Molecular Cell, Vol. 11, 1139–1150, May, 2003, Copyright ©2003 by Cell Press

Type IV Pilin Structure and Assembly: X-Ray and EM Analyses of *Vibrio cholerae* Toxin-Coregulated Pilus and *Pseudomonas aeruginosa* PAK Pilin

Lisa Craig,[1] Ronald K. Taylor,[3] Michael E. Pique,[1] Brian D. Adair,[2] Andrew S. Arvai,[1] Mona Singh,[1,7] Sarah J. Lloyd,[1] David S. Shin,[1] Elizabeth D. Getzoff,[1] Mark Yeager,[2,4] Katrina T. Forest,[5] and John A. Tainer[1,6,*]

[1]Department of Molecular Biology and
The Skaggs Institute for Chemical Biology
[2]Department of Cell Biology
The Scripps Research Institute
La Jolla, California 92037
[3]Department of Microbiology and Immunology
Dartmouth Medical School
Hanover, New Hampshire 03755
[4]Division of Cardiovascular Diseases
Scripps Clinic
La Jolla, California 92037
[5]Department of Bacteriology
University of Wisconsin–Madison
Madison, Wisconsin 53706
[6]Life Sciences Division
Lawrence Berkeley National Laboratory
Berkeley, California 94720

Summary

Pilin assembly into type IV pili is required for virulence by bacterial pathogens that cause diseases such as cholera, pneumonia, gonorrhea, and meningitis. Crystal structures of soluble, N-terminally truncated pilin from *Vibrio cholera* toxin-coregulated pilus (TCP) and full-length PAK pilin from *Pseudomonas aeruginosa* reveal a novel TCP fold, yet a shared architecture for the type IV pilins. In each pilin subunit a conserved, extended, N-terminal α helix wrapped by β strands anchors the structurally variable globular head. Inside the assembled pilus, characterized by cryo-electron microscopy and crystallography, the extended hydrophobic α helices make multisubunit contacts to provide mechanical strength and flexibility. Outside, distinct interactions of adaptable heads contribute surface variation for specificity of pilus function in antigenicity, motility, adhesion, and colony formation.

Introduction

The virulence factors of bacterial pathogens are a current focus in cell biology and medicine due to increasing antibiotic resistance. Type IV pili are key virulence factors for many important human pathogens including *Vibrio cholerae*, *Pseudomonas aeruginosa*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, and *Escherichia coli*. These filaments mediate bacterial motility on solid surfaces (twitching motility), host cell adhesion, microcolony formation, bacteriophage adsorption, and natural transformation. The type IV pili are assembled from thousands of copies of a single protein subunit, pilin. As they occur on the surfaces of virtually all Gram-negative bacteria, type IV pili and their pilin subunits represent a potentially general antibacterial target.

The toxin-coregulated pili (TCP) are type IV pili of *V. cholerae*. They are essential for colonization of the human intestine (Herrington et al., 1988; Tacket et al., 1998) and are receptors for CTXϕ, the bacteriophage that carries genes for cholera toxin (Waldor and Mekalanos, 1996). TCP promote bacterial interactions, allowing *V. cholerae* cells to establish microcolonies on the intestinal epithelia (Taylor et al., 1987; Chiang et al., 1995; Kirn et al., 2000), where they secrete cholera toxin and cause severe diarrhea (reviewed in Kaper et al., 1995). The type IV pili of *P. aeruginosa* promote twitching motility (Bradley, 1980) and attachment to host cells, enabling this opportunistic pathogen to colonize the lungs of cystic fibrosis patients and immunosuppressed individuals (Woods et al., 1980). The *Neisseria* pili mediate attachment to host cells and formation of microcolonies, colonization events that can lead to gonorrhea (*N. gonorrhoeae*) (Swanson, 1983; Virji and Heckels, 1983; Merz and So, 2000) or meningitis (*N. meningitidis*) (Nassif et al., 1994). Elucidating the molecular fold and architecture of these essential virulence factors should provide a foundation for understanding host-pathogen interactions and the mechanisms of twitching motility, adherence, and microcolony formation and would aid the design of bacterial vaccines and therapeutic agents capable of inhibiting pilus function.

Type IV pili are extremely long (1–4 μm), thin (50–80 Å), and flexible; yet these filaments are very strong, withstanding stress forces of greater than 100 pN (Merz et al., 2000; Maier et al., 2002). The type IV pilin subunits share N-terminal sequence similarity, a C-terminal disulfide bond, and a conserved cellular machinery for assembly into pilus filaments, suggesting a common subunit structure and filament architecture. However, X-ray structural analyses of pilin subunits have proved extremely challenging because their tendency to form fibers impedes crystallization. To date, the atomic structure of gonococcal (GC) pilin from *N. gonorrhoeae* is the only available structure of a full-length type IV pilin (Parge et al., 1995). The conserved N-terminal 52 residues of GC pilin form an extended α-helical spine that protrudes from the globular head domain for half of its length and is predicted to mediate fiber formation via hydrophobic association (Folkhard et al., 1981; Watts et al., 1983; Parge et al., 1995). Recently, structures of the globular C-terminal domains of two closely related pilins from *P. aeruginosa* were determined by X-ray crystallography (Hazes et al., 2000) and NMR (Keizer et al., 2001) by removing the hydrophobic N-terminal segment that corresponds to the protruding half of the α helix (α1-N) in the GC pilin structure. Despite low sequence similarity, the globular head domains of the *Pseudomonas* pilins resemble GC pilin in fold and disulfide bond arrangement. Yet, distinct models for pilus assembly have been proposed from these similar pilin structures

*Correspondence: jat@scripps.edu
[7]Present address: Albert Einstein College of Medicine, The Bronx, New York 10461.

