohexamer is mimicking the native FliPR heterohexamer. As the sixth FliP subunit that completes the ring in the absence of FliR and FliR can both fill the same position in the complex, FliR may adopt a structure similar to FliP. How does the $SctR_6T_1$ complex, a minor species in the mass spectrum of *Y. enterocolitica* SctRST, fit into this model? Considering the instability of this sample, the lack of SctS in the complex and presence of SctS in *Y. enterocolitica* SctRST complexes containing only 5 SctR subunits and the evidence from many other homologous complexes, this observation can be regarded as an outlier likely resulting from the formation of a non-physiological oligomer from disassembled monomeric subunits.

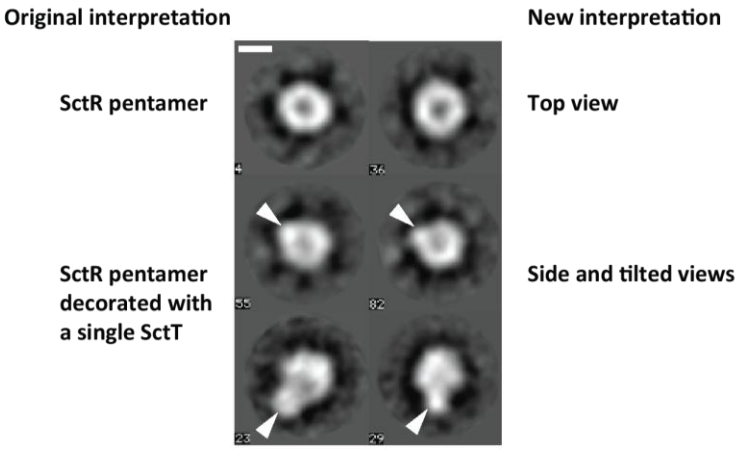


Figure 3.10 Re-interpretation of negative stain EM data.
Previously published 2D class averages of *S.* Typhimurium SPI-1 SctRT (Dietsche et al. 2016). The scale bar indicates 5 nm. The lower four averages were originally described as pentameric SctR rings decorated with an SctT subunit. A different interpretation, suggested by the new model of the architecture of FliPQR/SctRST, is that these averages represent different views of the same object, which is a heterohexameric ring of 5 SctR and 1 SctT.

3.8.2 Assembly of FliPQR

As FliP can form hexameric rings in the absence of FliR, the native system may need to prevent formation of these complexes which have been reported to cause toxicity by making the inner membrane permeable to ions (Dietsche et al. 2016; Ward et al. 2018). *In vivo* formation of large assembly intermediates containing FliP, likely corresponding to the $FliP_5Q_{\sim5}R_1$ complex revealed in this chapter, was strongly reduced in the absence of FliR (Fabiani et al. 2017), suggesting that in the native context FliP is stopped from

oligomerising before FliR is made. This would imply that FliR is the nucleus for export gate assembly. Nevertheless, FliP hexamers have been observed in complex with FliO in a recombinant context (Fukumura et al. 2017). The assembly of the non-native hexamers may occur only at a high concentration of FliP/SctR.

3.8.3 Effects of detergents on membrane proteins

The case of the export gate teaches that care needs to be taken to avoid formation of artefactual stoichiometries, in this case resulting from formation of a symmetric complex in place of the pseudosymmetric native complex due to detergent or expression artefacts.

A similar example of a membrane protein complex with controversial stoichiometry is the complex of ExbB and ExbD. ExbB can form pentameric and hexameric complexes (Celia et al. 2016; Maki-Yonekura et al. 2018) while a variable number of ExbD molecules have been observed inside the ExbB ring. The different stoichiometries have been proposed to be of functional relevance, however it is possible that at least one model is the result of artefacts of expression or detergent solubilisation. A related complex, the flagellar MotAB complex, is thought to be a complex of four MotA and 2 MotB subunits (Yonekura et al. 2011; Takekawa et al. 2016) despite the homology to the ExbBD complex. It is possible that expression and detergent solubilisation artefacts are partially responsible for the different observed stoichiometries.

Another case of controversial stoichiometry is found in a homomeric complex. The stoichiometry of the membrane protein complex of MscL can be manipulated using detergents, lipids and temperature and this has been directly observed with native mass spectrometry (Reading et al. 2015). It is not known which stoichiometry is the physiologically relevant one or whether different oligomeric states could be functionally relevant.

3.8.4 Future perspectives

The results presented in this chapter could form the basis for structural studies of the export gate complex using cryo-EM. In particular, the discovery of a more stabilising detergent, LMNG, with the potential to stabilise the sample when prepared for cryo-EM, might prove useful for further structural studies.

4 Structure of the export gate: membrane proteins outside the membrane

4.1 Introduction

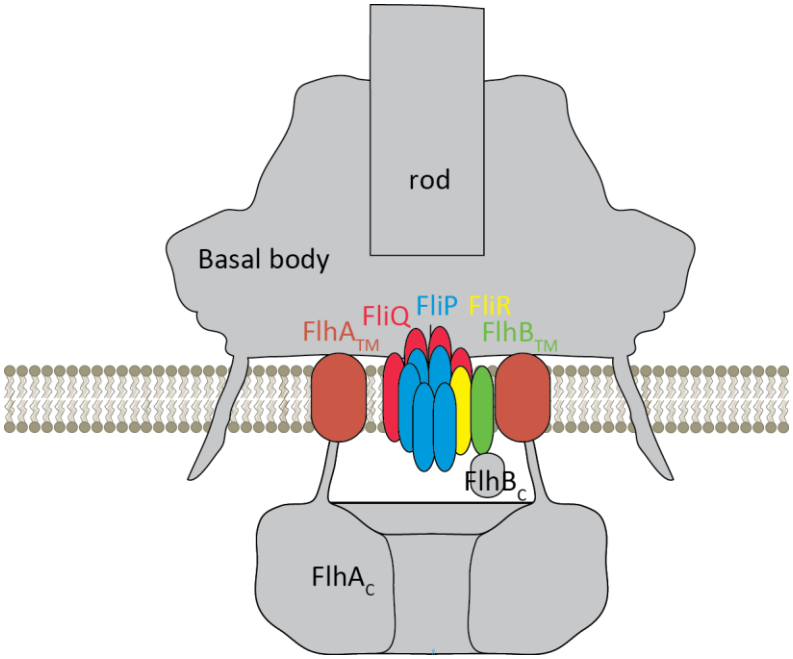


Figure 4.1 Cartoon showing the proposed location of FliPQR in the basal body. Model of the export gate in the basal body based on knowledge prior to work described in this chapter. The cartoon is based on the results in chapter 3 and cartoons in recent papers (Ohnishi et al. 1997; Lorenzen et al. 2007; Fukumura et al. 2014; Bechara et al. 2015).

In chapter 3 the architecture of the export gate complex, FliPQR, was analysed. These experiments not only generated and characterised valuable reagents, but also suggested an architecture consisting of a core $FliP_5R_1/SctR_5T_1$ heterohexameric ring surrounded by a ring of up to 5 FliQ/SctS subunits, which likely resides in the inner membrane surrounded by the basal body (Figure 4.1). In addition, experiments in chapter 3 showed that the FliPQR complex could be stabilised by extracting it from the membrane in the detergent LMNG rather than DDM and that LMNG stabilised the interaction between the regulatory protein FlhB and FliPQR, making LMNG the detergent of choice for further studies. In order to exploit these results, structural investigations using cryo-EM were begun.

4.2 Structure determination of FliPQR

4.2.1 Imaging of FliPQR in LMNG

Previous cryo-EM and crystallography studies in the lab had focused on *S.* Typhimurium SPI-1 SctRT (Dietsche et al. 2016) solubilised in the detergent DDM or the amphipol A8-35, but although data could be collected using Titan Krios microscopes (FEI), no high resolution volume could be calculated (experiments performed by Dr Patrizia Abrusci). As even data collected on the most advanced detectors and microscopes available did not lead to a high-quality reconstruction, a different sample was used here. Flagella rotate by applying high torque to the flagellar filament, and higher torque is correlated with additional subunits stabilising the nanomachine (Beeby et al. 2016). It is therefore possible that the flagellar EA has evolved to be more stable than injectisome homologues in order to withstand the forces resulting from rotation. Consistent with this hypothesis, higher FliQ/SctS stoichiometries were observed in some flagellar homologues prepared in LMNG using native mass spectrometry (chapter 3), in particular in *Vibrio* and *Pseudomonas* FliPQR, two species whose flagella rotate using higher torque than those of the common model organism *Salmonella* (Beeby et al. 2016; Kaplan et al. 2019). Nevertheless, *Salmonella* and *Escherichia* flagella are the most well studied (Berg 2003; Macnab 2003) and so *S*. Typhimurium FliPQR was chosen for structural studies in this chapter.

S. Typhimurium FliPQR was prepared in the detergents DDM and LMNG as before (section 2.5.3). Cryo-EM grids were prepared using samples Quantifoil Au 400 mesh holey carbon coated grids with a hole diameter of 1.2 µm and a hole spacing of 1.3 µm using a Vitrobot Mark IV (FEI) at 22 °C and 100% humidity. Initial screening showed that a sample concentration between 0.7 and 1 mg/ml resulted in a high particle density with sufficient distance between particles. Grids were imaged on a Talos Arctica electron microscope (FEI) operated at 200 kV using the program EPU (FEI) for automated data collection using a defocus range of 0.5 to 4 µm. The microscope was equipped with a Falcon 3 detector (FEI) operating in linear mode. The pixel size was 0.78 Å. Micrographs were pre-processed using the SIMPLE pipeline including an implementation of the program Unblur for motion correction and an implementation of CTFFIND4 for CTF estimation (Reboul et al. 2018). Particles were picked using Gautomatch (Urnavicius et al.

2015) and extracted in a 240x240 Å box. 2D class averaging was performed in SIMPLE (FliPQR DDM) or RELION1.4 (Scheres 2012) (FliPQR LMNG). Initial analysis showed that distinct particles could be observed with more detail in top views of the complex in LMNG samples compared to DDM when subjected to 2D classification. Some 2D averages of the top view of the LMNG sample featured six "swirls" which might correspond to the six subunits of the $FliP_5R_1$ ring. However, all averages from the LMNG sample looked like discs with varying levels of internal detail, while clearly different shapes could be seen in DDM (Figure 4.2). The circular average is likely to correspond to the top view of the membrane pore formed by FliPQR while the other view may correspond to a side view. The high resolution features of the complex in LMNG suggested that this sample was more likely to lead to a cryo-EM structure, however the lack of other views prevented an analysis in 3D.

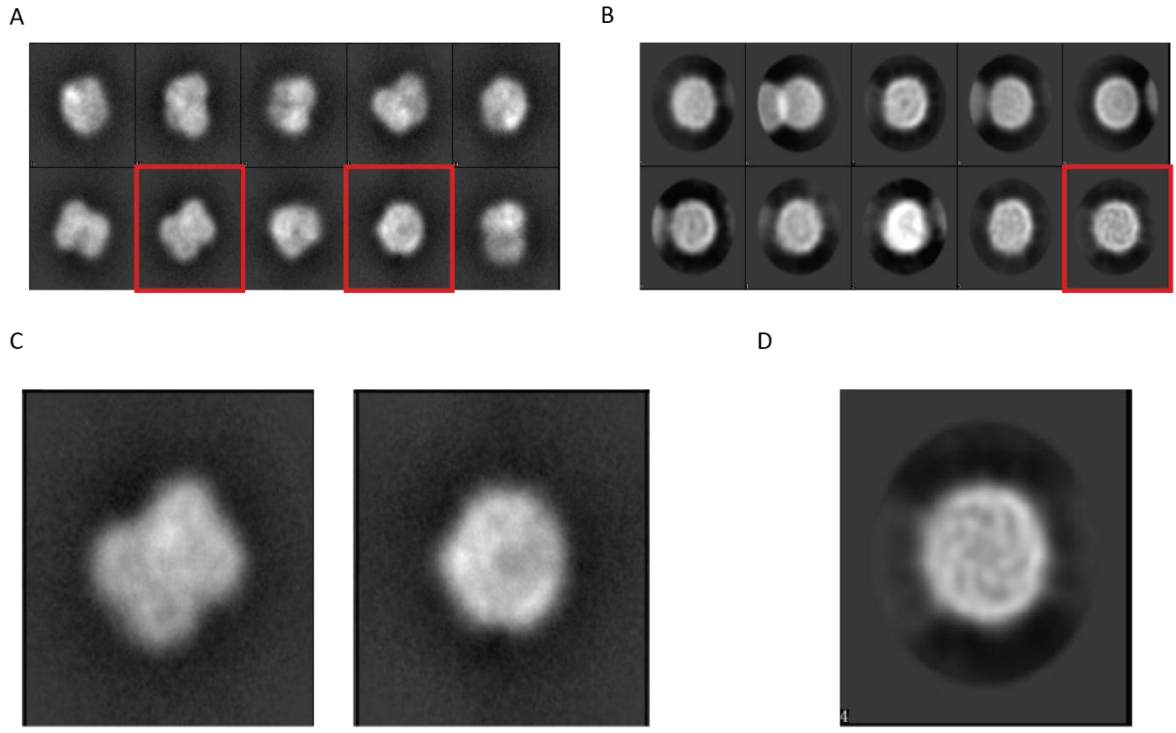


Figure 4.2 Initial cryo-EM characterisation of FliPQR.
A, 10 most populated 2D class averages of *S.* Typhimurium FliPQR in DDM calculated using SIMPLE PRIME2D. **B**, 10 most populated 2D class averages of *S.* Typhimurium FliPQR in LMNG calculated using RELION1.4. **C**, Close-up view of the highlighted averages from (**A**). **D**, Close-up view of the highlighted average from (**B**) in which swirls can be seen.

4.2.2 Initial imaging of FliPQR in complex with FlhB

In an attempt to both obtain side views by changing the surface of the protein and prepare a larger complex, co-purification of FlhB was attempted. The secretion specificity switch protein FlhB is made up of a predicted transmembrane domain of unknown structure and a cytoplasmic domain (Meshcheryakov et al. 2013) (chapter 1). Addition of these two domains to the FliPQR complex might change the orientation of the particles in the cryo-EM grid and the presence of the cytoplasmic domain outside the detergent micelle might aid image classification. Co-purification of substoichiometric amounts of FlhB with FliPQR was previously observed using proteomics (chapter 3), suggesting that preparation of a monodisperse FliPQR-FlhB complex is possible. For this purpose, an operon of FlhB and FlhA was placed in a pBAD plasmid with a C-terminal his tag on FlhA and was co-expressed with FliOPQR in a pT12 plasmid. The culture was supplemented with 0.1% rhamnose monohydrate and 0.1% arabinose and the appropriate antibiotics and the FliPQR-FlhB complex was expressed and purified using the same protocol as for FliPQR without FlhB using the detergent LMNG (section 2.5.3). The SEC profile of the resulting complex revealed a shoulder to the left hand side of the FliPQR peak (Figure 4.3A). SEC-MALS was used to further interrogate the shoulder of the peak (section 2.5.8). This revealed that although the mass of the complex corresponding to the shoulder was very similar to that of the main peak, it was slightly larger (about 8 kDa). SDS-PAGE analysis was also consistent with co-purification of FlhB (Figure 4.3B), with the full FliPQR-FlhB complex eluting in the shoulder of the peak.

Cryo-EM grids were made using different fractions across the peak and analysed on a Talos Arctica microscope (FEI). Quantifoil Au 400, R1.2/1.3 grids made with 0.9 mg/ml protein sample at 22 °C and 100% humidity (Table 2.10) from the fraction corresponding to the shoulder of the peak contained a mixture of views (Figure 4.3C, D) when analysed using SIMPLE and were selected for data collection on a Talos Arctica microscope (FEI). 11,650 movies were collected using a Falcon 3 detector as described above for FliPQR in LMNG (section 4.2.1). 2D classification of the particles was carried out in RELION-2.0 (Kimanius et al. 2016) and an ab initio initial model was calculated using SIMPLE (Reboul et al. 2018). This model was used as a reference for 3D classification in RELION-2.0 and

the particles in the highest resolution class were used for auto-refinement in RELION-2.0 resulting in a volume with a resolution of 7.8 Å using the FSC=0.143 criterion.

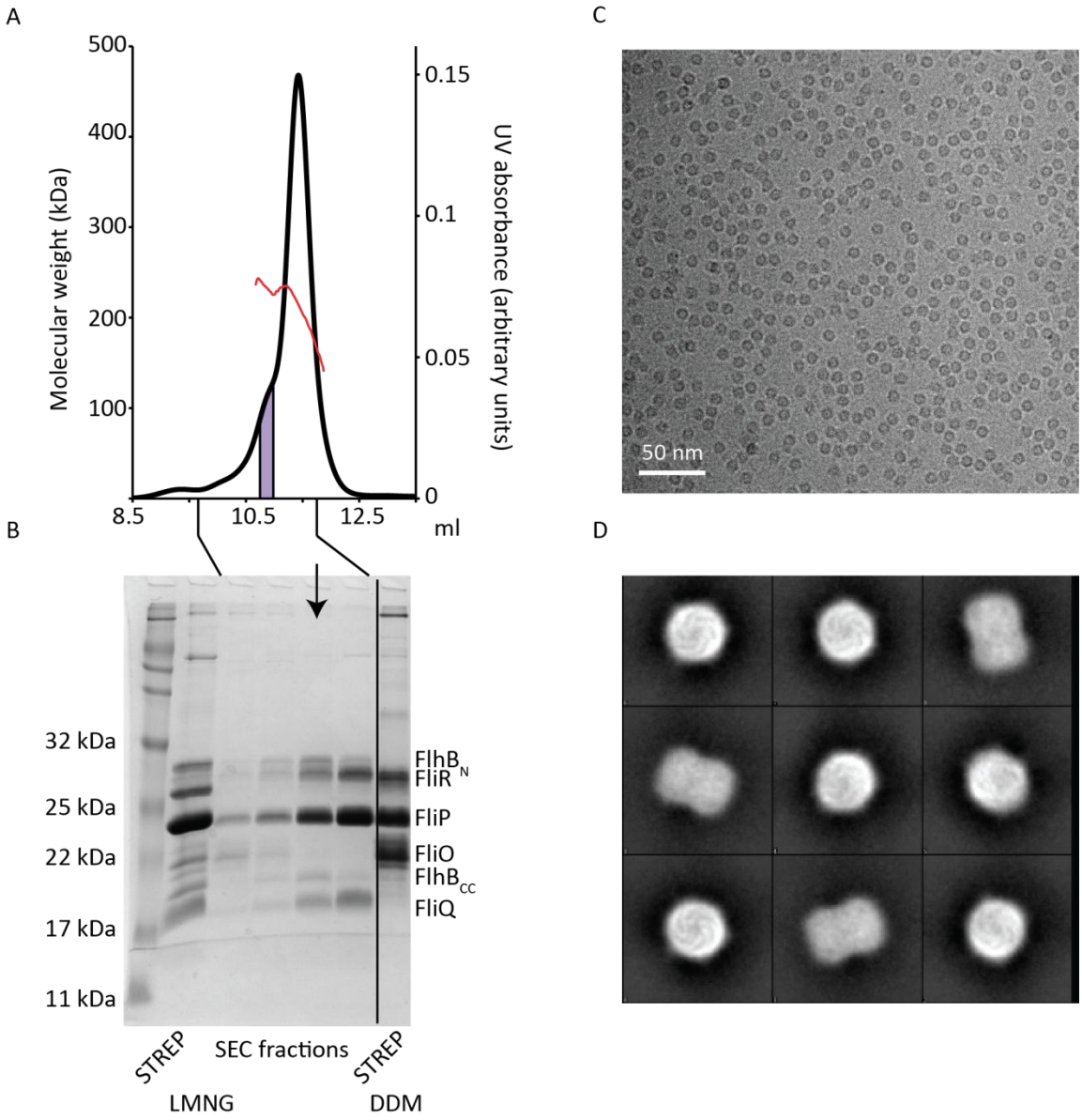
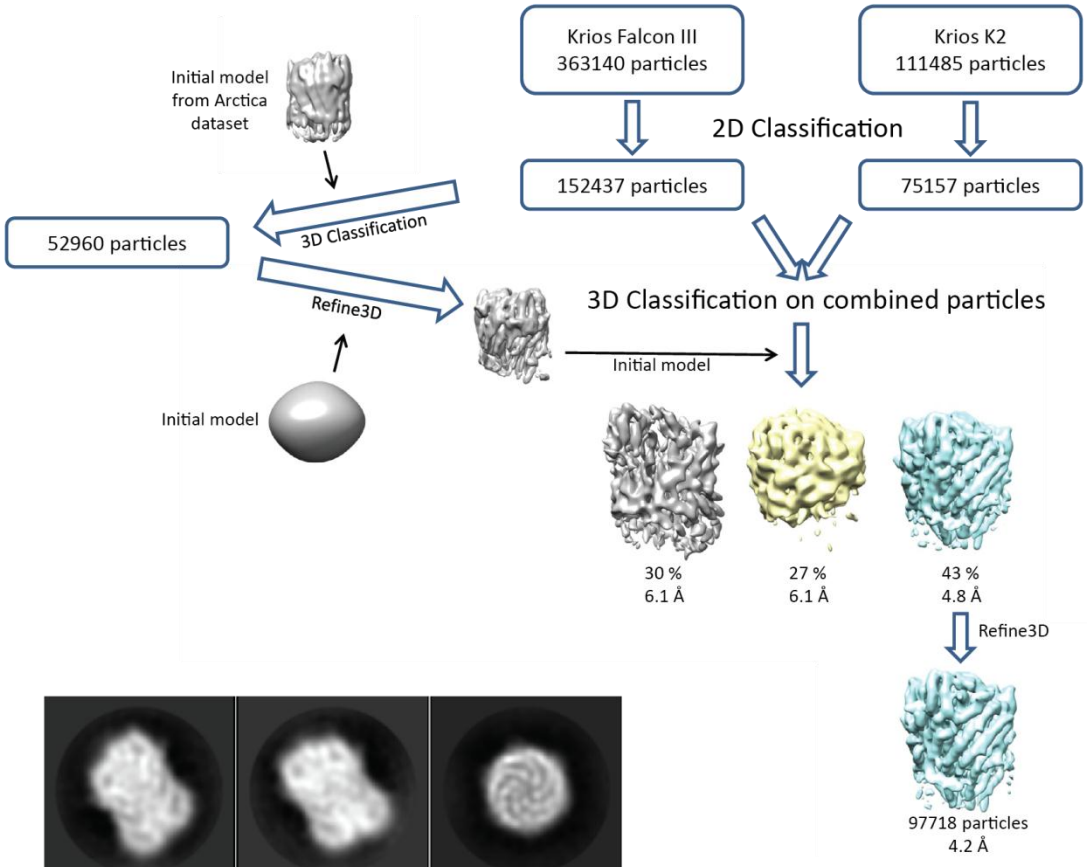


Figure 4.3 Co-purification of FlhB with FliPQR.
A, SEC-MALS of *S.* Typhimurium FliPQR-FlhB. The purple shading indicates the shoulder of the peak from which grids for cryo-EM were made. **B**, 15% SDS-PAGE analysis of the *S.* Typhimurium FliPQR-FlhB purification in LMNG (left) and DDM (right). Fractions from across the SEC peak were run on the gel. The arrow indicates the fraction equivalent to the shaded fraction in the SEC-MALS trace in (**A**). All fractions **C**, Representative micrograph of FliPQR-FlhB. **D**, 9 most populated 2D class averages of FliPQR-FlhB imaged using a Talos Arctica microscope (FEI) calculated using SIMPLE PRIME2D.

4.2.3 High resolution cryo-EM imaging of FliPQR-FlhB

More data were collected from grids made under identical conditions (section 4.2.2) on a Titan Krios microscope (FEI) operating at 300 kV using a Falcon 3 detector (FEI) and a K2 detector (Gatan) in counting mode and using EPU (FEI) for automated data collection. The magnification used for the Falcon detector was 96,000 and the magnification for the K2 detector was 165,000 resulting in a pixel size of 0.86 Å for both data sets in the setup of the microscope in the Oxford cryo-EM facility (COSMIC). The datasets were pre-processed as described for the Arctica data (section 4.2.1) using a SIMPLE implementation of Unblur and CTFFIND4 and particles were picked using Gautomatch (Urnavicius et al. 2015) and extracted in a 256x256 pixel box.

A



B

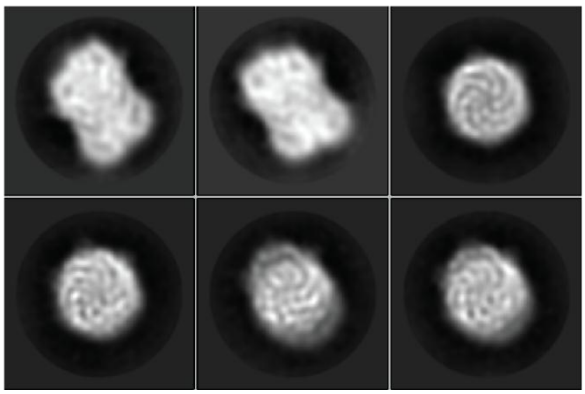


Figure 4.4 Single particle analysis of *S.* Typhimurium FliPQR FlhB in LMNG.
A, Flowchart describing the data processing strategy. **B**, Selected 2D class averages of the particles calculated in RELION-2.0.

~360,000 Falcon III particles and ~110,000 K2 particles were classified in 2D in RELION-2.0 (Kimanius et al. 2016) and the datasets were merged (Figure 4.4A, B). The cone-shaped Arctica model (section 4.2.2) was used as the initial model for 3D classification in RELION-2.0 of the Falcon3 particles and the highest resolution class was subjected to auto-refinement using a featureless ovoid object of similar dimensions as the initial model. The resulting volume, in which tube-like densities could be resolved, was used as the initial model for 3D classification of the entire dataset. High resolution features, including clear tubes corresponding to alpha helices, were observed in the most populated class, while only noisy, amorphous objects were reconstructed in the other classes. The highest resolution class was subjected to auto-refinement resulting in a 4.2 Å volume after post-processing using a soft mask in RELION-2.0 (Figure 4.4A and Table 4.1). The resolution of the model was calculated using the gold standard FSC=0.143 criterion and further assessed using ResMap (Swint-Kruse and Brown 2005) (Figure 4.5). The data were processed in RELION by Dr Steven Johnson.

Data collection and processing	
Voltage	300 kV
Electron exposure	48 $e^{-}/Å^2$ (K2), 50 $e^{-}/Å^2$ (Falcon 3)
Micrographs	2,088
Particles (total)	474,625
Particles (final)	97,718
Sampling	0.86 Å
Defocus range	0.5-4 µm
Refinement	
Resolution	4.2 Å
FSC threshold	0.143
Symmetry	C1
Map sharpening B factor	-227 $Å^2$

Table 4.1 Imaging and data processing statistics.

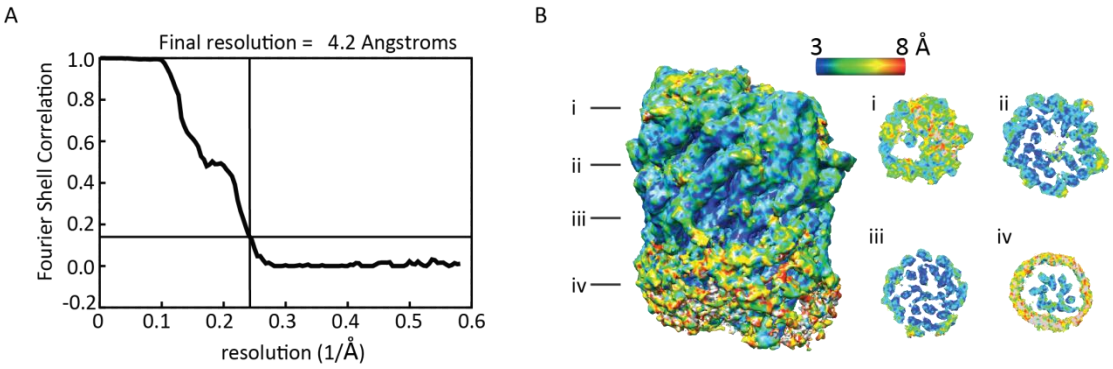


Figure 4.5 Resolution of the cryo-EM volume of *S.* Typhimurium FliPQR-FlhB. **A**, FSC curve highlighting the 0.143 threshold. **B**, ResMap representation showing variation of resolution from the core to the edge of the map.

4.2.4 Re-processing of the cryo-EM data in RELION-3.0

Following the release of RELION-3.0 (Zivanov et al. 2018; Zivanov et al. 2019), the cryo-EM data used to determine the structure of FliPQR were re-analysed. Motion correction was carried out using MotionCor2 as implemented in RELION-3.0 and CTFs were corrected using CTFFIND4 implemented in SIMPLE (Reboul et al. 2018). The previously selected particle set yielding the 4.2 Å structure was re-extracted from the micrographs using a 288x288 pixel box and after particle polishing, CTF refinement, auto-refinement and post-processing in RELION-3.0 the final resolution was 3.65 Å (Figure 4.6A) using an updated pixel size of 0.85 Å for the entire dataset. The B-factor for map sharpening was -80 $Å^2$. The final FSC curve was strongly improved and alpha helices and sidechains were more well defined (Figure 4.6B, C).

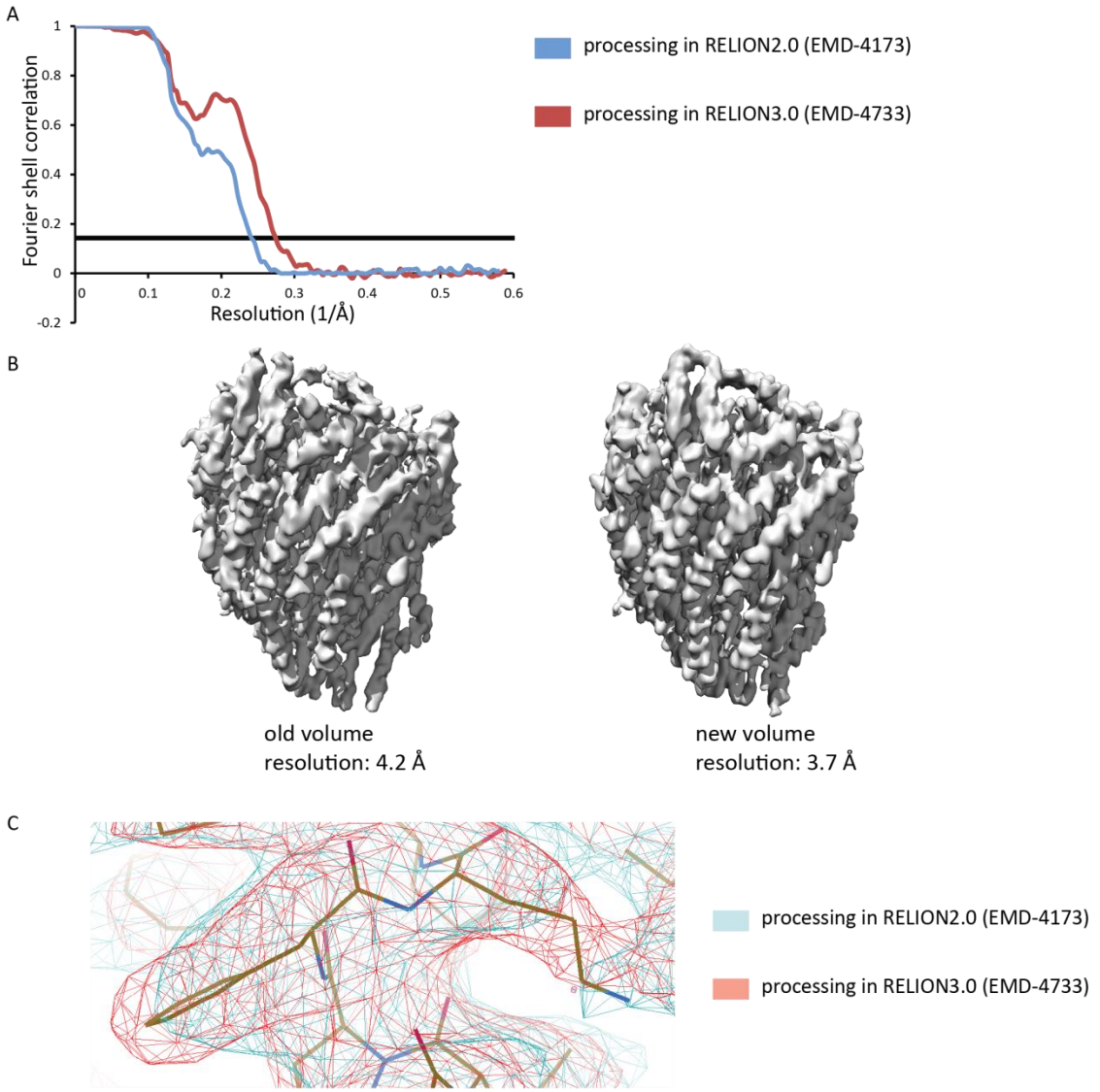


Figure 4.6 Improved cryo-EM map of FliPQR.
A, Solvent-corrected FSC curve of the *S.* Typhimurium FliPQR cryo-EM data processed in RELION-3.0 (red) and RELION-2.0 (blue). The 0.143 cut-off is shown in black. **B**, Comparison of the volumes calculated in RELION-3.0 (left) and RELION-2.0 (right). **C**, Zoomed in view of Phe221 and Lys222 of $FliP_1$.

4.2.5 Model building

The resulting map revealed a complex with a diameter of around 100 Å and a height of around 120 Å (Figure 4.7A). The volume is made up of a central six objects, five of which contain an external elaboration pointing away from centre the structure. Based on the native mass spectrometric analysis (chapter 3) and the fit of the crystal structure of a small fragment of FliP (Fukumura et al. 2017) into a part of the density (Figure 4.7B), these were assigned as FliP, while the sixth one was assigned as FliR. On the outside of each FliP, except the one to the right of FliR, a helical hairpin was observed, which was

assigned as FliQ. No density corresponding to FlhB was apparent. 5 copies of the crystal structure of the FliP "periplasmic" domain could be fit into the periplasmic density of FliP (Figure 4.7B,C), revealing that this part of FliP does not dimerise, disproving the previously proposed trimer of dimers model (Fukumura et al. 2017).

As there was no structure of a homologue of any of the components of the complex in the PDB, model building was started using models for the subunits derived from evolutionary co-variation (Ovchinnikov et al. 2015). Models were downloaded from the GREMLIN database (https://gremlin2.bakerlab.org/meta.php) and fit into the density. Although the predicted models locally fit into the map, the helices were straighter than in the map (Figure 4.8A,B). The predicted models were used as the basis for building the models of the FliPQR complex in Coot (Emsley et al. 2010).

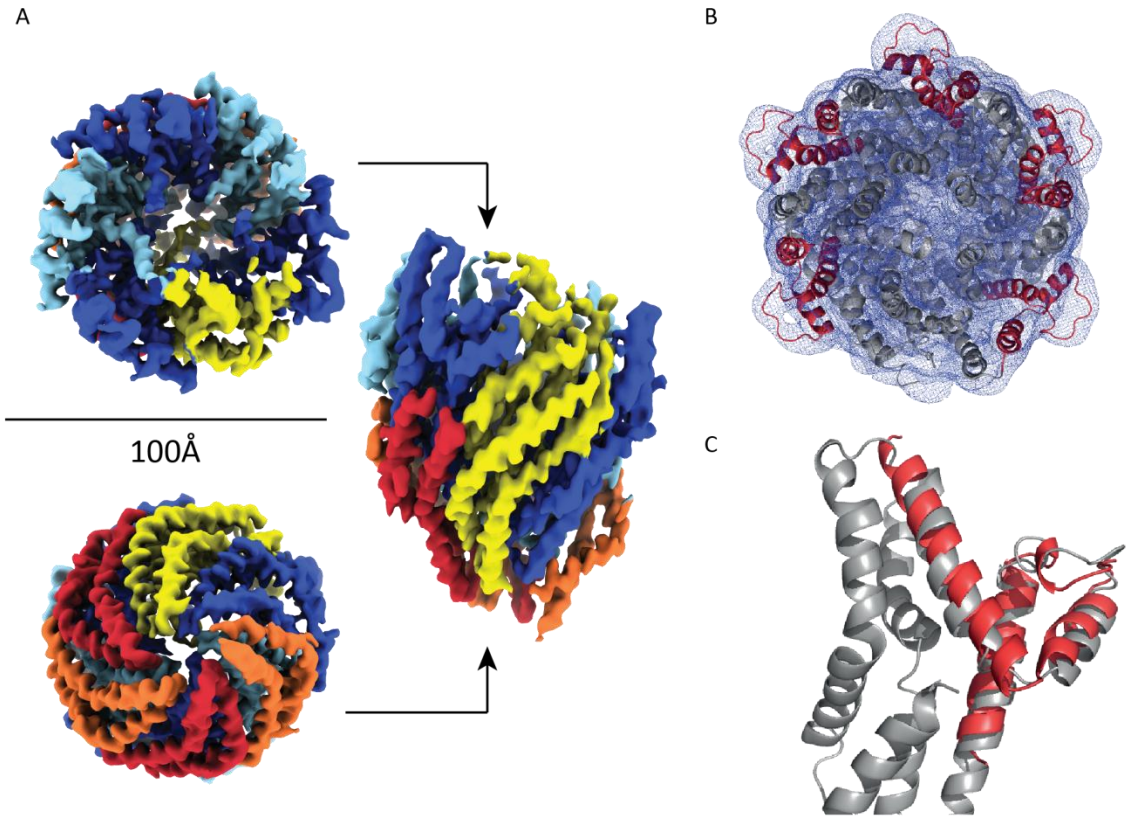


Figure 4.7 Cryo-EM map of the *S.* Typhimurium FliPQR complex.
A, The 4.2 Å map coloured according to subunit. FliP is light and dark blue, FliQ is orange and red and FliR is yellow. **B**, Five copies of the structure of a water soluble fragment of *Thermotoga* FliP (PDB: 5H72) were fit into the map. **C**, Comparison of 5H72 with the final structure of full-length *S.* Typhimurium FliP.

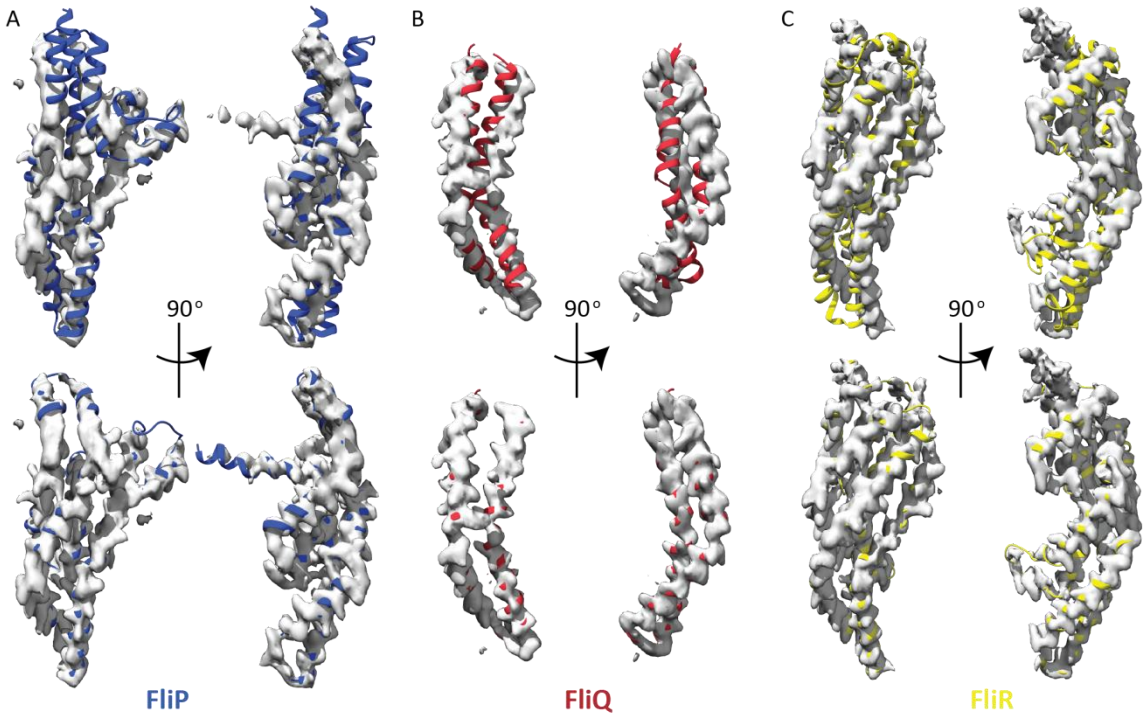


Figure 4.8 Fit of the GREMLIN models into the map.
A, **B**, **C**, Fit of the GREMLIN momdels of FliP, FliQ and FliR in the respective densities (top) and fit of the final models in the same density (bottom).

Hydrogens in the GREMLIN models were deleted in the CCP4 program pdbset (Winn et al. 2011) and the sequences were manually mutated from the *E. coli* sequence to the *S.* Typhimurium sequence in Coot. The helices of the modified GREMLIN model were fit into the density separately as rigid bodies. Subsequently, the side chains were moved into density using the Coot tool real space refine. Finally, the model geometry was refined using multiple iterations of the program phenix.real_space_refine (Afonine et al. 2018) and manual adjustment of Ramachandran plot and rotamer outliers in Coot. This model was originally built using the RELION-2.0 volume. When this model was re-refined in the RELION-3.0 map using the most recent version of phenix.real_space_refine (Afonine et al. 2018) model statistics were improved (Table 4.2), increasing confidence in the structure.

	RELION-3.0 model	RELION-2.0 model
CC_{Mask}	0.7981	0.7236
B factors	62 Å^2	227 Å^2
R.m.s. deviations		
Bond lengths	0.005 Å	0.01 Å
Bond angles	0.78°	1.18 °
Validation		
MolProbity score	2.2	2.6
Clashscore	13	25.7
Poor rotamers	0.15%	0.1%
Ramachandran plot		
Favoured	89.8%	83%
Allowed	9.8%	16.3%
Disallowed	0.4%	0.7%

Table 4.2 Model refinement statistics for FliPQR.

4.3 Structure of the FliPQR complex

4.3.1 Structural analysis of the subunits

The structure revealed that five copies of FliP and one copy of FliR form a heterohexameric, circular structure in the core of the complex. Surrounding the FliPR core are four copies of FliQ, one subunit in front of the outside face of four of the FliP subunits (Figures 4.7 and 4.9). This was consistent with the native mass spectrometry results (chapter 3). Unexpectedly, at 120 Å the complex was much taller than a typical membrane which, depending on lipid composition can have a thickness of the hydrocarbon core in the range of 25 to 35 Å (Nagle and Tristram-Nagle 2000). Furthermore, the three subunits did not adopt a canonical transmembrane fold, which would be expected to be a parallel arrangement of helices in the plane of the membrane. All three subunits form much more extended structures. FliQ consists of a single hairpin of two long, highly kinked helices, FliR of three such hairpins and FliP of two (Figure 4.9). The hairpins are offset from each other, as previously predicted by co-evolution analysis (Taylor et al. 2016). In addition, there is a small periplasmic insertion in FliP, part of which was previously crystallised (Figure 4.7C).

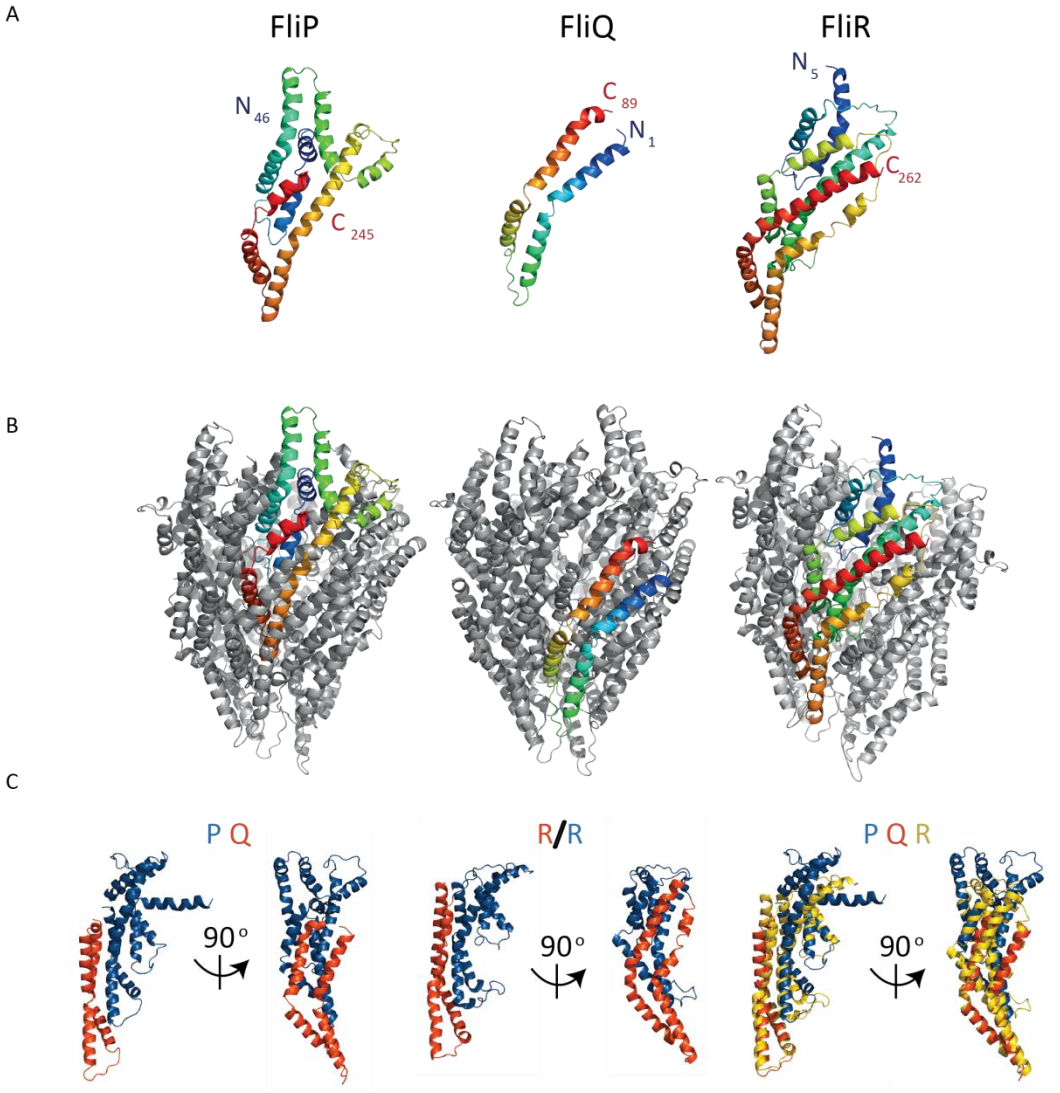


Figure 4.9 Structures of FliP, FliQ and FliR.
A, Structure of a single copy of each subunit, **B**, Structures in the context of the export gate complex. **C**, Structures of a FliP subunit with an associated FliQ subunit (left), structure of FliR coloured in orange and blue according to similarity to FliP (blue) and FliQ (orange) (middle) and overlaid structures of one FliP, one FliQ and one FliR subunit.

Closer inspection of the structure of FliR revealed that it is a fusion of the structures of FliP and FliQ (Figure 4.9C). Four of five FliP subunints are associated with a FliQ on the outside of the complex, generating a unit made up of three hairpins. The hairpin of FliR on the outside of the complex is the equivalent of a FliQ hairpin and the two FliR hairpins on the inside are the equivalent of a FliP subunit. FliP only differs through the presence of the periplasmic insertion between the two hairpins. This makes the FliPQR complex a pseudohexamer of six FliR-like, three-hairpin objects. The structural similarity of the FliP-FliQ pair and FliR might be an explanation for the observation of FliP hexamer formation in the absence of FliR (chapter 3).

4.3.2 FliPQR is a helical assembly

While analysing the map and structure, it was noticed that FliPQR does not form a flat ring. Instead, the structure is helical. As mentioned above, FliPQR can be regarded as a hexamer of 5 FliP-FliQ pairs, one of which lacks the FliQ subunit, and 1 FliR. These units are arranged along a right-handed helical axis, forming a “split-washer” shape. FliR is the topmost subunit and the FliP and FliQ subunits can be numbered from top to bottom (Figure 4.10). FliR separates the top and bottom FliP subunits.

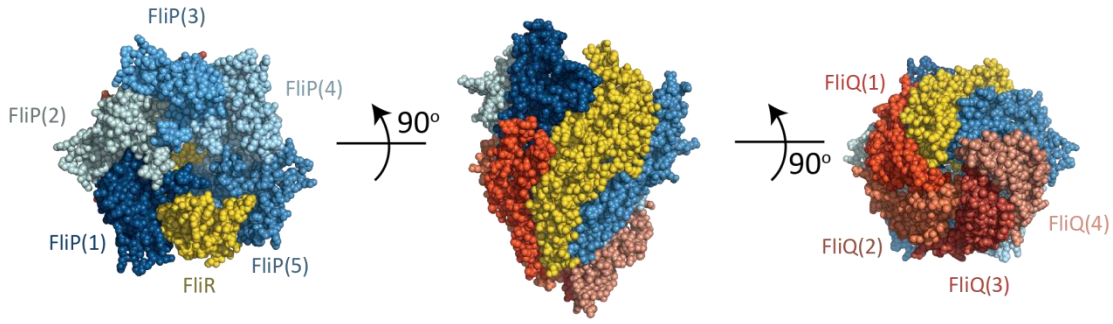


Figure 4.10 FliPQR is a helical complex.
Structure of the FliPQR complex with labels indicating the numbers of the subunits starting from the top of the complex (in the periplasm).

	Subunits per turn	Axial rise (Å)	reference
FliR to FliP (1)	5.1	5.6	this study
FliP (1) to FliP (2)	5.6	4.9	this study
FliP (2) to FliP (3)	5.8	3.2	this study
FliP (3) to FliP (4)	6	2.3	this study
FliP (4) to FliP (5)	6	3.8	this study
Average	5.7	4	this study
Flagellar filament	5.5	4.7	(Yonekura et al. 2003)
Flagellar hook	5.6	4.2	(Matsunami et al. 2016)
Flagellar rod	5.6	4.1	(Fujii et al. 2017)
T3SS needle	5.6	4.3	(Cordes et al. 2003)

Table 4.3 Helical parameters of FliPQR.

Closer analysis of the helical parameters of the complex revealed that the subunits per turn and the axial rise parameters between subunits are very close to that of previously reported filamentous structures built by type 3 secretion systems (Table 4.3). The flagellar rod (Fujii et al. 2017) and the injectisome needle (Cordes et al. 2003) are helical filaments that originate within the basal body, but it was not known how these helical structures are built on the circularly symmetric basal body (Worrall et al. 2016). The flagellar protein FliE has been proposed to be an adapter between the circular and helical symmetry (Minamino et al. 2000), and while no injectisome protein has been proposed to carry out this function, it has been associated with the cup and socket feature of the basal body (Marlovits et al. 2006). The shared helical parameters of FliPQR and the helical filaments leads to the intriguing proposal that the pseudo-hexameric 5 to 1 export gate complex, templates the 5.5 subunits per turn helical symmetry of the filament as subunits are secreted through the export gate and assemble onto its periplasmic surface. This implied that FliPQR is directly connected to the flagellar filament in a single, long helix and would rotate with the flagellum.

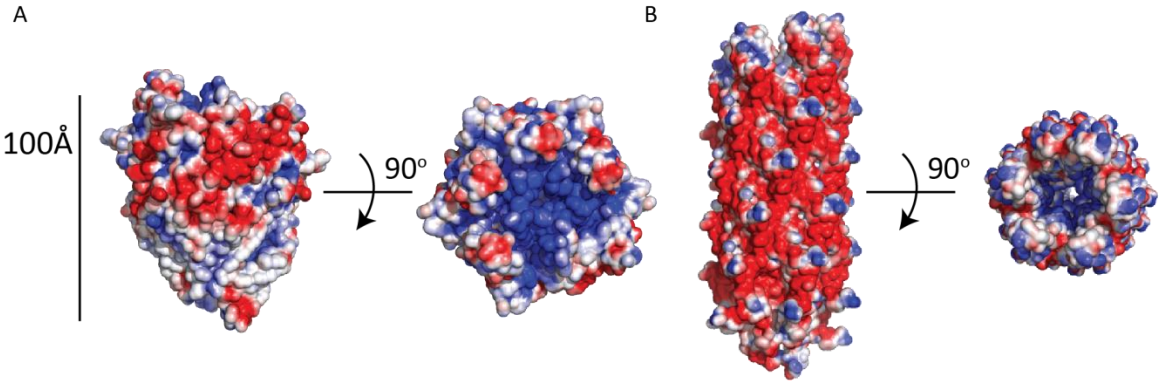


Figure 4.11 The electrostatic surfaces of export gate and needle are matched.
A, Surface of FliPQR coloured by electrostatic surface potential. The electrostatic potential surface was calculated using the APBS plugin for pymol (Baker et al. 2001). Red is negative charge, blue is positive charge and white is neutral. **B**, Surface of the needle (PDB: 6DV3) coloured by electrostatic surface potential.

Further support for the similarity of export gate and filament could be derived from an analysis of the electrostatics of the complex (Figure 4.11). This revealed a small negative charge on the outside surface of the complex and highly positively charged lumen. Similarly, the outside surface of the injectisome needle (Hu et al. 2018) is highly

negatively charged while there is a small positive charge on the inside, consistent with the idea that export gate and helical filament form a single conduit through which secreted proteins travel.

4.3.3 Hydrophobic and electrostatic interactions stabilise the FliPQR complex

Analysis of the interaction between FliPR and the FliQ subunits showed that their interface is highly hydrophobic and buries many of the residues previously predicted to be part of transmembrane helices (Figure 4.12).

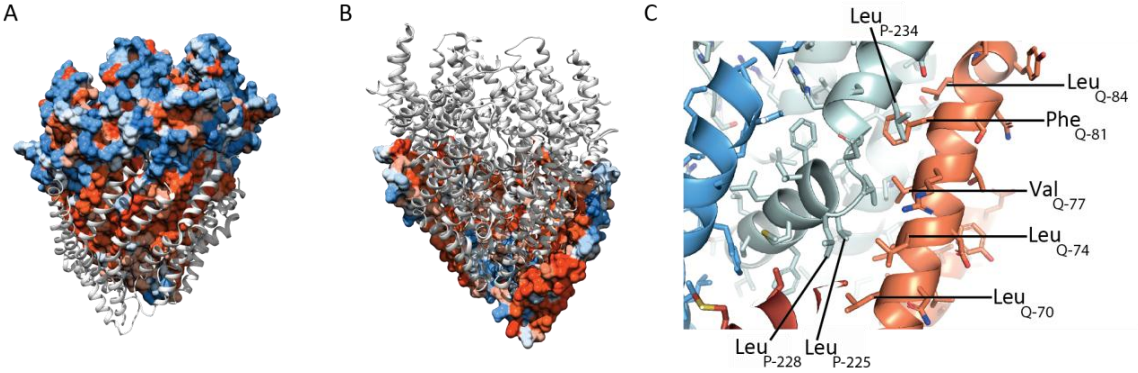


Figure 4.12 Hydrophobic interactions between FliPR and FliQ.
A, Surface of FliPR coloured according to hydrophobicity (orange=hydrophobic, blue=polar) and FliQ as grey ribbons. **B**, Surface of FliQ coloured according to hydrophobicity and FliPR as grey ribbons. The view of the complex is reversed to highlight the hydrophobic inside of the FliQ surface. **C**, Enlarged view of the interface between FliQ (orange) and FliP (blue) highlighting specific hydrophobic residues.

In addition, electrostatic interactions stabilising FliP and the interactions between the different FliQ subunits and FliR were found. Four charged residues in FliP (Arg66 and Glu178, Asp197 and Lys222) were found to form two salt bridges in the structures (Figure 4.13A). These residues were previously identified in a mutagenesis study of all perfectly conserved charges in the export apparatus and mutation of any of these caused strong impairment of motility (Erhardt et al. 2017; Ward et al. 2018). Similarly, FliQ Glu46 and Lys54 form salt bridges connecting adjacent FliQ subunits. Glu46 of $FliQ_4$ is lacking a neighbouring FliQ on one side and is instead forming a salt bridge with Arg206 of FliR (Figure 4.13B). Mutation of either of these residues in FliQ to alanine caused complete loss of motility (Erhardt et al. 2017). Basic character at Arg206 is highly

conserved (Erhardt et al. 2017) and the salt bridge with Glu46 of $FliQ_4$ is consistent with the native complex containing no more than four FliQ subunits.

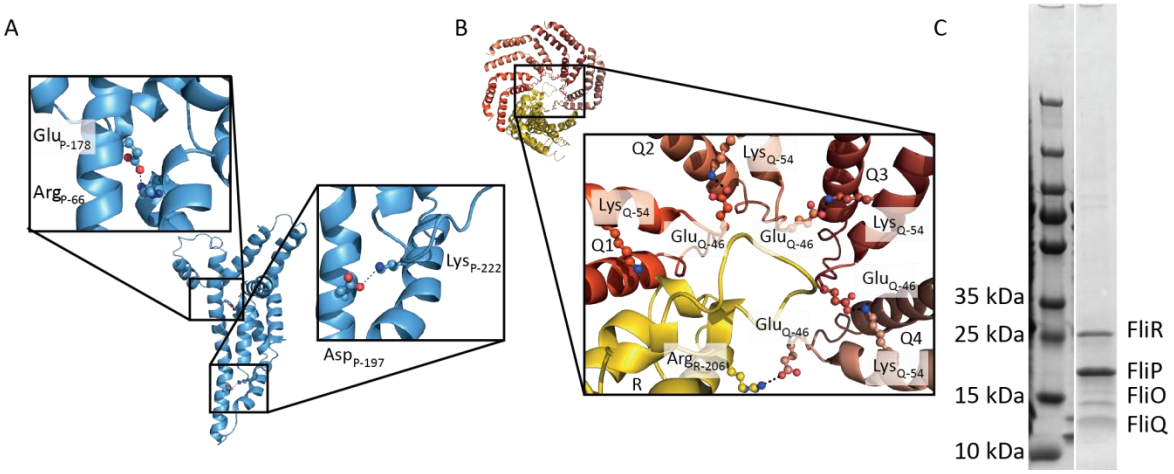


Figure 4.13 Salt bridges within the FliPQR complex.
A, Structure of FliP highlighting two salt bridges, **B**, Structures of the four FliQ subunits and FliR. Salt bridges connecting the FliQ subunits with each other and with Arg206 of the outer hairpin of FliR are highlighted. **C**, Pulldown of FliQ (E46A) with the FliPR complex. The strep eluate was run on a 4-20% gradient gel (BioRad).

Out of the three pairs of residues forming salt bridges, complete loss of motility when either residue of the salt bridge is mutated to alanine was only observed for the salt bridge formed by Lys54 and Glu46 in FliQ (Erhardt et al. 2017), therefore this salt bridge was selected for further investigation. FliQ Glu46 was mutated to alanine in the expression construct for *S.* Typhimurium FliOPQR and the complex was purified (section 2.5.3). Interestingly, FliQ could still be co-purified with FliP and FliR despite the loss of the salt bridge, suggesting that the loss of motility is not due to disrupted assembly of the complex but may instead be related to loss of function in the opening of the complex to allow passage of secretion substrates. A possibility that could not be excluded is that the mutant complex is not integrated into the flagellum due to quality control that could be carried out by the chaperone FliO (Barker et al. 2010; Fabiani et al. 2017).

