Structural Characterization of
*Vibrio cholerae* Toxin-Coregulated Pilus

by

Juliana Li

B.Sc., University of British Columbia, 2006

Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

in the
Department of Molecular Biology and Biochemistry
Faculty of Science

© Juliana Li 2012

SIMON FRASER UNIVERSITY
Spring 2012

All rights reserved.
However, in accordance with the Copyright Act of Canada, this work may be reproduced, without authorization, under the conditions for “Fair Dealing.” Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.
Approval

Name: Juliana Li
Degree: Doctor of Philosophy
Title of Thesis: Structural characterization of Vibrio cholerae toxin-coregulated pilus

Examining Committee:

Chair: Dr. Mark Brockman
Associate Professor
Department of Molecular Biology and Biochemistry

Lisa Craig
Senior Supervisor
Associate Professor
Department of Molecular Biology and Biochemistry

Rosemary Cornell
Supervisor
Professor
Department of Molecular Biology and Biochemistry

Michel Leroux
Supervisor
Professor
Department of Molecular Biology and Biochemistry

Edgar Young
Internal Examiner
Associate Professor
Department of Molecular Biology and Biochemistry

Mark Yeager
External Examiner
Professor, Department of Molecular Physiology and Biological Physics, University of Virginia

Date Defended/Approved: April 16, 2012
Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the “Institutional Repository” link of the SFU Library website (www.lib.sfu.ca) at http://summit.sfu.ca and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author’s written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, British Columbia, Canada

revised Fall 2011
Abstract

*Vibrio cholerae* are Gram-negative bacteria responsible for cholera, a severe and fatal gastrointestinal disease. The ability of *V. cholerae* and many other bacterial pathogens to cause disease is dependent on type IV pili. *V. cholerae* use toxin-coregulated pili (TCP) to colonize the human intestine. TCP are long, thin, flexible polymers of the TcpA subunit that self-associate to hold cells together in microcolonies, serve as the receptor for the cholera toxin bacteriophage CTXφ and secrete colonization factor proteins. To better understand TCP’s roles in pathogenesis, its structure was characterized using hydrogen/deuterium exchange mass spectrometry, computational modeling, electron microscopy (EM) and three-dimensional image reconstruction. The *V. cholerae* TcpA pilin crystal structure was docked into the pilus EM reconstruction to generate a pseudo-atomic resolution TCP structure. Tight packing of the hydrophobic N-terminal α-helices holds the pilin subunits together, but loose packing of the globular domains leaves gaps on the filament surface. These findings explain filament flexibility, suggest a molecular basis for pilus:pilus interactions and reveal a potential therapeutic target. TCP are members of the type IVb pilus subclass, which is distinguished from the type IVa subclass by differences in amino acid sequence, length and topology of the pilin globular domains. To understand the biological significance of the distinct pilin folds, circular dichroism spectroscopy was used to compare the stability of the *V. cholerae* type IVb TCP with that of the *Neisseria gonorrhoeae* type IVa gonococcal (GC) pilus together
with their pilin counterparts. We show that TcpA pilin monomers are more stable than GC pilins but surprisingly GC pili are more stable than TCP filaments. Thus, while the type IVb pilin fold appears to be more stable than the type IVa fold, differences in quaternary structures, including tighter packing and stacking of aromatic side chains appear to contribute to the extreme stability of the GC pili. The robustness of GC pili may be necessary to withstand high stress forces in the urogenital tract. This may also be a common feature of type IVa pili as an adaptation to the niches occupied by the bacteria, biological demands and functions of these filaments.

**Keywords:** Circular dichroism; electron microscopy; hydrogen/deuterium exchange mass spectrometry; *Neisseria gonorrhoeae*; type IV pili; *Vibrio cholerae*
For my family and mentors, who offered me support throughout the course of this thesis project.
Acknowledgements

Special thank you to my senior supervisor, Dr. Lisa Craig, for her relentless support and guidance and for the opportunities she has given me. I express gratitude to my supervisory committee, Dr. Rosemary Cornell and Dr. Michel Leroux, for their suggestions and encouragement throughout my graduate career. I thank Dr. Edgar Young and Dr. Mark Yeager, my internal and external examiners respectively, for their time and expertise. I acknowledge Dr. Virgil Woods, Jr. and Dr. Edward Egelman for providing me with training and assistance throughout our collaborations. I thank Dr. Ronald Taylor for insightful discussions, guidance in making tcpA mutations and comments on the deuterium-exchange mass spectrometry manuscript. I express my appreciation to Joseph Wall and Beth Lin (Brookhaven National Laboratory Scanning Transmission Electron Microscopy Facility) for STEM analysis and Tobias Fürstenhaupt (University of Calgary Microscopy and Imaging Facility) and Li Yang (Simon Fraser University Nano-imaging Facility) for technical support. I also thank Dr. Svetla Taneva for her training, guidance and insightful discussions. Finally, I thank past and present members of the Craig Laboratory.

This research would not have been possible without the financial assistance of Canadian Institutes of Health Research (CIHR), CIHR Frederick Banting and Charles Best Canada graduate scholarship – doctoral award, the Department of Molecular Biology and Biochemistry at Simon Fraser University
(teaching assistantship, graduate research fellowship, equipment and facilities) and the National Institutes of Health (NIH).
## Table of Contents

Approval ................................................................................................................ ii
Partial Copyright Licence ..................................................................................... iii
Abstract ................................................................................................................ iv
Dedication ............................................................................................................ vi
Acknowledgements ..............................................................................................vii
Table of Contents ................................................................................................. ix
List of Tables ......................................................................................................... xii
List of Figures ....................................................................................................... xiii
Glossary ..............................................................................................................xvi

1. **General Introduction** ................................................................................... 1
   1.1. *Vibrio cholerae* .............................................................................................. 1
   1.2. *V. cholerae* virulence factors .................................................................. 3
       1.2.1. Cholera toxin ...................................................................................... 4
       1.2.2. Soluble colonization factor, TcpF ....................................................... 6
       1.2.3. Toxin-coregulated pili ......................................................................... 7
   1.3. Type IV pilin and pili ...................................................................................... 9
       1.3.1. Type IVa and type IVb pilin subclasses ............................................ 10
       1.3.2. Structures of the type IV pilins .......................................................... 12
       1.3.3. Type IV pilus filament structure ....................................................... 15
       1.3.4. Type IV pilus assembly and disassembly ......................................... 23
   1.4. Research objectives ................................................................................... 28

2. **Vibrio cholerae** toxin-coregulated pilus structure analyzed by hydrogen/deuterium exchange mass spectrometry ................................................. 30
   2.1. Introduction ................................................................................................. 30
   2.2. Materials and Methods ............................................................................... 35
       2.2.1. Plasmids and bacterial strains .......................................................... 35
       2.2.2. Purification of toxin-coregulated pilus filaments ............................... 36
           2.2.2.1. Purification of TCP\(^{\text{H181A}}\) ........................................... 37
           2.2.2.2. Purification of TCP\(^{\text{WT}}\) ........................................... 37
       2.2.3. Purification of soluble, N-terminally truncated TcpA pilin ............... 38
       2.2.4. Chymotrypsin digestion .................................................................... 38
       2.2.5. Transmission electron microscopy ................................................... 39
       2.2.6. Hydrogen/deuterium exchange mass spectrometry ............................. 39
           2.2.6.1. Preparation of DXMS samples ............................................. 39
           2.2.6.2. Protein fragmentation, peptide identification and deuterium quantification ................................................................................. 40
       2.2.7. TCP model refinement ....................................................................... 42
       2.2.8. Introduction of mutations into the tcpA gene ..................................... 44
       2.2.9. Analysis of pilin expression and pilus assembly in *V. cholerae* tcpA mutants ..................................................................................... 45
2.2.10. CTXφ transduction assays ............................................................... 45
2.3. Results ........................................................................................................ 46
2.3.1. Suitability of TCP\textsuperscript{H181A} to study TCP structure ............... 46
2.3.2. Solvent accessibility of overlapping peptides in TCP\textsuperscript{H181A} filaments ................................................................. 49
2.3.3. α1 and the αβ-loop form the only subunit:subunit interfaces in the TCP filament ................................................................. 52
2.3.3.1. α1C ..................................................................................... 53
2.3.3.2. The αβ-loop ........................................................................ 53
2.3.3.3. α1N ..................................................................................... 54
2.3.4. The D-region is exposed in the TCP filament ................................... 55
2.3.5. Refinement of the TCP model based on DXMS results ................... 56
2.3.6. TcpA mutagenesis to test the refined TCP model ............................ 65
2.4. Discussion .................................................................................................. 68

3. Structure of the Vibrio cholerae type IVb toxin-coregulated pilus ......... 74
3.1. Introduction ................................................................................................. 74
3.2. Materials and Methods ............................................................................... 77
3.2.1. Bacterial strains and materials ............................................................. 77
3.2.2. TCP\textsuperscript{H181A} expression and purification ................................ 78
3.2.3. Transmission electron microscopy ...................................................... 78
3.2.4. Iterative helical real space reconstructions of the TCP filament ..... 79
3.2.5. Pseudo-atomic resolution structure of TCP ...................................... 83
3.2.6. Scanning transmission electron microscopy analysis ...................... 84
3.2.7. Generation of \textit{V. cholerae} tcpA\textsuperscript{E5A} mutant ................................ 84
3.2.8. Analysis of pilin and pilus expression and pilus function in TcpA mutants ................................................................................... 84
3.3. Results ........................................................................................................ 85
3.3.1. Negative stain reconstruction of TCP ............................................... 85
3.3.2. CryoEM reconstruction of TCP ......................................................... 87
3.3.3. Pseudo-atomic resolution structure of TCP ...................................... 89
3.3.4. Residue Glu5 is necessary for efficient TCP filament assembly ...... 95
3.4. Discussion .................................................................................................. 96

4. Stability comparison of type IVa and type IVb pilin and pili ................. 101
4.1. Introduction .................................................................................................. 101
4.2. Materials and Methods ............................................................................... 105
4.2.1. Bacterial strains and materials ............................................................. 105
4.2.2. ΔN-TcpA pilin and TCP\textsuperscript{H181A} expression and purification .... 106
4.2.3. GC pilus expression and purification .................................................... 106
4.2.4. Far-UV circular dichroism ................................................................. 107
4.2.5. Proteolytic digestion of GC pili and TCP and their pilin counterparts ................................................................................... 109
4.2.6. SDS-PAGE and immunoblot analysis .................................................. 109
4.2.7. Transmission electron microscopy ................................................. 109
4.3. Results ...................................................................................................... 110
  4.3.1. Type IVa and type IVb pilin globular domain stability comparison .......................................................... 110
  4.3.2. Type IVa and type IVb pilus thermal stability comparison .............. 113
  4.3.3. Type IVa and type IVb pilus polymerization and its effect on pilin secondary structure .......................................................... 116
  4.3.4. Type IVa and type IVb pilus chemical stability comparison .......... 119
  4.3.5. Type IVa and type IVb pilus protease susceptibility comparison ............................................................................. 122
  4.3.6. Type IVa and type IVb pilus stability in reducing agent comparison ............................................................................. 123
4.4. Discussion ................................................................................................ 125

5. General Discussion and Conclusion ...................................................... 133

References ...................................................................................................... 139
List of Tables

Table 2.1 Bacterial strains and plasmids ............................................................36
Table 3.1 Bacterial strains used in this study......................................................78
Table 4.1 Plasmids and strains used in this study ............................................106
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Transmission electron micrograph of <em>V. cholerae</em> cells.</td>
<td>2</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Infection process of <em>V. cholerae</em>.</td>
<td>4</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>Cholera toxin mechanism of action.</td>
<td>6</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>Type IV pilin amino acid sequence alignment of the N-terminal 55 residues.</td>
<td>11</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>Schematic comparison of type IVa and type IVb pilins.</td>
<td>12</td>
</tr>
<tr>
<td>Figure 1.6</td>
<td>Type IV pilin structures.</td>
<td>13</td>
</tr>
<tr>
<td>Figure 1.7</td>
<td>Secondary structure diagrams of type IVa and type IVb pilin.</td>
<td>15</td>
</tr>
<tr>
<td>Figure 1.8</td>
<td>Relationship between type IV pilins in a helically symmetrical pilus.</td>
<td>17</td>
</tr>
<tr>
<td>Figure 1.9</td>
<td>Docking of the GC pilin subunit into the TEM density map.</td>
<td>18</td>
</tr>
<tr>
<td>Figure 1.10</td>
<td>CryoEM reconstruction and pseudo-atomic resolution structure of the <em>N. gonorrhoeae</em> GC pilus.</td>
<td>20</td>
</tr>
<tr>
<td>Figure 1.11</td>
<td><em>V. cholerae</em> TcpA pilin structure and computational TCP filament model.</td>
<td>22</td>
</tr>
<tr>
<td>Figure 1.12</td>
<td>Type IV pilus assembly model.</td>
<td>25</td>
</tr>
<tr>
<td>Figure 1.13</td>
<td>Type IV pilus assembly and disassembly at the inner membrane.</td>
<td>27</td>
</tr>
<tr>
<td>Figure 1.14</td>
<td>Type IV pilus assembly, type IV pilus-mediated secretion and type II secretion.</td>
<td>28</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>DXMS analysis of His-ΔN-TcpA and TCP&lt;sup&gt;H181A&lt;/sup&gt;</td>
<td>41</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Comparison of morphology and protease susceptibility for TCP&lt;sup&gt;WT&lt;/sup&gt; and TCP&lt;sup&gt;H181A&lt;/sup&gt;.</td>
<td>48</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Overlapping peptides identified by pepsin digest and LC-MS.</td>
<td>50</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>Deuterium accessibility for peptides in the TCP&lt;sup&gt;H181A&lt;/sup&gt; filament and His-ΔN-TcpA pilin subunit.</td>
<td>51</td>
</tr>
</tbody>
</table>
Figure 2.5  Helical starts or paths of a pilus filament. ........................................ 57
Figure 2.6  Refined TCP filament model. .......................................................... 60
Figure 2.7  Amino acid sequence alignments of type IV pilin α1N and helical wheel representation of the amphipathic N-terminal segment. .......................................................... 64
Figure 2.8  Analysis of pilin expression, pilus assembly and pilus functions in TcpA mutants. .......................................................... 67
Figure 3.1  IHRSR algorithm cycle. ................................................................. 81
Figure 3.2  Negative stain TEM reconstruction and STEM analysis of TCP. ........................................................................... 87
Figure 3.3  CryoEM TCP reconstruction. .......................................................... 89
Figure 3.4  Pseudo-atomic resolution structure of the V. cholerae TCP filament. ........................................................................... 91
Figure 3.5  Comparison of the TCP filament model derived from the negative stain reconstruction with the negative stain and cryoEM reconstructions. .......................................................... 93
Figure 3.6  Effect of Glu5Ala substitution on TcpA expression and TCP assembly. .......................................................... 96
Figure 3.7  V. cholerae and N. gonorrhoeae pseudo-atomic resolution structures docked into their respective TEM densities. .......... 100
Figure 4.1  Characteristic far-UV CD spectra of poly-L-lysine in three different conformational states: α-helix, β-sheet and random coil. .......... 108
Figure 4.2  Thermal stability comparison for V. cholerae and N. gonorrhoeae type IV pilins and their respective pilus filaments. .......................................................... 112
Figure 4.3  CD spectra of GC pili after heat-treatment. ........................................ 115
Figure 4.4  Effect of βOG detergent treatment on pilin structure. .................... 119
Figure 4.5  Urea denaturation of TCP and GC pili. ........................................ 121
Figure 4.6  TEM images of urea-denatured wild type V. cholerae TCP .......... 122
Figure 4.7  Type IVa and type IVb pilin and pilus protease susceptibility comparison............................................................ 124

Figure 4.8  Tight packing of the pilin subunits and conserved α1N aromatic residues may stabilize GC pili................................. 127

Figure 4.9  Amino acid sequence alignment illustrating conserved aromatic residues in type IVa pilins. ........................................ 130
## Glossary

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>Autoagglutination</td>
</tr>
<tr>
<td>ADP</td>
<td>Adenosine diphosphate</td>
</tr>
<tr>
<td>Amp</td>
<td>Ampicillin</td>
</tr>
<tr>
<td>BFP</td>
<td>Bundle-forming pili</td>
</tr>
<tr>
<td>βOG</td>
<td>Octyl-β-D-glucopyranoside</td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>CD</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td>CFA/III</td>
<td>Colonization factor antigen III</td>
</tr>
<tr>
<td>CFU</td>
<td>Colony forming unit</td>
</tr>
<tr>
<td>CryoEM</td>
<td>Cryo-electron microscopy</td>
</tr>
<tr>
<td>CTF</td>
<td>Contrast transfer function</td>
</tr>
<tr>
<td>CTXφ</td>
<td>Cholera toxin phage</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DXMS</td>
<td>Deuterium exchange mass spectrometry</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EHEC</td>
<td>Enterohemorrhagic <em>Escherichia coli</em></td>
</tr>
<tr>
<td>EPEC</td>
<td>Enteropathogenic <em>Escherichia coli</em></td>
</tr>
<tr>
<td>ETEC</td>
<td>Enterotoxigenic <em>Escherichia coli</em></td>
</tr>
<tr>
<td>GC</td>
<td>Gonococcal</td>
</tr>
<tr>
<td>GnHCl</td>
<td>Guanidinium hydrochloride</td>
</tr>
<tr>
<td>His-tag</td>
<td>Hexahistidine tag</td>
</tr>
<tr>
<td>IMP</td>
<td>Integral membrane protein</td>
</tr>
<tr>
<td>IHRSR</td>
<td>Iterative helical real space reconstruction</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>Kn</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>LB</td>
<td>Luria broth</td>
</tr>
<tr>
<td>LC-MS</td>
<td>Liquid chromatography-mass spectrometry</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>PTA</td>
<td>Phosphotungstic acid</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>Sm</td>
<td>Streptomycin</td>
</tr>
<tr>
<td>T2SS</td>
<td>Type II secretion system</td>
</tr>
<tr>
<td>TCEP</td>
<td>Tris(2-carboxyethyl)phosphine</td>
</tr>
<tr>
<td>TCP</td>
<td>Toxin-coregulated pilus/pili</td>
</tr>
<tr>
<td>TEM</td>
<td>Transmission electron microscopy</td>
</tr>
<tr>
<td>Tₘ</td>
<td>Melting temperature</td>
</tr>
<tr>
<td>VPI</td>
<td>Vibrio pathogenicity island</td>
</tr>
</tbody>
</table>
1. General Introduction

1.1. *Vibrio cholerae*

*Vibrio cholerae* are rod-shaped, Gram-negative bacteria that colonize the small intestine causing the severe gastrointestinal disease cholera (Figure 1.1). They express a single polar flagellum on their surfaces for motility. These bacteria secrete cholera toxin, which is taken up by the epithelial cells of hosts, resulting in loss of fluids and ions and causing severe and often fatal diarrhea and dehydration within hours of onset (Slauch *et al.*, 1997). Cholera is less prevalent in industrialized countries with established sanitation systems but is endemic in developing countries of Asia and Africa including Bangladesh, Haiti, India, Somalia, and Zimbabwe, where access to clean water, proper sanitation and hygiene are lacking (Farmer *et al.*, 2011). Inadequate sanitation conditions allow the transmission of bacteria to new hosts by ingestion of contaminated food or water. Cholera remains a major health threat as limited access to medical care prevents prompt life-saving oral or intravenous rehydration therapy. The World Health Organization estimates millions of cholera cases and hundreds of thousands of deaths occur annually (http://www.who.int/mediacentre/factsheets/fs107/en/index.html), but these numbers are thought to be underreported.
Figure 1.1  Transmission electron micrograph of *V. cholerae* cells.

*Note.* Rod shaped, Gram-negative bacteria express bundled toxin-coregulated pilus filaments on their surfaces.

Pathogenic and non-pathogenic *V. cholerae* are naturally present in marine and freshwater environments where they attach to surfaces of algae, crustaceans, insects, plankton and plants (Butler and Camilli, 2005). Cholera disease is transmitted by ingestion of contaminated water and undercooked food. Water is further contaminated with high concentrations of *V. cholerae* bacteria in “rice-water” stools from infected individuals in regions where sewage is not properly treated. *V. cholerae* serogroups O1 and O139 can cause epidemics of cholera whereas hundreds of non-O1 and non-O139 serogroups only cause mild
gastroenteritis (Barua and Greenough, 1992). The O1 serogroup of *V. cholerae* is separated into classical and El Tor biotypes, which are responsible for all of the major cholera pandemics. The classical biotype was responsible for cholera outbreaks worldwide until the 1960s but was replaced with the El Tor biotype thereafter (Slauch *et al.*, 1997).

### 1.2. *V. cholerae* virulence factors

Genes encoding *V. cholerae* virulence factors are found in the *Vibrio* Pathogenicity Island (VPI), a virulence gene cluster, and the lysogenic bacteriophage CTXφ located on Chromosome 1, one of two circular chromosomes (Faruque and Mekalanos, 2003). Pathogenic strains of *V. cholerae* first acquired the VPI, which contains genes required for the production of the toxin-coregulated pilus (TCP) (Taylor *et al.*, 1987; Herrington *et al.*, 1988). TCP serves as a receptor for the filamentous cholera toxin phage φ (CTXφ). CTXφ carries the genes *ctxA* and *ctxB* encoding cholera toxin. The entire CTXφ genome can integrate into the *V. cholerae* chromosome and thus, non-pathogenic strains can become pathogenic by acquiring CTXφ as a lysogen (Waldor and Mekalanos, 1996). *V. cholerae* requires several virulence factors, including TCP and TcpF, a soluble colonization factor, to allow it to attach to and colonize epithelial cells of the small intestine, where it secretes cholera toxin (Figure 1.2). *V. cholerae* must first survive passage through the acidic environment of the stomach to infect the small intestines. In the human digestive
tract, environmental factors trigger a cascade of regulatory events that activate the expression of virulence genes (DiRita et al., 1991).

**Figure 1.2  Infection process of V. cholerae.**

*Note.* Upon ingestion, *V. cholerae* attach to intestinal epithelial cells via adhesins. Virulence factors TCP and TcpF mediate microcolony formation within the lumen of the small intestines. Cholera toxin is produced and secreted by the cells and taken up by the host resulting in the loss of fluids and electrolytes. Image courtesy of Lisa Craig, Simon Fraser University.

### 1.2.1. Cholera toxin

Upon colonization of the small intestine, *V. cholerae* expresses and secretes cholera toxin via the type II secretion system (T2SS) (Davis et al., 2000). This enterotoxin is an AB₅ toxin consisting of one active (A) subunit and five identical binding (B) subunits (Lonnroth and Holmgren, 1973). The B subunit homopentamer binds to GM1 ganglioside receptors of host intestinal epithelial
cells and the complex is taken up via endocytosis (King and Van Heyningen, 1973) (Figure 1.3). The A subunit is composed of A1 and A2 domains that are linked by a serine protease-sensitive cleavage site and a single disulfide bond (Gill and Rappaport, 1979). Upon dissociation by proteolysis and reduction, the A1 subunit catalyzes ADP-ribosylation of the alpha subunit of Gs ($G_{s\alpha}$), a guanine nucleotide binding regulatory protein. This stabilizes its GTP-bound form resulting in constitutive activation of adenylate cyclase and uncontrolled cyclic adenosine monophosphate (cAMP) production. cAMP activates protein kinase A (PKA), which phosphorylates cystic fibrosis transmembrane conductance regulator chloride ion channels causing them to open resulting in increased sodium conductance (Bubien et al., 1994; Levistre et al., 1995; Goodman and Percy, 2005). This leads to a massive efflux of fluids and electrolytes into the lumen of the small intestines resulting in severe watery diarrhea, dehydration and collapse of the circulatory system (Figure 1.3).
Figure 1.3   Cholera toxin mechanism of action.

