Understanding Type IV pilus-mediated secretion in *Vibrio cholerae*

by

Minh Nguyen

B.Sc. (Biochemistry), University of Texas at Austin, 2016

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Molecular Biology and Biochemistry
Faculty of Science

SIMON FRASER UNIVERSITY
Fall 2020

Declaration of Committee

Name:	**Minh Nguyen**
Degree:	**Master of Science in Molecular Biology & Biochemistry**
Thesis title:	**Understanding Type IV pilus-mediated secretion in *Vibrio cholerae***
Committee:	**Chair:** **Peter Unrau** Professor, Molecular Biology & Biochemistry
	Lisa Craig Supervisor Professor, Molecular Biology & Biochemistry
	Edgar Young Committee Member Associate Professor, Molecular Biology & Biochemistry
	Jenifer Thewalt Committee Member Professor, Physics, Molecular Biology & Biochemistry
	Julian Guttman Examiner Professor, Biological Sciences

Abstract

Bacterial pathogens depend on the expression of virulence factors that aid host infection. A mechanistic understanding of bacterial virulence can provide insights into novel antimicrobial targets and therapies. One virulence factor is Type IV pili (T4P), long thin filaments found on bacterial surfaces with roles in adhesion, DNA uptake and exoprotein secretion. The T4P system is closely related to the Type II secretion (T2S) system where periplasmic “pseudo-pili” exhibit a piston-like motion for exoprotein export. My research aims to understand T4P-mediated exoprotein secretion in the simple T4P system of *Vibrio cholerae*. I show that the exoprotein’s flexible N-terminal segment is the export signal, which may bind to minor pilin at the pilus tip for delivery across the secretin channel.

Keywords: Type IV pili; Type II secretion; *Vibrio cholerae*; toxin coregulated pilus; minor pilin; exoprotein

Dedication

To my family for their unconditional love. To my mentors, colleagues and friends for their guidance, advice and support.

Acknowledgements

My graduate journey in the Molecular Biology and Biochemistry Department at Simon Fraser University has been a blessing. Above all, I would like to express my deepest gratitude to my senior supervisor Dr. Lisa Craig. The work done here would not have been possible without her mentorship and guidance. I am grateful for the devotion and expertise she has shown for her students.

I would like to thank my committee members Dr. Edgar Young and Dr. Jenifer Thewalt for their time, valuable input and constructive feedback over the last few years. I am also thankful for my internal examiner Dr. Julian Guttman for his expertise and investment in my work.

During my time in the Craig lab, I have had the chance to meet and work with many talented individuals who contributed considerably to this work. To everyone who has made my graduate experience richer and more memorable, Dr. Subramania Kolappan, Bailey Burrell, Navi Garcha, Cole Swanson, Brenden Westman, Tony Harn, Nabeel Khan, John Zhang, Elaine Wang, Katie Danielson and Michele Wu. It has been a pleasure getting to know all of you.

Last but not least, I would like to extend my sincere appreciation to the staff and faculty members in the department for their assistance and support throughout the past few years.

Table of Contents

Declaration of Committee ii
Abstract iii
Dedication iv
Acknowledgements v
Table of Contents vi
List of Tables viii
List of Figures ix
List of Acronyms xi

Chapter 1. Introduction 1
1.1. Vibrio cholerae 1
1.1.1. Pathogenesis 1
1.1.2. Evolution of a human pathogen *V. cholerae* from a harmless aquatic microbe 3
1.2. Overview of bacterial secretion systems 4
1.2.1. Type II secretion system 4
1.2.2. Type IV pili system 5
1.3. Type IV pili 5
1.3.1. The pilus filament 6
1.3.2. Type IV pilus structure 7
1.3.3. Type IV pilus machinery 9
The Type IV pili machinery is structurally related to the Type II secretion system 10
1.3.4. Functions of the type IV pili 11
Similar to the Type II secretion system, the Type IV pili can also secrete substrates 12
1.3.5. Simple and complex T4P systems 15
1.3.6. The minor pilins 17
The complex T4P systems have several minor pilins 17
One single minor pilin drives both pilus assembly and retraction in the simple T4P systems 19
1.3.7. The T4P secreted exoproteins 23
V. cholerae exoprotein TcpF 24
ETEC exoprotein CofJ 24
Comparison between TcpF and CofJ 25
1.3.8. In the ETEC Type IV pili system, the minor pilin CofB binds the exoprotein CofJ 27
1.4. The T4P system is evolutionarily related to the T2S system 30
1.5. Hypotheses and research objectives 32

Chapter 2. Methods 34
2.1. Bacterial strains and growth media 34
2.2. Construction of vectors for various exoprotein clones 34
2.3. Electroporation transformation 35

2.4. Assessment of pili assembly by autoagglutination assay in *V. cholerae* 35
2.5. Exoprotein secretion assay by immunoblot.....35
2.6. Pull-down assay 36

Chapter 3. Results.....37
3.1. T4P-mediated secretion in *V. cholerae* and *E. coli* is specific for autologous exoproteins 38
3.2. The N-terminal segment of TcpF directs export in *V. cholerae*.....40
3.2.1. An exoprotein construct lacking the export sequence accumulates in the bacterial periplasm.....41
3.2.2. The TcpF N-term and C-term domains are secreted when possessing the export sequence 42
3.2.3. The export signal of TcpF and CofJ mediate heterologous secretion in *V. cholerae* and ETEC 43
3.3. Investigating the relationship between the exoprotein export signal and the minor pilin trimer 47
3.4. Investigating the relationship between exoprotein expression and pilus assembly... 50

Chapter 4. Discussion and Future Studies.....53

References.....56

Appendix A. List of bacterial strains, plasmids, primers and antibodies.....60

Appendix B. Sequence analysis of the major peaks obtained from Liquid Chromatography Mass Spectrometry for two TcpF samples.....62

List of Tables

Table 3-1 Summary of secretion phenotypes for all tested protein constructs 47

List of Figures

Figure 1-1 Human intestine infection by *Vibrio cholerae* 3
Figure 1-2 Electron micrograph of a *Vibrio cholerae* cell expressing its Toxin-Coregulated pili and a single flagellum 6
Figure 1-3 Structure of the major pilin subunit and the pilus 8
Figure 1-4 The Type IV pilus structure 9
Figure 1-5 Comparison between Type II secretion and Type IV pilus systems 10
Figure 1-6 Twitching motility mediated by the Type IV pilus 12
Figure 1-7 Cytoplasmic membrane translocation by the Sec machinery 14
Figure 1-8 Secretion by the Type II secretion system in *Vibrio cholerae* 15
Figure 1-9 Schematic of the operons encoding the simple Type IV pili systems in *Vibrio cholerae* and ETEC 16
Figure 1-10 Comparison between the simple and complex Type IV pili systems 17
Figure 1-11 GspI-GspJ-GspK complex in the ETEC Type II secretion system 18
Figure 1-12 Comparison between the *Vibrio cholerae minor pilin TcpB* and the ETEC minor pilin CofB 20
Figure 1-13 Crystal structures of the TcpB-C and ΔN-CofB trimers 21
Figure 1-14 Immunogold labeling of *V. cholerae* minor pilin TcpB at the pilus tip 22
Figure 1-15 Proposed model of Type IV pili-mediated CTXΦ uptake in *Vibrio cholerae* 23
Figure 1-16 Comparison between the *Vibrio cholerae* and ETEC exoproteins TcpF and CofJ 26
Figure 1-17 Comparison between the signal peptide and the unresolved segment of TcpF and CofJ 27
Figure 1-18 Crystal structure of the CofJ (1-24)-CofB complex 29
Figure 1-19 Proposed model of the CofJ-CFA/III pilus complex in ETEC 30
Figure 1-20 *Vibrio cholerae* Type II secretion and Type IV pili systems are both required for its pathogenesis 32
Figure 1-21 Proposed Type IV pili-mediated TcpF secretion in *Vibrio cholerae* 33
Figure 3-1 Schematic of TcpF and CofJ constructs used for secretion assays 38
Figure 3-2 *Vibrio cholerae* cannot secrete CofJ and *Escherichia coli* cannot secrete TcpF 40
Figure 3-3 A TcpF construct lacking its export sequence accumulates in the *Vibrio cholerae* periplasm 42
Figure 3-4 In *Vibrio cholerae*, secretion is determined by the export sequence of its wild-type exoprotein TcpF 44
Figure 3-5 In ETEC, secretion is determined by the export sequence of its wild-type exoprotein CofJ 45
Figure 3-6 Quantifying whole cell and supernatant fractions in our studied *Vibrio cholerae* and *Escherichia coli* mutant strains 46

Figure 3-7 TcpF appears as a doublet on polyacrylamide gel.....48
Figure 3-8 Liquid Chromatography Mass Spectrometry of two TcpF samples 49
Figure 3-9 Pull-down assay between His TcpB-C and two non-His-tagged TcpF proteins, one full-length and the other with degraded export signal 50
Figure 3-10 Pilus assembly in *Vibrio cholerae* as reflected by immunoblotting against the major pilin subunit TcpA in the culture supernatant 52

List of Acronyms

ADP	Adenosine diphosphate
Ap	Ampicillin
ATP	Adenosine triphosphate
CFA/III	Colonization factor antigen III
Cm	Chloramphenicol
CTD	C-terminal domain
CTXΦ	Cholera toxin phage
ES	Export sequence
ETEC	Enterotoxigenic *Escherichia coli*
HRP	Horseradish peroxidase
IM	Inner membrane
LB	Lysogeny broth
Ni-NTA	Nickel-nitrilotriacetic acid
NTD	N-terminal domain
NTS	N-terminal segment
OD_{600}	Optical density at 600 nm
OM	Outer membrane
PBS	Phosphate-buffered saline
rpm	Revolutions per minute
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Sec	General secretion pathway
SP	Signal peptide(s)
T2S	Type II secretion
T4P	Type IV pilus/pili
TAT	Twin arginine translocation
TCP	Toxin-coregulated pilus/pili
TEM	Transmission electron microscopy

Chapter 1.

Introduction

1.1. Vibrio cholerae

Vibrio cholerae is a Gram-negative, rod-shaped bacterium that is responsible for the diarrheal disease cholera. *V. cholerae* lives in aquatic environments and thrives on organic debris such as shells of crabs, shrimp and shellfish. Infection of *V. cholerae* typically stems from consuming raw or undercooked seafood or drinking contaminated water. Cholera infection is a leading cause of child morbidity in developing countries and regions with poor access to water, sanitation and hygiene infrastructure (Nelson, Harris, Morris, Calderwood, & Camilli, 2009). The world is currently experiencing the seventh cholera pandemic, which began in 1961. The World Health Organization (WHO, 2020) estimates a total 2.8 million cholera cases and over 90,000 deaths annually. Today, amidst the COVID-19 global pandemic, cholera continues to be a prominent threat in vulnerable communities, particularly those affected by natural disasters and humanitarian crises.

V. cholerae serogroups are distinguished based on their O antigen, a branched and variable carbohydrate present in the outer-membrane lipopolysaccharide. Among more than 200 serogroups, only two, O1 and O139, are responsible for human deaths. The O1 serogroup is further categorized into 2 distinct biotypes, classical and El Tor, based on their glucose-metabolization pathways. While the former is responsible for pandemics 1-6, the latter has become more prominent in the last two decades (Nelson et al., 2009). Most non-O1 and non-O139 serogroups are harmless aquatic microbes but have the potential to become pathogens by means of horizontal gene transfer.

1.1.1. Pathogenesis

Cholera infection typically begins with the ingestion of water or food contaminated with *V. cholerae*. Symptoms may take up to 72 hours to become apparent, ranging from stomach cramps to vomiting, diarrhea but in severe cases can rapidly progress to death if untreated. Pathogenesis of *V. cholerae* requires the expression of

two virulence factors: the toxin-coregulated pilus (TCP) and cholera toxin. Once inside the human host, *V. cholerae* adheres to the epithelial cells in the small intestine using their GlcNac-binding protein, GbpA (Kirn, Jude, & Taylor, 2005). *V. cholerae* further aggregates into microcolonies using their TCP (Taylor, Miller, Furlong, & Mekalanos, 1987). Colonization is followed by secretion of cholera toxin by the Type II secretion system. CT is a polymeric protein consisting of five B subunits and one enzymatically active A subunit. The B subunit pentamer recognizes and binds to GM_1 ganglioside, a common glycolipid found in eukaryotic cell membranes, while the A subunit is responsible for enzymatic activity by ribosylating ADP, causing fluid loss in the form of diarrhea characteristic of cholera infection (Finkelstein, 1996). At the molecular level, ADP ribosylation caused by the active A subunit of CT perpetually activates adenylate cyclase, leading to excess production of cyclic 5'-adenosine monophosphate (cAMP) (Cassel & Pfeuffer, 1978). The elevated cAMP levels hyper-activate proteinase A, which causes electrolyte imbalance and large efflux of water, resulting in massive diarrhea. In addition to the cholera toxin secreted by the Type II secretion system, *V. cholerae* also secretes another protein, TcpF, by its Type IV pili. TcpF was proposed to act as a colonization factor, aiding *V. cholerae* microcolony formation in human epithelial cells by an unknown mechanism (Kirn, Bose, & Taylor, 2003). Patients experiencing severe symptoms of cholera can lose up to 20 L of body fluid per day (Finkelstein, 1996). The most effective treatment for *V. cholerae* infection is rehydration with electrolytes. Quick diagnosis and proper treatment can reduce cholera mortality to under 1% (Dick, Guillerm, Moussy, & Chaignat, 2012).

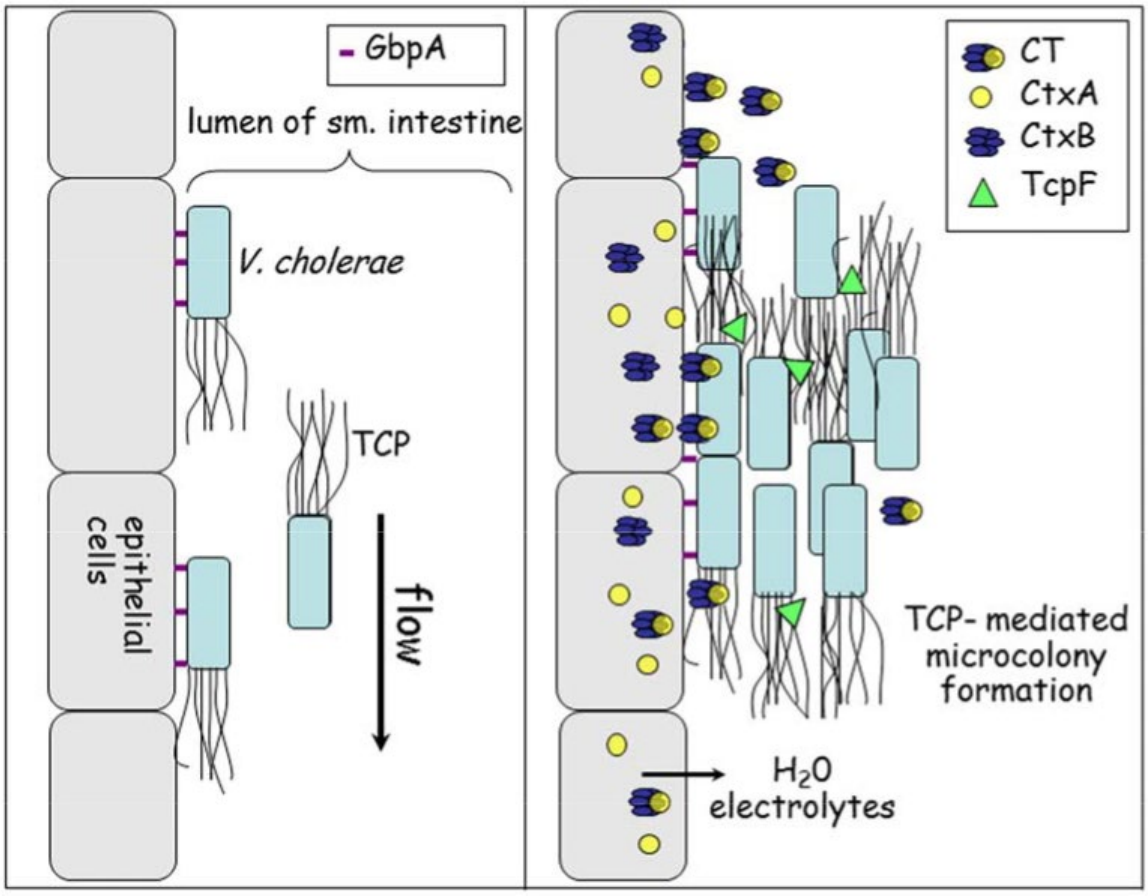

Figure 1-1 Human intestine infection by *Vibrio cholerae*
Pathogenic *V. cholerae* infection starts with cell adherence to the lumen of the small intestine. Once ingested into the stomach, *V. cholerae* attach to epithelial cell surface via their GbpA protein and form aggregates via their TCP. The colonization factor TcpF secreted by the TCP system system further promotes microcolony formation through an unknown mechanism. *V. cholerae* also possess a Type II secretion system responsible for the secretion of CT. Its B subunits binds GM_1 ganglioside to facilitate endocytosis of the active A subunit, causing elevated cAMP levels through a cascading response and massive water and electrolytes efflux characteristic of diarrhea. Image courtesy of Lisa Craig.

1.1.2. Evolution of a human pathogen *V. cholerae* from a harmless aquatic microbe

Only 2 out of over 200 known *V. cholerae* serogroups cause cholera, suggesting they acquired the virulence factors TCP and cholera toxin through means of horizontal gene transfer. Cholera toxin is encoded on the CTX element derived from the filamentous bacteriophage CTXΦ. CTXΦ selectively infects environmental *V. cholerae* strains displaying TCP (DiRita, Parsot, Jander, & Mekalanos, 1991). The CTX element is subsequently taken up into the bacterial cell for homologous recombination, ultimately turning *V. cholerae* into a pathogen responsible for global pandemics. The *tcp* operon is

encoded on a large gene cluster known as the *Vibrio* pathogenicity island (VPI-1) of unknown origin. VPI-1 also encodes a transcriptional factor ToxT, which activates expression of both virulence factors TCP and cholera toxin (Taylor, Miller, Furlong, & Mekalanos, 1986).

1.2. Overview of bacterial secretion systems

To enhance their environmental fitness in the competition for survival, microorganisms secrete substrates or exoproteins across their membranes into the environment. Secretion across the cytoplasmic membrane is facilitated by either the Sec machinery, which trasnlocates protein in an unfolded form, or by the twin arginine translocation (TAT), which translocates folded proteins. In Gram-negative bacteria which have a second, outer membrane, in addition to their inner membrane, secretion occurs either in a single step across both membrane barriers, or in two sequential steps: periplasmic translocation by the Sec or the TAT system followed by export across the outer membrane. Bacteria have several one-step secretion systems: Type I, Type III, Type IV and Type VI secretion (Green & Mecsas, 2016). The two-step secretion includes the Type II and Type V secretion systems. The Type III, Type IV and Type VI secretion systems have implications as bacterial injectisomes, acting as a needle and syringe to inject substrates from the bacterial cytoplasm into the host plasma membrane (Backert & Meyer, 2006; Buttner, 2012; Mougous et al., 2006). In certain Gram-positive species such as *Mycobacteria* and *Corynebacteria*, the presence of a mycomembrane, a heavily lipidated cell wall layer, requires specialized mycomembrane secretion in the form of the Type VII secretion system (Freudl, 2013).

1.2.1. Type II secretion system

The Type II secretion (T2S) system is present across the bacterial kingdom in pathogenic and non-pathogenic species alike. Generally referred to as the General Secretory Pathway (Gsp), the T2S system is capable of accommodating folded protein substrates with widely variable shapes, sizes and structures, from proteases, hydrolases, lipases to enzyme capable of breaking down complex carbohydrates (Korotkov, Sandkvist, & Hol, 2012). The bacterium *Shewanella oneidensis* uses its T2S system substrate to reduce insoluble metal oxides and perform electron transport (Shi et

al., 2008). Many bacterial pathogens use their T2S systems to export virulence factors into the extracellular space: *V. cholerae* secretes the folded/assembled cholera toxin hexamer, which causes a massive efflux of water and electrolytes from epithelial cells, and the plant pathogen *Erwinia carotovora* secretes a pectolytic enzyme responsible for soft rot disease in potatoes (Toth & Birch, 2005).

1.2.2. Type IV pili system

The T2S system is closely related to an important class of bacterial appendages, the Type IV pilus. The Type IV pili are present on the cell surface in both Gram-negative and Gram-positive bacteria alike, though they are more well-established in Gram-negative human pathogens such as *Vibrio cholerae* and enterotoxigenic *Escherichia coli*. The Type IV pili are adhesive and retractile, allowing them to facilitate many diverse functions (Craig, Forest, & Maier, 2019). Though the Type IV pili system can also secrete substrates, it is often overlooked in this capacity when compared to the well-established T2S system. Recent developments on the molecular architecture of the Type IV pilus machinery provide valuable insight into its role as an essential element in promoting bacterial pathogenesis.

1.3. Type IV pili

Type IV pili (T4P) are long thin filaments that are expressed on the cell surface of many bacteria (Fig 1-2). T4P are highly dynamic structures that extend from the inner membrane through the periplasm and across the outer membrane via a gated secretin channel (Craig et al., 2019). T4P are important appendages that mediate a variety of functions including but not limited to cellular adhesion, natural transformation, twitching motility and microcolony formation (Giltner, Nguyen, & Burrows, 2012). In pathogenic *V. cholerae* strains, the T4P are essential to virulence (DiRita et al., 1991), highlighting their potential as a desirable drug target for treatment development.

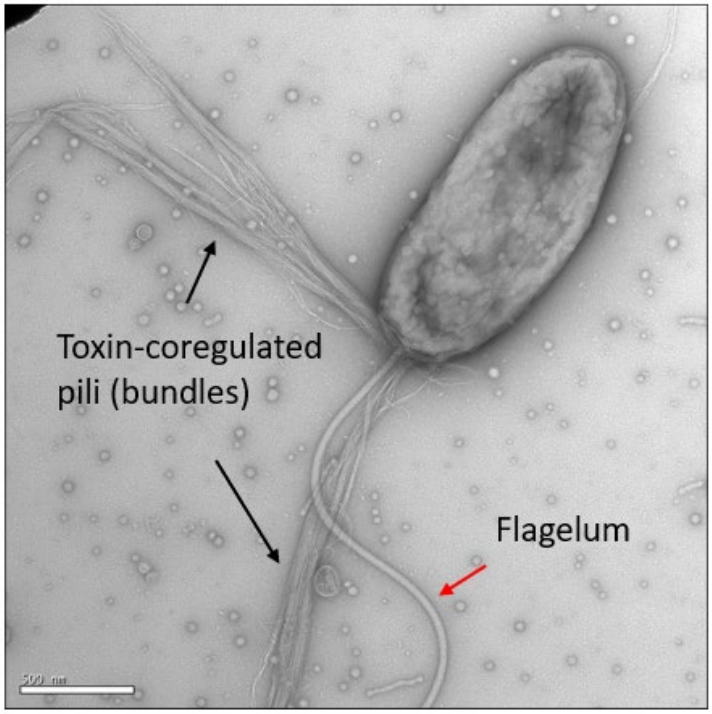

Figure 1-2 Electron micrograph of a *Vibrio cholerae* cell expressing its Toxin-Coregulated pili and a single flagellum

The Gram-negative bacterium, *V. cholerae*, displays a single polar flagellum and bundles of its Toxin-Coregulated Pili on the cell surface. Stained with 3% uranyl acetate and imaged on a Hitachi 8100 Scanning TEM by Lisa Craig.

1.3.1. The pilus filament

T4P are comprised of thousands of copies of the major pilin subunits, which form the body of these hair-like filaments, and a small number of minor pilins. Pili are flexible and elastic, with diameters of 6-8 nm. While typical pilus length is usually no more than a few microns, some can grow to over 20 µm, as seen for Longus, another T4P produced by ETEC (Giron, Levine, & Kaper, 1994). T4P are remarkably robust. *N. gonorrhoeae* GC pili are resistant in up to 8 M urea and temperatures of over 80^{o} C (Li, Egelman, & Craig, 2012). The T4P of extremophiles such as the hyper-thermophilic archaeon *Sulfolobus islandicus* show remarkable stability under intense physio-chemical conditions, resisting both protease digestion and heating in SDS (Wang et al., 2020). The resilient nature of the pilus filaments enables them to perform diverse functions in the extracellular space over considerable distances.