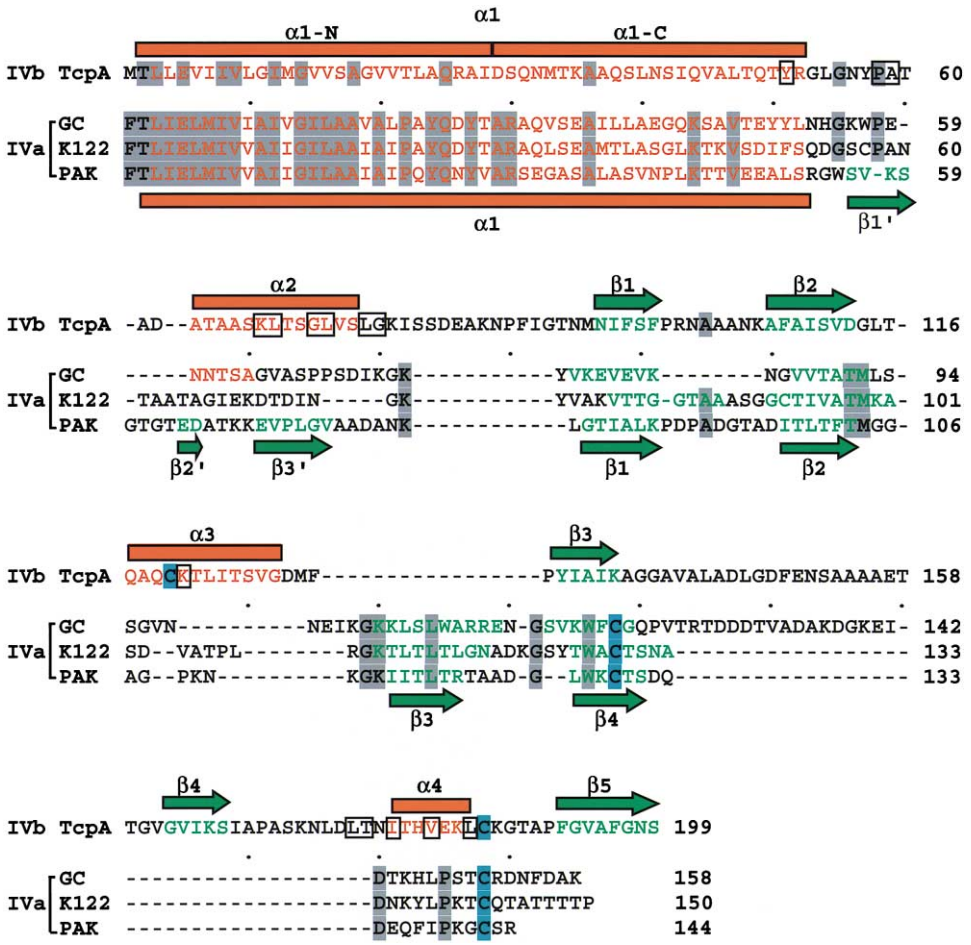


Figure 1. Sequence and Structural Alignment of the Four Type IV Pilins of Known Structure

TcpA (*V. cholerae* classical biotype, strain RT4236), GC (*N. gonorrhoeae* strain MS11), K122 (*P. aeruginosa* strain K122-4), and PAK (*P. aeruginosa* strain K). Sequences were aligned by structure superposition, and identical residues were shaded in gray only when their structures and/or positional orientations match as well. Sequence similarity between the type IVa and IVb classes exists primarily in the N-terminal 25 residues and is more pronounced within each class. Conserved cysteines are shaded in cyan. α helices (orange bars) and β strands (green arrows) are shown above the sequences for TcpA and below the sequences for PAK pilin with residues in all four sequences similarly colored by secondary structure. In the TcpA crystal structure and the K122-4 pilin NMR structure, α1-N (residues 1–28) is absent but is presumed to be helical, as implied by sequence similarity with the PAK and GC pilins. Residues involved in the hydrophobic interaction between subunits in the three-start helix are boxed. The sequences listed were obtained from Shaw and Taylor (1990) (TcpA), PDB file 1AY2 (GC), PDB file 1HPW (K122-4), and SWISS-PROT file P02973 (PAK).

(Parge et al. 1995; Hazes et al., 2000; Keizer et al., 2001; Chattopadhyaya and Ghose, 2002), suggesting the need for additional structural information from other full-length pilins, less-similar type IV pilin sequences, and intact pilus filaments.

A notable gap in our understanding of type IV pilus assembly is the absence of structural information for the type IVb subclass, which may be structurally distinct from the GC and *Pseudomonas* type IVa pilins. The type IVb pili are present on bacteria that colonize the human intestine and include TCP from *V. cholerae*, the bundle-forming pili from enteropathogenic *Escherichia coli* (Giron et al., 1991; Blank et al., 2000), longus and CFA/III from human enterotoxigenic *E. coli* (Giron et al., 1991; Taniguchi et al., 1995), and PilS from *Salmonella enterica* (Zhang et al., 2000). Type IVb pilins share features with the type IVa subclass, including assembly pathway, a C-terminal disulfide bond, and sequence similarity in the N-terminal 25 residues (Figure 1), but differ in having a longer leader sequence, a variable methylated N-terminal residue, and little or no similarity to the type IVa pilins in the C-terminal domain (Strom and Lory, 1993; Giron et al., 1997). There is, however, considerable sequence similarity *among* the type IVb pilins. Comparative structural analyses of these two pilin subclasses are therefore important to address questions regarding fold, assembly, and architecture among type IV pili.

Here we report the high-resolution X-ray crystal structures of the type IVb TCP subunit, TcpA, which is one of the largest type IV pilins (199 amino acids), and full length PAK pilin, one of the smallest (144 amino acids). By using X-ray crystallography and cryo-electron microscopy, we determined the helical symmetry and subunit organization for TCP. These analyses address fundamental questions regarding type IV pilus structure and assembly: Do the type IVb pilins differ in structure

Table 1. Crystallographic Data Collection and Analysis

Data Collection	Native TcpA	SeMet-TcpA			PAK C2	PAK $P2_12_12_1$
Beamline	SSRL 11-1	SSRL 9-2			SSRL 9-2	SSRL 9-1
Space group	$P6_3$	$P6_3$			C2	$P2_12_12_1$
Cell a, b, c (Å)	157.4, 157.4, 35.7	89.9, 89.9, 35.7			147.0, 44.4, 73.8	44.3, 67.8, 127.1
Cell α, β, γ (°)	90.0, 90.0, 120.0	90.0, 90.0, 120.0			90.0, 116.8, 90.0	90.0, 90.0, 90.0
Resolution (Å)	1.30	1.86			2.0	3.0
Wavelength (Å)	0.975913	0.979029	0.932183	0.979224	0.9791	0.785
Completeness (%)[a]	98.4/93.9	98.1/87.8	94.4/67.4	98.1/87.1	97.2/83.6	98.9/99.5
Observed reflections	573,481	112,246	121,301	112,213	324,152	95,883
Unique reflections	115,573	15,552	17,955	15,561	29,308	8,288
R_{sym} (%)[a,b]	7.3/24.5	7.2/21.9	7.2/26.9	6.3/22.2	7.9/32.9	9.5/59.0
I/σ[a]	5.1/1.8	15.8/3.1	14.9/2.4	15.8/3.0	14.8/2.4	16.0/2.4
Mosaicity	0.32	0.54	0.54	0.54	1.20	0.90
Refinement statistics						
Resolution limits (Å)	30–1.30	30–1.86			30–2.00	20–3.00
Molecules/A.U.	3	1			2	1
R_{cryst} (%)[c]	11.6	23.6			22.3	24.5
R_{free} (%)[d]	17.4	24.8			24.8	31.1
Average B factor (Å^2)	15.2	14.5			26.3	52.4

[a] Overall/last shell.
[b] R_{sym} is the unweighted R value on I between symmetry mates.
[c] $R_{cryst} = \Sigma hkl||Fobs(hkl)| - |Fcalc(hkl)||/\Sigma hkl|Fobs(hkl)|$.
[d] R_{free} is the cross validation R factor for 5% of reflections against which the model was not refined.

from the type IVa pilins? Is the extended N-terminal α helix a hallmark of type IV pilins? How does the hydrophobic N terminus act in concert with the globular head domain to promote pilin oligomerization? Do all type IV pilus filaments share the same helical symmetry and molecular architecture?

Results and Discussion

Novel *V. cholerae* TcpA Fold and Discrete Assembly Interfaces

To determine whether the sequence-distinct type IVb pilins share the same fold as the type IVa pilins from *P. aeruginosa* and *N. gonorrhoeae* (Figure 1), we solved the crystal structure of TcpA, the subunit of the type IVb pilus from *V. cholerae*, classical biotype. To provide sufficient quantities of protein for crystallization, we prepared recombinant TcpA in which residues 1–28 (α1-N) were replaced with an N-terminal histidine tag (total 21 residues), resulting in a soluble protein. Similar truncation of PAK pilin also produced soluble protein (Hazes et al., 2000), suggesting that the intact N terminus plays a key role in pilus assembly. The TcpA crystal structure was solved to 1.3 Å resolution by multiwavelength anomalous dispersion methods (MAD) with selenomethionine-labeled TcpA (Table 1).