Note. Cholera toxin A and B subunits are synthesized in the cytoplasm and transported to the periplasm by the Sec protein-translocation pathway. The cholera holotoxin is assembled in the periplasm and secreted by the type II secretion system. The B subunits facilitate binding to host GM1 receptors and the complex is internalized. The A subunit is dissociated into fragments A1 and A2. The A1 subunit catalyzes ADP-ribosylation of the alpha subunit of \( G_s \) (\( G_s^\alpha \)), resulting in constitutive activation of adenylate cyclase. Uncontrolled cAMP production and PKA activation causes cystic fibrosis transmembrane conductance regulator chloride ion channels to open. Sodium ions and water follow osmotic gradients, which leads to a massive efflux of fluids and electrolytes into the lumen of the small intestines resulting in severe watery diarrhea. Image courtesy of Lisa Craig, Simon Fraser University.

1.2.2.   Soluble colonization factor, TcpF

The \( tcp \) operon located on the VPI not only contains genes that encode for TCP biogenesis proteins but it also encodes for TcpF, a secreted colonization factor of unknown function and no known sequence homologs (Kirn et al., 2003).
TcpF is synthesized in the *V. cholerae* cytoplasm and translocated to the periplasm via the Sec secretion apparatus, which removes its signal sequence. Mature TcpF protein is folded in the periplasm and transported across the outer membrane in a TCP-dependent manner. *V. cholerae* lacking *tcpF* are able to produce functional pili but show a decrease in colonization of the infant mouse model compared to wild type (Taylor *et al.*, 1987; Kirn *et al.*, 2003). Functional TCP are necessary for TcpF secretion from the periplasm to the extracellular milieu. Both TcpA and TcpF antibodies are protective in the infant mouse model (Taylor *et al.*, 2004; Kirn and Taylor, 2005; Megli *et al.*, 2011). These studies show that TcpF is necessary for colonization and pathogenesis. The mechanism by which TCP secretes TcpF and the role TcpF plays in colonization have yet to be determined. Notably, the type IV pilus machinery responsible for the secretion of colonization factors share many similarities to the T2SS used by many bacteria to secrete various proteases, toxins and enzymes (Nunn, 1999; Sandkvist, 2001; Johnson *et al.*, 2006).

### 1.2.3. Toxin-coregulated pili

*V. cholerae* pathogenesis relies on bacterial colonization of the small intestinal epithelium and secretion of cholera toxin. A major virulence factor of *V. cholerae* is the TCP, a member of the type IVb pilus class. The genes encoding TCP assembly components are located within the *tcp* operon on the VPI of Chromosome 1. TCP are oligomers of thousands of pilin subunits encoded by the *tcpA* gene. These filamentous appendages are named “toxin-coregulated pili”
due to the coordinated regulation of cholera toxin genes (ctxA and ctxB) and the tcp gene cluster by the ToxT transcriptional activator. TCP filaments are approximately 80 Å in diameter and several microns in length. These filamentous appendages, expressed on the surfaces of V. cholerae cells, mediate microcolony formation and colonization of the small intestine (Kirn et al., 2000). TCP interact with each other forming rope-like bundles between the bacteria causing cells to agglutinate in microcolonies on the small intestinal epithelial cells, which localizes the secreted cholera toxin and protects the bacteria from host defenses. These pili also serve as receptors for CTXφ which can infect non-pathogenic V. cholerae and provide it with genes that encode cholera toxin (Waldor and Mekalanos, 1996). Thus, CTXφ transduction requires functional TCP in recipient bacteria. In addition to microcolony formation and phage transduction, TCP are necessary for the secretion of the soluble colonization factor, TcpF. (Kirn et al., 2003; Kirn and Taylor, 2005; Megli et al., 2011). V. cholerae strains that do not express TCP fail to colonize the infant mouse infection model (Taylor et al., 1987) or human volunteers (Herrington et al., 1988). The mechanisms by which TCP bind to each other and to CTXφ are not well understood, but are central to the V. cholerae life cycle and infection process.

The focus of this proposal is to examine the structure and assembly of TCP at an atomic level to understand their assembly process and their role in microcolony formation as a critical early step to V. cholerae pathogenesis. These
surface-displayed virulence factors are attractive targets for vaccines and antibacterial therapeutics due to their central role in pathogenesis. Thus, this knowledge can contribute to the development of antimicrobial agents for cholera.

1.3. Type IV pilin and pili

*V. cholerae* TCP belong to a large class of pili, the type IV pili, which are present on many Gram-negative bacteria including *V. cholerae*, *Neisseria gonorrhoeae* and *Pseudomonas aeruginosa* (Craig *et al.*, 2004). Microorganisms generate type IV pili on their cell surfaces to mediate diverse functions necessary for bacterial pathogenesis. For example, *N. gonorrhoeae* gonococcal (GC) pili are responsible for DNA uptake, host-cell adhesion, immune escape, microcolony formation, and twitching motility. Type IV pili of *P. aeruginosa* bind to DNA and can also mediate host-cell adhesion and twitching motility. These pili can tether to surfaces and retract into the bacterium for twitching motility (Merz *et al.*, 2000). Moreover, pilus retraction allows these filaments to uptake DNA upon binding.

Type IV pili are essential virulence factors consisting of thousands of copies of type IV pilin monomers. Typically, these filamentous appendages are several microns long but less than 10 nm in diameter. Type IV pili are classified based on amino acid sequence homology of the pilin subunit (Strom and Lory, 1993). Sequence conservation is found in the hydrophobic N-terminus (see Figure 1.4), while the C-terminal segment is much less conserved. All type IV
pilins have an N-methylated N-terminal residue, a conserved ~25 residue hydrophobic N-terminal segment, a pair of C-terminal cysteines forming a disulfide bridge and common components in the pilus biogenesis apparatus. The N-terminus was proposed to be the polymerization and membrane-anchor domain due to its hydrophobicity (Watts et al., 1982; Schoolnik et al., 1984; Parge et al., 1995; Craig et al., 2006; Craig and Li, 2008). Notably, residue Glu5 in α1N is conserved among type IV pilins and is necessary for pilus expression (Pasloske and Paranchych, 1988; Strom and Lory, 1991; Horiuchi and Komano, 1998; Aas et al., 2007; Li et al., 2012).

1.3.1. Type IVa and type IVb pilin subclasses

Type IV pilins are separated into type IVa and type IVb subclasses based on their amino acid sequence and length (Figure 1.5) (Strom and Lory, 1993; Giron et al., 1997; Craig et al., 2004). The type IVa pilins have a shorter leader sequence (~6 to 8 residues versus ~10 to 30 residues for the type IVb pilins) and their average mature amino acid sequence lengths are shorter than the type IVb pilins (~144 to 160 versus ~170 to 208 residues). While the N-methylated N-terminal residue differs among type IVb pilins, in type IVa pilins it is phenylalanine. The region between the conserved cysteines, the disulfide bridged region (the D-region), in type IVa pilins has an average length of 22 amino acids. In contrast, the D-region of type IVb pilins has an average length of 55 residues. Type IVb pili are associated with enteric pathogens whereas type IVa pili are found on bacteria that can colonize a range of hosts and tissues.
Interestingly, structural analysis of the type IV pilins revealed distinct protein folds for the type IVa and type IVb pilins (Parge et al., 1995; Craig et al., 2003; Craig et al., 2004; Craig and Li, 2008).

Figure 1.4 Type IV pilin amino acid sequence alignment of the N-terminal 55 residues.

Note. (A) Alignment of mature type IV pilin amino acid sequences to the N. gonorrhoeae GC pilin. Identical residues to the GC pilin are shaded dark and homologous residues are shaded light. Cysteine residues are underlined. Ng, N. gonorrhoeae strain MS11; Nm, N. meningitidis; Pa, P. aeruginosa; Dn, Dichelobacter nodosus; Mb, Moraxella bovis; Ec, Eikenella corrodens; EHEC, enterohemorrhagic Escherichia coli O157-H7; ntHi, Haemophilus influenzae; Cp, Clostridium perfringens; Ft, Francisella tularensis Schu S4; Vc, V. cholerae; ETEC, enterotoxigenic E. coli; Cr, Citrobacter rodentium; EPEC, enteropathogenic E. coli; St, Salmonella enterica serovar Typhi; Mm, Methanococcus maripaludis. MMP0236 is a putative type IV pilin from M. maripaludis and Nm PilX is a minor pilin subunit. (B) Alignment of mature type IVb pilins to the V. cholerae TcpA type IVb pilin. Conserved glycines are indicated by arrows. Craig, 2009 (p. 3); adapted and used with permission.
Figure 1.5  Schematic comparison of type IVa and type IVb pilins.

Note. The relative lengths of the leader sequence (black), α1 (grey), αβ-loop (green), β-sheet (grey) and the D-region (magenta), flanked by two conserved cysteines (cyan) are illustrated. The conserved, hydrophobic, α-helical N-termini are boxed in red. Craig et al., 2004 (p. 365); adapted and used with permission.

1.3.2. Structures of the type IV pilins

X-ray crystal structures of full-length pilin subunits have been determined for *N. gonorrhoeae* GC pilin (Parge et al., 1995; Craig et al., 2006) (Figure 1.6A), *P. aeruginosa* strain K (PAK) pilin (Craig et al., 2003) (Figure 1.6B) and *Dichelobacter nodosus* FimA pilin (Hartung et al., 2011). In addition, X-ray and nuclear magnetic resonance (NMR) structures are available for recombinant soluble pilins lacking the N-terminal ~28 residues: enteropathogenic *Escherichia coli* (EPEC) BfpA (Ramboarina et al., 2005), *Francisella tularensis* PilE (Hartung et al., 2011), *P. aeruginosa* K122-4 (Keizer et al., 2001; Audette et al., 2004), *P. aeruginosa* PAK (Hazes et al., 2000), *P. aeruginosa* strain Pa110594 (Nguyen et
Salmonella enterica serovar Typhi PilS (Xu et al., 2004) and V. cholerae TcpA (Craig et al., 2003; Lim et al., 2010) (Figure 1.6C) pilins. No structure of a full-length type IVb pilin has been solved to date.

Figure 1.6 Type IV pilin structures.

Note. Full-length crystal structures of (A) N. gonorrhoeae GC pilin and (B) P. aeruginosa PAK pilin and (C) N-terminally truncated V. cholerae TcpA pilin. The αβ-loop is coloured green, D-region is coloured magenta and the conserved cysteines linked by a disulfide bond are coloured in cyan and shown in ball and stick representation. Craig, 2009 (p. 4); adapted and used with permission.
The full-length type IVa pilin subunit structures from *N. gonorrhoeae*, *P. aeruginosa* and *D. nodosus* have conserved architectures consisting of an N-terminal α-helix, α1, that forms an extended stalk (α1N) and is embedded in a C-terminal globular domain for its second half (α1C) (Parge et al., 1995; Craig et al., 2003; Craig et al., 2006; Hartung et al., 2011). The truncated and full-length PAK pilin structures solved (Hazes et al., 2000; Craig et al., 2003) are essentially identical in the globular domain showing that the protruding N-terminus does not affect the globular domain fold. In the globular domain, α1C of all type IV pilins interacts with an anti-parallel β-sheet to form a structurally homologous hydrophobic core (Figure 1.6). A disulfide bridge in the C-terminal region links the D-region to the conserved core of the globular domain. The D-region and the αβ-loop (segment between α1 and the β-sheet) are variable in sequence, length and structure among the type IV pilins. Importantly, the topology of the globular domains differs between the two pilin subtypes: the β-sheet of type IVa pilins has continuous nearest-neighbour connectivity where each β-strand folds back forming hydrogen bonds with the previous strand in an antiparallel fashion, whereas the C-terminal β-strand of the type IVb pilin β-sheet forms the central strand (Figure 1.7) (Craig et al., 2004; Craig, 2009), suggesting they may have different biophysical properties.
1.3.3. Type IV pilus filament structure

Prior to knowledge of the pilin structures, the N-terminus of type IV pilins was hypothesized to be the polymerization domain based on its conserved amino acid sequence and hydrophobicity (Watts et al., 1982; Schoolnik et al., 1984). The first type IV pilus filament model proposed was that of the *N. gonorrhoeae* type IVa GC pilus by Parge and colleagues (1995) based on the crystal structure of the full-length GC pilin. In this computational model, the N-terminal α-helices of the pilin subunits are packed tightly, forming a helical array in the filament core and the β-sheets of neighbouring globular domains form a continuous β-sheet resulting in a relatively smooth filament surface.

In 2006, a 12.5 Å resolution type IVa *N. gonorrhoeae* GC pilus structure was solved by cryo-electron microscopy (cryoEM) and three dimensional image...
reconstruction (Craig et al., 2006). The GC pilin subunits within the GC filament relate to one another by a rise of 10.5 Å and a rotation of 100.8° for the right-handed one-start helix. The subunits are positioned equidistant from the filament axis, Z, in these helically symmetrical filaments (Figure 1.8A). The pilins are related by an axial rise and an azimuthal rotation about the helical axis of a primitive or one-start helix (Figure 1.8B). A single path connects the pilin subunits within the pilus filament in the one-start helix. To generate a pseudo-atomic resolution GC pilus structure, the 2.3 Å resolution GC pilin crystal structure was filtered to 12.5 Å and docked into the moderate resolution 3-dimensional electron density map generated by cryoEM and image reconstruction. The helical symmetry parameters (10.5 Å/100.8°) were subsequently applied to the single docked subunit resulting in the pseudo-atomic resolution pilus filament structure (Craig et al., 2006) (Figure 1.9 and Figure 1.10).
Figure 1.8  Relationship between type IV pilins in a helically symmetrical pilus.

Note.  (A) Subunits in a helically symmetrical filament are situated equidistant from the helix axis and relate to one another by an axial rise and a rotation about the helix axis.  (B) Spheres representing pilin subunits are shown joined by a primitive right-handed one-start helix where every subunit is connected along a single path in the filament. Craig, 2009 (p. 4); Figure (B) used with permission.
Figure 1.9  Docking of the GC pilin subunit into the TEM density map.

Note.  
(A) 2.3 Å GC pilin crystal structure. The αβ-loop is coloured green, the D-region is coloured magenta and the conserved cysteines linked by a disulfide bond are coloured in cyan and shown in ball and stick representation. (B) GC pilin subunit filtered to 12.5 Å resolution. (C) A single GC pilin subunit (magenta) docked into the GC pilus TEM density map (blue). The subunits relate to one another by a rotation of 100.8° and a rise of 10.5 Å. Image courtesy of Lisa Craig, Simon Fraser University.

In comparison to the computational GC pilus model (Parge et al., 1995), the interactions among the globular domains within the assembled filament structure are less extensive resulting in grooves on the pilus surface that may permit filament flexibility (Craig et al., 2006). These grooves are lined with positively charged residues, which may bind to the negatively charged DNA backbone mediating DNA uptake upon pilus retraction (see Section 1.3.4). The surface-exposed αβ-loop displays post-translational modifications that undergo
phase variation (Figure 1.10B, coloured blue) (Parge et al., 1995; Weiser et al., 1998; Hegge et al., 2004; Craig et al., 2006). Furthermore, the protruding D-region houses the “hypervariable loop” which contains highly variable amino acid sequences among Neisseria pilins (Figure 1.10B, coloured magenta). These surface features may explain the ability of Neisseria pathogens to evade immune system detection. Thus, the GC pilus structure helps to explain the unique biological features and functions of the GC pilus.
Figure 1.10 CryoEM reconstruction and pseudo-atomic resolution structure of the *N. gonorrhoeae* GC pilus.

*Note.* (A) Pseudo-atomic resolution structure (blue) docked in the CryoEM reconstruction (grey) (Craig *et al.*, 2006; Li *et al.*, 2012). (B) Surface representation of the pseudo-atomic resolution structure (PDB code 2HI2 (Craig *et al.*, 2006)). The αβ-loop is coloured green, D-region containing the “hypervariable loop” is coloured magenta and the two post-translational modifications on residues Ser63 and Ser68 are coloured blue.
Less is known about the type IVb pilus structure compared to the type IVa pilus. A computational model for the *V. cholerae* type IVb pilus, TCP, was proposed based on the N-terminally truncated TcpA pilin crystal structure, ΔN-TcpA, and packing arrangement in the crystallographic lattice, the TCP filament dimensions and helical symmetry parameters, rise and rotation per subunit for the 1-start helix, obtained by transmission electron microscopy (TEM) (Figure 1.11) (Craig *et al.*, 2003). The missing α1N of the ΔN-TcpA crystal structure was modeled using atomic coordinates for the corresponding region in the full length *P. aeruginosa* PAK pilin structure (Figure 1.11) (Craig *et al.*, 2003). The N-terminus is predicted to be α-helical in TcpA, due to its sequence homology with the PAK and GC pilins. The amino acid sequence similarity and identity of TcpA and PAK pilin α1N are 75% and 32% respectively. In this filament model, pilin subunits associate via interactions of the N-terminal α-helices, which form a hydrophobic core in the filament, and between the αβ-loops and D-regions of the globular domains (Figure 1.11). Notably, the helical symmetry of this TCP model for the left-handed one-start helix is 7.5 Å rise and 140° rotation which is significantly different than that of the GC pilus (a right-handed helix with a 10.5 Å rise and a 100.8° rotation). The computational TCP model had not been tested experimentally until this study.
Figure 1.11 V. cholerae TcpA pilin structure and computational TCP filament model.

Note. Left. N-terminally truncated 1.3 Å ΔN-TcpA crystal structure with the extended N-terminal α-helix, which was modelled onto the structure based on the coordinates of the P. aeruginosa PAK pilin full-length structure. Right. Computationally derived TCP filament model. The pilin subunits are related to one another by a 7.5 Å rise and a 140° rotation in the left-handed 1-start helix (Craig et al., 2003).

The differences in helical symmetry between the N. gonorrhoeae GC pilus reconstruction and the V. cholerae TCP computational model are surprising given the degree of conservation in the N-terminal polymerization domain. The
objective of this thesis project is to obtain a TEM reconstruction and pseudo-atomic resolution structure of *V. cholerae* TCP to further validate or disprove the computational model, to explain how structure defines TCP functions and to compare a prototypical type IVb pilus with that of a type IVa pilus to understand the pilus assembly mechanism, with implications for understanding pilus mediated secretion and other related secretion mechanisms.

1.3.4. Type IV pilus assembly and disassembly

The mechanism by which bacteria assemble long and very thin type IV pilus filaments on their surfaces is poorly understood. Type IV prepilin subunits are synthesized in the cytoplasm and are directed to the inner membrane via the Sec pathway where the N-terminal signal peptide is cleaved by a dedicated prepilin peptidase (Kaufman *et al.*, 1991; Strom *et al.*, 1993; Craig and Li, 2008). The extended hydrophobic half of α1 is thought to anchor the mature pilin in the inner membrane with the globular domain exposed in the periplasm.

A model for type IV pilus assembly is illustrated in Figure 1.12 based on previously-published data (Craig *et al.*, 2006; Craig and Li, 2008). The complex process of type IV pilus assembly requires at least eight proteins. Type IV pilus biogenesis occurs at the inner membrane, with the filament growing through the periplasm and across the outer membrane via a secretin channel (Bitter *et al.*, 1998; Collins *et al.*, 2001; Ayers *et al.*, 2010). We propose that upon translocation across the inner membrane, pilin subunits diffuse through the inner membrane and dock into an existing gap at the base of the growing pilus.
Subunits are attracted to the growing filament due to electrostatic attraction between a conserved negatively charged Glu5 residue and a positively charged N-terminal residue of the terminal subunit (Craig et al., 2006; Craig and Li, 2008). These exposed charges are unstable in the fatty acyl phase of the bilayer and thus, these complementary charges drive pilin docking. Pilus assembly requires a hexameric assembly ATPase localized to the cytoplasmic side of the inner membrane, which undergoes a conformational change upon ATP binding (Yamagata and Tainer, 2007). We propose that this change is transmitted through the conserved inner membrane protein (IMP) to extrude the growing pilus a short distance into the periplasm, ~8 to 10 Å corresponding to the rise of a single subunit in the 1-start helix (Craig et al., 2006; Yamagata and Tainer, 2007; Yamagata et al., 2012). This opens up a gap at the next site ~120° around the base of the filament for another pilin subunit to dock. As the filament extends, the globular domain of the newly added monomer shields the hydrophobic, extended N-terminal α-helices of the existing subunits of the filament from the aqueous environment of the periplasm and (subsequently the) extracellular space. (Craig et al., 2006; Craig and Li, 2008). With this model, the energy barrier for translocating a pilin subunit from its hydrophobic environment of the lipid bilayer to the hydrophobic environment of the filament is very small, with energy provided by ATP hydrolysis required only to extrude the filament.
Figure 1.12 Type IV pilus assembly model.

Note. (A) This generalized model is based on the type IVa *N. gonorrhoeae* pilus structure and common characteristics among all type IV pili. Pilin subunits, embedded in the inner membrane (IM) by their N-terminal α-helices, are added to three growing strands (each strand is coloured differently) at active sites powered by an ATPase. A single pilin subunit (red) is added to one strand of the growing filament via electrostatic interactions. ATP is hydrolyzed which leads to a “piston-like” movement of a conserved inner membrane protein (IMP) extruding the filament through a pore in the outer membrane and into the extracellular milieu as the filament lengthens. (B) Top view of the pilus assembly apparatus, looking down on the growing pilus filament, showing addition of pilin subunits iteratively to each of the 3 helical starts. The red subunit is added first, the yellow subunit next and the blue subunit last. Craig et al., 2006 (p. 660); Figure (A) adapted and used with permission.
Notably, active pilus disassembly or pilus retraction is a critical process for several bacteria that display type IV pili. Many type IV pili, including *N. gonorrhoeae* GC pili and EPEC bundle-forming pili, disassemble by using a retraction ATPase that drives depolymerization or disassembly of the pilus filament. Type IV pilus retraction is necessary to mediate many cellular functions such as DNA uptake, bacterial dissemination from microcolonies and twitching motility necessary for pathogenesis (Bieber *et al.*, 1998; Wolfgang *et al.*, 1998; Anantha *et al.*, 2000; Aas *et al.*, 2002). Twitching motility involves dynamic processes of extending of the pilus filaments by assembly, tethering to surfaces and retracting of the filaments by disassembly allowing pathogen movement. The disassembly mechanism is poorly understood but presumably, the pilin subunits translocate from the pilus filament back into the lipid bilayer (Figure 1.13) (Craig *et al.*, 2006) and are recycled. Notably, there is no evidence that *V. cholerae* TCP retract.
Figure 1.13 Type IV pilus assembly and disassembly at the inner membrane.

Note. Pilin subunits diffuse in the inner membrane and dock at the base of a growing pilus filament. Type IV pilus assembly begins at the inner membrane, grows into the periplasm, through the outer membrane and into the extracellular milieu. Some type IV pilus filaments have been shown to retract or disassemble and these subunits presumably return to the inner membrane where they are recycled. Image courtesy of Lisa Craig, Simon Fraser University.