1.3.2. Type IV pilus structure

The building block for the T4P, the major pilin, has a canonical structure. Type IV pilins have a conserved hydrophobic N-terminal α-helical spine (α1) and a globular C-terminal domain (Fig 1-3A). The N-terminal protruding half of α1, α1N, anchors the globular domain in the inner membrane prior to assembly and forms a helical array in the core of the intact pilus. The C-terminal globular domain is exposed to the periplasm and forms the exterior of the pilus filament (Craig et al., 2019). Pilin subunits are arranged in the filament along a right-handed helix with each subunit related to the next by an axial rise of ~10 Å and a rotation of ~100^{o} (Fig 1-3B) (Kolappan et al., 2016; Li et al., 2012; Wang et al., 2017). Pilus assembly occurs in the inner-membrane, where pilin subunits are recruited and docked at the base of the growing pilus, which is extruded outwards incrementally through the periplasm and across the secreting channel. Pilus assembly is driven in part by charge complementarity between the conserved Glu5 of the incoming major pilin and the positively-charged N-terminal amino nitrogen of the terminal subunit of the pilus (Fig 1-4A) (Craig & Li, 2008). This interaction neutralizes these opposite charges in the lipid bilayer and in the hydrophobic core of the assembled pilus. In contrast, the pilin globular domains make only minimal contacts with each other in the assembled pilus.

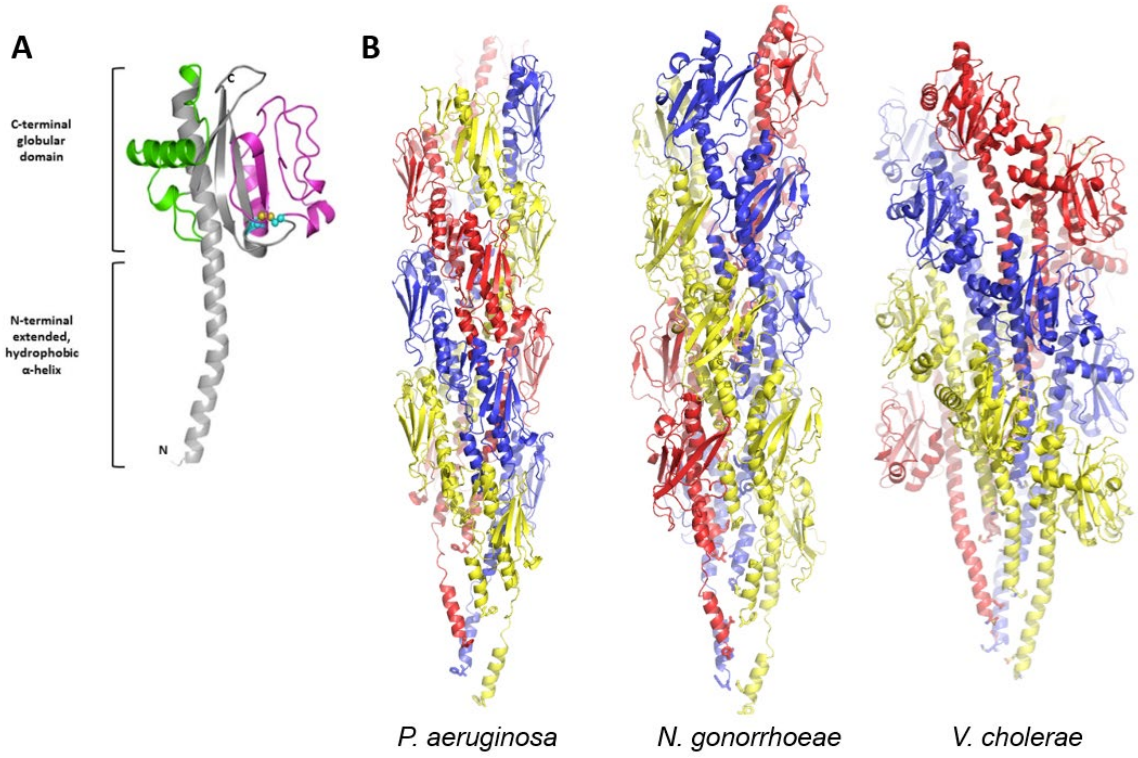

Figure 1-3 Structure of the major pilin subunit and the pilus
A. The major pilin subunit composes of the hydrophobic N-terminal α-helix and the C-terminal globular domain. **B**. Cryo-EM based pseudoatomic resolution structures of the Type IV pilus in *Pseudomonas aeruginosa*, *Neisseria gonorrhoeae* and *Vibrio cholerae*. Image courtesy of Lisa Craig.

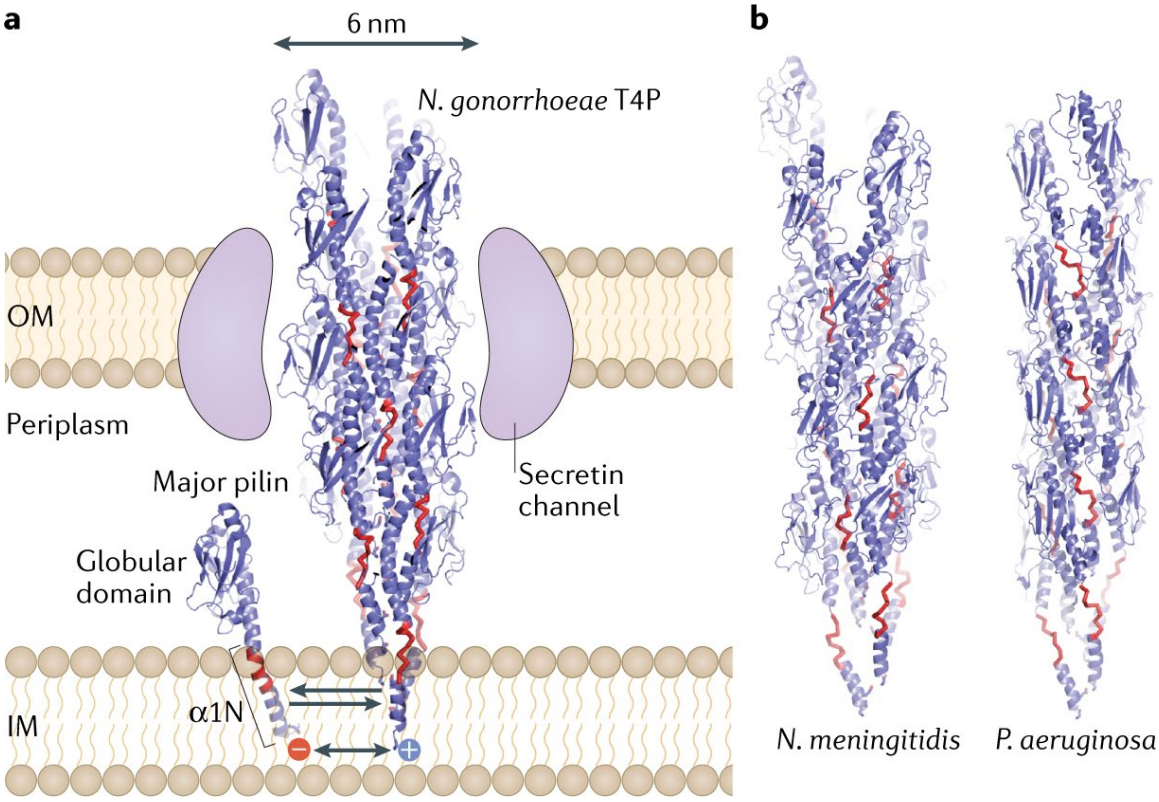

Figure 1-4 The Type IV pilus structure

a) Pilus assembly is driven by charge complementarity. The N-terminal protruding half of α1, α1N, anchors each pilin subunit in the inner membrane prior to pilus assembly. Charge complementarity between N1+ and Glu5- forms a salt bridge which mediate polymerization and neutralizes the charges in the hydrophobic pilus core. Each subunit recruited to the base of the growing pilus extends the filament outward incrementally. Through cycles of polymerization, the pilus extrudes across the periplasm to the extracellular space. **b)** Structure of the T4P from *Neisseria meningitidis* and *Pseudomonas aeruginosa*. Figure adapted from Craig et al., 2019 and used with permission from the publisher.

1.3.3. Type IV pilus machinery

In Gram-negative bacteria, the core T4P machinery comprises of a pilus, a cytoplasmic assembly ATPase, an inner-membrane platform and an outer-membrane gated secretin channel (Fig 1-5). The cytoplasmic assembly ATPase motor provides energy for pilus assembly in the form of ATP hydrolysis to extrude the pilus incrementally upon each pilin subunit addition (Craig et al., 2019; Mancl, Black, Robinson, Yang, & Schubot, 2016). The secretin channel is gated and multimeric, resembling a 12-fold symmetric cage-like structure that opens and closes to assist passage of pilus and exoprotein (Berry et al., 2012; Reichow, Korotkov, Hol, & Gonen, 2010). Some T4P systems possess a second retraction ATPase motor that drives pilus

depolymerization whereby pilin subunits diffuse from the base of the pilus back into the inner membrane, one subunit at a time (McCallum, Tammam, Khan, Burrows, & Howell, 2017). T4P systems lacking a retraction ATPase are still capable of retracting its pili, albeit at a lower force (Ellison et al., 2017; Ng et al., 2016).

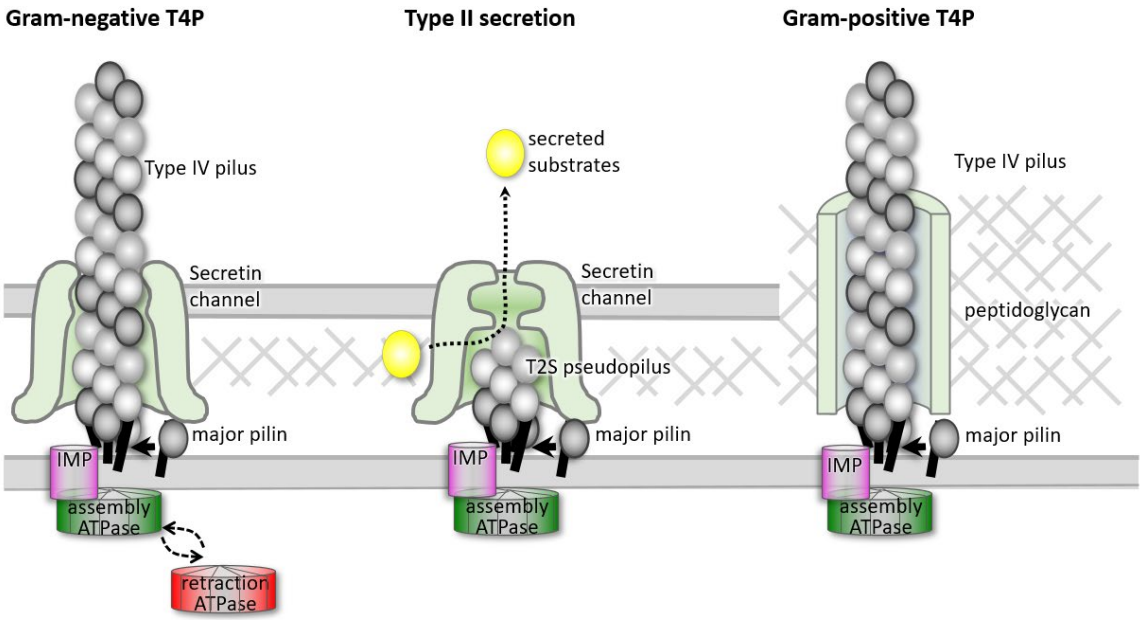

Figure 1-5 Comparison between Type II secretion and Type IV pilus systems
Type II secretion system and Type IV pilus in Gram-negative and Gram-positive bacteria. Several components are homologous across these systems, including the pilus, the inner-membrane platform (IMP), the cytoplasmic assembly ATPase and the secretin channel. Some T4P systems possess an additional retraction ATPase. Image courtesy of Lisa Craig.

The Type IV pili machinery is structurally related to the Type II secretion system

The T4P system is structurally similar to the T2S system. A complex multiprotein machinery consisting of 12 to 15 single-operon-encoded components, the T2S apparatus can be further categorized into 4 main compartments: the inner-membrane platform, the cytoplasmic ATPase, the pseudo-pilus and the outer-membrane secretin channel (Fig 1-5). Even though structurally related to a Type IV pilus, the Type II pseudo-pilus is limited to the bacterial periplasm. The pseudo-pilus is polymerized in the inner-membrane and extended into the periplasm where it is thought to act like a piston to “push” substrates across the secretin channel. The cytoplasmic ATPase provides energy for polymerization using ATP hydrolysis. The T2S pseudo-pilus is made up of the major and minor pilin subunits (Korotkov & Sandkvist, 2019). In theory less than a dozen pilins would be required to span the periplasmic space, which is only ~10 nm thickness

(Silhavy, Kahne, & Walker, 2010). It is thought that the pseudo-pilus is capped by a cluster of minor pilins, with a bulky minor pilin at the tip that blocks passage of the pseudo-pilus across the secretin channel (Korotkov & Hol, 2008; Sandkvist, 2001). Blockage of the pseudo-pilus stalls assembly and causes its spontaneous collapse. Repeated cycles of the pseudo-pilus quickly assembling and collapsing result in piston-like motion to secrete substrates (Korotkov et al., 2012).

1.3.4. Functions of the type IV pili

T4P are ubiquitous in bacteria and critical for pathogenesis. T4P were first described in *Pseudomonas aeruginosa* in the 1970s by David E. Bradley. Visualization by an electron microscope showed a reduction in the average pili length after RNA phage adsorption, suggesting pili are dynamic appendages (Bradley, 1972). Since then, the role of T4P has become a major focus for bacterial studies. T4P were identified as a key component for natural transformation in both Gram-negative and Gram-positive bacterial species. Apart from their functions in basic cellular needs such as motility, surface attachment and nutrition uptake, T4P have been implicated in promoting bacterial virulence. In 1973, Punsalang and Sawyer demonstrated in *Neisseria gonorrhoeae* the role of T4P in colonizing human epithelial cells (Punsalang & Sawyer, 1973). *N. gonorrhoeae* cells lacking pili were defective in epithelial cell adherence and more vulnerable to phagocytosis. The ability of *V. cholerae* to infect and form microcolonies in the human gut is dependent on the expression of its toxin-coregulated pili (TCP), a Type IV pilus (Taylor et al., 1987). Similar strategies have been used by *Neisseria meningitidis* where host cell attachment via its pili triggers bacterial aggregation (Chen & Seifert, 2011).

Many T4P are capable of twitching motility, a form of bacterial movement similar to the action of a grappling hook (Fig 1-6). Once attached to the surface, as the pili retract and exert a pulling motion, the bacteria appear to twitch. Bradley earlier observation about the decrease in average pili length of *Pseudomonas aeruginosa* after phage adsorption is in fact retraction. In this regard, the T4P machinery can be considered a molecular motor. A single *N. gonorrhoeae* T4P can produce a mechanical force of over 100 pN, marking it one of the strongest molecular motors known to date (Maier et al., 2002).

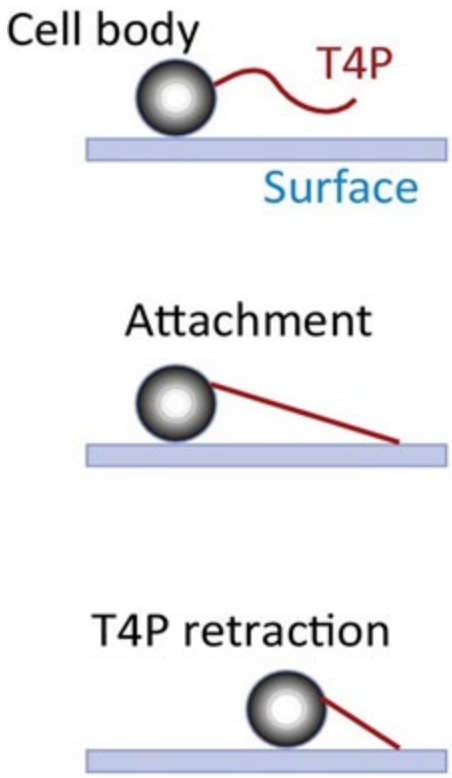

Figure 1-6 Twitching motility mediated by the Type IV pilus
By attaching to a surface and retract, the T4P act as a molecular motor to exert a pulling motion, causing the cell to twitch. Twitching motility is caused by cycles of T4P extension and retraction. Figure adapted from Maier et al., 2015 and used with permission from the publisher.

Similar to the Type II secretion system, the Type IV pili can also secrete substrates

T4P can also function as a secretion system. *Francisella tularensis*, the causative agent of the zoonotic disease tularaemia, and *Dichelobacter nodosus*, commonly known as the ovine foot rot pathogen, both secrete proteases via their T4P machinery (Hager et al., 2006; Han, Kennan, Parker, Davies, & Rood, 2007). The human pathogen *V. cholerae* secretes an exoprotein, TcpF, using its TCP (Kirn et al., 2003). *V. cholerae* mutants defective in TcpF secretion failed to colonize the infant mouse intestine. Enterotoxigenic *Escherichia coli* (ETEC), which causes traveler's diarrhea, secretes CofJ using its T4P known as the colonization factor antigen (CFA/III) (Yuen, Kolappan, Ng, & Craig, 2013). Despite much progress in understanding the structure and function of the T4P and the T2S systems, their underlying secretion mechanisms are poorly understood. Due to their evolutionary relatedness, elucidating this process in one system can provide valuable insight into the other.

In the Type II secretion system, secretion occurs in a two-step process

The T2S process can be separated into 2 steps, across 2 barriers, the inner membrane and the outer membrane sequentially. Together these are referred to as the General Secretion Pathway (Gsp). A common translocation mechanism across the inner-membrane utilizes the Sec system which is present ubiquitously across the bacterial kingdom, including both Gram-positive and Gram-negative species alike. By working sequentially with a second dedicated secretion system such as T2S, the Sec-dependent pathway indirectly contributes to bacterial virulence in many bacterial pathogens.

The Sec machinery consists of 3 components: a protein targeting guide SecB, a motor protein SecA, a membrane integrated conducting channel SecYEG and a signal peptidase (Fig 1-7). Proteins recognized by the Sec-dependent machinery are synthesized as precursors or “pre-proteins” having a short N-terminal signal peptide approximately 20 residues long. Evolutionarily conversed across many bacterial species, the signal peptides of different proteins share many common features, including a positive amino end, a central hydrophobic region and small aliphatic residues at the -3 and/or -1 region relative to the cleavage site (Paetzel, 2019). The hydrophobic stretch of the signal peptide is important for its insertion into the phospholipid bilayer and aliphatic residues confer recognition by the signal peptidase. A component of the Sec system, SecB, identifies the signal peptide and forms a complex with the precursor in the cytoplasm (Natale, Bruser, & Driessen, 2008). Serving as a chaperone, SecB keeps the pre-protein in its unfolded state and guides it to the membrane-bound conducting channel, the SecYEG translocon. A motor protein, the SecA ATPase, is associated with the translocon on the cytoplasmic side and provides energy for the translocation process by ATP hydrolysis (Green & Mecsas, 2016). An inner membrane-embedded protease, the signal peptidase, has its active site located on the periplasmic side. This protease cleaves off the membrane-anchored signal peptide, releasing the mature protein in the periplasm where it folds. In the enteric pathogens *V. cholerae* and ETEC, sequence analysis reveals the substrates secreted by their respective T4P systems also possess a short N-terminal segment about 20 residues long characteristic of the Sec-dependent signal peptide.

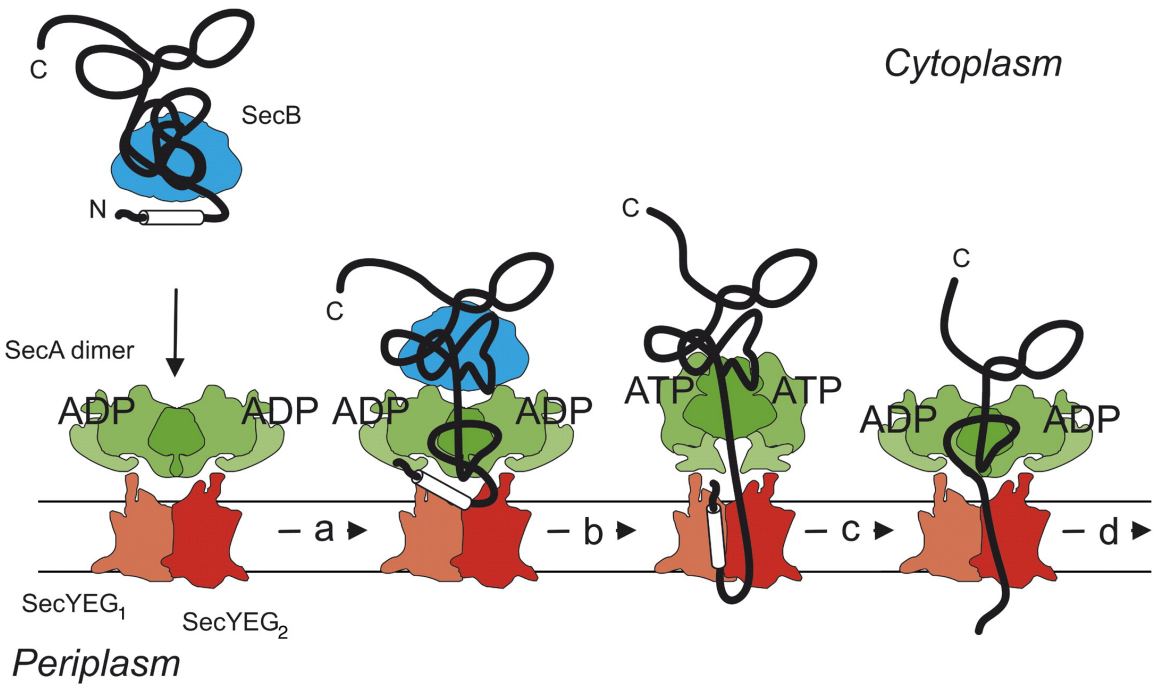

Figure 1-7 Cytoplasmic membrane translocation by the Sec machinery
Proteins destined for the periplasm or extracellular release are synthesized in the cytoplasm as precursors containing a cleavable signal peptide recognized by the SecB protein. SecB serves as a chaperone, keeping the precursor in its unfolded state, and delivers it to SecA. SecA is an ATPase motor that associates with the membrane-spanning SecYEG translocon. SecA provides energy for periplasmic translocation by ATP hydrolysis. Once in the periplasm, the protein adopts its native folded state. Figure adapted from Natale et al., 2008 and used with permission from the publisher.

Once processed and translocated into the periplasm by the Sec machinery, the substrate passage through the outer-membrane secretin channel is regulated by the T2S pseudo-pilus (Fig 1-8). Despite much progress in understanding the structures and functions of the T2S system over past decades, little is known about how substrate is selected for secretion across the outer membrane, what I refer to as “export”. Most likely folding of the substrate in the periplasm exposes yet a second signal that is recognized by the T2S system. Thus, substrates exported by the T2S system, referred to here as “exoproteins”, are synthesized in the cytoplasm with two discrete signals, the signal peptide (SP), recognized by the Sec machinery, and the export signal (ES) recognized by the T2S system. It is also not known which component of the T2S machinery is responsible for recognizing the exoprotein but two potential candidates are the pseudo-pilus and the secretin channel. Three components of the *Pseudomonas aeruginosa* T2S system Xcp, the secretin and two minor pilins of the pseudo-pilus tip, have been described to specifically interact with the secreted substrates (Douzi, Ball, Cambillau,

Tegoni, & Voulhoux, 2011). The cholera toxin secreted by *V. cholerae* was shown to bind to the periplasmic vestibule of its T2S secretin channel GspD (Reichow et al., 2011). Both observations are consistent with the proposed model where the pseudo-pilus acts as a piston to push substrates across the outer-membrane.