The TcpA structure reveals conserved architectural features plus key differences in fold and structure between TcpA and the type IVa pilins. The globular head domain of TcpA is comprised of an N-terminal α helix (α1-C, residues 29–52) and an antiparallel β sheet (Figures 1 and 2). Three strands of the β sheet interact with α1-C and an additional α helix (α3) to form a tight hydrophobic core within the globular head domain, while the last two strands of the sheet twist away from this core. The hydrophobic core of TcpA closely resembles the $\alpha\beta$ roll of GC and PAK pilin, particularly over α1-C, β1, and β2 (rmsd of 1.8 Å for TcpA on PAK pilin for 30 Cα atoms), and thus represents a conserved structural scaffold for the type IV pilins. The TcpA structure reveals an additional conserved feature: a disulfide linkage that joins a variable structure on one face of the globular head to the conserved α helix/β sheet scaffold.

Key structural differences between TcpA and the type IVa pilins occur in (1) the connectivity of the β sheet; (2) the nature and location of the disulfide linkage; and (3) the structural elements forming two discrete faces of the globular head, which lie on opposite sides of the N-terminal α helix (Figure 2) and are available for interaction with neighboring head domains in the assembled pilus. In the globular heads of GC and the *Pseudomonas* pilins, the N-terminal α helix (α1) is connected to a continuous four-stranded antiparallel β sheet, followed by a C-terminal loop of variable size. The TcpA structure is more complex, having an N-terminal α helix followed by a second, shorter α helix (α2) and then a five-stranded β sheet (Figures 1 and 2A–2C). The β sheet has two nonnearest neighbor repeats plus large insertions in three of its loops, including two additional α helices (α3 and α4). Unexpectedly, the C-terminal 9 residues form the central strand of the β sheet, β5. The β strands of TcpA are numbered β1 to β5 in sequence order (Figures 2A–2C), as was done for GC and PAK pilin (Figures 1, 2D, and 2F; and Parge et al., 1995, Hazes et al., 2000). This reveals the distinct β sheet topology of TcpA (β1→β2→β5→β3→β4) versus the type IVa pilins (β1→β2→β3→β4; Figure 2) and establishes a new protein fold for TcpA based on searches of all known protein structures (the DALI server; Holm and Sander, 1995).

TcpA is one of the largest type IV pilins, having 199 residues compared with 158, 144, and 150 residues for GC, PAK, and K122-4 pilin, respectively. This extra mass is primarily distributed in two regions forming surface-

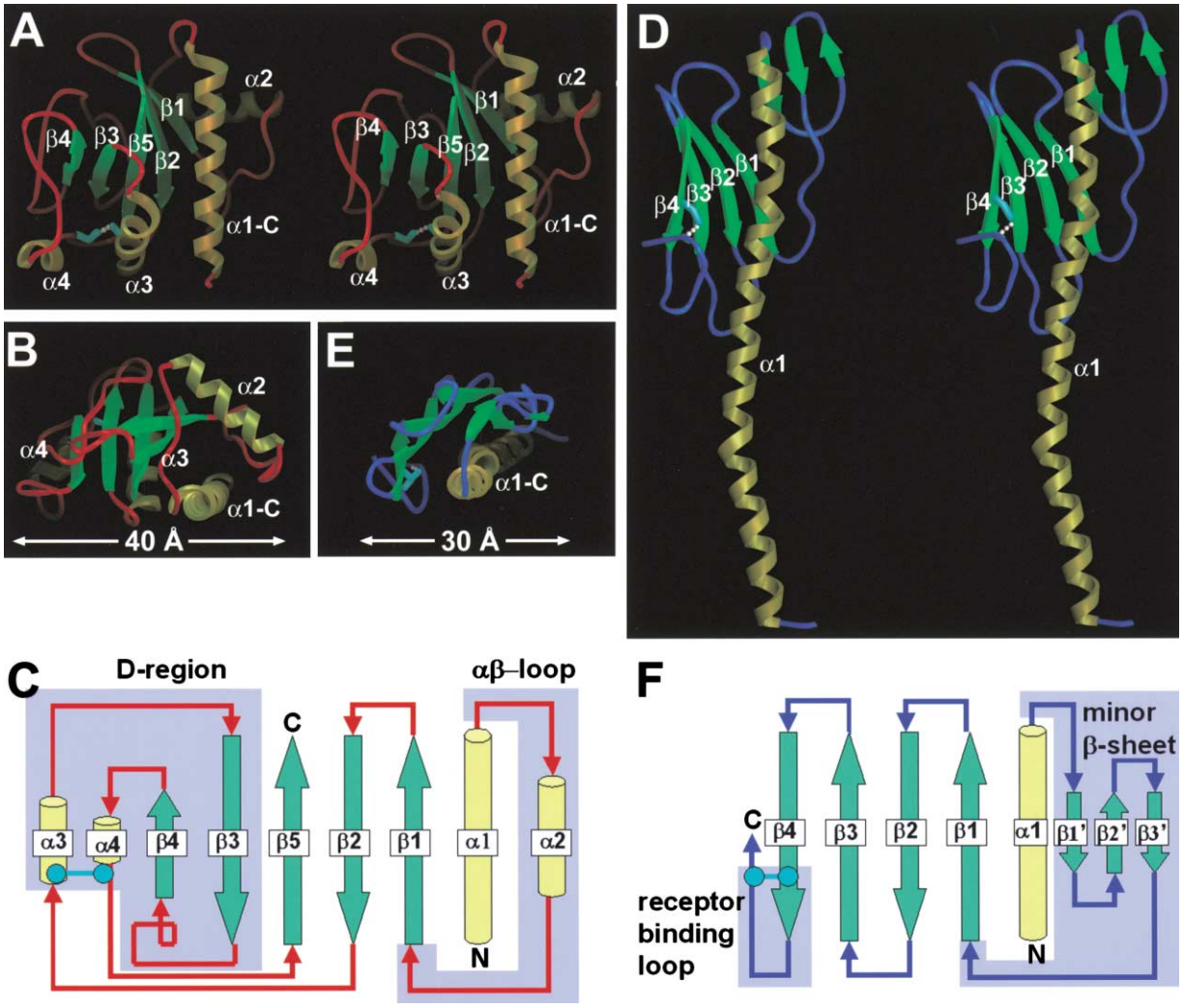


Figure 2. Comparison of Crystal Structures of *V. cholerae* TcpA and *P. aeruginosa* Strain K Pilin

(A) Stereo view of TcpA ribbon structure showing the truncated N-terminal α helix (yellow), β sheet (green), and disulfide bond (white spheres between the cyan side chains) seen in the type IVa pilins.

(B) Top view of TcpA, rotated 90° toward the viewer relative to (A).

(C) TcpA secondary structure diagram showing the connectivity of the β sheet. The αβ loop and D region are shown with a blue background.

(D) Stereo view of *P. aeruginosa* strain K (PAK) pilin ribbon structure showing the entire α1, which follows the same S-shaped curve as α1 of GC pilin.

(E) Top view of PAK pilin shown for size comparison with TcpA.