The type IV pilus machinery share many features with the T2SS (Nunn, 1999; Peabody et al., 2003; Ayers et al., 2010). The proteins associated with type IV pilus assembly share homology to those of the type II secretion apparatus (Figure 1.14). Both pilus assembly and T2SS machines consist minimally of a prepilin peptidase, pilins or pseudopilins, ATPases, an integral inner membrane protein (IMP), inner membrane anchoring proteins (IMAPs) and an outer membrane secretin. The type IV pilus machineries from *D. nodosus*, *F. tularensis*, and *V. cholerae* secrete enzymes, proteases and virulence factors (Kennan et al., 2001; Kirn et al., 2003; Hager et al., 2006) much like the T2SS,
suggesting that both systems evolved from a common ancestor. Thus, type IV pilus assembly and type II secretion are considered to be similar processes.

Figure 1.14 Type IV pilus assembly, type IV pilus-mediated secretion and type II secretion.

Note. Type IV pilus-mediated secretion presumably occurs by a mechanism comparable to the T2SS. Both systems require an assembly/secretion ATPase, an integral inner membrane protein (IMP), inner membrane anchoring proteins (IMAPs), a secretin and pilin or pseudopilin. It is hypothesized that the pilus or pseudopilus acts as a piston to push proteins through the outer membrane secretin (Sandkvist, 2001). Image courtesy of Lisa Craig, Simon Fraser University.

1.4. Research objectives

The disease cholera continues to be a major health treat in developing countries (Graves et al., 2010; Sinclair et al., 2011). Due to their exposure on bacterial surfaces and their role in colonization of hosts, TCP are attractive
targets for vaccines and antibacterial therapeutics that can contribute to preventing, treating and controlling the spread of the cholera disease. An atomic-level understanding of the type IVb TCP and its assembly process may provide important clues and targets for the design of anti-cholera pharmaceuticals. Moreover, type IV pilus studies are relevant to the type II secretion machinery (Peabody et al., 2003; Johnson et al., 2006; Ayers et al., 2010) and can contribute to treatment and prevention of diseases caused by pathogens that exploit these assemblies and advance our understanding of these critical molecular machines.

The specific aims of this thesis project are as follows:

**Aim 1** – To characterize the *V. cholerae* TCP by mapping the exposed and buried regions of the pilin subunits using hydrogen/deuterium exchange mass spectrometry (DXMS).

**Aim 2** – To obtain a pseudo-atomic-resolution structure of *V. cholerae* TCP using negative-stain and cryo-electron microscopy, helical image processing and fitting of the ΔN-TcpA crystal structure.

**Aim 3** – To investigate the relative stabilities of type IVa and type IVb pilin and pili using thermal and chemical denaturation together with circular dichroism spectroscopy and electron microscopy analysis to test the hypothesis that the type IVb pilin fold imparts greater stability than the type IVa pilin fold due to its C-terminal segment being buried in the β-sheet of the globular domain.
2. *Vibrio cholerae* toxin-coregulated pilus structure analyzed by hydrogen/deuterium exchange mass spectrometry


Contributions to research: Pilin and pilus protein expression and purification, chymotrypsin digestion, electron microscopy, deuterium exchange mass spectrometry, data analysis and co-wrote manuscript.

2.1. Introduction

Type IV pilus structures are of great interest due to their essential role in bacterial pathogenesis and their attractiveness as targets for vaccines and antimicrobial agents. These filaments are crucial to virulence for many pathogenic bacteria where they mediate a diverse array of functions, including DNA uptake, motility, protein secretion, microcolony and biofilm formation and immune escape (Kirn *et al.*, 2000; Merz *et al.*, 2000; Craig *et al.*, 2004; Hegge *et al.*, 2004; Jurcisek *et al.*, 2007). Type IV pili are polymers of pilin proteins expressed on many Gram-negative bacteria (Craig *et al.*, 2004) and at least one Gram-positive organism (Varga *et al.*, 2006). Type IV pilins are divided into type IVa and type IVb pilin subclasses based on differences in amino acid sequence, length and structure (Figure 1.4, Figure 1.5, Figure 1.6 and Figure 1.7). The pilin subunits share a conserved N-terminal α-helix that mediates filament assembly
but the amino acid sequences of the C-terminal globular domains vary considerably among type IV pilins and define the diverse and in some cases unique pilus functions (Craig et al., 2004; Craig and Li, 2008).

*V. cholerae* toxin-coregulated pili (TCP) are type IV pili that self-associate or self-aggregate, allowing the bacteria to clump together forming microcolonies that protect them from host defenses and concentrate their secreted cholera toxin (Taylor et al., 1987; Kirn et al., 2000). Pilus-mediated *V. cholerae* aggregation is essential for colonization of the human small intestine. Moreover, the TCP apparatus serves as a secretion system for TcpF, a protein factor necessary for efficient colonization (Kirn et al., 2003). TCP also act as highly specific receptors for the cholera toxin bacteriophage $\phi$ (CTX$\phi$), which can infect non-pathogenic *Vibrio* species and confer virulence by providing them with genes that encode the cholera toxin subunits (Waldor and Mekalanos, 1996).

Molecular structures of type IV pili can provide valuable insights into the mechanisms by which they accomplish their functions, and may help to identify specific targets for vaccines and therapeutics. These structures have proved challenging to solve due to the hydrophobicity of the full-length pilin subunits, and the size, flexibility and length heterogeneity of the pilus filaments. High resolution full-length and truncated type IV pilin monomer structures that lack the polymerization domain have been obtained by X-ray crystallography and NMR spectroscopy (Parge et al., 1995; Hazes et al., 2000; Keizer et al., 2001; Craig et al., 2003; Audette et al., 2004; Xu et al., 2004; Ramboarina et al., 2005; Craig et
al., 2006). These pilins consist of a conserved N-terminal helix, α1, and a globular C-terminal domain. The C-terminal half of α1, α1C, interacts with an anti-parallel β-sheet and a disulfide bridge in the C-terminal region to form the globular domain. The N-terminal half of α1, α1N, protrudes from the globular domain (see Figure 1.6A and B). Beyond the conserved structural core formed by α1 and the β-sheet, type IV pilin structures vary substantially in two regions: the αβ-loop that lies between α1C and the β-sheet, and the D-region, situated between two conserved, disulfide-bonded cysteines.

Only a single structure of an intact type IV pilus filament has been determined to date, the GC type IVa pilus from *N. gonorrhoeae* (Craig *et al.*, 2006), solved to 12.5 Å resolution by cryo-electron microscopy (cryoEM) and image reconstruction, combined with X-ray crystallography and computational docking (Figure 1.10A). In this structure, α1 forms a helical array in the filament core that anchors the globular domains on the pilus surface (Craig *et al.*, 2006). The globular domains are loosely packed resulting in a highly corrugated filament surface with grooves running between the globular domains that impart flexibility. Some of these grooves are lined with positively charged residues, which may allow DNA binding and uptake into the cell by pilus retraction. Post-translational modifications in the αβ-loop and a hypervariable segment in the D-region, both located on the pilus surface, may allow immune escape. Notably, the αβ-loops and D-regions of numerous other type IV pilins also mediate key roles in pilus functions (Craig *et al.*, 2004).
Less is known about type IVb pilins and pili. Although several structures are available of N-terminally truncated type IVb pilins, no full-length structure has been determined, and no TEM structure has been determined for any type IVb pilus. In the N-terminally truncated *V. cholerae* TcpA pilin structure, subunit packing within the TcpA crystal lattice provided a clue to the type IVb TCP filament architecture as the pilin subunits were arranged in a helical array with their α1Cs oriented into the core of the crystallographic filaments. In the absence of α1N, the ΔN-TcpA subunits were held together by interactions between the αβ-loops and D-regions. A computational model of the TCP filament was generated by integrating the TcpA structure and its helical packing in the crystal lattice with pilus dimensions and helical symmetry obtained by TEM (Craig *et al.*, 2003). In this TCP filament model, the TcpA subunits are held together by hydrophobic interactions among the N-terminal α-helices, which form a core in the filament and are shielded from solvent by the globular domains. As in the crystallographic filament, the αβ-loops and D-regions of the globular domains in adjacent pilin subunits were predicted to interact on the surface of the computational TCP model. Similar to GC pili, residues involved in pilus:pilus interactions were exposed on the surface of the filament (Kim *et al.*, 2000) despite the topology and structure of type IVa and type IVb pilins being different. Although the TCP model was consistent with existing biological data, it had not been experimentally validated and the compact packing of the subunits did not explain the flexibility of the filaments. Furthermore, the helical symmetry of GC pili and TCP differ (right-handed, 10.5 Å/100.8° versus left-handed, 7.5 Å/140°)
respectively) suggesting that their assembly mechanism may be different. Thus, I sought to characterize a type IVb pilus structure to understand the functions of these pili, to compare them with the type IVa pili and to examine the pilus biogenesis process.

To examine the structure of intact TCP filaments, I used hydrogen/deuterium exchange mass spectrometry (DXMS), a powerful technique that exploits amide hydrogen-deuterium exchange in protein to study conformational dynamics, protein structure and protein:ligand interactions (Woods and Hamuro, 2001). A number of hydrogens on a peptide are in equilibrium with protons in the solvent and are exchangeable with deuterium including amide hydrogens of peptide bonds and those on amino acid side chains (\(-\text{NH}_n\), \(-\text{OH}\) and \(-\text{SH}\)). Only peptide amide hydrogens exchange at measurable rates. All amino acids but proline have a single backbone amide hydrogen and thus exchange rates can be determined along the length of the protein backbone. Because the exchange rate for surface-exposed residues is higher than for residues buried in subunit:subunit interactions or in the secondary or tertiary structures of a protein, the rate of deuterium labeling of peptides in a protein will provide information on its location within a protein or protein complex.

In this study, the solvent accessibility of peptides in the TCP filament was compared with that of the soluble, N-terminally truncated TcpA monomer (Figure 2.1). Regions of TcpA that are shown by DXMS to be exposed in the monomer but are buried in the intact pilus filament are involved in subunit:subunit
interactions. The most solvent accessible residues of the monomer and the polymer become deuterated at very short incubation times (10 seconds) and the residues buried by protein folding will likely not show significant exchange even at the longest time points (50 minutes). The soluble pilin subunits should show consistently higher levels of deuteration for all surface exposed amides, as compared to the assembled pili. Amides that show rapid deuterium exchange for the soluble pilins, but slow exchange for the intact filaments are likely to be directly involved in subunit:subunit interactions in the filaments. These DXMS results were used as restraints for refining the computational TCP filament model (Craig et al., 2003). A new interaction interface was tested by altering key residues to disrupt pilus assembly. The resulting TCP model explains pilus flexibility and suggests a molecular mechanism for pilus:pilus interactions or pilus bundling in microcolony formation as well as other pilus functions. Unexpectedly, a segment of the N-terminal α-helix is exposed in a gap on the filament surface, instead of being buried in the filament core.

2.2. Materials and Methods

2.2.1. Plasmids and bacterial strains

V. cholerae strains RT4236, RT4225, RT4524, O395 and CL101, E. coli strain S17, plasmid pTK1 and TcpA-6 polyclonal rabbit antibody were gifts from Ronald Taylor, Dartmouth Medical School. E. coli Origami DE3 cells (Novagen) were transformed with the pET:15b vector (Novagen) containing the gene
encoding N-terminally truncated TcpA (ΔN-TcpA). This was used for expression of the His-ΔN-TcpA monomer containing an N-terminal His-tag/linker (Craig et al., 2003). Details of bacterial strains are described in Table 2.1.

**Table 2.1  Bacterial strains and plasmids**

<table>
<thead>
<tr>
<th>Bacteria</th>
<th>Description</th>
<th>Reference/Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>V. cholerae O395*</td>
<td>Wild type, classical biotype</td>
<td>(Taylor et al., 1987)</td>
</tr>
<tr>
<td>V. cholerae RT4225</td>
<td>TcpA H181A:toxT</td>
<td>(Kirn et al., 2000)</td>
</tr>
<tr>
<td>V. cholerae RT4236</td>
<td>pMT5:toxT</td>
<td>(DiRita et al., 1996)</td>
</tr>
<tr>
<td>V. cholerae RT4524</td>
<td>tcpA::lacZ</td>
<td>R. K. Taylor</td>
</tr>
<tr>
<td>V. cholerae ML10</td>
<td>TcpA C120A</td>
<td>This study</td>
</tr>
<tr>
<td>V. cholerae ML19</td>
<td>TcpA K68A</td>
<td>This study</td>
</tr>
<tr>
<td>V. cholerae ML21</td>
<td>TcpA R26E</td>
<td>This study</td>
</tr>
<tr>
<td>V. cholerae ML22</td>
<td>TcpA L76K</td>
<td>This study</td>
</tr>
<tr>
<td>V. cholerae ML23</td>
<td>TcpA E83R</td>
<td>This study</td>
</tr>
<tr>
<td>V. cholerae ML24</td>
<td>TcpA R26E/E83R</td>
<td>This study</td>
</tr>
<tr>
<td>V. cholerae ML10</td>
<td>pCTX-Knφ</td>
<td>(Kirn et al., 2003)</td>
</tr>
<tr>
<td>E. coli S17</td>
<td>λ.pir</td>
<td>(Skorupski and Taylor, 1996)</td>
</tr>
<tr>
<td>E. coli Origami(DE3)-pET:15b-TcpA</td>
<td>E. coli K12 gor-/trxB-, Δ1-28, hisTcpA</td>
<td>Novagen &amp; (Craig et al., 2003)</td>
</tr>
</tbody>
</table>

* all V. cholerae TcpA mutant strains are derived from O395, classical biotype

**2.2.2. Purification of toxin-coregulated pilus filaments**

*V. cholerae* strains expressing TCP\(^{WT}\) (RT4236) and TCP\(^{H181A}\) (RT4225) cells were grown for 16 hours under TCP-expressing conditions (Luria broth [LB: 1% tryptone, 0.5% yeast extract and 86 mM NaCl], pH 6.5, 30 °C) in a shaking incubator in media supplemented with 100 μg/ml ampicillin and 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Both strains contain the pMT5 plasmid with
an ampicillin resistance marker and an inducible toxT gene that encodes for a transcriptional activator of the tcp operon (DiRita et al., 1996). ToxT over-expression results in increased TCP expression.

2.2.2.1. Purification of TCP$^{H181A}$

*V. cholerae* RT4225(pMT5) cells were derived from classical *V. cholerae* O395 and have a mutation in the tcpA gene at the site encoding amino acid residue 181, resulting in a histidine-to-alanine change (H181A). This strain sheds TCP$^{H181A}$ into the culture supernatant, allowing purification by ammonium sulfate precipitation. Cells were grown overnight in pilus-inducing conditions in LB-Sm/Amp/IPTG. Cultures were centrifuged at 5,000 Xg to remove cells, and additional impurities were precipitated from the culture supernatant using 10% (W/V) ammonium sulfate. TCP$^{H181A}$ were precipitated with 30% ammonium sulfate. The pellet containing TCP$^{H181A}$ was resuspended in phosphate-buffered saline (PBS) with 10 mM ethylenediaminetetraacetic acid (EDTA) and dialyzed exhaustively to remove residual ammonium sulfate. All purification steps were carried out on ice or at 4 °C.

2.2.2.2. Purification of TCP$^{WT}$

*V. cholerae* RT4236(pMT5) cells over-expressing wild type TcpA were grown overnight in pilus-inducing conditions. The culture was centrifuged to pellet the bacteria, which were resuspended in PBS/EDTA. Pili were sheared from the surfaces of the bacteria by passing the bacteria through a 25-gauge needle and syringe followed by vigorous vortexing and homogenization using an Ultra-Turrax
T8.01 disperser (Ika). The pili were concentrated by centrifugation at 5,000 Xg, solubilized in PBS/EDTA and dialyzed exhaustively. This procedure yields a few hundred micrograms of wild type TCP (TCP\textsuperscript{WT}).

2.2.3. Purification of soluble, N-terminally truncated TcpA pilin

Soluble, His-ΔN-TcpA (residues 1-28 [α1N] were replaced with a histidine tag and linker) were expressed in E. coli Origami DE3 cells (Novagen) transformed with ΔN-TcpA-pET-15b and purified as described previously (Craig \textit{et al.}, 2003) with minor modifications. Cells were grown in 6 L of LB media at 37 °C to an optical density (OD) of 0.5 (OD\textsubscript{600}) at which point IPTG was added to a final concentration of 1 mM to induce pilin expression, and cells were grown overnight at 30°C. Cells in overnight cultures were sonicated to release the cytoplasmic proteins, which were then purified by cation exchange chromatography using a HiTrap SP HP column (GE Healthcare) followed by metal affinity chromatography using a HisTrap HP column (GE Healthcare) and size exclusion chromatography on a HiPrep 26/60 Sephacryl S-100 HR column (GE Healthcare). The pilin protein was dialyzed in PBS/EDTA and concentrated to 17.2 mg/ml using an Amicon stirred cell concentrator (Millipore) with a YM10 membrane.

2.2.4. Chymotrypsin digestion

Pilin and pilus protein concentrations were determined by the Bradford method (Sigma) (Bradford, 1976). Chymotrypsin diluted in PBS/EDTA was added
to pre-incubated His-ΔN-TcpA, TCP$^{H181A}$ and TCP$^{WT}$ samples at a chymotrypsin:protein mass ratio of 3:1 and incubated at 37 ºC while shaking. Aliquots were removed at specific time points, and proteolysis was quenched with phenylmethylsulfonyl fluoride (10 mM final concentration). These samples were analyzed by SDS-PAGE followed by Coomassie staining or immunoblotting using a polyclonal antibody against residues 174-199 of TcpA (Sun et al., 1991).

2.2.5. Transmission electron microscopy

Five µl samples of whole cell culture were applied to glow-discharged carbon-coated copper grids (Electron Microscopy Sciences), stained with 1% phosphotungstic acid, pH 7, and imaged on a FEI Tecnai F20 at 120 keV.

2.2.6. Hydrogen/deuterium exchange mass spectrometry

2.2.6.1. Preparation of DXMS samples

Ten µl of TCP$^{H181A}$ and His-ΔN-TcpA (2 mg/ml) was mixed with 10 µl of deuteration buffer (25 mM Tris-HCl, 150 mM NaCl in D$_2$O, pH 7.1) and incubated for varying amounts of time at 4 ºC (Figure 2.1). At time points between 10 seconds and 50 minutes, the “on-exchange” of deuterium was terminated by transferring 20 µl of sample to 40 µl of ice-cold quench solution (1.0 M tris-2-carboxyethyl phosphine [TCEP], 7.4 M guanidinium hydrochloride [GnHCl], pH 2.9). Quenched samples were then mixed with 140 µl of ice-cold 0.8% formic acid and immediately frozen on dry ice and stored at -80 ºC. This process locks the deuterium labels in place and simultaneously denatures and reduces the protein (Burns-Hamuro et al., 2005). All subsequent manipulations were
performed under quench conditions to minimize loss of deuterium labeling. Controls included a non-deuterated protein sample in 25 mM Tris-HCl, 150 mM NaCl in dH2O, pH 7.1 and an equilibrium-deuterated sample treated for 18 hours at 4 °C with 10 μl of D2O containing 0.8% formic acid.

2.2.6.2. Protein fragmentation, peptide identification and deuterium quantification

Frozen samples were transferred to dry ice in the sample basin of the autosampler module of the DXMS data acquisition apparatus at the University of California, San Diego (UCSD) and further processed as described previously (Burns-Hamuro et al., 2005; Brudler et al., 2006). Briefly, temperature control in the apparatus was maintained by storing valves, tubing, columns, and autosampler at 5 °C with columns immersed in ice water. Each sample was held on dry ice and then individually melted at 5 °C in the autosampler. Thawed samples were pumped over a protease column containing immobilized porcine pepsin (30 mg/ml, Sigma) coupled to a 20AL support from PerSeptive Biosystems, 66 μl column bed volume, flowing in 0.05% trifluoroacetic acid (TFA) at 100 μl/min. Extensively proteolyzed protein fragments were collected on a C18 high performance liquid chromatography (HPLC) column (Vydac) then eluted by a linear acetonitrile gradient. Column effluent was directed to the mass spectrometer (ESI LCQ Classic, Thermo Finnigan Inc.) (Figure 2.1).
Figure 2.1 DXMS analysis of His-ΔN-TcpA and TCP^{H181A}.

Note. (A) His-ΔN-TcpA pilin monomers and (B) TCP^{H181A} filaments were deuterated between 10 seconds to 50 minutes (1) and immediately quenched (2) at 0 °C, pH 2.9. (3) The quenching solution contains GnHCl and TCEP, which denatures the deuterium-labeled protein and reduces disulfide bonds respectively. (4) Peptides were digested by pepsin proteases and (5) analyzed by liquid chromatography-mass spectrometry (LC-MS). Image courtesy of Lisa Craig, Simon Fraser University.

Deuterium is a stable isotope of hydrogen that has an extra neutron, which increases its mass by approximately 1 atomic mass unit. Peptide fragments that acquire deuterium labeling at different time points are identified by mass spectrometry and the masses of a deuterated and a non-deuterated sample are compared using DXMS data reduction software (Woods and Hamuro, 2001). These data are used to determine the exchange rates for each peptide.
fragment in the full-length protein. By plotting the percentage of deuterium incorporated onto specific, short peptides as a function of time, one can assess the relative accessibility of the peptide’s amides to its aqueous environment. When identifying protein interaction interfaces within a complex, a difference of 10% or greater in deuteration levels between the monomer and protein complex is considered to be significant (Hamuro et al., 2004; Hsu et al., 2008).

Corrections for back-exchange were determined by employing the methods of Zhang and Smith (1993). The experimentally determined deuteron loss (back-exchange) from peptides in the equilibrium-deuterated reference protein samples was 15% at the beginning (minute 1.5) and 20% (minute 26.5) at the end of the liquid chromatography elution gradient employed for both reference and study proteins. Back-exchange losses for peptides from study proteins were determined by linear interpolation between 15% and 20%, based on peptide retention time.

2.2.7. TCP model refinement

The DXMS data were used as restraints to refine the existing TCP model by the method described in Craig et al., 2003. The full-length TcpA subunit was generated using the coordinates for the N-terminal 28 residues of the PAK pilin crystal structure (Craig et al., 2003) when superimposed on TcpA over Cα atoms 30 to 50. Three rotation angles, x, y and z, for the TcpA subunit, and its radial distance from the filament axis, were varied. Filament models were generated by systematically applying an axial translation (rise) and rotation (twist) to the newly
positioned TcpA subunit using the values determined for the original TCP model (7.5 Å and 140°, respectively for a left-handed 1-start helix). The goal was to maintain the overall architecture and helical symmetry of the original TCP model but to refine the subunit position based on the DXMS data.