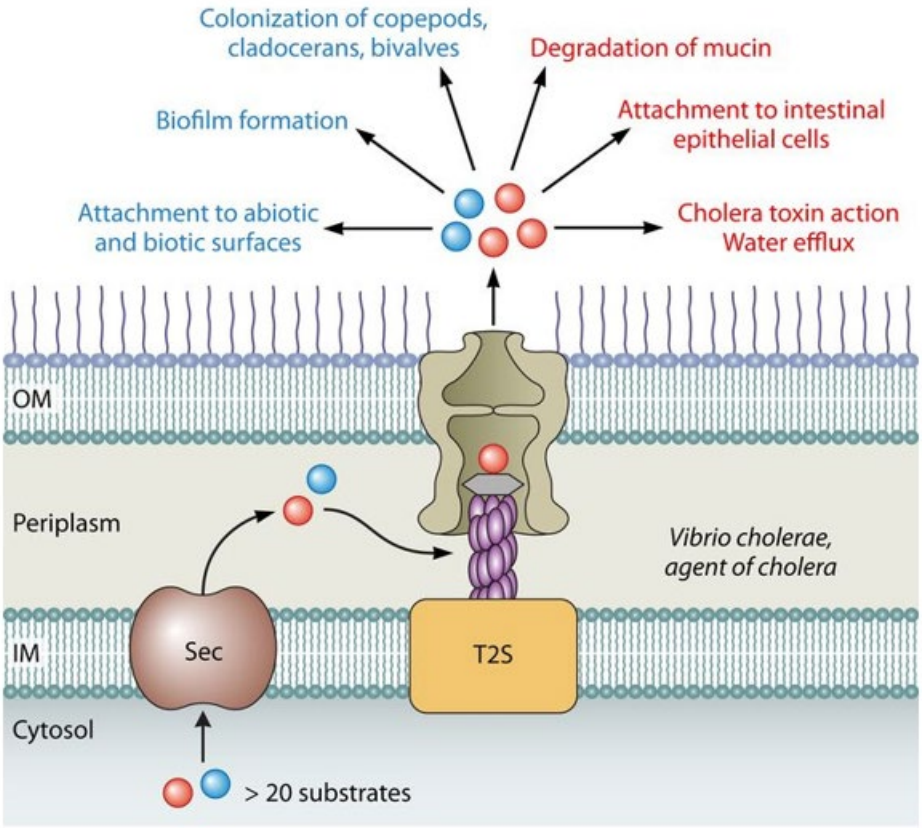

Figure 1-8 Secretion by the Type II secretion system in *Vibrio cholerae*
Substrate secretion by the T2S system in *V. cholerae*, the causative agent of cholera, occurs in a two-step process. Initial cytoplasmic translocation across the inner-membrane (IM) is dependent on the Sec system. Once in the periplasm, the substrate can be selected for export across the outer-membrane (OM) barrier by the T2S system. More than 20 substrates, including the cholera toxin, are exported via the T2S system in *V. cholerae*. Substrates can aid microcolony formation (blue) or confer bacterial pathogenesis (red). Figure adapted from Cianciotto et al., 2017 and used with permission from the publisher.

1.3.5. Simple and complex T4P systems

The T4P machinery spanning the bacterial envelope is further categorized into simple and complex systems. Though there are many features that vary from one T4P system to the next, and in fact these can be seen as a continuum, there are a few

features that allow a general classification of many T4P as simple or complex T4P (Craig et al., 2019). Simple systems have as few as 9 components which are all typically encoded on a single operon (Fig 1-9). Complex T4P machinery can have up to 40 components scattered throughout the entire bacterial chromosome. The type IV pilus filament is predominantly made up of the major pilin and a smaller number of the minor pilin subunits. The complex systems have several distinct minor pilins, whereas the simple T4P systems consist of one single minor pilin subunit which was shown to be essential for pilus assembly (Ng et al., 2016) (Fig 1-10). Although both can produce retractile pili, their mechanisms differ: while retraction is powered by a dedicated retraction ATPase motor in the complex system, it is driven by the minor pilin in the simple system (Ng et al., 2016). Simple T4P are best studied in enteric pathogens such as *V. cholerae* and ETEC. Complex T4P system encompasses a wide range of human and non-human pathogens such as *Pseudomonas aeruginosa*, *Dichelobacter nodosus*, *Myxococcus xanthus* and *Neisseria gonorrhoeae*. Both *V. cholerae* TCP and the ETEC CFA/III are simple T4P systems capable of retraction despite lacking a retraction ATPase (Craig et al., 2019). This is thought to be facilitated by incorporation of the minor pilin at the base of the growing pilus filament.

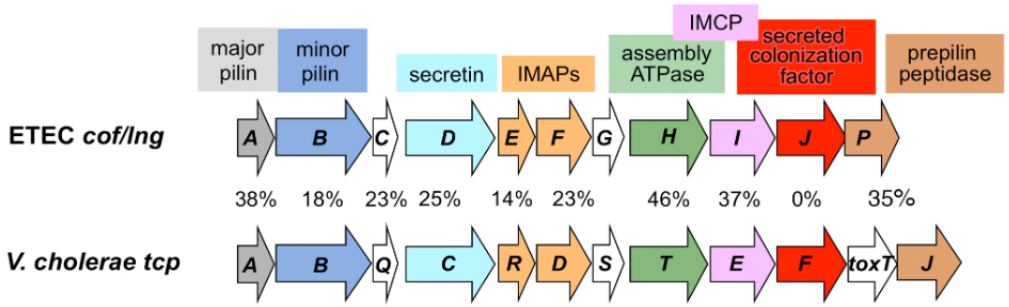

Figure 1-9 Schematic of the operons encoding the simple Type IV pili systems in *Vibrio cholerae* and ETEC

Above the operons, names of the T4P components corresponding to their respective encoding genes are shown. Sequence homology between various components the T4P systems of ETEC and *V. cholerae* is indicated. IMAPs, inner-membrane accessory proteins; IMCP, inner-membrane core proteins. Figure from (Yuen et al., 2013).

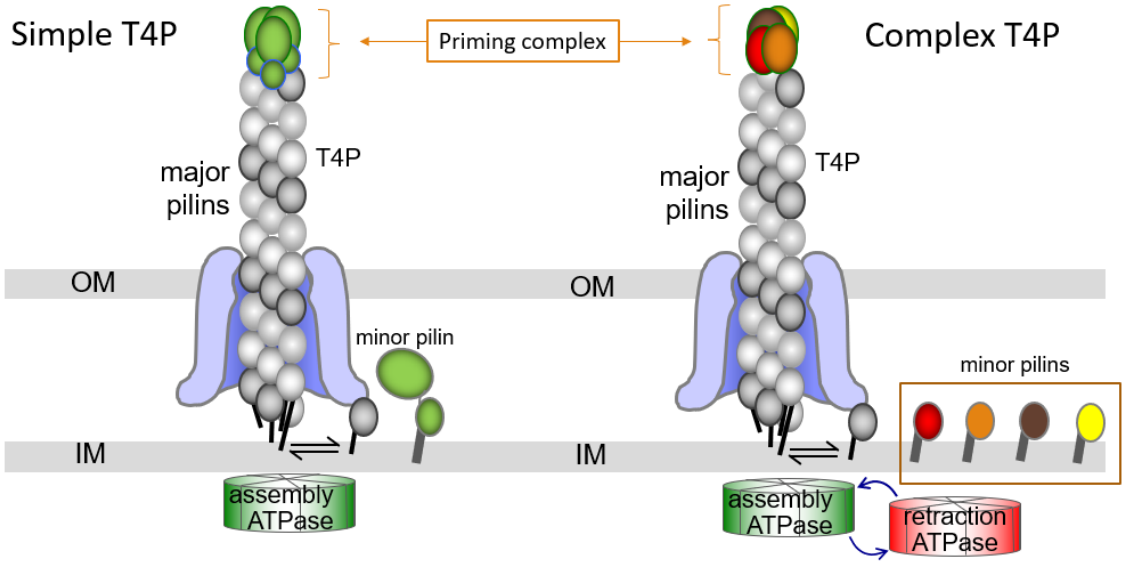

Figure 1-10 Comparison between the simple and complex Type IV pili systems
Simple T4P system has a single minor pilin which forms a homotrimer known as the priming complex at the pilus tip. While lacking a dedicated retraction ATPase, the simple T4P are retractile. Complex T4P system has a retraction ATPase and several minor pilins. Similar to the simple T4P, the minor pilins of the complex T4P are thought to form a tip-localized priming complex.

1.3.6. The minor pilins

Another common feature between the T4P and the T2S systems is the presence of a small number of minor pilins. While being at a lower abundance compared to the major pilins, the minor pilins play a significant role in the pilus functions. Mounting evidence describes the presence of the minor pilins at the pilus tip (Korotkov & Hol, 2008; Ng et al., 2016). Structurally, the minor pilins share the same canonical N-terminal hydrophobic α-helix as the major pilin, allowing them to be incorporated into the pilus filament. However, unlike the major pilins, the minor pilins also possess a C-terminal globular domain important for mediating the pilus functions.

The complex T4P systems have several minor pilins

Among the well-characterized pilus structure analysis is the minor pilin complex in the ETEC T2S system which can be used to confer some structural insight into the T4P minor pilins. The crystal structure of three minor pilins GspI, GspJ and GspK from the ETEC T2S system revealed a heterotrimer (Korotkov & Hol, 2008). These minor pilins were expressed recombinantly without their N-terminal α1Ns, and they are held together via their globular C-terminal domains, which would orient their absent α1Ns

staggered at one end of the complex. This GspIJK complex of the T2S system is thought to be structurally and functionally related to the tip-associated priming complex in the complex T4P system (Karuppiah, Thistlethwaite, & Derrick, 2016) (Fig 1-11). The largest minor pilin of the three, GspK, is the outermost pilin. GspK has a hydrophobic residue in place of the conserved Glu5 typically seen in other pilin subunits, consistent with it being the first pilin to assemble into the priming complex. As the first subunit in the pseudo-pilus, GspK does not need a Glu5 to neutralize the positive N-terminus of the previous pilin (Ng et al., 2016). Ng et al. later proposed the bulky C-terminal region of GspK blocks passage of the T2S pseudo-pilus through the secretin channel, restricting it to the periplasm. Interestingly, in systems possessing several minor pilins like those of the T2S and complex T4P systems, there is usually one minor pilin lacking the conserved Glu5, which may identify it as the outermost pilin and the first to assemble in the priming complex.

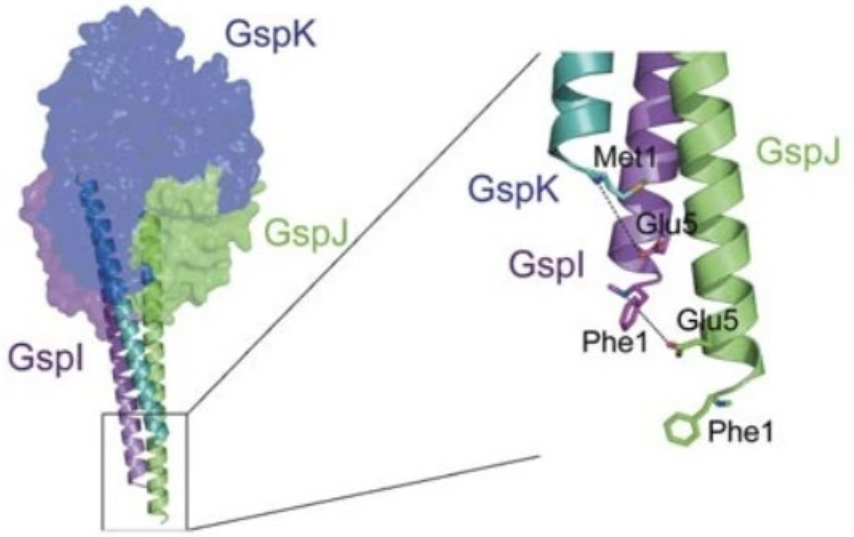

Figure 1-11 GspI-GspJ-GspK complex in the ETEC Type II secretion system
Predicted model of the GspI-GspJ-GspK heterotrimer showing their N-terminal α1-N and Glu5 interactions. Formation of the GspIJK complex in the inner-membrane is thought to initiate pseudo-pilus assembly in the ETEC T2S system. As the tip-most minor pilin, GspK does not have a Glu5 residue and is significantly larger than the other 2. Its N-terminus facilitates a salt-bridge interaction with the Glu5 of GspI. Figure adapted from Korotkov et al., 2008 and used with permission from the publisher.

One single minor pilin drives both pilus assembly and retraction in the simple T4P systems

Assembly and retraction require rapid extension and depolymerization of the pilus filament, respectively. In the simple TCP system of *V. cholerae*, the single minor pilin TcpB drives both of these antagonistic functions (Ng et al., 2016). The minor pilin CofB in the CFA/III system of ETEC most likely functions in a similar manner. Sequence alignment reveals TcpB and CofB share limited homology (Fig 1-12B). Structurally, the N-terminal pilin domain of TcpB resembles the major pilin TcpA from the same T4P system. However, TcpB has an extra C-terminal globular domain connected to the pilin domain via a flexible linker (Fig 1-12C). Both minor pilins are necessary for pilus assembly, and a CofB variant lacking the extended C-terminal domain is unable to rescue pilus assembly in a *cofB* mutant (Kolappan, Ng, Yang, Harn, & Craig, 2015).

Based on their roles in priming pilus assembly, TcpB and CofB were proposed to assemble at the pilus tips (Kolappan et al., 2015; Ng et al., 2016). The C-terminal domains of both TcpB and CofB form almost identical homotrimers in their crystal lattices (Fig 1-13) (Gutierrez-Rodarte, Kolappan, Burrell, & Craig, 2019; Kawahara et al., 2016). Kawahara et al proposed that the CofB homotrimer drives pilus assembly in the ETEC CFA/III system. In support of this model, the *V. cholerae* TCP minor pilin TcpB was shown to localize to the tip of the pilus (Fig 1-14) (Gutierrez-Rodarte et al., 2019). Tip-localized minor pilin and its homotrimeric formation suggest an assembly model for the simple T4P systems in which the minor pilin forms a homotrimer known as the priming complex at the pilus-tip which promotes pilus extension.

The *V. cholerae* minor pilin TcpB has been identified as the T4P component directly responsible for horizontal gene transfer by CTXΦ transduction (Gutierrez-Rodarte et al., 2019), introducing the gene encoding cholera toxin into what was previously a harmless aquatic microbe. CTXΦ infection of *V. cholerae* requires T4P (Waldor & Mekalanos, 1996) where the tip-associated minor pilin is recognized and bound by pIII, a coat protein exposed on the CTXΦ surface (Fig 1-15) (Gutierrez-Rodarte et al., 2019). Subsequent TcpB incorporation to the base of the pilus stalls assembly and triggers retraction, bringing the phage-encoded virulence genes into the bacterial periplasm as if it were an extension of the pilus. Integration of the virulence genes into the bacterial chromosome occurs through homologous recombination. In the competition for survival, horizontal gene transfer that confers adaptive advantage is

favored and spread throughout a species. For *V. cholerae*, acquiring the ability to produce cholera toxin greatly enhances its fitness. Profuse diarrhea caused by cholera toxin promotes the dissemination of both the bacterial host and its phage, explaining the global prevalence of infectious *V. cholerae* strains.

Interestingly, the pilus retraction mechanism facilitated by the minor pilin TcpB, and likely CofB, is dependent on the presence of their conserved Glu5 residue. A *V. cholerae tcpE5V* mutant is able to assemble pili at wild-type level but unable to retract them (Ng et al., 2016). It has been suggested that the low abundance minor pilins of the simple T4P systems not only can form priming complexes but can also incorporate into the base of the growing filament via their pilin domains (Li et al., 2012). Docking is facilitated by charge complementarity between the negative Glu5 of the minor pilin and the positive N-terminal amine of the terminal major pilin in the growing pilus filament. The protruding C-terminal domain of the minor pilin blocks passage of the pilus through the secretin channel, stalling pilus assembly and triggering depolymerization of the pilin subunits (Fig 1-15).

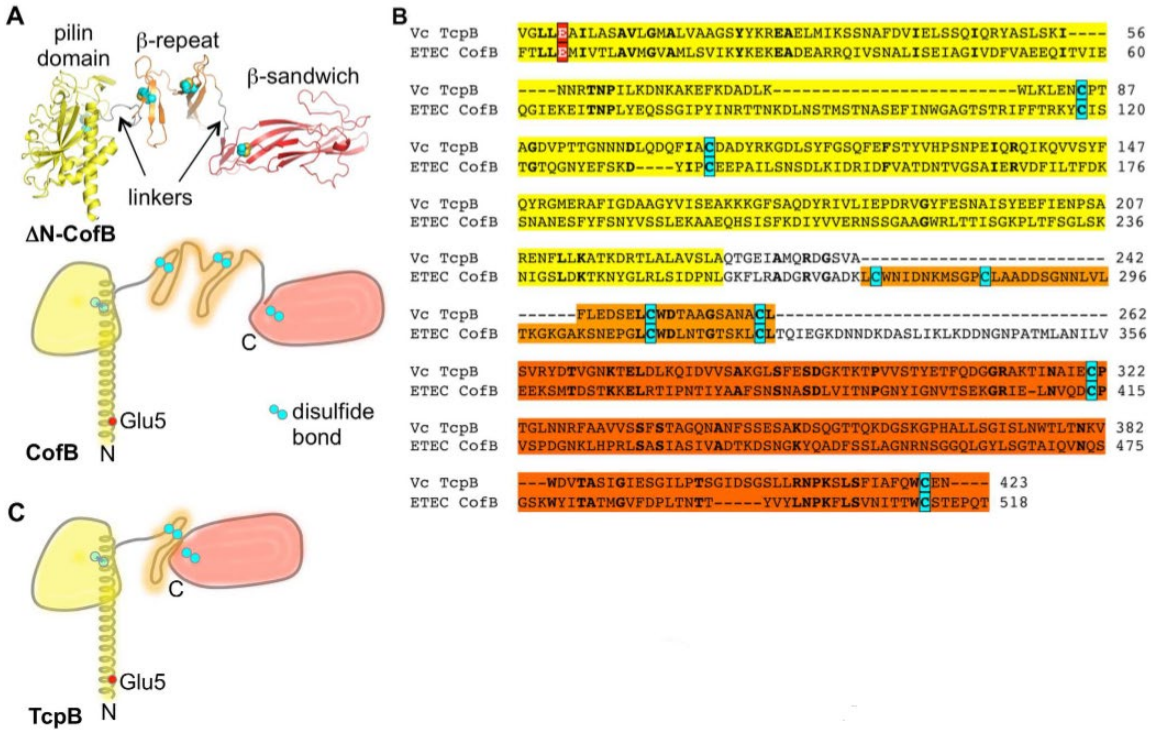

Figure 1-12 Comparison between the *Vibrio cholerae minor pilin TcpB* and the ETEC minor pilin CofB

A. Crystal structure of the N-terminally truncated CofB in cartoon representation (top) and a schematic with a full N-terminal α-helix (bottom). Cysteines are colored cyan. **B**. Sequence

alignment between TcpB and CofB by Clustal Omega and adjusted to align the cysteines. Conserved Glu5 is colored red. Identical residues are bold. Sequence is color-shaded based on corresponding representation in A. **C**. Proposed schematic of TcpB structure based on alignment with CofB. Figure from (Ng et al., 2016).

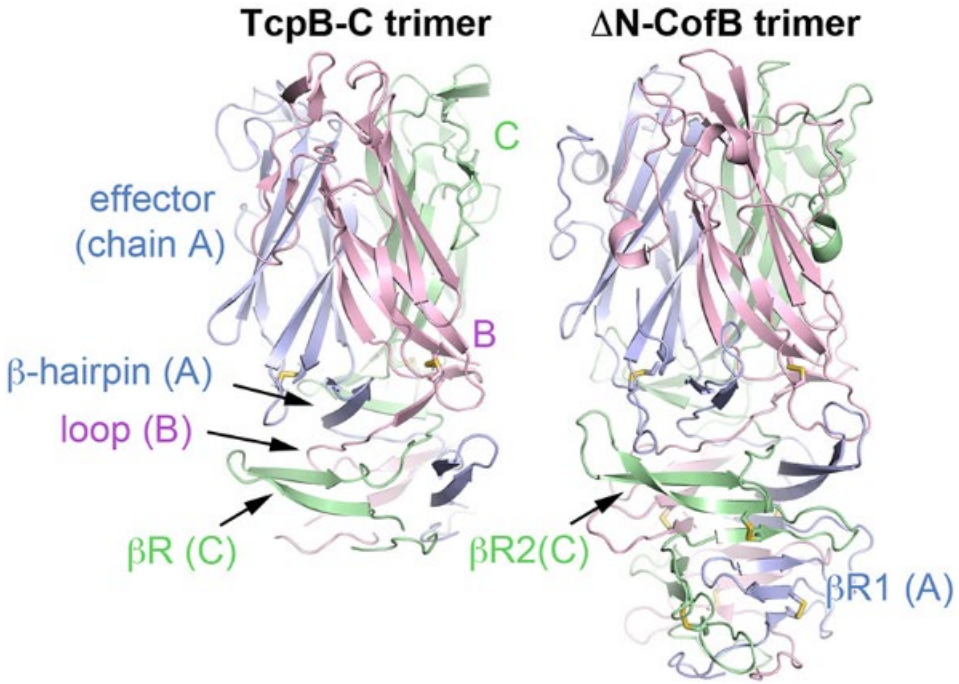

Figure 1-13 Crystal structures of the TcpB-C and ΔN-CofB trimers
The 3 subunits A, B and C of each trimer are colored differently. Only the β-repeats and the effector domains are shown in the ΔN-CofB trimeric structure. The hometrimeric formation in both TcpB-C and ΔN-CofB is stabilized mainly by hydrophobic interactions among the C-terminal β-sandwich domains (effector) and the β-repeat domains (βR). Figure adapted from Gutierrez-Rodarte et al., 2019 and used with permission from the publisher.

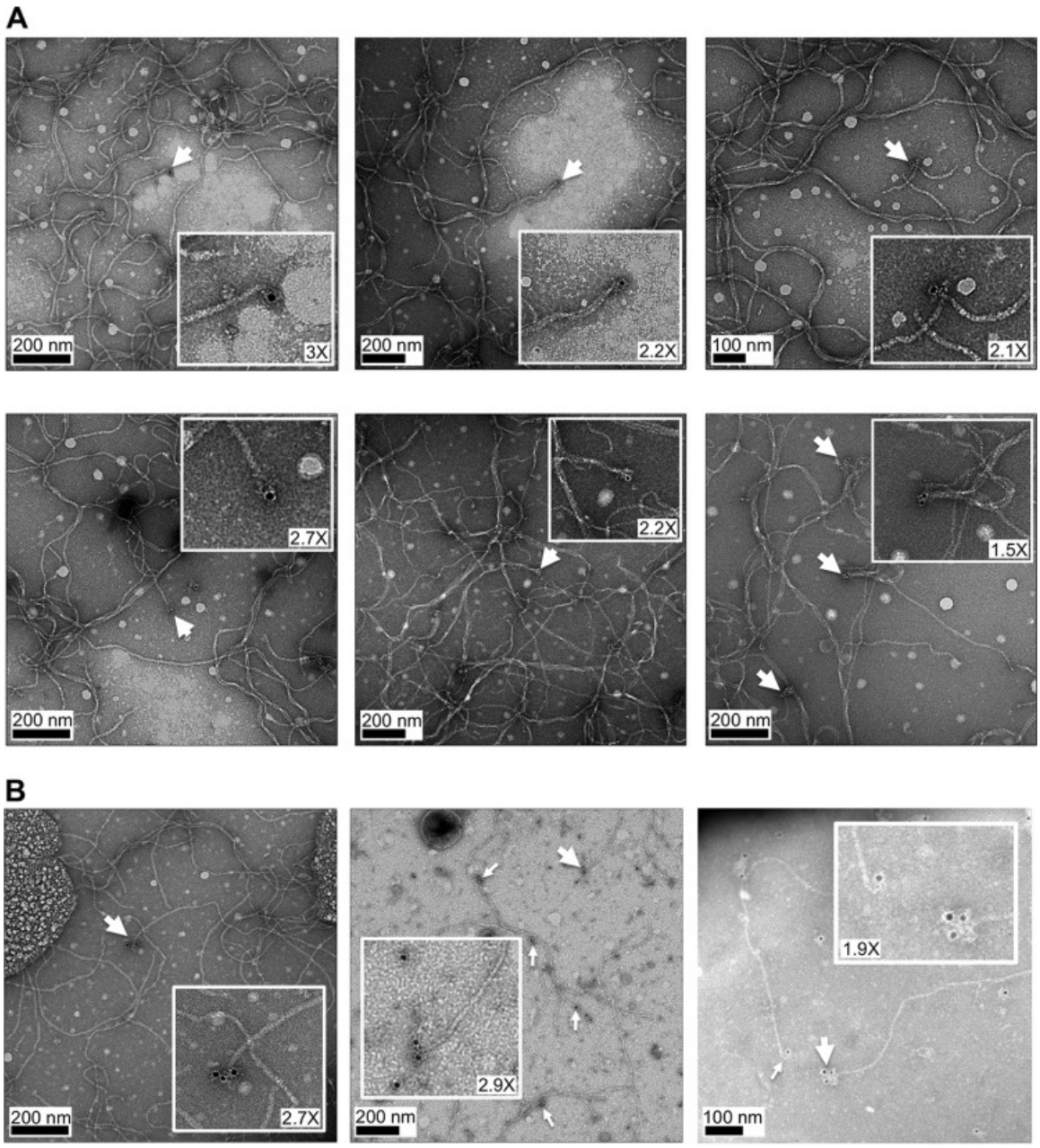

Figure 1-14 Immunogold labeling of *V. cholerae* minor pilin TcpB at the pilus tip
On copper grids, purified pili were attached and treated with chicken anti-TcpB antibody raised against ΔN-TcpB and donkey anti-chicken secondary antibody conjugated to gold nanoparticles. Grids were stained with 3% uranyl acetate and imaged by transmission EM. Presence of single (A) and multiple (B) gold particles at the tips of pili. Figure from (Gutierrez-Rodarte et al., 2019).