(F) PAK pilin secondary structure diagram showing the minor β sheet and receptor binding loop, which are located in the αβ loop and D region, respectively. The major structural differences between TcpA and PAK pilin lie in the connectivity of the β sheet, resulting in a new protein fold for TcpA, and in the size and secondary structure of the αβ loop and the D region.

exposed faces on either side of the N-terminal α1 helix in the globular head domain. These are (1) the loop connecting α1 with the β sheet, termed the αβ *loop*, and (2) the *D region*, delineated by the disulfide linkage (Figure 2). The αβ loop varies in size, conformation, and function among the type IV pilins, whereas the D region differences are most pronounced between TcpA and the type IVa pilins. In the αβ loop of TcpA, the four-turn α2 helix lies perpendicular to α1 and is flanked by extended segments (Figure 2A–2C). The αβ loop of GC pilin is extended and contains an O-glycosylated serine (Ser63) and a phosphoserine (Ser68), which are predicted to be surface exposed (Parge et al., 1995; Forest et al., 1999). In PAK pilin, the αβ loop forms the minor β sheet (Figures 2D–2F; and Hazes et al. 2000), and in K122-4 pilin, this region is poorly defined but has helical content (Keizer et al., 2001). Thus, in each of these pilins, the structurally variable αβ loop forms a discrete face or "edge" of the globular head (Figure 2) that is partially surface exposed and partially buried in the assembled pilus. This suggests that the αβ loop confers specificity of subunit interactions within the filament, as well as interactions between the pili and their environment.

The D region of TcpA comprises a second discrete face of the globular head domain, opposite the αβ loop (Figures 2A and 2B). Like the αβ loop, the D region exposes completely different structural elements in the various type IV pilins for interaction with neighboring

subunits: α helices and loops in TcpA; and β sheet and turns in GC, PAK, and K122-4 pilin (Figure 2). Differences in the D region edge between TcpA and the type IVa pilins result from substantial differences both in the local environment of the cysteines and in the conformation of the region they delineate (Figure 2). The disulfide bond of TcpA connects Cys120, which lies near the beginning of α3, to Cys186, which lies on a type II turn immediately following α4 (Figures 2A and 2C). The segment between the two cysteines comprises 65 residues (α3, β3, the β3-β4 flap, β4, and α4). In contrast, the disulfide bond of the type IVa pilins links the last strand of the β sheet (β4) to a peripheral C-terminal loop and delineates a much smaller segment, with 29, 12, and 12 residues for GC, PAK, and K122-4 pilin, respectively (Figures 1, 2D, and 2F). In the *Pseudomonas* pilins, the D region forms the receptor binding loop at the C terminus (Irvin et al., 1989; Lee et al., 1989), which is only exposed at the tips of the filaments (Lee et al., 1994). In GC pilin, the D region includes a hypervariable β hairpin important for antigenic variation (Seifert and So, 1984; Parge et al., 1995), followed by an extended loop. In TcpA the D region is predicted to be partially exposed on the TCP surface (Sun et al., 1997) and to contain residues important for pilus:pilus interactions as well as TCP structure stabilization, as determined by mutational analysis (Kirn et al., 2000). In TcpA and the type IVa pilins, the D region is stabilized by the disulfide bond, which links it to the α helix/β sheet scaffold. Thus, the structurally variable D region, like the αβ loop, evidently imparts specificity to both pilus assembly and function.

Interestingly, the D region in other type IVb pilins is of comparable size to the D region of TcpA. This observation, as well as the considerable sequence similarity observed among the type IVb pilins, suggests that the structure and topology of other type IVb pilins resemble TcpA rather than the type IVa pilins. Thus, our results argue that while they share an α helix/β sheet structural scaffold, the type IVa and IVb pilins have distinct globular folds with consequent differences in the secondary structural elements available for head group interactions and exposed on the surface of the assembled pili.

Full-Length PAK Pilin Structure

To test the hypothesis that α1-N and α1-C form an extended, continuous α helix (α1) in other type IV pilins, as seen for full-length GC pilin, we solved the crystal structure of full-length PAK pilin. Two independent PAK pilin space groups (C2 and $P2_12_12_1$ to 2.0 Å and 3.0 Å, respectively) revealed nearly identical structures (Table 1). The globular head domain of the full-length PAK pilin closely resembles the previously determined structure of N-terminally truncated PAK pilin (rmsd of 0.69 Å for all atoms over residues 30–144) (Hazes et al., 2000) (Figures 2D–2F), confirming that removal of α1-N and expression of PAK pilin in *E. coli* retains the globular head domain structure. Thus, our similar engineering and expression of TcpA is unlikely to have altered the structure of the TcpA head.

To examine the key issue of structural conservation of the N-terminal domain, which acts in pilus assembly, we compared the full-length PAK and GC pilin structures. Residues 1–52 of PAK pilin form a single extended ~80 Å long α helix that resembles α1 of GC pilin (Parge et al., 1995) in length and curvature (Figure 2D). The N-terminal half of the PAK pilin α helix (α1-N) protrudes from the head domain, forming a 41 Å long "handle," and the C-terminal half, α1-C, interacts intimately with the β sheet, forming the αβ roll fold. Both PAK and GC pilin have helix-perturbing residues at positions 22 (Pro) and 42 (Pro for PAK and Gly for GC pilin). These residues introduce two kinks in the helix, giving it an S-shaped curve. When GC pilin and the three independently solved full-length PAK pilin structures (two molecules in the asymmetric unit for the C2 space group and one molecule for the $P2_12_12_1$ space group) are superimposed over α1-C, flexibility is evident in the splaying of the N termini (by as much as 10 Å). In TcpA and the other type IVb pilins, the residues at positions 22 (Thr, Ile, Ser, or Phe) and 42 (Ser or Thr) are less likely to induce kinks. Indeed, in our TcpA structure, the C-terminal half of α1 (α1-C) is straight (Figures 2A and 2B). However, the considerable length and exposure of the protruding α1-N suggest flexibility, which would facilitate packing of the N-terminal α helix in the core of the pilus filament. On the basis of (1) the extended conformation of the N-terminal α helices in the full-length PAK and GC pilin structures, (2) the similar α1-C helices in the truncated structures of PAK, K122-4 pilin, and TcpA, and (3) the sequence similarity, we predict that the extended N-terminal α helix is a conserved structural feature of all type IV pilins. These results therefore support a conserved architecture for type IV pilins, whereby the extended N-terminal α helix and overlapping α helix/β sheet structural scaffold anchor the structurally variable and exposed regions of the globular head domain.

In most type IV pilins, the only charged residue in the first 25 amino acids is the conserved Glu5, which is essential for assembly of *Pseudomonas* pili (Pasloske et al., 1989; Strom and Lory, 1991; Macdonald et al., 1993). In our PAK pilin structure, the side chain of Glu5 points toward the N terminus and is positioned to form a salt bridge with the N-terminal amide of Phe1, suggesting that Glu5 may neutralize the positively charged N terminus in the hydrophobic pilus core. Other residues in the N terminus are tolerant to hydrophobic amino acid substitutions (Strom and Lory, 1991) but not to less conservative ones, which can affect pilus assembly, morphology, and function (Chiang et al., 1995; Park et al., 2001). These results, as well as the observation that removal of α1-N abolishes pilus formation (this work; Hazes et al. 2000; Keizer et al. 2001), support a key role for the structurally conserved hydrophobic N-terminal α helix in fiber assembly.