The new TCP models were generated by pivoting about Leu76-Cα, unlike the original modeling procedure, which oriented the TcpA subunit about a pivot point at Asp29-Cα for the x, y and z rotations. Pivoting about Leu76 exposes the D-region on the filament surface without forcing the αβ-loop into the center of the filament, consistent with the DXMS data. Hundreds of thousands of TCP models were created, most of which were automatically discarded due to Cα-Cα clashes or unacceptably large (>10 nm) filament diameters. The remaining ~24,000 models were assessed computationally based on having close distances between residues Leu76-Cα, located at the C-terminus of α2 in the center of the most buried region of the αβ-loop, and Arg26-Cα, located within a buried region of α1N, which were predicted to be nearby in the earlier filament model (Figure 1.11). An additional restraint was that Ile179-Cα on the exposed α4 in the D-region was no closer than 10 Å to Lys68-Cα on α2, to adequately expose this region. A model that conformed to the DXMS data was then refined manually, iteratively adjusting the radial distance and x, y and z rotations of the TcpA pilin subunit to optimize the subunit packing and minimize steric clashes.
2.2.8. **Introduction of mutations into the tcpA gene**

QuikChange site-directed mutagenesis (Stratagene) was used to generate missense mutations in the tcpA gene in plasmid pTK1, which was derived from the pKAS32 suicide vector (Skorupski and Taylor, 1996). *V. cholerae* TcpA mutants were generated by graduate student Mindy Lim using allelic exchange (Skorupski and Taylor, 1996). Donor *E. coli* S17 cells carrying pTK1-tcpA* (i.e. pTK1 with the mutated tcpA gene) were mated with *V. cholerae* recipient strain RT4524, and transconjugants were selected. The suicide vector carries the β-lactamase gene for ampicillin resistance and an ori<sup>R6K</sup> replication of origin, which requires the presence of the π protein for plasmid propagation. Since *V. cholerae* lack the π protein, only *V. cholerae* transconjugants that have integrated the plasmid into the genome by homologous recombination will be ampicillin resistant. Transconjugants were then screened for a second homologous recombination event where pTK1 is excised out of the chromosome, taking with it the endogenous wild type tcpA gene and leaving behind the mutated tcpA*. In RT4524, the endogenous tcpA gene has a segment replaced with a lacZ gene, allowing positive strains to be identified as light blue colonies on agar plates containing 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside. Strains that retained the endogenous tcpA-lacZ hybrid gene appear as dark blue colonies. All mutations were confirmed by sequencing the entire tcpA gene.
2.2.9. Analysis of pilin expression and pilus assembly in V. cholerae tcpA mutants

V. cholerae cells were grown overnight under pilus-inducing conditions (LB, pH 6.5, 30 °C) in 2 ml cultures and inspected visually for autoagglutination. Strains that express functional pili cause the V. cholerae cells to aggregate or autoagglutinate and fall to the bottom of culture tubes. Autoagglutination of mutant strains were scored relative to wild type classical V. cholerae strain O139. Cell densities were normalized based on optical density measurements and pilin production was assessed by sodium dodecyl sulfate polyacrylamide gels (SDS-PAGE) and Coomassie-staining or immunoblotting. Pilus assembly was assessed by physically shearing the pili from the surfaces of cells using an Ultra-Turrax T8.01 disperser (Ika) and analyzing the cell homogenate by SDS-PAGE and immunoblotting.

2.2.10. CTXφ transduction assays

Functional pilus expression was also assessed by a CTXφ transduction assay. V. cholerae CL101 cells produce CTXφ-Kn, in which the ctxA gene is replaced with the marker for kanamycin resistance (Waldor and Mekalanos, 1996). Strain CL101 was grown overnight under pilus inducing conditions and the cells were removed by centrifugation followed by filtration through a 0.22 μm pore bottletop filter (Nalgene). Culture supernatant containing CTXφ-Kn was mixed with equal volumes of V. cholerae overnight cultures and incubated at room temperature for 30 minutes. The mixtures were serially diluted and plated on LB-
kanamycin agar plates, incubated at 37 °C overnight, then counted for colony-forming units. *V. cholerae* cells expressing functional TCP are capable of CTXφ-Kn transduction and thus become kanamycin-resistant.

### 2.3. Results

#### 2.3.1. Suitability of TCP$^{H181A}$ to study TCP structure

To examine the *V. cholerae* TCP structure, DXMS was utilized to determine the surface accessibility of short overlapping peptides covering the entire length of the TcpA subunit. Rapid hydrogen/deuterium exchange of TCP peptides indicate regions that lie on the filament surface whereas those that exchange slowly are buried, either by the protein fold or by subunit:subunit interactions within the pilus filament. To differentiate between these two possibilities, hydrogen/deuterium exchange results of the intact TCP filament and the monomeric TcpA subunit were compared. Regions of the pilin protein that are buried by the protein fold should be similarly inaccessible to hydrogen/deuterium exchange in both the monomeric TcpA subunit and the intact pilus filament, whereas regions that are involved in subunit:subunit interfaces will be accessible in the TcpA monomer but not in the filament.

The soluble, N-terminally truncated TcpA pilin monomer used for DXMS analysis has its first 28 amino acid residues ($\alpha 1N$) replaced with a 21 residue segment composed of a hexahistidine tag and linker (His-$\Delta N$-TcpA, (Craig et al., 2003)). The TCP filaments used in this study have a modification at residue
His181 in the TcpA subunit, His181Ala (TCP$^{H181A}$), which causes them to be shed into the culture supernatant (Kirn et al., 2000) for large-scale purification. In contrast, the purity and yield of TCP$^{WT}$ is not sufficient for our studies. To test the suitability of TCP$^{H181A}$ to study TCP structure, the filaments were examined by TEM and subjected to proteolysis. TCP$^{H181A}$ filaments resemble TCP$^{WT}$ by negative stain TEM (Figure 2.2A and B) and are equally resistant to chymotrypsin proteolysis (Figure 2.2C) suggesting that they share the same quaternary structure and thus is an appropriate substitute for TCP$^{WT}$ in the structural studies. Importantly, TCP$^{H181A}$ do not bundle after purification by ammonium sulfate precipitation and concentration, thus the deuterium labeling is not inhibited by pilus:pilus interactions.
Figure 2.2 Comparison of morphology and protease susceptibility for TCP\textsubscript{WT} and TCP\textsuperscript{H181A}.

\textit{Note.} Transmission electron micrographs of (A) TCP\textsubscript{WT} and (B) TCP\textsuperscript{H181A} showing similar morphologies and bundling characteristics, though TCP\textsubscript{WT} bundle to a greater extent. (C) Coomassie-stained SDS-PAGE of TCP\textsuperscript{H181A} (lanes 1, 4 and 7), TCP\textsubscript{WT} (lanes 2, 5 and 8) and the His-\textDelta N-TcpA monomer (lanes 3, 6 and 9) incubated in the absence (0) and presence (+) of chymotrypsin at a chymotrypsin:protein mass ratio of 3:1. Both TCP\textsubscript{WT} and TCP\textsuperscript{H181A} are equally resistant to chymotrypsin proteolysis, remaining intact after 60 minutes of digestion. In contrast, the soluble pilin, His-\textDelta N-TcpA, was rapidly degraded at an early time point.
2.3.2. Solvent accessibility of overlapping peptides in TCP\textsuperscript{H181A} filaments

Initially, non-deuterated TCP\textsuperscript{H181A} filaments were used to optimize conditions for denaturing and reducing the pilins to produce a protease “fragmentation map” of overlapping peptides spanning the entire TcpA sequence. Reduced and denatured TCP\textsuperscript{H181A} filaments were digested on a pepsin column, the peptide fragments were separated, and their masses were determined using liquid chromatography-mass spectrometry (LC-MS). High-quality overlapping peptides were identified covering the entire TcpA sequence, with the exception of the first three amino acids (Figure 2.3A). Next, TCP\textsuperscript{H181A} filaments were incubated in deuterated buffer for 10 seconds to 50 minutes, then quenched to minimize back-exchange, and reduced and denatured under these optimized conditions. Figure 2.4A shows the percent deuteration averaged over six time points for select non-overlapping peptides covering the entire TcpA amino acid sequence. Highly deuterated peptides (>50%) represent solvent-exposed regions of the TcpA protein (coloured orange or yellow in Figure 2.4), whereas poorly deuterated peptides (<20%) represent regions that are buried, either by the protein fold or by pilin:pilin interactions within the TCP\textsuperscript{H181A} filament (coloured black or blue in Figure 2.4). These data show that much of α1 is buried in the TCP filament, as is the C-terminal half of α2 on the αβ-loop, the extended loop following α2, the β-sheet and α3 which is located within the D-region between two β-strands. Unexpectedly, a segment in α1N is solvent-exposed, as is the N-
terminal half of α2 and parts of the D-region, including α4 located at the periphery of TcpA.

Figure 2.3 Overlapping peptides identified by pepsin digest and LC-MS.

Note. Peptide maps of (A) the TCPH181A filament and (B) the His-ΔN-TcpA pilin subunit. Amino acid residues are numbered according to the mature full-length TcpA sequence.
Figure 2.4  Deuterium accessibility for peptides in the TCP^{H181A} filament and His-ΔN-TcpA pilin subunit.
Note. (A) Average deuteration levels for non-overlapping peptides in the TCP^{H181A} filament are shown. Bars span the length of each peptide fragment and are coloured according to their average deuteration level over six time points (10, 30, 100, 300, 1,000, and 3,000 seconds), reported as a percentage of a fully deuterated peptide as shown in the legend. The first two residues of each peptide are not shown as they are fully deuterated at all time points. The secondary structure of TcpA is indicated above the bar graph: blue coils represent α-helices and red arrows represent β-strands. (B) 1.3 Å His-ΔN-TcpA crystal structure. The N-terminal 28 residues not present in the truncated TcpA pilin structure were modeled onto the ΔN-TcpA structure using the coordinates for the PAK pilin N-terminus (Craig et al., 2003). TcpA is coloured as in (A) according to its deuterium accessibility in the TCP^{H181A} filament. Segments coloured in grey are regions not represented in the set of non-overlapping peptides. Residues Cys120 and Cys186, which delineate the D-region, are shown in a ball-and-stick representation. Structural regions and residues used for filament modeling and/or mutational analysis are labeled. (C) Comparison of deuteration levels for identical peptides in the His-ΔN-TcpA monomer (upper lines) and the TCP^{H181A} filament (lower lines), mapped onto the TcpA sequence. The pilin subunit and the filament have identical amino acid sequences for residues 29 to 199, with the exception of residue 181, which is an alanine in the TCP^{H181A} filament. Secondary and other structural features are indicated above the sequence. Peptides are shown for six time points: 10, 30, 100, 300, 1,000, 3,000 seconds, from top to bottom. Each peptide is labeled by the residues it spans, with its average percent labeling over the six time points shown in parentheses.

2.3.3. α1 and the αβ-loop form the only subunit:subunit interfaces in the TCP filament

The solvent accessibility of the TCP^{H181A} filament was compared to that of the His-ΔN-TcpA monomer to identify regions that were buried only within the pilus filament: these are the regions that form subunit:subunit interaction interfaces. Because the His-ΔN-TcpA pilin subunit lacks the N-terminal 28 residues, only residues 29-199 can be directly compared between the two samples. A complete overlapping fragmentation pattern was obtained for the His-ΔN-TcpA monomer using the same denaturation and reduction conditions as for the TCP^{H181A} filaments (Figure 2.3B). Fourteen peptides common to both the His-ΔN-TcpA monomer and the TCP^{H181A} filament were identified covering 80% of the shared amino acid sequence. Surprisingly, both samples showed very similar
patterns of deuteration over much of the protein length (Figure 2.4C), indicating that large regions of the protein are equally solvent-accessible in the monomer and filament forms, and thus are not involved in subunit:subunit interactions. Peptides that show similar low levels of deuteration in both the monomer and the filament indicate regions of TcpA that are buried by pilin protein folding rather than by subunit:subunit interactions within the filament form. Two regions showed substantially reduced levels of deuteration in the TCP filament compared to the TcpA pilin subunit, indicating subunit:subunit interfaces: residues 31-50, which covers most of α1C; and residues 68-83 of the αβ-loop.

2.3.3.1. α1C

Peptides 31-40 and 47-50 are moderately well-labeled in the His-ΔN-TcpA monomer (32.8% and 26.2% average deuteration, respectively), whereas corresponding peptides in the TCP^{H181A} filament are protected from hydrogen/deuterium exchange (16.7% and 10.2%, respectively) (Figure 2.4C). Although residues 41-46 cannot be directly compared due to the lack of a single common peptide, peptides in this region of the His-ΔN-TcpA pilin subunit are also moderately exposed, whereas peptides in this region of the TCP^{H181A} filament are only minimally deuterated for all time points measured (Figure 2.4A). Thus, α1C is buried by interactions with neighbouring subunits in the TCP filament.

2.3.3.2. The αβ-loop

The αβ-loop (residues 53-90) lies between α1C and the core β-sheet of the TcpA globular domain. A segment of the αβ-loop, between residues 68-83
which spans the C-terminal half of α2 followed by a short α-helical turn, is poorly labeled in the TCP\textsuperscript{H181A} filament relative to the His-ΔN-TcpA monomer. Two peptides are directly comparable for this region: peptides 68-73 and 76-83 are moderately well-labeled in the monomer (27.2% and 31.2% average deuteration, respectively) yet they are relatively inaccessible to deuterium exchange in the TCP\textsuperscript{H181A} filament (12.0% and 18.9%, respectively, Figure 2.4C). A peptide preceding and partially overlapping this segment, residues 66-69, is well labeled in both samples (41.0% average deuteration for the monomer, and 50.3% average deuteration for the TCP\textsuperscript{H181A} filament, Figure 2.4A), which suggests that the buried region of the αβ-loop begins at residue 70. This protruding region of the αβ-loop was previously predicted to interact with the D-region on neighbouring subunits in the published computational TCP model based on the interactions observed between His-ΔN-TcpA subunits in the crystal lattice (Craig et al., 2003). However, the DXMS results indicate that the D-region is well-labeled in both the subunit and filament. Thus, the D-region does not interact with the αβ-loop (see below), and residues ~70-83 must interact elsewhere with a neighbouring TcpA pilin subunit.

2.3.3.3. α1N

Although the deuterium accessibility of the N-terminal half of α1, α1N, in the TCP\textsuperscript{H181A} filament cannot be directly compared with that of the His-ΔN-TcpA monomer which lacks this segment, the results are nonetheless informative. Like α1C, much of α1N does not readily exchange with deuterium. Peptides spanning
residues 6-10 and 26-50 are almost completely inaccessible (Figure 2.4A and B). This is not surprising as the N-terminal α-helix is predicted to be buried within the filament core in the computational TCP model, and is buried in the *N. gonorrhoeae* GC pilus cryoEM structure (Craig *et al.*, 2003; Craig *et al.*, 2006). However, peptides spanning residues 13-23 are unexpectedly well-labeled (~50-65%), indicating that this segment of α1N is solvent-exposed. Thus, any model for the TCP filament assembly must bury most of the N-terminal α-helix, but expose residues 13-23 of α1N.

### 2.3.4. The D-region is exposed in the TCP filament

The D-region spans a significant portion of the TcpA pilin globular domain and is delimited by disulfide-bonded cysteines at residues 120 and 186 (Figure 2.4B). The computational TCP model predicted that α4 (residues 179-184), which lies within the D-region on a protruding edge of the globular domain, would interact with the αβ-loop of a neighbouring subunit (Craig *et al.*, 2003). However, the DXMS data show no such interaction, as a peptide covering α4 (residues 176-181) is the most highly-labeled peptide in the TCP\(^{H181A}\) filament (Figure 2.4A and B). The first third of the D-region (residues 120-138) forms part of the β-sheet and α3, which are buried in both the filament and the monomer. However, the second two thirds of the D-region (residues ~145-182) is the most solvent-accessible region of the protein in both samples (Figure 2.4) and is therefore not involved in subunit:subunit interactions. In fact, peptide 158-162 is even more readily-labeled in the TCP\(^{H181A}\) filament (51.2% average deuteration) than in the
monomer (39.3% average deuteration), as are peptides 147-155 and 165-176 flanking this region. This unexpected result may be an artifact of the His-tag/linker at the N-terminus of the His-ΔN-TcpA pilin monomer interacting with peptide 158-162 and its surrounding region and protecting them from hydrogen/deuterium exchange. In support of this, peptides spanning the His-tag/linker of the monomer are only moderately accessible to deuterium exchange (26.3% average deuteration), suggesting that this segment interacts with the globular domain of His-ΔN-TcpA. This interaction is not likely to be tight or well-ordered since the His-tag/linker segment was not resolved in the His-ΔN-TcpA pilin crystal structure (Craig et al., 2003).

2.3.5. Refinement of the TCP model based on DXMS results

The earlier TCP model (Craig et al., 2003) was derived from: (i) the soluble monomeric His-ΔN-TcpA crystal structure; (ii) the TEM-derived filament dimensions and helical symmetry parameters (Craig et al., 2003); and (iii) subunit:subunit interactions observed in the His-ΔN-TcpA crystal lattice (Figure 1.11, PDB code 1OR9 (Craig et al., 2003)). In this computational model, pilin subunits are arranged in a left-handed 1-start helix with a rise of 7.5 Å and a rotation of 140°. This helical symmetry is consistent with the 45 Å layer lines observed in the power spectra of TCP bundles from TEM micrographs: the 45 Å layer lines indicate a dominant 3-start helix with a 135 Å pitch, the vertical distance between a full turn of the helix, with each of the three helical strands separated by 45 Å (Craig et al., 2003). These subunits not only link by a single
path in the 1-start helix (Figure 2.5A), they also link by multiple paths in the 3-start helix (Figure 2.5B). The TCP filament has a maximum diameter of 88 Å measured from the outermost atom centers. The TcpA subunits are oriented such that their N-terminal α-helices form a hydrophobic core, leaving a small channel (<9 Å) running through the center of the filament axis. Subunits are held together by interactions among the N-terminal α-helices and by an interface between the αβ-loop of one subunit and the D-region of a neighbouring subunit that runs along the strands of the left-handed 3-start helix, creating a relatively smooth filament surface.

**Figure 2.5 Helical starts or paths of a pilus filament.**

*Note.* (A) Spheres representing pilin subunits are shown joined by a primitive one-start helix where all subunits are connected along a single path. The subunits relate to one another by an axial rise and an azimuthal rotation. (B) A symmetrical helical filament can possess many helical starts or paths. For example, the same helix is also represented as a 3-start helix where three separate paths trace the connectivity of all subunits in the filament. Craig, 2009 (p.4); adapted and used with permission.
DXMS analysis of *V. cholerae* TCP allowed us to refine our computational model (Craig *et al.*, 2003) and position the TcpA subunits more precisely within the filament. In the newly refined TCP model, the positions of the subunits have been adjusted such that a segment of α1N is exposed, the αβ-loops are turned slightly into the filament core and the D-regions are turned outward, creating a bumpy surface (see Figure 2.6B), consistent with the DXMS data. This TCP model, shown in Figure 2.6, differs from the previous model in the following ways: (i) the N-terminal α-helices now completely fill the filament core, leaving no space for a central channel or for water diffusion (Figure 2.6A and B); (ii) the filament is held together primarily by interactions between the N-terminal α-helices (Figure 2.6A to E), with much of the globular domain being solvent-exposed; (iii) no interaction interface is formed between the αβ-loop of one subunit and the D-region of a neighbouring subunit – instead, the αβ-loop interacts with α1N and α3 of a neighbouring subunit (Figure 2.6C and F); and (iv) the surface of the TCP filament is no longer smooth, but has deep cavities between the subunits, exposing a short segment of α1N, and the protruding D-region of each subunit forms bumps along the pilus surface (Figure 2.6B to D and G). The orientation of the subunits and the overall architecture of the TCP filament model resemble that of the *N. gonorrhoeae* type IV GC pilus cryoEM reconstruction (Craig *et al.*, 2006) in spite of them having substantially different helical symmetries (GC pili have a right-handed 1-start helix with a 10.5 Å rise and a 100.8° rotation).
The N-terminal α-helices pack remarkably well into the filament core, each contacting six neighbouring α1s without any steric clashes (Figure 2.6A to C). It is anticipated that the precise arrangement of the flexible N-terminal α-helices in the pilus interior will differ somewhat from the TCP model, which was generated using, for the most part, rigid body subunits, whose α1N segments have been modeled using the PAK pilin coordinates. Nonetheless, the general arrangement is likely to be maintained. Only minor adjustments were made to the dihedral angles of the two most N-terminal residues, Met1 and Thr2, in order to orient the positively charged N-terminal amide group of Met1 toward the negatively-charged Glu5-Oε1 of the neighbouring subunit in the 1-start helix (Figure 2.6E). Thus, this model supports the proposal that Glu5 is required in type IV pilins to neutralize the N-terminus in the otherwise hydrophobic interior of the pilus filament (Parge et al., 1995; Craig et al., 2006).
Figure 2.6  Refined TCP filament model.
Note. (A) Pseudo-atomic TCP filament model viewed as a ribbon diagram from the distal end of the filament shows the packing of the N-terminal α-helices. (B) Surface representation of the TCP model shows the N-terminal α-helices completely filling the core of the filament, leaving no solvent-filled channel for hydrogen/deuterium exchange. The bulky D-regions protrude from the surface of the filament giving it a bumpy architecture. (C) Side view of the TCP filament model as a ribbon and (D) surface representation. Residues 13-23 of α1N, which were readily labeled by deuterium and thus solvent-exposed, are coloured red. His181 is blue, Asp113 is yellow, Glu158 is orange, and Asp129 is magenta. Additional residues in the D-region (Asp175, Glu183 and Lys187) shown by Kirn and colleagues (2000) and (Ala138 and Ala156) shown by Lim and colleagues (2010) to be involved in pilus:pilus interactions are green and black respectively. (E) Close-up of the TcpA N-termini showing putative electrostatic interactions between the positively-charged N-terminal amine of Met1 and the negatively-charged Glu5 side chain of a neighbouring subunit. (F) Close-up of the α2:α1N/α3 interaction interface, formed between the αβ-loop of one subunit (green) and α1N and α3 of a neighbouring subunit (yellow). This is shown as if looking out from the core of the filament. (G) Close-up of the cavity on the TCP filament surface, which exposes the polar face of residues 13-23 of an amphipathic segment of the protruding N-terminal α-helix. Key residues, including the three glycines, are labeled.

Another interface holding the subunits together in the TCP filament is formed between segment 70-83 of the αβ-loop, and two segments of a neighbouring subunit: residues 23-34 of α1N and residues 122-126 of α3. This interaction is shown by the green and yellow subunits in Figure 2.6C and viewed from the core of the filament in Figure 2.6F. This interface is referred to as the α2:α1N/α3 interface. Residues 70-83 of the αβ-loop protrudes on one side of the TcpA globular domain and fits into a crevice in a neighbouring subunit, interacting with α1N as well as residues 29-34. α2 of one subunit also interacts with α3 of another subunit in the D-region at the bottom of the globular domain of the same neighbour. α3, consisting of residues 117-128, was not initially expected to form an interaction interface from the DXMS data as this region was poorly-labeled in both the His-ΔN-TcpA monomer and in the assembled filament (see peptides 112-122 and 125-131 in Figure 2.4C). However, assuming the N-
terminal His-tag/linker of monomeric His-ΔN-TcpA interacts with segment 158-162 on the outer face of the globular domain, it very likely crosses over α3 in order to reach that segment. Thus, it is reasonable to conclude that α3 is partially buried by the His-tag/linker in the His-ΔN-TcpA monomer and by the αβ-loop of a neighbouring subunit in the TCPH181A filament. Consistent with the DXMS data, the α2:α1N/α3 interface within the filament buries the solvent-inaccessible segment of the αβ-loop but leaves α4 solvent-exposed. Apart from α1C and discrete sections of the αβ-loop and α3, the globular domain does not interact with neighbouring subunits in the filament and is freely accessible to solvent.