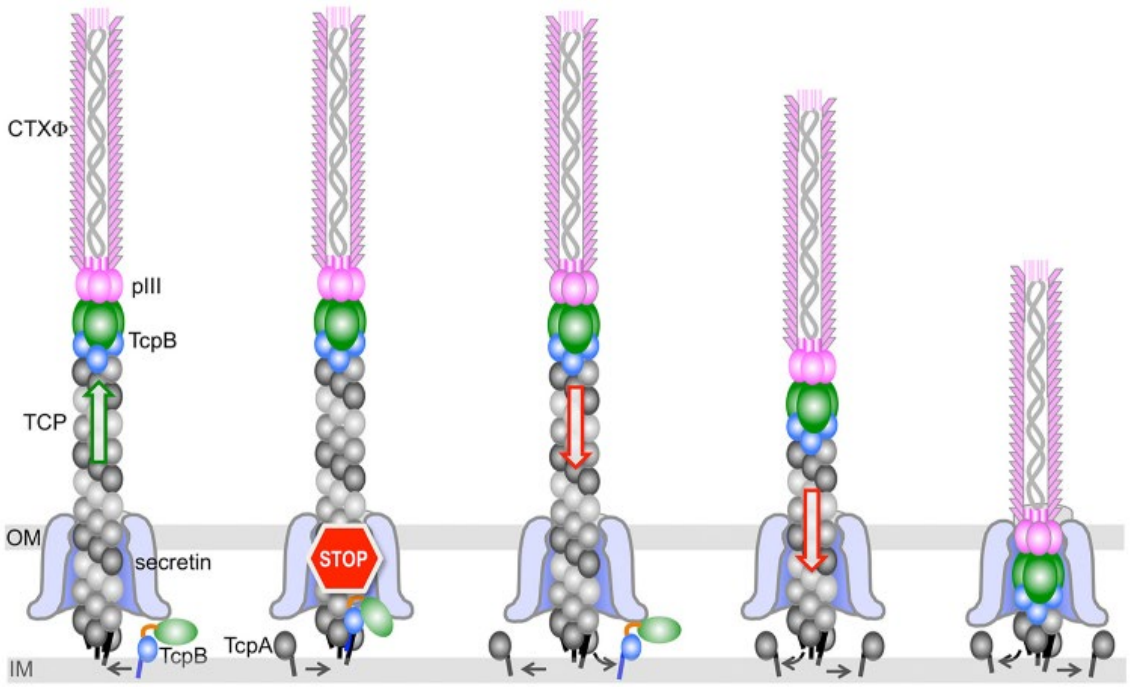

Figure 1-15 Proposed model of Type IV pili-mediated CTXΦ uptake in *Vibrio cholerae*

CTXΦ surface-exposed coat protein pIII binds the tip-localized minor pilin TcpB trimer of the growing *V. cholerae* TCP filament. Incorporation of a TcpB monomer at the base of the pilus filament stops pilus assembly. Subsequent depolymerization of the pilin subunits retracts the filament and draws the phage into the periplasm as if it were an extension of the pilus. TcpB pilin domain and effector domain are colored blue and green, respectively. Figure from (Gutierrez-Rodarte et al., 2019).

1.3.7. The T4P secreted exoproteins

As mentioned, several T4P have roles in exoprotein secretion. The strong similarities between the T4P and T2S systems suggest that their secretion mechanisms are similar. The exoproteins secreted by the T4P systems have been linked predominantly to infectious pathogens. *Francisella* virulence involves its T4P-mediated secretion of 2 chitinases (Hager et al., 2006). Extracellular protease secretion in *Dichelobacter nodosus* is dependent on its T4P biogenesis (Han et al., 2007). Secretion of two colonization factors from the well-established enteric pathogens, *V. cholerae* and ETEC, the causative agents of diarrhea, is mediated by their respective T4P systems (Kirn et al., 2003; Yuen et al., 2013).

V. cholerae exoprotein TcpF

The *V. cholerae* T4P secreted exoprotein TcpF is made up of 317 amino-acid residues with no apparent sequence homology to other known proteins. TcpF has been shown to be important for bacterial colonization of the infant mouse intestine (Kirn et al., 2003). While a *V. cholerae* mutant lacking the *tcpF* gene can still assemble T4P, it is significantly defective in colonization in the infant mouse model, about 100,000-fold less efficient compared to the wild-type strain (Kirn & Taylor, 2005). Secreted TcpF from a wild-type strain cannot rescue the colonization phenotype of this mutant. In fact, the colonization defect between a *V. cholerae* mutant lacking TcpF and another lacking its T4P is comparable. Interestingly, *in vitro* assays revealed the T4P morphology is independent of TcpF secretion, as shown by the bundled, laterally associated pilus fibres expressed by the *V. cholerae* Δ*tcpF* mutant (Kirn et al., 2003). Taken together, these observations suggest *V. cholerae* T4P and its exoprotein TcpF are both required for efficient colonization. However, the molecular mechanism by which TcpF facilitates *in vivo* microcolony formation remains unknown, though it is likely that the T4P and TcpF aid colonization in different ways.

The crystal structure of TcpF at 2.4 Å shows the exoprotein is made up of 2 discrete domains, the C-terminal domain (CTD) and the N-terminal domain (NTD) joined via a linker (Fig 1-16) (Megli et al., 2011). The first 25 residues of the N-terminal segment were not resolved is the crystal structure due to disorder. Two cysteine residues, C34 and C47, form a disulfide bond in the NTD of TcpF. The NTD and CTD come together and form a curved elongated protein. Each domain makes minimal contact with the other. The CTD of TcpF was crystallized on its own and its structure matches that within the full-length protein (Megli et al., 2011). Monoclonal antibody that binds the cleft between the two domains block *V. cholerae* colonization. TcpF shares no identifiable structural homology to other known proteins and its function in *V. cholerae* colonization remains unknown.

ETEC exoprotein CofJ

CofJ is similar in length to TcpF and is encoded on the ETEC *cof* operon in the same position as the *tcpF* gene in the *V. cholerae tcp* operon, but these proteins share no sequence similarity. CofJ is secreted by the ETEC CFA/III system as a potential membrane-bound factor (Yuen et al., 2013). Yuen et al. proposed CofJ association with

the phospholipid bilayer promotes ETEC pathogenesis by forming ring-like openings to induce host membrane leakage. The crystal structure of CofJ shows a single domain predominated by β-sheet secondary structures, overall forming an elongated β-sandwich core (Fig 1-16). The two N-terminal and C-terminal ends of CofJ are covalently connected by a disulfide bond between two terminal cysteine residues, C26 and C323.

To help elucidate the role of CofJ in bacterial pathogenesis, Yuen et al. identified its structural homologues, with top matches being the pore-forming toxins perfringolysin O from *Clostridium perfringens* and intermedilysin from *Streptococcus intermedius*. Both perfringolysin O and intermedilysin are secreted monomeric proteins that oligomerize to form ring-like openings on eukaryotic cell membranes and cause membrane leakage (Tweten, 2005). In support of CofJ-membrane association, direct interaction between CofJ and cultured epithelial cells, HeLa and Caco2, has been observed (Yuen et al., 2013). Furthermore, binding assay revealed bands corresponding to the molecular masses of CofJ dimer and trimer in the presence of Caco2 epithelial cells, suggesting the tendency for CofJ to oligomerize. Taken together, these observations support the potential role of CofJ as a pore-forming toxin and a putative colonization factor in the ETEC CFA/III system.

Comparison between TcpF and CofJ

At first glance, the crystal structures of the TcpF and CofJ indicate no structural similarity (Fig 1-16). However, their origins suggest otherwise. TcpF and CofJ are encoded in syntenic positions on the respective T4P operons of *V. cholerae* and ETEC. *V. cholerae* and ETEC are both enteric Gram-negative pathogens and the causative agents for diarrhea. Furthermore, both TcpF and CofJ are secreted by their corresponding simple T4P systems. While TcpF shares no homology to any known proteins, it has been demonstrated to be a necessary component for pathogenic *V. cholerae* colonization in infant mice (Kirn et al., 2003). Based on the observation that CofJ shares structural homology with 2 pore-forming toxins perfringolysin O from *Clostridium perfringens* and intermedilysin from *Streptococcus intermedius,* CofJ has been implicated in eukaryotic membrane association, likely by oligomerization to form ring-like openings (Yuen et al., 2013). How TcpF and CofJ determine *in vivo* microcolony formation remains a mystery, though it is likely that both share some overlapping molecular features. The overall dimensions of TcpF and CofJ are similar (Fig 1-16), but

narrower than the pilus filament (80 Å), and thus could be accommodated in the secretin channel. The N-terminal ~23 amino acid segment of both TcpF and CofJ mature proteins are unresolved in their crystal structures (Megli et al., 2011; Yuen et al., 2013). To determine whether this segment in CofJ was absent due to proteolysis or was simply disordered, Yuen et al. performed liquid chromatography-mass spectrometry which confirmed its presence in the crystal structure, ruling out its proteolysis. Interestingly, comparison between the N-terminal region of TcpF and CofJ reveals the presence of several serine and threonine residues in both flexible segments. In *V. cholerae*, this segment was shown to be essential for TcpF secretion by the T4P system (Krebs, Kirn, & Taylor, 2009). Furthermore, the tyrosine at position 5 within this segment was identified as TcpF secretion determinant (Megli & Taylor, 2013).

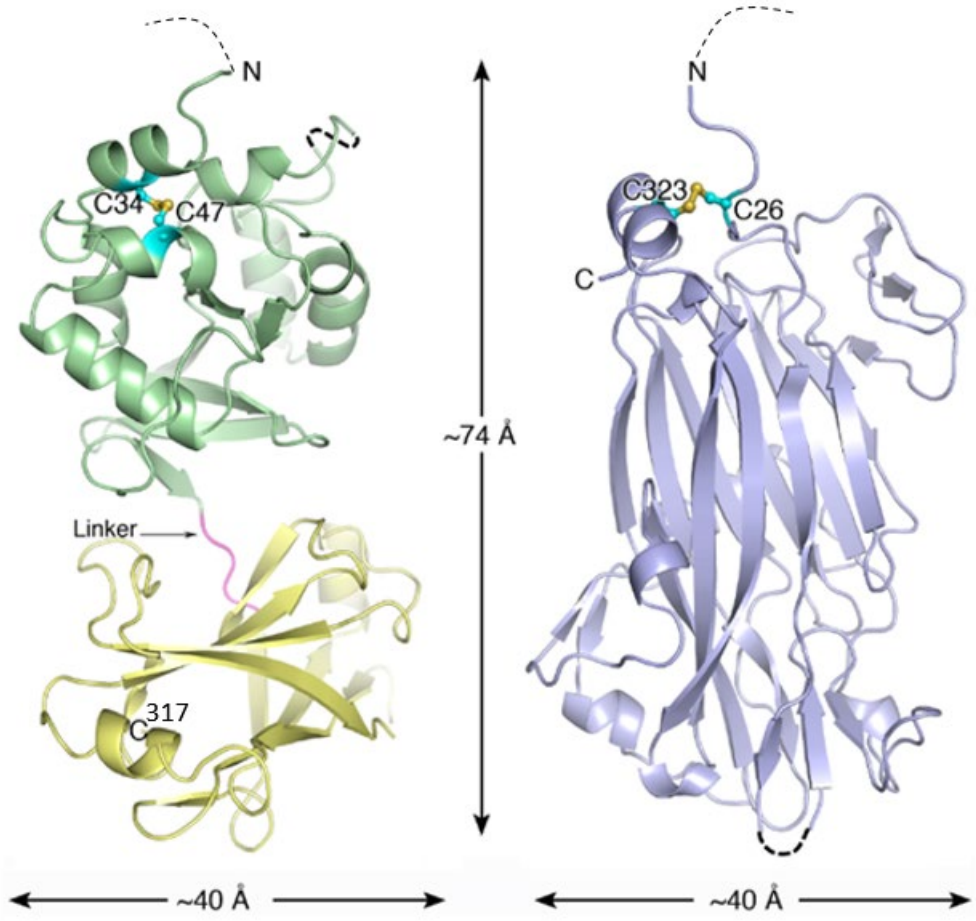

Figure 1-16 Comparison between the *Vibrio cholerae* and ETEC exoproteins TcpF and CofJ

The approximate dimensions of the 2 proteins are indicated. Although CofJ (right) and TcpF (left) structures suggest no similarity, their dimensions are comparable. The N-terminal domain of TcpF (green) is connected to its C-terminal domain (yellow) via a flexible linker (pink). Cysteine

residues forming disulfide bonds are indicated. Dashed lines represent unresolved regions. Figure from (Yuen et al., 2013).

TcpF periplasmic translocation involves the Sec-machinery (Kirn et al., 2003) which recognizes and cleaves the signal peptide in the N-terminus of the TcpF pre-protein. CofJ is most likely processed in a similar manner by the Sec-machinery due to its characteristic Sec-dependent signal peptide, with a positively charged amino end, a central hydrophobic segment, and a Glycine and Alanine at the -3 and -1 positions relative to the cleavage site (Paetzel, 2019) (Fig 1-17). Following the signal peptide is the unresolved flexible segment in TcpF and CofJ crystal structures (Megli et al., 2011; Yuen et al., 2013). Recent evidence from Oki et al. suggests this segment mediates interaction between the exoprotein CofJ and the tip-localized minor pilin CofB in the ETEC CFA/III system (Oki et al., 2018).

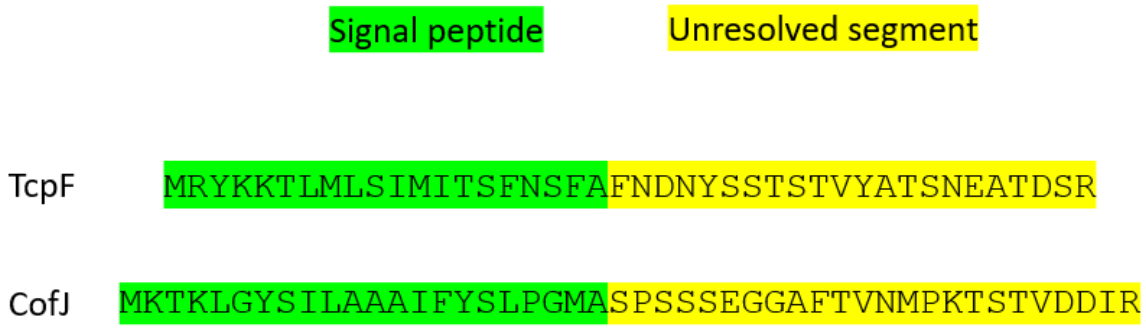

Figure 1-17 Comparison between the signal peptide and the unresolved segment of TcpF and CofJ

TcpF and CofJ have a characteristic Sec-dependent signal peptide (green) in the pre-proteins, with a positively charged amino end, a central hydrophobic segment and small aliphatic residues close to the cleavage site. In the TcpF and CofJ crystal structures, there is an unresolved flexible segment (yellow) in the N-terminus of the mature proteins (Megli et al., 2011; Yuen et al., 2013). CofJ unresolved segment is suggested to be required for interaction with the tip-localized minor pilin CofB in ETEC (Oki et al., 2018).

1.3.8. In the ETEC Type IV pili system, the minor pilin CofB binds the exoprotein CofJ

It has been established that in the simple T4P systems of *V. cholerae* and ETEC, the minor pilins form a homotrimer, termed the priming complex, at the pilus-tip that is essential for pilus functions (Kawahara et al., 2016; Kolappan et al., 2015). To accomplish exoprotein secretion, the pilus exerts piston-like motions to push substrates across the outer membrane. At some point during the process, the pilus-tip makes contact with the exoprotein, suggesting the tip-localized minor pilin as a potential

exoprotein binding partner. Recently, in the ETEC CFA/III system, emerging evidence has arisen in support of this interaction model. Oki et al. showed via a pull-down assay that the ETEC minor pilin CofB binds to its cognate exoprotein CofJ (Oki et al., 2018). They solved a crystal structure of the CofB trimer bound to three copies of a synthetic peptide representing the flexible N-terminal segment (24 residues) that was not resolved in CofJ crystal structure (Fig 1-18A). Particularly, it seems that CofJ residues 5 to 15 are important for recognition by the CofB trimer due to its well-defined electron density (Fig 1-18B). Presumably this flexible segment is stabilized when bound to the CofB trimer by hydrophobic and hydrogen-bonding interactions, wedged into the hydrophobic groove between two CofB C-terminal β-sandwich domains. This observation is consistent with a previous study identifying the segment rich in serines and threonines in the CofJ N-terminus important for its secretion (Krebs et al., 2009), likely due to the hydrogen-bonding capacity from the side-chain hydroxyl groups (Fig 1-18B). CofJ residue 10, a phenylalanine, also seems to facilitate hydrophobic interaction with the CofB trimer. Interestingly, Oki et al. also showed that a CofJ variant lacking this segment, ΔN24-CofJ, was unable to bind the CofB trimer.

Further sedimentation analysis by Oki et al. demonstrated CofJ-CofB complex exists as a heterotetramer in solution with one CofJ and three CofB molecules. This led to the proposed model of CofJ interaction with the CFA/III pilus by binding to the the tip-associated CofB trimer (Fig 1-19A). We further hypothesize that the CofJ-CofB interaction must be transient to allow for CofJ dissociation from the pilus, resulting in its secretion. CofJ exposed tyrosine cluster is thought be involved in target cell recognition (Fig 1-19B), consistent with earlier prediction about CofJ potential role as a pore-forming toxin on eukaryotic cell membrane (Yuen et al., 2013). Further understanding of the T4P-mediated secretion in *V. cholerae* and ETEC could provide valuable insight into the underlying mechanism in related secretion systems.

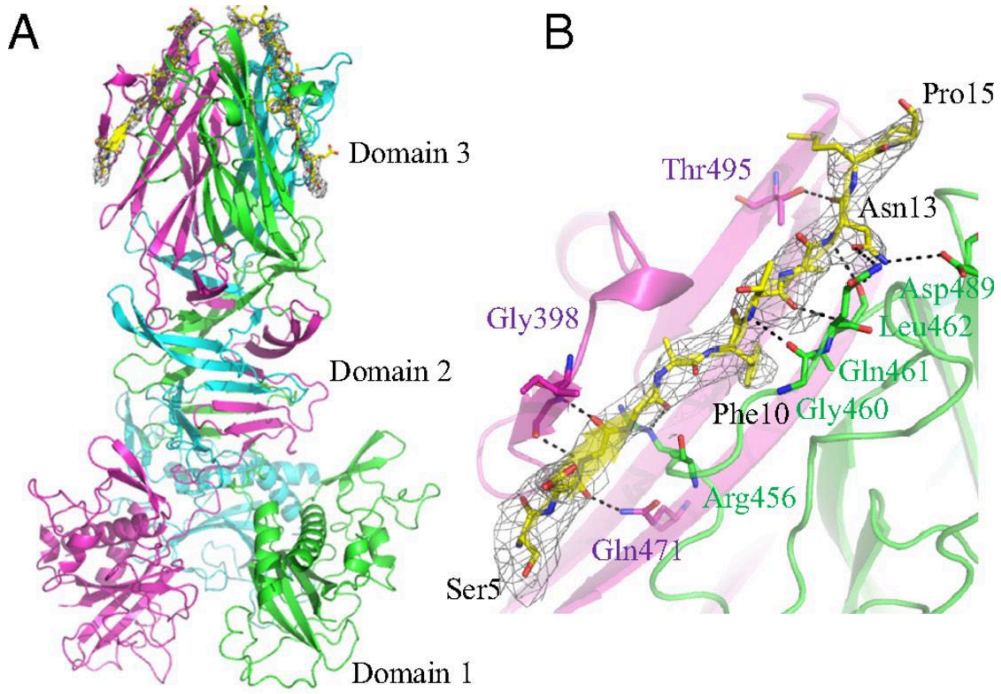

Figure 1-18 Crystal structure of the CofJ (1-24)-CofB complex
A. Side view cartoon representation of the CofB trimer (cyan, magenta and green) with bound CofJ (1-24) shown as yellow sticks. In the CofB trimer, domain 1 corresponds to the N-terminal pilin domain, domain 2 represents the β-repeat and domain 3 comprises the C-terminal β-sandwich effector. **B**. Close up view of the binding interface between CofJ (5-15) and domain 3 of CofB trimer. Color schematics is based on A. Dashed lines are hydrogen bonds. CofB residues that participate in hydrogen bonding are shown in sticks. Figure from (Oki et al., 2018).