Electron Microscopy Reveals a Distinct TCP Assembly

To examine pilus assembly, we recorded electron micrographs of both negatively stained and frozen-hydrated TCP filaments (Figures 3A and 3C). Fourier transforms of the pili produced layer lines corresponding to a three-start helical fiber with a 45 Å pitch (Figures 3B and 3D). Thus, the subunits in the TCP filaments are organized along three identical protein strands that form a helical twist around the fiber axis and have an axial separation of 45 Å (Figure 3E). Layer line intensities were strongest

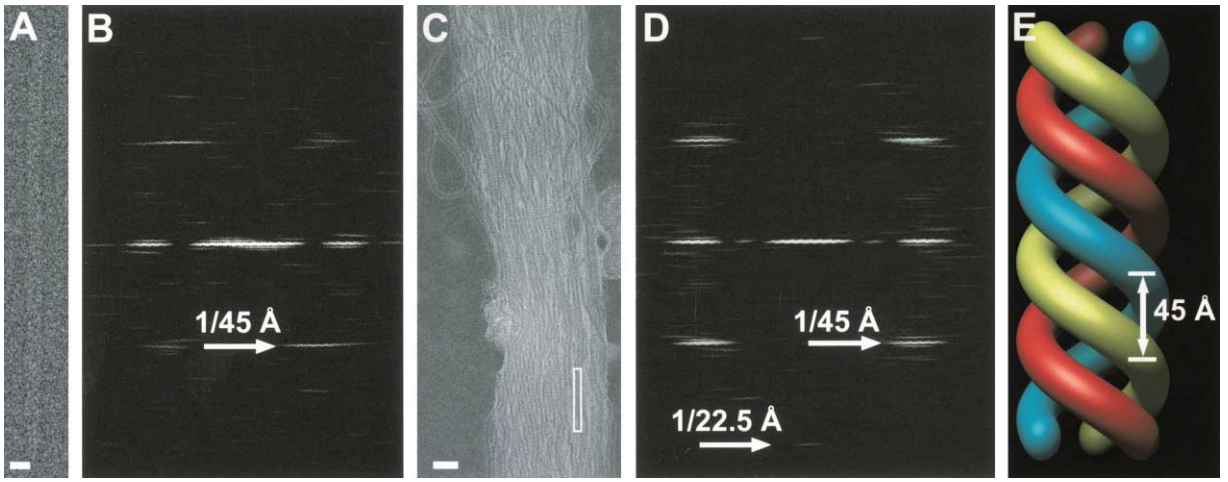


Figure 3. EM Analysis of TCP Reveals a Three-Start Helix with a 45 Å Pitch

(A) Image of a frozen-hydrated TCP filament and (B) computed Fourier transform showing the J_3 layer line at 1/45 Å. (C) Bundle of negatively stained TCP filaments and (D) computed Fourier transform of a single filament within the bundle as indicated by the box in (C). Scale bar: 100 Å for (A) and 500 Å for (C). The Fourier transforms show the J_3 layer line at 1/45 Å representing the three-start helix, as well as a weaker near-meridional layer line at 1/22.5 Å. The helical pitch is determined by measuring the vertical distance between the peak of the layer line and the equator (horizontal line at the center of the Fourier transform). Distances are in reciprocal space units ($Å^{-1}$). The upper left quadrant is identical to the lower right quadrant, and the layer lines on either side of the meridian represent the near and far sides of the filament; thus, only one layer line is labeled in each transform. (E) Left-handed representation of a three-start helix with each start shown in a different color. The 45 Å pitch is the axial distance between two helical strands.

for single filaments in bundles of negatively stained TCP (Figure 3D). For some TCP filaments, a meridional layer line was observed at 1/22.5 Å from the equator (Figure 3D), which is exactly half the value of the three-start helix pitch and may correspond to the rise of the subunits. A dominant three-start helix has been observed for other filamentous assemblies, including the Hib pili from *Haemophilus influenzae* type b (Mu et al., 2002), and is consistent with this being a particularly stable form of subunit packing in a filament. Notably, this three-start helical symmetry for TCP differs from the dominant one-start symmetry determined for PAK pili by X-ray fiber diffraction (Folkhard et al. 1981). Differences in helical symmetry between these two pili likely result from differences in the dimensions, surface structure, and chemistry of the globular head, as shown for the $\alpha\beta$ loop and D region (Figures 1 and 2), which confer different subunit interactions and packing arrangements on the filaments. Our discovery of three-start helical symmetry for TCP indicates that the visually similar type IVa and type IVb pilus subclasses represent distinct architectural assemblies to accommodate the distinct sequence families.

Crystallographic Packing Implications for the Helical Architecture of TCP Filaments

Unlike the full-length GC experiments done in detergent (Parge et al. 1995) and the truncated PAK pilin experiments done in high salt (Hazes et al., 2000), the TcpA crystals were grown without detergent and at near physiological conditions (see Experimental Procedures), which allowed the head domains to interact as they might within the assembled TCP filaments. The TcpA crystal packing therefore provided clues to the naturally preferred subunit interactions. The TcpA subunits in the $P6_3$ crystal lattice pack as helical filaments (Figures 4A–4C). Furthermore, the N-terminal α helices face the center of the filament and are oriented roughly parallel to the helical axis (Figures 4A–4C), as indicated by fiber diffraction for *P. aeruginosa* pili (Folkhard et al., 1981). Similar to the TCP filaments, these "crystallographic fibers" have a dominant three-start helical assembly with the subunits in each of the helical starts at the same level and 3-fold symmetric about the fiber axis. Each strand of the three-start helix has six subunits per turn, and the subunits are related by an azimuthal rotation (twist) of 60° about the fiber axis and an axial translation (rise) of 17.9 Å along the fiber axis (Figures 4B and 4C). This three-start helical fiber has a pitch of 35.7 Å and a maximum diameter of 100 Å compared to the intact TCP filaments, which have a 45 Å pitch and a diameter of ~80 Å from EM. In the crystallographic fiber, the N-terminal 28 residues of TcpA (α1-N) were replaced with a 21 residue disordered histidine tail, which occupies a central 25 Å channel in the crystallographic fiber. In the native TCP filaments, this channel is expected to be filled by hydrophobic N-terminal α helices that guide and stabilize fiber assembly.

In the TcpA crystal packing, two discrete interfaces connect subunits in the crystallographic fibers: (1) a polar interface between the stacked subunits directly in line along the fiber axis (Figure 4C) and (2) a hydrophobic interface between subunits in each strand of the left-handed three-start helix (Figures 4C and 4D). The polar interface, which knits the three-start helical strands together, varies among the three molecules in the asymmetric unit, suggesting that it is unique to TcpA crystal packing. In contrast, the hydrophobic interface connecting subunits within each strand of the three-start helical assembly is consistent among the three molecules in the asymmetric unit and is side chain mediated. These

subunits are joined by hydrophobic interactions between residues on and around helix α2 in the αβ loop of one subunit, and α3 and α4 in the D region of a neighboring subunit, plus two hydrogen bonds (Figure 4D). This interface buries 370 Å^2 of surface area and results in a contiguous mass distribution along the strands of the left-handed three-start helix.

The similar dimensions and symmetry of the crystallographic fibers and TCP filaments suggest that the *V. cholerae* pilus architecture resembles the crystallographic fibers. The residues involved in the three-start interaction in the TcpA crystallographic fiber are conserved among pathogenic isolates of *V. cholerae* (Boyd and Waldor, 2002), supporting a common architecture for TCP variants. The hydrophobic patches joining subunits within the three-start helical strands occur at the two sites with the greatest structural variation among the type IV pilins, the αβ loop and the D region (Figures 1, 2, and 4D). The interaction surface created by the D region is stabilized by the conserved disulfide bond, which links this edge to the α helix/β sheet scaffold through α3. The structurally variable surfaces formed by the αβ loop and D region may also mediate head:head interactions in other type IV pili, explaining the symmetry differences observed for TCP and PAK pili and providing a molecular basis for the pilus assembly specificity (Elleman and Peterson, 1987). Our data suggest that, in addition to solvent exclusion of the hydrophobic tail, pilus assembly is driven by specific interactions between the globular head domains, providing both assembly specificity and added stability.