The DXMS data indicate that instead of the pilin subunits snugly fitting together to form a smooth, continuous filament surface, the TcpA globular domains are loosely packed, leaving gaps on the TCP surface (Figure 2.6C, D and G). Each gap is flanked by four globular domains, forming a cavity that exposes segment ~11-29 of α1N of the uppermost (i.e. most distal) of these subunits. Residues 13-23 are the most exposed in α1N, consistent with the DXMS results (Figure 2.4A and B). Although residues 24-29 are also exposed in the refined TCP model, they were not well-labeled, most likely because of interactions with α-helices and αβ-loops of neighbouring subunits. Only one face of the N-terminal segment is solvent exposed in the model, with the back face interacting with α-helices in the core of the filament. Thus, it is somewhat surprising that a segment in this region, residues 13-23, is so well-labeled (~50-64% deuteration, see Figure 2.4A). This may be in part due to the abundance of
residues with small side chains (Gly14, Ser17, Ala18, Gly19, Ala24), which would allow better access of solvent to the amide hydrogens (Figure 2.6G and Figure 2.7A). Also, the unusual presence of three glycines (Gly11, Gly14 and Gly19) within this extended α-helix may impart flexibility in this segment and disrupt the α-helical backbone bonding, increasing its solvent accessibility. Although there are many hydrophobic residues in α1N, residues 13-29 is within an amphipathic segment of the α-helix, with the polar face being solvent-exposed (Figure 2.7B). A comparison of the α1N segments of type IV pilins reveals that this amphipathic segment, including the preponderance of glycine residues, is unique to a subset of the type IV pilins, the type IVb pilins, which are found almost exclusively on enteric pathogens (Figure 2.7A). Thus, the pronounced cavity observed on the surface of the TCP filament may be a common structural feature of this class of pili, which include bundle-forming pili (BFP) from EPEC and colonization factor antigen III (CFA/III) pili from enterotoxigenic E. coli (ETEC) (Giron et al., 1997).
Figure 2.7 Amino acid sequence alignments of type IV pilin α1N and helical wheel representation of the amphipathic N-terminal segment.

Note. (A) Residues 1-30 are shown by their one-letter amino acid codes. Polar residues are shaded. Boxed residues indicate the highly deuterated segment of V. cholerae TcpA α1N. Vc, V. cholerae; Ec, ETEC (CofA and LngA) and EPEC (BfpA); Cr, Citrobacter rodentium; Ng, N. gonorrhoeae; Nm, N. meningitidis; Pa, P. aeruginosa; Dn, D. nodosus; Hi, H. influenzae. (B) Helical wheel diagram of residues 11-29 of TcpA. Polar and glycine residues are shaded. The polar face of the α-helix is exposed in the TCP filament model, as indicated.

Mutational analysis by Kirn and colleagues (2000) determined that several residues in the protruding D-region were involved in pilus bundling and microcolony formation, validating the refined TCP model. The side chains of these residues are among the most protruding side chains of the TcpA subunit (Asp175, Glu183 and Lys187, coloured green in Figure 2.6D), or lie within the cavity formed between the globular domains (Asp129, coloured magenta in Figure 2.6D). Furthermore, residues Ala138 and Ala156, coloured black in Figure 2.6D, were shown by Lim and colleagues (2010) to affect V. cholerae aggregation. It is proposed that pilus:pilus interactions may require the protruding D-regions of one filament to intercalate into the cavities between the subunits in an adjacent filament.
2.3.6. TcpA mutagenesis to test the refined TCP model

The accuracy of the TCP model was verified by generating missense mutations for select residues at the α2:α1N/α3 interaction interface. Arg26 is one of only three charged residues in α1N and is positioned in the refined model to form an intermolecular salt bridge with Glu83 on the αβ-loop of a neighbouring subunit (Figure 2.6F). To test the involvement of Arg26 and Glu83 at this subunit:subunit interface, Arg26 was changed to Glu (TCP\textsuperscript{R26E}) and Glu83 was changed to Arg (TCP\textsuperscript{E83R}), both individually and as a double mutation (TCP\textsuperscript{R26E/E83R}). The \textit{V. cholerae} mutants were assessed for (i) their ability to produce pilin protein by immunoblot analysis of whole cell lysates; and (ii) their ability to assemble TCP by TEM analysis and by the presence of TcpA protein in the supernatant of homogenized cells. Pilus morphology and bundling characteristics were also assessed by TEM. To test whether these mutants express functional pili, aggregation or “autoagglutination” of overnight cell cultures was observed, which indicates pilus:pilus interactions. Furthermore, phage transduction assays were conducted to assay pilus function, since TCP are the receptors for CTXφ. Pilin protein was observed at wild type levels in whole cell cultures for both the single \textit{V. cholerae} point mutants, TCP\textsuperscript{R26E} and TCP\textsuperscript{E83R}, indicating that these alterations do not disrupt the pilin fold (Figure 2.8A). This is in contrast to a pilin folding mutation, Cys120Ala (TCP\textsuperscript{C120A}), which disrupts the conserved disulfide bridge and abrogates pilin production. Pilus assembly was assessed by immunoblot and TEM analysis of the supernatant after shearing of the pili from the cells. Only a very faint pilin band was seen for
TCP^{E83R}, and no pilin band was observed in the homogenized cell supernatant for TCP^{R26E}, indicating that these mutants assembled little or no pili respectively (Figure 2.8A and B). Some TCP^{E83R} filaments were observed by TEM (Figure 2.8C), confirming low levels of pilus expression for this mutant strain. This significant reduction in pilus expression for the TCP^{R26E} and TCP^{E83R} mutants demonstrates a role for Arg26 and Glu83 in pilus assembly. Notably, the TCP^{R26E/E83R} double mutant partially rescued pilus assembly, producing approximately half the level of TCP^{WT} (Figure 2.8A), thus supporting a direct interaction between Arg26 and Glu83, as predicted in the refined TCP model (Figure 2.6F). Although pilus morphology appeared normal for TCP^{E83R} and TCP^{R26E/E83R} by TEM (Figure 2.8C and D respectively), the typical rope-like bundles seen for TCP^{WT} were not observed. Instead, TCP^{E83R} filaments formed tangled masses, and the TCP^{R26E/E83R} filaments twisted around each other and themselves more than wild type. Moreover, both mutant strains showed reduced autoagglutination and phage transduction beyond what could be accounted for by reduced pilus expression (Figure 2.8B). This suggests that in addition to affecting pilus assembly, these amino acid substitutions have changed the surface characteristics of the filaments impacting their function.
Figure 2.8 Analysis of pilin expression, pilus assembly and pilus functions in TcpA mutants.

Note. (A) Immunoblot showing TcpA pilin subunit expression in whole cell lysates (top panel) and TCP filament assembly, indicated by the presence of pilin protein in sheared cell homogenates (bottom panel). Wild type refers to *V. cholerae* strain O395, classical biotype. (B) Pilin and pilus expression were also assessed by autoagglutination and phage transduction. (C) TEM of TCP<sub>E83R</sub> and (D) TCP<sub>R26E/E83R</sub> from *V. cholerae* mutants.

Two other residues were tested for their involvement in pilus assembly. Leu76, a residue that is centrally positioned in the α2:α1N/α3 interface at the C-terminus of α2 (Figure 2.6F), was changed to lysine. The TCP<sub>L76K</sub> mutant produced pilin, but no pili were assembled, as determined by immunoblot.
analysis, TEM and functional studies (Figure 2.8A and B). These results implicate Leu76 in filament assembly, showing that the α2:α1N/α3 interface involves both hydrophobic and ionic interactions. Furthermore, Lys68 is located in the middle of α2 (Figure 2.6F) and was readily labeled in the DXMS studies (Figure 2.4A), and thus is not expected to interact with neighbouring pilin subunits in the filament. Consistent with this and the TCP filament model, modifying Lys68 to alanine had no effect on pilus assembly or pilus functions (Figure 2.8A and B).

2.4. Discussion

The DXMS results were combined with structural data to generate a new molecular model for the TCP filament, which is supported by mutational analyses. This structure is a solid helical filament with no central channel. The pilin subunits are anchored by hydrophobic interactions among the N-terminal α-helices and by hydrophobic and polar interactions between the αβ-loop and α1N/α3, which leaves much of the globular domain solvent-accessible. The surface of the filament is bumpy in appearance, as it is decorated with bulges from the protruding D-regions, and cavities between the loosely packed globular domains expose a segment of α1N. These structural features help explain how type IV pili can possess high tensile strength (> 100 pN (Maier et al., 2002)) yet be flexible. This structure also explains several pilus features and functions, and
has important implications for the design of antibacterial and anti-virulence therapies.

Residues implicated in TCP bundling and *V. cholerae* cell autoagglutination (Kirn *et al.*, 2000; Lim *et al.*, 2010) localize to the protruding D-regions and cavities, prompting the proposal of a mechanism for pilus:pilus interactions: the D-region bulges of one filament intercalate into the cavities of adjacent filaments and vice versa holding the filaments together. These interactions extend over long lengths of the pilus filaments (see Figure 2.2A and B) and would offer a strong cumulative force to hold the *V. cholerae* cells together in microcolonies on the intestinal epithelium. Such bulge-in-cavity interactions between pili may explain pilus-mediated aggregation for other bacteria.

Indirect evidence suggests that the repeating cavities on the TCP surface may also provide binding sites for pilus assembly proteins and inner membrane anchoring proteins. The pili used for the DXMS studies carry the H181A modification, which causes them to fall off cells. In spite of this, *V. cholerae* H181A mutants express wild type levels of pili, suggesting that TCP\(^{H181A}\) shear off the cells because they grow too long and not because they are less stable. Residue His181 is situated on a protruding segment of the D-region exposed at the edge of the cavity (coloured in blue in Figure 2.6D), and thus, may interact with ruler-like protein(s) that sense and control pilus length, or with an as-yet-unidentified type IV pilus retraction motor. Residues in the TCP cavities may also
bind exogenous agents such as CTX$\phi$. In support of this idea, a single residue substitution at Asp113, which lies deep in the cavity (coloured yellow in Figure 2.6D), disrupts autoagglutination and TCP-mediated CTX$\phi$ transduction, although the pili have wild type properties in every other respect (Lim et al., 2010).

The cavities generated by the loosely packed globular domains expose a segment of $\alpha$1N and provide a compression space to allow for the flexibility observed for TCP, GC pili and other type IV pili. The exposed segment of the N-terminal $\alpha$-helix and surface gap may also be present in other type IVb pilins as they share amphipathic, glycine-rich segments in $\alpha$1N. No full-length crystal structure of a type IVb pilin has been solved to date, and it is possible that the N-terminal $\alpha$-helical segment is distorted around the glycines. Notably, this segment is also exposed, though to a lesser degree, in grooves between subunits of the *N. gonorrhoeae* type IVa GC pilus (Craig et al., 2006). Thus, the exposed N-terminal $\alpha$-helices may represent conserved targets for therapeutics for bacterial pathogens that utilize type IV pili. Such agents would bind in the cavity, perhaps to accessible residues on $\alpha$1N, and block essential pilus functions such as bundling or retraction. This approach could be particularly effective for organisms that infect the lungs and gastrointestinal tract, such as *P. aeruginosa* and *V. cholerae* respectively, since pilus-specific agents could be inhaled or ingested for delivery to the infectious pathogens without having to travel through the bloodstream, thus increasing their efficacy and minimizing their toxicity. For pathogens expressing retractile pili, the bacteria could take up bacteriocidal or
bacteriostatic agents selectively during pilus retraction. Although pilus retraction has not been shown for *V. cholerae* TCP, type IV pilus retraction is a common feature for pili that mediate twitching motility and DNA uptake (Burrows, 2005).

Amino acid changes to residues in the TCP α2:α1N/α3 interface had a dramatic effect on pilus expression, confirming the contribution of this interaction in pilus assembly. However, if this interface were essential for pilus stability, the TCP^{E83R} mutant should not produce TCP filaments, but our results show some TCP^{E83R} expression. The α2:α1N/α3 interface is relatively small when compared to the extensive contacts among the N-terminal α-helices throughout the filament. Thus, it is remarkable that single mutants R26E, L76K, and E83R have such dramatic effects on pilus assembly. We propose that the α2:α1N/α3 interface is not required for filament stability, but is instead involved in pilus assembly by initiating the addition of each subunit to the growing pilus. During type IV pilus biogenesis, pilin subunits are translocated across the inner membrane and likely remain anchored in the membrane via their extended, hydrophobic α1N helices before being incorporated into a growing pilus filament. In our model, pilus assembly is driven partly by charge complementarity between the conserved Glu5 side chain of the membrane-localized pilin and the positively charged N-terminus of the terminal subunit within the growing filament (Parge *et al.*, 1995; Craig *et al.*, 2006). Mutational results from this study suggest that charge complementarity between Arg26 and Glu83 may also facilitate pilus assembly. These electrostatic attractions would help to facilitate transfer of the pilin subunit.
from the hydrophobic reservoir of the inner membrane to that of the pilus filament. Furthermore, since the L76K mutant does not produce pili, it is clear that overall chemical complementarity and not just electrostatic complementarity is imperative for pilus biogenesis. When complementarity between these two surfaces is lacking, pilus assembly may occur but is much less efficient, explaining the dramatic reduction in pilus expression for the α2:α1N/α3 interface mutants. Once the pilus is assembled, it is the extensive hydrophobic interactions among the N-terminal α-helices that hold the subunits in place to provide structural stability.

The newly refined TCP model elucidates how type IV pili can assemble by a conserved mechanism and possess similar architectures yet display astonishing functional diversity. Presumably all type IV pili are held together by hydrophobic interactions among their conserved N-terminal α-helices. The β-sheets of the globular domains provide a physical link between the α-helical core of the filament and the exposed surface, which displays substantial sequence, length and structural variability. The arrangement of the pilin globular domains on the filament surface presents even more structural variation, resulting in unique chemical and shape properties that define the diverse and often unique functions of type IV pili necessary in the bacterial life cycle and in pathogenesis. Cavities and grooves around the globular domains allow for filament flexibility and offer unique interaction surfaces. The DXMS data suggest that the N-terminal α-helix, in addition to its well-established role in pilus assembly, also defines the filament
surface and hence its functions. This new TCP model provides testable hypotheses to relate type IV pilus structure to the assembly mechanism and its many functions, and offers the possibility of novel strategies for antibacterial therapeutic agents.
3. Structure of the *Vibrio cholerae* type IVb toxin-coregulated pilus


Contributions to research: TCP protein expression and purification, electron microscopy, three-dimensional image reconstruction of TCP, scanning transmission electron microscopy analysis, autoagglutination assay, phage transduction assay and co-wrote manuscript.

Accession numbers: The TCP TEM reconstructions have been deposited in the Electron Microscopy Data Bank under accession numbers EMDB 1954 and EMDB 1955.

3.1. Introduction

Type IV pili are filamentous appendages found on many bacterial pathogens (Craig *et al.*, 2004). These filaments are polymers of the pilin subunit, a 15-20 kDa protein that is comprised of a conserved, protruding ~25 residue α-helical N-terminus and a less conserved globular domain with a cysteine pair. Type IV pilins are divided into type IVa and type IVb subclasses based on differences in amino acid sequence, length and structure. Type IV pili are essential for bacterial virulence and their functions vary from one organism to another. *V. cholerae* type IVb TCP self-associate to hold bacteria together in microcolonies, representing a key step in bacterial colonization of the small intestine to cause the gastrointestinal disease cholera (Taylor *et al.*, 1987; Kirn *et
al., 2000; Lim et al., 2010). They secrete a soluble colonization factor, TcpF, to facilitate bacterial pathogenesis (Kirn et al., 2003; Kirn and Taylor, 2005; Megli et al., 2011). TCP also serve as high affinity receptors for the cholera toxin phage (CTXφ) (Waldor and Mekalanos, 1996). Other bacterial pathogens, including N. gonorrhoeae, P. aeruginosa, Haemophilus influenzae, S. Typhi, enterotoxigenic and enteropathogenic Escherichia coli (ETEC and EPEC, respectively) use type IV pili for colonization and invasion of human hosts.

Type IV pilins have an unusual N-methylated N-terminal residue and a glutamate at residue 5. Prepilin subunits are synthesized with an N-terminal signal sequence that is cleaved by a dedicated prepilin peptidase, a bifunctional enzyme that also catalyzes the N-methylation step, in the inner membrane (Kaufman et al., 1991; Strom and Lory, 1991; Zhang et al., 1994). Signal peptide removal requires a glycine residue at the -1 position of the prepilin, whereas N-methylation is dependent on residue Glu5 (Pasloske and Paranchych, 1988; Macdonald et al., 1993; Aas et al., 2007). Methylation is not necessary for pilus assembly as a P. aeruginosa prepilin peptidase mutant that can cleave but is incapable of pilin methylation is nonetheless able to produce functional pili (Pepe and Lory, 1998). In contrast, residue Glu5 is required for pilus assembly as some Glu5 point mutants are unable to produce pili (Pasloske and Paranchych, 1988; Strom and Lory, 1991; Horiuchi and Komano, 1998; Aas et al., 2007). Type IV pilus biogenesis is a complex process that requires eight or more proteins, many of which are conserved in different bacterial species, as well as in the related
T2SS (Figure 1.14) (Sandkvist, 2001; Ayers et al., 2010). The pilus assembly process is poorly understood.

A pseudo-atomic resolution structure of the *N. gonorrhoeae* GC pilus (Figure 1.10), belonging to the type IVa pilus subclass, was derived by computationally docking the full-length PilE pilin crystal structure into a 12.5 Å resolution cryo-electron microscopy (cryoEM) pilus reconstruction (Figure 1.9). This structure shows the pilin subunits arranged in a helical manner, held together mainly by hydrophobic interactions among the buried N-terminal α-helices (Craig et al., 2006). Each subunit is related to the next by a 10.5 Å axial rise and a 100.8° azimuthal rotation, following the path of a right-handed 1-start helix. This arrangement positions each conserved Glu5 residue within salt-bridging distance of the positively charged N-terminal amine of a neighbouring subunit to neutralize these two charges in the otherwise hydrophobic environment of the filament core. Although no type IVb pilus structure has been determined to date, computational *V. cholerae* TCP models have been proposed based on electron microscopy-derived symmetry parameters, packing of the pilin subunit, TcpA, in the crystal lattice, DXMS results identifying exposed and buried regions of the subunits and site-directed mutagenesis data (Chapter 2; Craig et al., 2003; Li et al., 2008; Lim et al., 2010). The arrangement of the pilin subunits within these TCP filament models is similar to that for the GC pilus, where the N-terminal α-helices form a stable helical array in the core of the filament and the protruding globular domains are exposed on its surface. Although TCP and GC
pili have a dominant 3-start helix, the helical symmetry of the TCP models is markedly different from that of the GC pilus: TcpA subunits are related by a 7.5 Å rise and a 140° rotation in a left-handed 1-start helix. If pilus filaments are assembled by sequential addition of each subunit, the opposite handedness for the primitive 1-start helices for GC pilus and TCP suggests distinct assembly mechanisms for these pili. To better understand the type IV pilus structure-function relationships and their assembly mechanism, I determined a TEM reconstruction and a pseudo-atomic resolution structure of the *V. cholerae* TCP.

### 3.2. Materials and Methods

#### 3.2.1. Bacterial strains and materials

Bacterial strains used in this study are listed in Table 3.1. The TcpA-6 polyclonal rabbit antibody specific for TcpA residues 174-199 (Sun *et al.*, 1990) was a gift from Ronald Taylor (Dartmouth Medical School). Concentrations of antibiotics used were as follows: ampicillin (Amp), 100 μg/ml; kanamycin (Kn), 45 μg/ml; streptomycin, (Sm), 100 μg/ml. Isopropyl β-D-1 thiogalactopyranoside (IPTG) was used at a final concentration of 0.4 mM for pilus induction.
Table 3.1  Bacterial strains used in this study

<table>
<thead>
<tr>
<th>Bacteria</th>
<th>Description</th>
<th>Reference/Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> S17</td>
<td>λ.pir</td>
<td>(Skorupski and Taylor, 1996)</td>
</tr>
<tr>
<td><em>V. cholerae</em> O395</td>
<td>Classical Ogawa, pMT5:toxT, SmR</td>
<td>(DiRita <em>et al.</em>, 1996)</td>
</tr>
<tr>
<td><em>V. cholerae</em> SJK70</td>
<td>O395, tcpA&lt;sup&gt;E5A&lt;/sup&gt;, pMT5:toxT</td>
<td>This study</td>
</tr>
<tr>
<td><em>V. cholerae</em> SJK70</td>
<td>O395, ΔtcpA</td>
<td>R. K. Taylor</td>
</tr>
</tbody>
</table>

3.2.2.  **TCP<sup>H181A</sup> expression and purification**

For TCP purification, *V. cholerae* classical strain RT4225 expressing TCP<sup>H181A</sup> was grown in 800 ml of liquid LB media under pilus-inducing conditions as described in section 2.2.2.1.

3.2.3.  **Transmission electron microscopy**

For the negative stain reconstruction, 5 μl of purified TCP (0.4 mg/ml) were applied to carbon-coated copper grids (Electron Microscopy Sciences) and stained with 1% phosphotungstic acid (PTA), pH 7. TCP were imaged on a FEI Tecnai F20 SUPERTWIN electron microscope at 120 keV, and images were collected on SO163 film (Kodak) at 51,000X magnification in low dose mode at the Simon Fraser University Nano-Imaging Facility. For cryoEM, 5 μl of purified TCP (1 mg/ml) was applied to glow-discharged C-flat holey carbon grids (Protochips Inc.) for five minutes. Grids were blotted for 3.5 seconds using a
Vitrobot Mark III Vitrification Robot (FEI) at 4 °C, 100% humidity and plunge-frozen in liquid ethane cooled with liquid nitrogen. Grids were stored in liquid nitrogen and transferred to a Gatan 626 cold stage for TEM imaging at 200 keV on a FEI Tecnai F20 TWIN electron microscope at the University of Calgary Microscopy and Imaging Facility. Micrographs were collected on SO163 film (Kodak) at 50,000X magnification in low dose mode.

3.2.4. Iterative helical real space reconstructions of the TCP filament

Electron micrographs of TCP were digitized at 2.5 Å/pixel using a Nikon Super Coolscan 8000 scanner and the iterative helical real space reconstruction (IHRSR) algorithm was used to generate the negative stain and cryoEM reconstructions (Egelman, 2000) (see Figure 3.1). For the negatively stained TCP images, straight filament segments of variable lengths were selected from the digitized images using the Helixboxer program from the EMAN package (Ludtke et al., 1999). To determine the helical symmetry of the filaments, 200 pixel x 200 pixel non-overlapping particles were generated from the TCP segments and an averaged power spectrum, a diffraction pattern of TCP filaments containing amplitude but no phase information, was determined from 976 particles. A minimum of two sets of layer lines is required to determine the rise and rotation per subunit of a 1-start helix. These approximate helical symmetry parameters are used in the IHRSR algorithm. The power spectrum in Figure 3.2A shows a strong layer line at ~1/(43 Å) corresponding to a 3-start helix and a weaker layer line at ~1/(110 Å) corresponding to a 4-start or 5-start helix.
Initial helical symmetry parameters for the primitive 1-start helix were determined by assigning a Bessel order \( n \) or “start” for each layer line (Stewart, 1988; Egelman, 2010). This resulted in four different combinations: \( n = -3, +4 \); \( n = -3, -4 \); \( n = -3, +5 \) and \( n = -3, -5 \). IHRSR runs were performed using these different starting symmetries with 8,034 overlapping particles (200 pixels in length with a 190-pixel overlap; 80 pixels in width, padded to 200 pixels). Only the starting symmetry corresponding to the -3-start, +4-start combination converged to a stable solution, having a rise of 8.4 Å and a rotation of 96.8° for the 1-start helix. To determine the resolution of the reconstruction, the TCP particles were randomly divided into two data sets each containing the same number of images and two volumes were generated. The resulting independently obtained reconstructions have a Fourier shell correlation 0.5 at 21 Å.
Figure 3.1 IHRSR algorithm cycle.