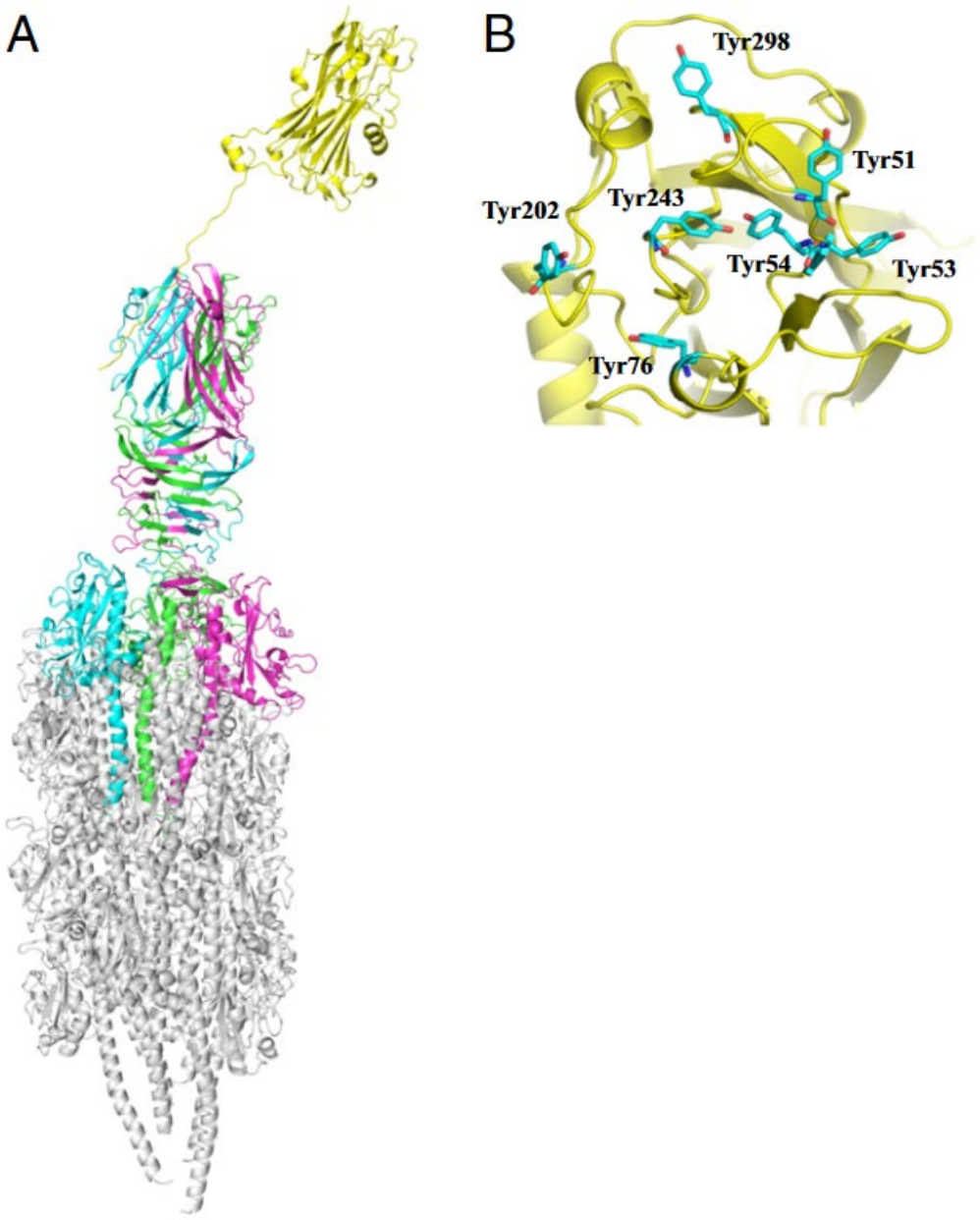

Figure 1-19 Proposed model of the CofJ-CFA/III pilus complex in ETEC
A. Side view cartoon representation of the CofJ-CFA/III pilus complex by superimposing the CofJ (1-24)-CofB crystal structure on the known CFA/III pilus structure. The 3 CofB monomers are colored cyan, green and magenta. The CofJ monomer is depicted in yellow. **B**. Close-up view of CofJ from A, showing seven aromatic tyrosine residues. Figure from (Oki et al., 2018).

1.4. The T4P system is evolutionarily related to the T2S system

The T4P system is closely related to the T2S system which has been well-established in its role as a dedicated secretion system. While the Type II secreted

substrates are not only diverse in their functions but also in their shapes and sizes, the less well-characterized Type IV secreted substrates are linked predominantly to bacterial pathogenesis. Both the T4P and T2S systems share many homologous components, including the cytoplasmic ATPase motor, the inner-membrane platform, the pilus and the outer-membrane secretin channel. The characteristic difference lies in their pilus filaments. While the T2S pseudo-pili are limited to the periplasm for the sole purpose of aiding secretion by piston-like motions (Sandkvist, 2001), the T4P system has long and laterally associated surface-displayed pilus bundles that regulate DNA uptake, twitching motility, surface adhesion as well as secretion (Craig et al., 2019). Pilus assembly in both systems is driven by their respective cytoplasmic ATPase motors in the form of ATP hydrolysis. Lack of a dedicated retraction ATPase motor links the simple T4P system to the T2S system where retraction was proposed to be mediated by the minor pilins (Korotkov & Hol, 2008; Ng et al., 2016). Exoproteins secreted by both systems must pass through 2 barriers, the inner-membrane and the outer-membrane sequentially. The presence of a cleavable N-terminal signal peptide on the pre-secretory protein aids its translocation to the periplasm via the Sec-machinery, after which it can be processed for secretion via the outer-membrane secretin channel. In some enteric Gram-negative pathogens, the T4P and T2S systems both contribute to bacterial pathogenesis (Fig 1-20). The pathogenic *V. cholerae* and ETEC use their T2S systems to secrete cholera toxin and heat-labile toxin, respectively (Sandkvist et al., 1997; Tauschek, Gorrell, Strugnell, & Robins-Browne, 2002). These enterotoxins induce a similar cascading response in the host cell, causing massive water efflux by raising cellular cyclic cAMP levels. Though poorly understood, the pathogenesis of ETEC and *V. cholerae* has been shown to be aided by their respective T4P-secreted exoproteins CofJ and TcpF.

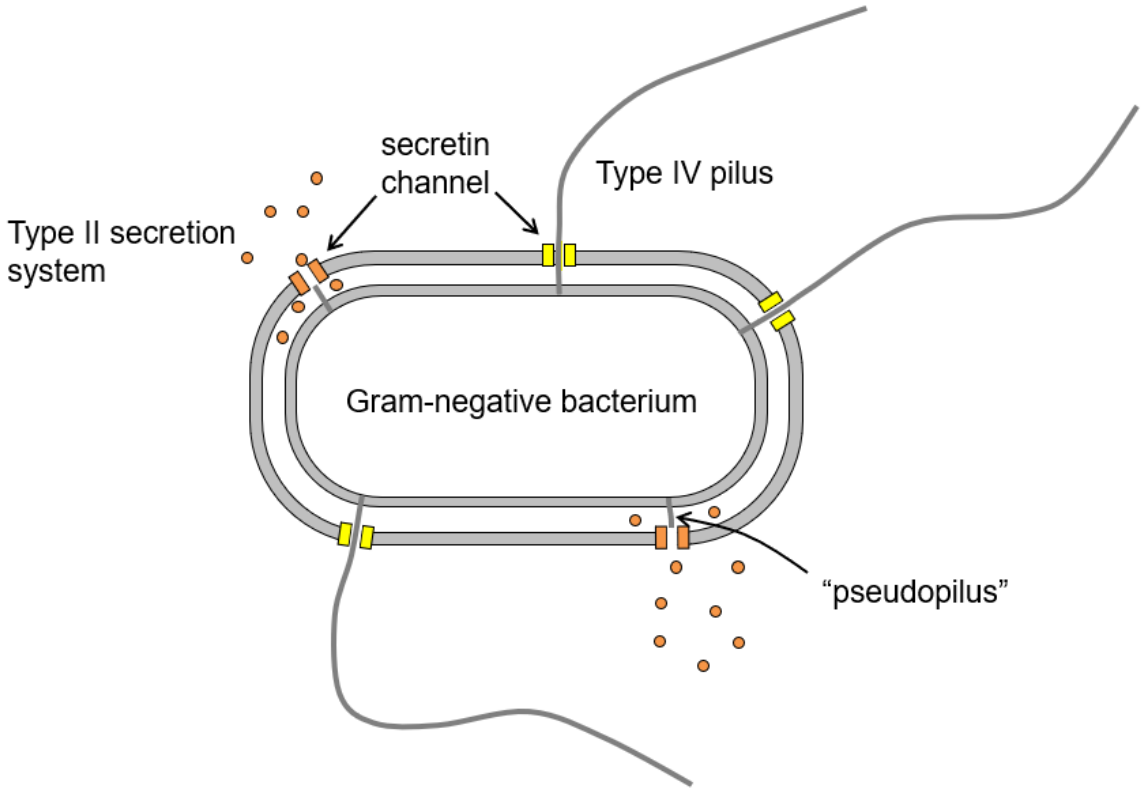

Figure 1-20 ***Vibrio cholerae*** **Type II secretion and Type IV pili systems are both required for its pathogenesis**

V. cholerae T2S system utilizes a pseudo-pilus to secrete cholera toxin (orange) across the outer-membrane secretin channel. On the other hand, the presence of its surface-displayed T4P is essential for microcolony formation in the host epithelial cells. Image courtesy of Lisa Craig.

1.5. Hypotheses and research objectives

The T4P are utilized by many species across the bacterial kingdom to perform diverse functions aiding their survival and pathogenesis. Among its numerous roles contributing to bacterial virulence is substrate secretion, an often-overlooked aspect when compared to the well-established T2S system. Understanding how the T4P machinery selectively chooses its secreted targets may provide a way to predict and exploit the closely related T2S system. The overarching goal of this thesis is to understand exoprotein secretion in the simple T4P system of *V. cholerae*. I hypothesize that the flexible N-terminal domain of the exoproteins, TcpF in *V. cholerae* and CofJ in ETEC, is the secretion determinant that transiently interacts with the minor pilin trimer located at the pilus tip for export via the secretin channel.

The specific objectives of this thesis are:

Aim 1 – To test the role of the exoprotein N-terminus as an export signal (ES).

Aim 2 – To investigate the relationship between pilus assembly and exoprotein expression.

Aim 3 – To test whether there exists an interaction between the secreted exoprotein and the tip-associated minor pilin trimer.

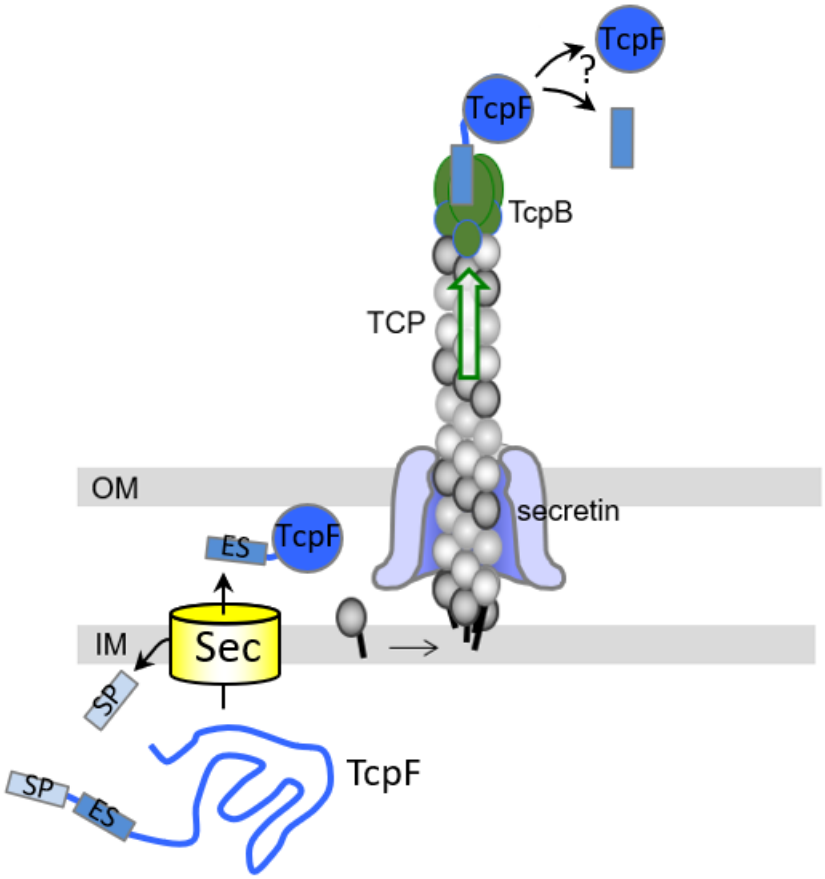

Figure 1-21 Proposed Type IV pili-mediated TcpF secretion in *Vibrio cholerae*
The unfolded exoprotein precursor TcpF is synthesized in the cytoplasm with intact signal peptide (SP) and export sequence (ES). Upon translocation across the inner-membrane (IM), the SP is processed and cleaved by the Sec machinery. In the periplasm, TcpF adopts its native folded structure and binds to the tip of the pilus via ES interaction with the minor pilin TcpB and is delivered across the outer-membrane (OM). TcpF dissociates from the pilus once in the extracellular space. The ES may be removed to release exoprotein.

Chapter 2. Methods

2.1. Bacterial strains and growth media

Bacterial strains, plasmids, and primers used in this study are described in Table 1. All strains were kept at -80° C in lysogeny broth (LB) containing 15% glycerol. *V. cholerae* strains were grown in classical TCP-promoting conditions: LB pH 6.5 at 30° C with aeration and vertical rotation at 60 rpm (Taylor et al., 1987). *E. coli* strains were grown in LB at 37° C with agitation at 250 rpm. Where necessary, strains were grown with added antibiotics appropriate for plasmid selection at the following concentrations: ampicillin (Ap), 100 ug/ml; chloramphenicol (Cm), 20 ug/ml. DNA manipulations were performed based on standard molecular and genetic techniques. Plasmid expression of exoprotein was induced using either arabinose or rhamnose at the indicated concentrations.

2.2. Construction of vectors for various exoprotein clones

To create various exoprotein derivatives for the purpose of secretion characterization, gene fragments encoding desired constructs of TcpF (*V. cholerae*), CofJ (ETEC) were PCR-amplified for insertion into vector plasmid pJMA10.1. The vector p*cof* contains the entire ETEC *cof* operon inserted into the cloning vector pACYC184 (ATCC) between restriction enzyme sites EcoNI and EagI-HF. p*cof* was transformed into electro-competent *E. coli* HB101 cells, and subsequent transformants were selected with Cm. The *cofJ* gene was disrupted in p*cof* using the γ-RED recombinase system used previously to generate the p*cof*Δ*cofA* construct (Yuen et al., 2013). Other vectors in this study, including pJMA10.1, were derived from the pBAD22 backbone in which the arabinose inducible *ara*C promoter had been replaced with the rhamnose-inducible promoter P_{rhaB} from *E. coli* BL21 chromosomal DNA. pJMA10.1 contains an ApR marker. Genes *tcpF*, *cofJ*, and their corresponding signal peptides (SP) and export sequences (ES) were PCR-amplified from ETEC 31-10 and *V. cholerae* O395 genomic DNA with the Q5 DNA polymerase (New England Biolabs). The gene fragment encoding the SP and ES of TcpF was PCR-amplified with primers TcpF-F1-FP and TcpF-F1-RP. The gene fragment encoding the SP and ES of CofJ was PCR-amplified with primers CofJ-F1-FP and CofJ-F1-RP. Depending on the particular construct design, the PCR product

was purified, digested with KpnI, XbaI or HindIII, and ligated into pJMA10.1 at the corresponding restriction sites with T4 DNA ligase. All constructs were verified by DNA sequencing (GENEWIZ). Transformants grown on LB-Cm/Ap or LB/Ap plates were subsequently chosen for secretion assay.

2.3. Electroporation transformation

Electro-competent cells are prepared prior and made ready for transformation by thawing on ice for 15 minutes. Purified plasmid (5 ug) is pipetted to thawed electro-competent cells and electroporated at 1.7 kV (E. coli) or 2.5 kV (*V. cholerae*). Electroporated cells are transferred to 700 ul of LB broth in a 1.5 ml Eppendorf tube, shaken at 37° C for 1 hour and grown on LB plates containing appropriate antibiotics. The *E. coli* mutant lacking the *cofJ* gene, HB101 p*cof*Δ*cofJ* and *V. cholerae* mutant lacking the *tcpF* gene, O395 Δ*tcpF* were studied strains for our secretion assays.

2.4. Assessment of pili assembly by autoagglutination assay in *V. cholerae*

V. cholerae cells were grown in 2ml LB with antibiotics if needed for 5 hours at 37° C with vigorous shaking. Cell concentrations were normalized to the 600 nm optical density (OD_{600}) of 0.01. Normalized cultures were diluted 1/1000 in 3ml LB pH 6.5 in clear glass tubes and grown for 15 hours in the Ferris wheel at 30° C with a vertical rotation of 60 rpm. Where appropriate, arabinose or rhamnose was added to complement exoprotein expression at the indicated induction levels. Cultures were allowed to settle at room temperature for 30 minutes before visual assessment of the autoagglutination phenotype and measurement of the OD_{600} readings. Experiments were performed in triplicate.

2.5. Exoprotein secretion assay by immunoblot

Fractions of overnight cell cultures were analyzed by SDS-PAGE and immunoblotting to assess exoprotein expression and secretion for the various exoprotein constructs. Samples were centrifuged at 8000 rpm to obtain visible cell pellets. Supernatant (SUP) fractions to assess exoprotein secretion were collected and filtered using a 0.22 µm syringe-drive filter (Pall) to remove remaining cells. Whole cell (WC)

fractions to assess total exoprotein production were obtained by resuspending cell pellets in 1x Phosphate-buffered Saline (PBS). Each fraction was mixed with sample buffer (62.5 mM Tris pH 6.8, 5% β-mercaptoethanol, 2.5% SDS, 10% glycerol, 0.002% bromophenol blue) and heated at 94° C for 10 minutes. 16 µl sample was loaded onto 15% SDS polyacrylamide gels. 5 µl WesternC ladder (Bio-Rad) was used as molecular markers. After running at 130 V for 1 hour 45 minutes, proteins were transferred onto a polyvinylidene difluoride (PVDF) membrane (Bio-Rad) in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol) at 4° C in a wet transfer apparatus (Bio-Rad). The membrane was blocked overnight at 4° C in blotto (5% skim-milk in Tris-buffered saline, 0.1% Tween). Exoproteins were detected with rabbit polyclonal antibodies raised against full-length TcpF or CofJ. Goat anti-rabbit secondary antibodies conjugated to horseradish peroxidase (HRP) (Jackson ImmunoResearch) were used to bind primary antibodies. The protein ladder was detected using Strep-Tactin-HRP conjugate (IBA). HRP signal was developed using the SuperSignal West Pico substrate (Thermo Fischer). Blots were visualized by chemiluminescence with the ChemiDoc MP Imaging System (Bio-Rad).

2.6. Pull-down assay

Purified recombinant protein pairs that were used in the pull-down assays: ΔN-TcpB, TcpB C-terminal domain (234-423) with an N-terminal His tag, degraded TcpF with a His tag, TcpF-ΔES and full-length TcpF. Protein pairs of interest were incubated at 37° C for 1 hour, loaded onto Ni-NTA agarose beads (QIAGEN) equilibrated with buffer (10 mM Tris-HCl pH 8 100 mM NaCl). Beads were washed 3 times with buffer containing 10 mM imidazole. Bound proteins were eluted from the beads with buffer containing 500 mM imidazole. Collected fractions were visualized by SDS-PAGE and Coomassie staining.

Chapter 3. Results

While the crystal structures of TcpF and CofJ suggest no similarity, each possesses a flexible N-terminal segment of ~23 residues with implication for T4P-mediated secretion. The crystal structure of the complex between the CofB trimer and the first 24 residues of CofJ suggests this segment is important for CofJ-CofB interaction (Oki et al., 2018). To determine the role of this flexible segment in the T4P-mediated secretion, I created several different TcpF and CofJ constructs by cloning their respective genes into the pJMA10.1 expression vector. The resulting vectors containing the appropriate genes were introduced into relevant *V. cholerae* and ETEC mutants by electroporation. Secretion assays were performed to collect desired fractions: whole cell represents total exoprotein in the cells, supernatant indicates total secreted exoprotein. Where appropriate, the WC fraction was treated with the antibiotic polymyxin B to isolate the periplasmic fraction from the spheroplast. Exoproteins were identified and quantified in these fractions by SDS-PAGE and immunoblotting.

Exoprotein secretion requires its transport across 2 membrane barriers and thus, I propose that they possess two distinct secretion signals, the signal peptide, SP, which is recognized by the Sec machinery, and the export signal, ES, which is recognized by the T4P machinery. TcpF and CofJ both have a canonical SP that is positively charged at the N-terminus, hydrophobic in the middle and possesses small aliphatic residues at the C-terminus (Paetzel, 2019). TcpF was shown to be translocated to the periplasm by the Sec-machinery (Kirn et al., 2003). We propose that the ES immediately follows the SP in the pre-proteins (Fig 3-1A). Together, these secretion signals are referred to as the N-terminal segment (NTS). SP is cleaved from the pre-proteins by the signal peptidase during translocation across the inner membrane (Fig 3-1B), and the new N-terminus, ES, is recognized by the T4P machine for translocation of the exoprotein through the secretin channel in the outer membrane and into the extracellular space. I refer to this second step as “export”. All exoprotein constructs were created based on this model (Fig 3-1B). Deletion constructs lack the SP, ES or both (the NTS). Swap constructs have their native NTS replaced by that of their exoprotein “partner” from the heterologous T4P system: the NTS of *V. cholerae* TcpF replaced by that of ETEC CofJ and vise versa. Though the “mature” exoprotein technically includes the ES, I use this term here to describe the exoprotein minus its secretion signals.

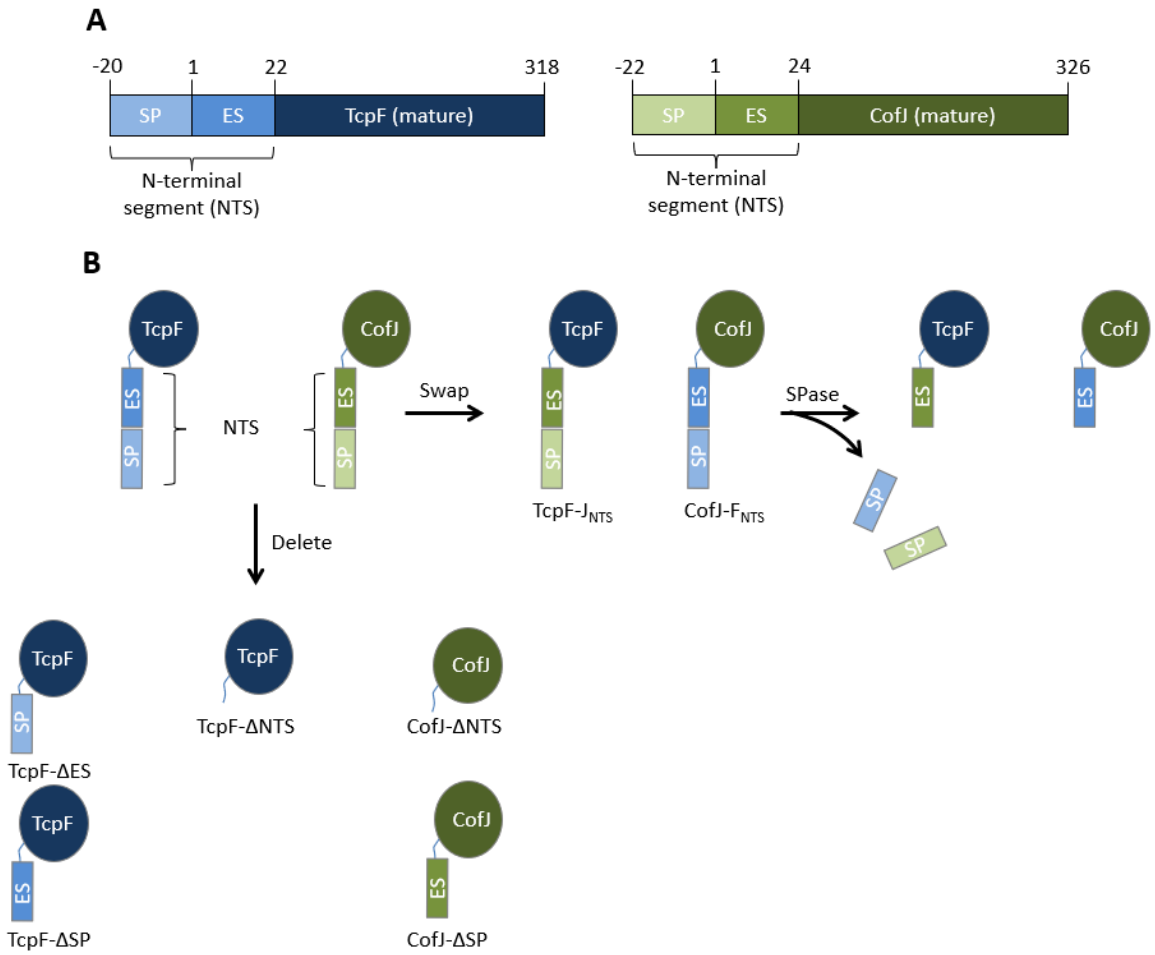

Figure 3-1 Schematic of TcpF and CofJ constructs used for secretion assays
A. TcpF and CofJ pre-proteins are shown as bars with indicated signal peptide (SP) and putative export sequence (ES). The N-terminal segment (NTS) encompasses both the SP and ES. **B**. Different TcpF and CofJ constructs used for secretion assay. SPase indicates SP process and cleavage by the Sec-machinery.