Structurally Implied TCP Assembly

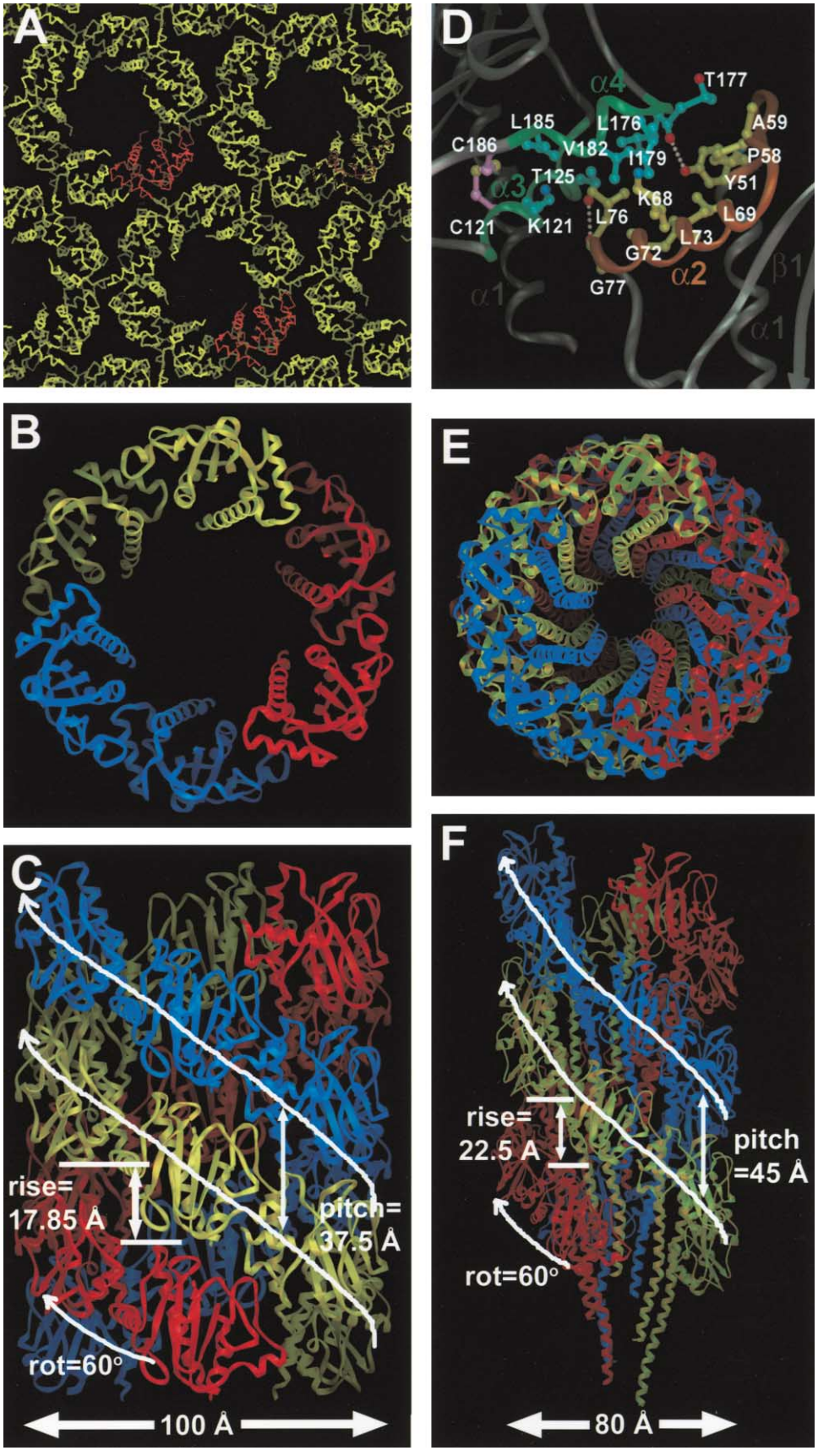
We computationally built TCP filaments from the dimensions and symmetry determined by EM and the subunit interactions revealed by the crystallographic fibers. Millions of TCP models were generated and evaluated by the following criteria: ~80 Å fiber diameter, ~45 Å pitch for a three-start helix, proximity of selected positive to negative charges, burial of hydrophobic residues, minimal steric clashes, and reasonable packing (see Supplemental Materials at http://www.molecule.org/cgi/content/full/11/5/1139/DC1). Of the vast conformational space of subunit orientations sampled by our search, models having a left-handed three-start helix with a 22.5 Å rise and a 60° counter-clockwise rotation best fit the above criteria. This independently derived 22.5 Å rise matches the meridional 1/22.5 Å layer line in the Fourier transforms of negatively stained TCP bundles (Figure 3D). Models were further refined by selecting those that retained subunit interactions similar to those observed in the crystallographic fibers.

Our preferred TCP model places the N-terminal α helices at about 10° angles to the fiber axis to form a central ~45 Å diameter core (Figures 4E and 4F). The remaining small central channel (~5 Å in diameter) is too narrow for double- or single-stranded DNA passage, as was proposed by Karaolis et al. (1999). The only steric clashes occur between terminal atoms in side chains of Arg26 and Ile179 and backbone atoms, and could be resolved in a flexible TCP filament. The maximum atom-to-atom fiber diameter of the TCP model, 88 Å, resembles the 80 Å diameter of the TCP filaments determined by EM, given that (1) the EM measurements reflect an average fiber diameter and not the distance between the flexible outermost atoms, and (2) backbone and side chain flexibility would allow a more compact structure. Importantly, the nature and spacing of lateral interactions between subunits along the three-start helix of the crystallographic fiber are maintained. The small increase in distance between the subunits aligned axially along the filament, compared with the crystallographic fiber (Figure 3C), is consistent with the three-start strands being cemented together by the hydrophobic N-terminal α helices, and thus not requiring axial connections. This axial space between subunits may contribute to the functionally important flexibility of the TCP filaments.

The azimuthal rotation for each subunit in the three-start strand of the TCP model matches the crystallographic fiber: 60°, giving six subunits per turn (Figure 4F). The increased pitch of the three-start helix, from 37.5 Å in the crystallographic fiber to 45 Å in the TCP model, involves an increase in the subunit rise, from 17.9 Å to 22.5 Å, and slight rotation of the subunit clockwise. The increase in rise moves the subunits closer to the center of the fiber relative to the crystallographic fiber, allowing contacts among the hydrophobic α-helical tails (missing in the crystallographic fiber) and a smaller diameter fiber consistent with the EM measurements. The extended conformation of the N-terminal α helices allows them to act as reinforcing rods in the pilus filament, making contact with a maximal number of neighboring subunits (nine in our model) and burying an impressive ~5000 Å^2 of surface area per α helix. This proposed assembly helps to explain the surprising mechanical strength of the pili, which can withstand stress forces of greater than 100 pN (Merz et al., 2000; Maier et al., 2002) despite their narrow diameter.