4.3.4 Structural analysis of the interaction with the membrane

As the dimensions of both the individual subunits and the assembled complex were not consistent with a location in the membrane, the interaction with the membrane was further analysed. A band of hydrophobic residues could be seen at the bottom of the complex (Figure 4.14A), this was also the location of the detergent belt seen in the cryo-EM volume (Figure 4.14B).The width of the hydrophobic band is around 30 Å, consistent with the size of a lipid bilayer (Nagle and Tristram-Nagle 2000). The topology of the subunits in relation to the lipid bilayer was redrawn based on the location of the detergent belt, demonstrating that many of the predicted transmembrane helices are in fact not membrane associated (Figure 4.14C).

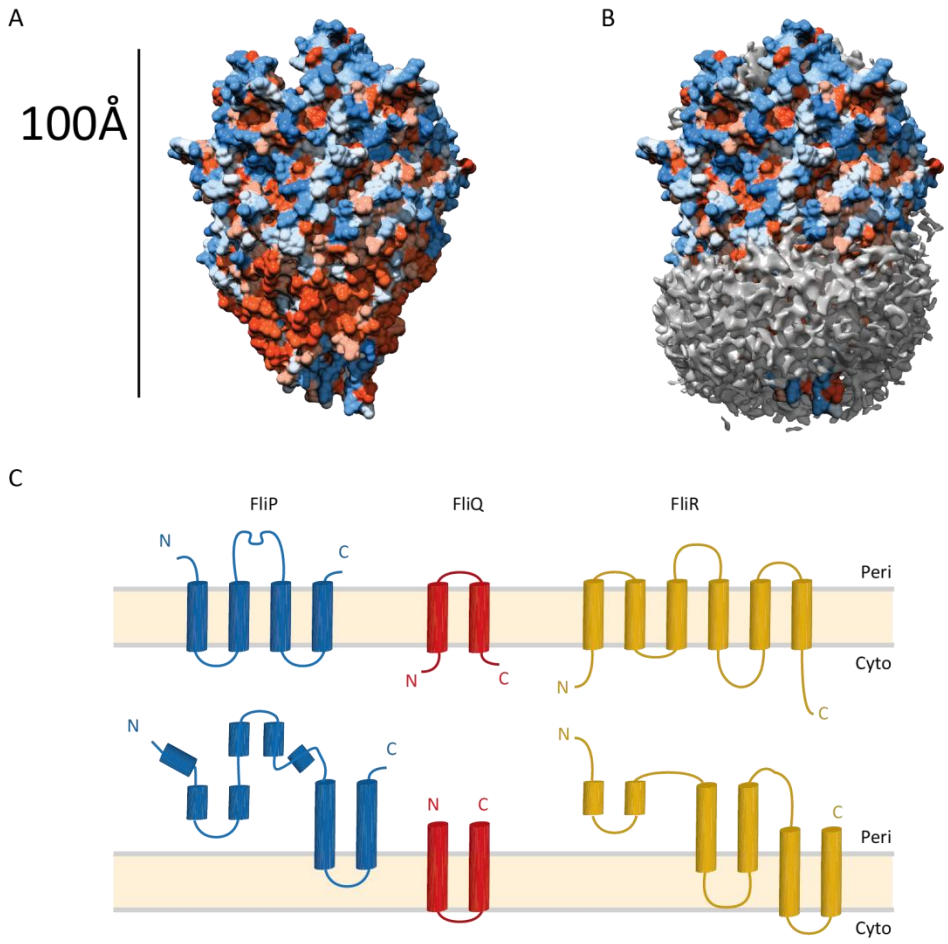


Figure 4.14 Topology of FliP, FliQ and FliR in the membrane.
A, Surface of the structure of FliPQR coloured according to hydrophobicity. Orange=hydrophobic, blue=polar. **B**, Surface as in (**A**) with the detergent belt of the cryo-EM map. The detergent belt map was generated as a difference map of the cryo-EM map minus a 10 Å map generated from the atomic model in Chimera (Pettersen et al. 2004). **C**, Topology of the FliP, FliQ and FliR proteins as predicted (Erhardt et al. 2017) (top) and as observed in the structure (bottom).

4.4 Placement of the structure in the basal body

Due to the unusual geometry of the FliPQR complex and its helical nature the localisation in the assembled nanomachine was further investigated. As mentioned in the introduction, FliPQR was considered to be located in a specialised patch of inner membrane inside the basal body (Macnab 2004). Recently, high resolution reconstructions of the basal body of the vT3SS have become available (Worrall et al. 2016), enabling fitting of the FliPQR model into those maps. The basal body has long been recognised to house a central structure known as cup and socket (Schraidt and Marlovits 2011), but its exact composition and function were not known due to the poor resolution in this area of the older reconstructions. Comparing the most recent high-resolution reconstruction to the FliPQR structure revealed that the export gate complex matches the cup and socket density well (Figure 4.15A). Surprisingly, this placed even the lowest point of the complex above the predicted plane of the inner membrane.

In order to further confirm this extra-membrane localisation of a membrane protein, collaborators in Professor Samuel Wagner's group in Tübingen performed structure based *in vivo* cross-linking and cysteine mutant cross-linking experiments in the *S.* Typhimurium SPI-1 vT3SS. Residue Glu138 of the basal body subunit SctJ was mutated to the unnatural amino acid *p*Bpa and could be cross-linked to the FliPQR homologues SctR, SctS and SctT (Figure 4.15A, B). In addition, they purified *S.* Typhimurium SPI-1 basal bodies and treated them with the chemical cross-linker DSS followed by mass spectrometric detection of cross-linked peptides. This revealed that SctR could be cross-linked to a periplasmic residue in the secretin SctC.

The localisation of the export gate in the cup and socket density, which has previously been suggested as an anchor for the rod and needle in vT3SS (Schraidt and Marlovits 2011), suggests a mechanism for how the needle may cause activation of secretion in response to target recognition. The direct contact between rod and top of the export gate could facilitate opening of the gate.

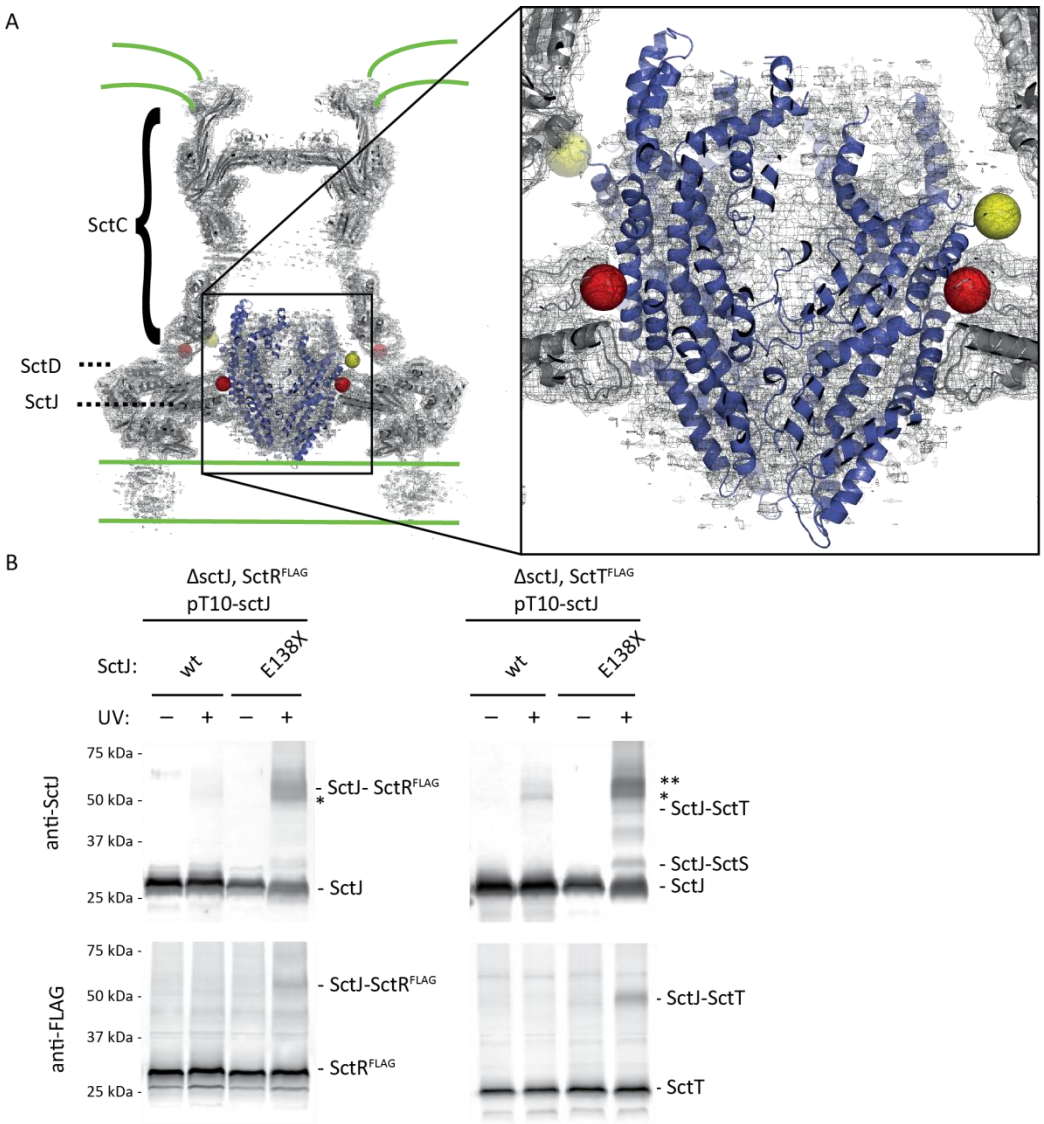


Figure 4.15 FliPQR is an extra-membrane core component of the basal body.
A, FliPQR (blue cartoon) fit into the cryo-EM map of the vT3SS basal body (grey mesh and cartoon) (Worrall et al. 2016). Yellow spheres indicate residues of SctR that could be cross-linked to the basal body and red spheres indicate residues in SctC/SctJ that could be cross-linked to SctRST. **B**, Collaborators' western blot data showing *in vivo* photocrosslinking of SctR and SctT to SctJ. There was also a potential cross-link between SctJ and SctS (top right gel), but this could not be confirmed due to difficulties in raising antibodies against SctS or introducing an affinity tag.

As a high resolution structure of the flagellar basal body is not available, it was not possible to analyse how FliPQR may interact with FliF, the flagellar equivalent of SctDJ (Bergeron 2016), but a similar mode of interaction is likely. It is not clear how the helical pseudosymmetry of FliPQR interacts with the circular symmetry of the basal body.

4.5 Interactions between the export gate and other proteins

4.5.1 Bioinformatic analysis of the interaction of FliPQR with the rod

If the export gate templates the helical filaments, as proposed above, the top of the export gate complex would be expected to interact with the rod (flagella) or inner rod (injectisome), which is the first helical filament to assemble. Indeed, such an interaction has been shown previously using *in vivo* cross-linking (Dietsche et al. 2016) between the rod protein SctI and the FliPR homologue SctRT. Mapping the points of cross-linking onto the structure revealed that they were all located at the top of the complex, consistent with a model in which the rod protein assembles onto the export gate (Figure 4.16A).

The identity of the SctI homologue in flagella is less clear as low sequence conservation complicated direct identification of homologues in the flagellar system. The sequence of SctI from *S. flexneri* was submitted to the jackhmmer server (Potter et al. 2018). After 13 iterations of searching the uniprot database, a hit annotated as the flagellar protein FliE from *Actinobacterium* (uniprot entry: A0A1Q7AW81_9ACTN) with an E value of 0.00092 was found. Furthermore, the secondary structure predicted for SctI and FliE is similar (Figure 4.16B). Evolutionary co-variation of FliE showed a clear antiparallel interaction between the two predicted helices, suggesting that they form a hairpin and results for SctI are similar, despite a small number of homologous sequences (Figure 4.16C). SctI has previously been suggested to have a fold similar to the needle subunit SctF (Cordes et al. 2003), which forms a hairpin of alpha helices (Loquet et al. 2012).

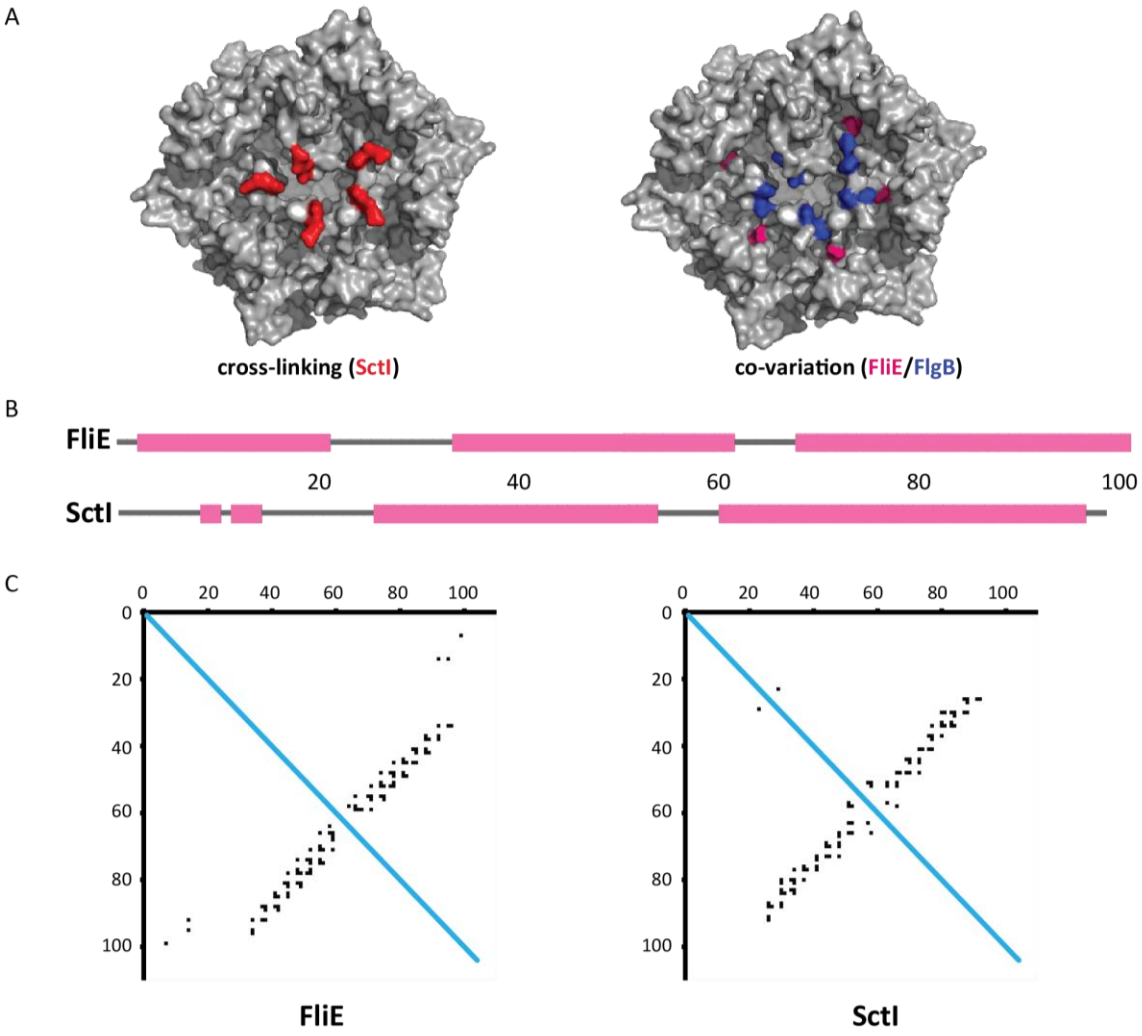


Figure 4.16 FliE is the flagellar equivalent of SctI.
A, Top view of the FliPQR structure (surface) with residues highlighted that have been shown to cross-link (Dietsche et al. 2016) to SctI (left, red) and highlighting of residues that co-vary (right) with FliE (pink) or FlgB (blue) with a probability score of at least 0.9, calculated using the GREMLIN server. **B**, Secondary structure prediction of FliE (*S.* Typhimurium) and SctI (*S. flexneri*) using PSIPRED (Buchan et al. 2013). Numbers indicate the residue number in the sequence. **C**, Evolutionary co-variation within the same sequences calculated using RaptorX (Wang et al. 2016). Predicted contacts with a probability score greater than 0.5 are displayed. The multiple sequence alignment for FliE contained 2089 sequences, and the multiple sequence alignment for SctI contained 114 sequences.

In addition to FliE, the flagellar rod is made up of FlgB, FlgC and FlgF (proximal rod) and FlgG (distal rod) (Fujii et al. 2017). Cryo-tomography of flagella in various stages of assembly suggests that the proximal rod is assembled sequentially, starting with FliE and continuing with FlgB and then FlgC (Zhao et al. 2013), consistent with an interaction between FliE and the export gate. Thus evolutionary co-variation between FliPQR and

FliE as well as the other flagellar rod components FlgB and FlgC was calculated using the GREMLIN server (Ovchinnikov et al. 2014). In order to avoid false positives, only contacts with a probability score greater than 0.9 were considered. This revealed contacts between the N-terminus of FliP and both FliE and FlgB (Table 4.4).

Residue in FliP	Residue in FliE
V42	G85
L58	V99
Residue in FliP	Residue in FlgB
L45	V133
T44	V133

Table 4.4 Evolutionary co-variation between FliPQR and rod proteins.

All residues of FliPQR that were predicted to interact with the flagellar rod/injectisome inner rod were at the periplasmic, top part of the complex (Figure 4.16A). This was consistent with the model that the rod assembles onto the export gate. Co-variation was observed with both FliE and FlgB. A genetic interaction between FliE and FlgB has been shown (Meshcheryakov et al. 2013) and interestingly mutation of FliE Val99, close to the C-terminus, disrupted motility. This residue co-varies with FliP, suggesting that mutations here could affect motility through disruption of the export gate – rod interaction. These results suggested that FliE is not an adapter between the helical filament and the circularly symmetric basal body, but instead is the first subunit of the flagellar rod.

4.5.2 Bioinformatic analysis of the interaction of FliPQR with FlhB

The switch protein, FlhB, is another component of the export apparatus. Its cytoplasmic domain interacts with secretion substrates just before they are unfolded and enter the secretion channel (Evans et al. 2013), but the structure and location of its transmembrane domain are not known.

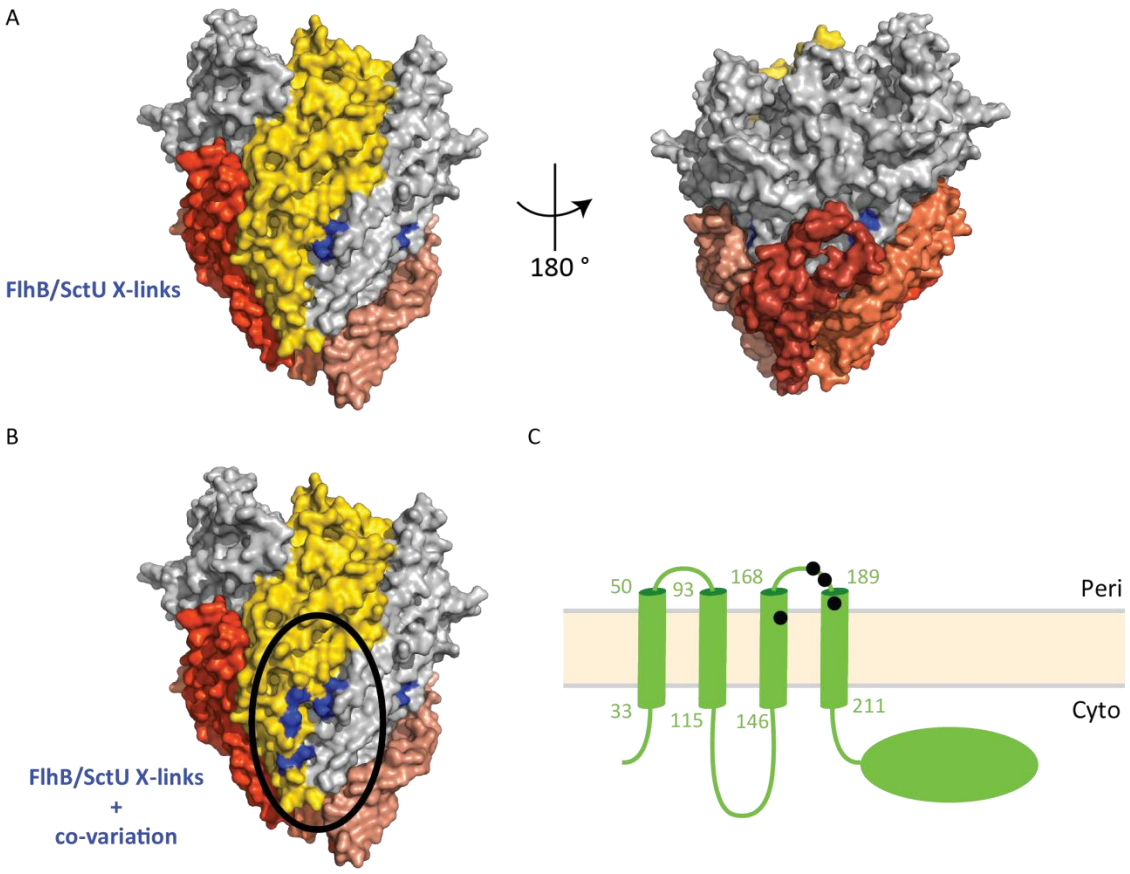


Figure 4.17 Interaction site of FlhB on the FliPQR complex.
A, Surface of the FliPQR complex. Yellow = FliR, shades of red = FliQ, grey = FliP. The residues shown to interact with FlhB by *in vivo* cross-linking (Dietsche et al. 2016) are highlighted in blue. **B**, Surface representation of the FliPQR complex as in (**A**) with the additional residues of FliR listed in Table 4.5 also highlighted in blue. The proposed binding site of FlhB is highlighted by a black ellipse. **C**, Topology of FlhB. The transmembrane helices were predicted using TMHMM (Krogh et al. 2001) and the soluble domain was placed in the cytoplasm (Evans et al. 2013). Black dots indicate the approximate predicted positions of the residues listed in Table 4.5.

In vivo cross-linking has been used to demonstrate an interaction between the vT3SS homologue of FliP (SctR) and FlhB (SctU) (Dietsche et al. 2016). Two residues, Met224 and Val227 in *S.* Typhimurium FliP numbering, were found to cross-link with the homologues of both FlhB and FliQ. Due to the stoichiometry of FliP, there are five potential interaction sites in the complex (Figure 4.17A), but most of these are occupied by a FliQ subunit. Evolutionary co-variation between FliPQR and FlhB was calculated as for FliE above and high probability contacts were only found in FliR (Table 4.5). Mapping of these four additional contacts between the two proteins onto the structure of the complex revealed that the interaction site of FlhB is on the surface of the complex where FliR and $FliP_5$ meet (Figure 4.17B). The residues in FlhB predicted to bind this site are at

the predicted periplasmic termini of helix 3 and 4 (Figure 4.17C). FlhB has been predicted to be a four helix bundle (Taylor et al. 2016). An interaction between FlhB and FliR has been hypothesised for a long time due to the existence of a genetic fusion of the two proteins in *Clostridium* (Macnab 2004). The cytoplasmic domain of FlhB, $FlhB_C$, is thought to interact with substrates before they are secreted by the T3SS (Evans et al. 2013) and its transmembrane domain may interact with the hydrophobic surface of the FliPQR complex.

Residue in FliR	Residue in FlhB
L188	A182
T192	A186
A200	T160
L199	V190

Table 4.5 Evolutionary co-variation between FliPQR and the switch protein FlhB.

4.6 The export gate structure is closed at multiple points

The export gate is thought to be the channel through which secretion substrates pass the inner membrane (Dietsche et al. 2016; Fukumura et al. 2017). The structure is consistent with the export gate being part of the secretion channel, but the putative path of export substrates is blocked at multiple points (Figure 4.18A). The lumen inside the complex is constricted by the N-terminal helices of FliP at the top, however the conformation of the FliP N-terminus may change once the rod assembles on top of it. In addition, there are three closure points at the bottom of the complex (Figure 4.18B): At the cytoplasmic entrance into the export gate the “Q-latch” is formed by a loop of FliQ and is held closed by a network of salt bridges. Above this, highly conserved methionine residues of FliP form a gasket, blocking the channel to leakage of ions. These methionines may block leakage in the open state of the complex while an unfolded polypeptide chain is secreted through the gasket. Finally, a long loop of FliR (residues 106 to 123) plugs the channel. Suppressor mutations that restore motility in a FlhA mutant background have been isolated in two sites in FliR (Hara et al. 2011). Mutation of Gly103 and Gly117, close to or within the FliR plug respectively (Figure 4.18C), may allow secretion by impairing the gating function carried out by the FliR plug. The residues of

the methionine gasket have been characterised (Ward et al. 2018). Deletion of one of the methionine residues increased conductance of ions, consistent with the central localisation observed here.

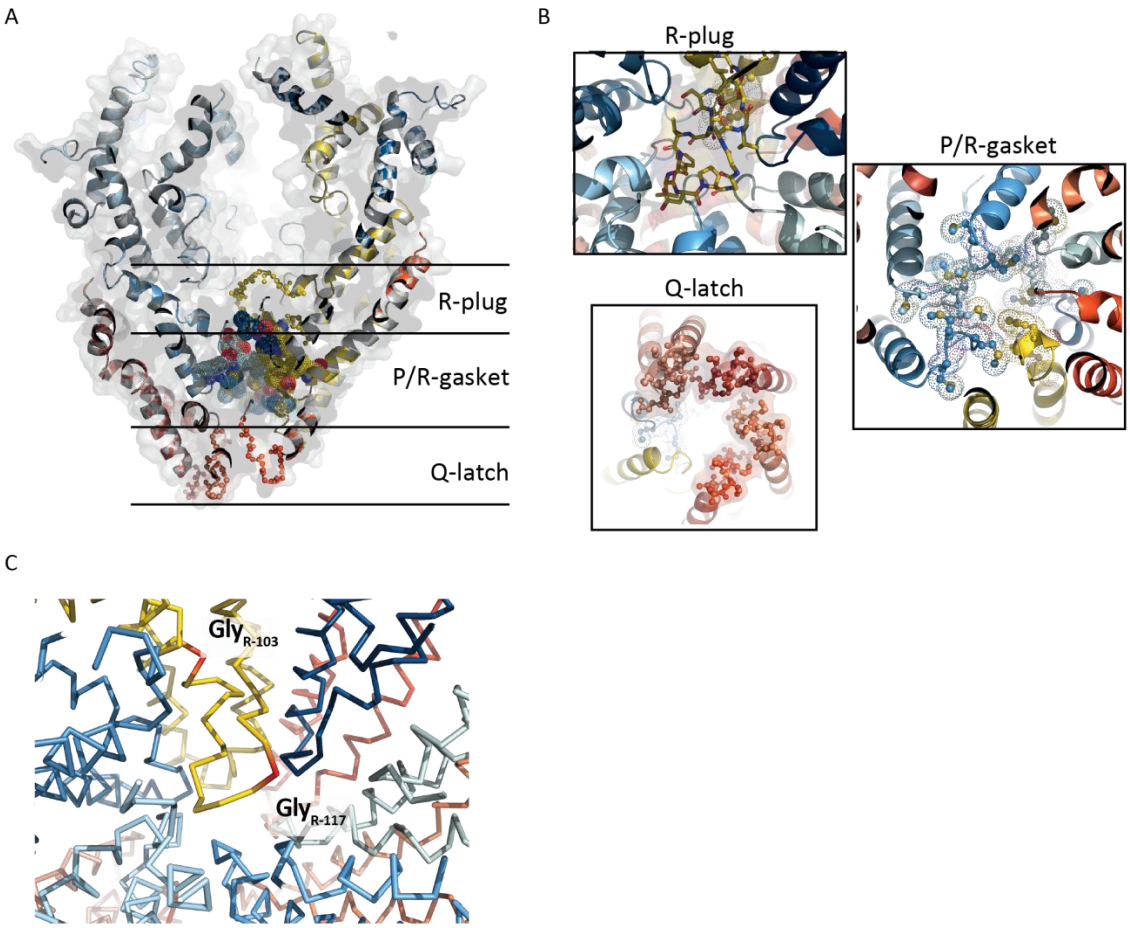


Figure 4.18 The FliPQR complex is gated at multiple sites.
A, Section of the FliPQR complex showing multiple constriction points. **B**, Zoomed view of the three constrictions. Residues FliR 106-125 (R-plug), FliP 209-211 and FliR 214 (P/R-gasket) and FliQ 40-46 (Q-latch) are highlighted. **C**, FliR plug with residues Gly103 and Gly117 highlighted in red.

The degree to which the aperture of the complex is closed is also determined by the highly kinked helices in each subunit (Figure 4.19A). Should these bends relax, the complex would open. These helices were less kinked in the structures predicted by GREMLIN (Figure 4.8) and replacing the structures modelled based on the cryo-EM density with the GREMLIN models did produce an open aperture (Figure 4.19B). However, the structure of the open state remains to be determined experimentally.

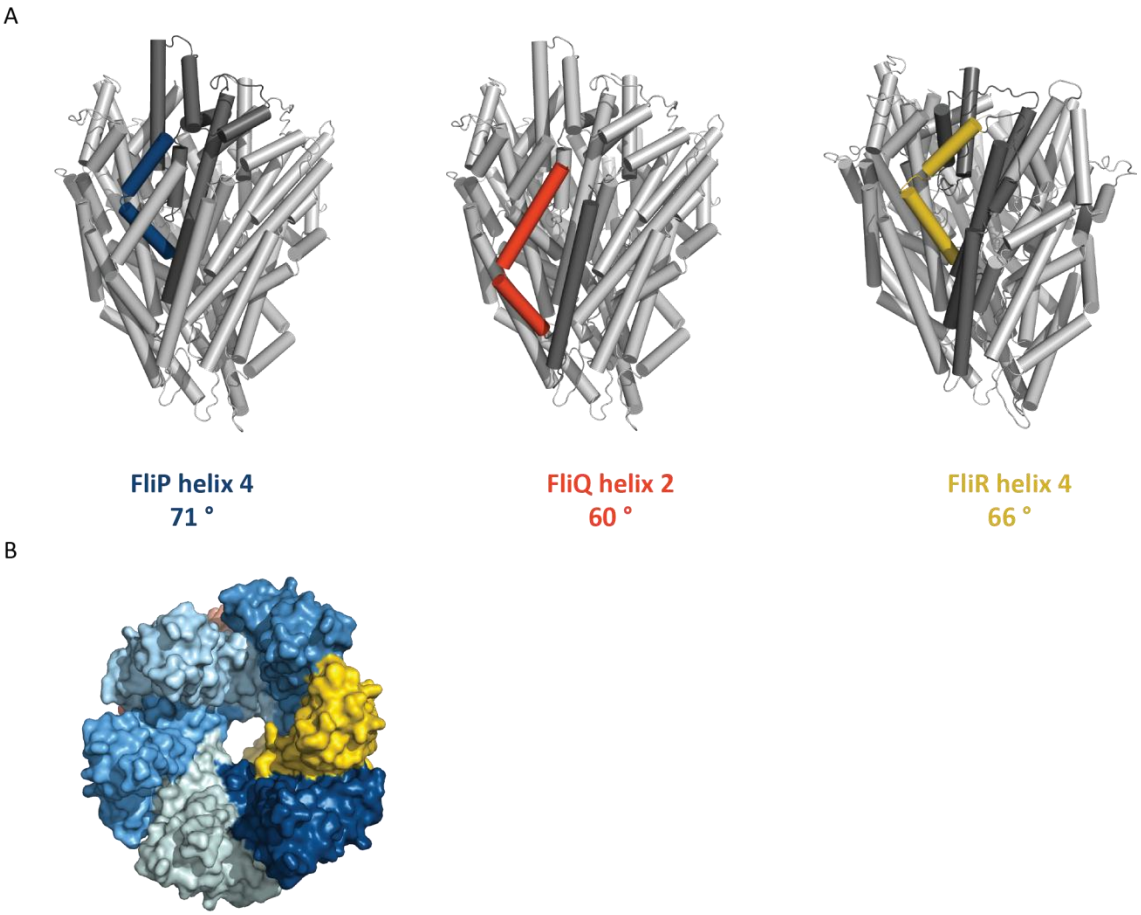


Figure 4.19 Kinked helices in FliPQR hold the complex closed.
A, Structure of FliPQR (cylinder helices) highlighting the angle to which helix 4 in FliP (left), helix 2 in FliQ (middle) and helix 4 in FliR (right) are bent. **B**, Replacement of the structures of FliP, FliQ and FliR with the models predicted by GREMLIN (Ovchinnikov et al. 2015) produces a possible model for an open state.

When overexpressed, SctR or FliP alone have been reported to allow passage of ions, resulting in toxicity (Dietsche et al. 2016; Ward et al. 2018), presumably by forming a pore, but this is not consistent with the closed structure of the FliPQR complex. In order to test whether the full export gate is closed while it resides in the membrane, prior to assembly into the basal body, a simple test of toxicity was designed. *E. coli* Mt56 cells carrying pT12 plasmids containing export gate genes were grown at 37 °C, 180 rpm in 100 ml of TB media supplemented with kanamycin (60 µg/ml) up to an OD600 of 0.6. Each culture was split into two halves in different flasks and expression in one half was induced using 0.1% (w/v) rhamnose monohydrate. Consistent with previous results (Dietsche et al. 2016), expression of *S.* Typhimurium SPI-1 SctRST and *S.* Typhimurium FliOPQR caused a growth defect (Figure 4.20A, B).

However, recombinant overexpression may introduce artefacts. The RBS of the first gene in the operon in both expression plasmids was an artificial, strong RBS, which might result in an excess of gene product from the first gene (Lim et al. 2011), potentially resulting in excess FliO, which is thought to be toxic (McMurry et al. 2004) or excess SctR, which may form a non-physiological complex like the FliP hexamer (Dietsche et al. 2016; Fukumura et al. 2017). Therefore, the entire FliLMNOPQR operon was expressed with the native RBS upstream of FliL (Figure 4.20B). No toxicity was observed. FliPQR could be purified successfully from this construct and its nMS spectrum was identical to FliPQR purified from the FliOPQR construct (Figure 3.4B and Table 3.1). This suggests that the FliPQR complex in the membrane, an assembly intermediate that can remain in the membrane until the basal body assembles around it, is closed and does not allow leakage of ions. A complication could be that this experiment does not exclude the possibility that FliL, FliM or FliN keep the FliPQR pore closed, but no interaction between FliPQR and FliL has previously been proposed and FliM and FliN are cytoplasmic proteins (McDowell et al. 2016).

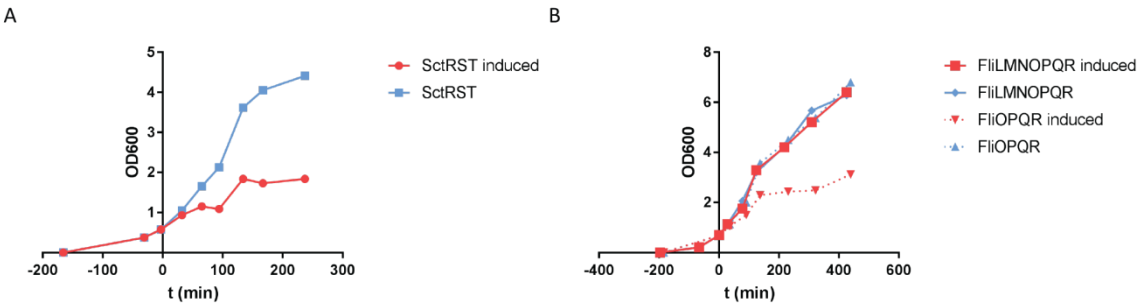


Figure 4.20 Toxicity of different export gate constructs.
A, Growth curves of cells expressing *S. flexneri* SctRST. **B**, Growth curves of cells expressing *S.* Typhimurium FliOPQR or the entire *S.* Typhimurium FliLMNOPQR operon.

4.7 Discussion

In this chapter the structure of the flagellar export gate from *S.* Typhimurium was determined, revealing new insights into the function and assembly of type 3 secretion but also providing implications for membrane protein complexes in general.

The structure of FliPQR was determined from a sample containing FlhB. Unfortunately, the FlhB protein was not visible in the cryo-EM density map resulting from this sample. This could have been due to the data analysis, i.e. the computational selection of particles could have excluded particles that include the FlhB subunit. However, despite extensive re-processing of the data, no class of particles with additional density compared to the FliPQR map could be identified. Another possibility is that the binding of FlhB to FliPQR is weakened during the grid preparation process and the particles fall apart, leaving behind only the stable FliPQR complex. It is further possible that particles including FlhB have a higher affinity to the carbon of the cryo-EM grids and are not entering the holes. If FlhB is not part of the complex, why were top and side views observed when only top views could be seen in a sample made without FlhB (Figure 4.2)? The sample was made from the left hand side shoulder of the FliPQR peak eluting at 10.8 ml on a Superdex 200 increase 10/300 column (GE). It is possible that the detergent concentration is different in this fraction than in the main peak, which elutes at 11.5 ml. 1% (w/v) LMNG in the absence of protein forms micelles eluting at 12 ml.

Although highly unusual, the overall structure is consistent with native mass spectrometry data and confirms the model constructed based on it (chapter 3). The stoichiometry of the *S*. Typhimurium sample could not previously be determined, but here four FliQ subunits were observed in addition to the five FliP and one FliR subunits in the complex.

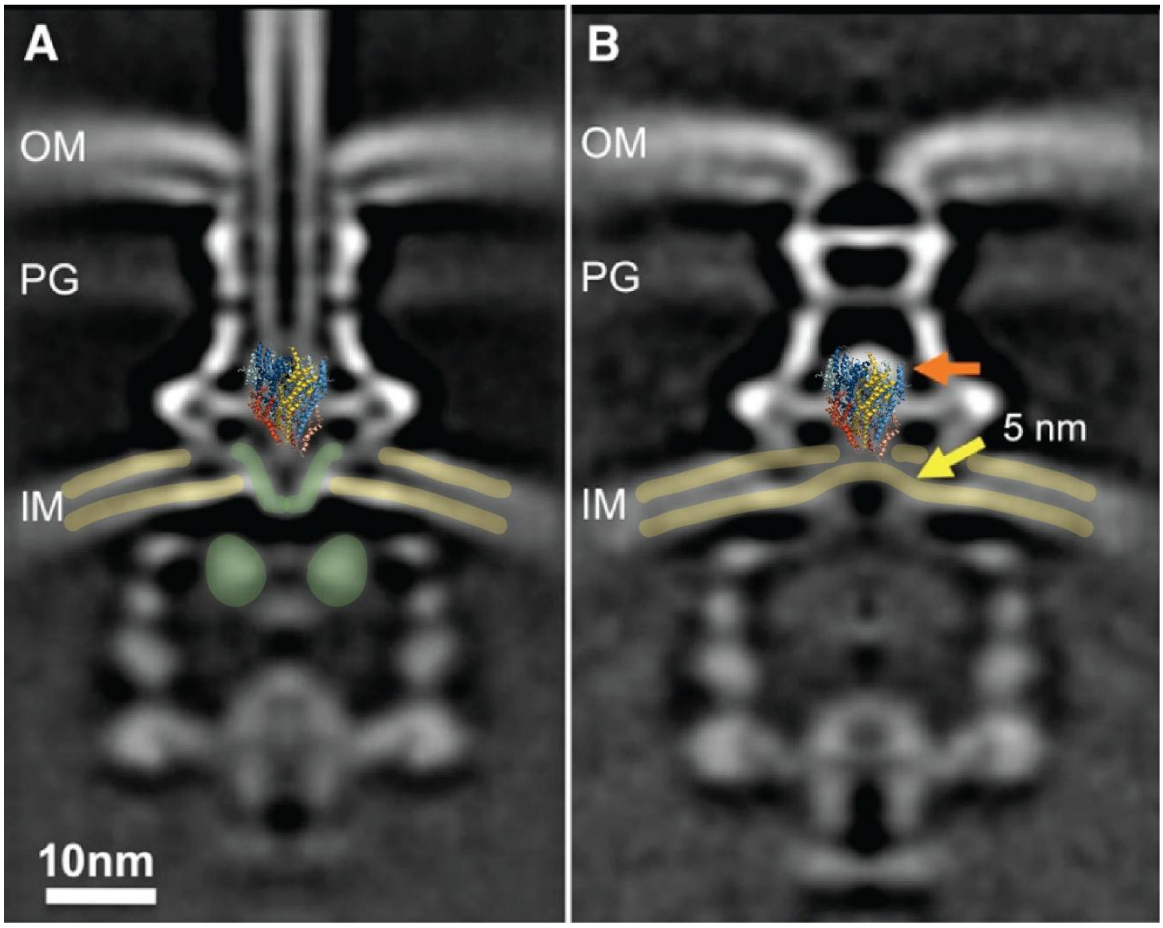


Figure 4.21 Re-interpretation of previous tomographic data.
FliPQR in the wild type (**A**) and ΔSctV (**B**) injectisome. Green shading indicates the position of SctV and yellow shading indicates the inner membrane. Figure taken from (Hu et al. 2017).

Unexpectedly, FliPQR is not a flat pore in the membrane, but a helical complex held in the periplasm by the basal body around it (Figure 4.10 and Figure 4.15). Although the helicity of the complex explains how the helicity of the helical filaments is templated on the export gate (Table 4.3) it is not clear how exactly FliPQR interacts with the circular basal body and if any of the hydrophobic surface that is covered by the membrane prior to assembly is buried by interactions with the basal body or another protein. Inspection of previous tomographic data in light of the results from this chapter suggests when the EA protein FlhA (SctV in vT3SS) is not present, the membrane reaches up to contact FliPQR, burying the exposed hydrophobic surface of the FliQ ring (Figure 4.21). However, in the presence of SctV, a V-shaped density can be seen underneath FliPQR. This may represent the SctV transmembrane domain and may bind to FliPQR/SctRST. As the export gate is likely to expand when the channel opens, rather than direct binding of the

two proteins, there may also be lipids in between FliQ/SctS and the FlhA/SctV transmembrane domain to allow more flexibility. No co-evolutionary contacts could be found between FliPQR and FlhA using the GREMLIN server (Ovchinnikov et al. 2015) and *in vivo* cross-linking experiments carried out in Samuel Wagner's laboratory did not identify any clear interaction.

The most unexpected aspect of the structure is that the subunits are not canonical transmembrane proteins. The structure resolves the conflicting older reports of topology (see Introduction), but despite the previous uncertainty it is very unusual for topology prediction programs like TMHMM (Krogh et al. 2001) and TOPCONS (Tsirigos et al. 2015), used for prediction of the subunit structures (Taylor et al. 2016), to predict multiple transmembrane helices with high confidence in a protein that has none, implying that topology predictions of membrane proteins known to be part of a larger complex should be interpreted with more care than is currently done. This incorrect prediction is partially explained by the many hydrophobic residues that are buried by protein-protein interactions within the complex (Figure 4.12).

The complex is unlikely to be inserted into the membrane in an assembled state. What is the conformation of the subunits immediately after synthesis? The structures of each monomer extracted from the complex could not be fit into a 30 Å wide lipid bilayer without also burying some hydrophilic patches (Figure 4.22A). This is partially due to the offset between the hairpins making up each structure (Figure 4.9). This offset has previously been predicted using evolutionary co-variation data (Taylor et al. 2016), but as noted by the authors in that study, it is difficult to reconcile the resulting structure with a position in the inner membrane. It is possible that following synthesis the monomers adopt a more canonical transmembrane fold in which the FliP and FliR hairpins are not offset from each other to the same degree as observed in the FliPQR complex structure. Although it is challenging to place the monomer structures in the membrane, the $FliP_5R_1$ complex is a possible intermediate as its hydrophobic surface can be placed in the membrane. As the hydrophobic surface around the FliPR and FliPQR complexes is not planar, these assembly intermediates may be at an angle relative to the membrane.

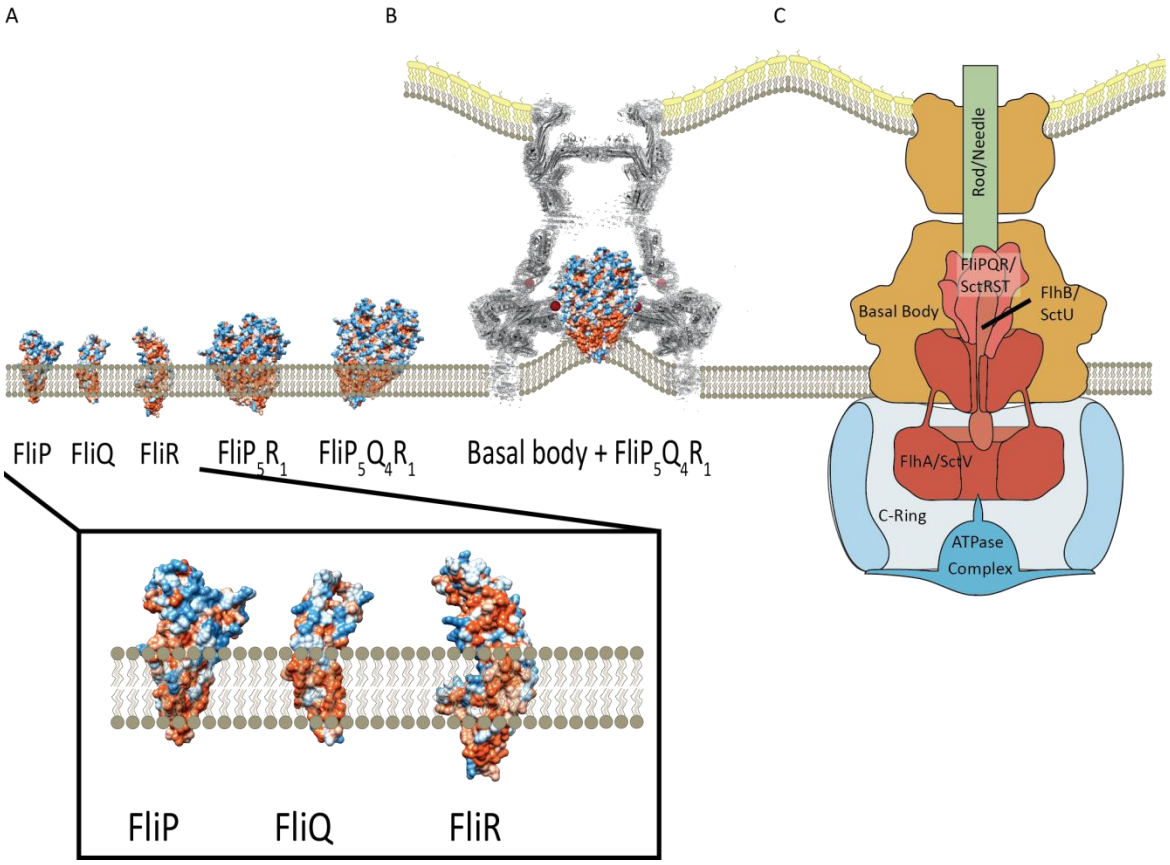


Figure 4.22 Assembly of the export gate.
A, Structures of isolated FliP, FliQ, FliR, $FliP_5R_1$ and $FliP_5Q_4R_1$ placed in the inner membrane in the same conformation as they adopt in the FliPQR complex. Bottom: Close-up view of the protomers of FliP, FliQ and FliR from the FliPQR complex structure tentatively placed in a 30 Å lipid bilayer. **B**, FliPQR placed in a reconstruction of the basal body (Worrall et al. 2016) in the absence of the other EA components, illustrating the bending of the inner membrane to cover the exposed hydrophobic surface of FliPQR (Figure 4.21). **C**, Cartoon model of the assembled injectisome with the export gate complex in the centre of the basal body.

Once FliPQR is fully assembled it is still in the inner membrane. The large periplasmic part of the structure would not fit through the bottom of an assembled basal body (Figure 4.22B,C), implying that the export gate cannot be inserted into pre-assembled basal bodies, and this has been demonstrated experimentally (Wagner et al. 2010). One hypothesis for insertion of the export gate in the basal body is therefore that coordinated assembly of the basal body around the export gate, which may nucleate basal body assembly, would lift the export gate out of the plane of the inner membrane. This would also be consistent with the order of assembly observed in the flagellum (Jones and Macnab 1990).

The structure determined here is closed. In particular, a previously identified methionine rich loop (Ward et al. 2018) is forming a gasket at the centre of the pore, a plug is formed by a loop of FliR and a Q-latch is formed by the cytoplasmic loops of the FliQ subunits and is tied together by salt bridges (Figure 4.18). Mutation of either of the residues in the FliQ salt bridge causes loss of motility (Erhardt et al. 2017), however, the same mutation had no effect when purifying FliPQR E46A. As mutation does not appear to disrupt assembly, it may instead prevent motility by disrupting the regulation of the Q-latch.

How might the complex open? There is no feature of the structure that clearly suggests a function in opening in response to a signal such as binding of substrates or recognition of a target host cell. One possibility was that the complex is permanently opened once the rod assembles on top of it and that subsequently the gating function is carried out at another point in the nanomachine, but a later cryo-EM volume of the needle complex, the basal body with an attached rod and needle, showed that the export gate is also closed in this environment (Hu et al. 2018). Another possibility is that the other EA proteins, FlhB and FlhA, have a role in opening. Unlike FliPQR, FlhB and FlhA have cytoplasmic domains that are known to bind to substrates, so it is possible that the membrane domains of these proteins undergo conformational changes in response to substrate binding and would then cause opening in FliPQR.

5 The structure of the export gate of injectisomes and flagella is highly conserved

5.1 Introduction

As described in chapter 1, the export apparatus proteins of type three secretion are highly conserved. Following the demonstration of the conserved stoichiometry of the export gate, made up of FliPQR/SctRST, in chapter 3 and the determination of the structure of FliPQR from *Salmonella* Typhimurium in chapter 4, it was predicted that homologues of FliPQR from other species and other systems would exhibit a very similar structure. In order to test this, the conservation of the structure was investigated by obtaining the structures of homologues of FliPQR. In addition to studying conservation, it was hoped that the structures of different homologues might reveal the complex in a different state or elucidate the structures of additional components present in the samples.

5.2 Sequence conservation of FliPQR/SctRST

The sequences of FliPQR and SctRST homologues were compared taking into account the structure of *S.* Typhimurium FliPQR (Figures 5.1 to 5.3). The FliR sequence is the most variable, while the middle of FliQ and the C-terminal half of FliP are highly conserved. As the variability of the sequence is high, especially that of FliR/SctT (Table 5.1), the structural conservation of FliPQR and SctRST was studied.

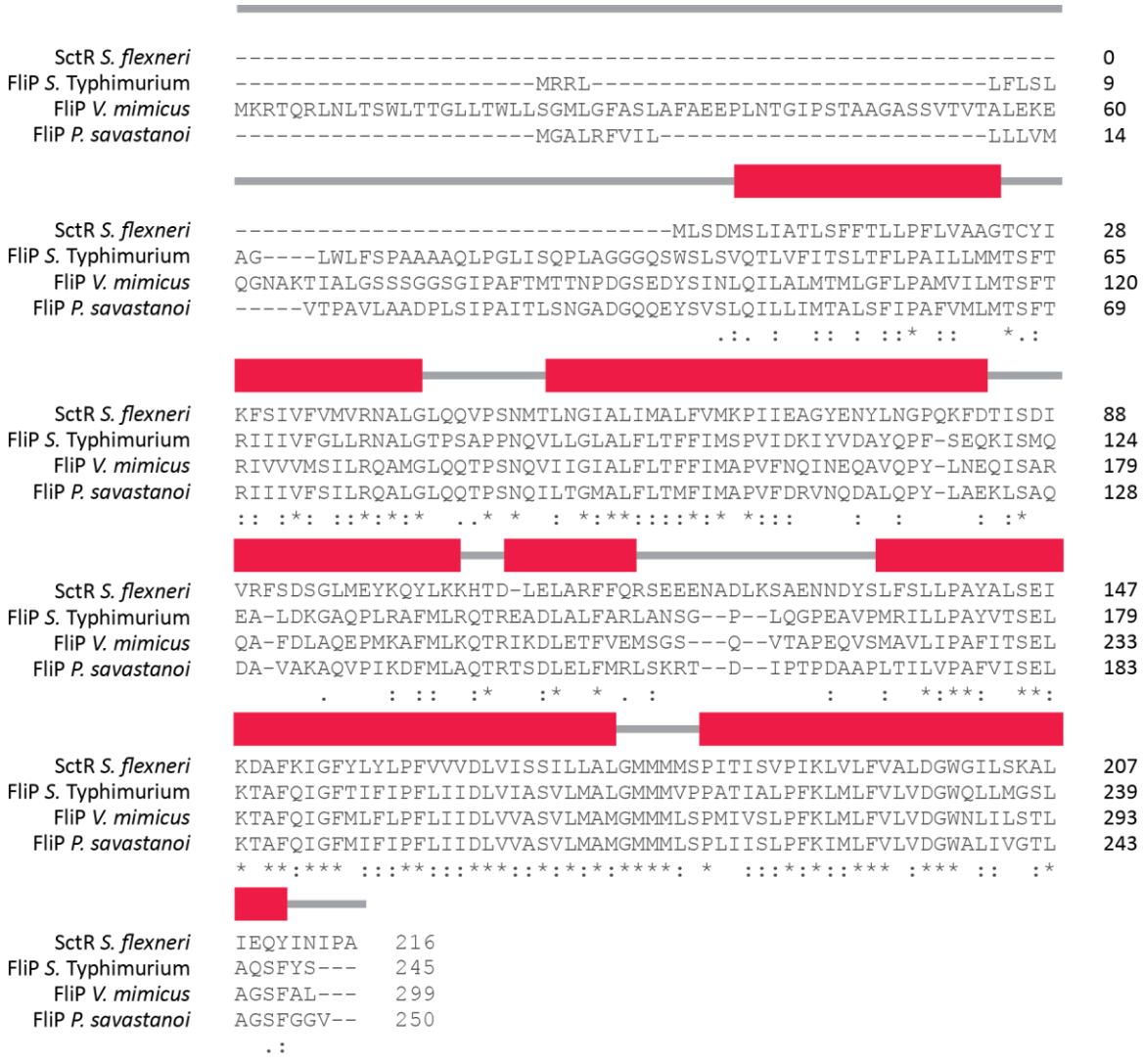


Figure 5.1 Alignment of FliP sequences.
Sequence alignment was calculated using Clustal omega (Sievers et al. 2011). Alpha helical secondary structure is shown in red based on the structure of *S*. Typhimurium FliPQR.

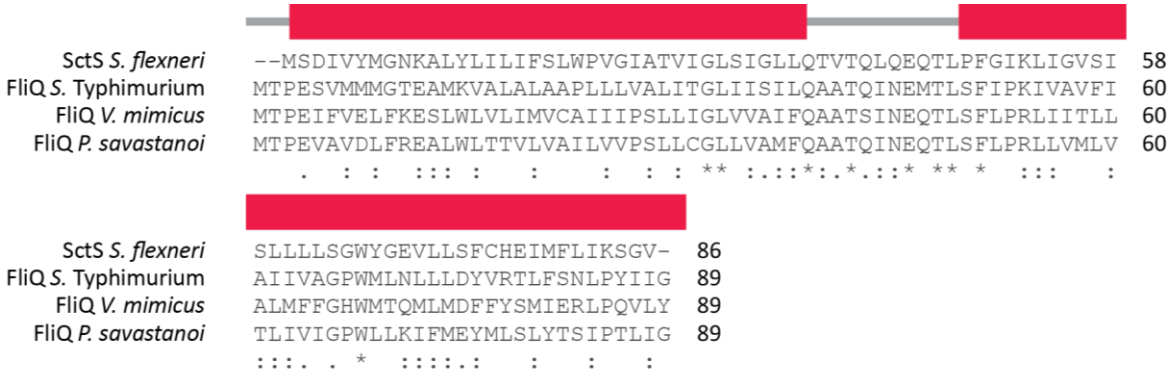


Figure 5.2 Alignment of FliQ sequences.
Sequence alignment was calculated using Clustal omega (Sievers et al. 2011). Alpha helical secondary structure is shown in red based on the structure of *S*. Typhimurium FliPQR.

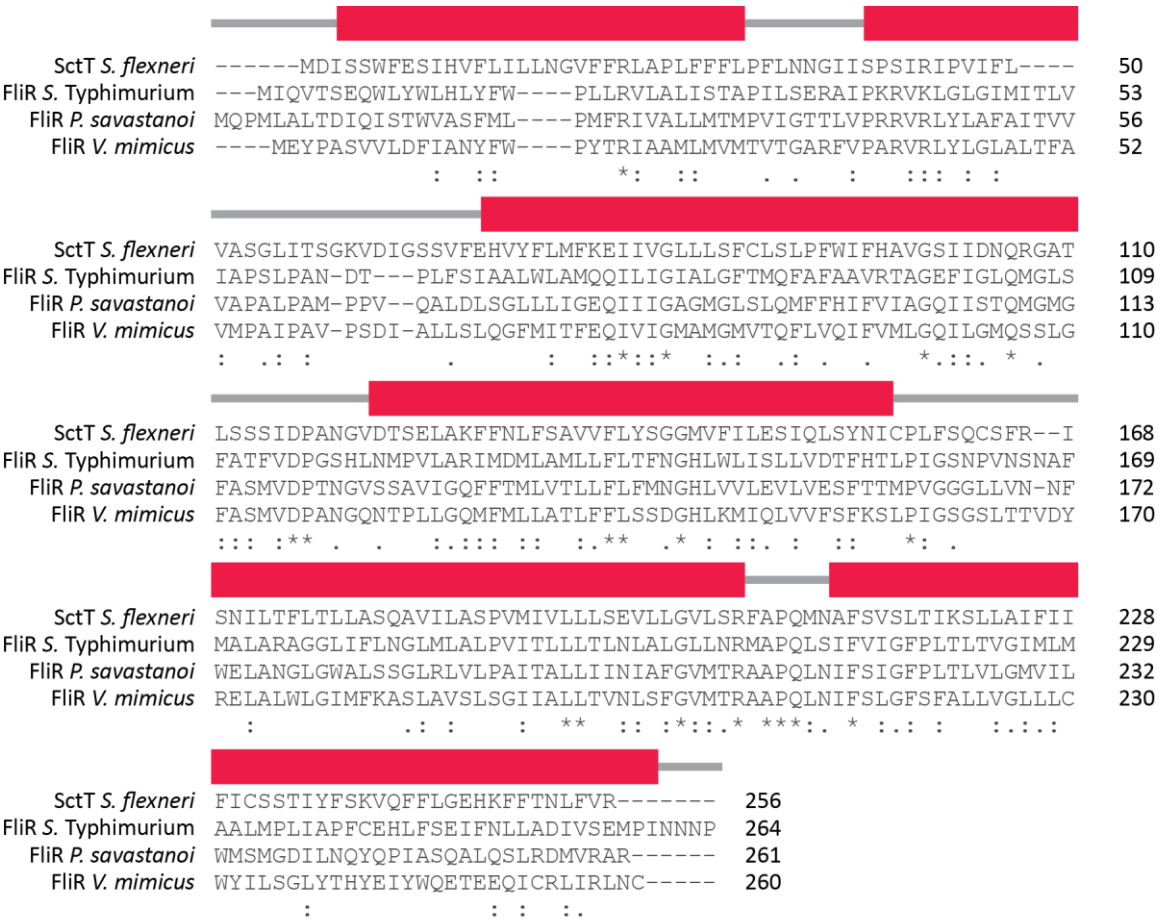


Figure 5.3 Alignment of FliR sequences.
Sequence alignment was calculated using Clustal omega (Sievers et al. 2011). Alpha helical secondary structure is shown in red based on the structure of *S*. Typhimurium FliPQR.