3.1. T4P-mediated secretion in *V. cholerae* and *E. coli* is specific for autologous exoproteins

Though neither structural similarity nor sequence homology exists between the *V. cholerae* exoprotein TcpF and the ETEC exoprotein CofJ, they are both transported from the periplasm to the extracellular space – exported – via their T4P machinery (Kirn et al., 2003; Yuen et al., 2013). To determine whether secretion of a substrate is specific to its native T4P system, I tested the ability of *V. cholerae* to secrete CofJ and vice versa. The ETEC CFA/III system is expressed in a non-pathogenic *E. coli* strain, HB101

(referred to in the figures as Ec), by cloning the entire *cof* operon into plasmid pAYCYC184, p*cof* (Yuen et al., 2013). Cells were grown overnight under pilus-inducing conditions in LB pH 6.5 at 30° C with 60 rpm (*V. cholerae*) or normal LB at 37° C with 250 rpm (*E. coli*). Whole cell and supernatant fractions were collected and subsequently analyzed on SDS-PAGE and immunoblotting. Presence of TcpF and CofJ on immunoblots were assessed and quantified with polyclonal anti-TcpF and anti-CofJ antibodies, respectively.

Wild-type (WT) *V. cholerae* grown overnight exported the majority of its protein into the culture supernatant (sup), with little remaining in the whole cell fraction (wc) (Fig 3-2A). In wild-type *E. coli*, CofJ is present in the culture supernatant although much of it remains in the cell fraction (Fig 3-2B). When expressed ectopically in the mutant *V. cholerae* Δ*tcpF*, while half of TcpF was secreted, CofJ accumulated in the whole cell (Fig 3-2A). When expressed ectopically in the mutant *E. coli* Δ*cofJ*, TcpF was found mainly in the whole cell (Fig 3-2A). These observations from *V. cholerae* and ETEC are consistent with specific exoprotein recognition by their respective T4P systems.

However, I did not observe CofJ secretion in the *E. coli* Δ*cofJ* mutant. I was unable to determine the underlying cause but most likely it was due to instability resulted from co-expression of 2 different plasmids, p*cof*Δ*cofJ* and p*cofJ*. p*cof*Δ*cofJ* carries the entire ETEC *cof* operon with a deletion in the *cofJ* gene. Notably, my secretion assay for ETEC all involved double plasmid transformation into *E. coli* expression strain HB101 but complete CofJ absence in the supernatant only occurred in the presence of the *pcofJ* plasmid.

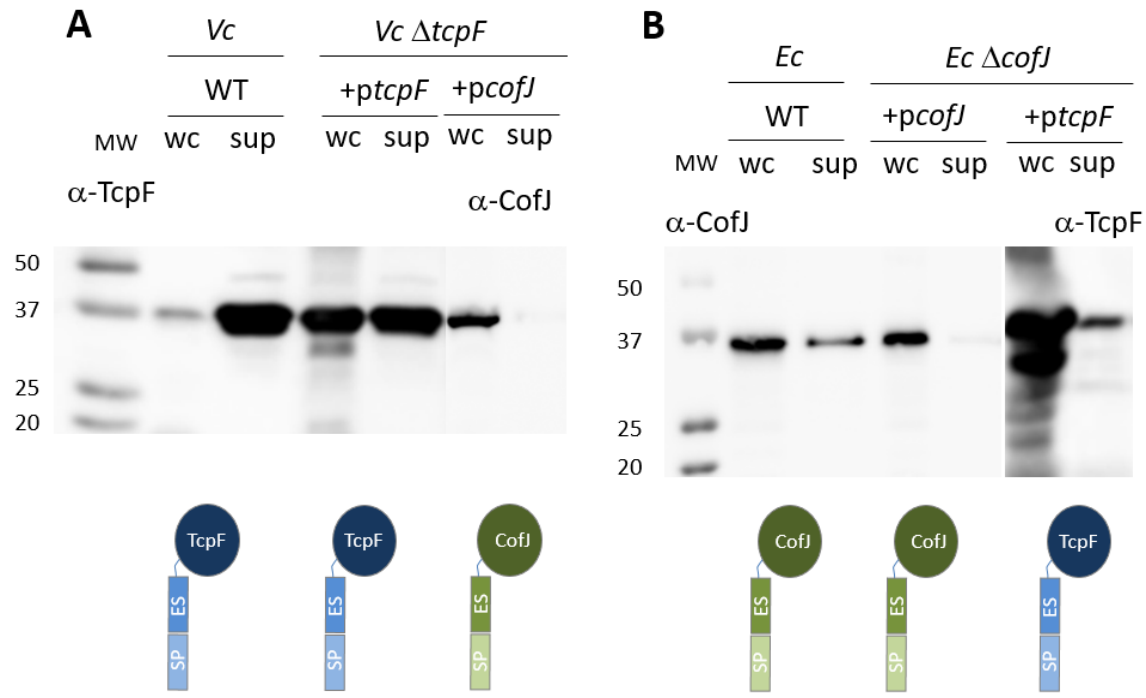

Figure 3-2 *Vibrio cholerae* cannot secrete CofJ and *Escherichia coli* cannot secrete TcpF

A. Immunoblot showing exoprotein presence in whole cell (wc), supernatant (sup) in *V. cholerae*. *V. cholerae* can secrete TcpF but not CofJ. **B**. Immunoblot showing exoprotein presence in whole cell (wc), supernatant (sup) in *E. coli*. *E. coli* can secrete CofJ but not TcpF. TcpF presence in *E. coli* supernatant is likely due to leakage from over-expression. Expression of exoproteins was induced at 0.001% rhamnose or arabinose. WT = wild-type. MW = molecular weight ladder.

3.2. The N-terminal segment of TcpF directs export in *V. cholerae*

In Gram-negative bacteria, exoprotein secretion occurs in one or two-step. Two-step secretion requires transport across the inner and outer-membrane sequentially. TcpF and CofJ carry a short segment in the N-terminus that is characteristic of the SP, indicating both are processed and cleaved by the Sec-system for periplasmic translocation. We hypothesize that after periplasmic translocation, the new N-terminus of the mature exoprotein serves as the second signal sequence, ES, which is recognized by the T4P system for export to the extracellular space.

3.2.1. An exoprotein construct lacking the export sequence accumulates in the bacterial periplasm

Following our hypothesis that the ES represents the second putative signal essential for exoprotein export across the outer-membrane barrier, I predicted that an exoprotein variant possessing its SP but lacking its ES would accumulate in the bacterial periplasmic space. I deleted the nucleotides encoding this segment in a plasmid-encoded *tcpF* gene (Fig. 3-3A) and expressed this *tcpF*-ΔES construct in a *V. cholerae* Δ*tcpF* strain. Cells were grown overnight under pilus-inducing conditions, after which desired fractions were collected and analyzed on SDS-PAGE and immunoblots. While wild-type TcpF appeared mostly in the culture supernatant, the TcpF-ΔES mutant remained mostly in the whole cell (Fig 3-2B). To determine if TcpF was stuck in the cytoplasm or accumulated in the periplasm, the whole cell fraction was treated with polymyxin B to induce membrane leakage for isolation of the periplasmic (peri.) and spheroplast (spher.) fractions. The majority of TcpF-ΔES was found in the periplasm, suggesting translocation of this exoprotein variant across the inner-membrane, but not the outer membrane. This observation suggests that ES is responsible for extracellular secretion of TcpF.

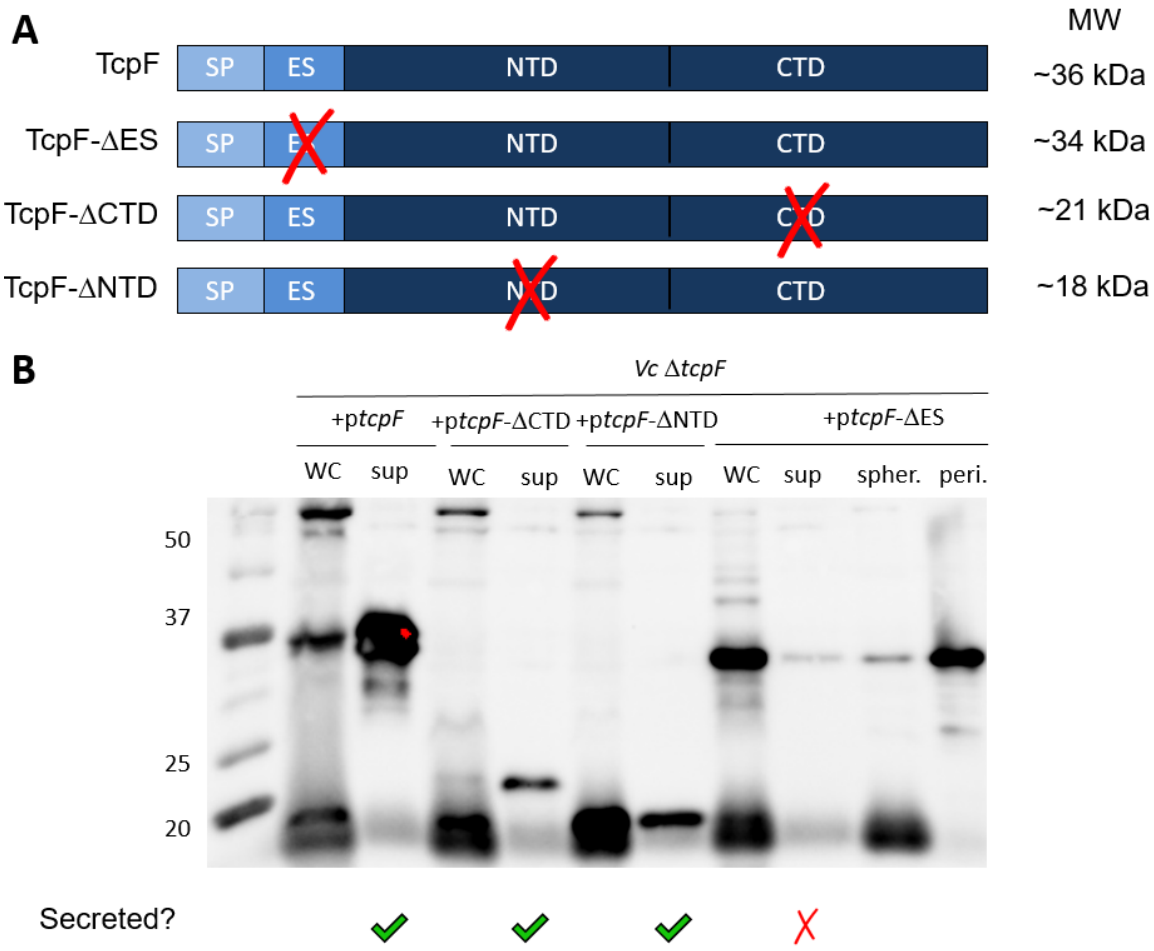

Figure 3-3 A TcpF construct lacking its export sequence accumulates in the *Vibrio cholerae* periplasm

A. Schematic of the 3 TcpF constructs tested in B with indicated molecular weight (MW). Created by Nabeel Khan, an undergraduate student in the Craig lab. **B**. Immunoblot showing exoprotein construct presence in whole cell (wc), supernatant (sup), spheroplast (spher.) and periplasm (peri.). The TcpF construct lacking its ES accumulates in the periplasm. The other 3 are secreted. Expression of all constructs were induced at 0.001% rhamnose. MW = molecular weight. NTD = N-terminal domain. CTD = C-terminal domain.

3.2.2. The TcpF N-term and C-term domains are secreted when possessing the export sequence

Having shown that the ES is necessary for TcpF secretion I next asked whether the ES alone is sufficient for this process, or whether other parts of TcpF might also be required. Thus, I generated two constructs carrying SP and ES, one possessing only the N-terminal domain (NTD) and the other with only the C-terminal domain (CTD), TcpF-ΔNTD and TcpF-ΔCTD, respectively (Fig. 3-3A). Domain deletion can impose structural risk by compromising proteins stability, leading to misfolded aggregates destined for cellular degradation. However, previous study demonstrates that the N-terminal domain

of TcpF can be expressed separately as a stable globular domain (Megli et al., 2011). This serves as our foundation for TcpF domain deletion in this analysis. The *tcpF*-ΔNTD and *tcpF*-ΔCTD constructs were expressed in the Δ*tcpF* mutant, grown overnight and the whole cell and supernatant fractions were immunoblotted with polyclonal anti-TcpF antibody raised against the intact exoprotein (Fig 3-3B). Mature TcpF-ΔNTD and TcpF-ΔCTD have molecular masses of 20.8 kDa and 17.5 kDa, respectively. Both were expressed well and found mainly in the supernatant, indicating secretion across both the inner-membrane and outer-membrane similar to wild-type TcpF. These results confirm that the ES of TcpF is sufficient for its outer-membrane export in the *V. cholerae* T4P system.

3.2.3. The export signal of TcpF and CofJ mediate heterologous secretion in *V. cholerae* and ETEC

Since the ES appears to be necessary and sufficient for TcpF export, and export is mediated by the T4P system, I wondered whether the ES is recognised by a component of the *V. cholerae* T4P machinery, and is thus specific for its native T4P system. To answer this, I tested the ability of *V. cholerae* to export CofJ into the culture supernatant, and of *E. coli* to export TcpF. As shown above, *V. cholerae* Δ*tcpF* secreted TcpF, whereas CofJ remained in the whole cell fraction (Fig 3-2A). Unfortunately for the reciprocal experiment, the positive control, ectopic expression of CofJ in *E. coli* Δ*cofJ* showed reproducibly poor secretion (Fig 3-2B). However, in this same strain, TcpF was expressed at a very high level but only a small proportion was released into the supernatant, suggesting TcpF secretion was very inefficient in *E. coli*. The presence of TcpF in the *E. coli* supernatant was likely due to membrane leakage from overexpression. These observations support the hypothesis that the *V. cholerae* and *E. coli* T4P-mediated secretion is specific for their cognate exoprotein.

Having shown that neither *V. cholerae* nor ETEC can efficiently secrete its non-native exoprotein, I hypothesized that the ES is specifically recognized by its cognate T4P system. To test this, I swapped both the SP and ES (i.e. the N-terminal segment, NTS) of TcpF with CofJ and vice versa to obtain 2 recombinant constructs: TcpF-J_{NTS} and CofJ-F_{NTS}, (Fig 3-1B). Because the Sec system is conserved across many species in the bacterial kingdom, I predicted that the SPs of TcpF and CofJ would be interchangeable and recognized by both *V. cholerae* and *E. coli* Sec systems. Thus,

each exoprotein should be translocated into the periplasm regardless of the bacterium, but only proteins bearing the cognate ES will be secreted into the extracellular space.

The recombinant exoproteins were expressed well in their heterologous systems, and were detected in both the whole cell and supernatant fractions, however the relative amounts indicated that extracellular secretion is very inefficient without the cognate ES. In *V. cholerae* Δ*tcpF*, CofJ-F_{NTS} was found mainly in the culture supernatant while TcpF-J_{NTS} accumulated mostly in the whole cell (Fig 3-4). Similarly, while CofJ-F_{NTS} accumulated mostly in the whole cell of *E. coli* Δ*cofJ*, the majority of TcpF-J_{NTS} was found in the supernatant (Fig 3-5). These data support our hypothesis that the *V. cholerae* and ETEC T4P systems specifically recognize the ES of their cognate exoprotein.

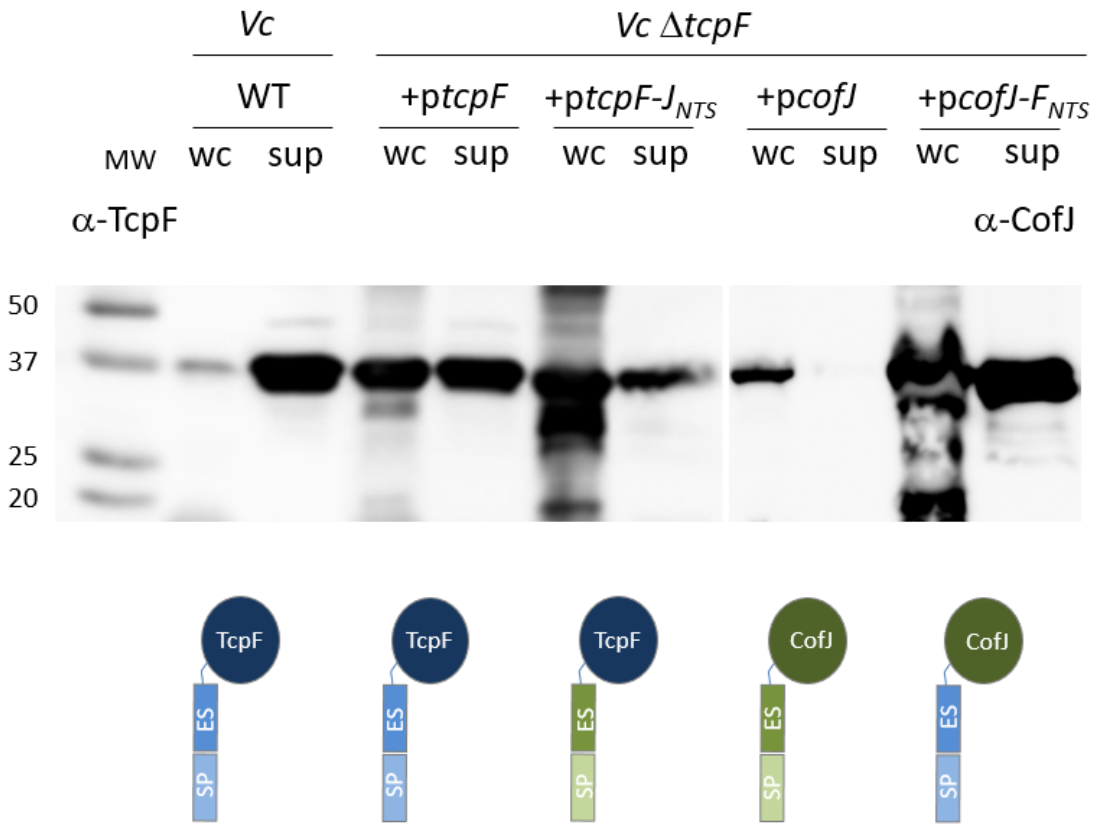

Figure 3-4 In *Vibrio cholerae*, secretion is determined by the export sequence of its wild-type exoprotein TcpF

Immunoblot showing exoprotein construct presence in whole cell (wc), supernatant (sup) in *V. cholerae*. Among the 2 wild-type and 2 recombinant exoproteins tested, *V. cholerae* can secrete TcpF and CofJ-F_{NTS}. TcpF-J_{NTS} presence in sup is likely due to membrane leakage from over-expression. Expression of all constructs were induced at 0.001% rhamnose. WT = wild-type. MW = molecular weight ladder.

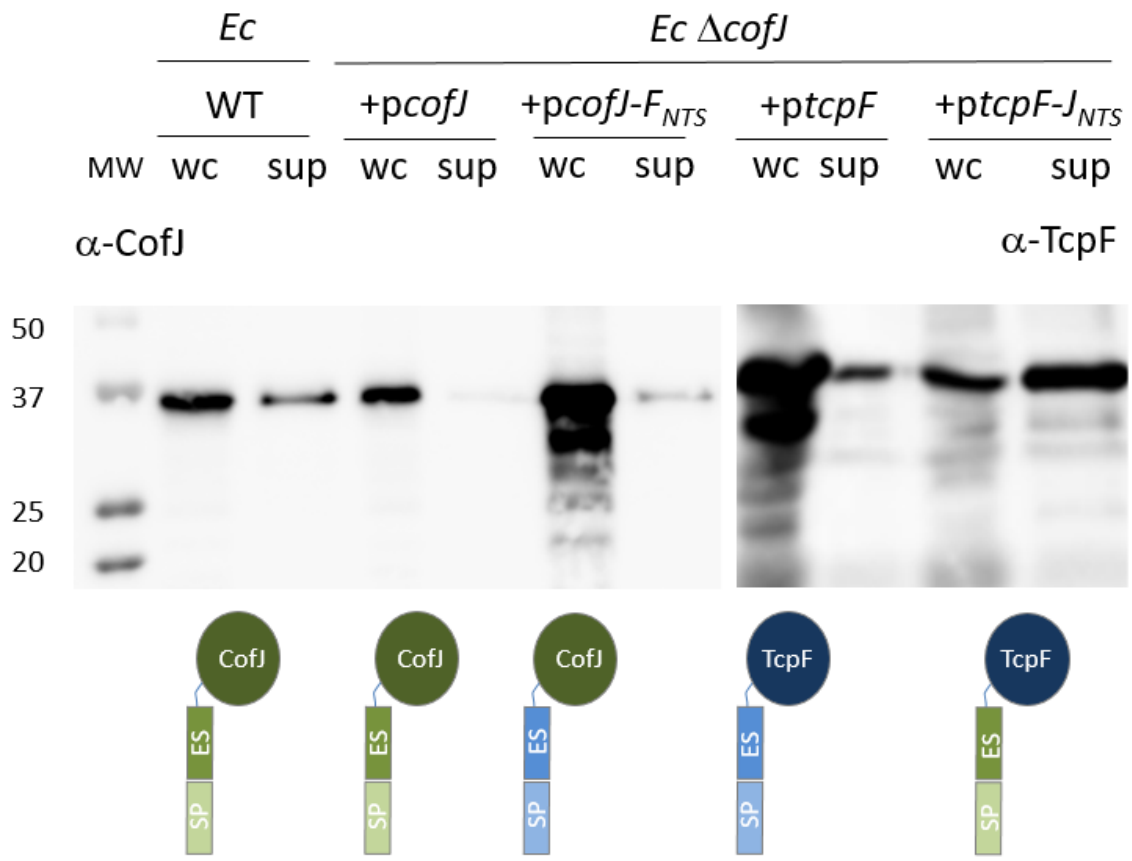

Figure 3-5 In ETEC, secretion is determined by the export sequence of its wild-type exoprotein CofJ

Immunoblot showing exoprotein construct presence in whole cell (wc), supernatant (sup) in *E. coli*. Among the 2 wild-type and 2 recombinant exoproteins tested, *E. coli* can secrete CofJ and TcpF-J_{NTS}. TcpF presence in the sup is likely due to membrane leakage from over-expression. Absence of CofJ in the sup of strain *E. coli* Δ*cofJ* is unexpected but likely due to instability from co-expression of 2 different plasmids. Expression of all constructs were induced at 0.001% rhamnose. WT = wild-type. MW = molecular weight ladder.

To quantify secretion phenotypes of tested exoprotein constructs, we repeated the aforementioned immunoblots in triplicate and plotted whole cell and supernatant fractions in our studied *V. cholerae* and *E. coli* strains (Fig 3-6). Roughly 90% of CofJ and 80% of TcpF remained in the whole cell of *V. cholerae* Δ*tcpF* and *E. coli* Δ*cofJ* respectively. Between the two recombinant exoproteins, only one of each was secreted in significant amount in each species: about 65% of CofJ-F_{NTS} was secreted in *V. cholerae* Δ*tcpF* and 60% of TcpF-J_{NTS} was secreted in *E. coli* Δ*cofJ*.