Functional Implications of Pilin Structure and Assembly

To evaluate the functional implications of our structural results, we mapped the amino acid positions corresponding to well-characterized mutations within the D region that were shown to affect pilus structure and/or function. Previous mutagenesis experiments (Kirn et al., 2000) defined point mutations that affected pilus morphology and assembly as *structural* mutants and those that affected *V. cholerae* autoagglutination, colonization, and transduction efficiency as *interaction* mutants (herein referred to as *functional* mutants to avoid confusion with regions involved in subunit interactions). The structural mutations cluster in the first two thirds of the D region, between Cys120 and Lys165, and also include the second Cys of the disulfide bond (Cys186). In three dimensions, this segment encompasses α3, β4, and β5, which are integral to the globular head domain (Figures 2A–2C, and 5A). Three structural mutations that completely abolished pilus assembly (Cys120→Ser, Lys121→Ala, and Cys186→Ser) (Kirn et al. 2000) form key contacts in the D region of TcpA: the conserved Cys are required for disulfide bond formation; and Lys121 on α3 forms hydrogen bonds with the main chain oxygens of Val182 and Leu185, two residues on α4 involved in hydrophobic interactions between subunits in the TcpA crystallographic fiber (Figures 1 and 4D). Five additional


A
B
C
rise=
17.85 Å
pitch=
37.5 Å
rot=60°
100 Å
D
α4
T177
L185
L176
A59
C186
V182
I179
P58
T125
α3
Y51
K68
L69
C121
K121
L76
L73
G72
α2
β1
G77
α1
α1
E
F
rise=
22.5 Å
pitch
=45 Å
rot=60°
80 Å

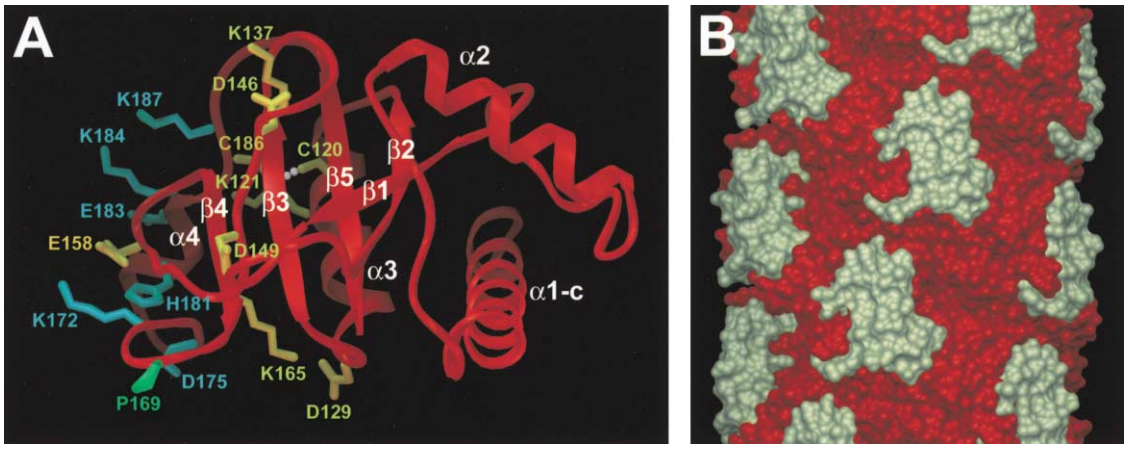


Figure 5. Results of Mutational and Epitope Mapping Studies Are Consistent with the Three-Start Helix TCP Model

(A) Top view of TcpA showing the location of structural (yellow) and functional (blue) residues. Pro69 (green) has both structural and functional properties.

(B) Surface representation of a section of the TCP model showing that overlapping protective epitopes (Sun et al., 1997), colored off-white, map to a large surface-exposed patch on the TcpA subunit.

Ala mutations in the D region alter pilus assembly or morphology (Asp129, Asp146, Asp149, Glu158, and Lys165; Figure 5A). The side chains of these residues form hydrogen bonds with the TcpA main chain, implying a role in folding of the globular head domain, and/or are positioned to interact with neighboring subunits in the TCP model, suggesting a role in TCP assembly. Thus, mutation of "structural" residues evidently alters pilus assembly and morphology by destabilizing the D region and disrupting interactions between subunits and/or by disrupting the TcpA fold.

Significantly, the functional mutations, which cluster in the last third of the D region and are believed to be surface exposed (Kirn et al. 2000), are localized to pilin subunit interface regions but are exposed for pilus interactions. These residues (Lys172, Asp175, His181, Glu183, Lys184, Lys187; Figure 5A) are on and adjacent to α4, which is involved in subunit interactions in the crystallographic fiber and in our TCP model. Yet these charged residues are distinct from the hydrophobic residues involved in subunit interactions in that they extend outward to the surface of the TCP filament, making minimal contact with neighboring subunits. This is in contrast to the hydrophobic residues, which form a buried cluster within the subunit:subunit interface of the three-start helical strands in both the TcpA crystallographic fiber and the TCP model (Figure 4D). Thus, the functional domain of the D region contains residues important for pilus function as well as residues involved in pilus assembly. Mutation of the exposed functional residues could alter the surface properties of the pilus fiber and prevent interactions with other pili, bacteriophage, or host cells. The locations of specific residues within the globular head of TcpA and within the TCP filament model therefore explain and extend our understanding of their roles implied by the mutational analyses.

To further test the functional implications of our structural results, we mapped the epitopes for three protective TCP antibodies (Sun et al., 1990, 1991). These overlapping epitopes, which comprise residues 145–168, 157–182, and 174–199, form a patch on the TCP surface that represents ~33% of the surface area and encompasses all of the functional domain residues (Figure 5B). Interestingly, many of the exposed residues in each of the protective epitopes are hydrophobic, with a high Ala content, as is much of the TCP surface, which suggests that antibody:pilus and pilus:pilus interactions are primarily entropically driven. Thus, mutation of a single

Figure 4. TcpA Crystal Lattice and Structure-Based Model of TCP Filament

(A) Crystal lattice showing TcpA subunits arranged in hexagonally packed fibers. The three molecules in the asymmetric unit are colored red and appear at different levels.

(B) A single crystallographic fiber showing that the N-terminal α helices face the center of the fiber and have the same polarity. The three strands of the helical fiber are colored red, blue, and yellow. Two subunits are shown for each start, with the second one rotated 60° counterclockwise relative to the first and translated 17.85 Å along the fiber axis (i.e., into the page).

(C) Side view of a single crystallographic fiber showing the left-handed three-start helix, fiber dimensions, and helical symmetry. Six subunits are shown in one complete turn for each helical strand of the three-start helix assembly.

(D) Hydrophobic interface connecting subunits within each strand of the left-handed three-start helices. Side chains on α2 of one subunit are shown as yellow ball-and-stick representations on an orange ribbon, and side chains on α3 and α4 are shown in cyan on a green ribbon. Relevant oxygen atoms are colored red, and nitrogens are blue. In addition to the hydrophobic interactions, two hydrogen bonds link the subunits (Tyr51:OH to Leu176:O, and Thr125:OH to Leu76:O) as indicated by white spheres. The disulfide-bound cysteines are colored pink with yellow sulfur atoms.

(E and F) Top view (E) and side view (F) of the structure-based TCP model derived from the symmetry determined by EM analysis and the packing arrangement seen in the crystallographic fibers. An extended α-helical tail has been added to the N terminus using the coordinates of the PAK pilin α1-N. The dimensions are shown for comparison with the crystallographic fiber in (C).

charged residue on the fiber surface could dramatically affect the surface character and, hence, the ability of the pilus to autoagglutinate and form microcolonies.