Note. The IHRSR procedure starts by incrementally rotating a solid cylinder and generating 60 reference projections. These projections are correlated with the raw filament segments to acquire an in-plane rotation angle, an x-shift, a y-shift, and an azimuthal angle. The aligned images from the TCP filaments, representing all views, are then used to produce a three-dimensional reconstruction by back projection, and helical symmetry is imposed. This new reconstruction is used as the reference volume for the subsequent cycle. This process is repeated iteratively until a stable structure is obtained. Egelman, 2000 (p. 227); adapted and used with permission.

For cryoEM, TCP filaments were flash-frozen on C-flat grids in vitreous ice in physiological buffer without heavy atom stains to mimic the native state of the protein (Figure 3.3A) and imaged. The contrast transfer function (CTF) was determined for each micrograph using the carbon film surrounding the hole, with the Bshow program in the Bsoft package (Heymann, 2001). Micrographs were corrected for phase reversals by multiplying by the CTF. Individual, straight TCP
segments were boxed using the Helixboxer software in EMAN (Ludtke et al., 1999), and a dataset containing 2,234 overlapping particles with a sampling size of 2.54 Å/pixel were generated (200 pixels in length with a 190-pixel overlap; 80 pixels in width, padded to 200 pixels). An averaged power spectrum was generated for all 2,234 particles as above. To generate three-dimensional reconstructions, particles were decimated to 5.08 Å/pixel, and IHRSR was run using the 2,234 overlapping particles and starting symmetries approximating that of the negative stain reconstruction. None of the IHRSR runs converged after 100 cycles, suggesting heterogeneity in the filaments. To reduce heterogeneity, the particles were sorted by pitch for the 3-start helix based on the observed layer line in the averaged power spectrum. Almost half of the particles had pitches in the range of 126 to 138 Å. One thousand TCP particles with pitches between 126 to 132 Å were used for IHRSR with a starting symmetry of an 8.2 Å rise and a 96.2° rotation. The IHRSR run converged to a final symmetry of 8.5 Å rise, 96.7° rotation for the cryoEM reconstruction. The final volume was corrected for the CTF using a Weiner filter (Frank and Penczek, 1995). The reconstruction has a resolution of 27 Å based on a Fourier shell correlation of 0.5.

Since a relatively small dataset was used for the TCP cryoEM reconstruction above, a larger set of images was collected and processed as described above in an attempt to obtain a higher resolution structure. In this case a total of 24,816 overlapping TCP particles were selected (200 pixels in length with a 190-pixel overlap; 80 pixels in width, padded to 200 pixels), and IHRSR
was run using starting symmetries approximating that of the negative stain reconstruction. None of the IHRSR runs converged after 100 cycles as was observed for IHRSR runs using smaller datasets. The particles were then sorted according to the 3-start helical pitch as described above. Again, almost half of the particles had a 3-start helical pitch in the range of 126 Å to 138 Å. Various combinations of these particles were used to generate three dimensional reconstructions using IHRSR: 10,415 particles with a pitch of 126 Å to 138 Å; 7,189 particles with a pitch of 126 Å to 132 Å; and 3,780 particles with a pitch of 132 Å. Each reconstruction converged to a symmetry very close to that of the negative stain reconstruction, but no improvement was achieved in terms of resolution or molecular detail.

3.2.5. Pseudo-atomic resolution structure of TCP

The 1.3 Å crystal structure of the soluble, N-terminally truncated TcpA pilin (Craig et al., 2003) was manually fit into the moderate resolution negative stain reconstruction using Chimera (http://www.cgl.ucsf.edu/chimera), to generate a pseudo-atomic resolution TCP model. The N-terminal 28 residues (α1N) were modeled onto the truncated structure using the coordinates for full-length P. aeruginosa PAK pilin (Craig et al., 2003). The pseudo-atomic resolution TCP structure was created by applying the helical symmetry parameters from the negative stain reconstruction, 8.4 Å rise and 96.8° rotation, to the docked subunit. While the overall orientation of the pilin subunit was clear, its precise position was not well-defined within the TEM volume due to the limited resolution
of the reconstruction. Thus, the fit was adjusted and the models were generated iteratively until a desirable model was obtained, which fit the following criteria: (i) was free of major steric clashes; (ii) was consistent with published structural data (close contacts between the N-terminal amine and Glu5 in neighbouring subunits and between Arg26 and Glu83 (Li et al., 2008)); and (iii) resembled the negative stain reconstruction when filtered to 21 Å resolution.

3.2.6. **Scanning transmission electron microscopy analysis**

Purified TCP\textsuperscript{H181A} were freeze-dried at the Brookhaven National Laboratory STEM facility. Digital dark-field images were collected and the data were analyzed using the PCMass program. TCP data were normalized to tobacco mosaic virus using a mass per unit length of 13.1 kDa/Å.

3.2.7. **Generation of *V. cholerae tcpA\textsuperscript{E5A} mutant**

The TcpA\textsuperscript{E5A} mutant was generated by undergraduate student Sumaiya Islam using the allelic exchange procedure (Lim et al., 2010).

3.2.8. **Analysis of pilin and pilus expression and pilus function in TcpA mutants**

After overnight growth under pilus-inducing conditions, TcpA expression was assessed by SDS-PAGE and immunoblot analysis of whole cell cultures with polyclonal antibody TcpA-6 (Sun et al., 1990). Pilus assembly was assessed by shearing the pili from the surfaces of the cells to release the filaments into the culture supernatant and by TEM of cell cultures grown under pilus-inducing
conditions. Autoagglutination and phage transduction assays were also completed to assess pilus assembly and function. Two-ml *V. cholerae* cultures were allowed to settle for 10 minutes at room temperature and then scored visually for autoagglutination. For phage transduction assays, overnight cultures of *V. cholerae* strains were incubated for 30 minutes at room temperature with equal volumes of filtered supernatant from overnight cultures of *V. cholerae* CL101 cells, which produce CTX-Kn phage (Waldor and Mekalanos, 1996; Kim et al., 2000). Cells were serially diluted and plated on LB-Kn agar plates, grown overnight at 37 °C and colony-forming units were counted. In parallel, *V. cholerae* strains were incubated with equal volumes of sterile media, serially diluted and grown on LB-Sm plates to determine the quantity of input bacteria and transduction frequency. Transduction efficiency was reported as a ratio of the transduction frequencies of the test strain and the wild type O395 strain.

### 3.3. Results

#### 3.3.1. Negative stain reconstruction of TCP

To obtain an atomic-level understanding of TCP structure, a negative stain reconstruction of TCP from *V. cholerae* strain RT4225 was determined using the IHRSR method (Egelman, 2000). After purification and dialysis in PBS/EDTA, these pili bundle and self-associate to a lesser extent compared to wild type TCP. The yield and purity of these filaments are also superior to wild type TCP making them more amenable for structural studies. Negatively stained TCP were
imaged by TEM (Figure 3.2A), micrographs were digitized and segments of filaments were selected for analysis. An averaged power spectrum was determined for ~1000 non-overlapping TCP particles. A dominant layer line was observed at ~1/(43 Å) corresponding to a 3-start helix and a weaker layer line at ~1/(110 Å) corresponding to a 4-start or 5-start helix (Figure 3.2A). The approximate helical symmetry (i.e. axial rise and azimuthal rotation per subunit in the 1-start helix) was determined for each possible combination of helical starts, and IHRSR was performed on 8,034 overlapping particles using these starting symmetries. This resulted in a single solution with a rise of 8.4 Å and a rotation of 96.8° for the 1-start helix. The TCP reconstruction has 3.7 subunits per helical turn, giving a mass per unit length of 2.4 kDa/Å, based on the 20.3 kDa TcpA subunit. This value was also determined independently by scanning transmission electron microscopy (STEM, Figure 3.2B). The reconstruction has a right-handed 1-start helix (+1), a left-handed 3-start helix (-3) as well as a right-handed 4-start helix (+4) shown in Figure 3.2C, although the mirror image (-1, +3, -4) was equally plausible from the IHRSR results. The resolution of the TCP reconstruction is approximately 21 Å based on Fourier shell correlation analysis.
3.3.2. CryoEM reconstruction of TCP

To improve the resolution of the pilus structure, a cryoEM reconstruction of TCP was determined. An averaged power spectrum from 2,243 overlapping particles showed a single layer line at \( \sim 1/(43 \, \text{Å}) \) corresponding to a 3-start helix with a pitch of 129 Å (Figure 3.3A). The IHRSR algorithm (Egelman, 2000) was applied to this relatively small data set using a starting symmetry of 8.4 Å rise and 96.8° rotation, obtained from the negative stain reconstruction, but no stable convergence was observed over 100 IHRSR cycles, suggesting that the TCP filaments may be heterogeneous in structure, as has been seen for other pili (Trachtenberg et al., 2005; Wang et al., 2006; Galkin et al., 2009). Thus, the particles were sorted based on the pitch of the 3-start helix (108 Å to 156 Å) and a smaller data set containing 1,000 particles with 3-start pitches of 126 Å to 132 Å.
Å was used for IHRSR, with a starting symmetry of 8.2 Å rise and 96.2° rotation. This cryoEM reconstruction converged to a stable solution at symmetry values of 8.5 Å rise, 96.7° rotation, which matches the helical symmetry of the negative stain reconstruction (Figure 3.3B and Figure 3.5). However, this cryoEM reconstruction had a lower resolution of 27 Å compared to the negative stain reconstruction as determined by Fourier shell correlation at 0.5 and did not provide additional molecular detail. In an attempt to improve the quality of the cryoEM reconstruction, many more TCP images were collected and digitized and IHRSR was performed on 24,816 particles, but none of these runs converged to a stable solution. Next, particles were sorted according to the pitch of the 3-start helix as described above, and several new reconstructions were generated using particles with the following 3-start helical pitches: 126 Å to 138 Å and 126 Å to 132 Å. Although some of these reconstructions used a 10-fold increase in the number of particles used, the reconstructions were no better than the original 1,000-particle dataset cryoEM reconstruction. Presumably, heterogeneity in the TCP filaments evident in the frozen-hydrated filaments but not in the negatively stained filaments limits the cryoEM reconstruction resolution.
3.3.3. Pseudo-atomic resolution structure of TCP

Both the negative stain and the cryoEM TCP reconstructions have an undulating exterior surface with a solid core devoid of a channel, with defined density for the pilin globular domains (Figure 3.2C, Figure 3.3B and Figure 3.5A). To obtain a pseudo-atomic resolution structure of the TCP filament, the high-resolution ΔN-TcpA crystal structure, which lacks the protruding N-terminal 28 residues (α1N) (Craig et al., 2003), was manually fitted into the moderate-resolution negative stain TCP reconstruction. Full-length TcpA was
superimposed on the docked ΔN-TcpA (Figure 3.4A), and the pseudo-atomic resolution TCP filament model was generated by applying the helical symmetry parameters for the negative stain TEM reconstruction to the single TcpA subunit (Figure 3.4B and C). Although this 21 Å resolution reconstruction does not reveal secondary structural elements that would accurately facilitate the fitting process, fitting was guided by the asymmetry of the globular domain. The globular domain of TcpA fit well into the right-handed 1-start reconstruction but poorly into its mirror image, thus defining the reconstruction as a right-handed 1-start helix with subunits related to another by an 8.4 Å axial rise and a 96.8° azimuthal rotation. Notably, this helical symmetry more closely matches that of the \textit{N. gonorrhoeae} GC pilus, which has a right-handed 1-start helix with a 10.5 Å rise and a 100.8° rotation (Craig \textit{et al.}, 2006), compared to previously-published computational TCP models (Craig \textit{et al.}, 2003; Li \textit{et al.}, 2008; Campos \textit{et al.}, 2011).
Figure 3.4  Pseudo-atomic resolution structure of the *V. cholerae* TCP filament.

**Note.**  
(A) Side and end views of a single, full-length TcpA pilin subunit docked into the negative stain TEM reconstruction. The relationship between the pilin subunits in the right-handed 1-start helix is indicated.  
(B) Side and end views of the TCP filament model shown as a ribbon representation, with the left-handed 3-start helical strands coloured in red, yellow and blue.  
(C) Side and end views of TCP illustrating the surface exposure of the D-region (magenta) and the involvement of the αβ-loop (green) in subunit:subunit interactions. Met1, Glu5, Arg26 and Glu83 are shown in stick representation.  
(D) Close-up of the stabilizing salt bridge between Arg26 on α1N of one subunit and Glu83 in the globular domain of a neighbouring subunit.  
(E) A segment of α1N is exposed between gaps formed by loose packing of the globular domains.  
(F) Close up of the N-termini showing the potential charge neutralization between Glu5 on one subunit and the N-terminal amine on an adjacent subunit in the 1-start helix.

The pseudo-atomic resolution TCP model fits well into the reconstruction volume, with the N-terminal α-helices of TcpA filling the core of the filament.
without steric clashes. Filament models with channels or hollow lumens were considered but eliminated due to diameter constraints and steric hindrance. Furthermore, most of α1 were poorly deuterated when analyzed by DXMS (Li et al., 2008), suggesting that the filament core is not solvent-exposed. In the TCP model, α1 of TcpA has the same S-shaped curve seen for P. aeruginosa PAK, N. gonorrhoeae GC and D. nodosus FimA pilin, which have kinks at Pro22 and Gly42 (Parge et al., 1995; Craig et al., 2003; Craig et al., 2006; Hartung et al., 2011). Although TcpA does not have helix-disrupting residues at these positions, it has glycines at residues 11, 14 and 19 and is hence expected to be curved or flexible. A number of TCP filament models were also generated using an idealized straight α-helix for α1N, but these models produced significant steric clashes and were ruled out. Thus, the curvature in α1 appears to enable packing in the pilus filament. The pseudo-atomic resolution TCP filament model was filtered to 21 Å resolution for comparison with the TEM reconstructions. The model resembles both the negative stain and the cryoEM reconstructions (Figure 3.5A), although its accuracy is limited by the moderate resolution of the reconstructions. Contour plot cross-sections of the 21 Å-filtered pseudo-atomic resolution TCP model were compared with those of the negative stain and the cryoEM reconstructions (Figure 3.5B). These contour plots show good correspondence of the TcpA subunit positions and connectivity between the TCP filament model and the reconstructions. Moreover, the power spectrum of the model (Figure 3.5C) matches the averaged power spectrum for the negatively
stained TCP particles (Figure 3.2A), both having equivalent 1/(43 Å) and 1/(110 Å) layer lines, providing additional validation for the model.

Figure 3.5  Comparison of the TCP filament model derived from the negative stain reconstruction with the negative stain and cryoEM reconstructions.

Note.  (A) Negative stain and cryoEM reconstructions compared with the pseudo-atomic resolution TCP filament model filtered to 21 Å resolution.  (B) Cross-sections of contour plots for the TCP reconstructions and filtered TCP model.  A distance of 5 Å along the filament axis separates each section.  (C) Power spectrum of the TCP filament model (58 subunits in length) showing layer lines at 1/(43 Å) for n=-3 and 1/(110 Å) for n=+4.
In this new TCP structure, the pilin subunits are arranged in a helical array with the N-terminal α-helices, α1, twisting around the filament axis to form a solid hydrophobic core (Figure 3.4). The interactions among the N-terminal α-helices hold the subunits together while the globular domains are loosely packed on the filament surface. Although the helical symmetry of the TCP reconstruction is substantially different from previously-published computational models for TCP (Craig et al., 2003; Li et al., 2008; Campos et al., 2011), the architecture and subunit:subunit interaction interfaces are very comparable. The TcpA pilin packing and orientation in the new TCP structure is consistent with the DXMS data, which mapped exposed and buried regions of the pilin subunits within the filament. In both models, the solid filament core is filled by interacting N-terminal α-helices and the globular domains form the outer surface of the filament. The D-regions (coloured magenta in Figure 3.4C), containing residues implicated in TCP-mediated microcolony formation (Kirn et al., 2000; Lim et al., 2010), protrude from the filament surface, and the αβ-loops (coloured green in Figure 3.4C) are oriented inward and interact with α1N of a neighbouring subunit. Residue Glu83 in the αβ-loop is positioned to form a salt bridge with Arg26 of α1N in the neighbouring subunit in the 3-start helix, as proposed for the DXMS-refined TCP model and supported by site-directed mutational analysis (Li et al., 2008) (Figure 3.4C and D). Grooves and cavities formed between the globular domains expose residues 14 to 29 in α1N, which was well-labeled in the DXMS study (Figure 3.4E). The TCP TEM reconstruction positions the conserved Glu5 residue of the N-terminal α-helix within salt-bridging distance of the positively-
charged N-terminal amine of Met1 of a neighbouring subunit in the 1-start helix, as observed with the DXMS-based TCP model (Figure 3.4F).

3.3.4. Residue Glu5 is necessary for efficient TCP filament assembly

To test the requirement of Glu5 in TCP filament biogenesis, the tcpA gene in V. cholerae was mutated, resulting in a Glu5Ala substitution. TcpA<sup>E5A</sup> pilin expression, pilus assembly and pilus function were assessed (Figure 3.6). The molecular weight and total pilin expression in overnight cultures was comparable for V. cholerae-tcp<sup>A<sub>E5A</sub></sup> and the wild type parental strain O395 indicating that TcpA<sup>E5A</sup> is produced and correctly processed by the prepilin peptidase, TcpJ. TcpA<sup>E5A</sup> has a slightly higher electrophoretic mobility than TcpA<sup>WT</sup>, which may be caused by the Glu5Ala substitution and loss of the N-methylation at Met1 combined, as was observed for a Glu5Val change for P. aeruginosa PAK pilin (Strom and Lory, 1991). After growth under pilus-inducing conditions, TcpA<sup>E5A</sup> levels in the sheared cell supernatant were greatly reduced relative to TcpA<sup>WT</sup> indicative of inefficient pilus assembly (Figure 3.6A). No Tcp<sup>E5A</sup> were observed by TEM and V. cholerae-tcp<sup>A<sub>E5A</sub></sup> cells autoagglutinated poorly relative to the wild type strain expressing TcpA<sup>WT</sup> (Figure 3.6B). Moreover, V. cholerae-tcp<sup>A<sub>E5A</sub></sup> CTX<sub>ψ</sub> transduction was 50-fold lower than wild type O395, but higher than a ΔtcpA deletion mutant. Thus, TCP is produced at very low levels by V. cholerae-tcp<sup>A<sub>E5A</sub></sup>, suggesting that Glu5 is required for efficient TCP assembly, as has been
shown for other type IV pili (Pasloske and Paranchych, 1988; Strom and Lory, 1991; Horiuchi and Komano, 1998; Aas et al., 2007).

**Figure 3.6 Effect of Glu5Ala substitution on TcpA expression and TCP assembly.**

Note. 

(A) Immunoblot of TcpA from wild type *V. cholerae* O395 and tcpA mutants showing total TcpA in whole cell cultures and TcpA pilin sheared from assembled pili. MWM, molecular weight markers; kDa units. The 37 kDa Coomassie-stained OmpU protein is shown as a loading control in the bottom panel. (B) Results of TEM analysis, autoagglutination and phage transduction assays to assess the presence of functional pili in *V. cholerae* O395 and tcpA mutant strains.

3.4. Discussion

TEM reconstructions of *V. cholerae* TCP and a pseudo-atomic resolution filament model were generated. The pilin subunit orientation and contacts in this new TCP model are similar to those of previously-published computational models (Craig et al., 2003; Li et al., 2008), but the helical symmetry is very different, and provides a better match with that of the GC pili (Craig et al., 2006). The old models were generated based on crystallographic packing of His-ΔN-
TcpA, TCP filament dimensions and knowledge that TCP had a dominant 3-start helix with a pitch of approximately 135 Å (Craig et al., 2003; Li et al., 2008; Campos et al., 2011). The positions of the subunits along these 3-start helical strands were modelled based on these data. Filament models were generated computationally and evaluated based on a diameter of approximately 80 Å, a 3-start helical pitch of approximately 135 Å and key subunit:subunit interaction interfaces determined by DXMS. The TCP TEM density, however, clearly defines the positions of the subunits within the filament, although their precise orientation must be inferred. The TCP reconstruction and resulting pseudo-atomic resolution filament model demonstrate the importance of acquiring accurate helical symmetry parameters: axial rise and azimuthal rotation of subunits in the 1-start helix and its hand. Incorrect symmetry values can lead to false filament models that nonetheless agree with existing genetic, immunological and biophysical data, as illustrated by earlier TCP models (Craig et al., 2003; Li et al., 2008; Campos et al., 2011). Obtaining a higher resolution structure would improve the accuracy of the TCP filament model, but filament heterogeneity makes this a challenging task.

The newly-determined TCP symmetry and architecture are comparable to that of GC pili (Craig et al., 2006) (Figure 3.7). In both filaments, extensive interactions between pilin subunits occur among α1 segments, between globular domains of neighbouring subunits, and between globular domains and α1N segments throughout the length of the filament. Each pilin subunit interacts with
six neighbouring subunits. These complex contact interfaces have not been accurately determined for either pilus subtype as the moderate resolution reconstructions are not sufficient to accurately fit the N-terminal α-helices and the full-length TcpA pilin structure is unavailable. Nonetheless, both pilins fit well as rigid bodies into their respective TEM densities and these pseudo-atomic resolution pilus filament models provide a strong basis for understanding pilin:pilin interactions, pilus structure and assembly.