FliP/SctR	*Salmonella*	*Vibrio*	*Pseudomonas*	*Shigella*
Salmonella	100%	49.59%	54.73%	34.78%
Vibrio	49.59%	100%	60.48%	34.47%
Pseudomonas	54.73%	60.48%	100%	32.69%
Shigella	34.78%	34.47%	32.69%	100%
FliQ/SctS	*Salmonella*	*Vibrio*	*Pseudomonas*	*Shigella*
Salmonella	100%	33.71%	40.45%	25.58%
Vibrio	33.71%	100%	49.44%	24.42%
Pseudomonas	40.45%	49.44%	100%	22.09%
Shigella	25.58%	24.42%	22.09%	100%
FliR/SctT	*Salmonella*	*Vibrio*	*Pseudomonas*	*Shigella*
Salmonella	100%	32.95%	37.74%	18.55%
Vibrio	32.95%	100%	41.25%	20%
Pseudomonas	37.74%	41.25%	100%	22.49%
Shigella	18.55%	20%	22.49%	100%

Table 5.1 Sequence identities between FliP, FliQ and FliR of *Salmonella*, *Pseudomonas* and *Vibrio* and SctRST of *Shigella*.
Sequence identities were calculated using Clustal omega (Sievers et al. 2011).

5.3 Cryo-EM imaging of homologues of FliPQR

5.3.1 Structure of *Vibrio* FliPQR

The first candidate sample for the study of the conservation of the export gate structure was the FliPQR homologue from *Vibrio mimicus* ($FliPQR_{Vm}$). In addition to enabling confirmation of the conservation of the structure, the structure of $FliPQR_{Vm}$ was of interest due to the presence of an additional domain in FliP (Figure 5.4). Finally, ongoing imaging efforts focused on the complex of FliPQR and FlhB and the *V. mimicus* complex was most promising (chapter 6). The structure of the FliPQR complex in the absence and presence of FlhB would allow comparison of any changes that FlhB might cause in the complex.

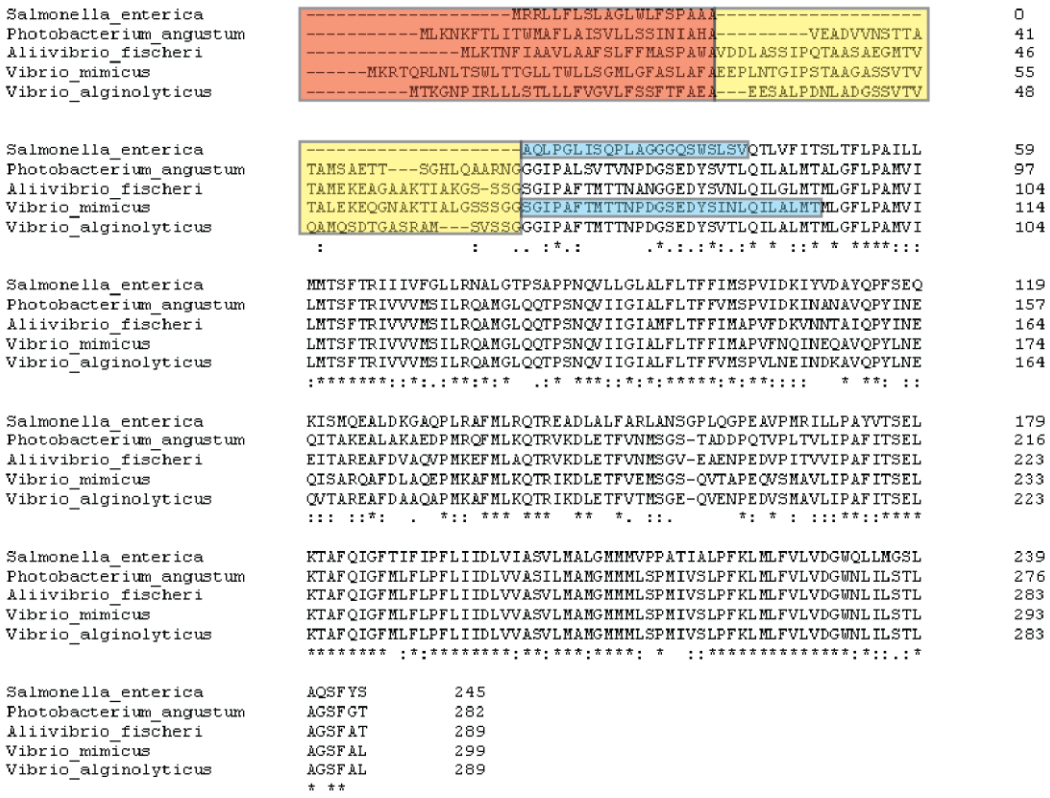


Figure 5.4 Sequence alignment of FliP sequences.
Sequences were downloaded from uniprot and an alignment was calculated using Clustal omega (Sievers et al. 2011). Red shading indicates the predicted signal sequences calculated using SignalP (Almagro Armenteros et al. 2019), yellow the N-terminal domain of the *Vibrionales* order, blue the N-terminal residues of FliP that could not be modelled due to lack of ordered density in the cryo-EM data.

The *V. mimicus* FliPQR complex was prepared in the detergent LMNG as described above (chapter 3 and section 2.5.3). Grids (Au 300 mesh Quantifoil grids) were prepared and imaged in the same way as *Salmonella* FliPQR grids (section 4.2.1). Initial imaging and 2D classification in SIMPLE showed that this sample of $FliPQR_{Vm}$ suffered from strong preferential orientation (Figure 4.2), but swirls could be seen in some classes, suggesting that a structure could be calculated if side views could be obtained.

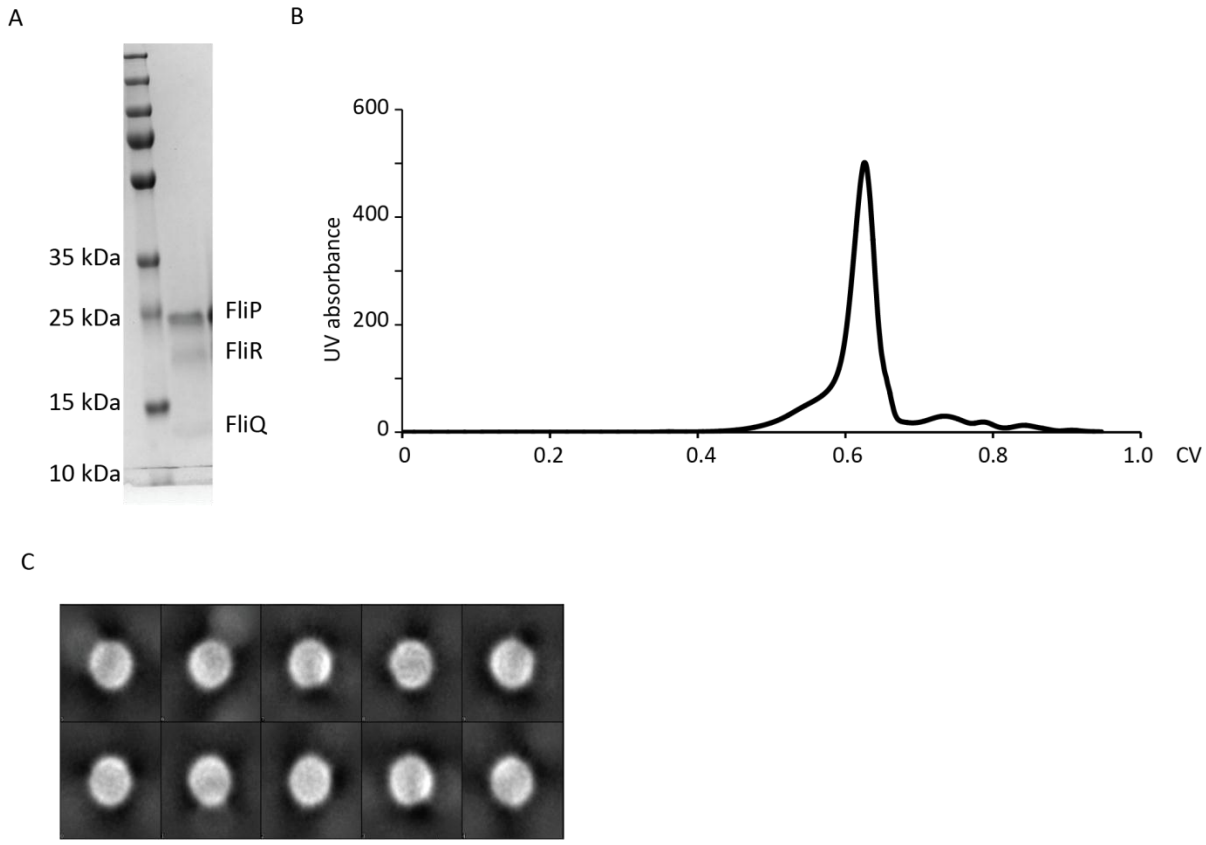


Figure 5.5 Initial cryo-EM characterisation of FliPQR.
A, 4-20% gradient gel of the purified *V. mimicus* FliPQR sample in LMNG. **B**, SEC of the *V. mimicus* FliOPQR purification on a Superose 6 increase 10/300 column. **C**, 10 most populated 2D class averages of *V. mimicus* FliPQR in LMNG calculated using SIMPLE PRIME2D (Reboul et al. 2018).

Anecdotal evidence suggests that adding fluorinated detergents to samples in detergent can result in changed orientations and this sample was used to test the versatility of this strategy. This experiment and the subsequent data processing in RELION was carried out by Dr Justin Deme. In order to achieve a wide range of views of the complex, fluorinated Fos-Choline 8 (f-FC8) (Anatrace) was titrated into the sample, which resulted in a lower particle density, requiring higher protein concentrations for samples with f-FC8. 3 µl of complex purified in LMNG (Figure 4.2) at 1 mg/ml (no f-FC8) or 2.7 mg/ml (supplemented with 50 µM, 500 µM or 3 mM f-FC8) were applied to glow-discharged holey carbon-coated grids (Quantifoil 300 mesh, Au R1.2/1.3) in a Vitrobot Mark IV (FEI). After a wait time of 10 seconds, grids were blotted for 3 s at 100% humidity at 4 °C and frozen in liquid ethane. The grids were imaged using EPU (FEI) on a Titan Krios microscope (FEI) equipped with a K2 detector (Gatan) (Table 5.2).

Data collection and processing	
Voltage	300
Electron exposure	48 $e^{-}/Å^2$
Micrographs	1,387 (0 mM f-FC8), 468 (0.05 mM f-FC8), 1,900 (0.5 mM f-FC8), 3,632 (3 mM f-FC8)
Particles (total)	1,050,955
Particles (final)	243,489
Sampling	0.822 Å
Defocus range	0.5-4 µm
Refinement	
Resolution	4.1 Å
FSC threshold	0.143
Symmetry	C1
Map sharpening B factor	-214 $Å^2$

Table 5.2 Cryo-EM data collection and processing statistics for FliPQR$_{Vm}$.

Pre-processing of the data was carried out in SIMPLE (Reboul et al. 2018) using SIMPLE implementations of Unblur and CTFFIND4 for motion correction and CTF estimation respectively. Particles were picked using SIMPLE and extracted in a 256x256 pixel box. The data were analysed in RELION-2.0 (Kimanius et al. 2016) (Figure 5.6A), revealing an increasing proportion of side and tilted views as the concentration of f-FC8 was increased (Figure 5.6B), confirming that using fluorinated detergents is a valid strategy for obtaining a broader distribution of views of membrane proteins solubilised in LMNG.

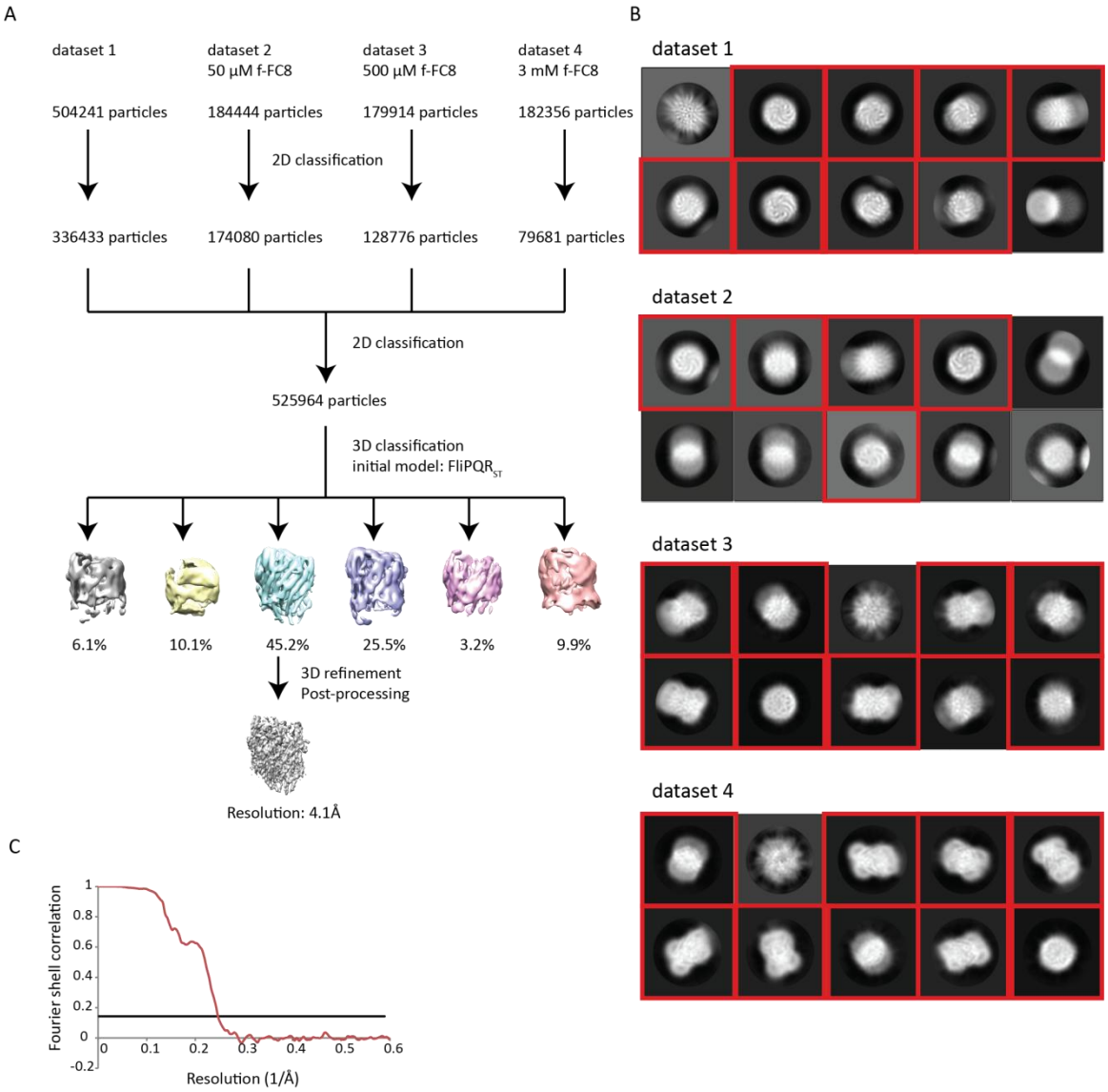


Figure 5.6 Cryo-EM data processing of FliPQR$_{Vm}$.
A, Flowchart describing the data processing strategy. **B**, 10 most populated 2D class averages of the four datasets. Red boxes highlight the selected class averages. **C**, Solvent-corrected FSC curve highlighting the 0.143 cut-off.

Class averages with high-resolution features were selected and noisy averages were removed. The selected particles resulting from the different concentrations of fluorinated detergent were merged and subjected to another round of 2D classification. The most detailed and least noisy averages were selected and the particles were subjected to 3D classification using the map of *S.* Typhimurium FliPQR (FliPQR$_{ST}$) (chapter 4) low-pass filtered to 60 Å as an initial model. The class with the best connectivity and accuracy of angles and rotations estimated by RELION was selected and

auto-refinement and post-processing of the corresponding particles resulted in a 4.1 Å map (Figure 5.6C).

An atomic model was built in Coot (Emsley et al. 2010). The structure of the FliPQR-FlhB complex from *V. mimicus* (chapter 6) was fit into the density and the coordinates for FlhB were deleted. The model was improved using the real space refine function in Coot and then refined using the program phenix.real_space_refine (Afonine et al. 2018) (Table 5.3).

CC_{Mask}	0.8448
B factors	102 $Å^2$
R.m.s. deviations	
Bond lengths	0.007 Å
Bond angles	0.87°
Validation	
MolProbity score	2.25
Clashscore	18.28
Poor rotamers	0.07%
Ramachandran plot	
Favoured	91.9%
Allowed	7.8%
Disallowed	0.39%

Table 5.3 Model refinement statistics for $FliPQR_{Vm}$.

The final model of the $FliP_5Q_4R_1$ complex was observed to possess the same architecture as the previously described $FliPQR_{ST}$ (Figure 5.7A). The loop of $FliQ_4$ could be observed but it was not possible to model residues 40 to 48 (Figure 5.7B). No density was observed for the N-terminal domain of FliP. The structure was highly similar to $FliPQR_{ST}$ with an RMSD of 1.57 Å over 1,483 residues. As in $FliPQR_{ST}$, a central $FliP_5R_1$ complex is surrounded by FliQ subunits (Figure 5.7). The complex is closed.

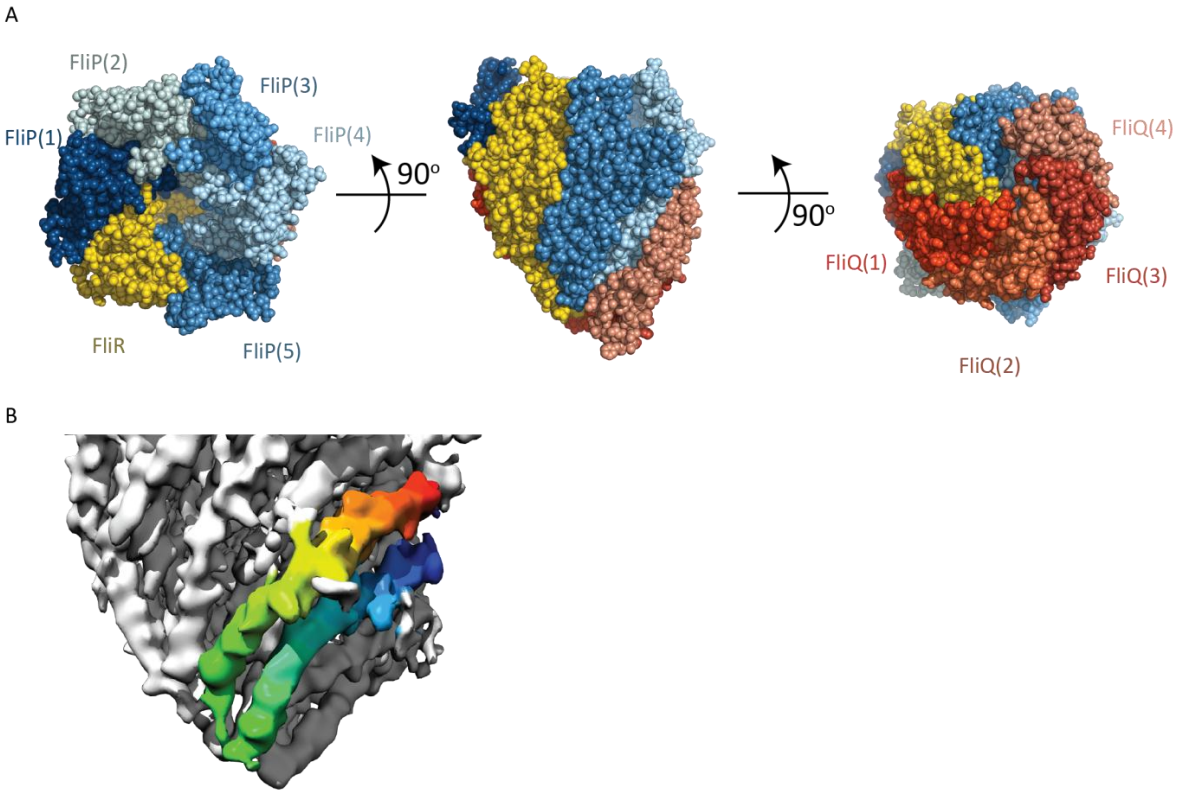


Figure 5.7 Structure of FliPQR$_{Vm}$.
A, Structure of FliPQR from *V. mimicus*. **B**, Cryo-EM map highlighting FliQ$_4$ in rainbow colouring.

5.3.2 Structure of *Pseudomonas* FliPQR

As described in chapter 3, *P. savastanoi* FliPQR overexpressed as a FliOPQR operon in *E. coli* and purified in the detergent LMNG was a mixture of FliP$_5$Q$_5$R$_1$ and FliP$_5$Q$_4$R$_1$ complexes. As only four FliQ subunits per complex were observed in the *S.* Typhimurium and *V. mimicus* structures (Figures 4.7A and 5.4A), it was possible that a potential fifth FliQ dissociated due to extraction in detergent and that the larger *P. savastanoi* FliP$_5$Q$_5$R$_1$ complex represented the physiologically relevant, full export gate complex. In order to gain further insight into how FliQ$_5$ binds the *P. savastanoi* FliPQR complex (FliPQR$_{Ps}$) its cryo-EM structure was determined.

The FliPQR$_{Ps}$ complex was purified in DDM and LMNG as described above (section 2.5.3). Interestingly, co-purification of FliQ was similar in the two detergents (Figure 5.8A), suggesting that FliQ$_{Ps}$ binds the complex more tightly than FliQ$_{ST}$, which dissociated easily in DDM (chapter 3). Upon SEC (Figure 5.8B), a larger proportion of the sample was

observed to elute in the void peak (~0.35 CV), which also contains the FliOPQR assembly intermediates (Figure 3.3), in DDM compared to LMNG, which suggests that LMNG is the more appropriate detergent to use also with this sample.

Initial screening of cryo-EM grids in DDM and LMNG indicated that the LMNG sample produced higher quality micrographs and that the particles adsorbed to the carbon surface outside the holes, requiring higher protein concentrations to saturate carbon binding. Therefore, the purified complex in LMNG was concentrated to 4.1 mg/ml and 3 µl were applied to a 300 mesh Au R1.2/1.3 grid (Quantifoil). After a 10 second wait time the grid was blotted and frozen using a Vitrobot Mark IV (FEI) at 100% humidity at 22 °C (Table 2.10). Data were collected using a Titan Krios microscope (FEI) equipped with a K2 detector (Gatan) using EPU software (FEI) for automated data collection (Table 5.4). Micrographs were pre-processed using the SIMPLE pipeline (Reboul et al. 2018), including implementations of Unblur for motion correction and CTFFIND4 for CTF estimation. A total of more than 500,000 particles were picked using SIMPLE, extracted in a 256x256 box and processed in RELION-2.0 (Kimanius et al. 2016) (Figure 5.9A). A mixture of top and side views was observed in 2D class averages (Figure 5.9B). Particles belonging to class averages with detailed features and little noise were selected and subjected to 3D classification using the structure of $FliPQR_{ST}$ low-pass filtered to 30 Å as an initial model. The class with the best connectivity in the density and the best estimated accuracy of rotations and translations in RELION was selected and a 4.1 Å map (Figure 5.9C) could be calculated when these particles were subjected to auto-refinement and post-processing. When RELION-3.0 was released, motion correction was repeated using MotionCor2 as implemented in RELION-3.0 (Zivanov et al. 2019) and the previously identified set of particles was re-extracted in a 288x288 pixel box and refined again. After CTF refinement, particle polishing and post-processing in RELION-3.0 the resolution of the final map was 3.5 Å.

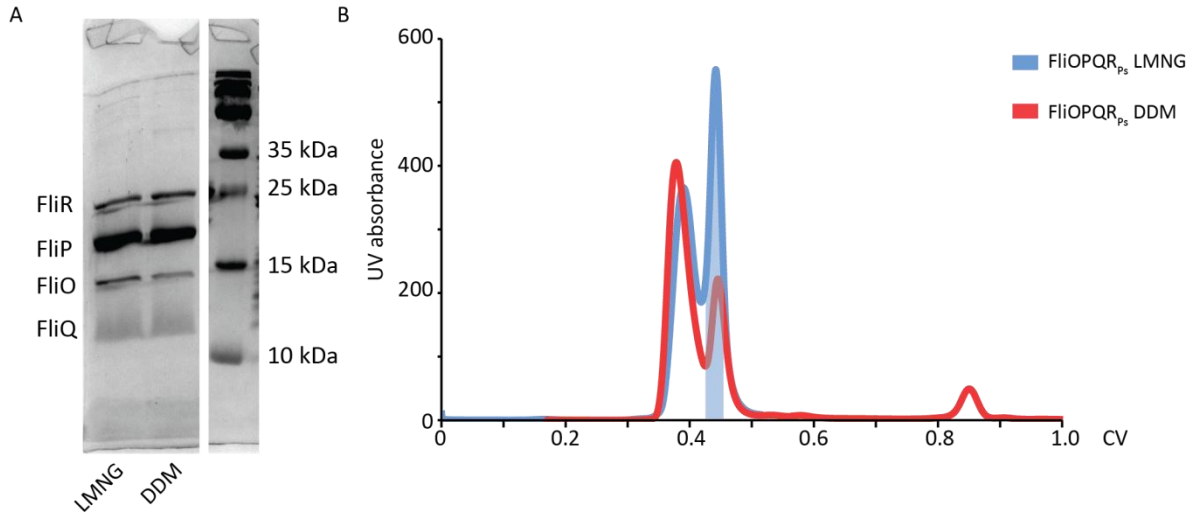


Figure 5.8 Purification of FliOPQR$_{Ps}$.
A, 15% SDS-PAGE analysis of FliOPQR strep eluates purified using the detergent LMNG or DDM. **B**, SEC profile on an Superdex 200 increase 10/300 column (GE) of FliOPQR$_{Ps}$ using TBS buffer containing 0.01% (w/v) LMNG or 0.02% (w/v) DDM. The shaded area indicates the fractions that were pooled for further analysis.

Data collection and processing	
Voltage	300
Electron exposure	48 e^{-}/Å^2
Micrographs	2,476
Particles (total)	503,177
Particles (final)	97,987
Sampling	0.822 Å
Defocus range	0.5-4 µm
Refinement	
Resolution	3.5 Å
FSC threshold	0.143
Symmetry	C1
Map sharpening B factor	-101 Å^2

Table 5.4 Cryo-EM data collection and processing statistics for FliPQR$_{Ps}$.

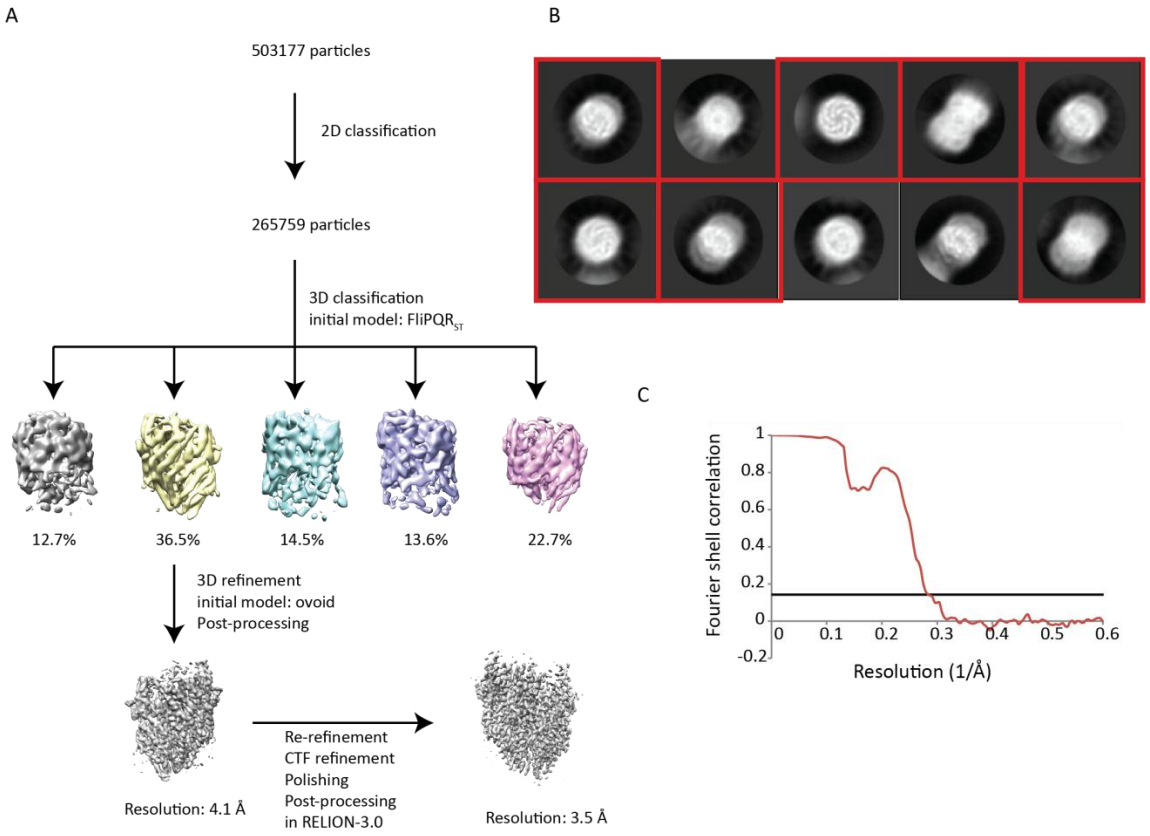


Figure 5.9 Cryo-EM data processing of FliPQR$_{Ps}$.
A, Flowchart describing the data processing strategy. **B**, 10 most populated 2D class averages. Red boxes highlight the selected class averages. **C**, Solvent-corrected FSC curve highlighting the 0.143 cut-off.

An atomic model of the complex was built in Coot (Emsley et al. 2010) by fitting the FliPQR$_{ST}$ coordinates into the density and manually mutating the sequence to the *P. savastanoi* sequence. A FliQ model for the fifth FliQ subunit was copied and moved into the appropriate density. The model was improved using the real space refine function in Coot and the geometry was refined using iterations of the program phenix.real_space_refine (Afonine et al. 2018) and manual editing in Coot (Table 5.5).

The structure of FliPQR$_{Ps}$ revealed a highly conserved architecture of the complex (RMSD=1.61 Å over 1,494 residues compared to FliPQR$_{ST}$ and an RMSD of 1.30 Å over 1,459 residues compared to FliPQR$_{Vm}$). Density for a fifth FliQ subunit was observed, but the density of FliQ$_5$ was weaker than that of the rest of the complex, so that density corresponding to FliQ$_5$ is highly disconnected at a high contour level (Figure 5.10B), while the densities of FliQ$_1$ to FliQ$_4$ are all of similar strength. This suggests

substoichiometric occupancy of $FliQ_5$, consistent with the mass spectrometry results showing a mixture of $FliP_5Q_4R_1$ and $FliP_5Q_5R_1$ (Figure 3.8).

CC_{Mask}	0.7676
B factors	37 $Å^2$
R.m.s. deviations	
Bond lengths	0.006 Å
Bond angles	0.88°
Validation	
MolProbity score	2.54
Clashscore	18
Poor rotamers	2.22%
Ramachandran plot	
Favoured	90.8%
Allowed	8.4%
Disallowed	0.8%

Table 5.5 Model refinement statistics for $FliPQR_{Ps}$.

$FliQ_5$ binds $FliP_5$ in the same fashion as the other FliP-FliQ pairs with additional lateral contacts to $FliQ_4$, including a salt bridge between $FliQ_4$ Glu46 and $FliQ_5$ Arg54 (Figure 5.10C). However, the conservation of basic character in residue 206 (*S.* Typhimurium numbering, Arg208 in *P. savastanoi*) in the C-terminal hairpin of FliR, which forms a salt bridge with $FliQ_4$ that is incompatible with a $FliQ_4$-$FliQ_5$ salt bridge suggests that there is no fifth FliQ subunit (Figure 4.13B). In $FliPQR_{Ps}$ FliR Arg208 is blocked from contacting $FliQ_4$ Glu46 by the presence of $FliQ_5$ Arg54 and FliR Arg208 points at the area where the loop of $FliQ_5$ would be expected. These residues of $FliQ_5$ (36-51) could not be modelled as there was no cryo-EM density. Given these observations, it is more likely that the native stoichiometry of FliQ is four rather than five, which would explain why basic character in FliR residue Arg208 is conserved. $FliQ_5$ may only bind to the complex in the absence of FlhB, which is predicted to bind a similar site on the FliPQR complex (Figure 4.17).

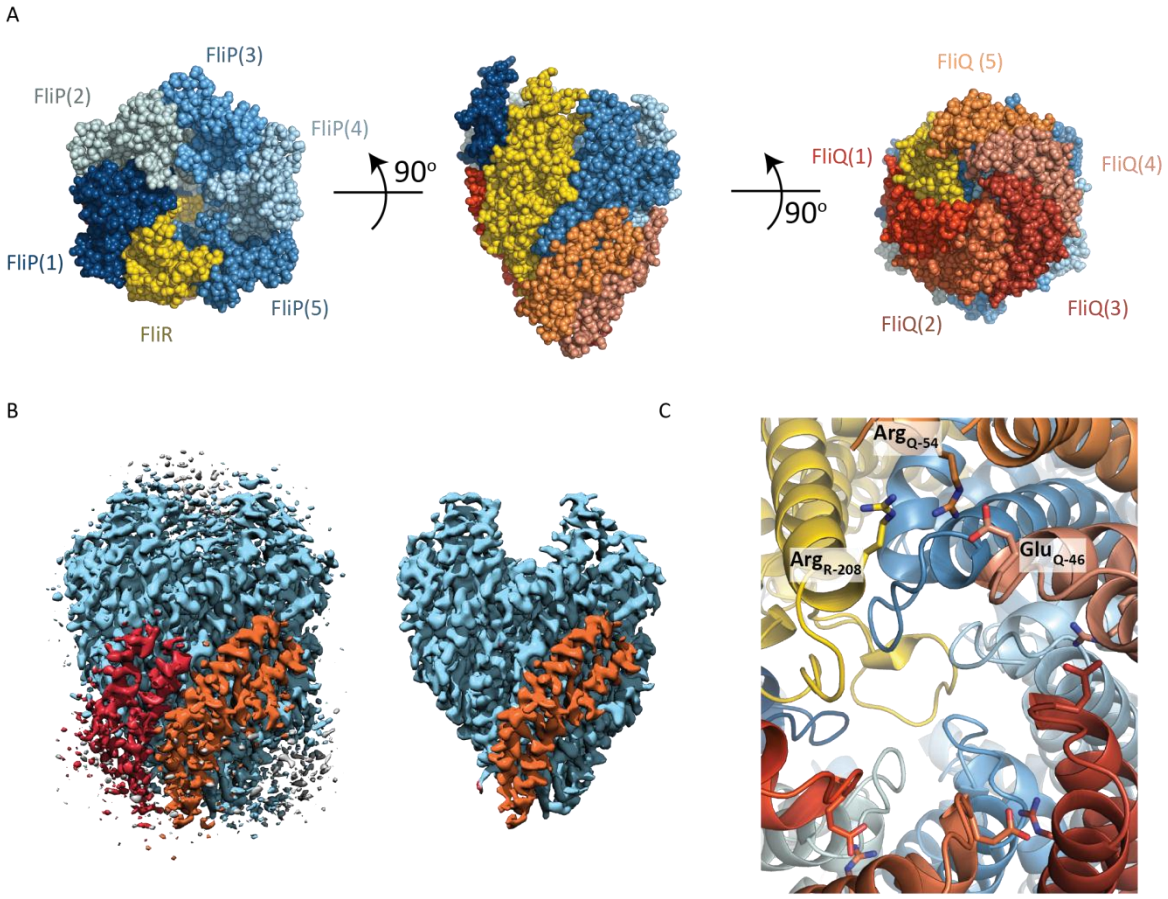


Figure 5.10 Structure of FliPQR$_{Ps}$.
A, Structure of FliPQR from *P. savastanoi* revealing a fifth FliQ subunit. **B**, The final cryo-EM volume highlighting FliQ$_4$ in orange and FliQ$_5$ in red. Left: unedited map, right: map displayed using the Chimera (Pettersen et al. 2004) feature "Hide Dust". Hide Dust hides disconnected, small parts of a volume. **C**, A salt bridge between FliQ$_4$ and FliQ$_5$ blocks the formation of a salt bridge between FliR and FliQ$_4$. The protein subunits are coloured as in (**A**).

5.4 Conservation of the structure

Despite the variability of some parts of the sequence (Figures 5.1 to 5.3), the structures of the FliPQR complexes are highly similar. In addition, each subunit is also highly conserved (Figure 5.11).

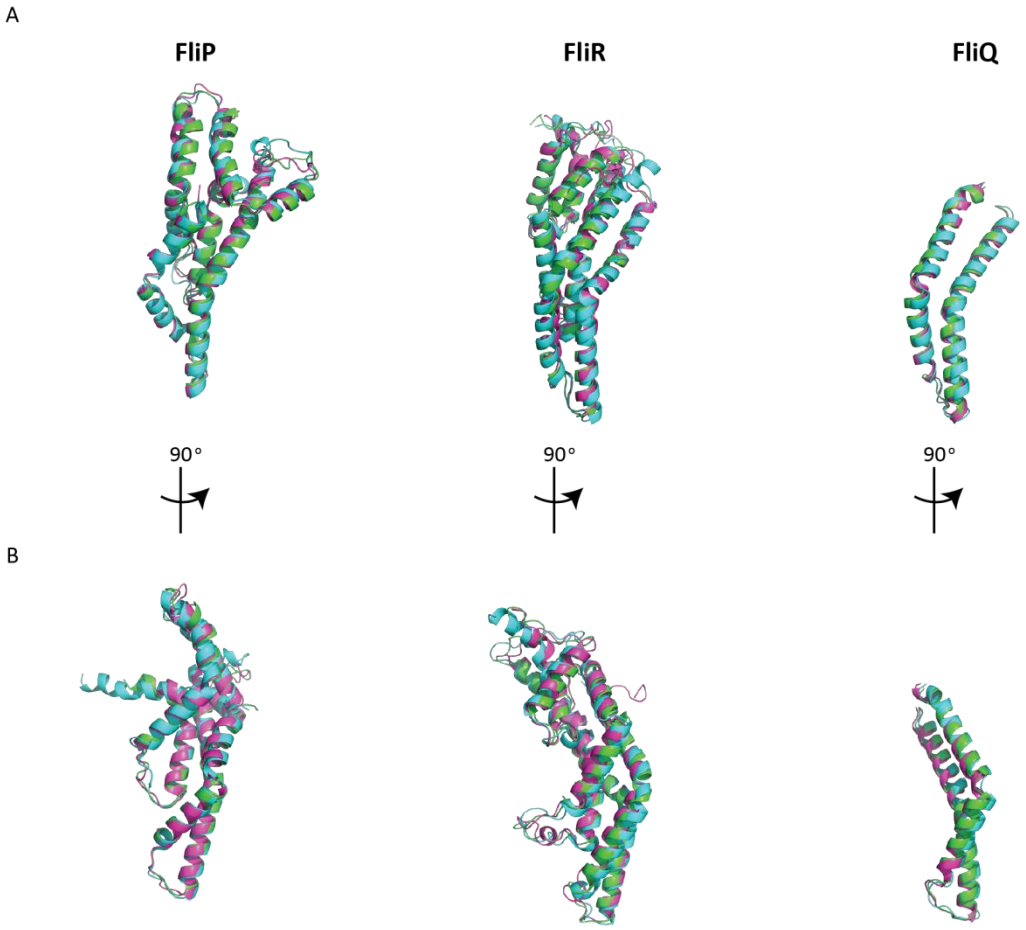


Figure 5.11 Conservation of the subunit structures.
A, Overlay of $FliP_1$, FliR and $FliQ_1$ seen from the outside the complex. Green: $FliPQR_{Ps}$, Blue: $FliPQR_{ST}$, Magenta: $FliPQR_{Vm}$. **B**, Side views of the same subunits.

When the sequence conservation was mapped onto the structure using the ConSurf server (Ashkenazy et al. 2016) most of the conserved residues were seen to localise to the bottom part of the complex (Figure 5.12), which is embedded in the inner membrane prior to assembly of the basal body. The periplasm-exposed surface is not conserved. In addition, the periplasmic residues predicted to be in contact with the rod proteins are also conserved (Figure 5.12C). The hydrophobic parts of the structures are likely to be important in assembly of the complex and may play a part in opening of the complex, explaining the conservation in this region. A cluster of highly conserved residues is created by the loops of the FliQ subunits at the entrance of the complex facing the cytoplasm.

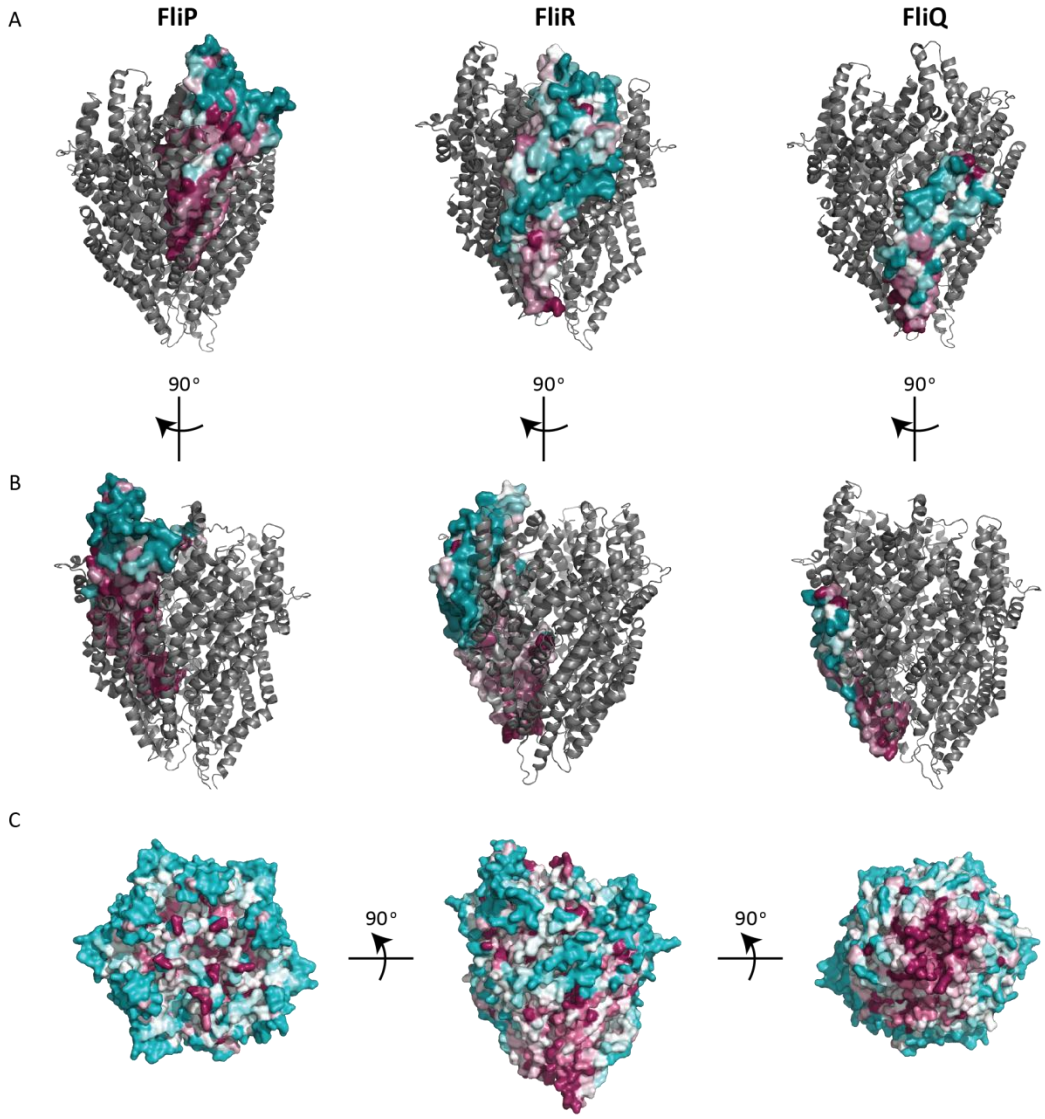


Figure 5.12 Consurf analysis the FliPQR complex.
A, B, ConSurf (Ashkenazy et al. 2016) analysis of each subunit of the complex. In all three cases the parts of the structures that are periplasmic prior to basal body assembly are less conserved while the hydrophobic parts of the structures are highly conserved. **C**, Consurf analysis of the entire complex showing the highly conserved bottom of the complex that is inserted in the membrane before assembly into the basal body. Isolated residues at the periplasmic face of the complex that interacts with the rod proteins are also highly conserved. Blue colour indicates low conservation, magenta indicates high conservation.

5.5 Structure of *Shigella* SctRST

In order to confirm structural conservation of the export gate complex in injectisomes, the vT3SS homologue of FliPQR, SctRST, was the final structural target. However, obtaining a sample for cryo-EM proved more difficult than for FliPQR. Of the SctRST homologues that were cloned, SctRST from *Y. enterocolitica* and *S.* Typhimurium SPI-1

could only be prepared at very low yields in LMNG (<10 µg and around 15 µg respectively per litre of culture using the detergent LMNG and protocol 2.5.3). It was not clear from SDS-PAGE analysis whether purification had been successful (Figure 5.13A). SctRST from *S.* Typhimurium SPI-2, *S. flexneri* and from *P. savastanoi* (yield around 100 µg, 200 µg and 30 µg per litre of culture respectively) could be produced in larger amounts, suitable for structural studies. Bands corresponding to all three components of the complex could be seen by SDS-PAGE (Figure 5.13B).

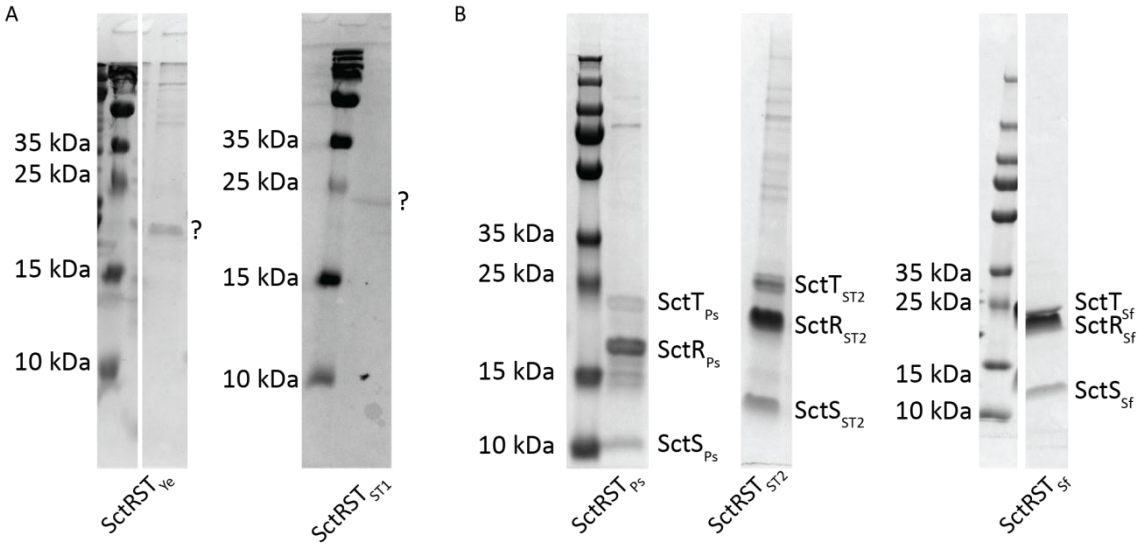


Figure 5.13 Purification of SctRST.
A, SDS-PAGE analysis of the strep eluates of SctRST from *Y. enterocolitica* ($SctRST_{Ye}$) and *S*. Typhimurium SPI-1 ($SctRST_{ST1}$). **B**, SDS-PAGE analysis of the strep eluates of SctRST from *P. savastanoi* ($SctRST_{Ps}$), *S.* Typhimurium SPI-2 ($SctRST_{ST2}$) and *S. flexneri* ($SctRST_{Sf}$). $SctRST_{Ps}$ and $SctRST_{ST2}$ were run on the same gel. The $SctRST_{Sf}$ gel is the same as in Figure 3.8.

When SctRST from *S.* Typhimurium SPI-2 or *P. savastanoi* was analysed by cryo-EM, no intact particles were observed, suggesting that these homologues are highly sensitive to the air-water interface (D'Imprima et al. 2019). SctRST from *S. flexneri* ($SctRST_{Sf}$) could be used to produce cryo-EM grids of sufficient quality for imaging, but no high-quality data could be obtained due to a large proportion of broken particles. Grids were prepared with $SctRST_{Sf}$ concentrated to 3.4 mg/ml, 5.7 mg/ml, 16.8 mg/ml and 8.4 mg/m (diluted 16.8 mg/ml sample) using Quantifoil 1.2/1.3 gold holey carbon grids, 300 mesh, and a Vitrobot Mark IV (FEI) at 100% humidity at room temperature. Good quality micrographs

were obtained from the sample at 8.4 mg/ml (Figure 5.14A) and this grid was imaged on a Titan Krios microscope (FEI) equipped with a K2 camera (Gatan) using EPU (FEI) for automated data acquisition (Table 5.6).

Motion correction, CTF estimation and particle picking were carried out using the SIMPLE pipeline (Reboul et al. 2018). Particle classification and refinement (Figure 5.14B,C) was performed in RELION-2.0 (Kimanius et al. 2016), revealing a mixture of top and side views. The data were processed in RELION by Dr Steven Johnson. The 2D class averages with the most detail and the least noise were selected and the corresponding particles subjected to 3D classification using the previous map of $FliPQR_{ST}$ as an initial model. The most populated class corresponded to a model with secondary structure detail and a similar overall shape compared to FliPQR and this class was selected for auto-refinement and post-processing in RELION, resulting in a 4.2 Å map.

Data collection and processing	
Voltage	300
Electron exposure	48 $e^{-}/Å^2$
Micrographs	3,617
Particles (total)	775,073
Particles (final)	212,561
Sampling	0.822 Å
Defocus range	0.5-4 µm
Refinement	
Resolution	3.5 Å
FSC threshold	0.143
Symmetry	C1
Map sharpening B factor	-111 $Å^2$

Table 5.6 Cryo-EM data collection and processing statistics for $SctRST_{Sf}$.

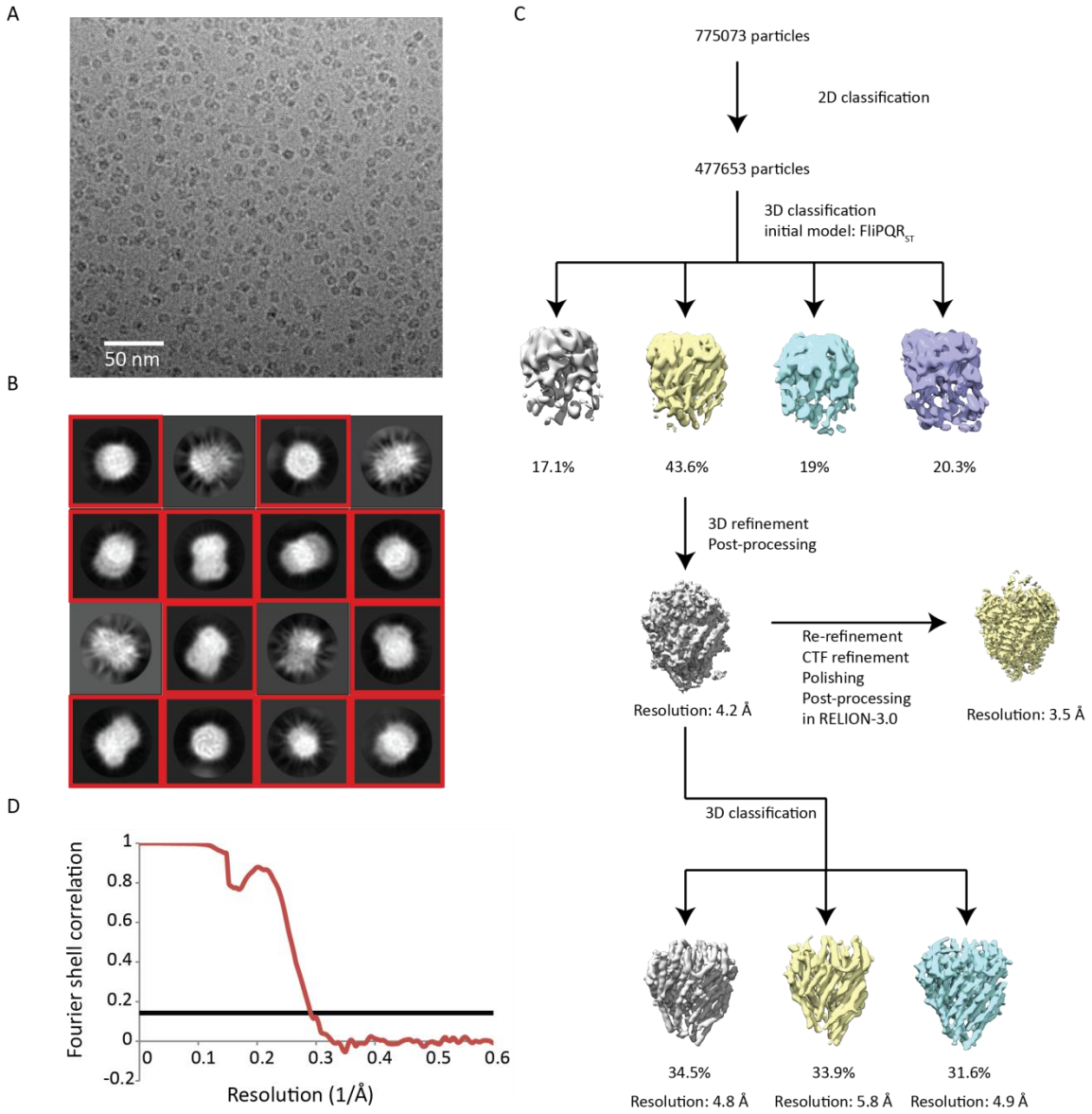


Figure 5.14 Cryo-EM data processing of SctRST$_{Sf}$.
A, Representative micrograph. Defocus=-3.9µm. **B**, 16 most populated 2D class averages. Red boxes highlight the selected class averages. **C**, Particle processing strategy. **D**, Solvent-corrected FSC curve of the final model highlighting the 0.143 cut-off.

Following the release of RELION-3.0 (Zivanov et al. 2018; Zivanov et al. 2019), the micrographs were processed again using MotionCor2 implemented in RELION-3.0 and particles were re-extracted in a 288x288 pixel box. After particle polishing, CTF refinement, auto-refinement and post-processing in RELION-3.0 the final resolution was 3.5 Å (Figure 5.14D and Table 5.6). An atomic model was built automatically using buccaneer (Cowtan 2006) and modified manually using Coot (Emsley et al. 2010). The model was refined using the program phenix.real_space_refine (Afonine et al. 2018) (Table 5.7).

The overall structure is very similar to the structure of $FliPQR_{ST}$ (Figure 5.15). Like FliPQR, $SctRST_{Sf}$ forms a single turn of a right-handed helix and is made up of a central SctRT complex and four peripheral SctS subunits. The outermost hairpin of SctT resembles SctS and binds to $SctS_1$ and the two inner hairpins resemble SctR. The overall RMSD of the $FliPQR_{ST}$ and $SctRST_{Sf}$ is 2.25 Å over 1,281 residues. The individual subunits are also highly conserved (Figure 5.15B), with an RMSD of 1.83 Å over 148 residues for $FliP_1$ and $SctR_1$, 2.42 Å over 230 residues for FliR and SctT and 1.16 Å over 79 residues for $FliQ_1$ and $SctS_1$.

CC_{Mask}	0.8014
B factors	50 $Å^2$
R.m.s. deviations	
Bond lengths	0.007 Å
Bond angles	1.26°
Validation	
MolProbity score	1.9
Clashscore	4.6
Poor rotamers	1.4%
Ramachandran plot	
Favoured	91.5%
Allowed	8.3%
Disallowed	0.2%

Table 5.7 Model refinement statistics for $SctRST_{Sf}$.

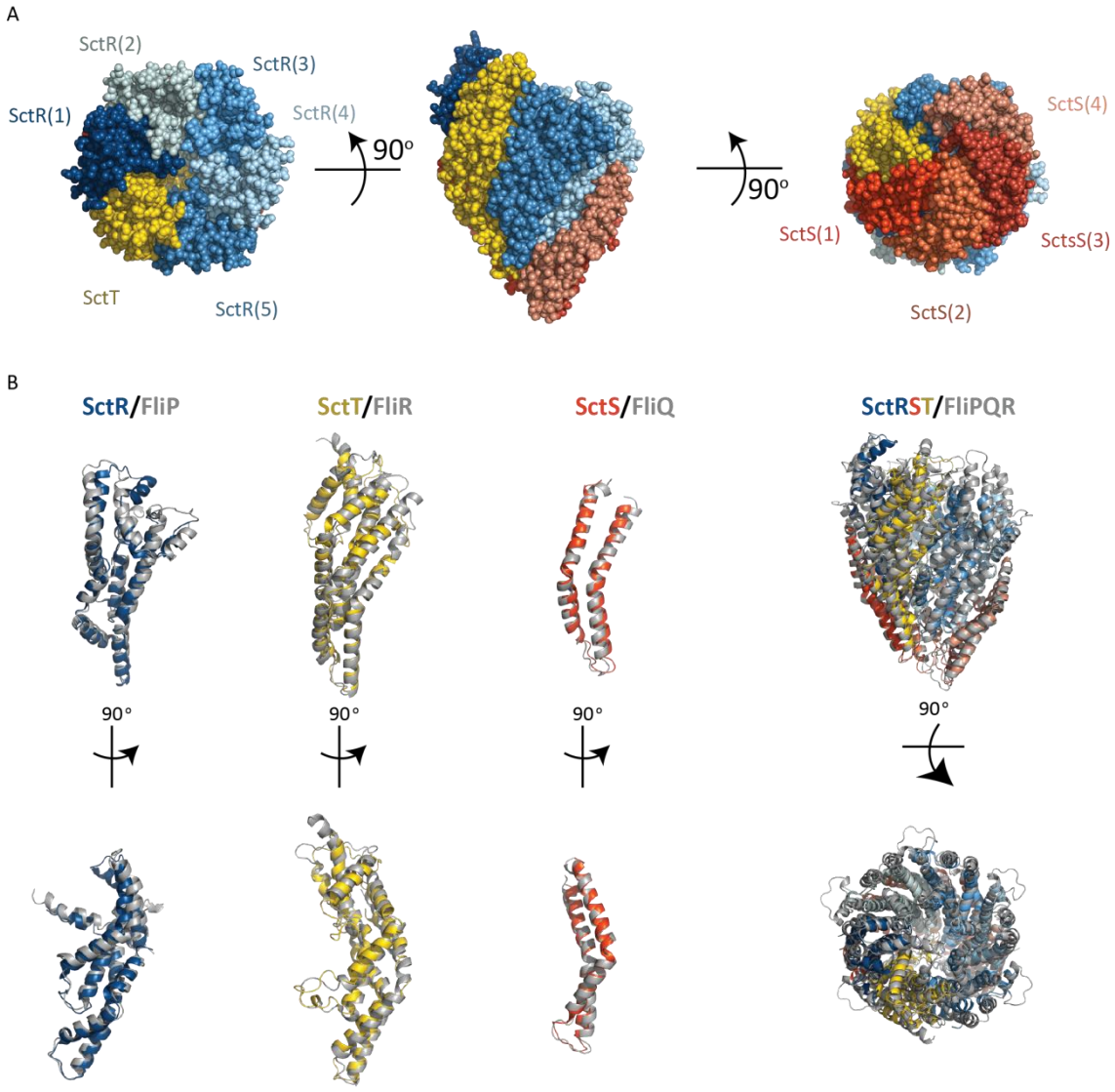


Figure 5.15 Structure of SctRST$_{Sf}$.
A, Structure of the SctRST$_{Sf}$ complex. **B**, Comparison of the structures of SctR (left), SctT (middle) and SctS (right) with their homologues from FliPQR$_{ST}$ and comparison of the structures of the assembled complexes (right).