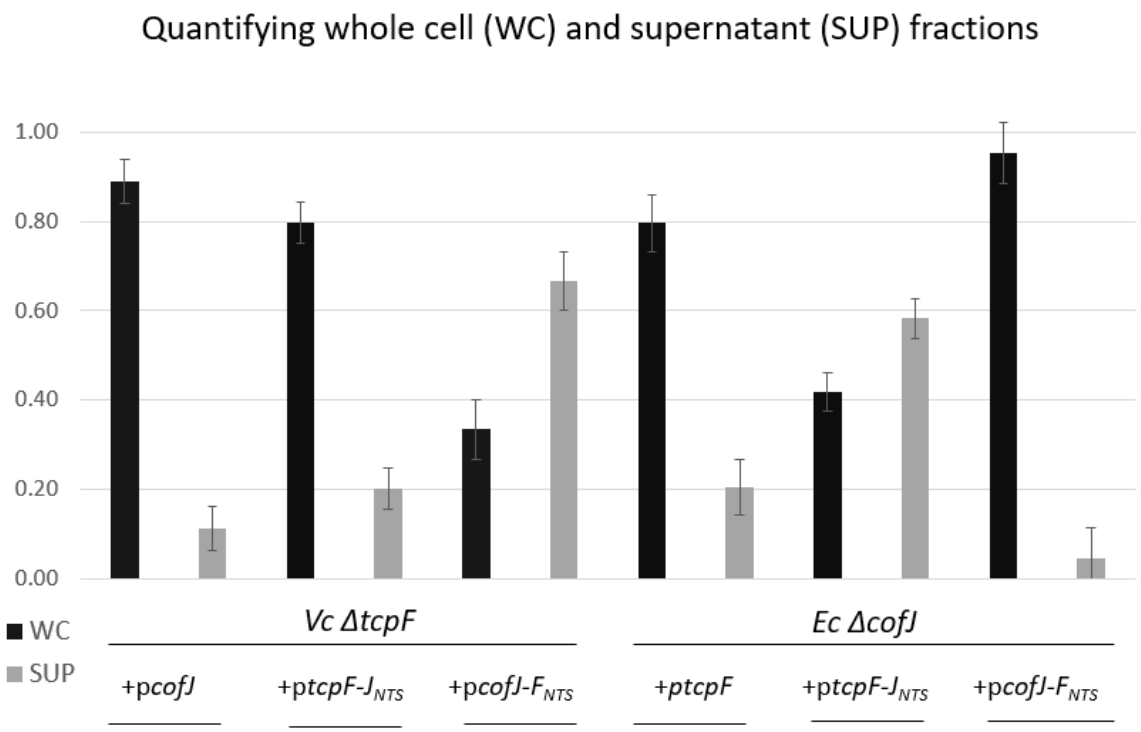

Figure 3-6 Quantifying whole cell and supernatant fractions in our studied *Vibrio cholerae* and *Escherichia coli* mutant strains

Immunoblots were performed in triplicate to collect data for quantification of whole cell and supernatant fractions. 90% of CofJ and 80% of TcpF remain in the whole cell of *V. cholerae* Δ*tcpF* and *E. coli* Δ*cofJ*, respectively. In *V. cholerae* Δ*tcpF*, 20% of TcpF-J_{NTS} and 65% of CofJ-F_{NTS} appear in the supernatant. In *E. coli* Δ*cofJ*, 60% of TcpF-J_{NTS} and 5% of CofJ-F_{NTS} appear in the supernatant. WC = whole cell. SUP = supernatant. Exoprotein expression were induced with either arabinose or rhamnose.

To confirm the importance of the SP for exoprotein secretion, I generated four exoprotein constructs lacking SP with and without ES: TcpF-ΔSP, TcpF-ΔNTS, CofJ-ΔSP and CofJ-ΔNTS. However, none was expressed in either *V. cholerae* or *E. coli* (data not shown). Since SP cleavage by the Sec-machinery is a precursor for protein translocation and folding in the periplasm (Green & Mecsas, 2016), I expect that these proteins remained in the cytoplasm as unfolded pre-proteins and were eventually targeted for proteolysis, explaining their lack of expression.

Two other non-exoproteins, PilX, a pilin protein in *Neisseria meningitidis*, and FbpA, an ABC transporter periplasmic protein in *Neisseria gonorrhoeae*, were also used in our secretion assays but unfortunately, they were not expressed in either *V. cholerae* or *E. coli*. The secretion phenotypes for all protein constructs tested are summarized in table 1.

Table 3-1 Summary of secretion phenotypes for all tested protein constructs

Protein constructs		*Vibrio cholerae*		*Escherichia coli*	
		Whole cell (wc)	Supernatant (sup)	Whole cell (wc)	Supernatant (sup)
Native	TcpF	+	+	+	-*
	CofJ	+	-	+	+
Deletion	TcpF-ΔES	+	+	nt**	nt
	TcpF-ΔSP	-	-	nt	nt
	TcpF-ΔNTS	-	-	nt	nt
	CofJ-ΔSP	nt	nt	-	-
	CofJ-ΔNTS	nt	nt	-	-
Swap	TcpF-J_{NTS}	+	-	+	+
	CofJ-F_{NTS}	+	+	+	-
Non-exoprotein	FbpA	-	-	-	-
	FbpA-F_{NTS}	-	-	-	-
	FbpA-J_{NTS}	-	-	-	-
	PilX	-	-	-	-
	PilX-F_{NTS}	-	-	-	-
	PilX-J_{NTS}	-	-	-	-

* no exoprotein present in either the wc or sup fractions means that it was not detected/expressed

** nt = not tested

3.3. Investigating the relationship between the exoprotein export signal and the minor pilin trimer

Our results thus far indicate that the ~23 amino-acid N-terminal segment of the *V. cholerae* and ETEC exoprotein is the export sequence recognized by the T4P machinery. Pull-down assays and crystal structure analysis revealed that the ETEC minor pilin CofB trimer binds specifically to the ES of the CofJ exoprotein (Oki et al., 2018). This observation from Oki et al. provided the foundation for my approach to test for similar interaction in *V. cholerae*, between the minor pilin TcpB and the exoprotein TcpF. While this group attributed this interaction to epithelial cell adhesion, we propose that it facilitates exoprotein secretion. Since exoprotein is destined for secretion across the outer-membrane, it must dissociate from the pilus filament once in the extracellular milieu. Thus, exoprotein binding is transient and likely weak. Furthermore, we propose that the ES of TcpF binds to TcpB which delivers it across the outer membrane. Previous crystal structure analysis reveals the ES of TcpF is flexible and prone to degradation (Megli et al., 2011; Yuen et al., 2013). We predict full-length TcpF binds TcpB but a TcpF variant with degraded ES will not exhibit the same interaction. In our pull-down analysis, we used two TcpF proteins, one full-length with a theoretical mass of 35846 Da, and the other that had undergone proteolysis to a slightly smaller form.

We analyzed the full-length and proteolyzed TcpF samples on polyacrylamide gel stained with Coomassie. Both samples ran as a doublet with marked difference in the intensity of the two individual bands. Full-length sample appears as a dominant top band whereas proteolyzed sample shows a dominant bottom band, most likely due to its slightly lower molecular weight which corresponds to a TcpF variant with degraded ES. We further subjected these two TcpF samples to Liquid Chromatography Mass Spectrometry (LCMS) for an in-depth analysis. LCMS reveals molecular weight difference between our two TcpF proteins. The full-length sample shows a dominant peak at 35843, which matches the molecular weight of the full TcpF (Fig 3-8A). This peak is absent in the proteolyzed sample, which shows a dominant peak at 34291 matching the molecular weight of a TcpF variant lacking 14 residues in the N-terminus (Fig 3-8B). We refer to this major component of the proteolyzed TcpF sample as ΔES-TcpF. Other lower abundance peaks correspond to additional N-terminal truncations. These peaks are also present in the full-length sample but at much lower abundance than the full-length peak. Further analysis of TcpF sequences corresponding to these three peaks is provided in Appendix B.

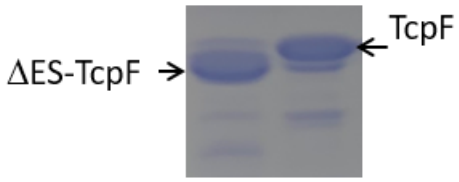

Figure 3-7 TcpF appears as a doublet on polyacrylamide gel
Even though both TcpF samples appear as a doublet on polyacrylamide gel stained with Coomassie, the two bands differ in intensity. Full-length TcpF sample shows a dominant top band. Due to proteolysis, ΔES-TcpF has a slightly lower molecular weight and shows a dominant bottom band. These two TcpF samples were analyzed further by liquid chromatography mass spectrometry.

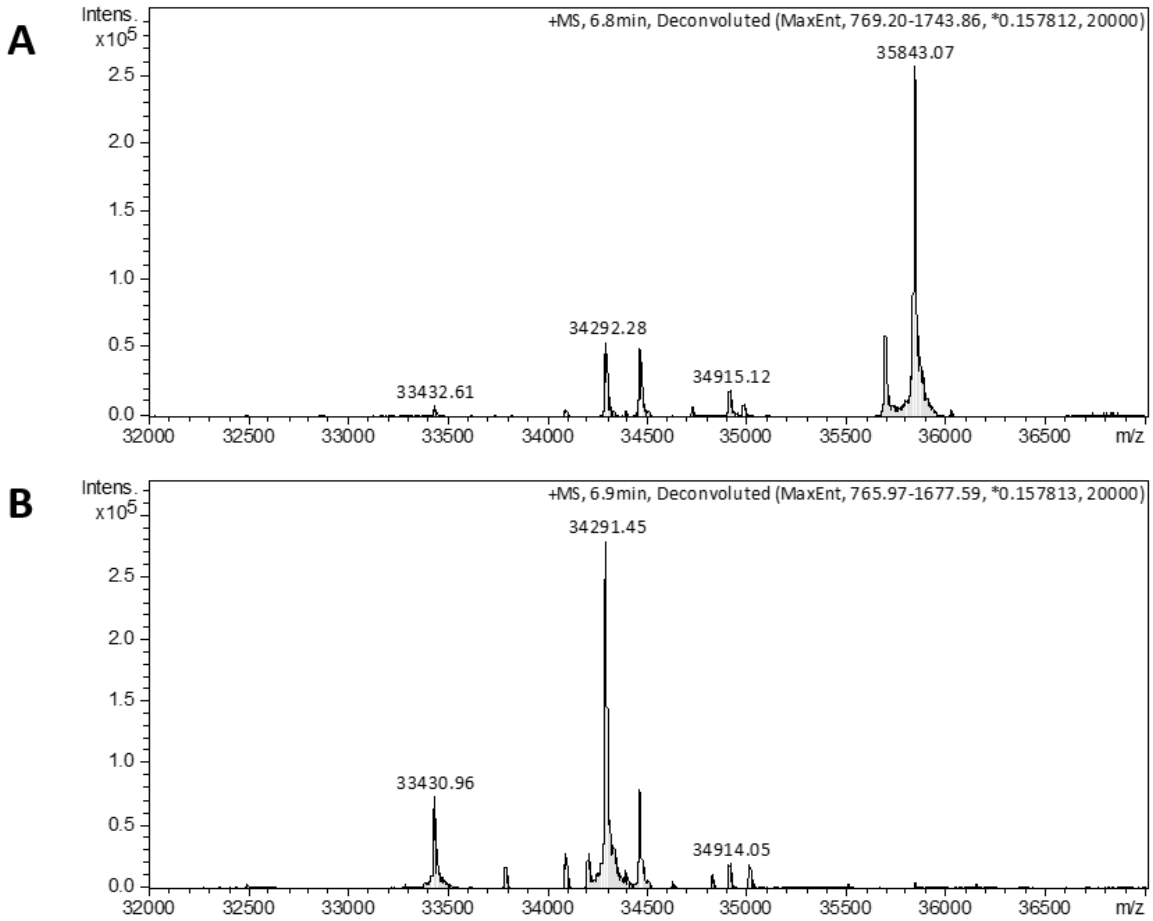

Figure 3-8 Liquid Chromatography Mass Spectrometry of two TcpF samples
A. Full-length sample shows a dominant peak at 35843 which matches the molecular weight of the full TcpF. **B**. Proteolyzed sample shows a dominant peak at 34291 which matches the molecular weight of a TcpF variant lacking 15 residues of its ES in the N-terminus. The three peaks in B (33431, 34291 and 34914) also appear in A but at significantly lower abundance.

We performed pull-down assays to test the interaction between *V. cholerae* minor pilin TcpB and its exoprotein TcpF. For the first assay (Fig 3-9A), the two target candidates were: (i) purified His-TcpB-C which is the TcpB C-terminal trimerization domain (234-423), this is the protein trimer that corresponds to the CofJ biding domain of the CofB trimer (Fig 1-18A); and (ii) non-His-tagged full-length TcpF. The protein pair of interest was incubated in a 1:1 ratio at 37° C for 1 hour and loaded onto Ni-NTA agarose beads. Beads were washed 3 times and bound proteins were eluted. Collected fractions were visualized by SDS-PAGE and Coomassie staining. The input fraction represents total protein added. The flow-through and pull-down fractions represent unbound and bound proteins, respectively. In the two control samples with one protein, either TcpF or His-TcpB-C, only His-TcpB-C was detected in the pull-down. When they

were incubated together, both His-TcpB-C and full-length TcpF appeared in the pull-down (Fig 3-9A), suggesting that His-TcpC binds to TcpF. For our second binding assay (Fig 3-9B), we used His-TcpB-C and non-His-tagged TcpF, which we had determined using LCMS to be ΔES-TcpF (Fig. 3-8). When these proteins were incubated together, ΔES-TcpF did not appear in the pull-down with His-TcpC, suggesting the ES of TcpF is required for TcpF-TcpB interaction.

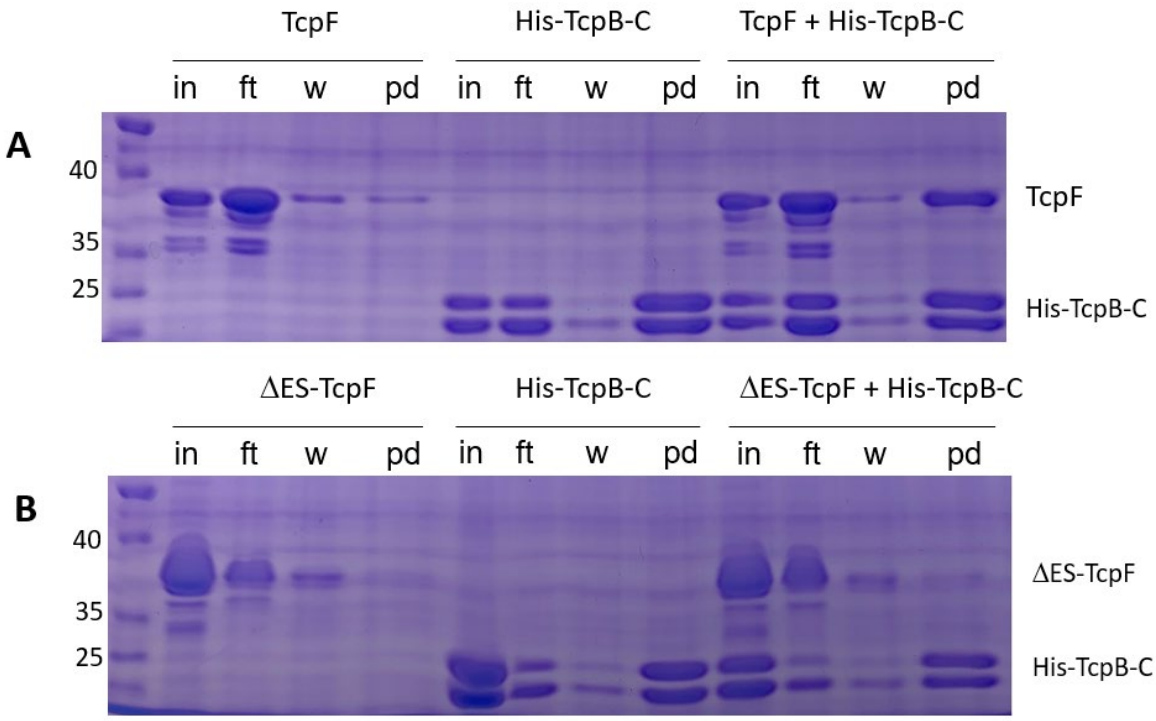

Figure 3-9 Pull-down assay between His TcpB-C and two non-His-tagged TcpF proteins, one full-length and the other with degraded export signal

His-TcpB-C is the trimerization domain of TcpB (234-423) with an N-terminal His tag. **A**. Pull-down between His TcpB-C and full-length TcpF. **B**. Pull-down between His-TcpB-C and ΔES-TcpF. Samples were incubated at 37° C for one hour and loaded onto Ni-NTA beads. Beads were washed three times and bound proteins were eluted. Collected fractions were separated by SDS-PAGE and detected by Coomassie stain. The position of each protein is indicated on the right. His TcpB-C appeared as 2 separate bands of equal intensity. In = Input. Ft = flow-through. W = wash. Pd = pull-down. Performed by Katie Danielson, an undergraduate student in the Craig lab.

3.4. Investigating the relationship between exoprotein expression and pilus assembly

Among the diverse functions regulated by the *V. cholerae* T4P is microcolony formation. In the human small intestine, pathogenic *V. cholerae* aggregates in microcolonies via contact among their T4P. *tcpF* is the only gene in the *tcp* operon that is not necessary for pilus assembly (Kirn et al., 2003) (Megli & Taylor, 2013). However,

TcpF is necessary for colonization of the infant mouse and hence, was referred to as a colonization factor (Kirn et al., 2003). Kirn et al. proposed that that the functions of TcpF and the pilus are separated but interrelated. We wonder if TcpF secretion might compete with pilus assembly since both TcpF and the pilus are extruded through the same secretin channel (Fig 1-20). To test this hypothesis, *V. cholerae* Δ*tcpF* cells were grown overnight under pilus-inducing conditions and exoprotein expression was induced at different levels. Whole cell (wc) and supernatant (sup) fractions were collected, run on SDS-PAGE and analyzed on immunoblots with anti-TcpF antibody to assess TcpF expression and secretion, and anti-TcpA antibody directed at the major pilin. The level of *V. cholerae* TcpA in the supernatant reflects pilus assembly (Ng et al., 2016).

For the purpose of controlled TcpF expression, p*tcpF* has an inducible *araC* promoter upstream of the *tcpF* gene, allowing intracellular TcpF production to be induced with arabinose. As expected, our immunoblot showed no detectable TcpF in the absence of the inducer, and increasing amounts of TcpF in both the whole cell and supernatant fractions with increasing arabinose concentrations. In line with the result from Megli et al., *V. cholerae* Δ*tcpF* showed TcpA levels in the supernatant comparable to that of wild-type strain (Fig 3-10), consistent with TcpF not being necessary for pilus assembly. In samples where TcpF expression was induced with arabinose, TcpA levels in the supernatant were also comparable to the wild-type strain regardless of TcpF expression levels, suggesting that overexpression of TcpF does not impact pilus assembly. Interestingly, unlike for wild type *V. cholerae*, TcpF was not secreted as efficiently when expressed ectopically, with only a portion of the total exoprotein being secreted at the higher expression levels, which are nonetheless lower than wild type TcpF levels (Fig 3-10).

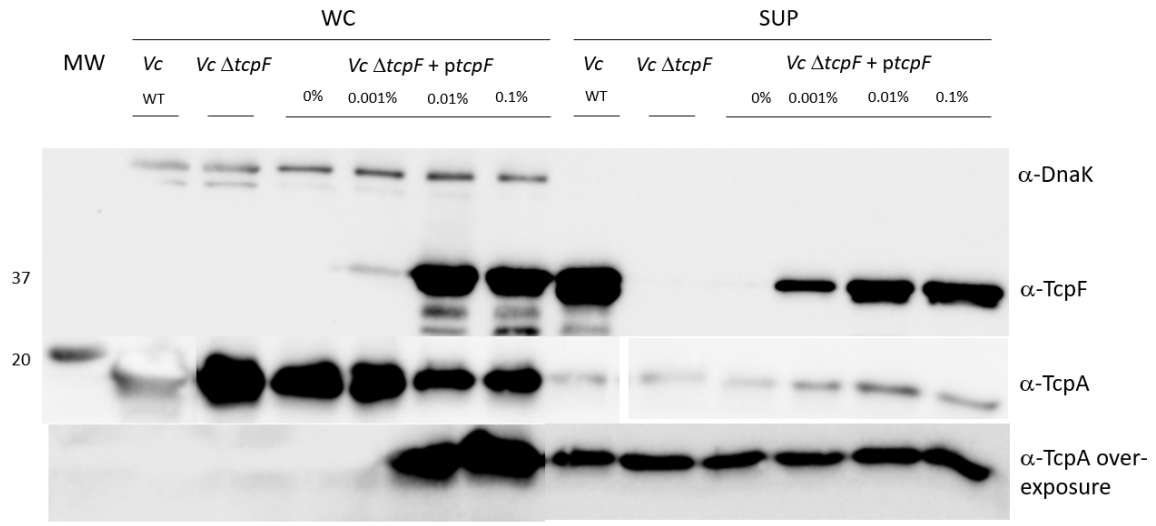

Figure 3-10 Pilus assembly in *Vibrio cholerae* as reflected by immunoblotting against the major pilin subunit TcpA in the culture supernatant

Immunoblot showing TcpF and TcpA presence in whole cell (wc), supernatant (sup) in *V. cholerae*. TcpF expression was induced in the mutant *V. cholerae* Δ*tcpF* with arabinose at the indicated levels. Different TcpF expression did not affect TcpA level in the sup. WT = wild-type. MW = molecular weight ladder.

Chapter 4. Discussion and Future Studies

The T4P systems demonstrate diverse capabilities across the bacterial kingdom, aiding both non-pathogenic and pathogenic species alike in their competition for survival. Evolutionarily related, the T2S systems and the T4P systems share many common functional and structural features. In this thesis, I studied the 2 simple T4P systems of *V. cholerae* and ETEC with an emphasis on exoprotein secretion. The exoproteins TcpF and CofJ are secreted by the *V. cholerae* and ETEC, respectively, as colonization factors aiding bacterial pathogenesis. The underlying mechanism by which these exoproteins promote cellular microcolony formation remains ambiguous. The crystal structures of both TcpF and CofJ have been solved but have provided few clues to their functional roles as colonization factors. While TcpF shares no homology to any known protein, CofJ bears some structural similarity to pore-forming toxins. The secretion mechanism by the respective *V. cholerae* TCP and the ETEC CFA/III systems is another area of uncertainty. Understanding secretion in these simple T4P systems will provide insights into secretion in the complex T4P systems as well as the T2S systems, and may suggest new antimicrobial strategies.

V. cholerae and ETEC are Gram-negative enteric pathogens and the causative agents for diarrhea. Each possesses a simple T4P system capable of secreting a particular exoprotein. I've shown that T4P-mediated secretion in *V. cholerae* and ETEC is specific. Particularly, *V. cholerae* only secretes the exoprotein native to its TCP system, TcpF. Similarly, ETEC CFA/III system only secretes CofJ. Despite many overlapping features between their simple T4P systems, this information suggests the presence of a unique feature in both exoproteins that is recognized only in one system. This observation provides a foundation for our work in elucidating their secretion mechanism.

I sought to test the role of the exoprotein N-terminus as the export sequence (ES) in the T4P-mediated secretion. In Gram-negative bacteria, transport across 2 barriers, the inner-membrane and outer-membrane, is regulated tightly. Periplasmic protein is typically synthesized in the cytoplasmic space as an unfolded precursor carrying a signal peptide (SP). Cleavage of the SP by the Sec-machinery is a means by which protein is translocated across the inner-membrane. Export across the outer-

membrane by the T2S system or the T4P systems might require a second signal, which we attribute to the ES of exoprotein. Together, the SP and ES allow exoprotein transport across the inner-membrane and outer-membrane sequentially. While SP cleavage by the Sec-machinery is well-established, exoprotein recognition by the T2S systems or the T4P systems is poorly understood. I established that in *V. cholerae*, an exoprotein mutant lacking its ES remains in the periplasm, consistent with it being delivered there by the Sec-machinery. Furthermore, I showed that two exoprotein variants, lacking a significant portion of their structures but with intact SP and ES, were secreted in *V. cholerae*, suggesting the requirement for the SP and ES, and only these, in T4P-mediated secretion. By swapping the SP and ES between the two *V. cholerae* and ETEC exoproteins, I demonstrated that unlike SP, the ES is only recognized by its native T4P system. Two other non-exoproteins, PilX, a pilin protein in *Neisseria meningitidis*, and FbpA, an ABC transporter periplasmic protein in *Neisseria gonorrhoeae*, were also used in our secretion assays but unfortunately, they were not expressed in either *V. cholerae* or *E. coli*. Important future work might involve more exoprotein constructs to underline the function of SP and ES across other T2S system and T4P systems. Additionally, instead of replacing both SP and ES, swapping ES alone should be sufficient to achieve heterologous secretion, placing a clearer emphasis on the outer-membrane export signal.