The TCP and GC pilus structures have comparable helical symmetries, both are right-handed 1-start helices with a rise per subunit of ~8-10 Å and a rotation of ~100° (Figure 3.7). In both structures, the conserved Glu5 residue is positioned to form a salt bridge with the positively-charged N-terminal amine of a neighbouring subunit in the 1-start helix in both pilus filament models. Site-directed mutagenesis data from this study show that this conserved Glu5 is necessary for efficient TCP assembly, as has been established for several other type IV pili (Pasloske and Paranchych, 1988; Macdonald et al., 1993; Horiuchi and Komano, 1998; Aas et al., 2007). These findings support a common type IV pilus filament assembly mechanism, where pilin subunits are added sequentially along the right-handed 1-start helix utilizing a conserved inner membrane assembly machinery. We propose that residue Glu5 not only neutralizes the positively-charged N-terminus in the otherwise hydrophobic core of the pilus, but also is a driving force for filament assembly (Figure 1.12). Pilin subunits are embedded in the inner membrane reservoir by their extended α1N. The
electrostatic attraction between the Glu5 residue of one subunit and the N-terminus of the terminal subunit in the growing filament drives the subunit to dock at the base of the filament (Craig et al., 2006; Craig and Li, 2008). Moreover, complementarity among the globular domains also mediates the translocation of pilins from the hydrophobic inner membrane into the pilus filament. Some functional pili are produced by the \textit{V. cholerae}\textsuperscript{tcpA\textsuperscript{E5A}} mutant, as they are able to autoagglutinate and transduce phage at low but detectable levels. Thus, Glu5 is essential for efficient TCP assembly. When Glu5 pilin variants are co-expressed heterologously along with wild type pilin, both \textit{P. aeruginosa} and \textit{N. gonorrhoeae} Glu5 and wild type pilin incorporate into mixed pilus filaments (Pasloske et al., 1989; Aas et al., 2007). However, the hybrid \textit{P. aeruginosa} pili had an unusual morphology, and the assembly of pili on the \textit{N. gonorrhoeae} strain required the retraction ATPase, PilT, to be disrupted (Aas et al., 2007). These data suggest that pilus assembly can occur without the Glu5-N1 interaction but is inefficient and the resulting pili may be abnormal. Interestingly, the related T2SS pseudopilins also have a glutamate at residue 5, and the crystal structure of a ternary complex between the globular domains of the GspI:GspJ:GspK minor pseudopilins from ETEC has quasi-helical symmetry (Korotkov and Hol, 2008). The GspI, GspJ and GspK subunits are arranged in a right-handed helix with an \(\sim 100^\circ\) rotation about the filament axis, comparable to GC pili and TCP. Furthermore, when the N-terminal \(\alpha\)-helices are modeled onto the globular domains in the complex, they are staggered by \(\sim 10\) Å and the Glu5 residues of GspI and GspJ are within 5-7 Å of the N-terminal amine of their
neighbouring subunits, GspK and Gspl, respectively. These findings support a common architecture and assembly mechanism for type IV pili and the related T2SS pseudopili.

Figure 3.7  V. cholerae and N. gonorrhoeae pseudo-atomic resolution structures docked into their respective TEM densities.

Note.  Both filaments are dominant left-handed 3-start helices. Helical symmetries for the right-handed primitive, 1-start helices are indicated.
4. Stability comparison of type IVa and type IVb pilin and pili


Contributions to research: TcpA pilin, TCP and GC pilus protein expression and purification, circular dichroism, protease digestion, electron microscopy and co-wrote manuscript.

4.1. Introduction

Type IV pili are multifunctional filaments essential for bacterial pathogenesis. In *V. cholerae*, the toxin-coregulated pili mediate microcolony formation, phage attachment, and colonization factor secretion to cause the disease cholera (Taylor *et al.*, 1987; Waldor and Mekalanos, 1996; Kirn *et al.*, 2000; Kirn and Taylor, 2005; Lim *et al.*, 2010). The gonococcal or GC pili play a fundamental role in *N. gonorrhoeae* colonization of the urogenital tract causing the sexually transmitted disease gonorrhea. GC pili facilitate microcolony formation and also attach directly to host cell receptors (Punsalang and Sawyer, 1973; Swanson, 1973; Park *et al.*, 2001; Higashi *et al.*, 2007). These filaments retract and thus the bacteria are capable of flagella-independent twitching motility and DNA uptake (Wolfgang *et al.*, 1998; Wolfgang *et al.*, 1998; Merz *et al.*, 2000). Many other bacterial pathogens, including *N. meningitidis, P. aeruginosa,*
*H. influenzae*, S. Typhi, ETEC and EPEC, use type IV pili for colonization and invasion of human hosts.

Type IV pilins are divided into two subgroups, type IVa and type IVb, based on amino acid sequence and length, and the N-methylated N-terminal residue (Strom and Lory, 1993; Craig *et al.*, 2004). The signal peptide, the mature protein and the D-region, bound by conserved disulfide-bonded cysteines, are longer in type IVb pilins. The first residue of the mature pilin protein is a phenylalanine in the type IVa pilins, but this residue varies in the type IVb pilins. The type IVa pilins are more highly conserved throughout their amino acid sequence, and share minimal sequence similarity to the type IVb pilins beyond the N-terminal ~25 residues. Genes encoding the type IVa pilus assembly components are scattered throughout the genome whereas the type IVb pilus assembly genes are generally clustered in a single operon (Pellicic, 2008). The type IVa pilins are found on a broad range of bacterial pathogens, including *N. meningitidis* and *N. gonorrhoeae*, which infect the brain and urogenital tract, respectively, and *P. aeruginosa*, which colonizes the lungs. On the other hand, type IVb pilins are found almost exclusively on enteric pathogens including *V. cholerae*, ETEC, EPEC, S. Typhi and *Citrobacter rodentium*.

Structural analysis revealed distinct folds for the type IVa and IVb pilin subtypes. Full length type IVa pilin crystal structures, PilE from *N. gonorrhoeae* (Parge *et al.*, 1995; Craig *et al.*, 2006), PAK pilin from *P. aeruginosa* (Craig *et al.*, 2003) and FimA from *D. nodosus* (Hartung *et al.*, 2011) show similar
architectures: a curved extended 53 residue N-terminal α-helix, α1, the second half of which is embedded in the C-terminal globular domain containing a central anti-parallel 4-stranded β-sheet. Structural differences among these type IVa pilins occur in the αβ-loop, which lies between α1 and the β-sheet, and in the D-region (Figure 1.6A and B). However, in all three of these type IVa pilin structures, the C-terminal loop, which follows the β-sheet and includes part of the D-region, lies at the periphery of the globular domain. This loop lies like a thumb at the edge of a closed fist, stabilized by non-covalent interactions as well as the disulfide bond to the β-sheet (Figure 1.7A). This topology is also evident in structures of recombinantly expressed N-terminally truncated type IVa pilins (Hazes et al., 2000; Audette et al., 2004; Nguyen et al., 2010) as well as the minor pilin PilX from N. meningitidis (Helaine et al., 2007). No full length structures of type IVb pilins have been solved, but several N-terminally truncated structures reveal an architecture similar to that of the type IVa pilins, with an extended N-terminal α-helix, α1C embedded in a globular domain containing a central anti-parallel β-sheet and a stabilizing disulfide bond (Craig et al., 2003; Xu et al., 2004; Ramboarina et al., 2005; Lim et al., 2010). Yet, the topology of the β-sheet of the type IVb pilins is distinctly different from type IVa pilins, having non-nearest neighbour connectivity with the C-terminus forming the central β-strand of the β-sheet (Figure 1.6C and Figure 1.7B). As with type IVa pilins, a disulfide bond secures a peripheral loop in the globular domain of type IVb pilins. The significance of these distinct topologies is unclear but may reflect adaptations to niches occupied by the bacteria and to the demands and functions of the pili.
Type IVa GC pili have been studied extensively to elucidate pilus assembly and disassembly. Pilus biogenesis requires an assembly ATPase situated on the cytoplasmic side of the inner membrane. Some type IV pili, such as GC pili, possess a second "retraction ATPase" that assists in pilus depolymerization or disassembly (Wolfgang et al., 1998; Wolfgang et al., 1998; Merz et al., 2000). GC pilus retraction is essential for twitching motility, natural transformation, intimate adhesion of *N. gonorrhoeae* to host cells, and cortical plaque formation and NF-κb activation in host cells (Merz et al., 2000; Aas et al., 2002; Higashi et al., 2009; Dietrich et al., 2011). Optical laser tweezers studies demonstrated that single GC pilus filaments are capable of enduring forces in excess of 100 pN (Merz et al., 2000; Maier et al., 2002) and pilus bundles retract with even greater forces (Biais et al., 2008). Similarly, this impressive tensile strength is also evident for type IVa pili from *P. aeruginosa* and *M. xanthus* pathogens (Winther-Larsen et al., 2007; Clausen et al., 2009). Pilus retraction is less well-established in the type IVb pili. In EPEC, disruption of BfpF, a putative retraction ATPase, results in type IVb pilus overexpression, increased aggregation and adherence, and reduced infectivity (Anantha et al., 1998; Bieber et al., 1998; Zahavi et al., 2011). On the contrary, no pilus disassembly ATPase has been identified for *V. cholerae* TCP and retraction has not been observed. Further studies of the type IVb pili are needed to determine if the pilus architecture and remarkable mechanical features of the type IVa pili are shared by all type IV pili. To examine the significance of structural differences between the type IVa and type IVb pilins with respect to pilus stability, the *N. gonorrhoeae*
and *V. cholerae* pilin and pili were compared by circular dichroism (CD) and transmission electron microscopy (TEM). We predicted that the buried C-terminal segment imparts a more stable protein fold for the type IVb pilins compared to the type IVa pilins, which has a peripheral C-terminal loop. The C-terminal β-strand of type IVb pilins is buried within an anti-parallel β-sheet stabilized by an extensive network of hydrogen bonds and hydrophobic interactions whereas the less ordered C-terminal loops of type IVa pilins are exposed on the pilin surface. Results are discussed in the context of specific type IV pilus functions and the environmental niches occupied by these bacterial pathogens.

### 4.2. Materials and Methods

#### 4.2.1. Bacterial strains and materials

Bacterial strains and plasmids used in this study are listed in Table 4.1. The TcpA-6 polyclonal rabbit antibody specific for TcpA residues 174-199 (Sun *et al.*, 1990) was a gift from Ronald Taylor (Dartmouth Medical School). Katrina Forest (University of Wisconsin-Madison) provided purified ΔN-PilE protein for the CD studies (Hansen *et al.*, 2007). Concentrations of antibiotics used were as follows: ampicillin (Amp), 100 μg/ml; kanamycin (Kn), 45 μg/ml; streptomycin, (Sm), 100 μg/ml. Isopropyl β-D-1 thiogalactopyranoside (IPTG) was used at a final concentration of 0.4 mM for pilus induction.
Table 4.1  Plasmids and strains used in this study

<table>
<thead>
<tr>
<th>Bacteria</th>
<th>Description</th>
<th>Reference/Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>N. gonorrhoeae</em> MS11</td>
<td>Variant C30</td>
<td>(Segal et al., 1985)</td>
</tr>
<tr>
<td><em>V. cholerae</em> O395</td>
<td>Classical Ogawa, pMT5:toxT, SmR</td>
<td>(DiRita et al., 1996)</td>
</tr>
<tr>
<td><em>V. cholerae</em> RT4225</td>
<td>O395, tcpA&lt;sup&gt;H181A&lt;/sup&gt;, pMT5:toxT</td>
<td>(DiRita et al., 1996; Kim et al., 2000)</td>
</tr>
<tr>
<td><em>E. coli</em> Origami(DE3)-pET:15b-TcpA</td>
<td><em>E. coli</em> K12 gor-/trxB-, Δ1-28, hisTcpA</td>
<td>Novagen &amp; (Craig et al., 2003)</td>
</tr>
</tbody>
</table>

4.2.2. ΔN-TcpA pilin and TCP<sup>H181A</sup> expression and purification

ΔN-TcpA was over-expressed in *E. coli* Origami cells (Novagen) as described in section 2.2.3. The hexahistidine tag (His-tag)/linker was removed from purified ΔN-TcpA by treatment with 1 unit of thrombin protease (GE Healthcare) per 500 μg of substrate at 4 °C for 16 hours. The His-tag/linker was separated from the thrombin-digested pilin protein using a HisTrap HP column. The eluent was dialyzed in PBS in the presence of 10 mM EDTA. Protein was concentrated using an Amicon stirred cell concentrator (Millipore) with a YM10 membrane when necessary. TCP<sup>H181A</sup> filament expression and purification was performed as described in section 2.2.2.1. The His181Ala substitution in TcpA causes the pili to fall off the cells into the culture supernatant.

4.2.3. GC pilus expression and purification

GC pili were expressed and purified as described in (Craig et al., 2006). Briefly, *N. gonorrhoeae* strain MS11 variant C30 cells were grown on GC
medium base plates for 24 hours at 37 °C in the presence of 5% CO₂. Cells were scraped from the plates and resuspended in 50 mM CHES buffer, pH 9.5. The pili were sheared from the surface of the cells by vigorous vortexing and isolated by centrifugation at 10,000 Xg. The pili were resuspended in CHES buffer and purified by dialysis into pilus buffer (50 mM Tris, 150 mM NaCl, 0.02% NaN₃, 1 mM dithiothreitol, pH 7.5) followed by centrifugation. Purified pili were resuspended in PBS.

4.2.4. Far-UV circular dichroism

CD experiments were conducted using a 1 mm path length quartz cuvette (Hellma) on a Jasco J-810 spectropolarimeter. Pilus samples were dialyzed exhaustively in salt-free phosphate buffer and diluted to 0.3 mg/ml for CD analysis. For urea denaturation, the pili were treated with 0 M, 2 M, 4 M, 6 M or 8 M urea for 1 hour at 22 °C prior to CD analysis. Secondary structure characteristics and folding properties of proteins can be determined by far-UV CD spectroscopy. α-helix, β-sheet and random coil secondary structural elements have characteristic CD spectra (see Figure 4.1). For the thermal denaturation studies, CD measurements were collected at the specified wavelengths while the pilin and pilus sample temperature was increased from 20 °C to 80 °C at a rate of 0.5 °C per minute. Octyl-β-D-glucopyranoside (βOG) detergent was added to a final concentration of 50 mM to dissociate the pilus filaments into pilin subunits, and CD spectra were collected immediately and after 24 hours of stirring at 4 °C. Far-UV CD spectra were collected between 190 nm
and 260 nm at 22 °C, 40 °C, 60 °C and 80 °C. All CD data were collected in duplicate. The buffer background spectra were subtracted from all pilin and pilus spectra, and duplicate data were averaged. The Spectra Analysis program within the Spectra Manager package (Jasco) was used to convert ellipticity values in millidegrees to mean residue molar ellipticity, expressed in units of degree*cm²/decimole.

Figure 4.1 Characteristic far-UV CD spectra of poly-L-lysine in three different conformations: α-helix, β-sheet and random coil.

Note. Greenfield and Fasman, 1969 (p. 4110); adapted and used with permission.
4.2.5. Proteolytic digestion of GC pili and TCP and their pilin counterparts

TCP and GC pili were treated with the proteases chymotrypsin, elastase, proteinase K, subtilisin A, thermolysin or trypsin at a protease:protein mass ratio of 1 at 37 °C in a shaking incubator. At various time points, aliquots were collected and proteolysis was quenched by boiling in Laemmli sample buffer (Laemmli, 1970) containing 10 mM phenylmethylsulfonyl fluoride. Protease susceptibility was assessed by Coomassie-stained SDS-PAGE.

4.2.6. SDS-PAGE and immunoblot analysis

Aliquots of whole cell culture (20 μl for Coomassie-stained gel analysis, and 5 or 10 μl for immunoblot analysis) and purified TCP (0.2 μl or 10 mg) were boiled in Laemmli sample buffer and analyzed by 15% SDS-PAGE. Gels were either stained with Coomassie Blue or immunoblotted to transfer protein from the gels to polyvinylidene fluoride membrane and probed using polyclonal antibody TcpA-6 (Sun et al., 1990). Bound antibody was visualized using goat anti-rabbit antibody conjugated to alkaline phosphatase (Bio-Rad), with 5-bromo-4-chloro-3'-indolyl phosphate p-toluidine salt/nitro-blue tetrazolium chloride (BCIP/NBT, Bio-Rad) as a substrate.

4.2.7. Transmission electron microscopy

To image heat-treated pili, TCP and GC pili at 0.3 mg/ml were subjected to heat for 5 minutes at the specified temperatures. Five μl of these samples were then applied to carbon-coated copper grids (Electron Microscopy Sciences)
and stained with 1% PTA. Pili were imaged in a Hitachi 8000 STEM at 200 keV at the SFU Nano-Imaging Facility. Chemically-denatured pili were incubated in final concentrations of 0 M, 2 M, 4 M, 6 M or 8 M urea or 6 M guanidinium-HCl for 1 hour at 22 °C then applied to grids, stained and imaged as described above. To test the effect of disulfide bond reduction on pilus stability, TCP and GC pili were treated with 100 mM TCEP for 15 minutes at 22 °C before being applied to copper grids. TEM sampling bias was avoided by obtaining images representative of the entire sample.

4.3. Results

4.3.1. Type IVa and type IVb pilin globular domain stability comparison

The thermal stability of soluble *N. gonorrhoeae* PilE and *V. cholerae* TcpA as representatives of the type IVa and IVb pilins respectively, were compared to determine if the type IVb pilin fold imparts greater structural stability to the globular domain due to its C-terminus being buried in the β-sheet. Both pilins were N-terminally truncated lacking α1N: ΔN-PilE composed of residues 27-158 (Hansen *et al.*, 2007); and ΔN-TcpA composed of residues 29-199 (Craig *et al.*, 2003). First, a far-UV CD spectrum from 190 nm to 260 nm wavelength was collected at room temperature to obtain the CD profile for each protein and to determine a suitable wavelength at which to follow protein unfolding. The CD spectra are presented between 200 nm to 260 nm wavelength to eliminate data with significant spectral noise. The ΔN-TcpA has a CD spectrum typical of a
predominantly α-helical protein with negative peaks at 208 nm and 222 nm (Figure 4.1 and Figure 4.2A), whereas, the ΔN-PilE CD spectrum has an intense negative peak at 208 nm and a negative shoulder between 220 nm to 230 nm (Figure 4.2B). Pilin thermal denaturation was observed by CD spectroscopy at peak wavelength 222 nm for ΔN-TcpA and 208 nm for ΔN-PilE by heating the samples between 20 °C to 80 °C, the upper limit of our CD spectropolarimeter. Figure 4.2C shows that ΔN-TcpA unfolding occurs between ~47 °C and ~60 °C with a melting temperature (T_m) of 53.5 °C. On the contrary, ΔN-PilE unfolding immediately starts upon heat-treatment and continues to denature through 80 °C with no discrete T_m, indicating instability for this recombinant pilin (Figure 4.2D). Thus, as predicted, ΔN-TcpA is substantially more stable than ΔN-PilE, which is likely related to type IV pilin fold as both recombinant pilins lack the conserved, protruding N-terminal segment. This finding is supported by higher α-carbon B-factor values observed for the C-terminal loop segment in the PilE crystal structure (Craig et al., 2006), indicative of greater flexibility compared to the corresponding region in the TcpA crystal structures (Craig et al., 2003; Lim et al., 2010).
Figure 4.2 Thermal stability comparison for *V. cholerae* and *N. gonorrhoeae* type IV pilins and their respective pilus filaments.

Note. Far-UV CD spectra for (A) ΔN-TcpA pilin and TCP filaments and (B) ΔN-PilE pilin and GC pili. (C) Thermal denaturation of ΔN-TcpA and TCP and (D) ΔN-PilE and GC pili measured by CD spectroscopy at the specified wavelengths. The CD plots are averages of two independent experiments. (E) TEM images of negatively stained pili treated for 5 minutes at the temperatures indicated.
4.3.2. Type IVa and type IVb pilus thermal stability comparison

As GC pili have been demonstrated to possess remarkable tensile strength (Maier et al., 2002), their stability was compared to that of TCP using thermal denaturation and far-UV CD spectroscopy. Far-UV CD scans were performed for purified TCP and GC pili and compared to the CD spectra for the soluble pilin monomers. Again, the CD spectra are shown between 200 nm to 260 nm wavelength to eliminate low confidence data with significant spectral noise possibly due to light scattering of the long filaments. The TCP spectrum resembles that of the pilin monomer ΔN-TcpA (Figure 4.2A), though the intensity of the troughs at 208 nm and 222 nm are greater for the pili. This can be explained by the extended, 28-residue α-helix (α1N), which is absent in the soluble ΔN-TcpA pilin. Notably, the GC pilus spectrum is distinctly different from that of the ΔN-PilE monomer, with a much less intense negative trough at 208 nm and a second trough at 228 nm (Figure 4.2B). These data suggest that the GC pilin secondary structure differs in the soluble monomer and the filamentous forms.

Thermal denaturation of the two pilus subtypes, TCP and GC pili, was monitored at peak wavelengths of 222 nm and 228 nm respectively. Denaturation of TCP is highly cooperative, occurring between ~55 °C and ~64 °C with a T_m of 59.4 °C, indicating that the TCP filaments have greater stability than their ΔN-TcpA monomer counterparts (Figure 4.2C). On the other hand, no change was detected in the GC pilus CD spectrum between 20 °C and 80 °C at its peak
wavelength, 228 nm (Figure 4.2D), or at other wavelengths tested. This implies that these pilus filaments are stably folded over the duration of this heat-treatment. The stability of GC pili was further studied by obtaining CD scans after subjecting samples to temperatures between 22 °C and 80 °C for 5 minutes and 40 minutes incubation times. These data show that the characteristic GC pilus CD spectrum is observed at 22 °C, 40 °C and 60 °C but the 228 nm trough intensity is reduced at 80 °C for both incubation times (Figure 4.3), suggesting that protein unfolding occurs above 60 °C.
To determine the effects of heating on the morphology of both pilus types, the samples were examined by TEM. TCP filaments are abundant at 22 °C and 40 °C, although they look fragmented at 40 °C (Figure 4.2E), consistent with the CD results. No TCP were observed after treatment at 60 °C or 80 °C. In contrast,
heating GC pili to 60 °C appears to have no effect, and some filaments are even observed after incubation at 80 °C. Thus, the assembled GC pili are considerably more stable than TCP despite the ΔN-TcpA monomer being more stable than the ΔN-PilE monomer.

4.3.3. Type IVa and type IVb pilus polymerization and its effect on pilin secondary structure

To understand the effect of polymerization on the pilin secondary structure, both pilus samples were subjected to treatment with the non-ionic detergent βOG, which disrupts pilin:pilin interactions dissociating the filaments into pilin monomers. Importantly, this method does not unfold the pilin subunits, as it has been used for crystal structure determination of full-length pilins (Craig et al., 2003; Craig et al., 2006; Hartung et al., 2011). Far-UV CD scans were obtained for TCP and GC pili immediately after the addition of βOG (0 minutes) and after 24 hours of incubation with detergent at 4 °C (Figure 4.4). Results show that detergent treatment had an immediate effect on the CD spectra for both pilus samples and the spectra did not deviate considerably after 24 hours. Pilus filament dissociation was confirmed by TEM, which showed few pili in the 0 minute samples and no pili after 24 hours in βOG. Furthermore, βOG-treated pili filtered through an Amicon stirred cell (Millipore) with a 300 kDa molecular weight cut-off membrane, which normally retains untreated, assembled pili. The CD spectra for detergent-treated pili were overlaid on the spectra for untreated and heat-denatured pili (Figure 4.4, top panels) and soluble, N-terminally truncated
pilins (Figure 4.4, bottom panels). The negative intensity trough at 222 nm is observed for βOG-treated TCP as well as untreated TCP and ΔN-TcpA pilin subunits, indicating that all three samples share similar secondary structures. In contrast, the CD spectra for βOG-treated GC pili more closely resembles that of ΔN-PilE pilin rather than untreated GC pili which suggests a change in secondary structure for the PilE pilin subunit upon polymerization into pilus filaments. This is in agreement with the CD wavelength scans of untreated ΔN-PilE pilin and GC pili in Figure 4.2B, where disparity between 200 nm and 230 nm is evident.
Figure 4.4  Effect of βOG detergent treatment on pilin structure.