A difference between the FliPQR complexes and SctRST can be seen in the most periplasmic helices of FliP/SctR (Figure 5.16). In the SctRST complex these are bent towards the centre of the complex and the density was not clear enough in most SctR subunits to model them completely. This may be related to the difference between the rod protein that assembles on top of the export gate, SctI or FliE (Figure 4.16).

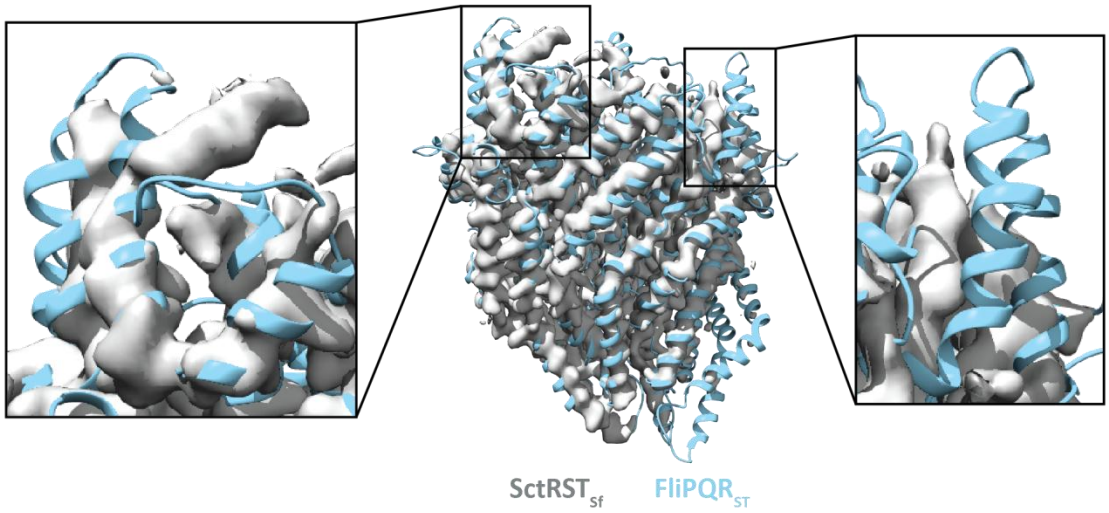


Figure 5.16 Comparison of the periplasmic domains of FliP and SctR.
Structure of FliPQR$_{ST}$ in the cryo-EM density of SctRST. Boxes highlight the difference in the periplasm between FliP$_1$ and SctR$_1$ (left) and FliP$_5$ and SctR$_5$ (right).

As the density corresponding to SctS$_4$ was weak, suggesting substoichiometric occupancy, further classification of the final set of particles (Figure 5.14C) was carried out. This revealed three different classes. These classes corresponded to different stoichiometries of the SctRST$_{Sf}$ complex with approximately one third of particles each classified into a SctR$_5$S$_2$T$_1$, SctR$_5$S$_3$T$_1$ and SctR$_5$S$_4$T$_1$ class (Figure 5.17). Interestingly, the "missing" SctS subunits in the SctR$_5$S$_2$T$_1$ and SctR$_5$S$_3$T$_1$ classes were removed from the outside the most cytoplasmic FliP, i.e. SctR$_5$S$_3$T$_1$ was missing SctS$_4$ and SctR$_5$S$_2$T$_1$ was missing SctS$_3$ and SctS$_4$, but there was no class in which an SctS subunit did not contact another SctS. This implies that SctS$_1$/FliQ$_1$, which is in contact with SctT/FliR, has the highest affinity for the complex and affinity decreases progressively for SctS$_2$/FliQ$_2$, SctS$_3$/FliQ$_3$ and SctS$_4$/FliQ$_4$. Consistent with this hypothesis, the single FliQ subunit that remained associated with FliPQR$_{ST}$ when the complex was extracted in DDM was most likely FliQ$_1$, as it was lost when FliR was truncated (chapter 3). Although it is possible that the different SctS/FliQ stoichiometries represent assembly intermediates, the sensitivity of this subunit to dissociation in detergents (chapter 3) suggests that they are products of dissociation.

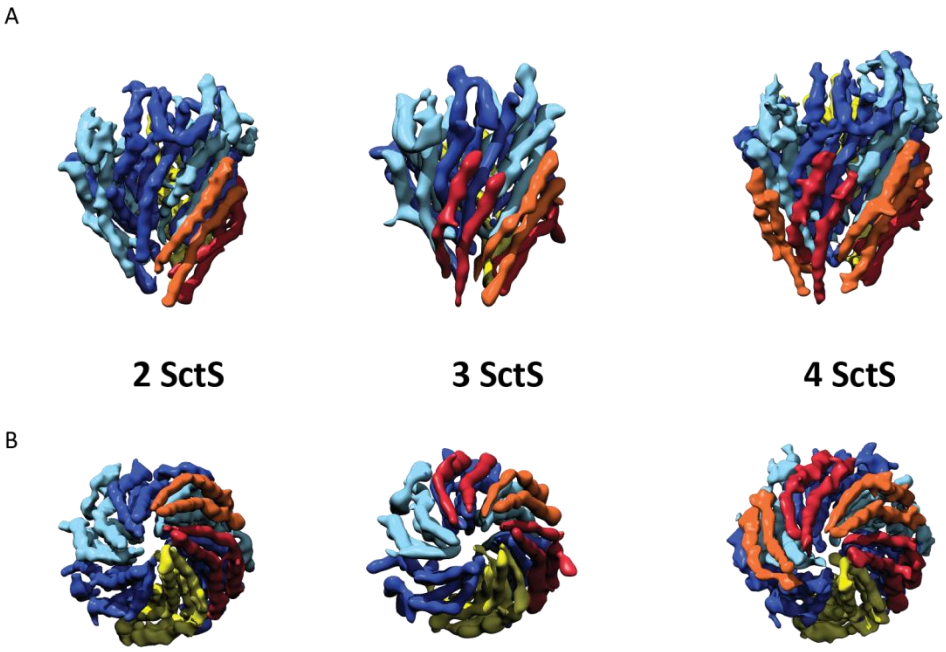


Figure 5.17 Classification of the SctRST$_{Sf}$ particles revealed states with different SctS occupancy.
A, Side view of the three different stoichiometries of the SctRST$_{Sf}$ complex. SctR is coloured blue/light blue, SctS is coloured red/orange and SctT is coloured yellow. **B**, View from the cytoplasm.

5.6 Assembly in the membrane

The observation of the progressive loss of SctS following the helical assembly suggested a possible assembly mechanism in which SctS/FliQ binds to an assembled SctRT/FliPR complex progressively, starting with SctS$_1$/FliQ$_1$. All intermediates from SctR$_5$T$_1$ to SctR$_5$S$_4$T$_1$ have been observed by nMS (chapter 3) and the classification of the SctRST$_{Sf}$ dataset (Figure 5.17) revealed the structures of the intermediates with two and three SctS subunits.

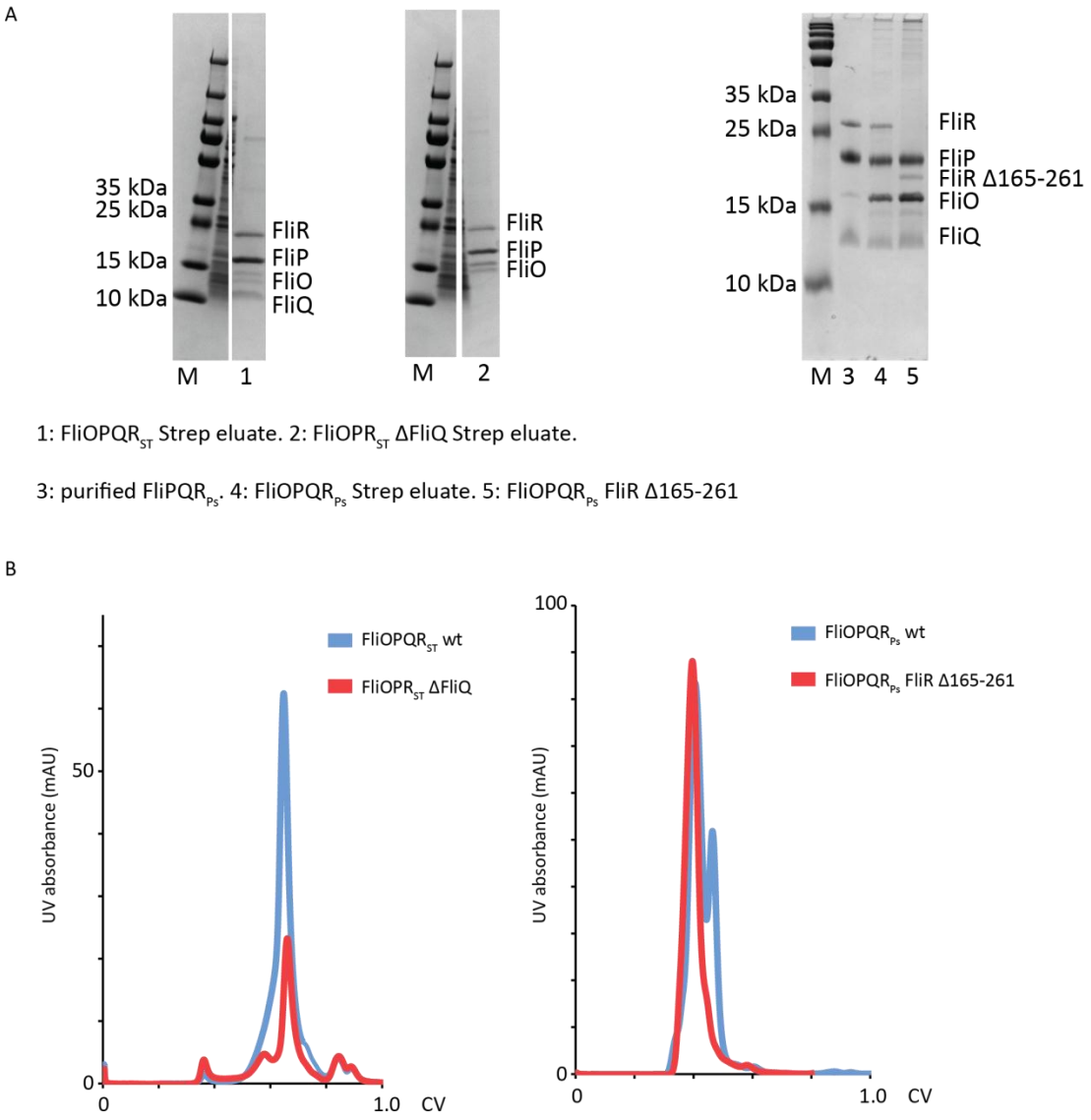


Figure 5.18 FliOPQR mutant pull-down experiments.
A, SDS-PAGE analysis of the purification of FliOPQR$_{ST}$ wild type (left), FliOPR$_{ST}$ ΔFliQ (centre), and FliOPQR$_{Ps}$ wild type and FliR Δ165-261 (right) in LMNG. **B**, Analytical SEC of FliOPQR$_{ST}$ wild type (blue) and ΔFliQ mutant (red) (left) and FliOPQR$_{Ps}$ wild type (red) and FliR Δ165-261 mutant (blue)

In order to further investigate the assembly pathway, mutant pulldown assays were designed. If FliQ/SctS assembles last, it is unclear why it is in the middle of the FliOPQR/SctRST operon (Figure 3.1). In order to test if FliPR/SctRT can assemble in the absence of FliQ/SctS, a deletion of FliQ/SctS was designed. As the sctS stop codon and the sctT start codon overlap in *S. flexneri*, the mutant was designed in the *S.* Typhimurium FliOPQR sequence. FliQ residues 6 to 84 were deleted, leaving the sequences adjacent to FliP and FliR unchanged in order to avoid affecting translation of these genes (i.e. creating a polar effect mutation). The mutated FliOPQR$_{ST}$ construct and

the wild type construct were expressed in BL21 cells and membranes were prepared following protocol 2.5.3. 9 ml of the membrane fractions of wild type and mutant were mixed with 1 ml 10% (w/v) LMNG and the rest was discarded. After stirring for one hour, the clarified solution was applied to a 5 ml StrepTrap column (GE), which was washed with TBS containing 0.01% (w/v) LMNG and eluted with the same buffer supplemented with 10 mM desthiobiotin. The eluted protein was concentrated and finally SEC was performed using a Superose 6 increase 10/300 column (GE) with the running buffer TBS containing 0.01% (w/v) LMNG. This revealed that the FliQ deletion caused a small reduction in yield, but an assembled FliPR complex could be pulled down (Figure 5.18), consistent with it being an intermediate in the assembly pathway.

As described above, FliR/SctT is a structural fusion of FliP/SctR and FliQ/SctS (Figure 4.9). This makes the C-terminal hairpin of helices in FliR/SctT the equivalent of a "$FliQ_0/SctS_0$". If FliQ/SctS only assembles once the FliPR/SctRT complex has formed, this $FliQ_0/SctS_0$ feature of the assembly intermediate may nucleate assembly of the FliQ/SctS subunits. If this is correct, deletion of this hairpin in FliR/SctT should prevent assembly of FliQ/SctS. A complication with a pull-down experiment using such a deletion mutant could be that FliQ/SctS becomes destabilised and is lost in detergent. This makes it challenging to distinguish failure of assembly of FliQ/SctS in the membrane, which was the question of this experiment, from later dissociation after extraction with detergent, which only tests the stability of the already assembled complex. Therefore, the most detergent-stable homologue, $FliPQR_{Ps}$ (Figure 5.8), was chosen for this experiment. For this pull-down protocol 2.5.3 was scaled down. 200 ml of TB supplemented with 0.1% (w/v) rhamnose monohydrate and 60 µg/ml kanamycin were inoculated with Bl21 cells carrying a pT12 plasmid encoding $FliOPQR_{Ps}$ wild type/FliR ΔG165-R261 with a C-terminal strep tag on FliR and incubated for 8 hours at 37 °C and 180 rpm. The harvested cells were lysed in an emulsiflex C5 homogeniser (Avestin) and the clarified lysate was spun in an ultracentrifuge (235,000x g, 1 hour, 4 °C) and the membranes were homogenised in TBS (8 ml per gram of membrane). LMNG was added to 1% (w/v) and after stirring for one hour the solution was clarified by centrifugation. Finally, the solubilised membranes were applied to StrepTrap columns (GE), which were washed with TBS containing 0.01% (w/v) LMNG and eluted with the wash buffer supplemented with 10 mM desthiobiotin.

Analytical SEC was performed using a Superdex 200 increase 10/300 column (GE) with TBS containing 0.01% (w/v) LMNG as the running buffer. Analysis of the strep eluate by SDS-PAGE revealed a FliQ band of similar intensity in both wild type and mutant (Figure 5.18A), despite the removal of almost the entire interface between $FliQ_1$ and FliR. This suggests that although FliR/SctT may contribute to assembly and stabilisation of FliQ/SctS, it is not essential, and assembly of FliQ/SctS instead depends on FliP/SctR. This is consistent with the retention of SctS binding in a smaller truncation of SctR (Figure 3.6C) in $SctRST_{Sf}$. Interestingly, only one peak was observed for the mutant by SEC (Figure 5.18B), but two in the wild type. Purification of FliOPQR leads to two SEC peaks, one corresponding to the FliOPQR assembly intermediate and one to the final FliPQR complex (Figure 3.2), which elutes around 0.5 CV on this column. In the case of the $FliR_{Ps}$ truncation mutant the entire sample appeared to remain bound to FliO in the FliOPQR intermediate complex during SEC as only one peak close to the void was observed. This suggests that the FliO chaperone was bound more tightly to the mutant complex and did not dissociate during SEC as easily as in the wild type (Figure 5.18), suggesting that the chaperone "recognises" that the complex is not complete. It is not known how FliO is released upon completion of the assembly. Assembly of FliP, FliQ and FliR is not sufficient for complete release of FliO (Figure 3.2), but it is possible that *in vivo* assembly of the basal body protein FliF and the EA protein FlhA around FliPQR facilitates chaperone release.

5.7 Discussion

In this chapter the structures of three homologues of FliPQR were described, revealing the high degree of conservation of this structure and demonstrating again that the fundamental principles of type three secretion are conserved in both injectisomes and flagella.

The structures of all export gate homologues were in the same conformational state, but a variable stoichiometry of FliQ/SctS was observed with up to 5 FliQ subunits in $FliPQR_{Ps}$ and a mixture of complexes containing 2, 3 or 4 SctS subunits in SctRST. Disassembly of the complex, which presumably is fully assembled in the bacterial inner membrane, proceeded following the helical pseudosymmetry of the complex (Figure 5.17) which

combined with analysis of the hydrophobic surface of the intermediates suggests an assembly pathway that mirrors disassembly (Figure 5.19).

The earlier intermediate states in the assembly pathway are highly speculative, as none of these sub-complexes have been observed, but using the structures of the subunits in the final complex a series of intermediate complexes can be constructed. A possible pathway, in which all intermediates expose hydrophobic surfaces that are compatible with insertion in the inner membrane, starts with assembly of SctR down the helix of $SctRST_{Sf}$, i.e. starting with $SctR_1$ and continuing with $SctR_2$ to $SctR_5$, followed by addition of SctT. As subunits assemble, the hydrophobic part of SctR is pushed out of the membrane and interactions with lipids are replaced by interactions with the next SctR subunit. In this way, a large proportion of the complex is periplasmic following assembly, but before incorporation in the basal body, despite each subunit being inserted into the inner membrane by the SEC machinery.

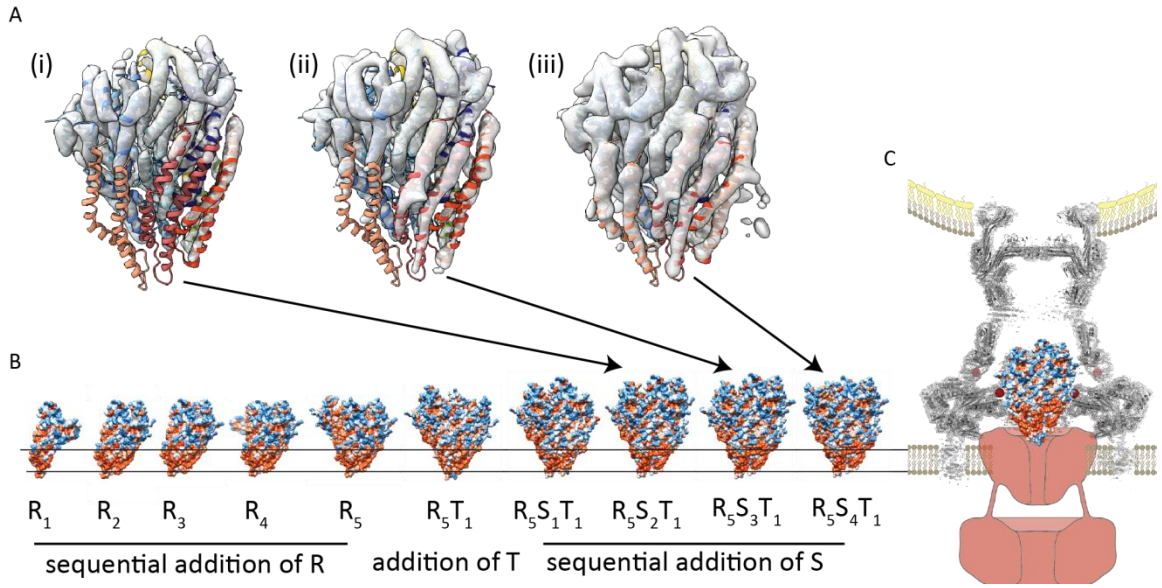


Figure 5.19 Assembly model of SctRST (FliPQR) in the membrane.
A, 3D classification results in three different assembly states of SctRST in which a different number of SctS subunits is bound: (i) $SctR_5S_2T_1$, (ii) $SctR_5S_3T_1$ and (iii) $SctR_5S_4T_1$. **B**, The height of the surface-exposed hydrophobic patches of hypothetical assembly intermediates is consistent with sequential assembly of SctR. As more subunits assemble from the membrane facing side of the complex, following the helical pseudosymmetry, part of the complex is pushed out of the membrane. **C**, The final step in the assembly is assembly of the basal body (grey density and cartoon) around SctRST, raising it above the membrane, while the EA protein SctV forms a channel in the inner membrane.

The main remaining question regarding the assembly of FliPR/SctRT is at what point in the pathway FliR/SctT is incorporated. Although analysis of the hydrophobic surface suggests that FliR/SctT could be assembled after a $FliP_5/SctR_5$ intermediate has formed, there are two complications:

1. When FliR and FliQ were dissociated from the complex using a harsh detergent, only FliP hexamers but no other FliP oligomers were obtained (Figure 3.4), suggesting $FliP_5$ is not stable. Furthermore, overexpression of FliP leads to formation of FliP hexameric rings (Fukumura et al. 2017). If a FliP/SctR oligomer begins to form in the membrane, how is assembly paused at the pentameric stage with a gap in the ring until FliR fills this gap instead of a sixth FliP subunit completing the ring? It is possible that the conserved operon structure of FliPQR/SctRST leads to tightly controlled relative rates of translation so that FliR/SctT is made just in time to fill the gap in the FliP/SctR pentamer before a sixth FliP/SctR subunit can complete the hexameric ring.

2. Blue-native PAGE of membranes from *Salmonella* has been used to show that assembly of FliO-FliP assembly intermediates is dependent on FliR (Fabiani et al. 2017), suggesting that at native expression levels FliP is not able to oligomerise without nucleation by FliR.

How might these complications be reconciled with the analysis of hydrophobic patches of subunits in the assembled complex? It is possible that FliR/SctT nucleates assembly of FliPR/SctRT in a different conformation from what is found in the final complex. The shape of FliR/SctT is not compatible with its insertion in the membrane as a monomer (chapter 4), but it would be expected that the translocon inserts the protein into the membrane in a more canonical transmembrane protein fold immediately after synthesis.

6 The membrane domain of the switch protein FlhB regulates FliPQR

6.1 Introduction

The structures of the FliPQR/SctRST complex described in chapters 4 and 5 are all closed. Fitting of these structures into maps of the assembled needle complex (Hu et al. 2018) shows that the complex is also closed in the context of the assembled nanomachine, raising the question of how the pore opens. One possibility is that other subunits of the export apparatus that bind secretion substrates in the cytoplasm could facilitate opening. One candidate for this possible function is the FlhB protein, whose C-terminal cytoplasmic domain, $FlhB_C$, is known to bind substrates (Evans et al. 2013). Interestingly, $FlhB_C$ is an autoprotease and cleavage in $FlhB_C$, splitting the sequence into $FlhB_{CN}$ and $FlhB_{CC}$, is required for switching of secretion from early to late substrates (Fraser et al. 2003; Zarivach et al. 2008). The N-terminal transmembrane domain of FlhB has not been structurally characterised, but many mutations affecting function have been identified in the N-terminus of FlhB (Minamino et al. 2003) and the linker between transmembrane and cytoplasmic domains, $FlhB_{CN}$ (Inoue et al. 2019a).

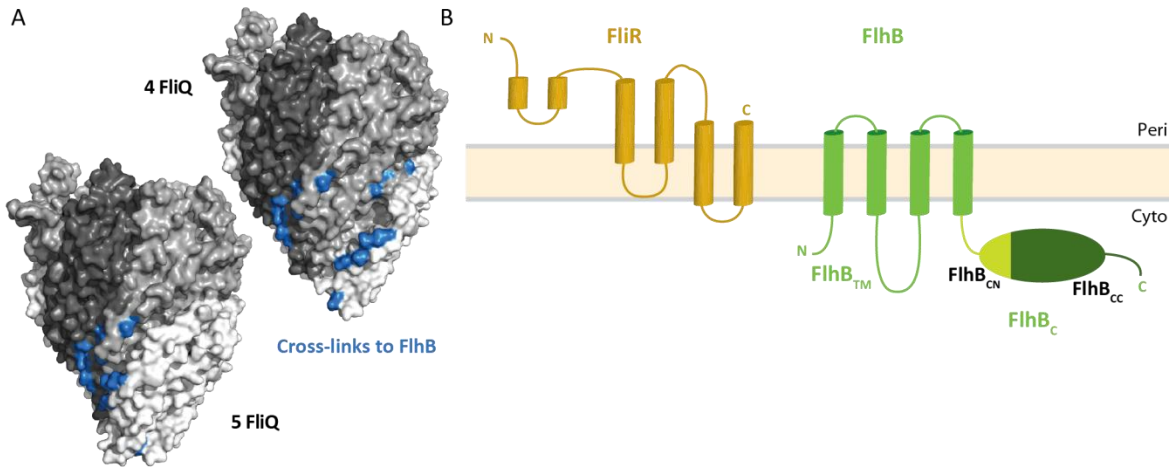


Figure 6.1 Contacts between FlhB and FliPQR.
A, Mapping of photocrosslinks between the FliPQR complex and FlhB (identified using the *S.* Typhimurium SPI-1 complex SctRSTU in Samuel Wagner's laboratory) on the structure of the FliPQR complex containing 4 or 5 FliQ subunits. **B**, Experimental topology of FliR (chapter 4) and predicted topology of FlhB (Erhardt et al. 2017).

The structures of FliPQR complexes were difficult to interpret in terms of FlhB binding due to differences in FliQ numbers which overlapped the predicted FlhB binding site.

Does FlhB bind to a relatively large site on the FliPQR complex predicted by mapping *in vivo* photocrosslinks (identified in Samuel Wagner's laboratory) on a structure with 4 FliQ subunits, or the more compact binding site when 5 FliQ subunits are present (Figure 6.1A)? Is the predicted topology of FlhB (Erhardt et al. 2017) correct, given that the C-terminus of FliR can be fused to the N-terminus of FlhB (Van Arnam et al. 2004), but both FliR termini are periplasmic whereas FlhB termini are predicted to be cytoplasmic (Figure 6.1B)? In order to address these issues the previous co-purification of FlhB (chapters 3 and 4), was optimised with the aim of finding the structure of the complex of FliPQR and FlhB.

6.2 Preparation of a pure complex of FliPQR and FlhB

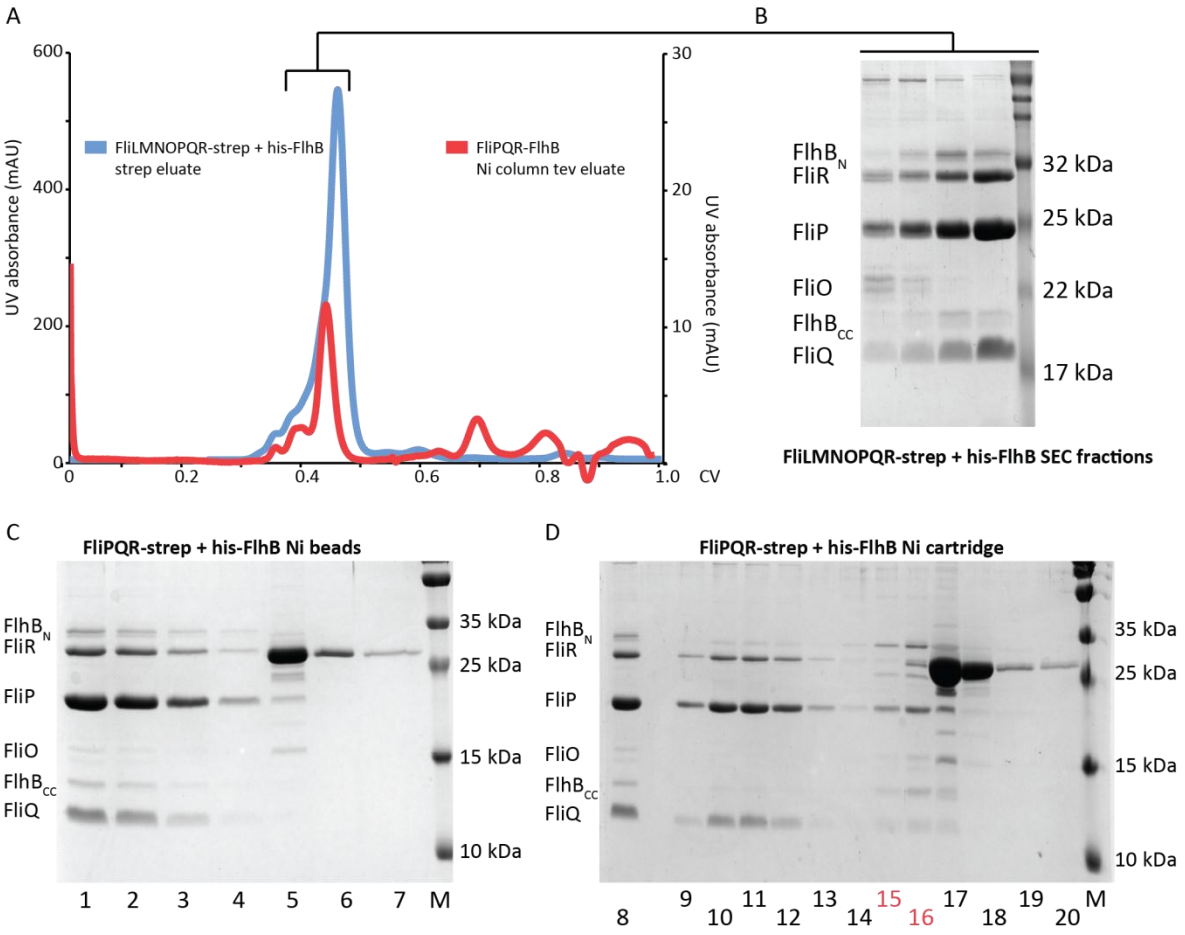
6.2.1 Co-expression of FliPQR and FlhB from different plasmids

When *S.* Typhimurium FliPQR was co-expressed with FlhB and FlhA, limited co-purification of FlhB was observed (chapter 4), but no density corresponding to FlhB was present in the cryo-EM volume determined using this sample. Due to the uncertainty as to why no FlhB was observed, the strategy for determining the structure of the FliPQR-FlhB complex was to first prepare a pure FliPQR-FlhB sample without contamination with FliPQR.

As the previously prepared FliPQR-FlhB complex eluted as a shoulder of the FliPQR peak in SEC (Figure 4.3), it was not possible to fully separate the two complexes by SEC. For purification of pure FliPQR-FlhB, both FliPQR and FlhB were tagged. The FliLMNOPQR operon in the pT12 vector with a C-terminal Strep tag on FliR was co-expressed with FlhB fused to an N-terminal His tag in a pBAD plasmid. Both tags contained a tev protease site. The proteins were purified (section 2.5.5) using the Strep tag and finally SEC (Figure 6.2A). As expected, a shoulder of the main peak contained the FlhB subunit (Figure 6.2B).

The three fractions containing the majority of FlhB were pooled and half of the sample (750 µl) was mixed with 400 µl of 50% (v/v) Ni-NTA beads (Qiagen) and gently stirred for 30 minutes. The beads were washed twice with TBS. Subsequently, protein was eluted by incubating with tev protease (0.75 ml TBS mixed with 0.25 ml of 1 mg/ml tev in 25%

(v/v) glycerol, at 4 °C) overnight in order to avoid destabilising the complex with a high concentration of a solute like imidazole. In order to test if any protein is very tightly bound, additional elution was carried out with 400 µl TBS containing 600 mM imidazole and finally 400 µl of 100 mM EDTA. This revealed that binding of the complex to nickel beads is very inefficient (Figure 6.2C) as the majority of protein was in the flow-through and wash fractions. It is possible that the incubation time with the beads was too short. In order to improve binding efficiency, a cartridge was used instead. The remaining 750 µl of the FliPQR and FliPQR-FlhB mixture were pooled with the flow-through and wash fractions from the nickel beads and applied to a 1 ml Ni-NTA cartridge (Qiagen). The column was washed with 2.5 ml of TBS without imidazole. Finally, 1 ml of tev protease (0.75 ml TBS mixed with 0.25 ml of 1 mg/ml tev in 25% (v/v) glycerol) was injected into the column and incubated overnight at 4 °C. Additional elution was carried out with 1 ml 300 mM imidazole in TBS and 1 ml 100 mM EDTA. SDS-PAGE analysis revealed that tev was sufficient to elute all the protein that had bound to the column, but a small proportion of the complex containing FlhB was in the flow-through (Figure 6.2D). The eluate was pooled and analysed by SEC (Figure 6.2A), revealing a single peak eluting at the same volume as the shoulder of the strep eluate, supporting that a pure FliPQR-FlhB complex had successfully been prepared. The elution volume of the peak was consistent with the shoulder of the peak resulting from the FliPQR and FliPQR-FlhB mixture. However, the yield was very low (<5 µg/litre of culture), complicating structural studies.



1: loaded sample, 2: flow-through, 3: wash 1, 4: wash 2, 5: tev eluate, 6: imidazole eluate, 7: EDTA eluate

8: loaded sample, 9-12: flow-through, 13-14: wash, 15-16: tev eluate, 17-18: imidazole eluate, 19-20: EDTA eluate

Figure 6.2 Purification of FliPQR-FlhB from *S.* Typhimurium.
A, SEC profile of purified FliPQR-FlhB after strep-pulldown using the tag on FliR (blue, left hand side axis), and after additional purification using the his-tag on FlhB (red, right hand side axis). The complex purified using both strep and his tag elutes at the same volume as the shoulder of the complex purified only using the strep tag. **B**, 15% SDS-PAGE gel analysis of the fractions from the SEC profile of FliLMNOPQR + FlhB (A). **C**, 15% SDS-PAGE gel analysis of purification of FliPQR-FlhB from the strep eluate using nickel beads showing presence of FlhB in the flow-through fractions. **D**, 15% SDS-PAGE gel analysis of purification of FliPQR-FlhB from the strep eluate using a nickel cartridge showing presence of FlhB in the eluate of the nickel column.

One problem in the first preparation of homogeneous FliPQR-FlhB was the very low yield. In order to attempt improving this, FlhB was fused to a C-terminal rather than an N-terminal his tag, however this did not improve the yield when purification was attempted at a small scale (data not shown). Therefore, the previous experiment using an N-terminal his tag was repeated and scaled up, but the yield dropped to less than 2

µg/litre of culture in two separate replication attempts. This may be due to the longer delays between steps when processing larger amounts of biomass.

As purification with multiple tags requires multiple steps of washing of the protein bound to a resin, which may strip essential lipids and disrupt a fragile membrane protein complex, a protocol with a single affinity purification step was designed. An N-terminal twin-strep tag was fused to FlhB in a pBAD plasmid and co-expressed with FliOPQR without a tag in the pT12 vector. The complex was expressed as described above for the his-tagged FlhB and purified using the same protocol as used for FliPQR (section 2.5.3), except that the ratio of LMNG to membrane was lowered as described above. 8 µg/litre of culture of FliPQR-FlhB complex could be purified (Figure 6.3), which was more than achieved previously but still very low.

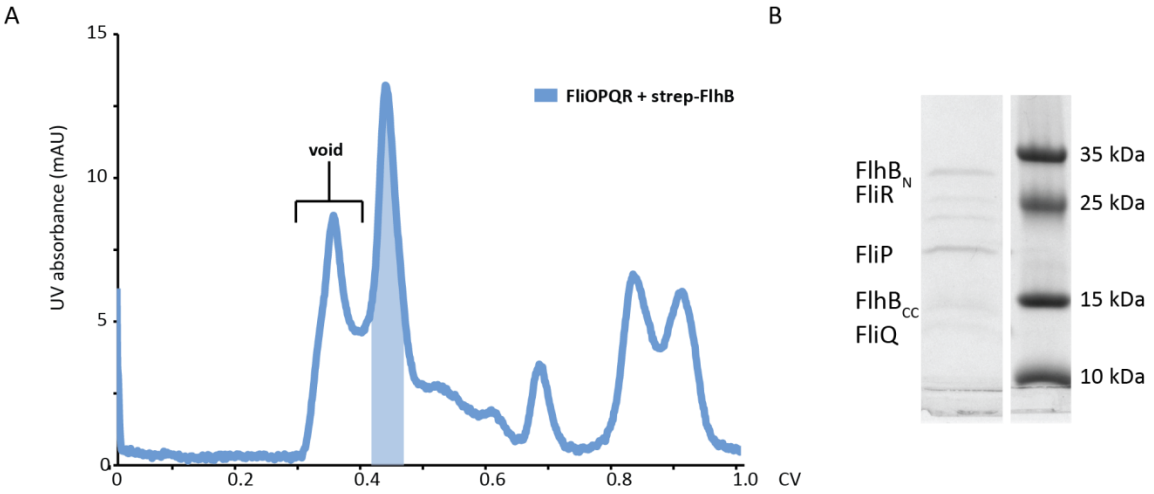


Figure 6.3 Purification of strep-tagged FliPQR-FlhB from *S.* Typhimurium.
A, SEC profile of the FliPQR-FlhB complex purified using a strep tag on FlhB. **B**, 4-20% SDS-PAGE gel analysis of the pooled peak fractions around 0.45 CV from (**A**), indicated by blue shading.

6.2.2 Expression of FliPQR-FlhB/SctRSTU from a single plasmid

One complication of the *S.* Typhimurium system is that FliOPQR and FlhB are found as part of separate operons, requiring co-expression from two different plasmids. It is possible that expression from a single operon, resulting in translation of all components from the same mRNA, would facilitate correct assembly. Therefore multiple homologous

sequences were selected for expression of FliOPQR-FlhB/SctRSTU from a single plasmid with a C-terminal twin strep tag on FlhB/SctU. The operon structure of *S. flexneri* and *S.* Typhimurium SPI-2 SctRSTU and *P. savastanoi* and *V. mimicus* FliOPQR-FlhB allowed cloning of all four/five genes into plasmid pT12 with a C-terminal twin-strep tag. In addition *C. sporogenes* FliOPQR-FlhB was cloned into the pT12 vector in the same way and an additional construct was made in which *C. sporogenes* FliO was replaced by *S.* Typhimurium FliO and the *S.* Typhimurium FliP signal sequence was fused to the mature sequence of *C. sporogenes* FliP, as expression in *E. coli* may not allow recognition of gram positive signal sequences. All six constructs were purified following the same protocol as was used for FliPQR/SctRST previously (section 2.5.3) but 1 ml of 10% (w/v) LMNG solution per 1.2 grams of membrane was used for extraction. The complex could only be purified using the *S.* Typhimurium SPI-2 SctRSTU and *V. mimicus* FliPQR-FlhB constructs. The final purification step was SEC using a Superose 6 increase 10/300 column (GE) (Figure 6.4A). In both cases a complex could be purified which appeared to contain FlhB/SctU (Figure 6.4B). The FliPQR complex prepared earlier (chapters 3-5) elutes around 0.55 CV and the FliOPQR intermediate complex elutes around 0.45 CV on this column (Superose 6 increase 10/300 (GE)) consistent with peak 1 corresponding to FliOPQR-FlhB and peak 2 corresponding to FliPQR-FlhB. Yields of just above 100 µg per litre of culture could be achieved for *V. mimicus* FliPQR-FlhB (peak 2). Additionally, the *P. savastanoi* construct yielded protein, but the majority eluted in peak 1, the putative elution volume of the intermediate FliOPQR-FlhB complex. Therefore the *V. mimicus* sample was prioritised.

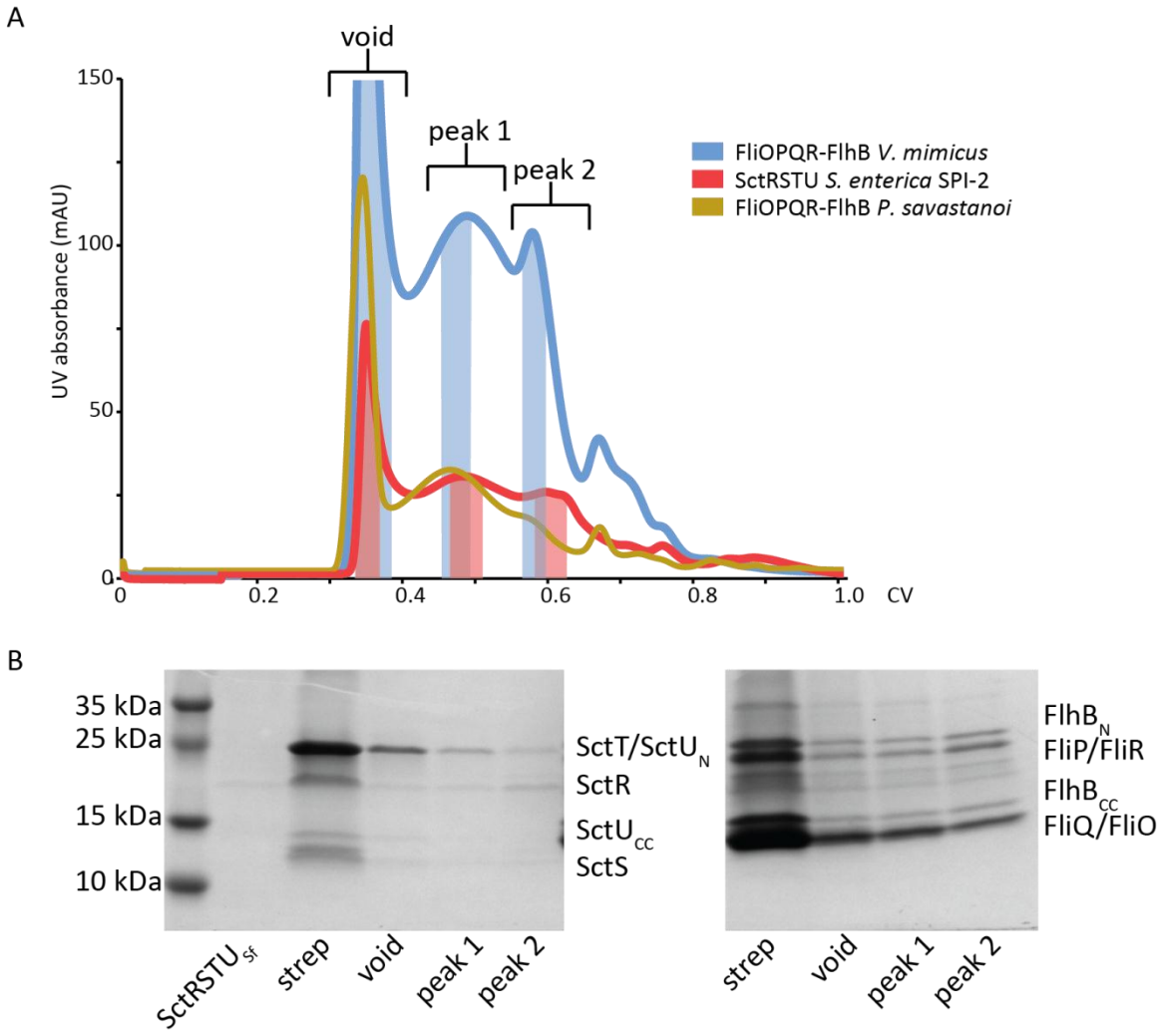


Figure 6.4 Purification of FliPQR-FlhB and SctRSTU.
A, SEC profile of purified FliPQR-FlhB and SctRSTU. **B**, 4-20% SDS-PAGE gel analysis of the strep eluate of *S. flexneri* SctRSTU ($SctRSTU_{Sf}$) as a representative example of a failed purification and the strep eluate and SEC peaks of SctRSTU from *S.* Typhimurium SPI-2 (left) and *V. mimicus* FliPQR-FlhB (right) as indicated in (**A**).

Initial cryo-EM imaging using sample from the peak 2 SEC fractions indicated that *S.* Typhimurium SPI-2 SctRSTU, like the SctRST complex from this system, dissociated during the cryo-EM grid preparation process and this sample was not characterised further. It may be possible to image this sample in the future after chemical cross-linking. When *V. mimicus* FliPQR-FlhB was imaged, a mixture of top and side views was observed and this sample was selected for data collection.

6.3 Structure determination of *Vibrio* FliPQR with FlhB

6.3.1 Structure of FliPQR with FlhB

Purified *V. mimicus* FliPQR-FlhB was concentrated to 2.5 mg/ml and 3 µl of sample were applied to a carbon-coated holey carbon grid (Au300, R1.2/1.3, Quantifoil) in a Vitrobot Mark IV (FEI) at 100% humidity at room temperature. After 10 seconds the grid was blotted for 3 seconds and plunged into liquid ethane cooled by liquid nitrogen. The relatively high sample concentration was chosen to achieve a high particle density (Figure 6.5A). A dataset of more than half a million particles (2,698 micrographs) was collected using a Titan Krios microscope (FEI) (Table 6.1).

Particles were extracted in a 256x256 pixel box. 2D classification in RELION-2.0 (Kimanius et al. 2016) revealed a mixture of top and side views (Figure 6.5B), despite the absence of fluorinated detergent, which was required to obtain side views of *V. mimicus* FliPQR in LMNG (chapter 5), indicating that binding of FlhB can cause a change in orientation of the particle. 3D classification was carried out using the structure of FliPQR from *S.* Typhimurium as an initial model (Figure 6.5C). In order to avoid overfitting, the particle alignment was only carried out using information up to 7 Å. As the particles were split into multiple classes with detectable detail but none with a clear density corresponding to FlhB, the three least noisy classes were selected and classified again. This led to an intermediate resolution reconstruction (4.8 Å after auto-refinement and post-processing in RELION-2.0) revealing additional density compared to the previous FliPQR data.

Additional data were collected from the same grid and following separate 2D classification the datasets were merged. Using the earlier intermediate resolution volume as an initial model, the particles were classified into four classes, two of which (class 2 and 3) contained few high-resolution features while the other two (class 1 and 4) were very similar reconstructions in which secondary structure elements could be seen. Auto-refinement and post-processing of just class 1 led to a 4.0 Å map, while combining class 1 and class 4 led to a 4.1 Å map. Once RELION-3.0 was released (Zivanov et al. 2018; Zivanov et al. 2019), the particles were further refined using CTF refinement, particle

polishing and motion correction in RELION-3.0, leading to a final 3.2 Å volume (Figure 6.5D).

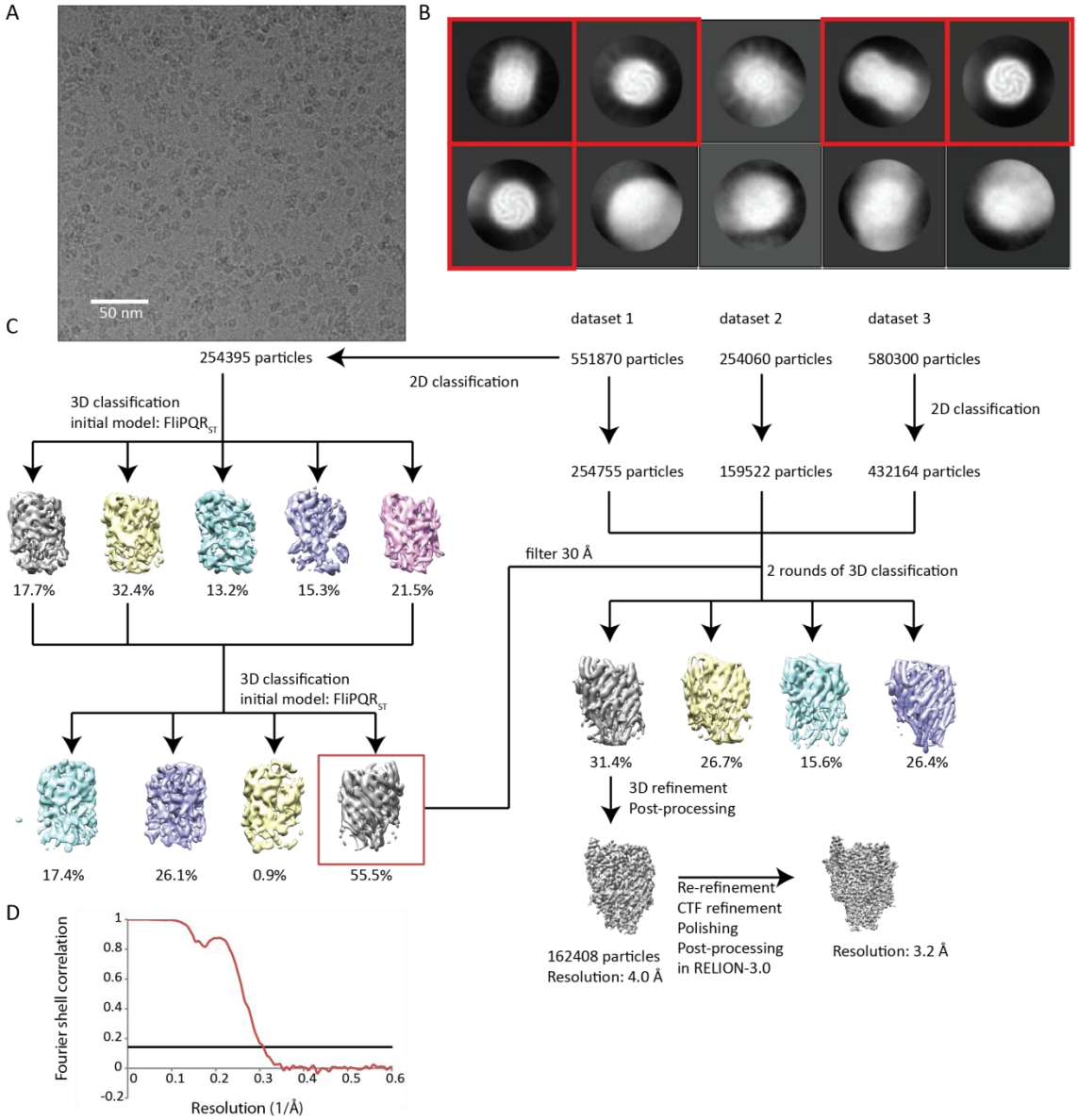


Figure 6.5 Structure determination of FliPQR-FlhB.
A, Representative micrograph (Defocus=3.7 µm). **B**, 10 most populated 2D class averages of dataset 3. Red boxes indicate selected classes. **C**, Particle classification strategy in RELION. **D**, Solvent corrected FSC curve of the final volume.

Data collection and processing	
Voltage	300
Electron exposure	48 $e^{-}/Å^2$
Micrographs	9,461
Particles (total)	1,386,230
Particles (final)	162,408
Sampling	0.822 Å
Defocus range	0.5-4 μm
Refinement	
Resolution	3.2 Å
FSC threshold	0.143
Symmetry	C1
Map sharpening B factor	-97 $Å^2$

Table 6.1 Cryo-EM data collection and processing statistics for FliPQR-FlhB.

The volume revealed an object similar to the previously determined structures of the FliPQR complex with an additional density contacting FliR and $FliP_5$ (Figure 6.6A). The high resolution of the map (Figure 6.6B) allowed the construction of an atomic model. The model was built in the map by placing the *P. savastanoi* FliPQR structure (chapter 5) in the density and manually mutating the sequence to the *V. mimicus* sequence in Coot (Emsley et al. 2010). Despite the apparently good fit of $FliQ_5$ in the map, analysis of the connectivity of the map (Figure 6.6) showed that the complex contains only four copies of FliQ but part of the FlhB structure resembles FliQ (see below). In order to obtain a model for FlhB, a structure was predicted *de novo* using the RaptorX server (Wang et al. 2017) and fitted into the density (Figure 6.6C). As the fit was locally good but globally poor, extensive manual editing of the co-ordinates was required. Short segments of the RaptorX model were fit into the density as rigid bodies using the position of bulky hydrophobic residues as a guide where the model was of sufficient quality. The remaining residues were built manually. The fit of the residues into the density was improved using real space refinement in Coot. Finally, the model geometry was improved through multiple iterations of the program phenix.real_space_refine (Afonine et al. 2018) and manual editing of rotamer and Ramachandran outliers in Coot (Figure 6.6D and Table 6.2).

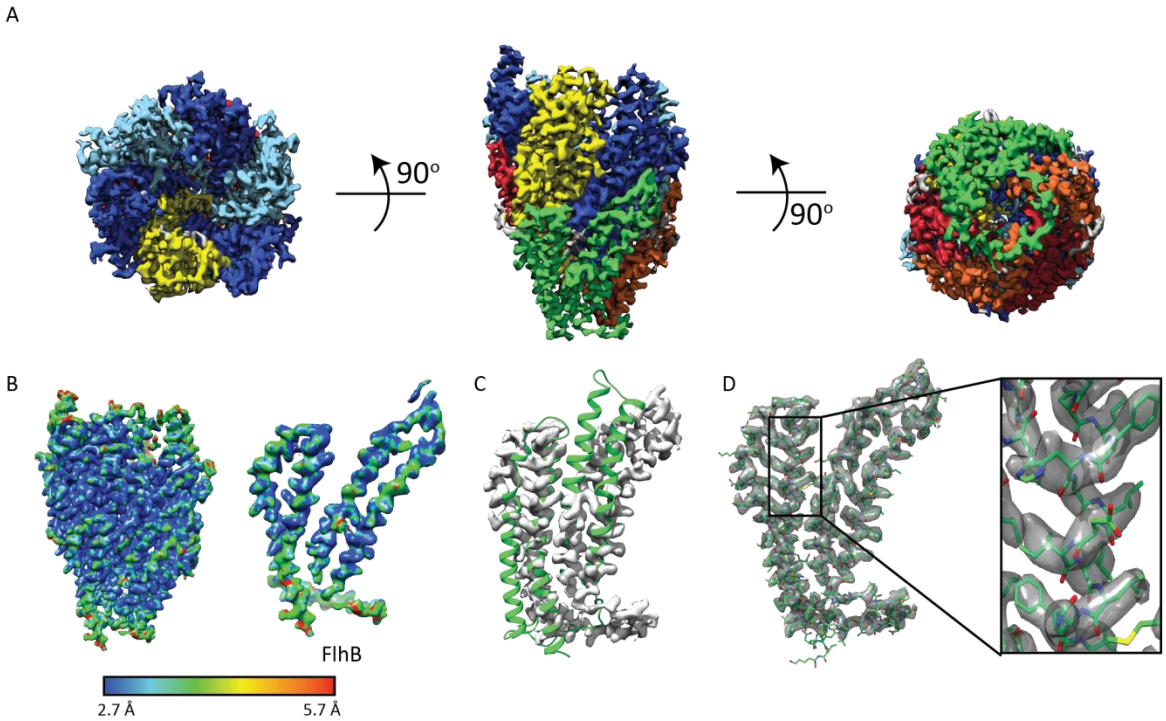


Figure 6.6 Cryo-EM volume of the FliPQR-FlhB complex.
A, Cryo-EM volume of the FliPQR-FlhB complex. FliP is coloured in shades of blue, FliR in yellow, FliQ in shades of red and FlhB in green. **B**, Variation of resolution in the map calculated with ResMap (Swint-Kruse and Brown 2005). **C**, Fit of the RaptorX model (Wang et al. 2017) of FlhB in the density. **D**, Final model of FlhB in the density.

CC_{Mask}	0.8117
B factors	43.55 $Å^2$
R.m.s. deviations	
Bond lengths	0.01 Å
Bond angles	0.93°
Validation	
MolProbity score	2.45
Clashscore	15.56
Poor rotamers	2.72%
Ramachandran plot	
Favoured	93.64%
Allowed	5.84%
Disallowed	0.52%

Table 6.2 Model refinement statistics.

The model (Figure 6.7A) that could be built into the density revealed the structure of the FlhB transmembrane domain ($FlhB_{TM}$, residues 29 to 221) (Figure 6.7B). The previously characterised cytoplasmic domain, $FlhB_C$, (Zarivach et al. 2008) could not be modelled due to a lack of interpretable density, suggesting that the position of $FlhB_C$ is not fixed with respect to FliPQR-$FlhB_{TM}$. The structure of FliPQR excluding FlhB is highly similar to the structure of $FliPQR_{Vm}$ (chapter 5) with an RMSD of only 0.651 Å over 1,531 residues. The topology of $FlhB_{TM}$ was highly unusual; both termini are close to each other in the middle of the structure.

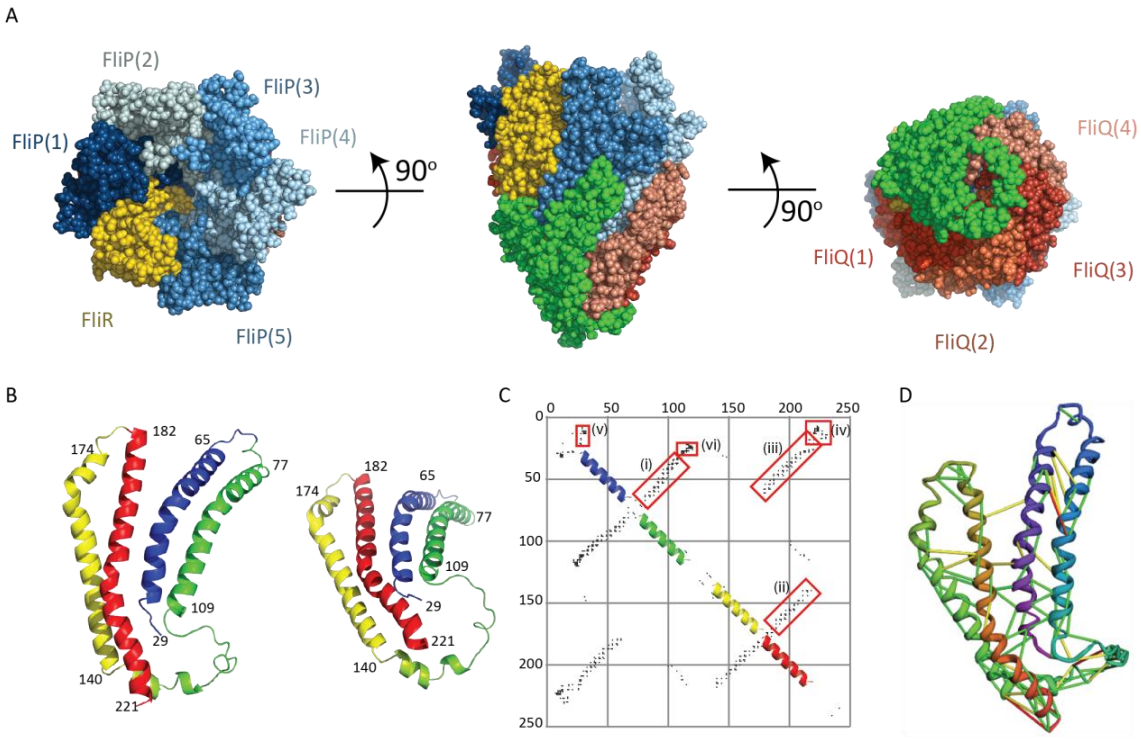


Figure 6.7 Structure of the FliPQR-FlhB complex.
A, Structure of the FliPQR-FLhB complex in which FliR is coloured yellow, FliP blue, FliQ red and FlhB green. **B**, Structure of $FlhB_{TM}$ with numbers indicating the residues at the termini of the four helices. **C**, Map of evolutionary co-variation calculated using RaptorX. Only contacts with a probability greater than 0.5 were plotted. Red boxes highlight the interaction between helix 1 and 2 (i), helix 1 and 4 (ii), helix 3 and 4 (iii), the N-terminus and $FlhB_{CN}$ (iv), within the N-terminus (v) and between the loop between helix 2 and 3 and the N-terminus (vi). **D**, Covariation within $FlhB_{TM}$ mapped onto the structure. Co-variation was calculated in GREMLIN and only contacts with a probability higher than 0.75 are displayed. Green lines indicate co-varying residues with a distance < 5 Å, yellow lines 5 – 10 Å and red > 10 Å.