In *V. cholerae*, exoprotein secretion and pilus assembly have been demonstrated as precursors for *in vivo* colonization in the infant mouse intestine (Kirn et al., 2003) yet the molecular mechanism is unclear. Recently, Oki et al. has indicated binding between the exoprotein CofJ and the minor pilin CofB in the ETEC CFA/III system. Furthermore, the solved crystal structure of the complex indicates recognition of CofJ N-terminal first 24 residues by the CofB trimer. Here we show that a similar interaction exists between *V. cholerae* exoprotein TcpF and its minor pilin TcpB, and that the ES of TcpF is required for this interaction. Going forwards, co-crystallization of the complex between TcpF and TcpB could provide more detailed insight into the nature of this binding, elucidating how release might occur once the exoprotein reaches the extracellular space. We further propose that association with the minor pilin places the exoprotein at the pilus tip where it is secreted as the pilus filament extends across the outer-membrane secretin channel. Once the pilus reaches the extracellular space, the exoprotein dissociates from the minor pilin, perhaps by proteolysis or simply due to the

interaction being transient, and the pilus retracts back into the cell to extrude another TcpF molecule.

I confirmed that exoprotein is not required for pilus assembly. A *V. cholerae* mutant defective in exoprotein production showed presence of the major pilin in the supernatant level comparable to a wild-type strain. Additionally, I showed that increased exoprotein expression has no effect on pilus assembly. This result was surprising given that TcpF is required for *V. cholerae* colonization in the infant mouse intestine (Kirn et al., 2003). Because a *V. cholerae* mutant lacking exoprotein expression is able to form microcolonies *in vitro* as assessed by pilus assembly, it is unclear what causes its *in vivo* colonization defect as demonstrated by Kirn et al.

In conclusion, our results have provided a framework for the presence of a signal sequence that regulates export across the outer-membrane barrier in the simple T4P systems of *V. cholerae* and ETEC. It is likely that exoprotein secretion in the T4P systems of two other pathogens, *Francisella tularensis* and *Dichelobacter nodosus* follows a similar mechanism, but whether this extends to the closely related T2S systems remains to be seen. A prevalent feature utilized across the bacterial kingdom, the T4P systems perform diverse functions aiding non-pathogenic and pathogenic species alike. Identification of a signal sequence for substrate secretion can provide valuable information for our understanding of bacterial activity and development of microbial treatments. Exploiting the signal sequence to create a synthetic compound that binds irreversibly to the minor pilin at the pilus tip could prove a beneficial tool to block substrate secretion and hinder bacterial pathogenesis.

References

Backert, S., & Meyer, T. F. (2006). Type IV secretion systems and their effectors in bacterial pathogenesis. *Curr Opin Microbiol, 9*(2), 207-217. doi:10.1016/j.mib.2006.02.008

Berry, J. L., Phelan, M. M., Collins, R. F., Adomavicius, T., Tonjum, T., Frye, S. A., . . . Derrick, J. P. (2012). Structure and assembly of a trans-periplasmic channel for type IV pili in *Neisseria meningitidis*. *PLoS Pathog, 8*(9), e1002923. doi:10.1371/journal.ppat.1002923

Bradley, D. E. (1972). Evidence for the retraction of *Pseudomonas aeruginosa* RNA phage pili. *Biochem Biophys Res Commun, 47*(1), 142-149. doi:10.1016/s0006-291x(72)80021-4

Buttner, D. (2012). Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. *Microbiol Mol Biol Rev, 76*(2), 262-310. doi:10.1128/MMBR.05017-11

Cassel, D., & Pfeuffer, T. (1978). Mechanism of cholera toxin action: covalent modification of the guanyl nucleotide-binding protein of the adenylate cyclase system. *Proc Natl Acad Sci U S A, 75*(6), 2669-2673. doi:10.1073/pnas.75.6.2669

Chen, A., & Seifert, H. S. (2011). Interactions with Host Cells Causes *Neisseria meningitidis* Pili to Become Unglued. *Front Microbiol, 2*, 66. doi:10.3389/fmicb.2011.00066

Craig, L., Forest, K. T., & Maier, B. (2019). Type IV pili: dynamics, biophysics and functional consequences. *Nat Rev Microbiol, 17*(7), 429-440. doi:10.1038/s41579-019-0195-4

Craig, L., & Li, J. (2008). Type IV pili: paradoxes in form and function. *Curr Opin Struct Biol, 18*(2), 267-277. doi:10.1016/j.sbi.2007.12.009

Dick, M. H., Guillerm, M., Moussy, F., & Chaignat, C. L. (2012). Review of two decades of cholera diagnostics--how far have we really come? *PLoS Negl Trop Dis, 6*(10), e1845. doi:10.1371/journal.pntd.0001845

DiRita, V. J., Parsot, C., Jander, G., & Mekalanos, J. J. (1991). Regulatory cascade controls virulence in *Vibrio cholerae*. *Proc Natl Acad Sci U S A, 88*(12), 5403-5407. doi:10.1073/pnas.88.12.5403

Douzi, B., Ball, G., Cambillau, C., Tegoni, M., & Voulhoux, R. (2011). Deciphering the Xcp *Pseudomonas aeruginosa* type II secretion machinery through multiple interactions with substrates. *J Biol Chem, 286*(47), 40792-40801. doi:10.1074/jbc.M111.294843

Ellison, C. K., Kan, J., Dillard, R. S., Kysela, D. T., Ducret, A., Berne, C., . . . Brun, Y. V. (2017). Obstruction of pilus retraction stimulates bacterial surface sensing. *Science, 358*(6362), 535-538. doi:10.1126/science.aan5706

Finkelstein, R. A. (1996). Cholera, *Vibrio cholerae* O1 and O139, and Other Pathogenic Vibrios. In th & S. Baron (Eds.), *Medical Microbiology*. Galveston (TX).

Freudl, R. (2013). Leaving home ain't easy: protein export systems in Gram-positive bacteria. *Res Microbiol, 164*(6), 664-674. doi:10.1016/j.resmic.2013.03.014

Giltner, C. L., Nguyen, Y., & Burrows, L. L. (2012). Type IV pilin proteins: versatile molecular modules. *Microbiol Mol Biol Rev, 76*(4), 740-772. doi:10.1128/MMBR.00035-12

Giron, J. A., Levine, M. M., & Kaper, J. B. (1994). Longus: a long pilus ultrastructure produced by human enterotoxigenic *Escherichia coli*. *Mol Microbiol, 12*(1), 71-82. doi:10.1111/j.1365-2958.1994.tb00996.x

Green, E. R., & Mecsas, J. (2016). Bacterial Secretion Systems: An Overview. *Microbiol Spectr, 4*(1). doi:10.1128/microbiolspec.VMBF-0012-2015

Gutierrez-Rodarte, M., Kolappan, S., Burrell, B. A., & Craig, L. (2019). The *Vibrio cholerae* minor pilin TcpB mediates uptake of the cholera toxin phage CTXphi. *J Biol Chem, 294*(43), 15698-15710. doi:10.1074/jbc.RA119.009980

Hager, A. J., Bolton, D. L., Pelletier, M. R., Brittnacher, M. J., Gallagher, L. A., Kaul, R., . . . Guina, T. (2006). Type IV pili-mediated secretion modulates *Francisella* virulence. *Mol Microbiol, 62*(1), 227-237. doi:10.1111/j.1365-2958.2006.05365.x

Han, X., Kennan, R. M., Parker, D., Davies, J. K., & Rood, J. I. (2007). Type IV fimbrial biogenesis is required for protease secretion and natural transformation in *Dichelobacter nodosus*. *J Bacteriol, 189*(14), 5022-5033. doi:10.1128/JB.00138-07

Karuppiah, V., Thistlethwaite, A., & Derrick, J. P. (2016). Structures of type IV pilins from *Thermus thermophilus* demonstrate similarities with type II secretion system pseudopilins. *J Struct Biol, 196*(3), 375-384. doi:10.1016/j.jsb.2016.08.006

Kawahara, K., Oki, H., Fukakusa, S., Yoshida, T., Imai, T., Maruno, T., . . . Nakamura, S. (2016). Homo-trimeric Structure of the Type IVb Minor Pilin CofB Suggests Mechanism of CFA/III Pilus Assembly in Human Enterotoxigenic *Escherichia coli*. *J Mol Biol, 428*(6), 1209-1226. doi:10.1016/j.jmb.2016.02.003

Kirn, T. J., Bose, N., & Taylor, R. K. (2003). Secretion of a soluble colonization factor by the TCP type 4 pilus biogenesis pathway in *Vibrio cholerae*. *Mol Microbiol, 49*(1), 81-92. doi:10.1046/j.1365-2958.2003.03546.x

Kirn, T. J., Jude, B. A., & Taylor, R. K. (2005). A colonization factor links *Vibrio cholerae* environmental survival and human infection. *Nature, 438*(7069), 863-866. doi:10.1038/nature04249

Kirn, T. J., & Taylor, R. K. (2005). TcpF is a soluble colonization factor and protective antigen secreted by El Tor and classical O1 and O139 *Vibrio cholerae* serogroups. *Infect Immun, 73*(8), 4461-4470. doi:10.1128/IAI.73.8.4461-4470.2005

Kolappan, S., Coureuil, M., Yu, X., Nassif, X., Egelman, E. H., & Craig, L. (2016). Structure of the *Neisseria meningitidis* Type IV pilus. *Nat Commun, 7*, 13015. doi:10.1038/ncomms13015

Kolappan, S., Ng, D., Yang, G., Harn, T., & Craig, L. (2015). Crystal Structure of the Minor Pilin CofB, the Initiator of CFA/III Pilus Assembly in Enterotoxigenic *Escherichia coli*. *J Biol Chem, 290*(43), 25805-25818. doi:10.1074/jbc.M115.676106

Korotkov, K. V., & Hol, W. G. (2008). Structure of the GspK-GspI-GspJ complex from the enterotoxigenic *Escherichia coli* type 2 secretion system. *Nat Struct Mol Biol, 15*(5), 462-468. doi:10.1038/nsmb.1426

Korotkov, K. V., & Sandkvist, M. (2019). Architecture, Function, and Substrates of the Type II Secretion System. *EcoSal Plus, 8*(2). doi:10.1128/ecosalplus.ESP-0034-2018

Korotkov, K. V., Sandkvist, M., & Hol, W. G. (2012). The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat Rev Microbiol, 10*(5), 336-351. doi:10.1038/nrmicro2762

Krebs, S. J., Kirn, T. J., & Taylor, R. K. (2009). Genetic mapping of secretion and functional determinants of the *Vibrio cholerae* TcpF colonization factor. *J Bacteriol, 191*(11), 3665-3676. doi:10.1128/JB.01724-08

Li, J., Egelman, E. H., & Craig, L. (2012). Structure of the *Vibrio cholerae* Type IVb Pilus and stability comparison with the *Neisseria gonorrhoeae* type IVa pilus. *J Mol Biol, 418*(1-2), 47-64. doi:10.1016/j.jmb.2012.02.017

Maier, B., Potter, L., So, M., Long, C. D., Seifert, H. S., & Sheetz, M. P. (2002). Single pilus motor forces exceed 100 pN. *Proc Natl Acad Sci U S A, 99*(25), 16012-16017. doi:10.1073/pnas.242523299

Mancl, J. M., Black, W. P., Robinson, H., Yang, Z., & Schubot, F. D. (2016). Crystal Structure of a Type IV Pilus Assembly ATPase: Insights into the Molecular Mechanism of PilB from *Thermus thermophilus*. *Structure, 24*(11), 1886-1897. doi:10.1016/j.str.2016.08.010

McCallum, M., Tammam, S., Khan, A., Burrows, L. L., & Howell, P. L. (2017). The molecular mechanism of the type IVa pilus motors. *Nat Commun, 8*, 15091. doi:10.1038/ncomms15091

Megli, C. J., & Taylor, R. K. (2013). Secretion of TcpF by the *Vibrio cholerae* toxin-coregulated pilus biogenesis apparatus requires an N-terminal determinant. *J Bacteriol, 195*(12), 2718-2727. doi:10.1128/JB.01122-12

Megli, C. J., Yuen, A. S., Kolappan, S., Richardson, M. R., Dharmasena, M. N., Krebs, S. J., . . . Craig, L. (2011). Crystal structure of the *Vibrio cholerae* colonization factor TcpF and identification of a functional immunogenic site. *J Mol Biol, 409*(2), 146-158. doi:10.1016/j.jmb.2011.03.027

Mougous, J. D., Cuff, M. E., Raunser, S., Shen, A., Zhou, M., Gifford, C. A., . . . Mekalanos, J. J. (2006). A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science, 312*(5779), 1526-1530. doi:10.1126/science.1128393

Natale, P., Bruser, T., & Driessen, A. J. (2008). Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. *Biochim Biophys Acta, 1778*(9), 1735-1756. doi:10.1016/j.bbamem.2007.07.015

Nelson, E. J., Harris, J. B., Morris, J. G., Jr., Calderwood, S. B., & Camilli, A. (2009). Cholera transmission: the host, pathogen and bacteriophage dynamic. *Nat Rev Microbiol, 7*(10), 693-702. doi:10.1038/nrmicro2204

Ng, D., Harn, T., Altindal, T., Kolappan, S., Marles, J. M., Lala, R., . . . Craig, L. (2016). The *Vibrio cholerae* Minor Pilin TcpB Initiates Assembly and Retraction of the Toxin-Coregulated Pilus. *PLoS Pathog, 12*(12), e1006109. doi:10.1371/journal.ppat.1006109

Oki, H., Kawahara, K., Maruno, T., Imai, T., Muroga, Y., Fukakusa, S., . . . Nakamura, S. (2018). Interplay of a secreted protein with type IVb pilus for efficient enterotoxigenic *Escherichia coli* colonization. *Proc Natl Acad Sci U S A, 115*(28), 7422-7427. doi:10.1073/pnas.1805671115

Paetzel, M. (2019). Bacterial Signal Peptidases. *Subcell Biochem, 92*, 187-219. doi:10.1007/978-3-030-18768-2_7

Punsalang, A. P., Jr., & Sawyer, W. D. (1973). Role of pili in the virulence of *Neisseria gonorrhoeae*. *Infect Immun, 8*(2), 255-263. doi:10.1128/IAI.8.2.255-263.1973

Reichow, S. L., Korotkov, K. V., Gonen, M., Sun, J., Delarosa, J. R., Hol, W. G., & Gonen, T. (2011). The binding of cholera toxin to the periplasmic vestibule of the type II secretion channel. *Channels (Austin), 5*(3), 215-218. doi:10.4161/chan.5.3.15268

Reichow, S. L., Korotkov, K. V., Hol, W. G., & Gonen, T. (2010). Structure of the cholera toxin secretion channel in its closed state. *Nat Struct Mol Biol, 17*(10), 1226-1232. doi:10.1038/nsmb.1910

Sandkvist, M. (2001). Type II secretion and pathogenesis. *Infect Immun, 69*(6), 3523-3535. doi:10.1128/IAI.69.6.3523-3535.2001

Sandkvist, M., Michel, L. O., Hough, L. P., Morales, V. M., Bagdasarian, M., Koomey, M., . . . Bagdasarian, M. (1997). General secretion pathway (eps) genes required for toxin secretion and outer membrane biogenesis in *Vibrio cholerae*. *J Bacteriol, 179*(22), 6994-7003. doi:10.1128/jb.179.22.6994-7003.1997

Shi, L., Deng, S., Marshall, M. J., Wang, Z., Kennedy, D. W., Dohnalkova, A. C., . . . Fredrickson, J. K. (2008). Direct involvement of type II secretion system in extracellular translocation of *Shewanella oneidensis* outer membrane cytochromes MtrC and OmcA. *J Bacteriol, 190*(15), 5512-5516. doi:10.1128/JB.00514-08

Silhavy, T. J., Kahne, D., & Walker, S. (2010). The bacterial cell envelope. *Cold Spring Harb Perspect Biol, 2*(5), a000414. doi:10.1101/cshperspect.a000414

Tauschek, M., Gorrell, R. J., Strugnell, R. A., & Robins-Browne, R. M. (2002). Identification of a protein secretory pathway for the secretion of heat-labile enterotoxin by an enterotoxigenic strain of *Escherichia coli*. *Proc Natl Acad Sci U S A, 99*(10), 7066-7071. doi:10.1073/pnas.092152899

Taylor, R. K., Miller, V. L., Furlong, D. B., & Mekalanos, J. J. (1986). Identification of a pilus colonization factor that is coordinately regulated with cholera toxin. *Ann Sclavo Collana Monogr, 3*(1-2), 51-61. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/2892514

Taylor, R. K., Miller, V. L., Furlong, D. B., & Mekalanos, J. J. (1987). Use of phoA gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc Natl Acad Sci U S A, 84*(9), 2833-2837. doi:10.1073/pnas.84.9.2833

Toth, I. K., & Birch, P. R. (2005). Rotting softly and stealthily. *Curr Opin Plant Biol, 8*(4), 424-429. doi:10.1016/j.pbi.2005.04.001

Tweten, R. K. (2005). Cholesterol-dependent cytolysins, a family of versatile pore-forming toxins. *Infect Immun, 73*(10), 6199-6209. doi:10.1128/IAI.73.10.6199-6209.2005

Waldor, M. K., & Mekalanos, J. J. (1996). Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science, 272*(5270), 1910-1914. doi:10.1126/science.272.5270.1910

Wang, F., Baquero, D. P., Su, Z., Beltran, L. C., Prangishvili, D., Krupovic, M., & Egelman, E. H. (2020). The structures of two archaeal type IV pili illuminate evolutionary relationships. *Nat Commun, 11*(1), 3424. doi:10.1038/s41467-020-17268-4

Wang, F., Coureuil, M., Osinski, T., Orlova, A., Altindal, T., Gesbert, G., . . . Craig, L. (2017). Cryoelectron Microscopy Reconstructions of the *Pseudomonas aeruginosa* and *Neisseria gonorrhoeae* Type IV Pili at Sub-nanometer Resolution. *Structure, 25*(9), 1423-1435 e1424. doi:10.1016/j.str.2017.07.016

Yuen, A. S., Kolappan, S., Ng, D., & Craig, L. (2013). Structure and secretion of CofJ, a putative colonization factor of enterotoxigenic *Escherichia coli*. *Mol Microbiol, 90*(4), 898-918. doi:10.1111/mmi.12407

Appendix A.

List of bacterial strains, plasmids, primers and antibodies

	Description/nucleotide sequence	Source
Strains		
ETEC 31-10	Wild-type ETEC LT, CFA/III	Honda et al. (1987)
ETEC 31-10P	CFA/III-negative plasmid less derivative of strain31-10; LT	Taniguchi et al. (2001)
***E. coli* HB101**	F-, *hsdS20*(rB-, mB-), *xyl5*, λ-, *recA13*, *galK2*, *ara14*, *supE44*, *lacY1*, *rpsL20*(strr), *leuB6*, *mtl-1*, *thi-1*	ThermoFisher
E. coli* HB101 p*cof	*E. coli* HB101 transformed with the p*cof* plasmid	Yuen et al. (2013)
E. coli* HB101 p*cof*Δ*cofJ	*E. coli* HB101 transformed with the p*cof*Δ*cofJ* plasmid	Yuen et al. (2013)
***V. cholerae* O395**	Classical O1, Ogawa, Sm^R	R. K. Taylor
V. cholerae* O395 Δ*tcpF	*V. cholerae* O395 with *tcpF* gene disrupted	R. K. Taylor
Plasmids		
pACYC184	p15A, Cm^R and Tc^R	ATCC
p*cof*	pACYC184 with *cof* operon cloned in Tc^R gene, Cm^R	Yuen et al. (2013)
p*cof*Δ*cofJ*	p*cof* plasmid with *cofJ* gene disrupted	Yuen et al. (2013)
p*JMA10.1*	pBAD22 derivative; *araC* replaced with *PrhaB* NcoI removed; *bla*, Ap^R	R. K. Taylor
p*tcpF*	pBAD22 plasmid containing the *tcpF* gene, *bla*, Ap^R	R. K. Taylor
p*cofJ*	pBAD22 plasmid containing the *cofJ* gene, *bla*, Ap^R	R. K. Taylor
p*tcpF-J*$_{NTS}$	p*tcpF* with the gene encoding the NTS of TcpF replaced by that of CofJ	This study
p*cofJ-F*$_{NTS}$	p*cofJ* with the gene encoding the NTS of CofJ replaced by that of TcpF	This study
p*tcpF*-ΔES	p*tcpF* plasmid with the gene encoding ES deleted	This study
p*tcpF*-ΔNTD	p*tcpF* plasmid with the gene encoding NTD deleted	This study
p*tcpF*-ΔCTD	p*tcpF* plasmid with the gene encoding CTD deleted	This study
Primers		
TcpF-F1-FP	CGGGGTACCATGAGATATAAAAAAACCTTAATGTTATCAATCATG	This study

TcpF-F1-RP	GCTCTAGAATCCGTAGCTTCATTAGACG	This study
TcpF-F2-FP	GCTCTAGAGGGAGTGAGCACCTTAG	This study
TcpF-F2-RP	GGAAGCTTTTATTTAAAGTTCTCTGAATATGCTTTGCTATAC	This study
CofJ-F1-FP	CGGGGTACCATGAAAACAAAACTCGGGTATAGCATTTTAGC	This study
CofJ-F1-RP	GCTCTAGAATCATCTACAGTAGAGGTCTTAGGC	This study
CofJ-F2-FP	GCTCTAGAGGTTGCCCAACTTTGGAAACACCA	This study
CofJ-F2-RP	GGAAGCTTTTAATCAAGGCCACAAGCCTTC	This study
Antibodies		
Goat anti-rabbit HRP	Goat polyclonal antibody raised against rabbit IgG, conjugated to HRP	Jackson ImmunoResearch
Rabbit anti-TcpF	Rabbit polyclonal antibody raised against TcpF	R. K. Taylor
Rabbit anti-TcpA	Rabbit polyclonal antibody raised against TcpA peptide (174-199)	R. K. Taylor
Rabbit anti-CofJ	Rabbit polyclonal antibody raised against CofJ peptide (164-185)	Pacific Immunology
Rabbit anti-HSP70 (DnaK)	Rabbit polyclonal antibody raised against DnaK peptide (1-100)	MyBioSource
Mouse anti-His	Mouse monoclonal antibody raised against His tag	Applied Biological Materials (ABM)

Appendix B.

Sequence analysis of the major peaks obtained from Liquid Chromatography Mass Spectrometry for two TcpF samples

Table B-1 Major peaks and intensity as reported by Liquid Chromatography Mass Spectrometry for the full-length and proteolyzed TcpF samples

TcpF sequence	TcpF		Proteolyzed TcpF	
	Mass (m/z)	Abundance (intensity)	Mass (m/z)	Abundance (intensity)
1-318	35843	254749		
9-318	34914	18931	34914	19235
15-318	34292	53407	34292	278278
23-318	33431	7304	33431	73200

* TcpF sequences corresponding to each major peak is indicated on the left

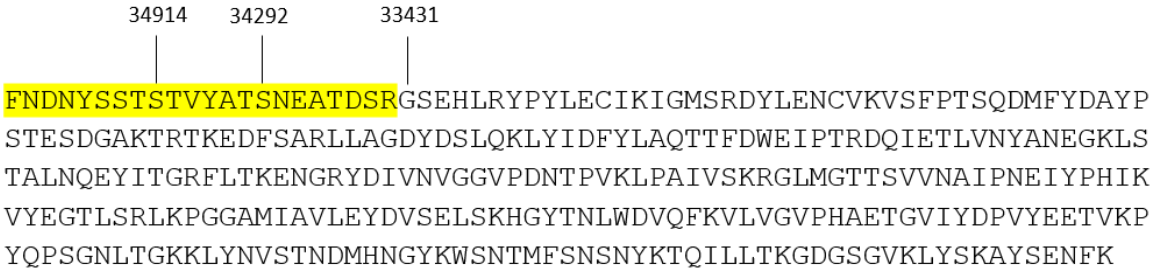

Figure B-1 Sequence of the *Vibrio cholerae* exoprotein TcpF
The putative export signal is indicated in yellow. The three peaks of our proteolyzed TcpF sample obtained from Liquid Chromatography Mass Spectrometry represent three TcpF variants lacking variable segments of the export signal. The 34914, 34292 and 33431 peaks correspond to the absence of 8, 14 and 22 residues in the N-terminus, respectively.