Note. Pili were treated with βOG at a final concentration of 50 mM to dissociate the filaments into pilin subunits and the CD spectra were obtained immediately (0 minutes) and after 24 hours. (A) CD spectra of βOG-treated TCP superimposed on the spectra of untreated or heat-denatured TCP (top graph) and ΔN-TcpA (bottom graph). (B) CD spectra of βOG-treated GC pili overlaid on the spectra of untreated or heat-denatured GC pili (top graph) and ΔN-PilE (bottom graph). CD data are averages of two independent experiments.

4.3.4. Type IVa and type IVb pilus chemical stability comparison

GC pilus and TCP filaments were mixed with the chemical denaturant urea to further compare the relative stabilities of the type IVa and type IVb pili. Pili were subjected to 0, 2 M, 4 M, 6 M and 8 M urea for 1 hour and then analyzed by CD spectroscopy and TEM. Figure 4.5 shows CD scans for urea-denatured pili between 200 nm and 260 nm wavelengths. The 222 nm intensity is reduced upon treatment of TCP with 2 M urea relative to the untreated sample and the distinctive α-helical CD spectrum is absent at urea concentrations of 4 M or greater, demonstrating protein unfolding (Figure 4.5A). TCP are observed by TEM after treatment with 2 M urea but no pili are seen at higher urea concentrations (Figure 4.5C), consistent with the CD data. In contrast, the secondary structure of GC pili appears unaffected by up to 6 M urea. Remarkably, the GC pili negative peak at 228 nm is present at ~50% of its original intensity at 8 M urea (Figure 4.5B). The abundant GC pili observed by TEM were morphologically indistinguishable from untreated pili for all urea concentrations tested (Figure 4.5C), which demonstrates the astonishing sturdiness of these filaments. Interestingly, CD and TEM analysis show that GC pili are completely denatured using 6 M guanidinium hydrochloride.
To ensure that the instability of the TCP filaments relative to GC pili was not due to the His181Ala substitution in TcpA, the susceptibility of TCP\textsuperscript{H181A} and wild type TCP to urea was compared. Wild type TCP were isolated from classical V. cholerae O395 liquid cultures by shearing and precipitation. While the yield and purity of the wild type TCP are inferior compared to TCP\textsuperscript{H181A}, an adequate amount was obtained for chemical denaturation and TEM analysis. Wild type and TCP\textsuperscript{H181A} were treated with various concentrations of urea and analyzed by TEM. Wild type TCP are stable in 2 M urea but are denatured in 4 M urea, similar to urea-treated TCP\textsuperscript{H181A}, indicative of comparable stabilities between these filaments (Figure 4.6). Moreover, TCP\textsuperscript{H181A} and wild type TCP are equally resistant to chymotrypsin proteolysis (Li et al., 2008) and both mediate phage uptake indicating that they are functional (Kirn et al., 2000). Therefore, TCP\textsuperscript{H181A} are suitable replacements for wild type TCP in structural studies.
Figure 4.5  Urea denaturation of TCP and GC pili.
Far-UV CD scans for (A) *V. cholerae* TCP and (B) *N. gonorrhoeae* GC pili after treatment with urea at the concentrations specified. CD plots are averages of two independent experiments. (C) Negatively-stained TEM images of urea-treated pili.

**Figure 4.6** TEM images of urea-denatured wild type *V. cholerae* TCP.

(A) Untreated wild type TCP. Wild type TCP treated with (B) 2 M urea and (C) 4 M urea for 1 hour. Scale bars, 500 nm.

### 4.3.5. Type IVa and type IVb pilus protease susceptibility comparison

Protease susceptibility was utilized to test the relative stabilities of TCP, GC pili, ΔN-TcpA and ΔN-PilE. Both pilins were degraded by all proteases tested within one hour of treatment at 37 °C using a high protease:pilin mass ratio of 1:1 (Figure 4.7). The pilus filaments are significantly more stable relative to the pilins as longer protease incubations were required. Little or no proteolysis was observed after a 1-hour treatment for some proteases indicating that the quaternary structure of the pilus protects the pilin subunits from degradation. The TCP filaments were more prone to degradation than GC pili for all but one of the proteases tested. As with the thermal and chemical denaturation studies, these proteolysis results imply that GC pili are considerably more stable than TCP.
4.3.6. Type IVa and type IVb pilus stability in reducing agent comparison

The Cys120-Cys186 disulfide bond is necessary for stable TcpA pilin expression and TCP assembly (Kirn et al., 2000; Li et al., 2008; Lim et al., 2010). To test the role of the conserved disulfide bond in pilus stability, reducing agent, TCEP, was added to TCP and GC pili at a final concentration of 100 mM. No pili were observed by TEM after a 5 minute treatment at 22 °C with TCEP, demonstrating the importance of the disulfide bond for filament stability. Presumably, disruption of this interaction destabilizes the pilin subunits and leads to unfolding and dissociation.
Figure 4.7  Type IVa and type IVb pilin and pilus protease susceptibility comparison.

Note.  
Coomassie-stained SDS-PAGE analysis of TCP, GC pili, ΔN-TcpA pilin and ΔN-PilE pilin digested with a panel of proteases at 37 °C for the times specified using a protein:protease mass ratio of 1:1. Molecular weights are 20.3 kDa for full-length TcpA pilin and TCP, 17.2 kDa for full-length PilE pilin and GC pili, 18.0 kDa for ΔN-TcpA and 15.0 kDa for ΔN-PilE pilin. Bands seen in ΔN-PilE sample digested with chymotrypsin at 1, 6 and 24 hour time points are chymotrypsin self-degradation products that resolve at ~14 kDa.
4.4. Discussion

This study reveals that two prototypes for the type IVa and type IVb pilus subclasses have marked differences in stability. The *V. cholerae* type IVb pilin ΔN-TcpA subunit is more stable than the *N. gonorrhoeae* type IVa pilin ΔN-PilE subunit, which was expected based on their globular domain topology differences. However, GC pilus filaments are more robust than TCP filaments, which cannot be attributed to the subunit stability. Both pilus samples are more stable than their soluble pilin counterparts as protein oligomers tend to be more robust than their monomeric equivalents (Ali and Imperiali, 2005).

This study shows that GC pili are considerably more resistant than TCP to thermal and chemical denaturation and to proteolysis despite the fact that they are thinner than TCP (~60 Å versus ~80 Å in diameter for GC pili and TCP, respectively). Since ΔN-PilE is less stable than ΔN-TcpA, the higher stability observed for GC pili relative to TCP is not due to subunit stability but instead arise from the quaternary assembly of the PilE subunits in GC pili. The pseudo-atomic resolution structure of TCP shows that the pilin α1 segments are not packed as tightly within the core of the filament as they are in GC pili (Figure 4.8A). The larger filament diameter and higher tilt angle of TcpA relative to the filament axis give TCP a larger diameter α-helical core than GC pili. Another feature contributing to pilus stability may be the packing of the subunits on the
surface of the filament. Although the TcpA globular domain is much larger than PilE (171 versus 100 residues), its bulky D-region protrudes from the filament surface, which increases the filament diameter instead of packing neighbouring subunits more tightly against one another (Figure 4.8A and B). TCP has a sizable cavity, measuring \( \sim 16 \) Å in width and \( \sim 25 \) Å in length, between neighbouring subunits that exposes a segment of \( \alpha 1N \) (Figure 4.8C). This is supported by DXMS labeling of \( \alpha 1N \) in the TCP filament (Li et al., 2008). GC pili, however, have a much narrower, \( \sim 5 \) to \( 8 \) Å cleft between pilin subunits due to the peripherally-located pilin C-terminus, which partially bridges the gap between the globular domains (Figure 4.8C). In TCP, the C-terminal loop of TcpA is embedded in the globular domain rather than filling the gaps between subunits leaving a substantial cavity. Thus, the distinct type IV pilin topologies may define pilus stability due to differences in D-region bulk distribution and the position of the C-terminal loop. Lastly, PilE has three aromatic residues, Phe1, Tyr24, Tyr27, that have side chains oriented into the core of the GC pilus and positioned to stack. The Phe1 side chain of one subunit is in close proximity to and positioned to insert between Tyr24 and Tyr27 of a neighbouring subunit in the GC pilus, which would contribute to the higher binding energy for pilin:pilin interactions (Figure 4.8D). These aromatic residues are conserved in the type IVa pilins but are not present in most type IVb pilins including \( V. \) cholerae TcpA (Figure 4.8E and Figure 4.9). Finally, we predicted that the unique topology of the type IVb pilins would make this subclass less dependent on the disulfide bond, because the hydrogen bonds within the \( \beta \)-sheet stabilizes the C-terminus. However, TCP
completely disintegrated after treatment with a reducing agent as did GC pili, establishing that the disulfide bond for both pilus types are essential for stability.

Figure 4.8 Tight packing of the pilin subunits and conserved α1N aromatic residues may stabilize GC pili.
Note.  
(A) Top view of TCP and GC pili showing differences in diameter and packing of the α-helices in the filament core. The globular domain is illustrated for a single subunit coloured in yellow. The conserved cysteines, Cys120 and Cys186 in TcpA, and Cys121 and Cys151 in PilE, are coloured cyan. The C-terminal segment beyond the second cysteine is coloured red. (B) Side view of TCP and GC pili with a single subunit coloured as in (A) showing differences in the globular domain mass distribution and position of the C-terminal segment. The C-terminal residue is indicated. (C) TCP and GC pili possess gaps between globular domains exposing a segment of α1N (circled) shown in surface representation. The cavities in TCP are ~16 Å wide and ~25 Å high, and in GC pili are ~7 Å wide by ~32 Å wide. (D) Left. Cross-section of the α1 segments in the core of the GC pilus filament showing the close proximity of aromatic residues Phe1, Tyr24 and Tyr27. Phe1 of one subunit (sticks coloured in yellow) and Tyr24 and Tyr27 of a neighbouring subunit (sticks coloured in cyan) may stack to increase the binding energy of the filament. Phe1 of the yellow subunit is the only residue visible in this view. Each subunit has a different colour, with oxygen atoms coloured in red. Right. Side view of the GC pilus illustrating the yellow and cyan pilins from the left panel, which are separated by an axial distance of 42 Å. (E) Left. Cross-section of the α1 segments in the core of the TCP filament showing residues Met1, Ala24 and Ala27. Each subunit is coloured differently, with oxygen atoms coloured in red. Right. Side view of the yellow and cyan pilins from the left panel of TCP, which are separated by 34 Å.

The abundant and tight interactions among PilE pilin subunits explain the robust stability of GC pili. The strength observed for GC pili may be a common feature of the type IVa pili, since GC, PAK, and M. xanthus pili have been shown to withstand high stress forces (Merz et al., 2000; Maier et al., 2002; Winther-Larsen et al., 2007). The requirement for high filament stability and tensile strength likely depends upon the biological demands of the pili. GC pili are necessary for N. gonorrhoeae adhesion to host cells in the urogenital tract, where the pathogens encounter high urine flow rates. Furthermore, GC pili retract, and their role in twitching motility (Higashi et al., 2007; Higashi et al., 2009) would also subject the filaments to high stress. In contrast, V. cholerae form TCP-mediated microcolonies in protected crypts in the lining of the small
intestine (Yamamoto and Yokota, 1988; Olivier et al., 2007) and TCP have not been shown to be retractile. Thus, these filaments are not exposed to the high flow forces seen by *N. gonorrhoeae*. Type IVb pilins and the related T2SS pseudopilins lack residues Tyr24 and Tyr27, which are conserved in the type IVa pilins. It is possible that these residues are only necessary for filaments that are subjected to mechanical stress. Thus, rather than adapting a more stable fold to provide mechanical stability, *N. gonorrhoeae* and other type IVa pili evolved residues that hold the subunits together more tightly. Enteric bacterial pathogens may express type IVb pili instead of type IVa pili simply because they acquired the genes horizontally from other gut bacteria or phage and these were sufficient to mediate colonization in the intestinal environment.
Figure 4.9 Amino acid sequence alignment illustrating conserved aromatic residues in type IVa pilins.

Note. Aromatic residues Phe1, Tyr24 and Tyr27 are highly conserved in type IVa pilins but not type IVb pilins. Identical and similar residues to the *N. gonorrhoeae* PilE pilin are shaded. Residues 1, 24 and 27 are boxed. Ng, *N. gonorrhoeae* strain MS11; Nm, *N. meningitidis*; Pa, *P. aeruginosa* (top, strain K; bottom, strain K122-4); Dn, *D. nodosus*; Mb, *Moraxella bovis*; EHEC, enterohemorrhagic *E. coli* O157-H7; Mx, *M. xanthus*; Ec, *Eikenella corrodens*; ntHi, nontypeable *H. influenzae*; Ft, *F. tularensis* Schu S4; Cp, *C. perfringens* (a Gram-positive bacterium). Nm PilX is a minor pilin subunit. Vc, *V. cholerae*; ETEC, enterotoxigenic *E. coli*; Cr, *C. rodentium*; EPEC, enteropathogenic *E. coli*; ST, S. Typhi; and R64, Incl1 plasmid R64.

The GC pilus stability would appear to counteract pilus retraction, which involves disassembly or depolymerization of the filaments. This process, like denaturation or mechanical shearing and breakage of the pilus filaments, involves disruption of non-covalent bonds between pilin subunits. However, dissociation of pili and unfolding of pilin subunits in an aqueous environment, as is the case for thermal or chemical denaturation, is a very different process requiring considerably more energy input than that of pilus retraction, where the
pilin subunits translocate one subunit at a time from the hydrophobic environment of the pilus filament to the hydrophobic environment of the inner membrane in a reversal of the pilus assembly process (Figure 1.13). We showed that both GC pili and TCP dissociate immediately upon treatment with detergent, leaving the monomeric subunits intact and stably folded. Thus, the observed stability of pili in an aqueous environment is not inconsistent with their ability to retract. Additional studies to characterize TCP and other type IVb pili with respect to their stability, retraction capability and range of functions are necessary to elucidate the implications of their unique structural features.

The CD spectra for the pilus filaments and their respective pilin subunits were expected to be comparable, and this was the case for TCP and ΔN-TcpA. However, the CD spectra for GC pili and ΔN-PilE were notably different. This result was supported by CD analysis of detergent-solubilized pilin from GC pili, where this CD profile more closely resembles that of ΔN-PilE than of intact pili. This change in CD spectra for βOG-treated GC pili cannot be due to protein unfolding, as the full-length PilE pilin crystal structure was determined under comparable conditions (Craig et al., 2003; Craig et al., 2006). Thus, these results suggest that the secondary structure of PilE changes upon polymerization into filaments. Notably, an α-helical to β-strand variation was shown for the Shigella flexneri and S. Typhi type III secretion needle protomers in going from its monomeric to polymerized form (Poyraz et al., 2010). This conformational change in PilE upon polymerization is likely to be a small modification that is not
apparent at the 12.5 Å reconstruction resolution, as the monomeric PilE structure fits well into the 12.5 Å GC pilus cryoEM reconstruction (Craig et al., 2006). Additional studies and higher resolution pilus filament structures are necessary to fully comprehend these intriguing results.

In conclusion, results from this study indicate that the type IVb pilin fold is more stable than the type IVa pilin fold, but subunit stability is not synonymous with pilus filament stability. Both GC pili and TCP have comparable helical symmetries and architectures where the pilin subunits are held together and stabilized by extensive hydrophobic interactions. Conserved aromatic residues located at the α1N interaction interfaces of GC pili, together with more compact subunit packing governed by the shape and distinct type IVa pilin fold, may explain the extraordinary stability of GC pili and may be generalizable for all type IVa pili.
5. General Discussion and Conclusion

Successful colonization of the host intestinal epithelium is essential for *V. cholerae* bacterial pathogenesis. Efforts to determine an atomic structure of the key virulence factor of the type IVb TCP from *V. cholerae* are aimed at better understanding the mechanism behind its functions to ultimately prevent and control cholera outbreaks. To understand type IV pilus functions, type IV pilus assembly and type II secretion for bacterial pathogens, many high-resolution structures have been solved for pilin monomers but until this work, only a single pilus structure was available, for the type IVa *N. gonorrhoeae* GC pilus.

Here we used DXMS to map the exposed and buried regions of the *V. cholerae* TcpA in the TCP filament (Li et al., 2008). These experimental data showed that a segment of α1N is exposed in cavities between loosely packed pilin subunits. Consistent with previously-published data, residues involved in pilus:pilus interactions and autoagglutination are located on the protruding D-regions and in the exposed cavities (Kirn et al., 2000; Lim et al., 2010). This model was validated by site-directed mutagenesis of TcpA. Selected residues were mutated at the α2:α1N/α3 interface affecting pilus assembly confirming the involvement of this interface in pilus filament. Results from this study suggest that pili may self-associate via intercalation of D-region bulges of one filament into the
cavities of an adjacent filament.

To further characterize the *V. cholerae* classical biotype TCP, we generated a three dimensional reconstruction using the IHRSR algorithm (Egelman, 2000; Egelman, 2010). The 1.3 Å ΔN-TcpA pilin (Craig *et al.*, 2003) was docked into the three dimensional reconstructions with some certainty to generate a pseudo-atomic resolution structure of the TCP filament. Although the helical symmetry of the TCP reconstructions are different from the previously-published TCP models (Craig *et al.*, 2003; Li *et al.*, 2008; Campos *et al.*, 2011), the pilin:pilin interactions and pilin orientation within these filaments are consistent with DXMS, scanning transmission electron microscopy analysis and site-directed mutagenesis data (Kirn *et al.*, 2000; Li *et al.*, 2008; Lim *et al.*, 2010). The pseudo-atomic resolution structure of TCP shows that the pilin subunits of these helical filaments are held together by the interactions among the N-terminal α-helices while the globular domains are loosely packed leaving gaps permitting filament flexibility and exposing a short segment of α1N. The N-terminal α-helices fill the core of the filament leaving no channel. Pilus stability is attained by extensive hydrophobic interactions between the N-terminal α-helices. The αβ-loop and D-region of the pilin subunits protrude out from the surface exposing this structurally and functionally variable section allowing pilus:pilus interactions.

The type IVb TCP structure was compared with the type IVa GC pilus structure (Craig *et al.*, 2006). Both the GC pilus and TCP TEM reconstructions indicate a highly corrugated surface with bulges and grooves decorating the
filament surface. Thus, the protruding D-region appears to be a trait shared by all type IV pilins. In fact, both N. gonorrhoeae and V. cholerae form microcolonies suggesting that pilus:pilus interactions involving intercalation of bulges and grooves on the filament surface is also common among type IV pili. Both the TCP and GC pilus reconstructions converged to similar symmetry values and are right-handed 1-start helices; TCP: 96.8° rotation and 8.4 Å rise, GC pilus: 100.8° rotation and 10.5 Å rise suggesting that type IVa and type IVb pilins share a common assembly mechanism, as was predicted. If pilus filaments are assembled by sequential addition of subunits, the opposite handedness for the primitive 1-start helix (GC pilus EM structure: right handed (Craig et al., 2006); earlier computational TCP models: left handed (Craig et al., 2003; Li et al., 2008)) imply different assembly mechanisms despite conserved pilus assembly machineries (Figure 1.12).

The pilus filament structures provide clues to the pilus assembly mechanism. We hypothesize that electrostatic attractions drive TcpA pilin monomers from the inner membrane reservoir to the growing pilus filament (Craig et al., 2006; Craig and Li, 2008). Prior to their incorporation into filaments, pilin monomers are anchored in the inner membrane by their extended hydrophobic α-helices. Type IV pilin subunits dock into an existing gap at the base of the filament and are added one after another in a clockwise direction. The subunits are extruded approximately 8-10 Å from the inner membrane by the coordinated action of the assembly ATPase and the IMP (Figure 1.12). Residue
Glu5 in the protruding, hydrophobic N-terminal α-helix, is conserved among type IV pilins and is essential for efficient type IV pilus assembly of bacterial pathogens. The new TCP filament structures show a staggered arrangement of the TcpA pilin monomers within the filaments which allow each negatively charged Glu5 side chain to interact with the positively charged N-terminal residue. The formation of these salt bridges neutralizes the unfavourable charges in the hydrophobic core of the filament validating our hypothesis. In support of this model, a Glu5Ala substitution in TcpA resulted in inefficient TCP assembly similar to *N. gonorrhoeae*, *P. aeruginosa* and IncI1 plasmid R64 pili (Pasloske and Paranchych, 1988; Strom and Lory, 1991; Macdonald *et al.*, 1993; Horiuchi and Komano, 1998; Aas *et al.*, 2007). Notably, residue Glu5 is also present in the related type II secretion pseudopilins and the ETEC pseudopilins are related by a rotation of ~100° rotation in a right-handed helix (Korotkov and Hol, 2008) suggesting a common architecture and mechanism of assembly for type IV pili and type II secretion pseudopili.

Type IVa pili are found on a diverse group of human pathogens whereas type IVb pili are only present on enteric pathogens. The type IVa and type IVb pilins have similar overall architectures but they are distinguished by differences in amino acid sequence, length and topology of the globular domain. We proposed that the type IVb pilin fold, where the C-terminus is buried within the core of the β-sheet, is more stable than the type IVa fold and this might impart enhanced stability to the type IVb pili. To examine structural differences between
the type IVa and type IVb pilins and their impact on pilus architecture and stability, the relative stabilities were compared for the type IVa GC pilus from *N. gonorrhoeae* and the type IVb TCP from *V. cholerae*. Results indicate that GC pilus filaments are more stable than TCP despite their pilin subunits being less stable. TcpA and other type IVb pilins have larger and bulkier D-regions and the subunits are packed more loosely in the filament resulting in deeper grooves and cavities compared to type IVa pilins likely impacting their stability. Furthermore, the stacking of conserved aromatic residues within the core of the GC pili and other type IVa pili, that is not present in type IVb pili including TCP, contributes to type IVa pilus stability. Enteric pathogens utilize type IVb pili to mediate colonization possibly by acquiring the genes horizontally from other gut bacteria or phage whereas type IVa pili evolved to be more stable. The robustness of GC pili may be necessary to withstand high stress forces in the urogenital tract and during retraction for twitching motility and intimate adhesion (Higashi *et al.*, 2007; Higashi *et al.*, 2009). This robust stability may be a common feature of type IVa pili as an adaptation to the ecological niches of the bacteria and biological demands and functions of these filaments.

A high-resolution, full-length type IVb pilin structure and higher resolution TCP reconstruction would produce a more accurate filament model, but the TCP reconstructions determined here advance our understanding of pilus functions and the assembly mechanism. These results provide insights into pilus:pilus interactions and their role in *V. cholerae* microcolony formation, with implications
for other bacterial pathogens that rely on type IV pili or the related T2SS for
infection. The TCP models will aid the design of tcpA mutants that affect
pilus:pilus interactions or TCP filament assembly to further test the type IV pilus
assembly model. Further studies to determine atomic structures of the pilus
biogenesis machinery components are necessary to obtain a more
comprehensive understanding of pilus-mediated bacterial interactions and type
IV pilus assembly. By elucidating these key virulence mechanisms, one can
attempt to design strategies to control pathogenesis. For example, the exposed
N-terminal α-helices and neighbouring grooves of type IV pili may be conserved
targets for therapeutics. Antimicrobial peptides or agents may bind in the cavity
and block essential pilus functions such as bundling or retraction. Alternatively,
bacteriocidal or bacteriostatic agents may bind to pili and be taken up by the
pathogens upon pilus retraction. One can design antimicrobial agents by
computational modeling and docking methods to prevent the spread of infection
and ultimately the diseases caused by these pathogens.
References