Interestingly, in the RaptorX prediction of the structure (Figure 6.6C) the arrangement of the helices was overall correct, whereas a previous study using co-variational data predicted the protein to be a 4-helix bundle (Taylor et al. 2016). The authors of that study found that different programs for analysing co-variational data led to predictions of different architectures, including 4-helix bundles and open ("pretzel") structures, but human selection of models and modelling based on finding a consensus between the different methods used led to a final model that is a 4-helix bundle. In order verify that the final structure is consistent with all co-variational data, evolutionary co-variation within the domain was calculated using the RaptorX server (Wang et al. 2017) (Figure 6.7C) and contacts calculated in GREMLIN server (Ovchinnikov et al. 2015) (Figure 6.7D) were mapped onto the structure. This revealed contacts between helix 1 and 2, helix 3 and 4 and helix 1 and 4 but not between helix 2 and 3, consistent with the structure.

$FlhB_{TM}$ occupies the binding site of $FliQ_5$ in the structure of FliPQR from *Pseudomonas savastanoi* (chapter 5) and is very similar in structure to two copies of FliQ (Figure 6.8A). However, the topology is reversed in FlhB (Figure 6.8B), suggesting convergent evolution rather than homology to be the cause of the similarity. The overlay of FlhB and the modelled $FliQ_5$ and $FliQ_6$ also demonstrates that the helices of FlhB continue the spiral of the FliPQR complex (chapter 4). The helical rise between the FliQ subunits is perfectly mimicked by the offset between the two hairpins of the $FlhB_{TM}$ structure.

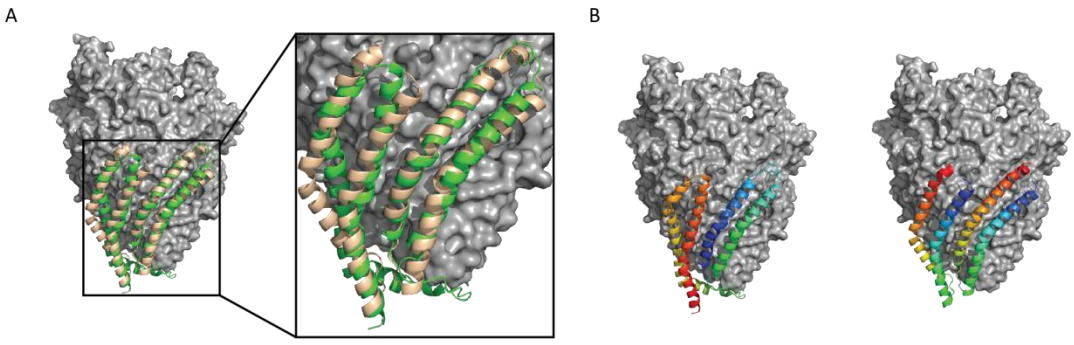


Figure 6.8 FlhB is a mimic of FliQ with inverted topology.
A, Structure of FliPQR-FlhB highlighting FlhB in green. A theoretical $FliQ_5$ and $FliQ_6$ (brown) were modelled by aligning $FliQ_2$ of a $FliQ_2$-$FliQ_3$-$FliQ_4$ object onto $FliQ_4$ of the structure. **B**, Structure of FliPQR-FlhB highlighting FlhB in rainbow colouring (left) and structure of FliPQR of the FliPQR-FlhB complex with a modelled $FliQ_5$ and $FliQ_6$ in rainbow colours.

The most surprising feature of the structure is the loop connecting helix 2 and 3 (residues 110 to 139), which will be referred to as $FlhB_L$ from now. $FlhB_L$ forms a ring shaped structure underneath the cytoplasmic entrance to the lumen of FliPQR, resembling a sphincter (Figure 6.9A, B). Interestingly, this loop contains the most highly conserved residues in the structure (Figure 6.9C).

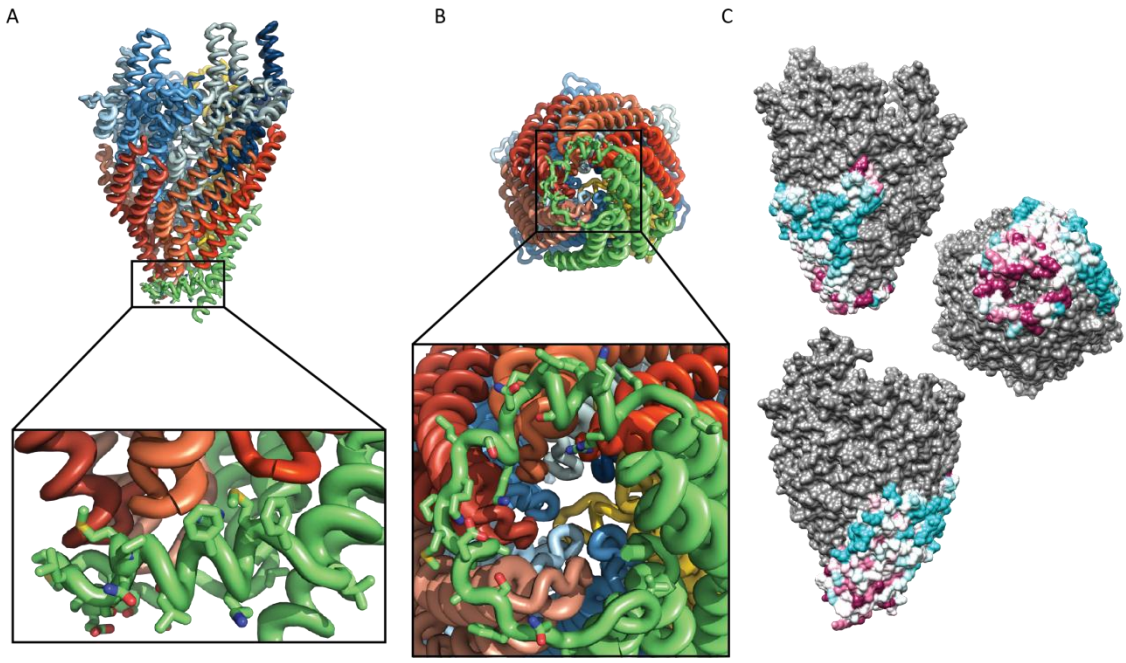


Figure 6.9 $FlhB_L$ forms a ring underneath the complex.
A, Side view of $FlhB_L$. $FlhB_L$ is wrapped around the bottom of the FliPQR complex which faces the cytoplasm. **B**, View from the cytoplasm. $FlhB_L$ forms a ring around the cytoplasmic entrance to the export gate. The majority of sidechains point either at the FliQ subunits or away from them, but not into the inside of the $FlhB_L$ ring. **C**, ConSurf (Ashkenazy et al. 2016) analysis shows that $FlhB_L$ is the most conserved part of $FlhB_{TM}$. Blue indicates low conservation, magenta indicates high conservation.

6.4 Functional studies of $FlhB_L$

The structure and position of $FlhB_L$ in the FliPQR-FlhB complex suggest functional importance. In order to further investigate the function of this loop, mutations were designed and assayed in the motile bacterium *E. coli* strain W. If the loop is blocking secretion and can be moved through hinge motions out of the secretion path in the open state, then mutations that remove the interactions between $FlhB_L$ and FliQ should promote secretion, presumably in an unregulated way. If however $FlhB_L$ is actively involved in opening, for example by adopting a more extended conformation and pulling

the complex open in the process of extending, then removal of the $FlhB_L$-FliQ interaction would prevent secretion as the export gate could not open in such a mutant.

Deletion mutants of fliOPQR and flhB in *E. coli* W cells were made using the Lambda Red system (Datsenko and Wanner 2000) (section 2.3.7) which could then be complemented with plasmids expressing mutated genes. *E. coli* W cells expressing the Lambda Red genes were made electrocompetent and transformed with a linear PCR product of a chloramphenicol cassette with ~50 bp overhangs which were homologous to the sequence surrounding the DNA sequence to be deleted. Integration of this cassette into the genome by the Lambda Red machinery causes the native genes to be replaced by the deletion cassette. FliOPQR was deleted entirely, as these genes are at the end of an operon (Archer et al. 2011). The flhB gene is part of the flhBAE operon, so in order to avoid polar effects on expression of downstream genes, all base pairs corresponding to the sequence between the first 4 and the last 11 residues were targeted, as previously described in *Salmonella* (Fraser et al. 2003).

After transformation of the PCR products, the cells were grown on agar plates containing chloramphenicol (20 μg/ml). The relatively low concentration of antibiotic was used as expression of the gene from the genome is likely to result in less gene product compared to a plasmid. Large and small colonies were observed after overnight growth. One large and one small colony was analysed by PCR of the flhB or fliOPQR genes with the aim of finding a size change following successful replacement by the chloramphenicol cassette. In this way a fliOPQR deletion mutant was obtained (Figure 6.10). As none of the colonies resulting from the flhB deletion experiment had lost motility, further clones were tested, but no deletion mutant could be isolated. Therefore, a linear piece of DNA containing the chloramphenicol cassette with 1000 bp flanking regions was generated by PCR and Gibson assembly using the NEBuilder HiFi assembly mastermix (NEB) and used for the Lambda Red mutagenesis. Two clones were obtained, both of which contained the deletion cassette (Figure 6.10). The selected deletion clones were transformed with plasmid pCP20, which contains a gene that excises the chloramphenicol cassette. pCP20 was removed by picking colonies and growing the cells at 37 °C. The pCP20 origin of replication is temperature sensitive.

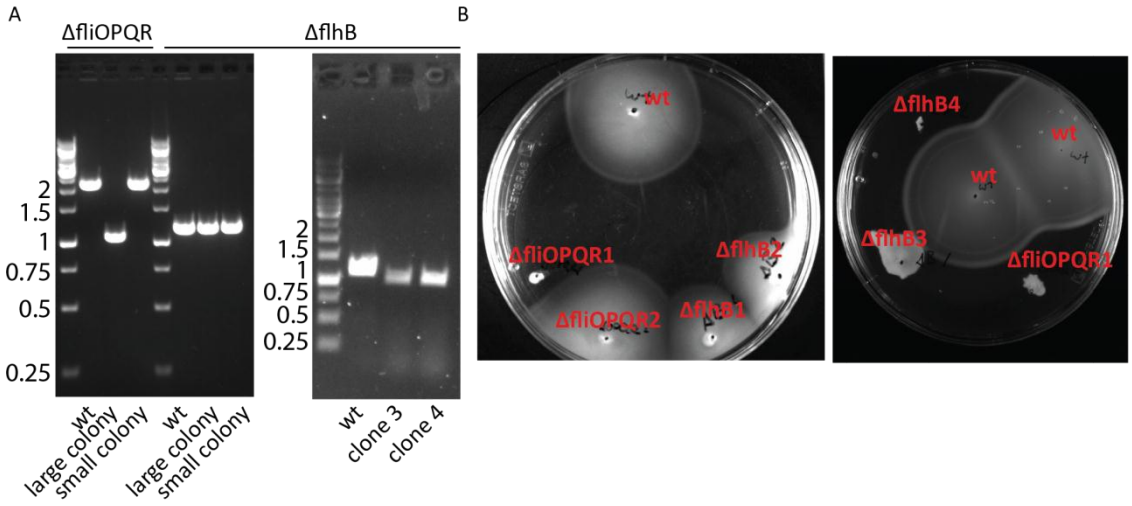


Figure 6.10 Deletion of fliOPQR and flhB in *E. coli* W.
A, PCR products of *fliOPQR* or *flhB* made from wild type cells or deletion mutants. The deletion mutant DNA migrates differently from the wild type as it is smaller. The size of molecular weight markers is indicated in kilobases. **B**, Motility of isolated clones in soft LB agar. Deletion mutants are not motile while isolated colonies in which the deletion was not successful remained motile, indicated by the large circles of motile cells growing around the inoculation site.

Mutations were designed in FliQ and $FlhB_L$ in order to disrupt the interaction between the two. Mutations in FliQ introduce four changes in the complex, one per subunit. Extensive contacts were seen between FliQ A41 and $FlhB_L$ and FlhB A121, M128 and M137 are conserved hydrophobic residues whose sidechains are inserted in between the gaps between FliQ subunits (Figure 6.11), making these residues targets for mutations that would disrupt the FliQ-$FlhB_L$ interaction.

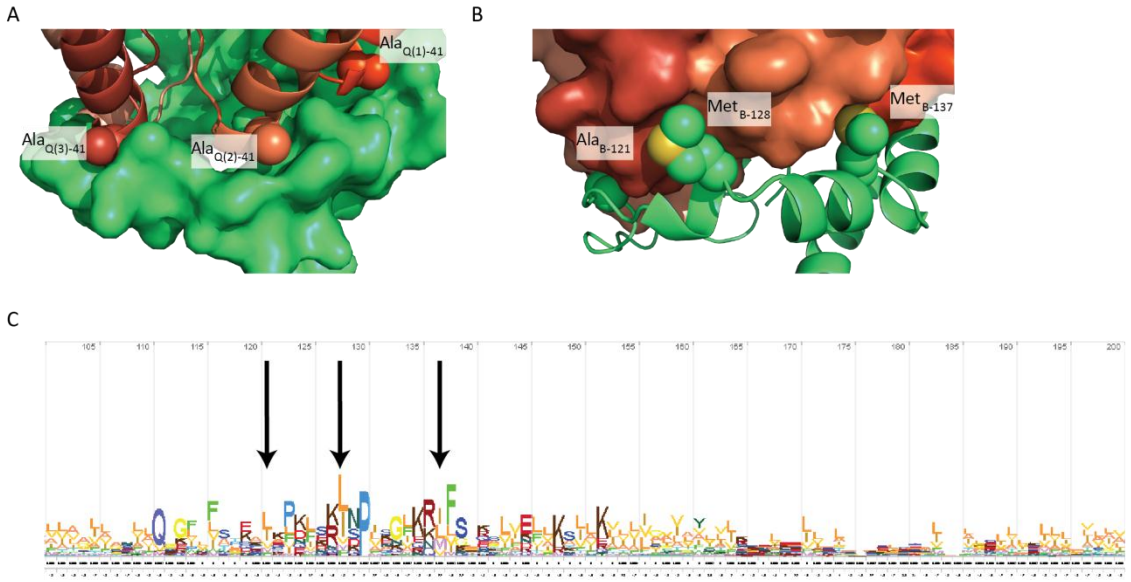


Figure 6.11 FliQ interaction with $FlhB_L$.
A, Interaction of FliQ A41 with $FlhB_L$. **B**, Interaction of FlhB A121, M128 and M137 with FliQ. **C**, Sequence logo of residue 101 to 200 of FlhB generated using the jackhmmer server (Potter et al. 2018) with the FlhB sequence from *V. mimicus*. Arrows indicate the three hydrophobic residues inserted between FliQ subunits in the structure.

The *E. coli* W deletion strains were transformed with pT12 (fliOPQR) or pBAD (flhB) plasmids containing the mutated genes of interest and their motility was assayed using soft agar plates (section 2.4.1). This revealed that mutation of FliQ A41 and T42 to a charged residue or deletion of A41 abolished motility, while mutation to the bulky hydrophobic residue tryptophan only reduced motility, in the case of A41W severely (Figure 6.12A). Similar results were obtained for mutations in $FlhB_L$; while alanine substitutions had only small effects, mutation of two of the $FlhB_L$ hydrophobic residues binding between FliQ subunits (L127 and M136 in *E. coli*) to aspartate abolished motility. Collaborators in Samuel Wagner's group at the University of Tübingen performed a systematic *in vivo* photocrosslinking study based on the structure by mutating residues in the FliQ and FlhB homologues SctS and SctU in the *S.* Typhimurium SPI-1 vT3SS to the large, bulky amino acid *p*Bpa. Mutations in SctS abolished secretion (SctS V39 is A41 in FliQ and SctS T40 is the equivalent of FliQ T42) while SctU was generally more tolerant to mutation (Figure 6.12B), consistent with motility assays in which two mutations were required to cause loss of motility. All sites in SctS could be cross-linked to SctU. Mapping of the identified residues onto the structure highlights the importance of the $FlhB_L$-FliQ interaction for the correct functioning of type 3 secretion (Figure 6.12C).

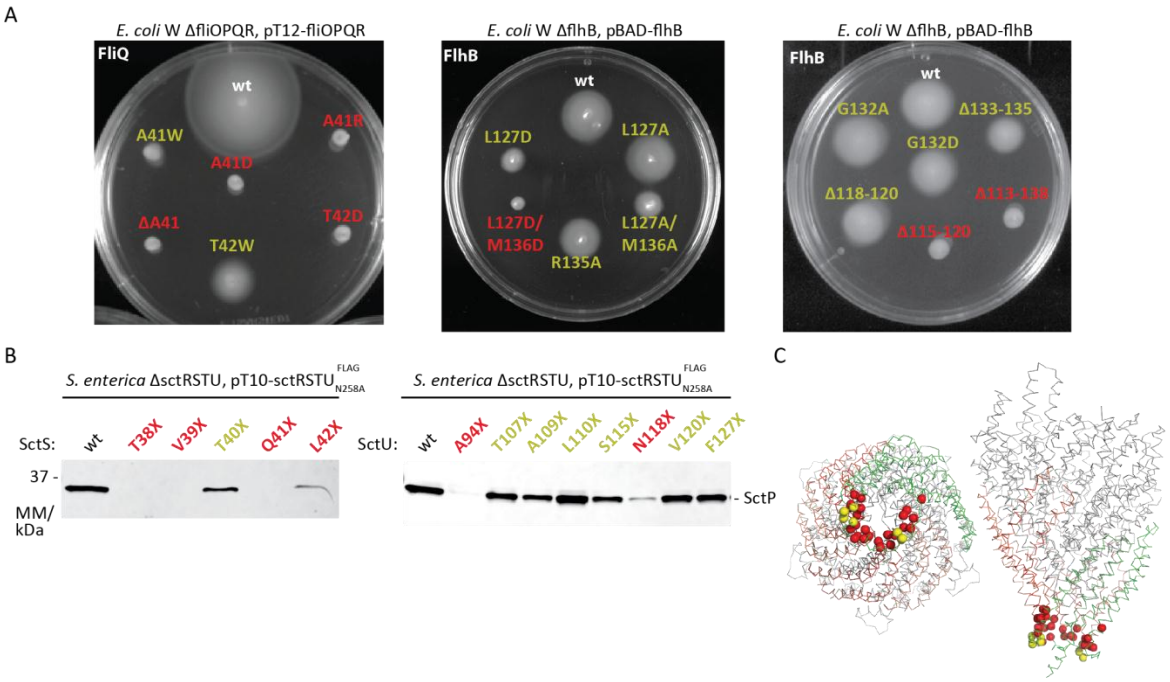


Figure 6.12 Analysis of mutations disrupting the interaction between FlhB$_L$ and FliQ. **A**, Representative swimming motility plates of *E. coli* W deletion mutants complemented with plasmid based genes containing the indicated mutation. Red indicates loss of function and yellow indicates no or only a small effect. Experiments were performed in triplicate. **B**, Secretion data with the indicated mutations in the FliQ and FlhB homologues SctS and SctU generated by collaborators in the Wagner laboratory. Substitution with the synthetic amino acid *p*Bpa is denoted with X. Secretion of SctP was monitored by Western blot. **C**, Mapping of the mutations onto the structure. Residues were labelled red if one substitution mutation caused loss of function even if mutation to a different amino acid had no effect.

The loss of function in many of the mutants suggests that the proper interaction between FliQ and FlhB$_L$ is highly important for the opening and closing of the export gate. Loss of function was observed when charges were inserted into the interaction, but substitution of interacting residues with hydrophobic amino acids had a smaller effect. Although it is possible that negative charge is not tolerated in the secretion process, the same effect was seen when FliQ A41 was mutated to the positively charged arginine or simply deleted, shortening the FliQ loop and removing the interaction site of FlhB$_L$. A potential alternative explanation for the loss of function in some of the mutants would be inability to assemble into the export gate complex. In order to test this, the *V. mimicus* FliPQR-FlhB expression construct in the pT12 vector was mutated to introduce the FliQ A41W and A41D mutations. The mutant complexes were purified using the same protocol as was used for the wild type (section 2.5.3) using the detergent LMNG.

This showed that both mutant complexes could be purified and the analytical SEC profile was very similar to that of the wild type (Figure 6.13).

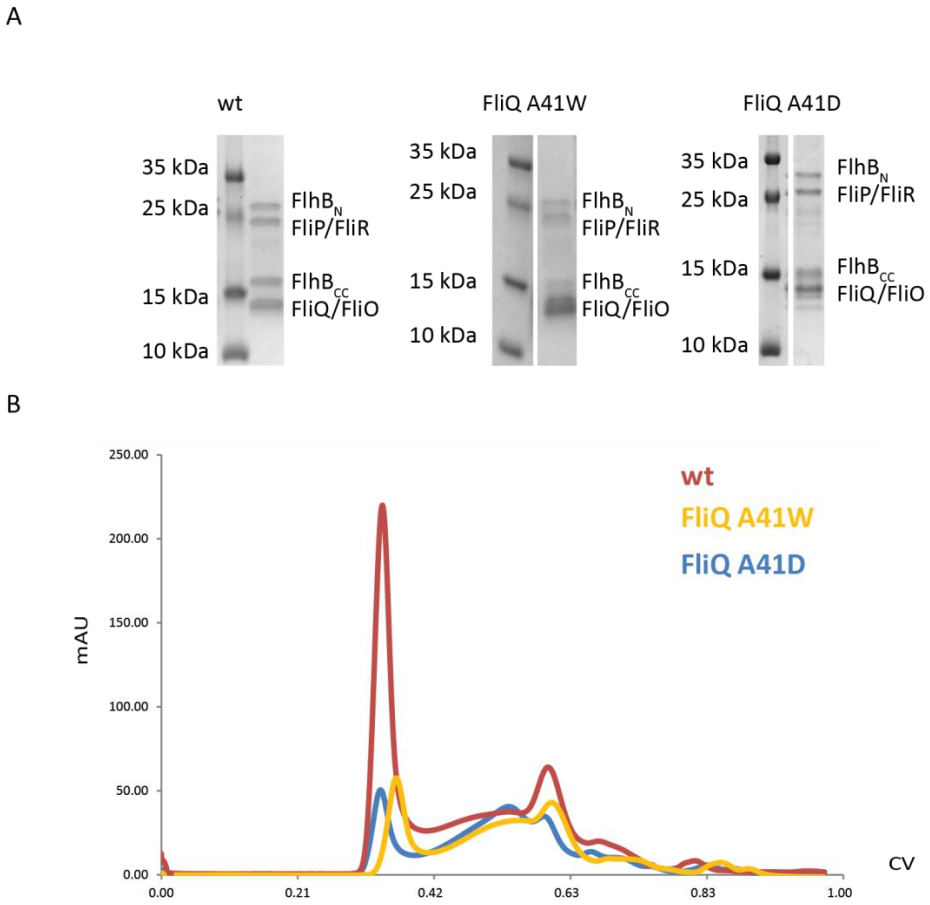


Figure 6.13 Pull-down assay of mutant FliPQR-FlhB complexes.
A, SDS-PAGE analysis of the strep eluates of wild type, FliQ A41W and FliQ A41D FliPQR-FlhB complexes from *V. mimicus* expressed in BL21. **B**, Analytical SEC of the three strep eluates using a Superose 6 increase 10/300 column (GE).

The mutational data suggests that it is more likely that the FlhB$_L$ loop extends and contracts as the FliPQR complex opens and closes in order to maintain the FliQ-FlhB$_L$ interaction throughout the opening and closing cycle, rather than undergoing hinge motions to move out of the secretion path. If FlhB$_L$ is merely required to move away from the centre of the FliPQR-FlhB opening when the complex is open, then pushing FlhB$_L$ away from FliQ through the described mutations would be expected to lead to a strongly secreting phenotype, so that a flagellum could likely still be built, resulting in motility. The loss of motility and secretion in many of the more disruptive mutations, such as residue Ala41 of FliQ, is better explained by a continued interaction of FlhB$_L$ and

FliQ through the secretion cycle. A potential caveat to this interpretation is that a displaced $FlhB_L$ could form a misfolded plug that physically blocks secretion. In order to test this possibility, the entire $FlhB_L$ (residues 113-138) was deleted and replaced with a short GGGSG linker, which also caused loss of motility (Figure 6.12A). If instead of hinge movements the FliQ-$FlhB_L$ interaction is maintained throughout secretion cycles then $FlhB_L$ would need to extend in the open state of the export gate complex. That would require a certain minimum length of $FlhB_L$. Although deletions of three residues had little effect, a deletion of six residues completely abolished secretion (Figure 6.12A), consistent with this idea of $FlhB_L$ extending and contracting as the complex cycles through open and closed states.

In addition, two more $FlhB_L$ residues were targeted. A G127D mutation in $SctU_L$ of *Yersinia* (Riordan and Schneewind 2008) (FlhB G132 in *E. coli*) prevented secretion, however mutation of this residue in *E. coli* W had no effect on motility (Figure 6.12A). In addition, a conserved positively charged residue (R135 in *E. coli*) was mutated, but this mutation had little effect on motility (Figure 6.12A).

6.5 Unidentified ligands

Further inspection of the cryo-EM map revealed a number of extended densities which most likely correspond to lipidic ligands. Although these were also seen in previous structures of FliPQR complexes (chapters 4 and 5), they were clearest in the FliPQR-FlhB map due to its higher resolution. These ligands are found on the outside the central FliPR complex. They are surrounded by two FliQ subunits each, for example helix 2 of $FliQ_1$ and helix 1 of $FliQ_2$ (Figure 6.14A-B). In addition, ligands are also seen in between FlhB helix 2 and $FliQ_4$ and between helices 1 and 4 of FlhB. Although the identity of these ligands is not known, they bind in a hydrophobic cavity and it was possible to hypothetically model a fatty acid in the density (Figure 6.14D). Even though the ligand may not be a fatty acid, it is clear that the density is too small to accommodate a phospholipid or LMNG, the detergent used for purification. The unexpected location of these ligands inside the complex suggests that they are of functional importance and become trapped during assembly. One possibility is that the ligands act like “grease” to facilitate the opening and closing motions of the complex.

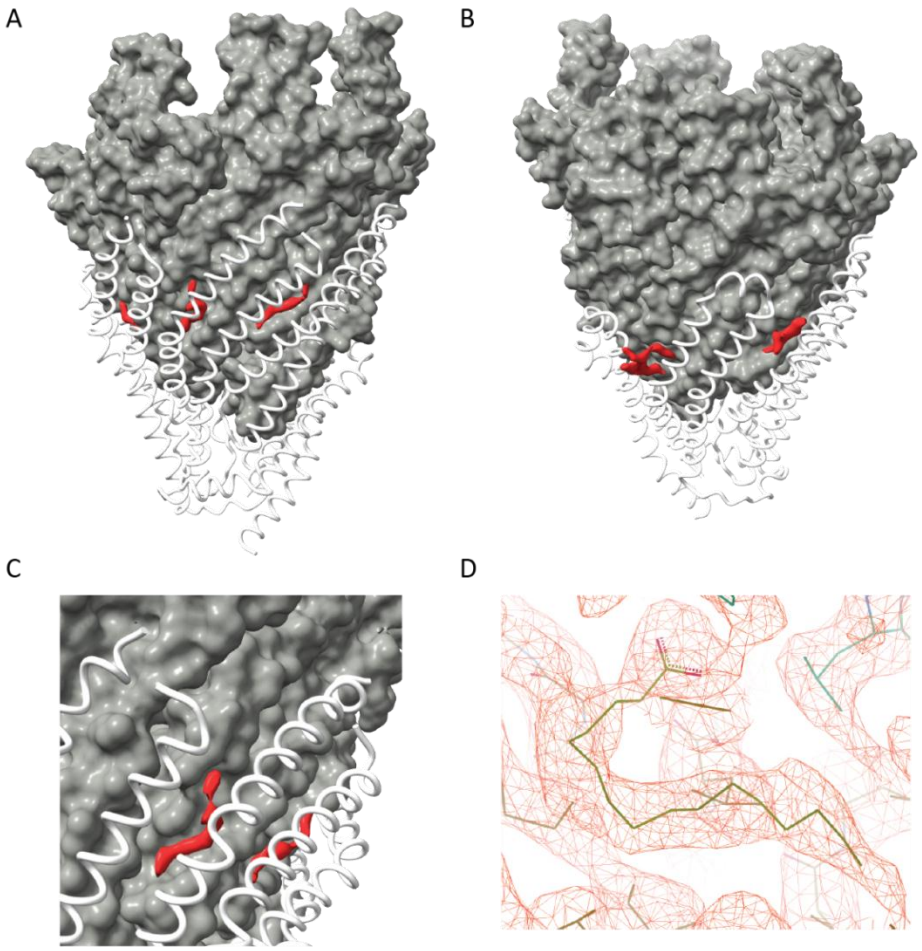


Figure 6.14 Unidentified ligands in the map.
A,B, Density (red) corresponding to ligands that were not modelled in the structure was seen on the outside surface of FliPR and covered by two FliQ subunits each. **C**, Zoomed view on one ligand. **D**, Hypothetical modelling of palmitoic acid in the density of one of the ligands.

6.6 Discussion

In this chapter the structure of the complete export gate, containing FliPQR and the FlhB hydrophobic domain ($FlhB_{TM}$), was described. The complex is in a closed conformation in the structure and highly similar to the previously described structures of FliPQR in the absence of FlhB, suggesting that FlhB is responsible for gating the export gate complex by triggering opening rather than holding it closed. $FlhB_{TM}$ is wrapped around the outside the FliPQR complex, mimicking the shape of FliQ but with an inverted topology. Interestingly, a loop connecting helix 2 and 3 forms a circular structure at the cytoplasmic entrance of the export gate ($FlhB_L$) and mutagenesis shows that the interaction between FliQ and $FlhB_L$ is required for motility and secretion.

6.7 Modelling of an extended FlhB structure

Only residues 29 to 221 of FlhB were modelled. However, additional low resolution density was observed at both termini. Although it was not possible to model the structure in this region, parts of the model predicted using RaptorX (Wang et al. 2017) (Figure 6.6C) could be connected to the experimental model and fit approximately into these densities as rigid bodies (Figure 6.15). This extends the model of FlhB to approximately residue 15 and 237. Despite the absence of information at resolution of single residues, it is clear that both N-terminus and C-terminus are located just underneath $FlhB_L$ and the entrance to the FliPQR-FlhB lumen. This is consistent with the strong contacts derived from evolutionary co-variation between the N-terminus and the N-terminal part of the cytoplasmic domain ($FlhB_{CN}$) (Figure 6.7C (iv)). Recently, a genetic interaction between E11 and E230, in the N-terminus of FlhB and $FlhB_{CN}$ respectively, has been identified in *S.* Typhimurium FlhB (Inoue et al. 2019a) and these two residues would be predicted to be close to each other based on the low-resolution density. The N-terminus of FlhB is also the location of many mutations that restore function when the flagellar ATPase is deleted (Minamino et al. 2003; Minamino and Namba 2008). It is possible that mutations close to the N-terminus of FlhB reduce the energy requirement of opening the FliPQR-FlhB complex by affecting the physically close $FlhB_L$ or helix 4 of $FlhB_{TM}$ (see next section).

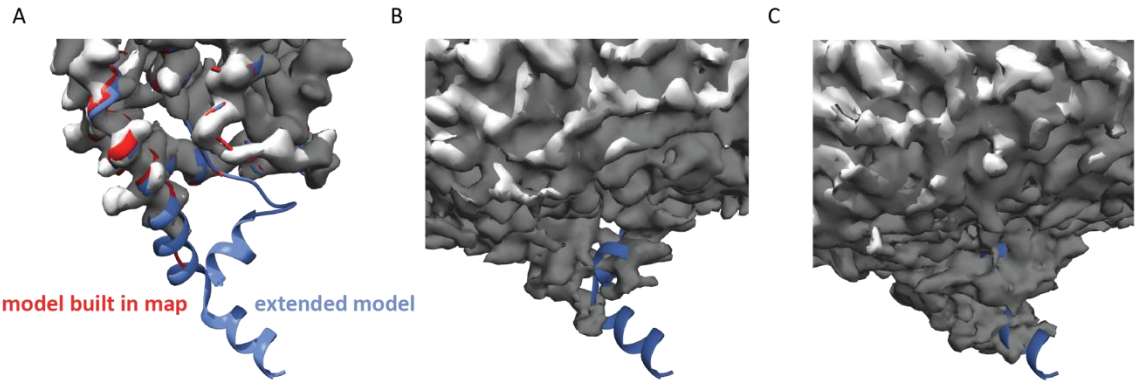


Figure 6.15 Location of the termini of $FlhB_{TM}$.
A,B,C, The model of the FliPQR-FlhB complex (red) and an extended model incorporating N-terminal and C-terminal parts of a predicted model from RaptorX (blue) in the refinement map of the FliPQR-FlhB complex shown at three different contour levels.

6.8 Opening mechanism of the export gate

When the structure of the FliPQR-FlhB complex was fit into a tomographic reconstruction of the injectisome (Figure 6.16A) the previously described $FlhB_C$ crystal structure (Meshcheryakov et al. 2013) could be placed in the map in a variety of positions of which two extremes are shown. There are 7 residues between the last residue of the structure described in this chapter and the crystal structure. No ordered density was observed from around the middle of the linker between $FlhB_{TM}$ and $FlhB_C$ (Figure 6.15), suggesting that this region can be flexible. If this linker is helical and pointing downwards, $FlhB_C$ could contact the cytoplasmic domain of the major export apparatus protein FlhA ($FlhA_C$), which binds secretion substrates (Bange et al. 2010; Xing et al. 2018), while a hinge motion in the linker would allow $FlhB_C$ to reach up to FliPQR. Although it is not known if $FlhB_C$ contacts either $FlhA_C$ or FliPQR directly, it is thought to interact with secretion substrates before they unfold and are funnelled into the export gate but after they interact with $FlhA_C$ (Evans et al. 2013).

In addition to directing substrates towards the export gate through different positions of $FlhB_C$, there must be a mechanism of opening of the FliPQR-FlhB complex in order to allow secretion. If opening of the export gate follows handover of substrates from $FlhA_C$ to $FlhB_C$, it is conceivable that $FlhA_C$ exerts a pulling force on $FlhB_C$ or the substrate bound to $FlhB_C$. Alternatively, the transmembrane domain of FlhA, which is a proton channel (Minamino et al. 2016), may apply force to the FliPQR-FlhB complex or both domains may work in concert. Whichever domain is the source of such a pulling force on FlhB, this force could then be propagated through the complex via the network of conserved charged residues that are buried between $FlhB_{TM}$ and FliPQR, which could then lead to conformational changes in the tightly connected $FlhB_L$ and the FliQ subunits (Figures 6.12 and 6.13), leading to opening of the complex. In particular, residue D208 connects helix 4 to helices 1 and 3 and a D208A mutation caused loss of motility in *Salmonella*, but motility was restored upon overexpression of FlhA (Erhardt et al. 2017), strengthening the link between opening of FliPQR-FlhB and FlhA. This model can also explain the requirement of the $FlhB_L$-FliQ interaction for motility and secretion (Figure 6.12), as when the interaction is disrupted, the conformational changes may not be fully transmitted from $FlhB_L$ to FliQ. In this model the mutations close to the N-terminus of

FlhB which suppress deletion of the ATPase (Minamino et al. 2003; Minamino and Namba 2008) may affect the dynamics of the pulling force as the N-terminal residues are closely associated with $FlhB_{CN}$ (Figure 6.15).

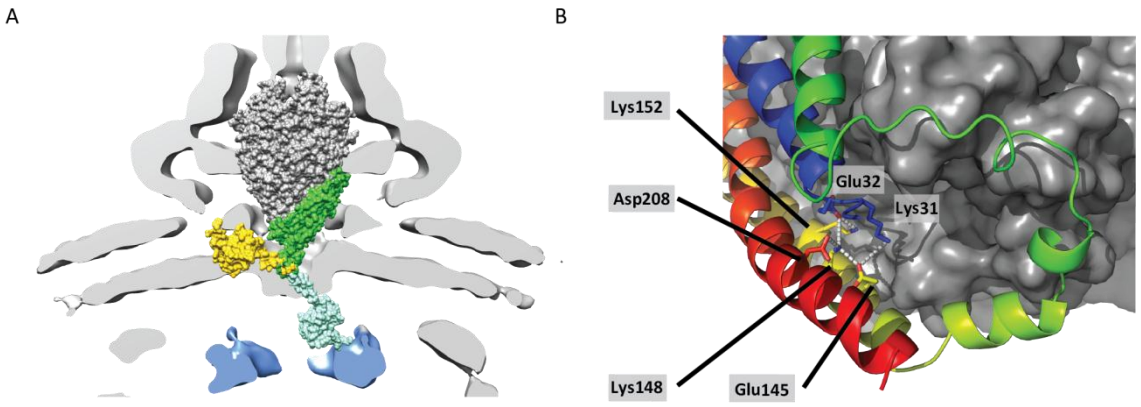


Figure 6.16 Potential mechanism of secretion through the export gate.
A, Position of $FlhB_C$ within the assembled *Salmonella* SPI-1 vT3SS (EMD-8544) (Hu et al. 2017). The structures of FliPQR (grey surface), the FlhB transmembrane domain (green) and the FlhB cytoplasmic domain in two different positions (yellow and cyan) were fit into the tomogram of the vT3SS. The density corresponding to the SctV cytoplasmic domain is coloured in blue. **B**, Network of buried charged residues in FlhB.

Although the results from this chapter suggests that $FlhB_{TM}$ is crucial in the gating of the export gate, $FlhB_C$ is most well known to be responsible for substrate specificity switching, which also requires proteolytic cleavage (Fraser et al. 2003; Zarivach et al. 2008; Evans et al. 2013; Monjaras Feria et al. 2015). The resulting C-terminal fragment has been reported to be secreted in *Yersinia* (Frost et al. 2012) but is encoded by a separate gene as a "spare part" in some species (Wand et al. 2006) and it was not known if FlhB plays any role in secretion after switching. The function of $FlhB_{TM}$ in gating of secretion studied here implies that FlhB is required for opening of the export gate even after switching. Future studies will be required to test if switching only requires a change in affinity of $FlhB_C$ to substrates in the cytoplasm or if $FlhB_{TM}$ is also involved.

7 Structural studies of the major export apparatus protein, FlhA

7.1 Introduction

The export apparatus of type three secretion is made up of five integral membrane proteins, FliPQR FlhBA in flagella and SctRSTUV in injectisomes. In chapters 3 to 6 the structure determination of the export gate FliPQR/SctRST and the membrane domain of FlhB has been described. The structures of the cytoplasmic domains of FlhB/SctU (Deane et al. 2008; Zarivach et al. 2008; Lountos et al. 2009; Meshcheryakov et al. 2013) and the nonameric FlhA/SctV ($FlhA_C$/$SctV_C$) (Bange et al. 2010; Saijo-Hamano et al. 2010; Worrall et al. 2010; Abrusci et al. 2013) have long been known and their interaction with secretion substrates is crucial for subsequent funnelling of the substrates through the export gate in an unfolded state (Evans et al. 2013; Xing et al. 2018). All structures of the export gate were in the closed state and FlhA/SctV has been hypothesised to be responsible for the opening of the complex (chapter 6). This may involve the structurally uncharacterised membrane domain of FlhA/SctV, which is physically close to FliPQR/SctRST in the assembled T3SSS nanomachine (chapters 4 and 6). The membrane domain is thought to be made up of around nine times eight transmembrane helices and contains a short stretch of soluble residues that are thought to form a soluble domain called the FHIPEP domain (McMurry et al. 2004) which has been shown to be important in protein export (Barker et al. 2016). Furthermore, FlhA has been shown to be a proton and sodium channel (Minamino et al. 2016) and conserved charged residues in FlhA are essential for motility (Erhardt et al. 2017). This potentially makes the structure of the membrane domain of FlhA the key to understanding the mechanisms of export gate opening, transfer of substrates from the cytoplasmic sorting platform (Lara-Tejero et al. 2011) into the secretion channel and coupling the pmf to type 3 secretion.

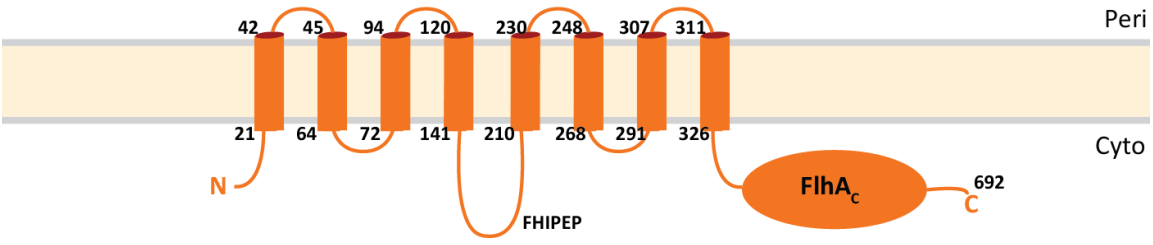


Figure 7.1 Predicted topology of FlhA.
Topology of FlhA following (Erhardt et al. 2017).

7.2 Screening of constructs for purification of assembled FlhA rings

7.2.1 Initial attempts of producing full-length FlhA oligomers

Previous experiments in the lab, carried out by Dr. Patrizia Abrusci, demonstrated the instability of full-length *S. flexneri* SctV oligomers. Recently, it has been demonstrated that when *S.* Typhimurium FlhA is recombinantly expressed, the subunits can be cross-linked with low efficiency using disulphide bonds in the FHIPEP domain (Erhardt et al. 2017), providing a possible mechanism through which the assembled nonameric protein might be stabilised. Therefore, these two cysteine mutations (D170C, R185C) were introduced in an expression plasmid of *S.* Typhimurium FlhA with a C-terminal strep tag in a pET vector. As the reported cross-linking efficiency was low, two further cysteine mutations in the cytoplasmic domain were designed using the previously published structures of $FlhA_C$ (Saijo-Hamano et al. 2010) and $SctV_C$ (Abrusci et al. 2013) and the server Disulfide by Design (Craig and Dombkowski 2013). When the $SctV_C$ structure (three chains) was submitted to the server, the residue pair Q501/Y514 was identified, but no evolutionary co-variation interaction was identified between them using the GREMLIN server (Ovchinnikov et al. 2014). Therefore, three $FlhA_C$ molecules were aligned with the nonameric $SctV_C$ structure and this trimer was submitted to the server, which identified the residue pair K423/D553, which also shows a co-variational interaction using GREMLIN.

The flhA gene (D170C, R185C, K423C, D553C) was cloned into a pET28 vector with a C-terminal twin strep tag. The plasmid was introduced into SHuffle cells (NEB), in which the cytoplasm is an oxidising environment, enabling the formation of cytoplasmic disulphide bonds (Lobstein et al. 2012). At the same time, the protein was also made in Mt56 cells (Baumgarten et al. 2017), a strain of BL21 in which the T7 polymerase is mutated resulting in slower transcription. Cells were grown at 37 °C in TB media with the appropriate antibiotics supplemented with 0.1 mM IPTG overnight and purified using the same protocol as was used for FliPQR (section 2.5.3) using the detergent LMNG. Due to the large number of tags expected on a nonameric complex, the elution from the StrepTrap column was carried out in two steps. First, protein was eluted using buffer containing 1 mM desthiobiotin. A second elution was carried out using buffer containing

10 mM desthiobiotin. It was expected that larger, fully assembled complexes would only elute at the higher concentration of desthiobiotin.

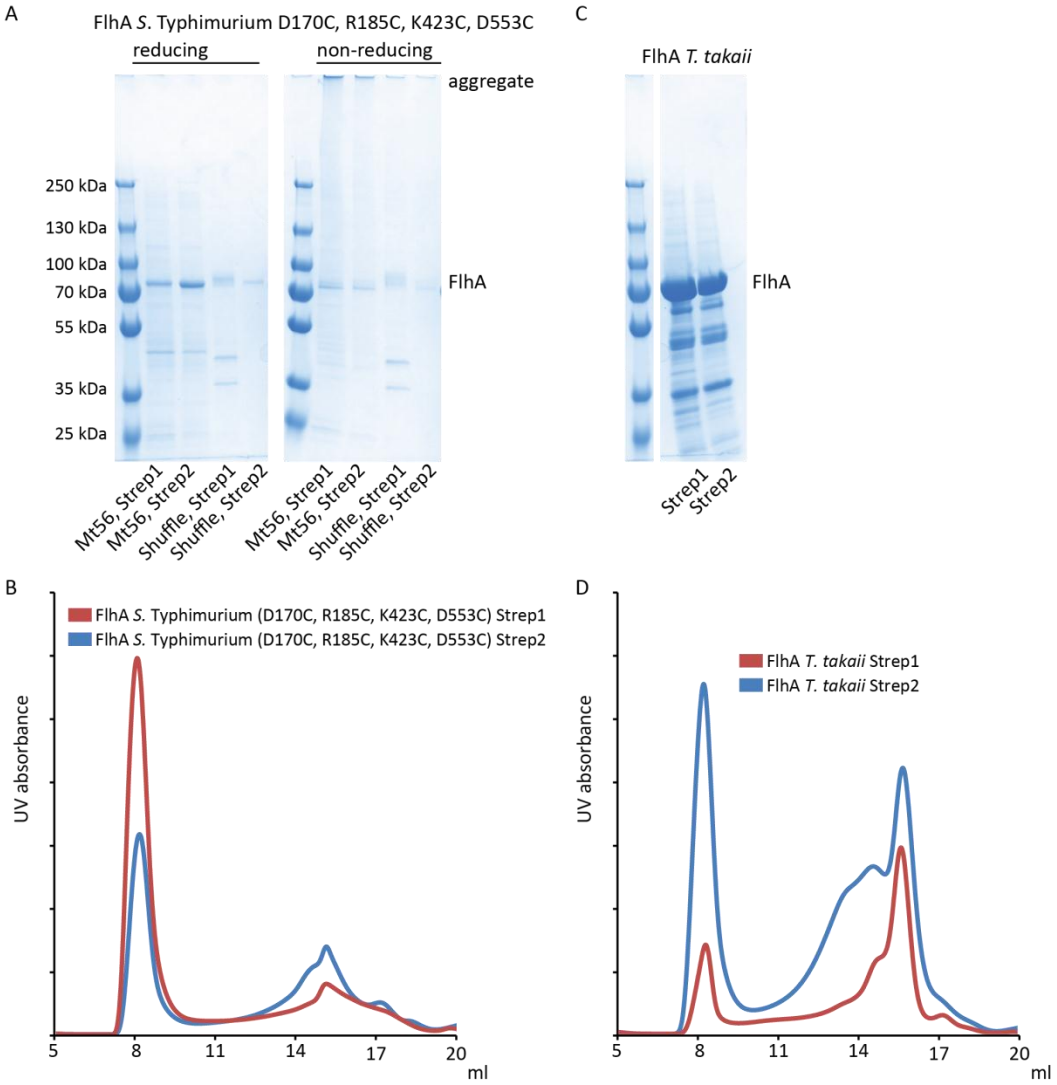


Figure 7.2 Purification of full-length FlhA.
A, SDS-PAGE analysis of the 1 mM desthiobiotin (Strep1) and 10 mM desthiobiotin (Strep2) Strep eluates of FlhA (*Salmonella*) produced in Mt56 cells or Shuffle cells. A band just above the 70 kDa marker was observed, consistent with FlhA. **B**, Analytical SEC of the Strep1 and Strep2 eluate of the FlhA mutant made in Mt56 cells using a Superose 6 increase 10/300 column (GE). **C**, SDS-PAGE analysis of the 1 mM and 10 mM desthiobiotin Strep eluates of FlhA (*T. takaii*). **D**, Analytical SEC of the two strep eluates from FlhA (*T. takaii*) using a Superose 6 increase 10/300 column (GE).

When the purified sample was analysed by SDS-PAGE, no ladder of cross-linked proteins was observed (Figure 7.2A, B). Under non-reducing conditions a band that did not enter the gel could be observed. This likely corresponds to highly cross-linked, misfolded

protein rather than assembled FlhA rings as when the samples were subjected to SEC using a Superose 6 increase 10/300 column (GE) only a void peak and a second peak eluting very late, around 15 ml, could be observed. The 260 kDa *V. mimicus* FliPQR-FlhB complex eluted at around 14 ml on the same column (chapter 6) while the molecular weight of a nonamer of FlhA, whose membrane domain surrounds FliPQR-FlhB, would be expected to be 670 kDa.

Another possibility for stabilisation of the full-length protein was to use the homologous protein from a thermophile bacterium. FlhA from the thermophile *Thermosulfidibacter takaii* (strain DSM 17441) was cloned into the same pET vector with a C-terminal twin strep tag and the plasmid was co-transformed with the plasmid pRARE2 (Novagen), a plasmid containing tRNAs uncommon in *E. coli*. The protein was expressed and purified in the same way as the *Salmonella* homologue, except that the culture was supplemented with chloramphenicol (30 µg/ml) to maintain pRARE2. Although large quantities of the protein could be purified (Figure 7.2C), the protein eluted from the SEC column as a mixture of species and the largest SEC peak was consistent with a monomeric protein. Interestingly, there appeared to be a larger proportion of oligomeric protein in the sample derived from the second elution of the strep column, using the higher concentration of desthiobiotin for elution (Figure 7.2D), although the void peak was also larger in this sample. This suggested that it might be possible to further optimise the purification of the FlhA complex from *T. takaii* to obtain the fully assembled nonamer, however, further stabilisation attempts using disulphide bonds (D162C, R178C which are the equivalent of the mutations D170C, R185C described in *Salmonella* (Erhardt et al. 2017)) or halving the amount of detergent used for extraction from the membrane were not successful in producing the assembled complex according to SEC. Purification of a stable oligomeric complex of FlhA was also unsuccessful using *Salmonella* FlhA with just two cysteine substitutions (D170C, R185C) or using the detergents thesit or GDN (Anatrace) (data not shown).

7.2.2 Stable FlhA rings fused to GFP

As initial attempts of producing the assembled FlhA complex were not successful and no disulphide bonds had formed, it was not certain if assembly of the complex in the

membrane had occurred. In order to assess assembly *in vivo*, *S.* Typhimurium FlhA was fused to a C-terminal GFP followed by a tev site and a twin strep tag in the pT12 vector. Observation of fluorescent spots has previously been used to show that fusion proteins of FlhA/SctV and a fluorescent protein form oligomeric assemblies in the membrane (Diepold et al. 2011; Li and Sourjik 2011; Morimoto et al. 2014).

8 µl of an overnight culture of BL21 cells expressing the GFP fusion construct were applied to a glass slide, covered with a coverslip and imaged using a Zeiss 880 inverted fluorescence microscope in which GFP fluorescence was triggered by excitation at 488 nm (section 2.4.2). Expression was induced using 0.01% (w/v) rhamnose monohydrate. As fluorescent spots localising to the cell envelope were observed (Figure 7.3A), this construct was taken forward for purification of protein.

S. Typhimurium FlhA-GFP in the pT12 vector was expressed and purified using the same protocol as was used for FliPQR (section 2.5.3) using the detergent LMNG. During the affinity purification step the StrepTrap column (GE) turned green. The total yield after elution from the StrepTrap column was about 11 mg/litre of culture (Figure 7.3B). For comparison, the *Salmonella* FlhA without GFP expressed from the pET vector could be purified with a yield of around 0.4 mg/litre and the *Thermosulfidibacter* FlhA without GFP had been purified with a yield of around 3 mg/litre of culture. The eluted protein was concentrated and subjected to SEC using a Superose 6 increase 10/300 column (GE). The protein eluted in a peak consistent with a large assembly with a long tail (Figure 7.3C). In order to test the stability of the large complex, 500 µl of the peak fraction were re-injected into the SEC column. The elution volume of this sample was shifted to the right (Figure 7.3C), suggesting that the assembly is not stable.

In order to further explore assembly of the nonameric FlhA complex in this fusion protein, SEC-MALS was carried out using 100 µl of purified protein at increasing concentrations (section 2.5.8). This showed that assembly was concentration dependent (Figure 7.3D). Due to the limited dynamic range of the light scattering detector it was not possible to test higher concentrations of sample, but the elution volume of the large

amount of highly concentrated protein taken from the StrepTrap eluate (500 µl at more than 10 mg/ml) is likely to correspond to the intact complex, eluting at 12 ml.

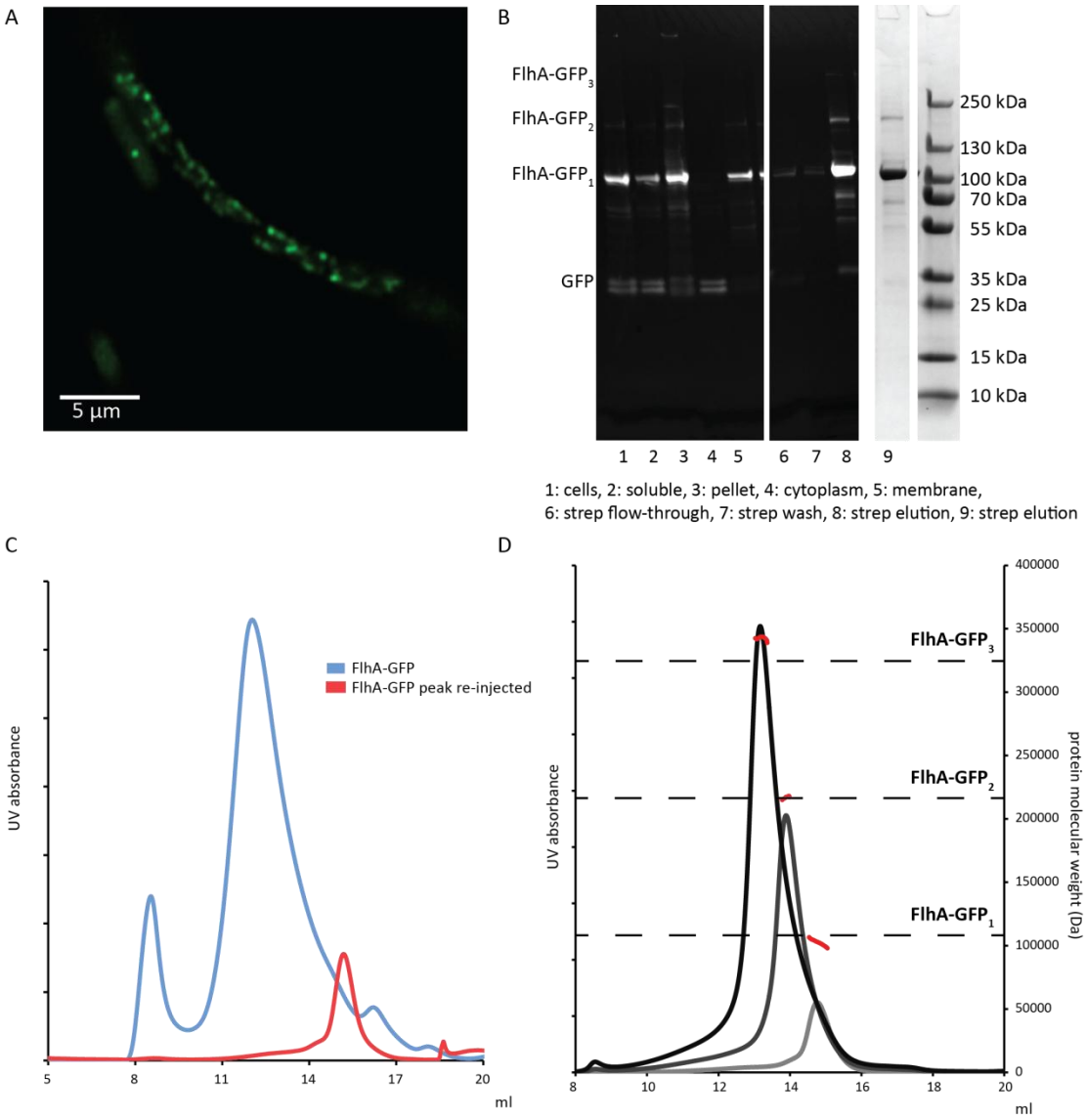


Figure 7.3 FlhA fusion to GFP.
A, Fluorescence image of BL21 cells expressing FlhA-GFP. **B**, Purification of FlhA-GFP. The gel was imaged using blue light (left). Lane 8 is also shown after the gel was stained (labelled lane 9, right). **C**, SEC of purified FlhA-GFP (blue). 500 µl of the 12 ml peak were re-injected into the column (red). The values in the curve shown were amplified ten-fold for comparison. **D**, SEC-MALS of FlhA-GFP at increasing concentrations (1.9, 7.6 and 15.3 mg/ml).

LMNG was chosen for the first purification of FlhA-GFP as this detergent is comparatively gentle (Chae et al. 2010) and had been used successfully for the structure determination of FliPQR and FlhB (chapters 3-6). LMNG was also successfully used for

the preparation of high-quality cryo-EM grids in other projects in the lab, making LMNG the detergent of choice for cryo-EM. However, FlhA-GFP was insufficiently stabilised by LMNG and it is not clear if the membrane domain is correctly folded if the complex is re-assembled from monomers following concentration. Therefore, different detergents were screened for the preparation of stable FlhA-GFP rings.

In order to test a variety of stabilising detergents, four times 10 ml of 10% (w/v) membranes from cells expressing FlhA-GFP were prepared as above. Different detergents were added to a final concentration of 1% (w/v) and the membranes were solubilised for one hour. Four detergents that have been reported to be more stabilising than the common detergent DDM were chosen for this experiment: GDN (glyco-diosgenin) (Chae et al. 2012), MNA-C12 (mannitol based amphiphile) (Hussain et al. 2016), DMNG (Chae et al. 2010) and Cy6NG (cymal-6 neopentyl glycol, Anatrace). The clarified solubilised membranes were applied to 5 ml StrepTrap columns (GE), which were washed with HBS containing 0.01% (w/v) detergent and eluted in the same buffer containing 10 mM desthiobiotin. All samples were immediately concentrated using a 1,000 kDa MWCO concentrator (GE) (the nonameric assembly of the fusion protein has a molecular weight of 966 kDa without considering bound detergent molecules and there is a large hole in the centre of the cytoplasmic domain (Abrusci et al. 2013)) and negative stain grids were prepared (section 2.5.11) using the samples retained by the concentrators after dilution in HBS containing 0.01% (w/v) detergent to 0.03 mg/ml. In addition, grids were made using a sample of the previously purified protein in the detergent LMNG diluted to 0.03 mg/ml in buffer containing detergent and in detergent free buffer. The aliquot of this sample had been concentrated to 15 mg/ml and flash frozen. In the absence of other detergents LMNG is thought to dissociate from a membrane protein very slowly (Chung et al. 2012) and LMNG micelles can be removed from the buffer using a detergent free glycerol gradient (Hauer et al. 2015) while keeping the membrane protein stable. If the 15 mg/ml sample contains intact rings, as suggested by SEC (Figure 7.3), and dissociation of the rings at low concentration is due to detergent, it may be possible to stabilise the assembly by diluting without detergent to avoid a potential destabilisation of the soluble domain by detergent (Yang et al. 2014).

Negative stain EM (section 2.5.11) of the FlhA-GFP samples revealed unclear objects in samples purified in MNA-C12, DMNG, Cymal6-NG or LMNG and intact particles in the GDN sample (Figure 7.4A). In addition, large amorphous objects were observed in the GDN sample. It is possible that these are patches of membrane containing FlhA molecules that were incompletely solubilised as it may not be possible to fully solubilise the *E. coli* membrane using a detergent as gentle as GDN. Interestingly, the highest number of intact particles was observed in the sample purified in LMNG but diluted in detergent free buffer.

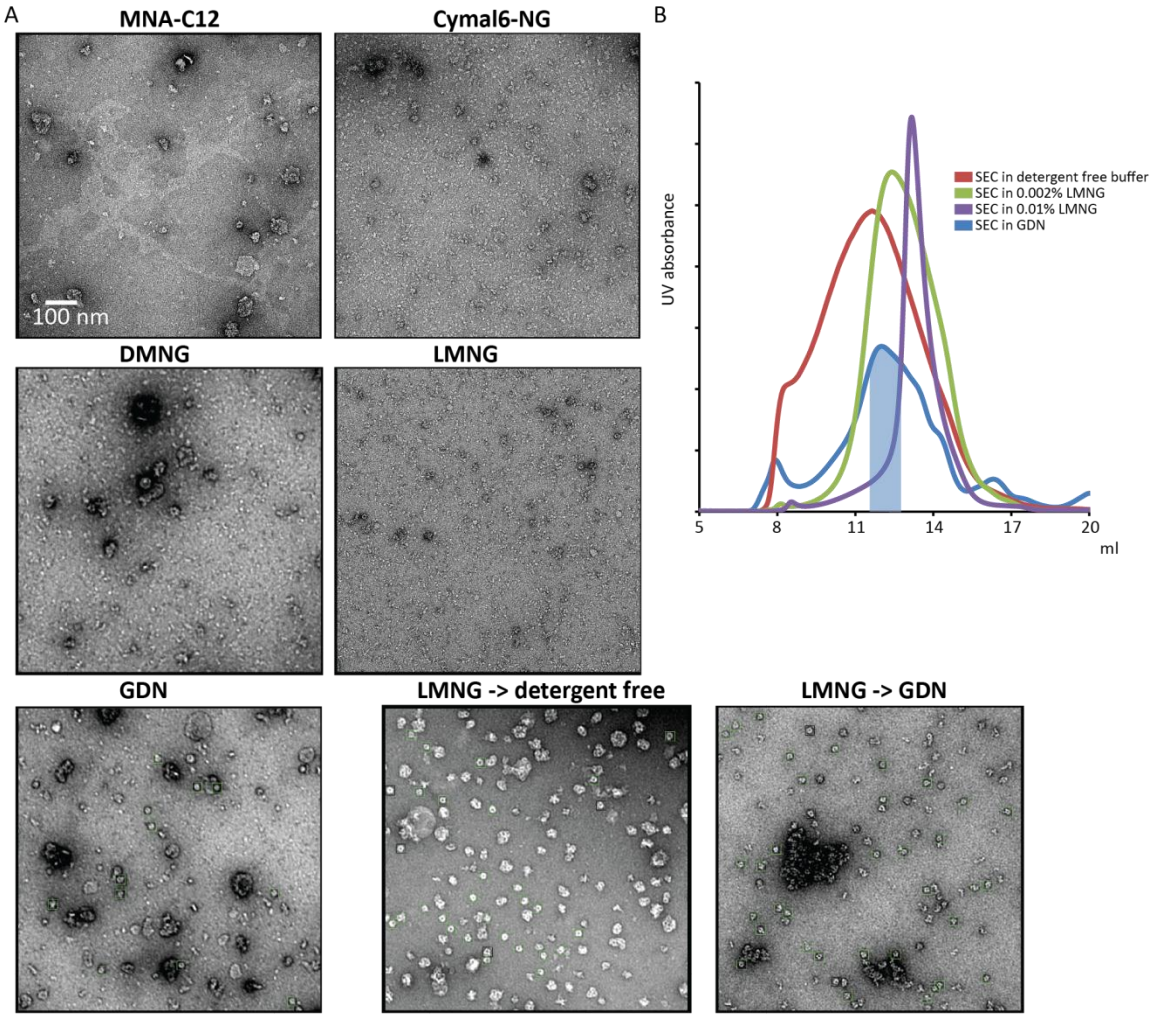


Figure 7.4 Stability of FlhA depends on detergent.
A, Negative stain EM micrographs of FlhA-GFP purified in the indicated detergent or purified in LMNG and diluted in detergent free buffer (bottom middle) or purified in LMNG and subjected to SEC in buffer containing GDN (bottom right). Intact particles are boxed (green). **B**, Analytical SEC of FlhA purified in 0.01% (w/v) LMNG using running buffer containing the indicated detergent. The fraction taken for grid preparation following SEC in GDN is shaded in blue.

In order to further analyse how the intact complex in LMNG at high protein concentrations can be stabilised, analytical SEC was carried out with 100 µl of 15 mg/ml FlhA in LMNG using a Superose 6 increase 10/300 column (GE) and TBS buffer containing no detergent, 0.002% (w/v) LMNG (2 times CMC), 0.01% (w/v) LMNG and 0.01% (w/v) GDN (using protein concentrated to 6 mg/ml) (Figure 7.4B). Interestingly, the protein eluted around 12 ml in GDN and in detergent free buffer, although the peaks were very broad, while addition of LMNG caused a dose dependent shift of the peak to the right. When a sample from the 12 ml peak in GDN was analysed by negative stain EM with grids prepared as above, intact rings could be observed (Figure 7.4A).

7.3 Cryo-EM analysis of *Salmonella* FlhA

As a very large oligomer of FlhA-GFP could be stabilised by SEC in the absence of detergent, the FlhA-GFP purified in LMNG, which had been frozen in aliquots at 15 mg/ml, was subjected to SEC using TBS without detergent or TBS containing 0.01% (w/v) GDN. Cryo-EM grids were made from fractions of the previously identified peak at 12 ml (section 2.5.12) using 300 mesh holey carbon copper grids (Quantifoil). 3 µl of sample were applied to the grid, blotted after 10 seconds at 100% humidity at 22 °C and frozen in liquid ethane using a Vitrobot Mark IV (FEI). In initial screening the ice in all grids made with the GDN sample was too thick to collect data. However, the ice on a grid of FlhA-GFP at 1.9 mg/ml in the absence of detergent in the buffer was of sufficient quality for imaging. 1,329 micrographs were acquired using a Titan Krios microscope (FEI) equipped with a K2 direct electron detector (Gatan). Motion correction and particle picking was carried out in SIMPLE (Reboul et al. 2018). The quality of the particles was poor and the majority of particles had dissociated (Figure 7.5A). Around 180,000 candidate particles were extracted and 2D class averaging was performed in RELION-3.0 (Zivanov et al. 2018). After 2D classification, top and side views were observed (Figure 7.5B). However, little detail could be resolved when a 3D volume was calculated from the data using RELION-3.0 (Figure 7.5C) using C9 symmetry. One likely cause of this is the small number of undamaged particles – fewer than 10,000 particles were selected after 2D classification as the majority of particles were damaged or aggregated. One feature that could be seen in 2D averages was a line on the periplasmic side of the membrane domain. It is not clear what part of the protein this feature corresponds to.

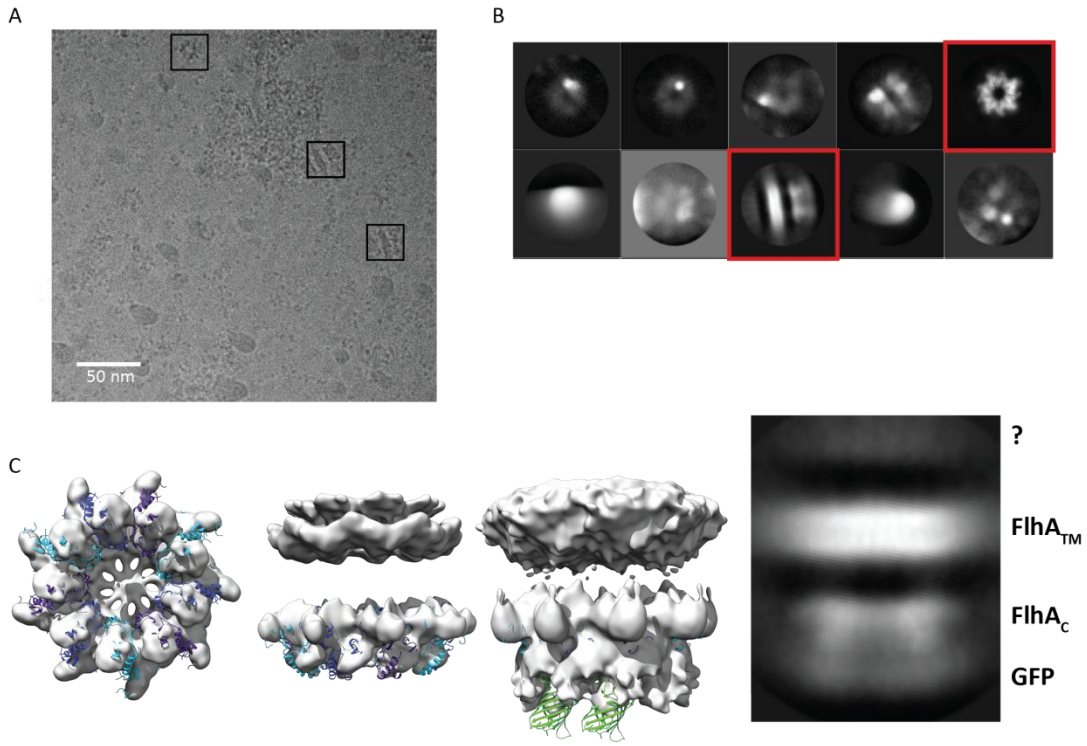


Figure 7.5 Cryo-EM of *Salmonella* FlhA.
A, Representative micrograph of *Salmonella* FlhA-GFP. Defocus = -3.4 µm. Three particles can be seen (boxed). **B**, 2D class averages. A top and a side view are highlighted. **C**, Low resolution 3D volume in which the structure of *S. flexneri* SctV (Abrusci et al. 2013) was fit (blue). At a lower contour level (right) density corresponding to GFP can be seen. Two GFP molecules were fit into the map (1GFL). Right: 2D class average of a side view of the complex annotated to show the position of the membrane domain (FlhA$_{TM}$), the cytoplasmic domain (FlhA$_C$), the GFP and a feature of unknown origin on the periplasmic side of the membrane domain.

As only a low resolution volume could be calculated from the data, the protein was cross-linked with the aim of stabilising the structure so as to avoid damage in the grid preparation process. Purified FlhA-GFP was subjected to SEC in HBS containing 0.01% (w/v) LMNG, the fractions containing protein were pooled and concentrated to 5.9 mg/ml. 400 µl of this sample were mixed with 2 µl of 25% (v/v) glutaraldehyde and incubated on ice for 30 minutes. 50 µl of 1 M Tris, pH 8 were added in order to quench remaining glutaraldehyde and aggregates were removed by centrifugation (17,000x g, 10 minutes). Three such reactions were set up and pooled after cross-linking. Intact nonameric complexes were isolated from the cross-linked sample through two rounds of SEC using TBS containing 0.01% (w/v) LMNG. The peak fractions eluting around 12 ml after the first SEC run were pooled, concentrated to 500 µl, and subjected to SEC again (Figure 7.6A, B). The fractions eluting around 12 ml were pooled and the protein was

concentrated. Grids were made in the same way as for the previous sample at a range of concentrations. A grid made with sample concentrated to 4.8 mg/ml was selected for data collection due its ice thickness and particle density.

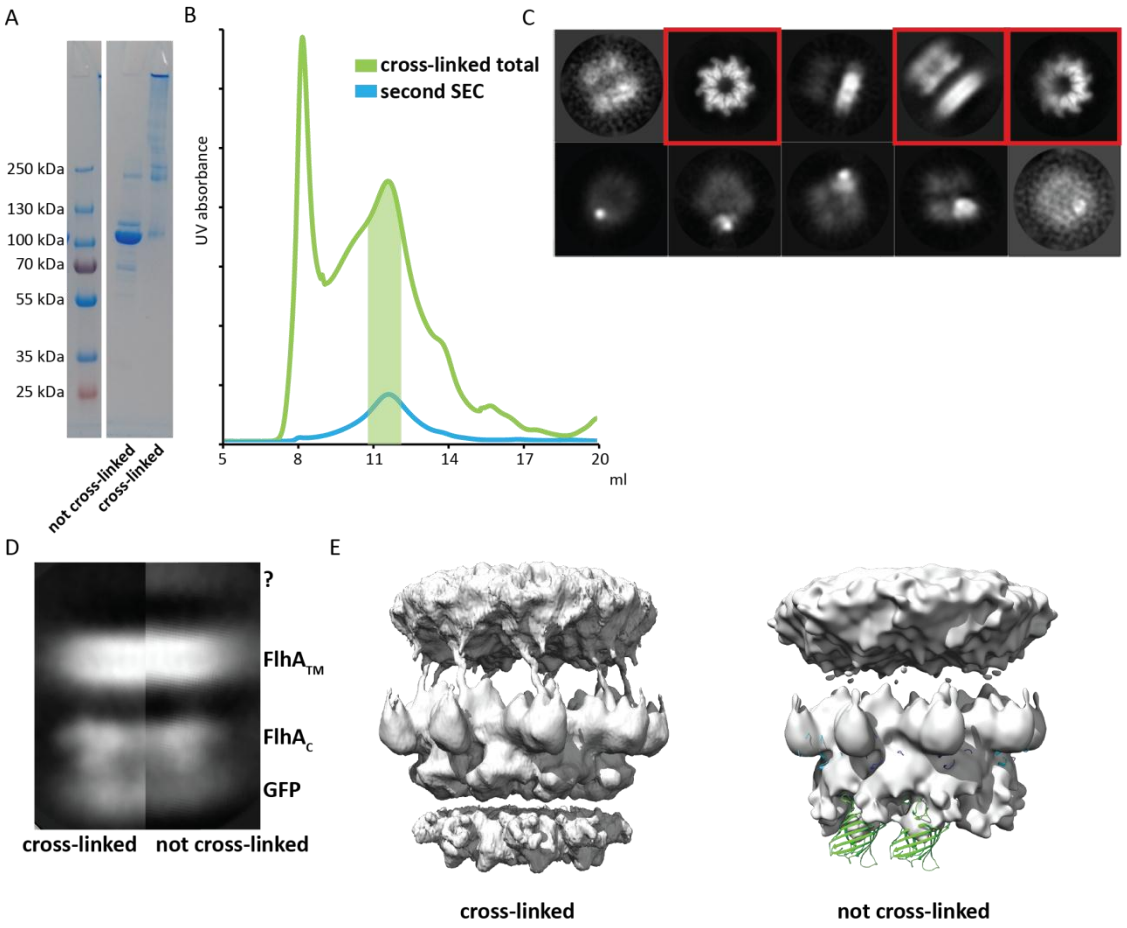


Figure 7.6 Cryo-EM analysis of cross-linked *S.* Typhimurium FlhA-GFP.
A, 4-20% SDS-PAGE (Bio-Rad) gel of FlhA before and after cross-linking. **B**, SEC profile of the cross-linked protein (green) and SEC profile after pooling the peak fractions (green shading) and repeating SEC. The column was a Superose 6 increase 10/300 column (GE). **C**, 2D class averages of the cross-linked sample. **D**, Comparison of the side views of FlhA imaged with and without cross-linking. Note the change in the density ascribed to GFP (bottom) and the density of unknown origin (top). **E**, Comparison of 3D volumes calculated from the data in RELION-3.0 (Zivanov et al. 2018).

The micrographs were processed as was done for the sample without cross-linking. 94,208 candidate particles were extracted from 1,328 micrographs. 2D class averaging in RELION-3.0 (Zivanov et al. 2018) revealed top and side views, including damaged complexes (Figure 7.6C). Unexpectedly, the feature on the periplasmic side of the membrane domain had disappeared (Figure 7.6D) and the density corresponding to GFP

had become larger and less ordered (Figure 7.6E) compared to the sample that had not been exposed to glutaraldehyde.

7.4 Screening of FlhA homologues

As these attempts at optimising the *Salmonella* FlhA ($FlhA_{ST}$) grids were not successful the decision was made to screen a number of homologous flhA/sctV genes for the production of a more stable complex. $FlhA_{ST}$ is very unstable and it may be difficult to completely optimise this sample for structure determination. However, the expression system using the GFP fusion was successful at producing a large quantity of protein. A screen was designed for small-scale purification of 13 different homologous FlhA or SctV complexes in the detergent LMNG followed by SEC. The presence of a sharp peak, likely around 12 ml on a Superose 6 increase 10/300 column (GE), would indicate a stable nonameric assembly that is resistant to dissociation by LMNG which was hoped would correlate with stability in the thin film of water on cryo-EM grids. This screen and the section 7.5 of this chapter were carried out in collaboration with an undergraduate project student, Mr. Jerry Cao.

In order to cover a wide range of sequence space, the flhA genes from *Chlamydia trachomatis* ($flhA_{Ct}$), *Thermosulfidibacter takaii* ($flhA_{Tt}$), *Myxococcus xanthus* ($flhA_{Mx}$) the *Vibrio mimicus* polar flagellum ($flhA_{Vm}$) and the *Vibrio parahaemolyticus* lateral flagellum ($flhA_{Vp}$) and the sctV genes from *Yersinia enterocolitica* ($sctV_{Ye}$), *Shigella flexneri* ($sctV_{Sf}$), *Pseudomonas savastanoi* ($sctV_{Ps}$), *Chlamydia trachomatis* ($SctV_{Ct}$), *Myxococcus xanthus* ($sctV_{Mx}$) and the *Salmonella* Typhimurium SPI-1 ($sctV_{ST1}$) and SPI-2 ($sctV_{ST2}$) systems were cloned as GFP fusions into the pT12 vector in the same way as $FlhA_{ST}$. A short linker without a cleavage site between GFP and FlhA was chosen in order to avoid potential proteolytic activity of cellular proteases. All 13 plasmids were transformed into BL21 cells for expression. $FlhA_{Tt}$, $FlhA_{Mx}$, $SctV_{Ct}$ and $SctV_{Mx}$ are derived from organisms only distantly related to *E. coli* and in order to avoid reduced expression due to codon usage differences the corresponding plasmids were transformed into Rosetta2 cells which carry the plasmid pRARE2 (Novagen), which contains the genes of the 7 least common tRNAs in *E. coli*.

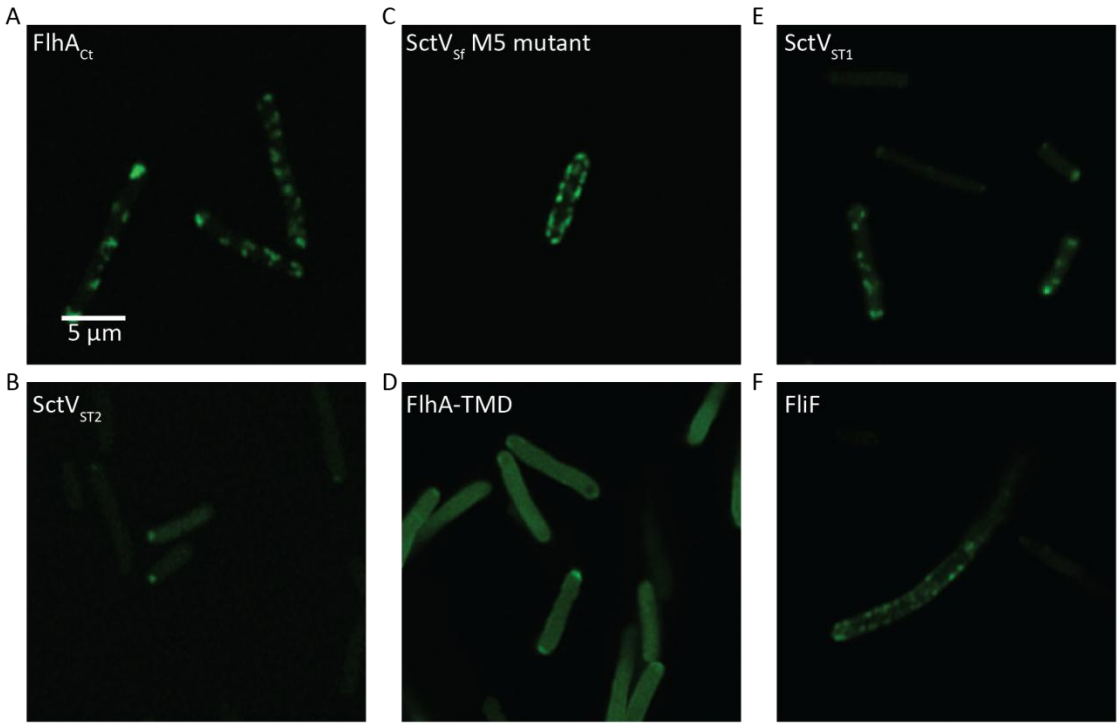


Figure 7.7 Live cell imaging of BL21 cells expressing FlhA/SctV-GFP fusion proteins. Fluorescence micrographs of BL21 cells expressing GFP fusions of FlhA$_{Ct}$ (**A**), SctV$_{ST2}$ (**B**), SctV$_{Sf}$ M5 mutant (**C**), FlhA$_{ST}$ transmembrane domain (**D**), SctV$_{ST1}$ (**E**) and FliF (**F**).

The first step of the screen was to assay expression and complex assembly in the cells using fluorescence microscopy. 10 ml overnight cultures of all strains were grown using 2YT media complemented with 0.01% (w/v) rhamnose monohydrate, kanamycin (30 µg/ml) and chloramphenicol (34 µg/ml) for Rosetta2 cells. 8 µl of the culture were applied to a glass slide and GFP fluorescence was imaged in a fluorescence microscope (Zeiss) (section 2.4.2). In the majority of cases, spots localising to the cell envelope were observed (Figure 7.7A and Table 7.1), but in two cases there were only large spots at the cell poles or at the centre of elongated cells (Figure 7.7B). Inclusion bodies in BL21 are known to localise to the cell poles (Lee et al. 2008; Lindner et al. 2008). These two sequences were excluded from the next step of the screen. As a negative control, two further GFP fusion proteins were imaged. Ring formation has been proposed to be disrupted in the M5 mutant of SctV$_{Sf}$, in which five alanine mutations delete all intermolecular salt bridges (Abrusci et al. 2013), and deletion of the cytoplasmic domain is thought to prevent assembly (Diepold et al. 2011). Therefore, the M5 mutant and the membrane domain of FlhA$_{ST}$ were cloned into the pT12 vector with a C-terminal GFP tag. Interestingly, cells expressing the M5 mutant were indistinguishable from the wild type

(Figure 7.7C). Cells expressing only the membrane domain fused to GFP exhibited uniform green colour and inclusion bodies (Figure 7.7D). The successful complex formation in the M5 mutant may be due to stabilisation of the interactions between protomers by the GFP fusion. An unexpected distribution of cells was observed in some homologues: while some cells exhibited multiple spots corresponding to complexes in the membrane, others only contained inclusion bodies (Figure 7.7E). As a positive control, an *S.* Typhimurium FliF-GFP fusion was expressed from the pT12 vector in the same way as the FlhA fusion proteins. Formation of fluorescent spots by FliF fused to fluorescent proteins has been reported previously (Li and Sourjik 2011; Morimoto et al. 2014). These cells looked very similar to the cells expressing FlhA (Figure 7.7F). In all cases the cells were about 1 µm in width but many µm long. It is possible that expression of the membrane protein affects cell morphology. Elongated cell morphology in stationary phase has been reported in mutants of the sigma factor RpoS which governs stationary phase (Lange and Hengge-Aronis 1991). It is not known whether the rpoS gene is active in the lab strain of BL21 cells.

Next, the 11 FlhA/SctV proteins for which fluorescent spots were observed were purified on a small scale. The proteins were expressed and purified as described above (section 2.5.3) using the detergent LMNG. Large amounts of protein could be purified from most homologues except $SctV_{Ps}$, in which case the StrepTrap column did not turn green but the flow through was green under blue light. The eluates were concentrated to 10 mg/ml and SEC was performed with 100 µl of the concentrated sample as well as 100 µl of a 1 mg/ml dilution of the concentrated sample in order to test if assembly is concentration dependent. For testing whether the more gentle detergent GDN was required for stability, SEC was performed both using buffer containing LMNG and GDN (0.01%) using samples at both 1 and 10 mg/ml.

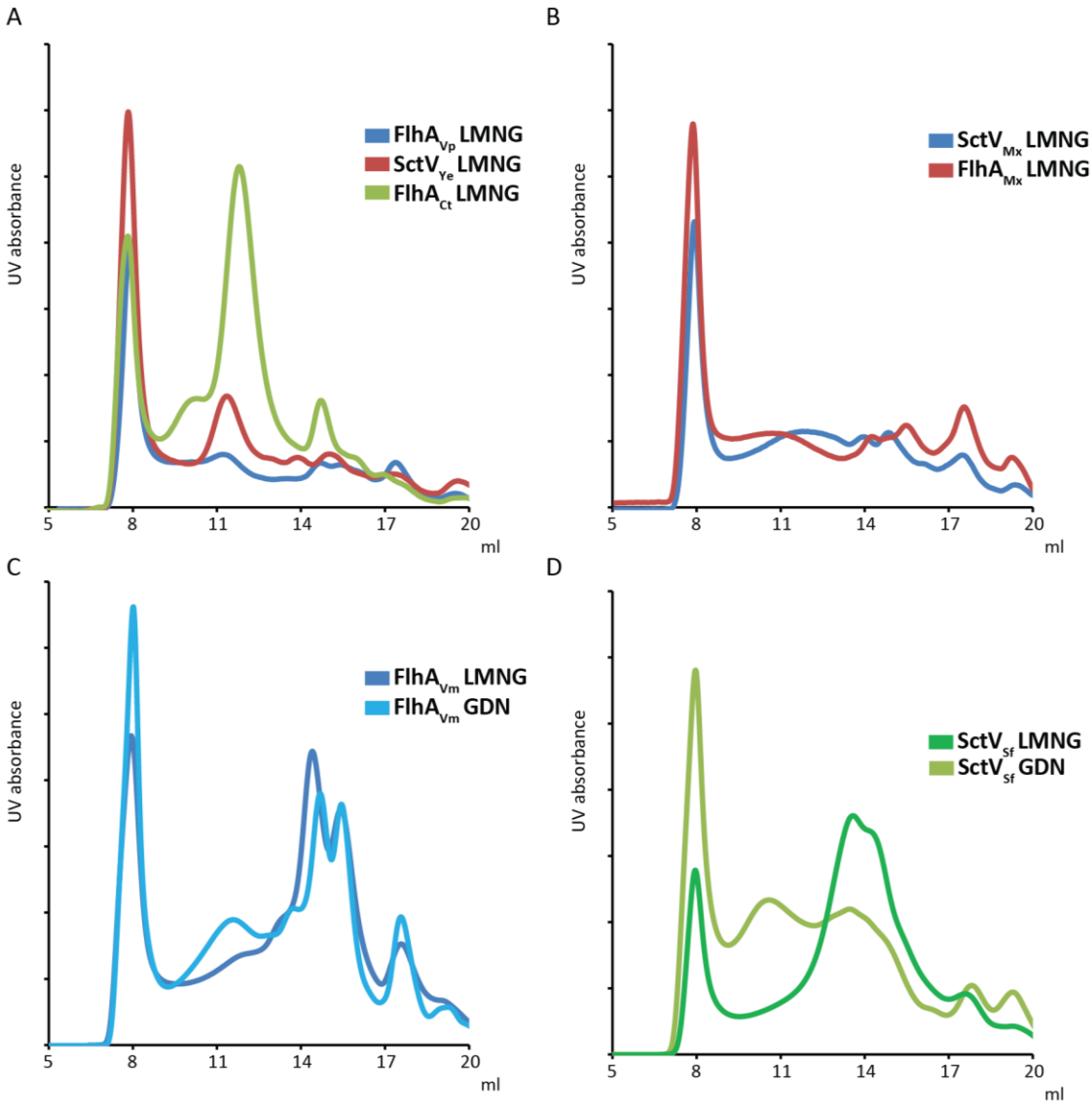


Figure 7.8 Screening FlhA/SctV-GFP assembly by SEC.
A, SEC of GFP fusions of FlhA$_{Vp}$, SctV$_{Ye}$ and FlhA$_{Ct}$ at 1 mg/ml in TBS containing 0.01% (w/v) LMNG using a Superose 6 increase 10/300 column (GE). **B**, SEC of GFP fusions of SctV$_{Mx}$ and FlhA$_{Mx}$ at 1 mg/ml in LMNG. **C**, SEC of FlhA$_{Vm}$ at 10 mg/ml in TBS containing 0.01% (w/v) LMNG or GDN. **D**, SEC of SctV$_{Sf}$ at 10 mg/ml in TBS containing 0.01% (w/v) LMNG or GDN.

The SEC elution volume of the major peak except the void peak was recorded and compared to FlhA$_{ST}$ (Table 7.1). Peaks close to the previously identified elution volume of 12 ml of the nonameric complex (Figure 7.4A) were regarded as indicating an intact complex. Three homologues, FlhA$_{Vp}$, FlhA$_{Ct}$ and SctV$_{Ye}$, were selected due to the peaks observed around 12 ml (Figure 7.8A) while FlhA$_{Mx}$ and SctV$_{Mx}$ samples were not selected because their peaks around 12 ml were very broad (Figure 7.8B). Concentration dependent assembly which was stabilised by GDN was most significant in FlhA$_{ST}$ as

described above, but partial stabilisation of a larger oligomer by GDN was also observed in $FlhA_{Vm}$ and $SctV_{Sf}$ (Figure 7.8C, D).

			SEC elution volume (ml)			
Homologue	Fluorescence	Yield (mg/litre culture)	LMNG, 1 mg/ml	LMNG, 10 mg/ml	GDN, 1 mg/ml	GDN, 10 mg/ml
$FlhA_{ST}$	Spots	10	14.08	12.47	14.5	12.19
$FlhA_{Vp}$	Spots	7	11.31	11.43	11.54	10
$FlhA_{Vm}$	Spots	4.5	14.59	14.43	14.97	15.06
$FlhA_{Tt}$	Spots	2	14.58	14.29	14.84	14.71
$FlhA_{Mx}$	Spots	5	10.73	10.5	11.24	11.28
$FlhA_{Ct}$	Spots	5	11.85	11.95	11.74	11.89
$SctV_{Ye}$	Spots	4.5	11.42	11.61	11.33	11.51
$SctV_{Sf}$	Spots	7	14.75	13.59	14.61	10.61
$SctV_{Ps}$	Spots	0				
$SctV_{Ct}$	Inclusion bodies					
$SctV_{Mx}$	Spots	3.5	11.79	11.78	11.23	10.73
$SctV_{ST1}$	Spots	4.5	14.84	14.81	15.43	14.97
$SctV_{ST2}$	Inclusion bodies					

Table 7.1 Results of the homologue screen.
Green indicates a positive result, after which the sample was analysed further.

7.5 Purification and Cryo-EM imaging of selected FlhA/SctV homologues

In the screen a short, uncleavable linker was used to fuse GFP to FlhA/SctV in order to avoid the small possibility that a longer linker could become the target of cellular proteases. However, the ring of GFP next to the ring of the cytoplasmic domain was not well ordered in one of the previous samples (Figure 7.6E). There is a small possibility that the presence of additional density that is not ordered relative to the density of interest would complicate the analysis of the cryo-EM data. Therefore, the three selected genes were cloned into the pT12 vector without GFP and with a tev site between FlhA/SctV

and GFP. The proteins were purified in the same way as the GFP fusion constructs in the screen using 4 L of culture.

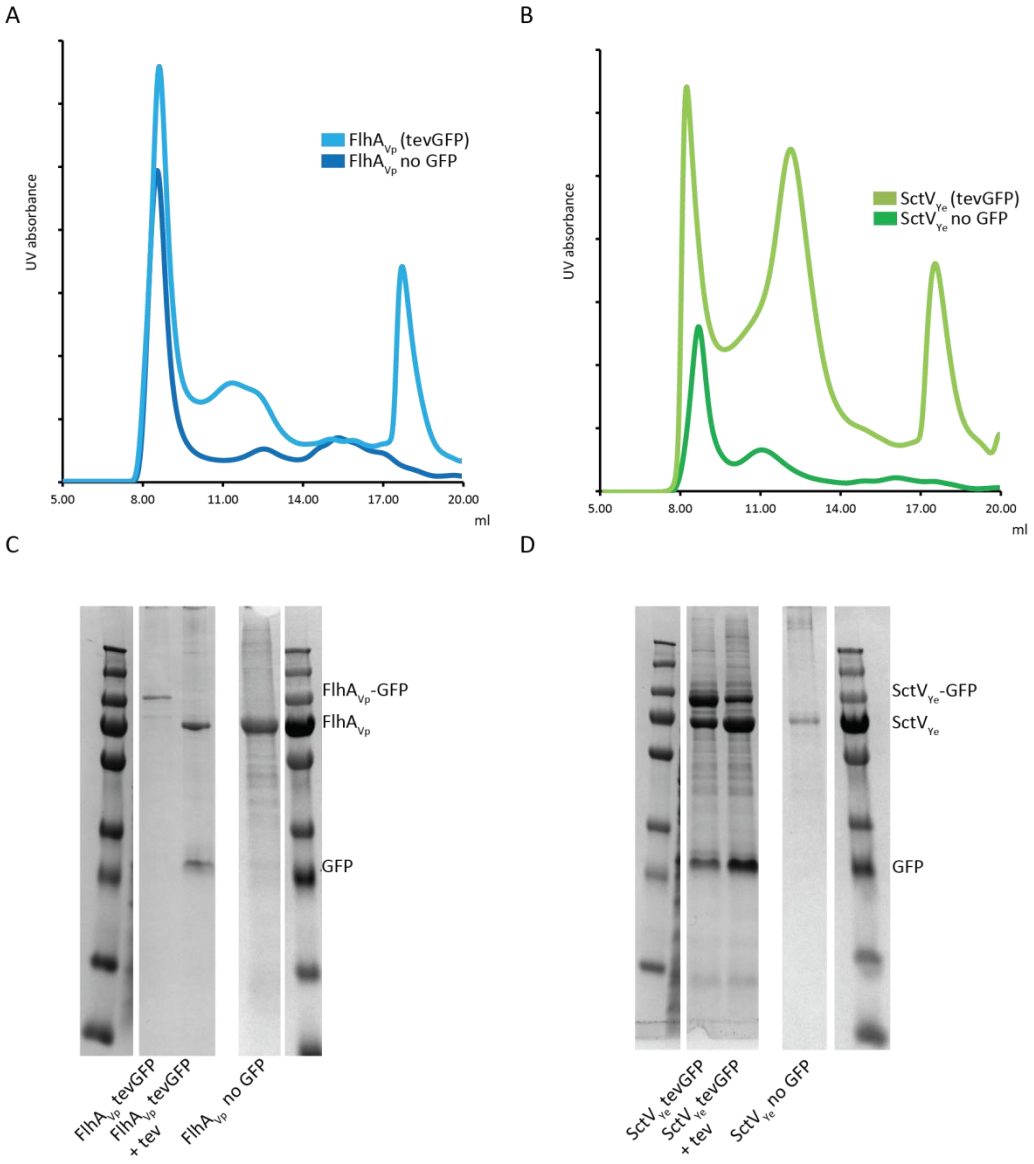


Figure 7.9 Assembled FlhA/SctV complexes lacking GFP.
A, SEC using a Superose 6 increase 10/300 column (GE) of purified FlhA$_{Vp}$ prepared by tev cleavage of the GFP fusion protein (light blue) or by expressing the protein without GFP (dark blue). **B**, SEC of purified SctV$_{Ye}$ prepared by cleavage of GFP (light green) or expression without GFP (dark green). **C**, SDS-PAGE analysis of strep eluates of FlhA$_{Vp}$-GFP before and after tev cleavage and FlhA$_{Vp}$ made without GFP. **D**, SDS-PAGE analysis of strep eluates of SctV$_{Ye}$-GFP before and after tev cleavage and SctV$_{Ye}$ made without GFP.

No protein was detected in the StrepTrap (GE) eluate when purification of either of the FlhA$_{Ct}$ constructs was attempted, suggesting that the stability of this homologue is limited. FlhA$_{Vp}$ and SctV$_{Ye}$ without GFP could both be purified, but there was a large

proportion of monomeric protein in $FlhA_{Vp}$ without GFP (Figure 7.9A and Table 7.2) and the yield of $SctV_{Ye}$ without GFP was very low compared to the GFP fusion protein (Figure 7.9B). However, purification of $FlhA_{Vp}$ and $SctV_{Ye}$ fused to GFP via a tev site was successful (Figure 7.9C, D). 1 mg of tev protease was added to the eluate and the protein was digested overnight at 4 °C. Subsequently, the sample was subjected to SEC using TBS containing 0.01% (w/v) LMNG (Figure 7.9A) on a Superose 6 increase 10/300 column.

Prior to further structural studies, the fractions corresponding to the peaks around 12 ml were pooled and analysed by SEC-MALS (section 2.5.8) (Table 7.2). This confirmed the presence of monomeric protein when $FlhA_{Vp}$ was expressed without GFP and this construct was excluded from further characterisation due to potential instability of the complex. A potential dimer of nonamers was observed in the $SctV_{Ye}$ complex expressed without GFP, and this larger complex was also present as a small contaminant in $SctV_{Ye}$ after tev cleavage of the GFP fusion. This suggests that the presence of GFP can inhibit dimerization, implying that dimerization occurs via the cytoplasmic domain to which the GFP was fused. This contaminant was also observed in $FlhA_{Ct}$-GFP. The completely dimerised $SctV_{Ye}$ without GFP was excluded from further studies and the fractions corresponding to the 12 ml peak of $SctV_{Ye}$ after tev cleavage of GFP and $FlhA_{Ct}$-GFP were selected for EM studies. The $FlhA_{Vp}$ complex after tev cleavage of GFP eluted close to 12 ml, but the calculated stoichiometry was 24.4. This was partially due to the processing of the MALS data suggesting that no detergent was bound to the protein. A technical replicate of the SEC-MALS experiment confirmed the result. It is possible that the MALS result is incorrect, if the real extinction coefficient of the protein differs from the theoretical value that was used in the data processing. Despite the uncertainty surrounding this sample, it was taken forward for initial imaging by EM in order to confirm whether the intact complex could be observed.

Homologue	Construct	Elution volume (ml)	Stoichiometry (protein conjugate analysis)
$FlhA_{Vp}$	FlhAtev(GFP-strep)	11.5	24.4
	FlhA-strep	12.3	12.5
		15.8	1.2
$FlhA_{Ct}$	FlhA-GFP-strep	10.5	18
		12.1	8.5
$SctV_{Ye}$	SctVtev(GFP-strep)	10.8	20.3
		12.2	9.6
	SctV-strep	11.1	19.8

Table 7.2 Summary of SEC-MALS results of selected constructs.
GFP-strep in brackets indicates a sample made as a GFP fusion from which GFP had been cleaved off.

Cryo-EM grids were made of all three selected samples ($FlhA_{Ct}$-GFP, $FlhA_{Vp}$ after tev cleavage of GFP and $SctV_{Ye}$ after tev cleavage of GFP) at a range of concentrations (section 2.5.12). 3 µl of sample were applied to a Quantifoil 300 mesh gold grid (copper grids for $SctV_{Ye}$) which was blotted after 5 seconds and plunge frozen in liquid ethane cooled by liquid nitrogen using a Vibtrobot Mark IV (FEI). Initial screening of grids showed that the ice thickness in all of the $FlhA_{Ct}$-GFP grids was too thick for imaging, but grids of $FlhA_{Vp}$ (at 0.8 mg/ml) and $SctV_{Ye}$ (at 1 mg/ml) of sufficient quality could be obtained and were imaged on a Titan Krios microscope (FEI) equipped with a K2 direct electron detector (Gatan). EPU (FEI) was used for automated data collection. CTF correction was carried out using CTFFIND4 implemented in SIMPLE (Reboul et al. 2018), motion correction using RELION's implementation of MotionCor2 (Zivanov et al. 2019) and particles were picked in SIMPLE (Reboul et al. 2018). The particles were classified in 2D in RELION-3.0 (Zivanov et al. 2018).

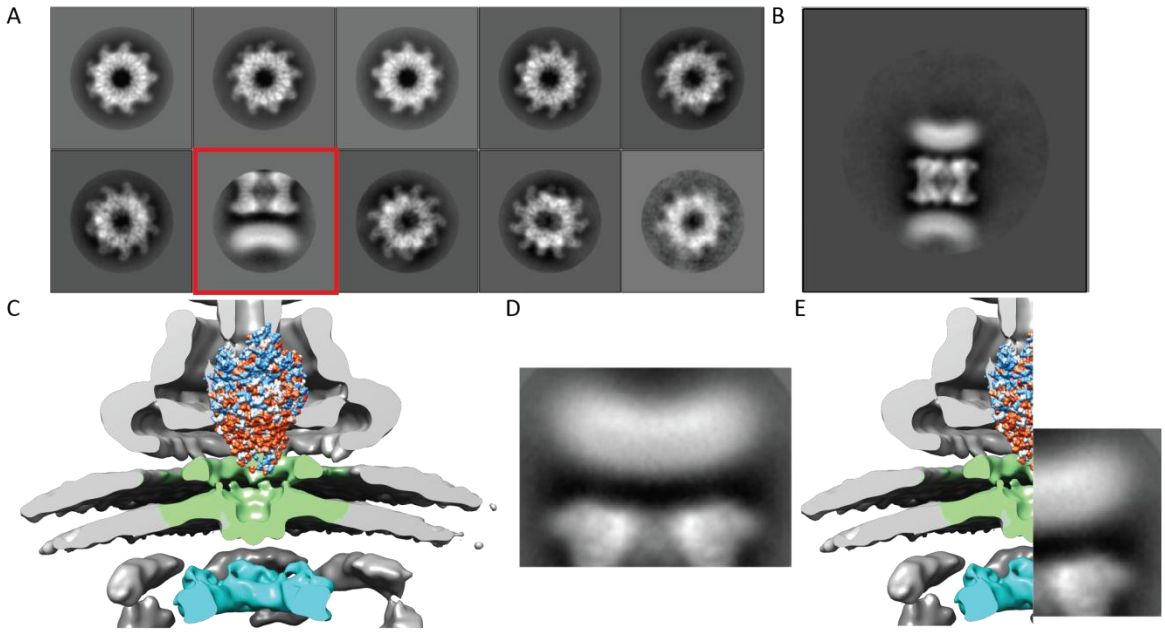


Figure 7.10 Cryo-EM imaging of $SctV_{Ye}$.
A, 2D class averages of $SctV_{Ye}$ following cleavage of GFP. **B**, Representative 2D class average of the $SctV_{Ye}$ side view extracted in a bigger box. **C**, The export gate coloured by hydrophobicity in a tomographic reconstruction of the vT3SS (EMD-8544). The SctV cytoplasmic ring is coloured in blue and the transmembrane ring is coloured in green. **D**, Zoomed view of the 2D class average representing the side view of $SctV_{Ye}$. **E**, Comparison of the tomographic reconstruction and the side view of $SctV_{Ye}$.

2D classification of $SctV_{Ye}$ particles (around 100,000 candidate particles from 4,854 micrographs) revealed class averages with apparent high resolution detail (Figure 7.10A), but the direction in which the outermost subdomain of $SctV_{Ye}$ pointed was not the same at every point of the ring, consistent with projections of a dimer of nonamers. Furthermore, the side view of the particles revealed a very large, symmetric structure in place of the cytoplasmic domain. These particles were selected and re-extracted in a bigger box and classified again, revealing the particles to be dimers of nonamers of $SctV_{Ye}$ (Figure 7.10B). In SEC experiments this species was just a minor contaminant (Figure 7.9 and Table 7.2). It is not clear why mostly dimers of nonamers were observed by cryo-EM; it is possible that this species is more stable. Interestingly, the membrane domain of $SctV_{Ye}$ was not flat, as previously observed for the $FlhA_{ST}$ membrane domain (Figure 7.5), but curved. The density ascribed to the membrane domain of SctV in tomographic reconstructions (Figure 7.10C and chapter 4) is less wide than the large cytoplasmic domain but it is taller due to a cytoplasmic extension. The curvature of the purified $SctV_{Ye}$ membrane domain is more consistent with the shape of this domain in

the *in situ* structure (Figure 7.10D, E) than the previously observed flat $FlhA_{ST}$ membrane domain.

This dataset is currently being processed in RELION-3.0 (Zivanov et al. 2018) by Dr Steven Johnson. Using a mask around the dimer of nonamers of the cytoplasmic domain and applying D9 symmetry led to a 3.7 Å resolution (FSC=0.143) volume (Figure 7.11A) of the cytoplasmic domain. Further processing is ongoing. Preliminary analysis reveals extensive interactions between the linker between cytoplasmic and membrane domains, $SctV_L$, and the neighbouring subunit (Figure 7.11B).

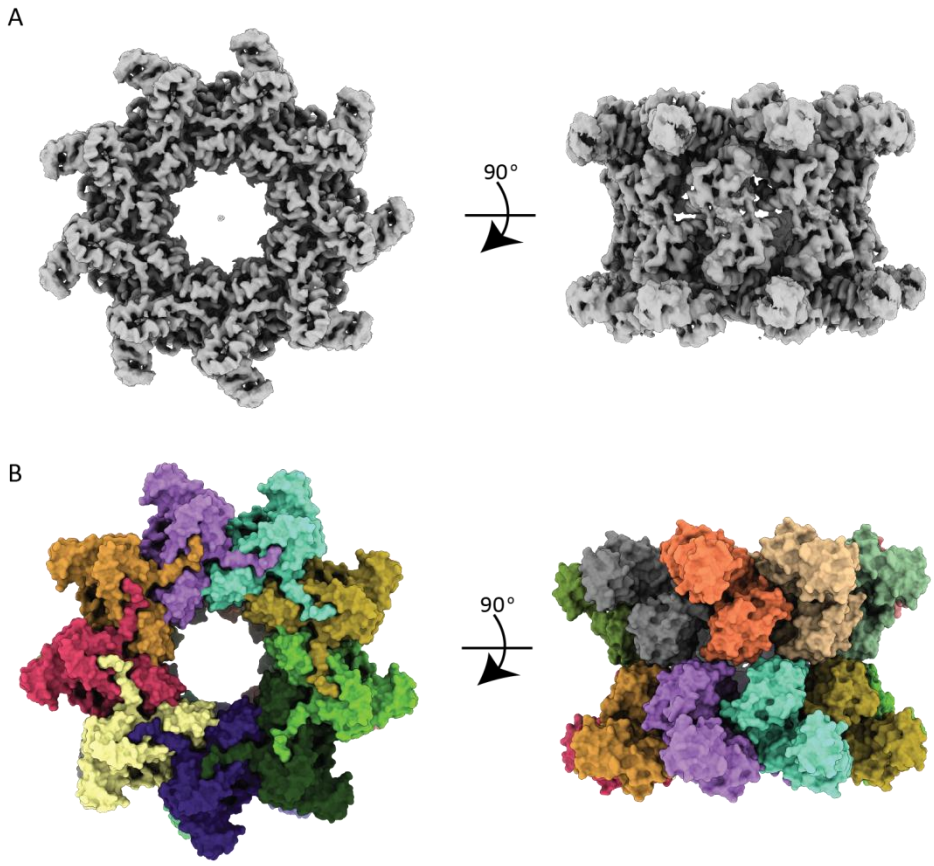


Figure 7.11 Structure of the cytoplasmic domain of $SctV_{Ye}$.
A, Unsharpened map of the $SctV_{Ye}$ cytoplasmic domains. **B**, Preliminary structure of $SctV_{Ye}$ (surface representation).

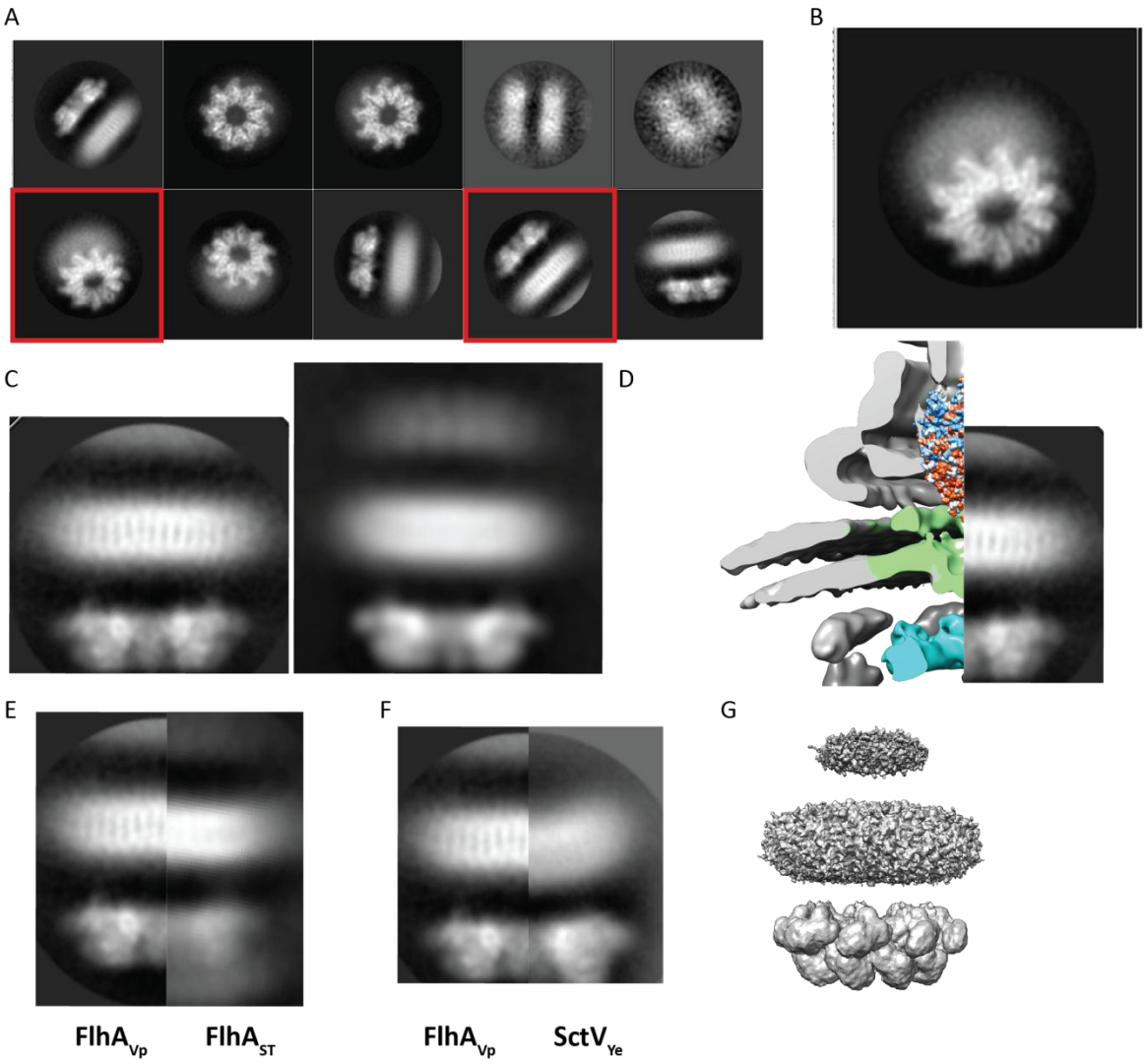


Figure 7.12 Cryo-EM imaging of FlhA$_{Vp}$.
A, 2D class averages of FlhA$_{Vp}$ calculated in RELION after selecting intact particles in the SIMPLE "clean-up" 2D classification. Red boxes indicate a selected tilted view and a side view. **B**, Zoomed view of the tilted view of FlhA$_{Vp}$. **C**, Zoomed view of the side view of FlhA$_{Vp}$ (left) and side view resulting from 2D class averaging with a larger mask (right). **D**, Comparison of the side view to the tomographic reconstruction of the injectisome (EMD-8544). **E**, Comparison with FlhA$_{ST}$ without cross-linking. **F**, Comparison with SctV$_{Ye}$. **G**, The 3D volume of FlhA$_{Vp}$ calculated from the data only has high resolution features in the cytoplasmic domain.

The FlhA$_{Vp}$ dataset was classified first using an unpublished "clean-up" 2D classification, a variation on the SIMPLE 2D classification (Reboul et al. 2018), and the selected particles were processed in RELION-3.0 (Zivanov et al. 2018). Top, side and tilted views of the complex could be observed (Figure 7.12A). However, there was no detail in the membrane domain in the tilted views (Figure 7.12B). In total, around 187,000 candidate particles were picked from 4,131 micrographs (Table 7.3). After "clean-up" 2D in SIMPLE,

only 22,077 particles were selected. The membrane domain of $FlhA_{Vp}$ imaged under these conditions was flat (Figure 7.12C), and while the size is approximately consistent with the tomographic volume, the flat shape is not (Figure 7.12D). It is possible that the overall conformation of the membrane domain changes following extraction with detergent. Interestingly, striations could be seen in the side views, which may correspond to transmembrane helices. However, this structure cannot be seen in the tilted view of the complex (Figure 7.12B). The unexplained density on the periplasmic side of the membrane domain was again observed as in $FlhA_{ST}$ and unlike $SctV_{Ye}$ (Figure 7.12E, F). Subsequent 3D classification and refinement only revealed detail in the cytoplasmic domain (Figure 7.12G).

Data collection and processing	
Voltage	300
Electron exposure	48 $e^{-}/Å^2$
Micrographs	4,131
Particles (total)	186,670
Particles (final)	8,916
Sampling	0.822 Å
Defocus range	0.5-4 µm
Refinement	
Resolution	3.8 Å
FSC threshold	0.143
Symmetry	C9
Map sharpening B factor	-134 $Å^2$

Table 7.3 Cryo-EM data collection and processing statistics for $FlhA_{Vp}$.

Following 2D classification all particles were aligned using RELION auto-refine, revealing good detail in the cytoplasmic domain, $FlhA_C$, but none in the transmembrane domain. Subsequently, 3D classification was carried out with local angular searches using a mask around either the cytoplasmic or the transmembrane domain (Figure 7.13). This revealed only noise in the transmembrane domain except one class in which the cytoplasmic domain was revealed by misaligned particles. Classification with a mask around $FlhA_C$, selection of the particles belonging to the highest resolution class and

auto-refinement and post-processing led to a 3.8 Å map (Figure 7.13 and Figure 7.14A). C9 symmetry was applied throughout. At lower contour levels a low resolution disc-shaped density could be seen in place of the transmembrane domain (Figure 7.14B).

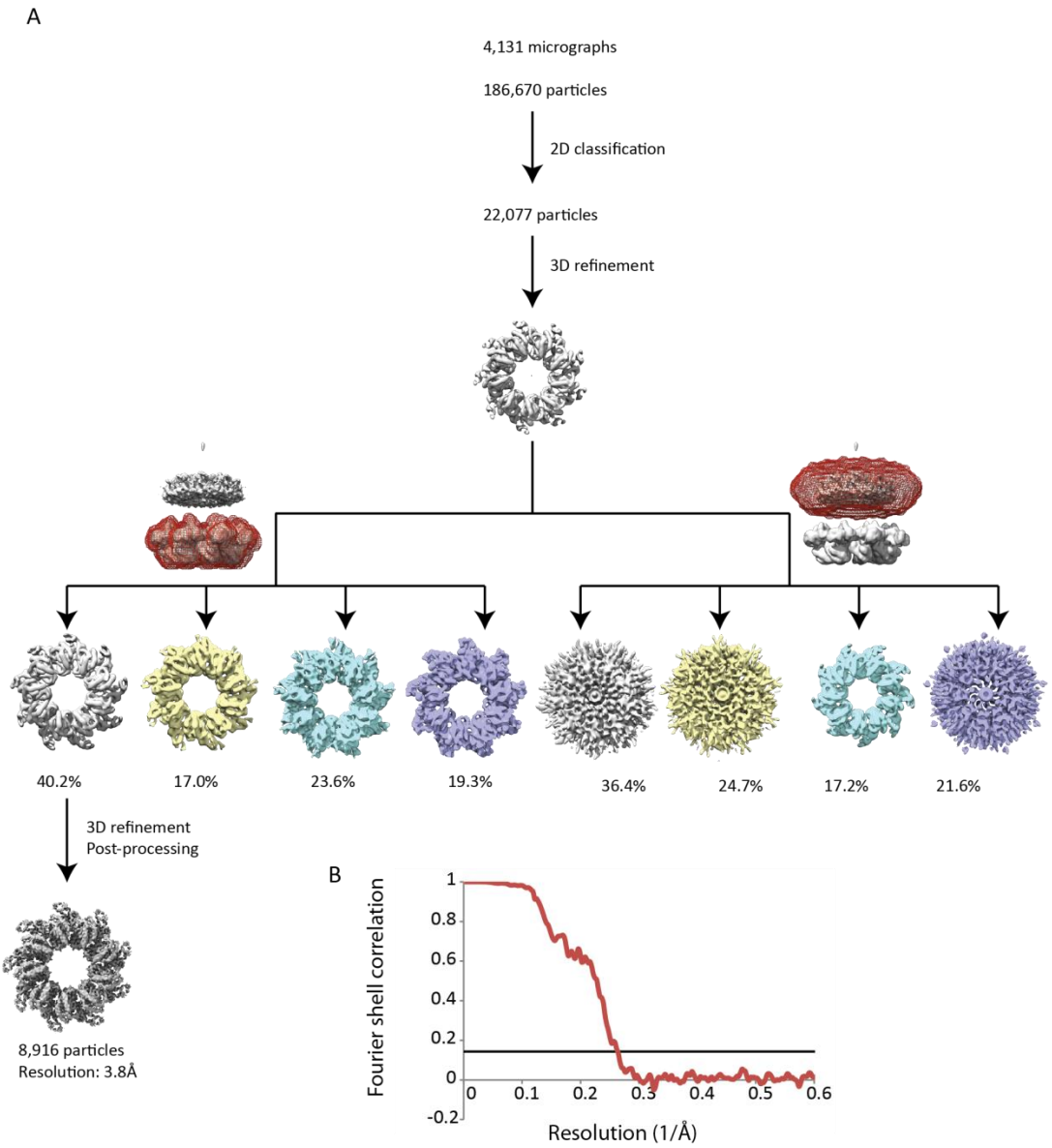


Figure 7.13 Data processing strategy for $FlhA_{Vp}$.
A, Data processing strategy flowchart. The volumes are shown from the perspective of the cytoplasm. The mask used for 3D classification is shown in red around a side view of the volume. **B**, Solvent-corrected FSC curve.

A model of $FlhA_C$ was built in Coot (Emsley et al. 2010) using the published structure of the monomeric $FlhA_C$ of *Salmonella* (Saijo-Hamano et al. 2010) as a guide. A single chain was built, copied 8 times and moved into the densities corresponding to the other

subunits. The model was refined using iterations of phenix.real_space_refine (Afonine et al. 2018) and manual adjustments in Coot (Table 7.4 and Figure 7.14).

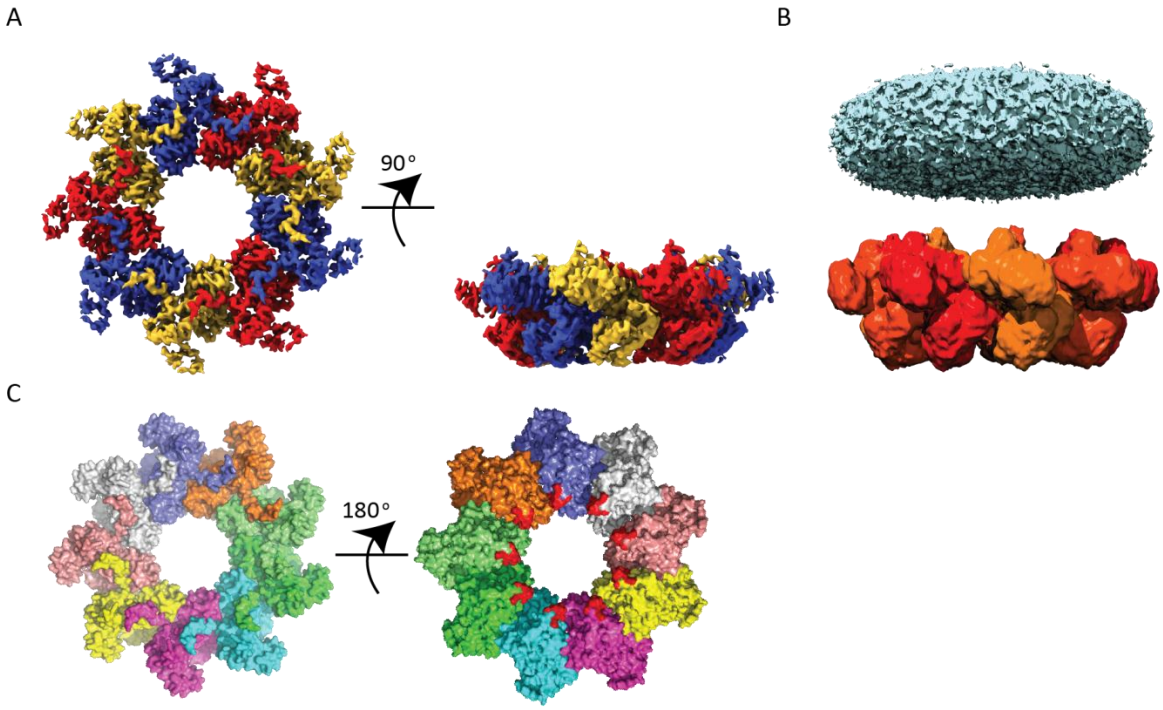


Figure 7.14 Structure of the cytoplasmic domain of $FlhA_{Vp}$.
A, Cryo-EM map of the cytoplasmic domain of $FlhA_{Vp}$. **B**, Cryo-EM map as in A at a lower contour level showing the noisy density corresponding to the membrane domain (blue). **C**, Structure of $FlhA_{Vp}$ coloured by chain. A cytoplasmic loop not present in $FlhA_{ST}$ is highlighted in red.

As expected from earlier crystal structures (Saijo-Hamano et al. 2010; Abrusci et al. 2013), $FlhA_C$ formed a nonameric ring in which the N-terminal linker that connects $FlhA_C$ to the transmembrane domain binds the top of the neighbouring subunit in the ring (Figure 7.14C), but the cryo-EM structure is the first observation of the nonamer in solution. A loop at the cytoplasmic side of $FlhA_C$ is not present in the *Salmonella* FlhA structure.

CC_{Mask}	0.7717
B factors	69.93 $Å^2$
R.m.s. deviations	
Bond lengths	0.006 Å
Bond angles	1.105°
Validation	
MolProbity score	1.98
Clashscore	6.4
Poor rotamers	0.84%
Ramachandran plot	
Favoured	86.51%
Allowed	12.35%
Disallowed	1.14%

Table 7.4 Model refinement statistics.

The structure of the $FlhA_{Vp}$ monomer is highly similar to $FlhA_{ST}$ (RMSD= 1.99 Å over 326 residues) (Figure 7.15A). FlhA/SctV can adopt an open and closed structure, and the linker between cytoplasmic and membrane domain ($FlhA_L$/$SctV_L$) binds the neighbouring subunit in the open form (Inoue et al. 2019b). The structures of both $FlhA_{Vp}$ and $SctV_{Ye}$ are in the open state (Figure 7.15A, B) confirming that the interaction between linker and neighbouring subunit occurs in the assembled nonamer in the open state. This $SctV_{Ye}$ structure is the first structure of an SctV protein in the open state. Biochemical data supports the idea that this interaction with the linker is important for assembly of the nonameric complex (Terahara et al. 2018). However, this interaction between $FlhA_C$ and $FlhA_L$ had only been observed in crystal structures of the monomeric protein (Saijo-Hamano et al. 2010), in which the interactions in the crystal mimic the native interaction in the nonamer, but not in the structure of the cytoplasmic domain of $SctV_{Sf}$, which crystallised in the closed state (Abrusci et al. 2013). In the $FlhA_{Vp}$ and $SctV_{Ye}$ structures $FlhA_L$/$SctV_L$ can be seen to extensively contact the neighbouring subunit (Figure 7.15C, D). The residue Trp350 in $FlhA_L$ binds a hydrophobic groove in the neighbouring subunit (Figure 7.15E) and this residue has been implicated in ring formation (Terahara et al. 2018).

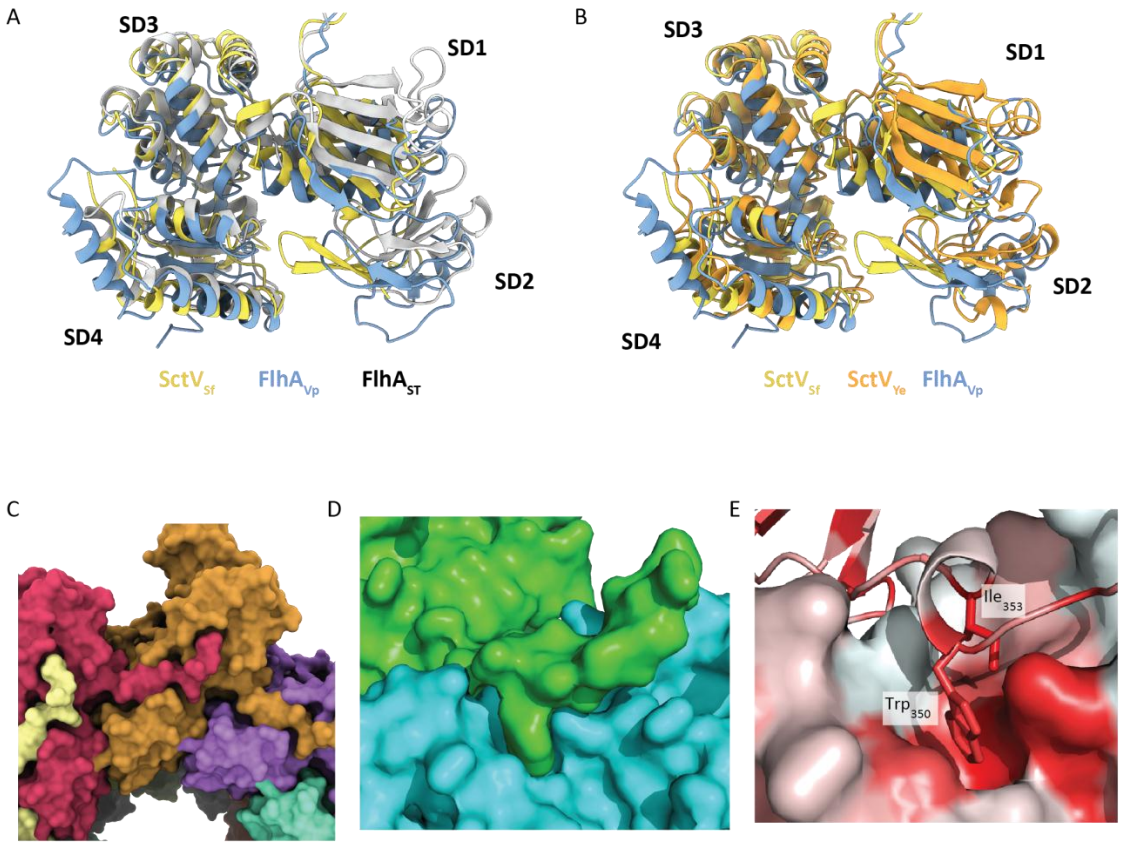


Figure 7.15 Structure of FlhA$_C$/SctV$_C$.
A, Comparison of a single chain of FlhA$_{Vp}$ (light blue) with FlhA$_{ST}$ (grey, PDB 3A5I) and SctV$_{Sf}$ (yellow, PDB 4A5P). **B**, Comparison of a single chain of FlhA$_{Vp}$ (light blue) with SctV$_{Sf}$ (yellow, PDB 4A5P) and SctV$_{Ye}$ (orange). **C**, Close-up view of the interaction between *Yersinia* SctV$_L$ and the neighbouring subunit. **D**, Zoomed view of the interaction between *Vibrio* FlhA$_L$ of chain A and chain B. **E**, Interaction of bulky hydrophobic residues in the linker with a hydrophobic groove in the neighbouring subunit in the structure of FlhA$_{Vp}$. Hydrophobic residues are coloured red and hydrophilic residues white.

As a high resolution structure could be calculated of one domain of the FlhA$_{Vp}$ complex, it was possible to subtract this volume from the particles and reveal the isolated membrane domain (Figure 7.16A) (Scheres 2016). 2D classification of the subtracted particles showed that although the striations could still be seen in the side views (Figure 7.16B), there was no detail in the top view of the membrane domain. This is consistent with the lack of detail in the membrane domain in tilted views (Figure 7.12B). High resolution information in only one direction would not be sufficient to reconstruct a 3D volume.

A possible explanation for the observation of high resolution features in one view but not the other could be that particles in a certain orientation in the ice are partially damaged due to contact with the air water interface. In an attempt to stabilise the sample, $FlhA_{Vp}$ was cross-linked. $FlhA_{Vp}$ was produced as before by tev cleavage of the C-terminal GFP fusion (Figure 7.9). The running buffer for SEC was HBS containing LMNG rather than TBS. The peak fractions were pooled and cross-linked in the same way as $FlhA_{ST}$ before (section 7.3) using glutaraldehyde. In addition to the final sample which was cross-linked for 30 minutes, samples were also taken after 10 and 20 minutes to monitor the cross-linking reaction (Figure 7.16C). Following cross-linking, a second SEC step was performed, as before. The peak fractions were taken and grids were made without concentrating the sample, which was at 0.9 mg/ml. Grids were made and imaged in the same way as for the sample without cross-linking. 2D classification in RELION-3.0 (Zivanov et al. 2018) revealed particles of similar shape as without cross-linking. Less detail could be seen in the cytoplasmic domain, but there were no more features in the membrane domain than before (Figure 7.16E, F).

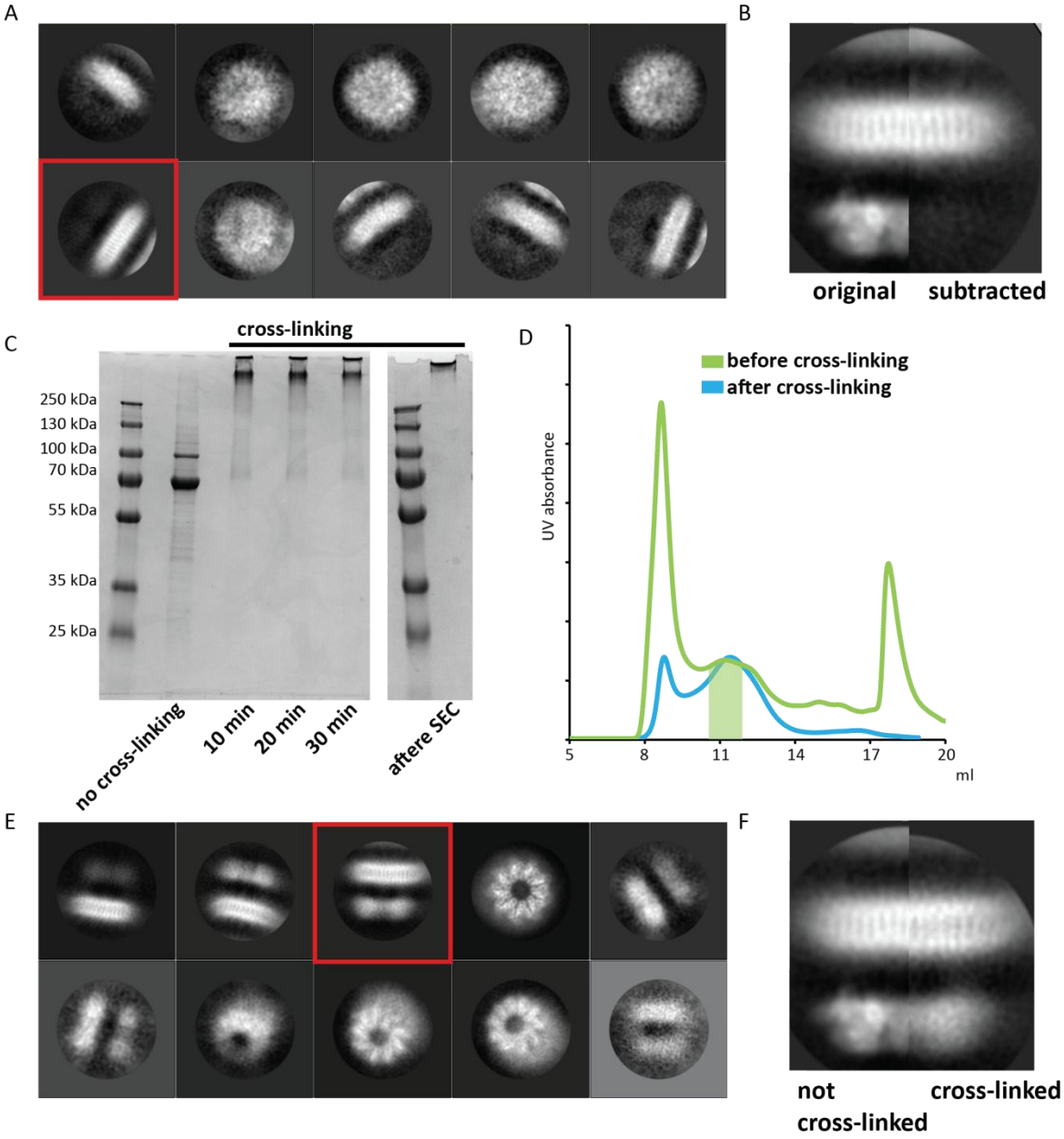


Figure 7.16 Additional cryo-EM analysis of FlhA$_{Vp}$.
A, 2D class averages of particles selected for the final reconstruction of the 3.8 Å volume following partial signal subtraction of the cytoplasmic domain. **B**, Comparison of the side view of the complex before and after particle subtraction. **C**, SDS-PAGE analysis of the cross-linking reaction of FlhA$_{Vp}$. **D**, SEC using a Superose 6 increase 10/300 column of one third of the sample before cross-linking (green) and the entire sample after cross-linking (blue). **E**, 2D class averages of the cross-linked FlhA$_{Vp}$. **F**, Comparison of the cross-linked and not cross-linked side view of the complex.

7.6 Preparation of larger complexes

The genomic context of $FlhA_{Vp}$, $FlhA_{Ct}$ and $SctV_{Ye}$ was explored in order to find other proteins that might be easily co-expressed as an operon simply by cloning a larger piece of DNA into the expression vector. This analysis revealed that $FlhA_{Vp}$ is downstream of FliPQR-FlhB and $SctV_{Ye}$ is downstream of SctXY (YscXY), which are known to interact with SctV (Diepold et al. 2011). No fliO gene could be found in this cluster of flagellar genes. Co-purification of FliPQR would produce a more biologically relevant complex and might stabilise the membrane domain of FlhA in the conformation in which it interacts with other membrane proteins. Co-purification of SctXY would enable the structural characterisation of the interaction of these secretion substrates with the export apparatus and their presence at the cytoplasmic domain of $SctV_{Ye}$ might prevent dimerization (Figure 7.10B), allowing the imaging of this protein without the requirement for the GFP fusion to block dimerization.

Purification of FliPQR-$FlhBA_{Vp}$-tevGFP using the same protocol as used for the construct of just $FlhA_{Vp}$ (section 2.5.3) produced pure $FlhA_{Vp}$ without co-purification of FliPQR (Figure 7.17A, B). Interestingly, the $FlhA_{Vp}$ complex eluted as a single, sharp peak around 12 ml unlike the broad peak observed previously, however this was not reproducible.

When $SctXYV_{Ye}$-tevGFP was purified in the detergent LMNG, using the same protocol as for $SctV_{Ye}$ alone (section 2.5.3), a band between 10 and 15 kDa could be observed in addition of $SctV_{Ye}$ (Figure 7.17A). Proteomics confirmed that it corresponded to SctY with some SctX also present. The purification of $SctXYV_{Ye}$ was scaled up and the tev cleavage step was skipped. The StrepTrap (GE) eluate was split in two and SEC was performed using TBS containing 0.01% (w/v) LMNG or 0.01% (w/v) GDN using a Superose 6 increase 10/300 column (GE). The peak around 12 ml was pooled and subjected to GraFix (Stark 2010) by performing gradient centrifugation. The purified protein (200 µl at 3 mg/ml) was deposited on top of a 5 ml sucrose gradient (15-40% (w/v) sucrose/0-0.2% (v/v) glutaraldehyde in TBS containing 0.01% (w/v) LMNG or GDN). The gradient was spun in a SW55Ti rotor at 35,000 rpm for 15.5 hours. The migration of the protein in the gradient was assessed by eye using fluorescence of GFP. The gradient was fractionated and the green fractions were pooled (Figure 7.17C), neutralised by

adding Tris, pH 8 to 100 mM and subjected to SEC again to remove the sucrose, which can reduce contrast in cryo-EM (De Carlo et al. 1999).

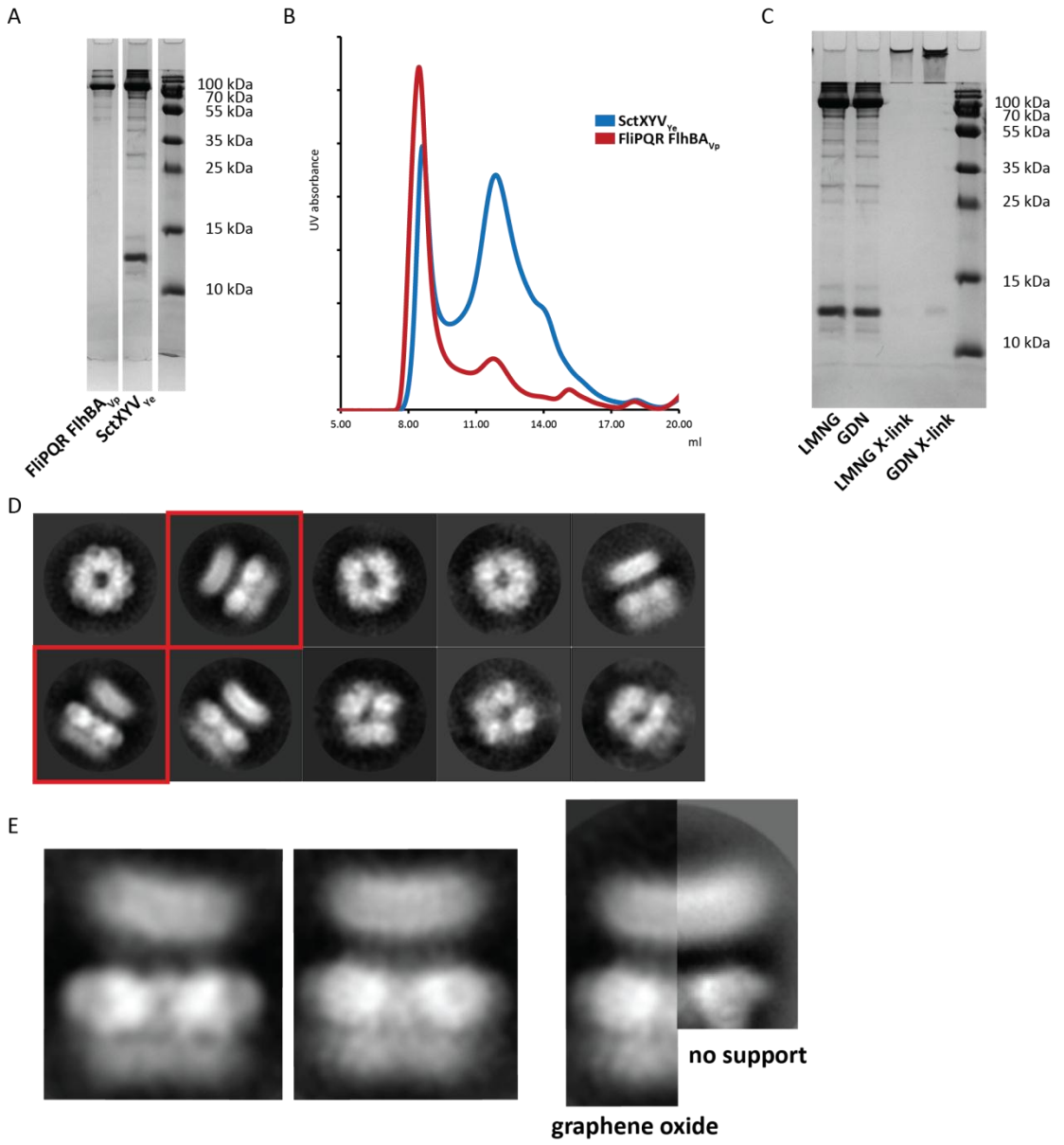


Figure 7.17 Purification and cryo-EM analysis of FliPQR-FlhBA$_{Vp}$ and SctXYV$_{Ye}$.
A, SDS-PAGE analysis of the strep eluate of FliPQR-FlhBA$_{Vp}$-GFP and SctXYV$_{Ye}$-GFP. **B**, Analytical SEC of SctXYV$_{Ye}$ (blue) and FliPQR-FlhBA$_{Vp}$ (red). **C**, SDS-PAGE analysis of SctXYV$_{Ye}$ in LMNG or GDN before and after GraFix. **D**, 2D class averages of SctXYV$_{Ye}$-GFP in GDN following GraFix on graphene oxide. **E**, Comparison of different side views (left) and comparison with SctV$_{Ye}$ imaged without a graphene oxide support layer (right).

Finally, cryo-EM grids were made as above (section 2.5.12) of both cross-linked SctXYV$_{Ye}$ in LMNG and GDN either using standard copper Quantifoil, 300 mesh grids or the same

grids with a layer of graphene oxide. If particles adsorb onto the graphene oxide surface, they should be protected from the air-water interface upon blotting. For graphene oxide grids, sample at 0.1 mg/ml was applied to the grid in a Vitrobot Mark IV (FEI) and after a wait time of 60 seconds blotted for 3 seconds and plunge frozen in liquid ethane. The humidity was 100% at 4 °C.

Initial screening of the grids using an Arctica microscope (FEI) showed that LMNG was incompatible with graphene oxide as there was large amount of amorphous objects on the grids, possibly related to LMNG micelles. Particle density on the grids without the graphene oxide support was low and there was limited aggregation. The ice on the grids of the LMNG sample without graphene oxide was too thick to image. Only the graphene oxide grid of $SctXYV_{Ye}$ in GDN was chosen for imaging. 1,832 micrographs were collected using a Titan Krios microscope (FEI) equipped with a K2 detector (Gatan). The micrographs were processed as above in SIMPLE (Reboul et al. 2018) and RELION-3.0 (Zivanov et al. 2019). Although individual particles were visually appealing, 2D class averaging in RELION-3.0 (Zivanov et al. 2019) revealed little high resolution detail (Figure 7.17D). Interestingly, there were two types of side views. One was similar to the side view of $SctV_{Ye}$ observed above (Figure 7.10), while the other was at an angle to the cytoplasmic domain. A low resolution volume of the cytoplasmic domain could be calculated when applying C9 symmetry (Figure 7.18A) into which the structure of $FlhA_{Vp}$ could be fit. At a lower contour level a ring of GFP could be seen on the cytoplasmic side of the cytoplasmic domain of $SctV_{Ye}$ (Figure 7.18B), suggesting that GFP is more ordered in this sample than in the cross-linked $FlhA_{ST}$ (Figure 7.6).

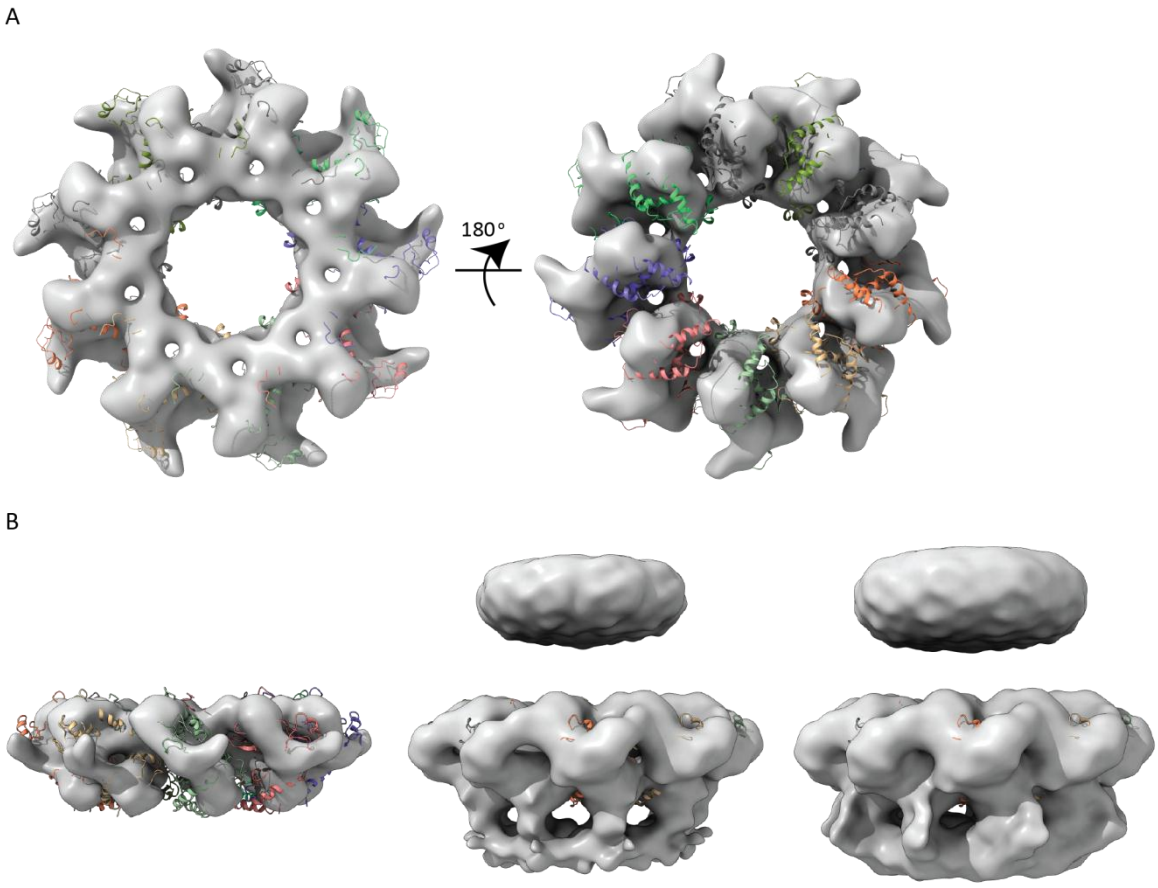


Figure 7.18 SctXYV$_{Ye}$ cryo-EM volume derived from sample on a graphene oxide support layer.
A, Low resolution volume of the SctXYV$_{Ye}$-GFP cytoplasmic domain into which the structure of the FlhA$_{Vp}$ cytoplasmic domain was fit. **B**, Side view of the volume at decreasing contour levels shows a ring formed underneath the cytoplasmic domain where GFP would be expected.

It has recently been proposed that the closed state of FlhA$_C$/SctV$_C$ binds to early substrates such as FlgD, but the affinity was low in pull-down assays (Inoue et al. 2019b). It is possible that the affinity is lower in the monomeric FlhA$_C$. Unpublished data from Jerry Cao indicated that FlhA$_C$ from *V. parahaemolyticus* can form the nonameric complex even in the absence of the membrane domain. Therefore, FlhA$_C$ from *V. parahaemolyticus* was cloned into the pT12 vector with a C-terminal tev site fused to GFP. FlgD from the same flagellar cluster was also cloned into the pT12 vector including the stop codon, making the expressed protein untagged. The two proteins were expressed separately in TB media containing 0.2% (w/v) rhamnose monohydrate and 60 µg/ml kanamycin. The cells were harvested and pooled after overnight growth at 37 °C, 180 rpm and lysed in a C5 homogeniser (Avestin). The clarified lysate was applied to a 5

ml StrepTrap column (GE), which was then washed with TBS and eluted with TBS containing 10 mM desthiobiotin. Finally, the protein was further purified by SEC using a Superose 6 increase 10/300 column (GE). SDS-PAGE analysis revealed no co-purification of FlgD (Figure 7.19A). Prior to further studies with the purified $FlhA_C$, tev cleavage was carried out to remove GFP using 1 mg of tev. Surprisingly, SEC following tev cleavage showed a very large shift of approximately 6 ml to approximately the elution volume of GFP, indicating dissociation of the protein into most likely monomers (Figure 7.19B). This further indicates that GFP stabilises the nonameric assembly.

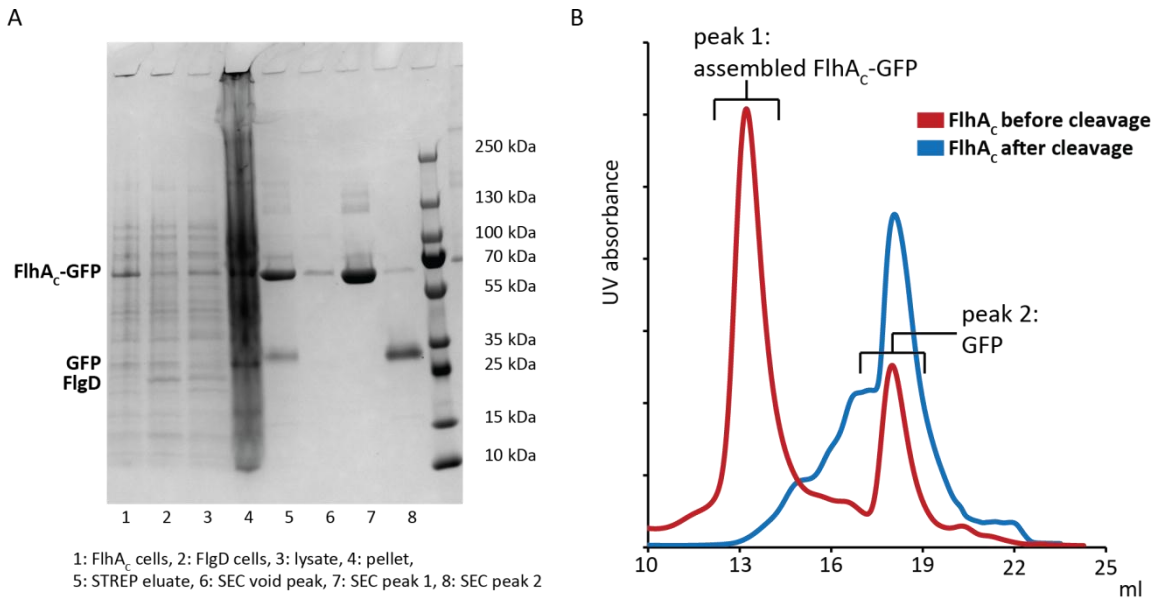


Figure 7.19 $FlhA_C$ nonamerisation can be stabilised by GFP.
A, Purification of the cytoplasmic domain of $FlhA_{Vp}$, $FlhA_C$. **B**, SEC of $FlhA_C$ before (red) and after (blue) cleavage of GFP.

7.7 Discussion

Full-length FlhA/SctV has long been regarded as a highly fragile complex. In this chapter, preparation of an intact nonameric complex was demonstrated through the use of GFP fusion proteins. This allowed the determination of the structure of the cytoplasmic domain of FlhA from *V. parahaemolyticus*. It shows for the first time a nonameric structure of the cytoplasmic domain of FlhA in the open state. It confirms that the linker which connects cytoplasmic and membrane domains, $FlhA_L$, is attached to the neighbouring subunit through a hydrophobic pocket, which has previously been

suggested (Terahara et al. 2018). The importance of hydrophobic interactions with the linker may explain the detergent sensitivity of the protein. The specific structure of FlhA from the lateral flagellar system of *V. parahaemolyticus* also illustrates that there is a short insertion in subdomain SD4 in this homologue of FlhA. There are four different T3SS in *V. parahaemolyticus*, two injectisomes and two flagellar systems (Makino et al. 2003). Therefore, higher specificity of recognising substrates may be required in this species. The additional loop in SD4 points at the inside of the ring and it could be speculated that it may have a role in selecting substrates.

The cytoplasmic domain of $FlhA_{Vp}$ was in the open state. Recently, the interaction between FlhA/SctV and late substrates, which bind to the export apparatus in complex with chaperones, has been studied crystallographically (Xing et al. 2018). Chaperones were found to bind the open state of FlhA in a construct of the cytoplasmic domain lacking the linker to the membrane domain. In addition, it has been suggested that FlhA only binds late substrates, like flagellin, in the open state while the closed state is responsible for secretion of early substrates like FlgD (Inoue et al. 2019b). The structure of $FlhA_{Vp}$ is consistent with the modelled structures of the nonameric complex in the open state. In particular, the structure of the linker contacting the neighbouring subunit confirms this stabilisation of the nonamer in the open state. The structure of the complex of FlhA and the chaperone-substrate complex can now be modelled using the nonameric structure of $FlhA_{Vp}$ (Figure 7.20).

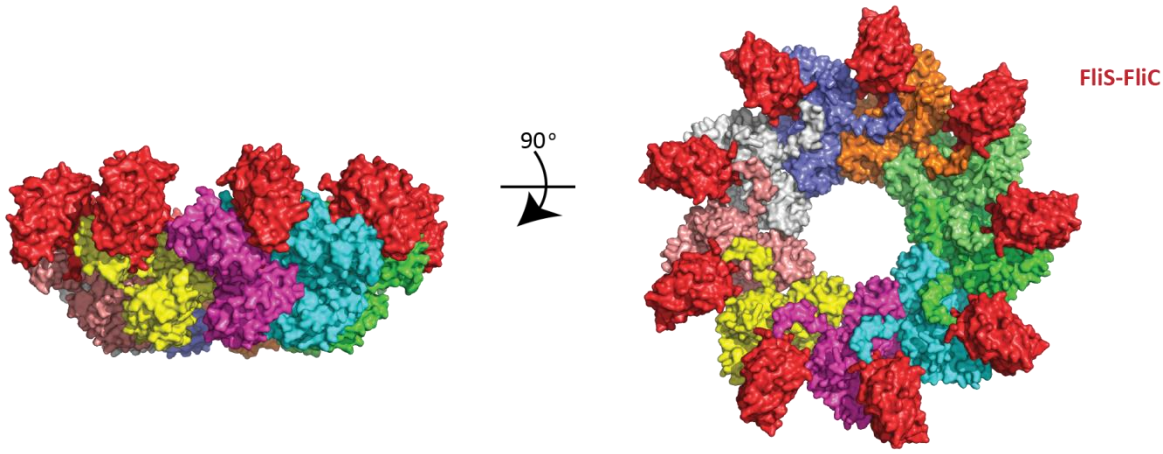


Figure 7.20 Modelling of substrate binding to $FlhA_{Vp}$.
Structure of the $FlhA_{Vp}$ cytoplasmic domain (coloured by chain) with a modelled FliS-FliC (chaperone and flagellin) fusion construct taken from PDB: 6CH3 (Xing et al. 2018).

Although it was not possible to determine the structure of the membrane domain, different shapes corresponding to this domain were observed in the different samples. In both $FlhA_{ST}$ and $FlhA_{Vp}$ a flat membrane domain was seen and in the side view of $FlhA_{Vp}$ there were consistent striations in different samples (Figure 7.16). Interestingly, the shape of the membrane domain was curved in $SctV_{Ye}$ (Figure 7.10). This may reflect a more native overall conformation, but no high resolution detail could be discerned. Furthermore, a third conformation could be seen in a cross-linked $SctV_{Ye}$ sample on a graphene oxide support in which the membrane domain was tilted at an angle with respect to the cytoplasmic domain (Figure 7.17). It is not clear if all these are possible conformations or artefacts that could be caused for example by extraction in detergent. In addition, a density of unknown origin was observed at a defined distance on the periplasmic side of the membrane domain in both samples in which the membrane domain was flat, $FlhA_{ST}$ and $FlhA_{Vp}$ (Figure 7.12), but the meaning of this is not clear.

Despite these differences, the distance between the membrane domain and the cytoplasmic domain was similar in all samples and matched the distance observed in tomograms. This suggests that although the membrane domain may be damaged at some point in the purification and grid preparation process, preventing structure determination, there is some mechanism for setting the distance between the two domains and, presumably, the membrane domain has to be folded correctly for it to function even in the detergent solubilised complex. A cyclical movement of the cytoplasmic domain to and from the membrane has been proposed as a mechanism of facilitating secretion (Erhardt et al. 2017). The observation that the linker between the two domains is not entirely flexible and instead maintains a set distance between the domains supports the idea that the distance can be controlled. However, the possibility of a second state in which the distance is reduced remains hypothetical.

How might the structure of the membrane domain be obtained? The membrane domain in the $FlhA_{Vp}$ sample appears to be less structured in some views. Additional stabilisation of this domain may allow structure determination. Ideally, the membrane domain would be embedded in lipid, for example by preparing a sample in nanodiscs or extracting the protein from the membrane using co-polymers in the absence of detergents (Autzen et

al. 2019), but this may not be possible due to the large size of the membrane domain of FlhA/SctV. Further strategies include preparing grids at a higher speed using the Spotiton technology which can avoid destabilisation of the protein at the air-water interface (Noble et al. 2018). Already purified membrane proteins may be stabilised by the addition of solutes such as novel calixarene based reagents (Hardy et al. 2018) while more common reagents such as trehalose (Nguyen et al. 2018) or glycerol (Majewski et al. 2019) have also been used. In particular, trehalose has been described as a remarkable substance due to its property of reducing radiation damage while maintaining contrast (De Carlo et al. 1999). The $SctV_{Ye}$ cytoplasmic domain dimerised when GFP was removed and the structure of the membrane domain could not be determined from this sample (Figure 7.10). Although imaging of a GFP fusion on graphene oxide avoided dimerisation, no high resolution detail could be seen in this context (Figure 7.18). It may be possible to obtain high resolution data by preparing a grid as in the first $SctV_{Ye}$ sample but employing a GFP fusion to block dimerisation.

Whichever strategy may be used for preparing future samples, the observations about distance between the domains and the shape of the membrane domain suggest that structure determination is possible using the constructs developed here that for the first time allow relatively straightforward production of large quantities of nonameric FlhA/SctV.

8 Conclusions

8.1 Summary

The aim of this thesis was to provide structural understanding of how secretion of proteins across the inner membrane is achieved in type three secretion, the bacterial protein secretion mechanism found in the molecular syringes known as injectisomes (vT3SS) and the bacterial flagellum (fT3SS). These nanomachines are associated with bacterial motility and pathogenicity in bacteria responsible for diseases including plague, typhoid fever and some forms of dysentery. Despite description of the structure of a large number of components of injectisomes and flagella (Portaliou et al. 2016) the crucial step of crossing the inner membrane remained mysterious.

The focus of this work was on the five putative membrane proteins that have been associated with the step of crossing the inner membrane, SctRSTUV in vT3SS and FliPQR FlhBA in fT3SS, collectively known as the export apparatus. Only the substrate binding, cytoplasmic domains of SctV/FlhA (Bange et al. 2010; Saijo-Hamano et al. 2010; Worrall et al. 2010; Abrusci et al. 2013) and SctU/FlhB (Deane et al. 2008; Zarivach et al. 2008; Lountos et al. 2009; Meshcheryakov et al. 2013) and a fragment of SctR/FliP (Fukumura et al. 2014; Fukumura et al. 2017) had been characterised. The membrane domains of the export apparatus were assumed to form a pore in the inner membrane through which substrates can be secreted, but it was not known which proteins would make up the pore and whether it was gated. SctRST/FliPQR in particular had been difficult to study directly due to their predicted membrane topology of being almost completely embedded in the membrane (Erhardt et al. 2017). In a first step the FliPQR and SctRST complexes from multiple systems were produced by expressing the proteins from operons, as they are found in genomes (Macnab 2004), and characterised by native mass spectrometry (chapter 3). This suggested a conserved architecture of a heterohexameric ring made up of five FliP/SctR and one FliR/SctT surrounded by a number of more fragile FliQ/SctS subunits, explaining earlier results that had led to differing models for fT3SS and vT3SS (Dietsche et al. 2016; Fukumura et al. 2017).

Mass spectrometric analysis also showed that the complex was stabilised best in the detergent LMNG and using this detergent proved instrumental in producing a sample for structure determination by cryo-EM (chapter 4). The structure of FliPQR confirmed the mass spectrometry model of the unusual structure of a ring formed by five FliP and one FliR. Surprisingly, this ring was a pseudosymmetric, helical complex made up of very extended subunits, producing a large structure on the periplasmic side of the membrane despite having been widely regarded as canonical transmembrane proteins (Taylor et al. 2016; Erhardt et al. 2017; Ward et al. 2018). Hydrophobic residues predicted to form transmembrane helices are instead buried by protein-protein interactions, illustrating the dangers of assessing membrane protein topology of isolated subunits of a complex. Five copies of FliP and one FliR, which is a structural fusion of FliP and FliQ, form a single turn of helix with helical parameters similar to the helical needle and filament of injectisome and flagellum respectively. Comparison with reconstructions of the assembled nanomachine showed that the FliPQR complex is not present in the inner membrane following incorporation into the basal body (Figure 8.1). Instead, it is held in the periplasm by contacts with other basal body proteins. This location combined with the helicity of the complex suggest that the helical rod, the first part of the helical filament to assemble inside the basal body, is assembled directly onto FliPQR/SctRST which templates the rod's helical parameters.

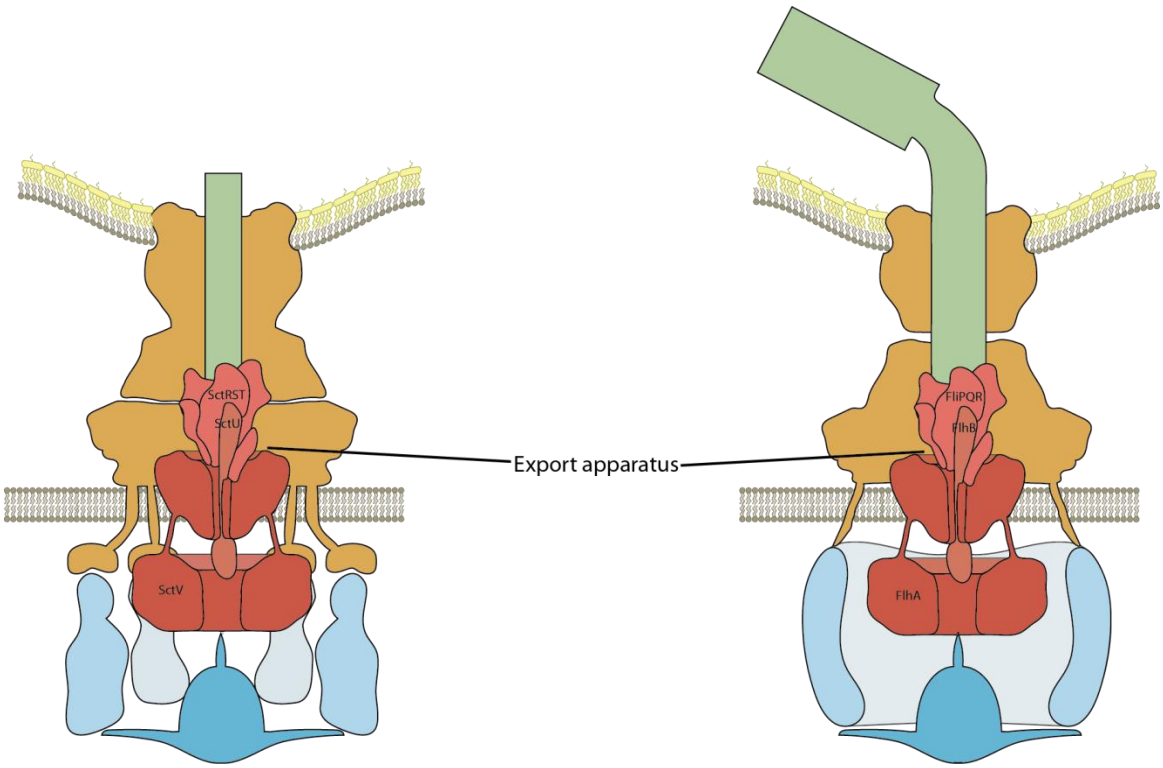


Figure 8.1 New model of the export apparatus in the T3SS of injectisomes and flagella. Cartoon model showing the location of the export apparatus proteins as established in this work.

The conservation of this unusual complex was further explored by determining the structures of two FliPQR homologues and one SctRST complex, revealing the high degree of conservation of the complex (chapter 5). Observation of different occupancy of SctS and analysis of the hydrophobic surfaces of the subunits suggested that as the complex assembles in the inner membrane following its helicity, the subunits are slowly pushed out of the membrane as protein-lipid interactions are replaced by protein-protein interactions.

As all structures of FliPQR/SctRST were in the closed state, the interaction with the substrate specificity switch protein FlhB/SctU was studied next in the hope of gaining insight into the mechanism of opening (chapter 6). The structure of FlhB in complex with FliPQR revealed that the FlhB transmembrane domain is made up of four helices which instead of forming a bundle, as previously suggested (Taylor et al. 2016), are a structural mimic of two FliQ subunits. Like FliQ, the FlhB helices are wrapped around the outside the FliPR core of the FliPQR-FlhB complex following its helical parameters. As in the

previous structures, the complex is closed. Unexpectedly, a loop of FlhB formed a circular structure around the entrance to the FliPQR lumen facing the cytoplasm. Mutagenesis supports the importance of the tight interaction between FliQ and the FlhB loop for opening of the complex and analysis of the mutant data and salt bridges in the FliPQR-FlhB complex suggest that a pulling force on FlhB could cause opening of the channel. Such a force might be applied by the final remaining export apparatus protein, FlhA.

The membrane domain of the fifth export apparatus protein, FlhA/SctV, has been implicated in powering secretion by harnessing the pmf (Minamino and Namba 2008; Minamino et al. 2016). The cytoplasmic domain of FlhA/SctV is known to form a nonameric complex (Abrusci et al. 2013), but study of the assembled full-length protein has been hampered by its fragility. GFP fusions of FlhA allowed production of large amounts of full-length protein, which assembled in a concentration dependent manner, and screening revealed that stably assembled complexes could be produced from sequences derived from multiple different systems (chapter 7). Although structure determination of the membrane domain was not successful, cryo-EM data provided intriguing observations about the shape of the membrane domain and its fixed distance from the cytoplasmic domain. In addition, the structure of the nonameric cytoplasmic domain was solved, illustrating the stabilisation of the structure by the interaction of the linker between the domains with the neighbouring subunit's cytoplasmic domain. These results will form the basis of future efforts to study the structure of the membrane domain of FlhA/SctV.

8.2 Evolutionary origin of FliPQR-FlhB/SctRSTU

The unusual architecture of the FliPQR-FlhB/SctRSTU export gate raises the question of why the complex is not built in a simpler way. For example FliP is known to form hexameric rings in the absence of FliR (Fukumura et al. 2017). Due to the structural similarity of a FliP-FliQ pair and FliR a similar complex could be formed by a hexamer of FliR, but analysis of a wide range of genome sequences failed to find a species which possesses only a FliR homologue, which could potentially form a primordial hexameric complex (Fabiani et al. 2017).

One function of the asymmetry introduced by splitting FliR into FliP and FliQ may be to create a unique point in the complex for the interaction with a single FlhB subunit. Another reason for the complex assembly may be that a homohexamer may be predisposed to be circularly symmetric rather than helical, which would not allow the interaction with the helical rod and filament. Furthermore, the specific 5 to 1 stoichiometry at the core of the export gate allows the complex to adopt the helical parameters of the filamentous components of the secretion system, which is made up of a helix with around 5.5 subunits per turn. A perfect hexamer might template a 6 subunit per turn helix.

Notably, the role of the T3SS export apparatus proteins has recently been established in the biogenesis of nanotubes (Bhattacharya et al. 2019; Pal et al. 2019), suggesting that at least the complex of FliPQR-FlhBA/SctRSTUV complex may be an exchangeable module of different systems and may be older than the T3SS.

8.3 Assembly of FliPQR and FliO

FliO is a flagella-specific chaperone that is thought to be responsible for stability of FliP and assembly of the FliPQR complex (Barker et al. 2010; Fabiani et al. 2017), which is why it was co-expressed with FliPQR to produce the FliPQR samples in this thesis. However, it is not known why the flagellar export gate assembly is stabilised by this specific chaperone but the injectisome SctRST complex isn't and why a signal sequence is present in FliP but not SctR.

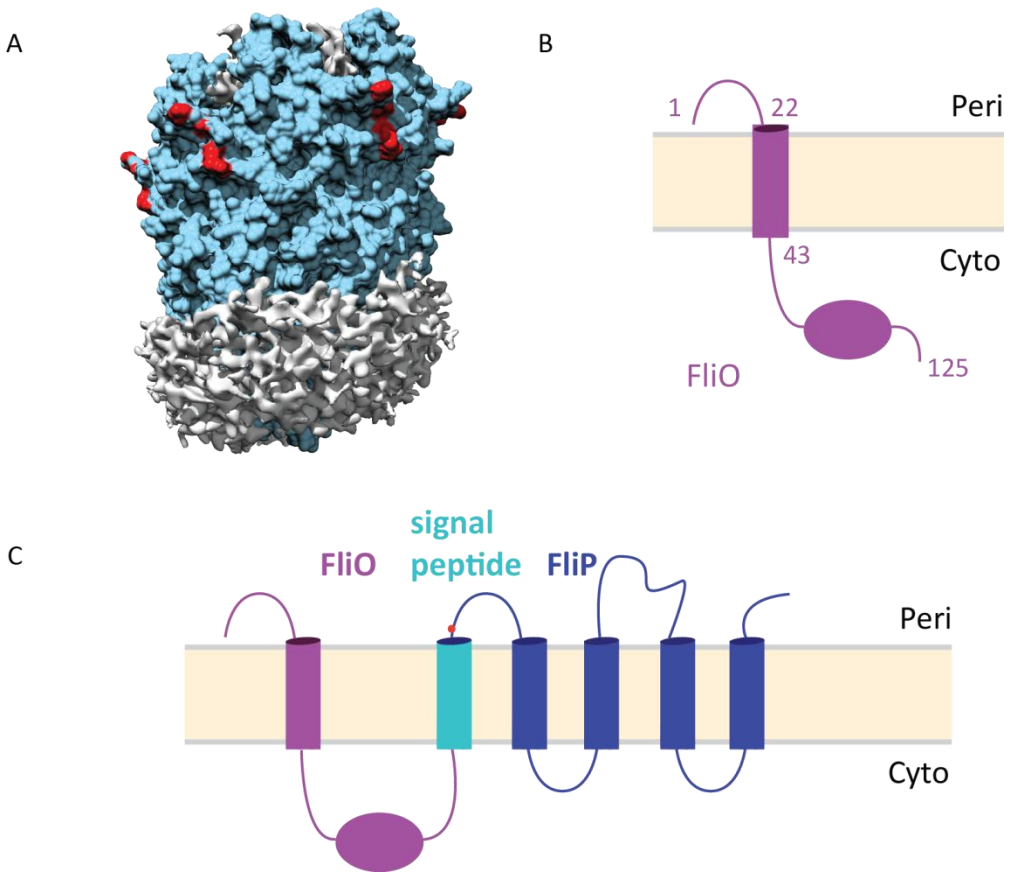


Figure 8.2 Interaction of FliO with FliP.
A, FliPQR (blue surface) structure with highlighting of the residues of FliP (red) shown to interact with FliO by *in vivo* cross-linking (Fukumura et al. 2017). The detergent belt is shown in grey. **B**, Experimental topology of FliO as determined by fusions to alkaline phosphatase and GFP (Barker et al. 2010). **C**, Predicted topology (Krogh et al. 2001) of the *Buchnera* FliO-FliP fusion protein. The parts of the sequence homologous to the *Salmonella* FliO protein, the FliP signal peptide and FliP are highlighted. A predicted signal peptide cleavage site (Nielsen 2017) is highlighted with a red dot.

FliO and FliP have been shown to interact by *in vivo* cross-linking (Fukumura et al. 2017), but when proposed interacting residues are mapped onto the structure of FliPQR (Figure 8.2A) it is not clear how interaction with FliO, which is predicted to only contain 22 periplasmic residues (Barker et al. 2010) (Figure 8.2B), is achieved. Furthermore, the FliO cytoplasmic domain alone can weakly complement a ΔFliO mutant, but FliP in the FliPQR structure is not exposed to the cytoplasm and the crosslinking residues are far above the plane of the membrane (Figure 8.2A, B). These inconsistencies suggest assembly intermediates with different structures and make FliO an interesting target for future structural studies.

In addition to this unclear interaction between FliO and FliP, fusion proteins of FliO and FliP are found in some species, most notably in the aphid symbiont *Buchnera aphidicola* (Macnab 2004). However, given the location of the export gate complex inside the basal body there would be no space for FliO in the same location (chapter 4), however, it is not known whether the *Buchnera* T3SS basal body could be unusually large to be able to accommodate the five FliO subunits, assuming the stoichiometry of the export gate is conserved in *Buchnera*. Analysis of the sequence of the fusion protein (Figure 8.2C) suggests that the signal peptide cleavage site remains conserved, but it is not known if the signal peptidase in this species can cleave the fusion protein between the predicted transmembrane helices 2 and 3. These predicted transmembrane helices are likely not canonical transmembrane helices, as observed in the structures of FliPQR and SctRST reported in this thesis. A different possibility is that there is an internal start codon resulting in the production of FliP without the fusion to FliO.

8.4 The FlhA/SctV membrane domain and opening of the export apparatus

The structure of the membrane domain of FlhA/SctV, which may be important for understanding how the export gate opens, remains as the main structural challenge. All structures of FliPQR/SctRST described in this thesis were in a closed state in which multiple closure points were observed (Figure 8.3). Although an open state could be modelled in which the long, kinked helices straighten, opening the entrance to the lumen of the complex towards secretion substrates in the cytoplasm, the structure of the open state remains to be determined experimentally. The structure of the FliPQR complex bound by FlhB suggests that a network of interactions in the complex allows force exerted by FlhA/SctV on the export gate complex to trigger opening. The membrane domain of FlhA/SctV may also carry out additional gating of the T3SS and is in close proximity to the FliPQR/SctRST complex (Figure 8.1).

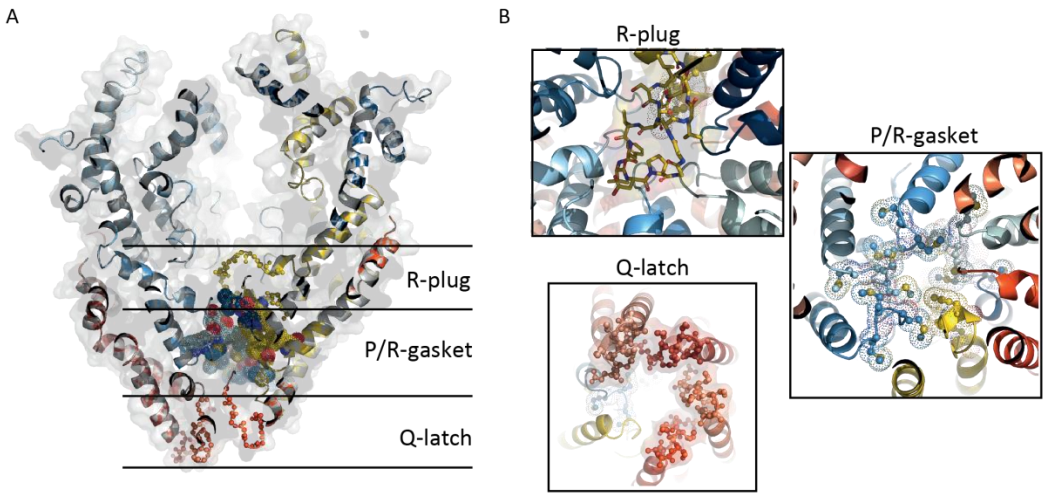


Figure 8.3 The export gate is closed at multiple points.
A, B, Closure points in the structure of FliPQR (chapter 4).

Furthermore, it is not known whether the export gate opens and closes dynamically in cycles of secretion or whether it remains open throughout secretion. Flagellar secretion has been proposed to function by forming a chain of substrates (Evans et al. 2013), which may block the export gate from closing once it has opened. Recently, an injection-diffusion model has gained favour (Renault et al. 2017), which might involve repeated opening and closing of the gate. However, a permanently open gate may facilitate the very high rate of protein secretion in T3SS. This structure of the open state remains to be elucidated. One possibility may be to employ a GFP fusion of a secretion substrate to block the export gate (Radics et al. 2014) and determine the structure of the export gate in the assembled needle complex, which is possible with current technology (Goessweiner-Mohr et al. 2019).

The differences in shape of the FlhA/SctV membrane domain in different samples that were imaged (chapter 7) are likely due to the removal of the protein from the membrane by its solubilisation in detergent. Imaging of multiple homologues in membrane mimetic systems (Autzen et al. 2019), if it is possible to prepare such samples, may reveal the structure of the membrane domain. Other possibilities include using additives to stabilise the protein during the grid preparation process, increasing the speed of grid preparation or adsorbing the protein onto a surface (Glaeser 2018). A different form of stabilisation would come from the preparation of a larger complex, for

example the entire export apparatus or FlhA/SctV in complex with a basal body protein surrounding it.

8.5 Structure prediction of membrane protein complexes

The most unexpected aspect of the structure of FliPQR was that many of the predicted transmembrane helices were not found in the membrane (chapter 4). Instead, many hydrophobic residues are buried in protein-protein interactions. Although it is not clear how widespread such complexes are, the example of FliPQR/SctRST is sufficient to serve as a warning in the interpretation of topology predictions of membrane proteins unrelated to known structures. In addition, the many experimental studies interrogating the topology of the export gate subunits, for example (Berger et al. 2010; Dietsche et al. 2016), failed to find the correct topology of the subunits. One reason for this may be that the subunits cannot be analysed in isolation without taking into account the extensive interactions with other components of the complex. This may be relevant in other complex membrane proteins.

The unexpected structure and location of predicted transmembrane helices leads to the question of assembly and synthesis of such proteins. Are the stretches of hydrophobic residues recognised as potential transmembrane helices by human algorithms also recognised as transmembrane helices when the export gate subunits are synthesised and inserted into the membrane? One possibility is that through an unknown mechanism it is recognised that they are not transmembrane helices. Another is that they are inserted into the membrane and that this state represents a protein folding intermediate. The subunits may only adopt their final structures as they interact with each other and co-operatively assemble into a complex.

8.6 Conclusion

In summary, the structures in this thesis have upset the previous model of an export gate in a patch of inner membrane (Macnab 2004) and instead revealed its true localisation, with many implications for its function, in flagella and injectisomes. The membrane domain of the secretion specificity switch protein has been shown to have a

role in secretion distinct from substrate specificity and the foundations for further structural studies of the last remaining structurally uncharacterised domain in the export apparatus have been laid.

9 References

Abrusci P, Vergara-Irigaray M, Johnson S, Beeby MD, Hendrixson DR, Roversi P, Friede ME, Deane JE, Jensen GJ, Tang CM et al. 2013. Architecture of the major component of the type III secretion system export apparatus. *Nature structural & molecular biology* **20**: 99-104.

Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, Adams PD. 2018. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta crystallographica Section D, Structural biology* **74**: 531-544.

Aizawa SI. 2001. Bacterial flagella and type III secretion systems. *Fems Microbiol Lett* **202**: 157-164.

Aizawa SI, Dean GE, Jones CJ, Macnab RM, Yamaguchi S. 1985. Purification and characterization of the flagellar hook-basal body complex of Salmonella typhimurium. *Journal of bacteriology* **161**: 836-849.

Akiba T, Yoshimura H, Namba K. 1991. Monolayer crystallization of flagellar L-P rings by sequential addition and depletion of lipid. *Science* **252**: 1544-1546.

Allaoui A, Sansonetti PJ, Parsot C. 1992. MxiJ, a lipoprotein involved in secretion of *Shigella* Ipa invasins, is homologous to YscJ, a secretion factor of the *Yersinia* Yop proteins. *Journal of bacteriology* **174**: 7661-7669.

Almagro Armenteros JJ, Tsirigos KD, Sonderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature biotechnology* **37**: 420-423.

Anderson DM, Schneewind O. 1997. A mRNA signal for the type III secretion of Yop proteins by *Yersinia enterocolitica*. *Science* **278**: 1140-1143.

Archer CT, Kim JF, Jeong H, Park JH, Vickers CE, Lee SY, Nielsen LK. 2011. The genome sequence of *E. coli* W (ATCC 9637): comparative genome analysis and an improved genome-scale reconstruction of *E. coli*. *BMC genomics* **12**: 9.

Asakura S. 1970. Polymerization of flagellin and polymorphism of flagella. *Advances in biophysics* **1**: 99-155.

Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N. 2016. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic acids research* **44**: W344-350.

Ates LS, Houben EN, Bitter W. 2016. Type VII Secretion: A Highly Versatile Secretion System. *Microbiology spectrum* **4**.

Autzen HE, Julius D, Cheng Y. 2019. Membrane mimetic systems in CryoEM: keeping membrane proteins in their native environment. *Current opinion in structural biology*.

Auvray F, Thomas J, Fraser GM, Hughes C. 2001. Flagellin polymerisation control by a cytosolic export chaperone. *Journal of molecular biology* **308**: 221-229.

Babic A, Lindner AB, Vulic M, Stewart EJ, Radman M. 2008. Direct visualization of horizontal gene transfer. *Science* **319**: 1533-1536.

Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 10037-10041.

Bange G, Kummerer N, Engel C, Bozkurt G, Wild K, Sinning I. 2010. FlhA provides the adaptor for coordinated delivery of late flagella building blocks to the type III secretion system. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 11295-11300.

Barker CS, Inoue T, Meshcheryakova IV, Kitanobo S, Samatey FA. 2016. Function of the conserved FHIPEP domain of the flagellar type III export apparatus, protein FlhA. *Molecular microbiology* **100**: 278-288.

Barker CS, Meshcheryakova IV, Kostyukova AS, Samatey FA. 2010. FliO regulation of FliP in the formation of the *Salmonella enterica* flagellum. *PLoS genetics* **6**: e1001143.

Barnhart MM, Chapman MR. 2006. Curli biogenesis and function. *Annual review of microbiology* **60**: 131-147.

Bartsev AV, Deakin WJ, Boukli NM, McAlvin CB, Stacey G, Malnoe P, Broughton WJ, Staehelin C. 2004. NopL, an effector protein of *Rhizobium* sp. NGR234, thwarts activation of plant defense reactions. *Plant physiology* **134**: 871-879.

Baumgarten T, Schlegel S, Wagner S, Low M, Eriksson J, Bonde I, Herrgard MJ, Heipieper HJ, Norholm MH, Slotboom DJ et al. 2017. Isolation and characterization of the *E. coli* membrane protein production strain Mutant56(DE3). *Scientific reports* **7**: 45089.

Bechara C, Noll A, Morgner N, Degiacomi MT, Tampe R, Robinson CV. 2015. A subset of annular lipids is linked to the flippase activity of an ABC transporter. *Nature chemistry* **7**: 255-262.

Beckham KS, Ciccarelli L, Bunduc CM, Mertens HD, Ummels R, Lugmayr W, Mayr J, Rettel M, Savitski MM, Svergun DI et al. 2017. Structure of the mycobacterial ESX-5 type VII secretion system membrane complex by single-particle analysis. *Nature microbiology* **2**: 17047.

Beebe JM. 1941. The Morphology and Cytology of *Myxococcus xanthus*, N. Sp. *Journal of bacteriology* **42**: 193-223.

Beeby M, Ribardo DA, Brennan CA, Ruby EG, Jensen GJ, Hendrixson DR. 2016. Diverse high-torque bacterial flagellar motors assemble wider stator rings using a conserved protein scaffold. *Proceedings of the National Academy of Sciences of the United States of America* **113**: E1917-1926.

Berg HC. 1974. Dynamic properties of bacterial flagellar motors. *Nature* **249**: 77-79.

Berg HC. 2003. The rotary motor of bacterial flagella. *Annual review of biochemistry* **72**: 19-54.

Berg HC, Brown DA. 1972. Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking. *Nature* **239**: 500-504.

Berger C, Robin GP, Bonas U, Koebnik R. 2010. Membrane topology of conserved components of the type III secretion system from the plant pathogen *Xanthomonas campestris* pv. vesicatoria. *Microbiology* **156**: 1963-1974.

Bergeron JR. 2016. Structural modeling of the flagellum MS ring protein FliF reveals similarities to the type III secretion system and sporulation complex. *PeerJ* **4**: e1718.

Bergeron JR, Fernandez L, Wasney GA, Vuckovic M, Reffuveille F, Hancock RE, Strynadka NC. 2016. The Structure of a Type 3 Secretion System (T3SS) Ruler Protein Suggests a Molecular Mechanism for Needle Length Sensing. *The Journal of biological chemistry* **291**: 1676-1691.

Bhattacharya S, Baidya AK, Pal RR, Mamou G, Gatt YE, Margalit H, Rosenshine I, Ben-Yehuda S. 2019. A Ubiquitous Platform for Bacterial Nanotube Biogenesis. *Cell reports* **27**: 334-342 e310.

Bjornfot AC, Lavander M, Forsberg A, Wolf-Watz H. 2009. Autoproteolysis of YscU of *Yersinia pseudotuberculosis* is important for regulation of expression and secretion of Yop proteins. *Journal of bacteriology* **191**: 4259-4267.

Bock D, Medeiros JM, Tsao HF, Penz T, Weiss GL, Aistleitner K, Horn M, Pilhofer M. 2017. In situ architecture, function, and evolution of a contractile injection system. *Science* **357**: 713-717.

Bolhuis A, Mathers JE, Thomas JD, Barrett CM, Robinson C. 2001. TatB and TatC form a functional and structural unit of the twin-arginine translocase from *Escherichia coli*. *The Journal of biological chemistry* **276**: 20213-20219.

Bos KI, Schuenemann VJ, Golding GB, Burbano HA, Waglechner N, Coombes BK, McPhee JB, DeWitte SN, Meyer M, Schmedes S et al. 2011. A draft genome of *Yersinia pestis* from victims of the Black Death. *Nature* **478**: 506-510.

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

Burkinshaw BJ, Deng W, Lameignere E, Wasney GA, Zhu H, Worrall LJ, Finlay BB, Strynadka NC. 2015. Structural analysis of a specialized type III secretion system peptidoglycan-cleaving enzyme. *The Journal of biological chemistry* **290**: 10406-10417.

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

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

Castellani A, Chalmers AJ. 1919. *Manual of Tropical Medicine, 3rd ed.* Williams Wood and Co., New York.

Celia H, Noinaj N, Zakharov SD, Bordignon E, Botos I, Santamaria M, Barnard TJ, Cramer WA, Lloubes R, Buchanan SK. 2016. Structural insight into the role of the Ton complex in energy transduction. *Nature* **538**: 60-65.

Chae PS, Rasmussen SG, Rana RR, Gotfryd K, Chandra R, Goren MA, Kruse AC, Nurva S, Loland CJ, Pierre Y et al. 2010. Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nature methods* **7**: 1003-1008.

Chae PS, Rasmussen SG, Rana RR, Gotfryd K, Kruse AC, Manglik A, Cho KH, Nurva S, Gether U, Guan L et al. 2012. A new class of amphiphiles bearing rigid hydrophobic groups for solubilization and stabilization of membrane proteins. *Chemistry* **18**: 9485-9490.

Chandran Darbari V, Waksman G. 2015. Structural Biology of Bacterial Type IV Secretion Systems. *Annual review of biochemistry* **84**: 603-629.

Cheung M, Shen DK, Makino F, Kato T, Roehrich AD, Martinez-Argudo I, Walker ML, Murillo I, Liu X, Pain M et al. 2015. Three-dimensional electron microscopy reconstruction and cysteine-mediated crosslinking provide a model of the type III secretion system needle tip complex. *Molecular microbiology* **95**: 31-50.

Chung KY, Kim TH, Manglik A, Alvares R, Kobilka BK, Prosser RS. 2012. Role of detergents in conformational exchange of a G protein-coupled receptor. *The Journal of biological chemistry* **287**: 36305-36311.

Coburn B, Sekirov I, Finlay BB. 2007. Type III secretion systems and disease. *Clinical microbiology reviews* **20**: 535-549.

Cordes FS, Komoriya K, Larquet E, Yang S, Egelman EH, Blocker A, Lea SM. 2003. Helical structure of the needle of the type III secretion system of *Shigella flexneri*. *The Journal of biological chemistry* **278**: 17103-17107.

Cowtan K. 2006. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta crystallographica Section D, Biological crystallography* **62**: 1002-1011.

Craig DB, Dombkowski AA. 2013. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC bioinformatics* **14**: 346.

D'Imprima E, Floris D, Joppe M, Sanchez R, Grininger M, Kuhlbrandt W. 2019. Protein denaturation at the air-water interface and how to prevent it. *eLife* **8**.

Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 6640-6645.

Davis BR, Fanning GR, Madden JM, Steigerwalt AG, Bradford HB, Jr., Smith HL, Jr., Brenner DJ. 1981. Characterization of biochemically atypical *Vibrio cholerae* strains and designation of a new pathogenic species, *Vibrio mimicus*. *Journal of clinical microbiology* **14**: 631-639.

De Carlo S, Adrian M, Kalin P, Mayer JM, Dubochet J. 1999. Unexpected property of trehalose as observed by cryo-electron microscopy. *Journal of microscopy* **196**: 40-45.

Deane JE, Graham SC, Mitchell EP, Flot D, Johnson S, Lea SM. 2008. Crystal structure of Spa40, the specificity switch for the *Shigella flexneri* type III secretion system. *Molecular microbiology* **69**: 267-276.

Delalez NJ, Berry RM, Armitage JP. 2014. Stoichiometry and turnover of the bacterial flagellar switch protein FliN. *mBio* **5**: e01216-01214.

Delalez NJ, Wadhams GH, Rosser G, Xue Q, Brown MT, Dobbie IM, Berry RM, Leake MC, Armitage JP. 2010. Signal-dependent turnover of the bacterial flagellar switch protein FliM. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 11347-11351.

Deng W, Marshall NC, Rowland JL, McCoy JM, Worrall LJ, Santos AS, Strynadka NCJ, Finlay BB. 2017. Assembly, structure, function and regulation of type III secretion systems. *Nature reviews Microbiology* **15**: 323-337.

Devignat R. 1951. Varieties of *Pasteurella pestis*; new hypothesis. *Bulletin of the World Health Organization* **4**: 247-263.

Diepold A, Amstutz M, Abel S, Sorg I, Jenal U, Cornelis GR. 2010. Deciphering the assembly of the *Yersinia* type III secretion injectisome. *The EMBO journal* **29**: 1928-1940.

Diepold A, Armitage JP. 2015. Type III secretion systems: the bacterial flagellum and the injectisome. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* **370**.

Diepold A, Wagner S. 2014. Assembly of the bacterial type III secretion machinery. *FEMS microbiology reviews* **38**: 802-822.

Diepold A, Wiesand U, Cornelis GR. 2011. The assembly of the export apparatus (YscR,S,T,U,V) of the *Yersinia* type III secretion apparatus occurs independently of other structural components and involves the formation of an YscV oligomer. *Molecular microbiology* **82**: 502-514.

Dietsche T, Tesfazgi Mebrhatu M, Brunner MJ, Abrusci P, Yan J, Franz-Wachtel M, Scharfe C, Zilkenat S, Grin I, Galan JE et al. 2016. Structural and Functional Characterization of the Bacterial Type III Secretion Export Apparatus. *PLoS pathogens* **12**: e1006071.

Duan Q, Zhou M, Zhu L, Zhu G. 2013. Flagella and bacterial pathogenicity. *Journal of basic microbiology* **53**: 1-8.

Ehrenberg CG. 1838. Die Infusionsthierchen als vollkommene Organismen: ein Blick in das tiefere organische Leben der Natur.

Emerson SU, Tokuyasu K, Simon MI. 1970. Bacterial flagella: polarity of elongation. *Science* **169**: 190-192.

Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta crystallographica Section D, Biological crystallography* **66**: 486-501.

Engelmann TW. 1881. Neue Methode zur Untersuchung der Sauerstoffausscheidung pflanzlicher und thierischer Organismen. *Pflüger, Arch* **25**: 285-292.

Erhardt M, Hirano T, Su Y, Paul K, Wee DH, Mizuno S, Aizawa S, Hughes KT. 2010. The role of the FliK molecular ruler in hook-length control in *Salmonella enterica*. *Molecular microbiology* **75**: 1272-1284.

Erhardt M, Wheatley P, Kim EA, Hirano T, Zhang Y, Sarkar MK, Hughes KT, Blair DF. 2017. Mechanism of type-III protein secretion: Regulation of FlhA conformation by a functionally critical charged-residue cluster. *Molecular microbiology* **104**: 234-249.

Evans LD, Poulter S, Terentjev EM, Hughes C, Fraser GM. 2013. A chain mechanism for flagellum growth. *Nature* **504**: 287-290.

Everett KD, Bush RM, Andersen AA. 1999. Emended description of the order *Chlamydiales*, proposal of *Parachlamydiaceae* fam. nov. and *Simkaniaceae* fam. nov., each containing one monotypic genus, revised taxonomy of the family *Chlamydiaceae*, including a new genus and five new species, and standards for the identification of organisms. *International journal of systematic bacteriology* **49 Pt 2**: 415-440.

Fabiani FD, Renault TT, Peters B, Dietsche T, Galvez EJC, Guse A, Freier K, Charpentier E, Strowig T, Franz-Wachtel M et al. 2017. A flagellum-specific chaperone facilitates assembly of the core type III export apparatus of the bacterial flagellum. *PLoS biology* **15**: e2002267.

Fan F, Ohnishi K, Francis NR, Macnab RM. 1997. The FliP and FliR proteins of *Salmonella typhimurium*, putative components of the type III flagellar export apparatus, are located in the flagellar basal body. *Molecular microbiology* **26**: 1035-1046.

Ferreira JL, Gao FZ, Rossmann FM, Nans A, Brenzinger S, Hosseini R, Wilson A, Briegel A, Thormann KM, Rosenthal PB et al. 2019. gamma-proteobacteria eject their polar flagella under nutrient depletion, retaining flagellar motor relic structures. *PLoS biology* **17**: e3000165.

Ferris HU, Furukawa Y, Minamino T, Kroetz MB, Kihara M, Namba K, Macnab RM. 2005. FlhB regulates ordered export of flagellar components via autocleavage mechanism. *The Journal of biological chemistry* **280**: 41236-41242.

Francis NR, Irikura VM, Yamaguchi S, DeRosier DJ, Macnab RM. 1992. Localization of the *Salmonella typhimurium* flagellar switch protein FliG to the cytoplasmic M-ring face of the basal body. *Proceedings of the National Academy of Sciences of the United States of America* **89**: 6304-6308.

Fraser GM, Hirano T, Ferris HU, Devgan LL, Kihara M, Macnab RM. 2003. Substrate specificity of type III flagellar protein export in *Salmonella* is controlled by subdomain interactions in FlhB. *Molecular microbiology* **48**: 1043-1057.

Frederiksen W. 1964. A study of some *Yersinia pseudotuberculosis*-like bacteria (*Bacterium enterocoliticum* and *Pasteurella* X). *Proc XIV Scand Cong Pathol Microbiol*: 103-104.

Fronzes R, Remaut H, Waksman G. 2008. Architectures and biogenesis of non-flagellar protein appendages in Gram-negative bacteria. *The EMBO journal* **27**: 2271-2280.

Frost S, Ho O, Login FH, Weise CF, Wolf-Watz H, Wolf-Watz M. 2012. Autoproteolysis and intramolecular dissociation of *Yersinia* YscU precedes secretion of its C-terminal polypeptide YscU(CC). *PloS one* **7**: e49349.

Fujii T, Kato T, Hiraoka KD, Miyata T, Minamino T, Chevance FF, Hughes KT, Namba K. 2017. Identical folds used for distinct mechanical functions of the bacterial flagellar rod and hook. *Nat Commun* **8**: 14276.

Fukumura T, Furukawa Y, Kawaguchi T, Saijo-Hamano Y, Namba K, Imada K, Minamino T. 2014. Crystallization and preliminary X-ray analysis of the periplasmic domain of FliP, an integral membrane component of the bacterial flagellar type III protein-export apparatus. *Acta crystallographica Section F, Structural biology communications* **70**: 1215-1218.

Fukumura T, Makino F, Dietsche T, Kinoshita M, Kato T, Wagner S, Namba K, Imada K, Minamino T. 2017. Assembly and stoichiometry of the core structure of the bacterial flagellar type III export gate complex. *PLoS biology* **15**: e2002281.

Fukuoka H, Inoue Y, Terasawa S, Takahashi H, Ishijima A. 2010. Exchange of rotor components in functioning bacterial flagellar motor. *Biochemical and biophysical research communications* **394**: 130-135.

Galan JE, Curtiss R, 3rd. 1989. Cloning and molecular characterization of genes whose products allow *Salmonella typhimurium* to penetrate tissue culture cells. *Proceedings of the National Academy of Sciences of the United States of America* **86**: 6383-6387.

Galan JE, Waksman G. 2018. Protein-Injection Machines in Bacteria. *Cell* **172**: 1306-1318.

Gault J, Donlan JA, Liko I, Hopper JT, Gupta K, Housden NG, Struwe WB, Marty MT, Mize T, Bechara C et al. 2016. High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nature methods* **13**: 333-336.

Glaeser RM. 2018. Proteins, Interfaces, and Cryo-Em Grids. *Current opinion in colloid & interface science* **34**: 1-8.

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

Goguen JD, Yother J, Straley SC. 1984. Genetic analysis of the low calcium response in *Yersinia pestis* mu d1(Ap lac) insertion mutants. *Journal of bacteriology* **160**: 842-848.

Gorter E, Grendel F. 1925. On Bimolecular Layers of Lipoids on the Chromocytes of the Blood. *The Journal of experimental medicine* **41**: 439-443.

Goyal P, Krasteva PV, Van Gerven N, Gubellini F, Van den Broeck I, Troupiotis-Tsailaki A, Jonckheere W, Pehau-Arnaudet G, Pinkner JS, Chapman MR et al. 2014. Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG. *Nature* **516**: 250-253.

Green ER, Mecsas J. 2016. Bacterial Secretion Systems: An Overview. *Microbiology spectrum* **4**.

Guzman LM, Belin D, Carson MJ, Beckwith J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of bacteriology* **177**: 4121-4130.

Hakansson S, Schesser K, Persson C, Galyov EE, Rosqvist R, Homble F, Wolf-Watz H. 1996. The YopB protein of *Yersinia pseudotuberculosis* is essential for the translocation of Yop effector proteins across the target cell plasma membrane and displays a contact-dependent membrane disrupting activity. *The EMBO journal* **15**: 5812-5823.

Hall Z, Hernandez H, Marsh JA, Teichmann SA, Robinson CV. 2013. The role of salt bridges, charge density, and subunit flexibility in determining disassembly routes of protein complexes. *Structure* **21**: 1325-1337.

Hara N, Namba K, Minamino T. 2011. Genetic characterization of conserved charged residues in the bacterial flagellar type III export protein FlhA. *PloS one* **6**: e22417.

Hardy D, Desuzinges Mandon E, Rothnie AJ, Jawhari A. 2018. The yin and yang of solubilization and stabilization for wild-type and full-length membrane protein. *Methods* **147**: 118-125.

Hauer F, Gerle C, Fischer N, Oshima A, Shinzawa-Itoh K, Shimada S, Yokoyama K, Fujiyoshi Y, Stark H. 2015. GraDeR: Membrane Protein Complex Preparation for Single-Particle Cryo-EM. *Structure* **23**: 1769-1775.

Hay ID, Belousoff MJ, Lithgow T. 2017. Structural Basis of Type 2 Secretion System Engagement between the Inner and Outer Bacterial Membranes. *mBio* **8**.

Hayashi F, Smith KD, Ozinsky A, Hawn TR, Yi EC, Goodlett DR, Eng JK, Akira S, Underhill DM, Aderem A. 2001. The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* **410**: 1099-1103.

Hayward AC, Waterston JM. 1965. *Pseudomonas phaseolicola. CMI Descr Pathog Fungi Bact* **45**: 1-2.

Hayward RD, Koronakis V. 1999. Direct nucleation and bundling of actin by the SipC protein of invasive *Salmonella*. *The EMBO journal* **18**: 4926-4934.

Hernandez H, Robinson CV. 2007. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nature protocols* **2**: 715-726.

Hu B, Lara-Tejero M, Kong Q, Galan JE, Liu J. 2017. In Situ Molecular Architecture of the *Salmonella* Type III Secretion Machine. *Cell* **168**: 1065-1074 e1010.

Hu B, Morado DR, Margolin W, Rohde JR, Arizmendi O, Picking WL, Picking WD, Liu J. 2015. Visualization of the type III secretion sorting platform of *Shigella flexneri*. *Proceedings of the National Academy of Sciences of the United States of America* **112**: 1047-1052.

Hu J, Worrall LJ, Hong C, Vuckovic M, Atkinson CE, Caveney N, Yu Z, Strynadka NCJ. 2018. Cryo-EM analysis of the T3S injectisome reveals the structure of the needle and open secretin. *Nat Commun* **9**: 3840.

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

Hueck CJ. 1998. Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiology and molecular biology reviews : MMBR* **62**: 379-433.

Hughes KT. 2012. Flagellar hook length is controlled by a secreted molecular ruler. *Journal of bacteriology* **194**: 4793-4796.

Hussain H, Du Y, Scull NJ, Mortensen JS, Tarrasch J, Bae HE, Loland CJ, Byrne B, Kobilka BK, Chae PS. 2016. Accessible Mannitol-Based Amphiphiles (MNAs) for Membrane Protein Solubilisation and Stabilisation. *Chemistry* **22**: 7068-7073.

Ide T, Laarmann S, Greune L, Schillers H, Oberleithner H, Schmidt MA. 2001. Characterization of translocation pores inserted into plasma membranes by type III-secreted Esp proteins of enteropathogenic *Escherichia coli*. *Cellular microbiology* **3**: 669-679.

Iino T. 1969. Polarity of flagellar growth in *Salmonella*. *Journal of general microbiology* **56**: 227-239.

Imada K, Minamino T, Tahara A, Namba K. 2007. Structural similarity between the flagellar type III ATPase FliI and F1-ATPase subunits. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 485-490.

Imada K, Minamino T, Uchida Y, Kinoshita M, Namba K. 2016. Insight into the flagella type III export revealed by the complex structure of the type III ATPase and its regulator. *Proceedings of the National Academy of Sciences of the United States of America* **113**: 3633-3638.

Inoue Y, Kinoshita M, Namba K, Minamino T. 2019a. Mutational analysis of the C-terminal cytoplasmic domain of FlhB, a transmembrane component of the flagellar type III protein export apparatus in Salmonella. *Genes to cells : devoted to molecular & cellular mechanisms*.

Inoue Y, Ogawa Y, Kinoshita M, Terahara N, Shimada M, Kodera N, Ando T, Namba K, Kitao A, Imada K et al. 2019b. Structural Insights into the Substrate Specificity Switch Mechanism of the Type III Protein Export Apparatus. *Structure* **27**: 965-976 e966.

Johnson S, Fong YH, Deme J, Furlong E, Kuhlen L, Lea SM. 2019. Structure of the bacterial flagellar rotor MS-ring: a minimum inventory/maximum diversity system. *Biorxiv*.

Jones CJ, Homma M, Macnab RM. 1987. Identification of proteins of the outer (L and P) rings of the flagellar basal body of *Escherichia coli*. *Journal of bacteriology* **169**: 1489-1492.

Jones CJ, Macnab RM. 1990. Flagellar assembly in *Salmonella typhimurium*: analysis with temperature-sensitive mutants. *Journal of bacteriology* **172**: 1327-1339.

Jones CJ, Macnab RM, Okino H, Aizawa S. 1990. Stoichiometric analysis of the flagellar hook-(basal-body) complex of *Salmonella typhimurium*. *Journal of molecular biology* **212**: 377-387.

Josenhans C, Suerbaum S. 2002. The role of motility as a virulence factor in bacteria. *International journal of medical microbiology : IJMM* **291**: 605-614.

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

Kaplan M, Ghosal D, Subramanian P, Oikonomou CM, Kjaer A, Pirbadian S, Ortega DR, Briegel A, El-Naggar MY, Jensen GJ. 2019. The presence and absence of periplasmic rings in bacterial flagellar motors correlates with stator type. *eLife* **8**.

Kaplan M, Subramanian P, Ghosal D, Oikonomou CM, Pirbadian S, Starwalt-Lee R, Gralnick JA, El-Naggar MY, Jensen GJ. 2018. Stable sub-complexes observed in situ suggest a modular assembly pathway of the bacterial flagellar motor. *biorxiv*.

Kerschen EJ, Cohen DA, Kaplan AM, Straley SC. 2004. The plague virulence protein YopM targets the innate immune response by causing a global depletion of NK cells. *Infection and immunity* **72**: 4589-4602.

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

Kimanius D, Forsberg BO, Scheres SH, Lindahl E. 2016. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* **5**.

Kimbrough TG, Miller SI. 2000. Contribution of *Salmonella typhimurium* type III secretion components to needle complex formation. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 11008-11013.

Kitasato S. 1894. THE BACILLUS OF BUBONIC PLAGUE. *Lancet* **144**: 428-430.

Koronakis V, Sharff A, Koronakis E, Luisi B, Hughes C. 2000. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* **405**: 914-919.

Korotkov KV, Sandkvist M. 2019. Architecture, Function, and Substrates of the Type II Secretion System. *EcoSal Plus* **8**.

Korotkov KV, Sandkvist M, Hol WG. 2012. The type II secretion system: biogenesis, molecular architecture and mechanism. *Nature reviews Microbiology* **10**: 336-351.

Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of molecular biology* **305**: 567-580.

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

Kubori T, Shimamoto N, Yamaguchi S, Namba K, Aizawa S. 1992. Morphological pathway of flagellar assembly in *Salmonella typhimurium*. *Journal of molecular biology* **226**: 433-446.

Kudryashev M, Cyrklaff M, Wallich R, Baumeister W, Frischknecht F. 2010. Distinct in situ structures of the *Borrelia* flagellar motor. *Journal of structural biology* **169**: 54-61.

Kupferberg LL, Higuchi K. 1958. Role of calcium ions in the stimulation of growth of virulent strains of *Pasteurella pestis*. *Journal of bacteriology* **76**: 120-121.

Lange R, Hengge-Aronis R. 1991. Growth phase-regulated expression of bolA and morphology of stationary-phase *Escherichia coli* cells are controlled by the novel sigma factor sigma S. *Journal of bacteriology* **173**: 4474-4481.

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

Lara-Tejero M, Qin Z, Hu B, Butan C, Liu J, Galan JE. 2019. Role of SpaO in the assembly of the sorting platform of a *Salmonella* type III secretion system. *PLoS pathogens* **15**: e1007565.

Larsen SH, Adler J, Gargus JJ, Hogg RW. 1974a. Chemomechanical coupling without ATP: the source of energy for motility and chemotaxis in bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **71**: 1239-1243.

Larsen SH, Reader RW, Kort EN, Tso WW, Adler J. 1974b. Change in direction of flagellar rotation is the basis of the chemotactic response in *Escherichia coli*. *Nature* **249**: 74-77.

Lasica AM, Ksiazek M, Madej M, Potempa J. 2017. The Type IX Secretion System (T9SS): Highlights and Recent Insights into Its Structure and Function. *Frontiers in cellular and infection microbiology* **7**: 215.

Lauber F, Deme JC, Lea SM, Berks BC. 2018. Type 9 secretion system structures reveal a new protein transport mechanism. *Nature* **564**: 77-82.

Le Minor L, Popoff MY. 1987. Designation of *Salmonella enterica* sp. nov., nom. rev., as the Type and Only Species of the Genus *Salmonella*. *International journal of systematic and evolutionary microbiology* **37**: 465-468.

Lee KK, Jang CS, Yoon JY, Kim SY, Kim TH, Ryu KH, Kim W. 2008. Abnormal cell division caused by inclusion bodies in *E. coli*; increased resistance against external stress. *Microbiological research* **163**: 394-402.

Lee PC, Rietsch A. 2015. Fueling type III secretion. *Trends in microbiology* **23**: 296-300.

Lee VT, Schneewind O. 2001. Protein secretion and the pathogenesis of bacterial infections. *Genes & development* **15**: 1725-1752.

Leyton DL, Rossiter AE, Henderson IR. 2012. From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nature reviews Microbiology* **10**: 213-225.

Li H, Sourjik V. 2011. Assembly and stability of flagellar motor in *Escherichia coli*. *Molecular microbiology* **80**: 886-899.

Liko I, Hopper JT, Allison TM, Benesch JL, Robinson CV. 2016. Negative Ions Enhance Survival of Membrane Protein Complexes. *Journal of the American Society for Mass Spectrometry* **27**: 1099-1104.

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

Lim HN, Lee Y, Hussein R. 2011. Fundamental relationship between operon organization and gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 10626-10631.

Lin DY, Huang S, Chen J. 2015. Crystal structures of a polypeptide processing and secretion transporter. *Nature* **523**: 425-430.

Lindner AB, Madden R, Demarez A, Stewart EJ, Taddei F. 2008. Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 3076-3081.

Liu Z, Miyashiro T, Tsou A, Hsiao A, Goulian M, Zhu J. 2008. Mucosal penetration primes *Vibrio cholerae* for host colonization by repressing quorum sensing. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 9769-9774.

Lobstein J, Emrich CA, Jeans C, Faulkner M, Riggs P, Berkmen M. 2012. SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microbial cell factories* **11**: 56.

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

Lorenzen K, Versluis C, van Duijn E, van den Heuvel RHH, Heck AJR. 2007. Optimizing macromolecular tandem mass spectrometry of large non-covalent complexes using heavy collision gases. *Int J Mass Spectrom* **268**: 198-206.

Lountos GT, Austin BP, Nallamsetty S, Waugh DS. 2009. Atomic resolution structure of the cytoplasmic domain of *Yersinia pestis* YscU, a regulatory switch involved in type III secretion. *Protein science : a publication of the Protein Society* **18**: 467-474.

Low HH, Gubellini F, Rivera-Calzada A, Braun N, Connery S, Dujeancourt A, Lu F, Redzej A, Fronzes R, Orlova EV et al. 2014. Structure of a type IV secretion system. *Nature* **508**: 550-553.

Lowy J, Hanson J. 1965. Electron Microscope Studies of Bacterial Flagella. *Journal of molecular biology* **11**: 293-313.
Lynch MJ, Levenson R, Kim EA, Sircar R, Blair DF, Dahlquist FW, Crane BR. 2017. Co-Folding of a FliF-FliG Split Domain Forms the Basis of the MS:C Ring Interface within the Bacterial Flagellar Motor. *Structure* **25**: 317-328.
Lyons S, Wang L, Casanova JE, Sitaraman SV, Merlin D, Gewirtz AT. 2004. *Salmonella typhimurium* transcytoses flagellin via an SPI2-mediated vesicular transport pathway. *Journal of cell science* **117**: 5771-5780.
Macnab R, Koshland DE, Jr. 1974. Bacterial motility and chemotaxis: light-induced tumbling response and visualization of individual flagella. *Journal of molecular biology* **84**: 399-406.
Macnab RM. 1999. The bacterial flagellum: reversible rotary propellor and type III export apparatus. *Journal of bacteriology* **181**: 7149-7153.
Macnab RM. 2003. How bacteria assemble flagella. *Annual review of microbiology* **57**: 77-100.
Macnab RM. 2004. Type III flagellar protein export and flagellar assembly. *Biochimica et biophysica acta* **1694**: 207-217.
Majewski DD, Worrall LJ, Hong C, Atkinson CE, Vuckovic M, Watanabe N, Yu Z, Strynadka NCJ. 2019. Cryo-EM structure of the homohexameric T3SS ATPase-central stalk complex reveals rotary ATPase-like asymmetry. *Nat Commun* **10**: 626.
Majewski DD, Worrall LJ, Strynadka NC. 2018. Secretins revealed: structural insights into the giant gated outer membrane portals of bacteria. *Current opinion in structural biology* **51**: 61-72.
Maki-Yonekura S, Matsuoka R, Yamashita Y, Shimizu H, Tanaka M, Iwabuki F, Yonekura K. 2018. Hexameric and pentameric complexes of the ExbBD energizer in the Ton system. *eLife* **7**.
Maki-Yonekura S, Yonekura K, Namba K. 2010. Conformational change of flagellin for polymorphic supercoiling of the flagellar filament. *Nature structural & molecular biology* **17**: 417-422.
Makino K, Oshima K, Kurokawa K, Yokoyama K, Uda T, Tagomori K, Iijima Y, Najima M, Nakano M, Yamashita A et al. 2003. Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism distinct from that of V cholerae. *Lancet* **361**: 743-749.
Malakooti J, Ely B, Matsumura P. 1994. Molecular characterization, nucleotide sequence, and expression of the fliO, fliP, fliQ, and fliR genes of Escherichia coli. *Journal of bacteriology* **176**: 189-197.
Manson MD, Tedesco P, Berg HC, Harold FM, Van der Drift C. 1977. A protonmotive force drives bacterial flagella. *Proceedings of the National Academy of Sciences of the United States of America* **74**: 3060-3064.
Marketon MM, DePaolo RW, DeBord KL, Jabri B, Schneewind O. 2005. Plague bacteria target immune cells during infection. *Science* **309**: 1739-1741.
Marlovits TC, Kubori T, Lara-Tejero M, Thomas D, Unger VM, Galan JE. 2006. Assembly of the inner rod determines needle length in the type III secretion injectisome. *Nature* **441**: 637-640.
Marty MT, Baldwin AJ, Marklund EG, Hochberg GK, Benesch JL, Robinson CV. 2015. Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Analytical chemistry* **87**: 4370-4376.
Matsunami H, Barker CS, Yoon YH, Wolf M, Samatey FA. 2016. Complete structure of the bacterial flagellar hook reveals extensive set of stabilizing interactions. *Nat Commun* **7**: 13425.
Mattick JS. 2002. Type IV pili and twitching motility. *Annual review of microbiology* **56**: 289-314.
McDowell MA, Marcoux J, McVicker G, Johnson S, Fong YH, Stevens R, Bowman LA, Degiacomi MT, Yan J, Wise A et al. 2016. Characterisation of *Shigella* Spa33 and *Thermotoga* FliM/N reveals a new model for C-ring assembly in T3SS. *Molecular microbiology* **99**: 749-766.
McMurry JL, Minamino T, Furukawa Y, Francis JW, Hill SA, Helms KA, Namba K. 2015. Weak Interactions between *Salmonella enterica* FlhB and Other Flagellar Export Apparatus Proteins Govern Type III Secretion Dynamics. *PloS one* **10**: e0134884.
McMurry JL, Van Arnam JS, Kihara M, Macnab RM. 2004. Analysis of the cytoplasmic domains of *Salmonella* FlhA and interactions with components of the flagellar export machinery. *Journal of bacteriology* **186**: 7586-7592.
Meshcheryakov VA, Kitao A, Matsunami H, Samatey FA. 2013. Inhibition of a type III secretion system by the deletion of a short loop in one of its membrane proteins. *Acta crystallographica Section D, Biological crystallography* **69**: 812-820.
Michiels T, Cornelis GR. 1991. Secretion of hybrid proteins by the *Yersinia* Yop export system. *Journal of bacteriology* **173**: 1677-1685.
Michiels T, Wattiau P, Brasseur R, Ruysschaert JM, Cornelis G. 1990. Secretion of Yop proteins by *Yersiniae*. *Infection and immunity* **58**: 2840-2849.

Minamino T. 2018. Hierarchical protein export mechanism of the bacterial flagellar type III protein export apparatus. *Fems Microbiol Lett* **365**.
Minamino T, Gonzalez-Pedrajo B, Kihara M, Namba K, Macnab RM. 2003. The ATPase FliI can interact with the type III flagellar protein export apparatus in the absence of its regulator, FliH. *Journal of bacteriology* **185**: 3983-3988.
Minamino T, Gonzalez-Pedrajo B, Yamaguchi K, Aizawa SI, Macnab RM. 1999. FliK, the protein responsible for flagellar hook length control in *Salmonella*, is exported during hook assembly. *Molecular microbiology* **34**: 295-304.
Minamino T, Macnab RM. 1999. Components of the *Salmonella* flagellar export apparatus and classification of export substrates. *Journal of bacteriology* **181**: 1388-1394.
Minamino T, Macnab RM. 2000. Domain structure of *Salmonella* FlhB, a flagellar export component responsible for substrate specificity switching. *Journal of bacteriology* **182**: 4906-4914.
Minamino T, Morimoto YV, Hara N, Aldridge PD, Namba K. 2016. 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 pathogens* **12**: e1005495.
Minamino T, Namba K. 2004. Self-assembly and type III protein export of the bacterial flagellum. *Journal of molecular microbiology and biotechnology* **7**: 5-17.
Minamino T, Namba K. 2008. Distinct roles of the FliI ATPase and proton motive force in bacterial flagellar protein export. *Nature* **451**: 485-488.
Minamino T, Yamaguchi S, Macnab RM. 2000. Interaction between FliE and FlgB, a proximal rod component of the flagellar basal body of *Salmonella*. *Journal of bacteriology* **182**: 3029-3036.
Monjaras Feria JV, Lefebre MD, Stierhof YD, Galan JE, Wagner S. 2015. Role of autocleavage in the function of a type III secretion specificity switch protein in *Salmonella enterica* serovar Typhimurium. *mBio* **6**: e01459-01415.
Morgan JLW, Acheson JF, Zimmer J. 2017. Structure of a Type-1 Secretion System ABC Transporter. *Structure* **25**: 522-529.
Morimoto YV, Ito M, Hiraoka KD, Che YS, Bai F, Kami-Ike N, Namba K, Minamino T. 2014. Assembly and stoichiometry of FliF and FlhA in *Salmonella* flagellar basal body. *Molecular microbiology* **91**: 1214-1226.
Morimoto YV, Kami-Ike N, Miyata T, Kawamoto A, Kato T, Namba K, Minamino T. 2016. High-Resolution pH Imaging of Living Bacterial Cells To Detect Local pH Differences. *mBio* **7**.
Muhlenkamp M, Oberhettinger P, Leo JC, Linke D, Schutz MS. 2015. *Yersinia* adhesin A (YadA)--beauty & beast. *International journal of medical microbiology : IJMM* **305**: 252-258.
Nagle JF, Tristram-Nagle S. 2000. Structure of lipid bilayers. *Biochimica et biophysica acta* **1469**: 159-195.
Nans A, Kudryashev M, Saibil HR, Hayward RD. 2015. Structure of a bacterial type III secretion system in contact with a host membrane *in situ*. *Nat Commun* **6**: 10114.
Nguyen THD, Tam J, Wu RA, Greber BJ, Toso D, Nogales E, Collins K. 2018. Cryo-EM structure of substrate-bound human telomerase holoenzyme. *Nature* **557**: 190-195.
Nielsen H. 2017. Predicting Secretory Proteins with SignalP. *Methods Mol Biol* **1611**: 59-73.
Noble A, Wei H, Dandey V, Zhang Z, Tan Y, Potter C, Carragher B. 2018. Reducing effects of particle adsorption to the air-water interface in cryoEM. *biorxiv*.
Notti RQ, Bhattacharya S, Lilic M, Stebbins CE. 2015. A common assembly module in injectisome and flagellar type III secretion sorting platforms. *Nat Commun* **6**: 7125.
Nunoura T, Oida H, Miyazaki M, Suzuki Y. 2008. *Thermosulfidibacter takaii* gen. nov., sp. nov., a thermophilic, hydrogen-oxidizing, sulfur-reducing chemolithoautotroph isolated from a deep-sea hydrothermal field in the Southern Okinawa Trough. *International journal of systematic and evolutionary microbiology* **58**: 659-665.
Ochman H, Soncini FC, Solomon F, Groisman EA. 1996. Identification of a pathogenicity island required for *Salmonella* survival in host cells. *Proceedings of the National Academy of Sciences of the United States of America* **93**: 7800-7804.
Ohnishi K, Fan F, Schoenhals GJ, Kihara M, Macnab RM. 1997. The FliO, FliP, FliQ, and FliR proteins of *Salmonella typhimurium*: putative components for flagellar assembly. *Journal of bacteriology* **179**: 6092-6099.
Oomen CJ, van Ulsen P, van Gelder P, Feijen M, Tommassen J, Gros P. 2004. Structure of the translocator domain of a bacterial autotransporter. *The EMBO journal* **23**: 1257-1266.
Orth K. 2002. Function of the *Yersinia* effector YopJ. *Current opinion in microbiology* **5**: 38-43.
Ovchinnikov S, Kamisetty H, Baker D. 2014. Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *eLife* **3**: e02030.

Ovchinnikov S, Kinch L, Park H, Liao Y, Pei J, Kim DE, Kamisetty H, Grishin NV, Baker D. 2015. Large-scale determination of previously unsolved protein structures using evolutionary information. *eLife* **4**: e09248.

Pagel K, Hyung SJ, Ruotolo BT, Robinson CV. 2010. Alternate dissociation pathways identified in charge-reduced protein complex ions. *Analytical chemistry* **82**: 5363-5372.

Pal RR, Baidya AK, Mamou G, Bhattacharya S, Socol Y, Kobi S, Katsowich N, Ben-Yehuda S, Rosenshine I. 2019. Pathogenic *E. coli* Extracts Nutrients from Infected Host Cells Utilizing Injectisome Components. *Cell* **177**: 683-696 e618.

Palmer T, Berks BC. 2012. The twin-arginine translocation (Tat) protein export pathway. *Nature reviews Microbiology* **10**: 483-496.

Papanikou E, Karamanou S, Economou A. 2007. Bacterial protein secretion through the translocase nanomachine. *Nature reviews Microbiology* **5**: 839-851.

Park D, Lara-Tejero M, Waxham MN, Li W, Hu B, Galan JE, Liu J. 2018. Visualization of the type III secretion mediated *Salmonella*-host cell interface using cryo-electron tomography. *eLife* **7**.

Parkhill J, Wren BW, Thomson NR, Titball RW, Holden MT, Prentice MB, Sebaihia M, James KD, Churcher C, Mungall KL et al. 2001. Genome sequence of Yersinia pestis, the causative agent of plague. *Nature* **413**: 523-527.

Persson C, Nordfelth R, Andersson K, Forsberg A, Wolf-Watz H, Fallman M. 1999. Localization of the *Yersinia* PTPase to focal complexes is an important virulence mechanism. *Molecular microbiology* **33**: 828-838.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25**: 1605-1612.

Plano GV, Barve SS, Straley SC. 1991. LcrD, a membrane-bound regulator of the *Yersinia pestis* low-calcium response. *Journal of bacteriology* **173**: 7293-7303.

Portaliou AG, Tsolis KC, Loos MS, Balabanidou V, Rayo J, Tsirigotaki A, Crepin VF, Frankel G, Kalodimos CG, Karamanou S et al. 2017. Hierarchical protein targeting and secretion is controlled by an affinity switch in the type III secretion system of enteropathogenic *Escherichia coli*. *The EMBO journal* **36**: 3517-3531.

Portaliou AG, Tsolis KC, Loos MS, Zorzini V, Economou A. 2016. Type III Secretion: Building and Operating a Remarkable Nanomachine. *Trends in biochemical sciences* **41**: 175-189.

Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. 2018. HMMER web server: 2018 update. *Nucleic acids research* **46**: W200-W204.

Qin Z, Tu J, Lin T, Norris SJ, Li C, Motaleb MA, Liu J. 2018. Cryo-electron tomography of periplasmic flagella in *Borrelia burgdorferi* reveals a distinct cytoplasmic ATPase complex. *PLoS biology* **16**: e3000050.

Radics J, Konigsmaier L, Marlovits TC. 2014. Structure of a pathogenic type 3 secretion system in action. *Nature structural & molecular biology* **21**: 82-87.

Ramakrishnan G, Zhao JL, Newton A. 1991. The cell cycle-regulated flagellar gene flbF of *Caulobacter crescentus* is homologous to a virulence locus (lcrD) of *Yersinia pestis*. *Journal of bacteriology* **173**: 7283-7292.

Reading E, Walton TA, Liko I, Marty MT, Laganowsky A, Rees DC, Robinson CV. 2015. The Effect of Detergent, Temperature, and Lipid on the Oligomeric State of MscL Constructs: Insights from Mass Spectrometry. *Chemistry & biology* **22**: 593-603.

Reboul CF, Eager M, Elmlund D, Elmlund H. 2018. Single-particle cryo-EM-Improved ab initio 3D reconstruction with SIMPLE/PRIME. *Protein science : a publication of the Protein Society* **27**: 51-61.

Renault TT, Abraham AO, Bergmiller T, Paradis G, Rainville S, Charpentier E, Guet CC, Tu Y, Namba K, Keener JP et al. 2017. Bacterial flagella grow through an injection-diffusion mechanism. *eLife* **6**.

Riordan KE, Schneewind O. 2008. YscU cleavage and the assembly of Yersinia type III secretion machine complexes. *Molecular microbiology* **68**: 1485-1501.

Rocha JM, Richardson CJ, Zhang M, Darch CM, Cai E, Diepold A, Gahlmann A. 2018. Single-molecule tracking in live *Yersinia enterocolitica* reveals distinct cytosolic complexes of injectisome subunits. *Integrative biology : quantitative biosciences from nano to macro* **10**: 502-515.

Rollauer SE, Tarry MJ, Graham JE, Jaaskelainen M, Jager F, Johnson S, Krehenbrink M, Liu SM, Lukey MJ, Marcoux J et al. 2012. Structure of the TatC core of the twin-arginine protein transport system. *Nature* **492**: 210-214.

Rosqvist R, Forsberg A, Wolf-Watz H. 1991. Intracellular targeting of the *Yersinia* YopE cytotoxin in mammalian cells induces actin microfilament disruption. *Infection and immunity* **59**: 4562-4569.

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

Sachelaru I, Petriman NA, Kudva R, Koch HG. 2014. Dynamic interaction of the sec translocon with the chaperone PpiD. *The Journal of biological chemistry* **289**: 21706-21715.

Saijo-Hamano Y, Imada K, Minamino T, Kihara M, Shimada M, Kitao A, Namba K. 2010. Structure of the cytoplasmic domain of FlhA and implication for flagellar type III protein export. *Molecular microbiology* **76**: 260-268.

Sakai T, Miyata T, Terahara N, Mori K, Inoue Y, Morimoto YV, Kato T, Namba K, Minamino T. 2019. Novel Insights into Conformational Rearrangements of the Bacterial Flagellar Switch Complex. *mBio* **10**.

Sakazaki R, Iwanami S, Fukumi H. 1963. Studies on the Enteropathogenic, Facultatively Halophilic Bacteria, *Vibrio Parahaemolyticus*. I. Morphological, Cultural and Biochemical Properties and Its Taxonomical Position. *Japanese journal of medical science & biology* **16**: 161-188.

Salmond GP, Reeves PJ. 1993. Membrane traffic wardens and protein secretion in gram-negative bacteria. *Trends in biochemical sciences* **18**: 7-12.

Samatey FA, Matsunami H, Imada K, Nagashima S, Shaikh TR, Thomas DR, Chen JZ, Derosier DJ, Kitao A, Namba K. 2004. Structure of the bacterial flagellar hook and implication for the molecular universal joint mechanism. *Nature* **431**: 1062-1068.

Sargent F, Bogsch EG, Stanley NR, Wexler M, Robinson C, Berks BC, Palmer T. 1998. Overlapping functions of components of a bacterial Sec-independent protein export pathway. *The EMBO journal* **17**: 3640-3650.

Scheres SH. 2012. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of structural biology* **180**: 519-530.

Scheres SH. 2016. Processing of Structurally Heterogeneous Cryo-EM Data in RELION. *Methods in enzymology* **579**: 125-157.

Schmidt TG, Skerra A. 2007. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nature protocols* **2**: 1528-1535.

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

Shea JE, Hensel M, Gleeson C, Holden DW. 1996. Identification of a virulence locus encoding a second type III secretion system in *Salmonella typhimurium*. *Proceedings of the National Academy of Sciences of the United States of America* **93**: 2593-2597.

Shen DK, Blocker AJ. 2016. MxiA, MxiC and IpaD Regulate Substrate Selection and Secretion Mode in the T3SS of *Shigella flexneri*. *PloS one* **11**: e0155141.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* **7**: 539.

Sikora AE. 2013. Proteins secreted via the type II secretion system: smart strategies of *Vibrio cholerae* to maintain fitness in different ecological niches. *PLoS pathogens* **9**: e1003126.

Silverman M, Simon M. 1974. Flagellar rotation and the mechanism of bacterial motility. *Nature* **249**: 73-74.

Singer SJ, Nicolson GL. 1972. The fluid mosaic model of the structure of cell membranes. *Science* **175**: 720-731.

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

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

Spreter T, Yip CK, Sanowar S, Andre I, Kimbrough TG, Vuckovic M, Pfuetzner RA, Deng W, Yu AC, Finlay BB et al. 2009. A conserved structural motif mediates formation of the periplasmic rings in the type III secretion system. *Nature structural & molecular biology* **16**: 468-476.

Stark H. 2010. GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. *Methods in enzymology* **481**: 109-126.

Stock D, Namba K, Lee LK. 2012. Nanorotors and self-assembling macromolecular machines: the torque ring of the bacterial flagellar motor. *Current opinion in biotechnology* **23**: 545-554.

Suzuki H, Yonekura K, Namba K. 2004. Structure of the rotor of the bacterial flagellar motor revealed by electron cryomicroscopy and single-particle image analysis. *Journal of molecular biology* **337**: 105-113.

Swint-Kruse L, Brown CS. 2005. Resmap: automated representation of macromolecular interfaces as two-dimensional networks. *Bioinformatics* **21**: 3327-3328.

Takekawa N, Terahara N, Kato T, Gohara M, Mayanagi K, Hijikata A, Onoue Y, Kojima S, Shirai T, Namba K et al. 2016. The tetrameric MotA complex as the core of the flagellar motor stator from hyperthermophilic bacterium. *Scientific reports* **6**: 31526.

Taylor WR, Matthews-Palmer TR, Beeby M. 2016. Molecular Models for the Core Components of the Flagellar Type-III Secretion Complex. *PloS one* **11**: e0164047.

Terahara N, Inoue Y, Kodera N, Morimoto YV, Uchihashi T, Imada K, Ando T, Namba K, Minamino T. 2018. Insight into structural remodeling of the FlhA ring responsible for bacterial flagellar type III protein export. *Science advances* **4**: eaao7054.

Terashima H, Kawamoto A, Tatsumi C, Namba K, Minamino T, Imada K. 2018. In Vitro Reconstitution of Functional Type III Protein Export and Insights into Flagellar Assembly. *mBio* **9**.

Thomas DR, Francis NR, Xu C, DeRosier DJ. 2006. The three-dimensional structure of the flagellar rotor from a clockwise-locked mutant of *Salmonella enterica* serovar Typhimurium. *Journal of bacteriology* **188**: 7039-7048.

Thomas DR, Morgan DG, DeRosier DJ. 1999. Rotational symmetry of the C ring and a mechanism for the flagellar rotary motor. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 10134-10139.

Thomas S, Holland IB, Schmitt L. 2014. The Type 1 secretion pathway - the hemolysin system and beyond. *Biochimica et biophysica acta* **1843**: 1629-1641.

Torres-Vargas CE, Kronenberger T, Roos N, Dietsche T, Poso A, Wagner S. 2019. The inner rod of virulence-associated type III secretion systems constitutes a needle adapter of one helical turn that is deeply integrated into the system's export apparatus. *Molecular microbiology*.

Tsirigos KD, Peters C, Shu N, Kall L, Elofsson A. 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic acids research* **43**: W401-407.

Ueno T, Oosawa K, Aizawa S. 1992. M ring, S ring and proximal rod of the flagellar basal body of *Salmonella typhimurium* are composed of subunits of a single protein, FliF. *Journal of molecular biology* **227**: 672-677.

Urnavicius L, Zhang K, Diamant AG, Motz C, Schlager MA, Yu M, Patel NA, Robinson CV, Carter AP. 2015. The structure of the dynactin complex and its interaction with dynein. *Science* **347**: 1441-1446.

Van Arnam JS, McMurry JL, Kihara M, Macnab RM. 2004. Analysis of an engineered *Salmonella* flagellar fusion protein, FliR-FlhB. *Journal of bacteriology* **186**: 2495-2498.

Vazquez-Torres A, Xu Y, Jones-Carson J, Holden DW, Lucia SM, Dinauer MC, Mastroeni P, Fang FC. 2000. *Salmonella* pathogenicity island 2-dependent evasion of the phagocyte NADPH oxidase. *Science* **287**: 1655-1658.

Viprey V, Del Greco A, Golinowski W, Broughton WJ, Perret X. 1998. Symbiotic implications of type III protein secretion machinery in *Rhizobium*. *Molecular microbiology* **28**: 1381-1389.

Wagner DM, Klunk J, Harbeck M, Devault A, Waglechner N, Sahl JW, Enk J, Birdsell DN, Kuch M, Lumibao C et al. 2014. *Yersinia pestis* and the plague of Justinian 541-543 AD: a genomic analysis. *The Lancet Infectious diseases* **14**: 319-326.

Wagner S, Grin I, Malmsheimer S, Singh N, Torres-Vargas CE, Westerhausen S. 2018. Bacterial type III secretion systems: a complex device for the delivery of bacterial effector proteins into eukaryotic host cells. *Fems Microbiol Lett* **365**.

Wagner S, Konigsmaier L, Lara-Tejero M, Lefebre M, Marlovits TC, Galan JE. 2010. Organization and coordinated assembly of the type III secretion export apparatus. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 17745-17750.

Waksman G, Hultgren SJ. 2009. Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nature reviews Microbiology* **7**: 765-774.

Wallden K, Rivera-Calzada A, Waksman G. 2010. Type IV secretion systems: versatility and diversity in function. *Cellular microbiology* **12**: 1203-1212.

Wand ME, Sockett RE, Evans KJ, Doherty N, Sharp PM, Hardie KR, Winzer K. 2006. *Helicobacter pylori* FlhB function: the FlhB C-terminal homologue HP1575 acts as a "spare part" to permit flagellar export when the HP0770 FlhBCC domain is deleted. *Journal of bacteriology* **188**: 7531-7541.

Wang S, Li W, Liu S, Xu J. 2016. RaptorX-Property: a web server for protein structure property prediction. *Nucleic acids research* **44**: W430-435.

Wang S, Sun S, Li Z, Zhang R, Xu J. 2017. Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. *PLoS computational biology* **13**: e1005324.

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

Wattiau P, Cornelis GR. 1993. SycE, a chaperone-like protein of *Yersinia enterocolitica* involved in Ohe secretion of YopE. *Molecular microbiology* **8**: 123-131.

Weiner JH, Bilous PT, Shaw GM, Lubitz SP, Frost L, Thomas GH, Cole JA, Turner RJ. 1998. A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell* **93**: 93-101.

Wiles TJ, Mulvey MA. 2013. The RTX pore-forming toxin alpha-hemolysin of uropathogenic *Escherichia coli*: progress and perspectives. *Future microbiology* **8**: 73-84.

Williams AW, Yamaguchi S, Togashi F, Aizawa SI, Kawagishi I, Macnab RM. 1996. Mutations in fliK and flhB affecting flagellar hook and filament assembly in Salmonella typhimurium. *Journal of bacteriology* **178**: 2960-2970.

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A et al. 2011. Overview of the CCP4 suite and current developments. *Acta crystallographica Section D, Biological crystallography* **67**: 235-242.

Wood WB. 1966. Host specificity of DNA produced by *Escherichia coli*: bacterial mutations affecting the restriction and modification of DNA. *Journal of molecular biology* **16**: 118-133.

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

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

Xie L, Altindal T, Chattopadhyay S, Wu XL. 2011. Bacterial flagellum as a propeller and as a rudder for efficient chemotaxis. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 2246-2251.

Xing Q, Shi K, Portaliou A, Rossi P, Economou A, Kalodimos CG. 2018. Structures of chaperone-substrate complexes docked onto the export gate in a type III secretion system. *Nat Commun* **9**.

Xue C, Lam KH, Zhang H, Sun K, Lee SH, Chen X, Au SWN. 2018. Crystal structure of the FliF-FliG complex from *Helicobacter pylori* yields insight into the assembly of the motor MS-C ring in the bacterial flagellum. *The Journal of biological chemistry* **293**: 2066-2078.

Yan Z, Yin M, Xu D, Zhu Y, Li X. 2017. Structural insights into the secretin translocation channel in the type II secretion system. *Nature structural & molecular biology* **24**: 177-183.

Yang Z, Wang C, Zhou Q, An J, Hildebrandt E, Aleksandrov LA, Kappes JC, DeLucas LJ, Riordan JR, Urbatsch IL et al. 2014. Membrane protein stability can be compromised by detergent interactions with the extramembranous soluble domains. *Protein science : a publication of the Protein Society* **23**: 769-789.

Yersin A. 1894. La peste bubonique à Hong Kong. *Ann Inst Pasteur Paris* **8**: 662-667.

Yin M, Yan Z, Li X. 2019. Architecture of type VI secretion system membrane core complex. *Cell research* **29**: 251-253.

Yonekura K, Maki-Yonekura S, Homma M. 2011. Structure of the flagellar motor protein complex PomAB: implications for the torque-generating conformation. *Journal of bacteriology* **193**: 3863-3870.

Yonekura K, Maki-Yonekura S, Namba K. 2003. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* **424**: 643-650.

Young GM, Schmiel DH, Miller VL. 1999. A new pathway for the secretion of virulence factors by bacteria: the flagellar export apparatus functions as a protein-secretion system. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 6456-6461.

Yu XJ, Grabe GJ, Liu M, Mota LJ, Holden DW. 2018. SsaV Interacts with SsaL to Control the Translocon-to-Effector Switch in the *Salmonella* SPI-2 Type Three Secretion System. *mBio* **9**.

Yu XJ, Liu M, Matthews S, Holden DW. 2011. Tandem translation generates a chaperone for the *Salmonella* type III secretion system protein SsaQ. *The Journal of biological chemistry* **286**: 36098-36107.

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

Zarivach R, Vuckovic M, Deng W, Finlay BB, Strynadka NC. 2007. Structural analysis of a prototypical ATPase from the type III secretion system. *Nature structural & molecular biology* **14**: 131-137.

Zhao X, Zhang K, Boquoi T, Hu B, Motaleb MA, Miller KA, James ME, Charon NW, Manson MD, Norris SJ et al. 2013. Cryoelectron tomography reveals the sequential assembly of bacterial flagella in *Borrelia burgdorferi*. *Proceedings of the National Academy of Sciences of the United States of America* **110**: 14390-14395.

Zhou D, Galan J. 2001. *Salmonella* entry into host cells: the work in concert of type III secreted effector proteins. *Microbes and infection* **3**: 1293-1298.

Zhou D, Mooseker MS, Galan JE. 1999. Role of the *S. typhimurium* actin-binding protein SipA in bacterial internalization. *Science* **283**: 2092-2095.

Zilkenat S, Franz-Wachtel M, Stierhof YD, Galan JE, Macek B, Wagner S. 2016. Determination of the Stoichiometry of the Complete Bacterial Type III Secretion Needle Complex Using a Combined Quantitative Proteomic Approach. *Molecular & cellular proteomics : MCP* **15**: 1598-1609.

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

Zivanov J, Nakane T, Scheres SHW. 2019. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* **6**: 5-17.