Conclusions

Our detailed structural analyses of *V. cholerae* TCP and PAK pili have (1) identified a novel protein fold for the type IVb pilin TcpA; (2) established the extended N-terminal α helix and central α/β scaffold as conserved features of the type IV pilins; and (3) suggested a TCP filament architecture whereby subunits associate via hydrophobic interactions of the N-terminal α helices as well as between the variable αβ loop and D regions of the globular head domains. The structure-based TCP assembly model is generalizable for all type IV pili, with the conserved α helices forming the central hydrophobic core of the filament and maximizing subunit:subunit contacts, and the structurally variable regions forming specific subunit interactions and resulting in architectural variability among the different pili. In addition to mediating head domain interactions, the αβ loop and D region determine, to a large degree, the surface properties of the pilus filament, which are specifically adapted to the host cell environment and contribute to pathogen-specific biological functions such as motility, antigenicity, autoagglutination, and adhesion. Thus, conserved structural features of the type IV pilins produce the characteristic pilus thread-like appearance that combines strength and flexibility, while variable features impose distinct assembly architectures and in vivo functions.

Experimental Procedures

Cloning, Protein Expression, and Purification of TcpA

A detailed protocol is provided in the Supplemental Materials (at http://www.molecule.org/cgi/content/full/11/5/1139/DC1). Briefly, DNA encoding TcpA residues 29–199 was PCR-amplified from *V. cholerae* strain RT4236 DNA and ligated into the pCR 2.1-TOPO vector (Invitrogen). The TOPO vector containing the TcpA insert was amplified in OneShot Top 10 cells (Invitrogen), purified, digested, and ligated into pET-15b (Novagen) downstream of the N-terminal His-Tag. The resulting pET-15b-TcpA vector was transformed into *E. coli* Origami cells (Novagen), and protein expression was induced with isopropyl β-D-thiogalactopyranoside (IPTG, final concentration 0.4 mM) followed by overnight growth. TcpA was purified from the cell sonication supernatant by cation exchange chromatography and gel filtration and was concentrated using a Centriprep concentrator (Millipore) to 15 mg/ml in 50 mM HEPES (pH 7.0), 50 mM NaCl, 1 mM EDTA, and 1 mM β-mercaptoethanol. Selenomethionine-labeled TcpA (SeMet-TcpA) was prepared using the method of Van Duyne et al. (1993) and is described in Supplemental Materials (at http://www.molecule.org/cgi/content/full/11/5/1139/DC1).

Purification of PAK Pilin

PAK pili were expressed and purified from thin lawns of *P. aeruginosa* 2pfs cells (R. Irvin, U. Alberta) as described in Supplemental Materials (at http://www.molecule.org/cgi/content/full/11/5/1139/DC1). Pili were dissociated into pilin subunits by extensive filtration on an Amicon XM50 membrane in 100 mM Tris (pH 8.0), 20 mM NaCl, 1 mM DTT, and 1.5% *n*-octyl-β-D-glucopyranoside (βOG), whereby the pilin remained in the retentate.

Crystallization, Crystallographic Data Collection, Structure Determination, and Refinement of TcpA

TcpA crystals grew by sitting drop vapor diffusion by mixing 2 μl TcpA (15 mg/ml) with 2 μl reservoir solution (30% PEG 8000, 0.1 M imidazole [pH 8.0], 0.2 M NaCl, 6% glycerol) at room temperature. Crystals diffracted to 1.3 Å at beamline 11-1 at the Stanford Synchrotron Radiation Laboratories (SSRL). Data were processed using MOSFLM (Leslie, 1992). Experimental phases were obtained by multiwavelength anomalous dispersion (MAD) data from a single crystal of SeMet-TcpA collected at three wavelengths near the selenium absorption edge (Table 1). The SeMet-TcpA crystal diffracted to 1.8 Å on SSRL beamline 9.2. MAD data were processed using DENZO and SCALEPACK (Otwinowski and Minor, 1997).

The four expected selenium sites in SeMet-TcpA were located using SOLVE (Terwilliger and Berendzen, 1999) and used to calculate phases using SHARP (de la Fortelle and Bricogne, 1999). An interpretable electron density map was obtained after solvent flattening. The sequence of TcpA was fit manually using XFIT (McRee, 1999), and the resulting model was refined using CNS (Brünger et al., 1998) and XFIT. The refined model was used to solve the 1.3 Å crystal structure of TcpA by molecular replacement using AMoRe (Navaza, 2001). After refinement with SHELX-97 (Sheldrick, 1990), the final model contains 785 well-ordered water molecules and three glycerol molecules. PROCHECK analysis (Laskowski et al., 1993) revealed that 93% of all residues are in the most favored regions of the Ramachandron plot, and the remaining 7% are in allowed regions.

Crystallization, Crystallographic Data Collection, Structure Determination, and Refinement of PAK Pilin

PAK pilin crystals were grown in two space groups by sitting drop vapor diffusion by mixing 2.5 μl PAK pilin (15 mg/ml) with 2.5 μl reservoir buffer (15% methylpentanediol, 35% PEG 4000, 100 mM sodium citrate [pH 5.6], and 2.5 mM $MnCl_2$ for C2 crystals; 10% methylpentanediol, 35% PEG 4000, and 100 mM sodium citrate [pH 5.6] for $P2_12_12_1$ crystals). Data were processed using DENZO and SCALEPACK (Otwinowski and Minor, 1997), and the structures were solved by molecular replacement with the N-terminally truncated PAK pilin structure (Hazes et al., 2000) using AMoRe (Navaza, 2001). The structures were refined using CNS (Brünger et al., 1998) and manually fit using XFIT (McRee, 1999). As both structures are virtually identical, the higher resolution (2.0 Å) C2 structure is shown here. The final model contains 282 well-ordered water molecules. PROCHECK (Laskowski et al., 1993) indicated that 95% of the residues are in the most favored regions of the Ramachandron plot, with the remaining 5% in allowed regions.

Electron Microscopy and Image Analysis of TCP Filaments

Cryo-electron Microscopy

V. cholerae cells (strain RT4236) were grown overnight in TCP-expressing conditions (Kirn et al., 2000). Aliquots of culture (5 μl) were removed and applied to Quantifoil holey grids (Jena, Germany) that had been glow-discharged in the presence of amyl amine. After 1 min, grids were blotted for 5 s and then plunged into an ethane slush cooled with liquid nitrogen. Grids were transferred to a Gatan 626 cold stage, and images were recorded on a Philips/FEI CM120 electron microscope operating at 100 kEv in low dose mode at 45,000× magnification and a defocus of 1 μm.

Negatively Stained EM

Aliquots of *V. cholerae* RT4236 cultures (5 μl) were applied to Formvar-coated, glow-discharged copper grids. After 1 min, the sample was wicked off, and the grid was floated on a drop of 1% ammonium molybdate (pH 7.5) for 1 min, blotted, and air dried. Images were recorded as described above at a defocus of 0.5 μm.

Helical Image Analysis and Generation of TCP Three-Start Helix Assembly

Micrographs were digitized using a SCAI scanner (Zeiss), and Fourier transforms were computed from images of individual filaments using the MRC package (Crowther et al., 1996). The Bessel order, J_n, or "start" of the helix was determined using the relationship $2\pi Rr = n$, where R is the distance in reciprocal space between the meridian and the peak of the layer line, r is the outer radius of the filament, and n is a fixed value corresponding to the Bessel order. For the frozen-hydrated filaments, R = 1/58 Å, and r = 39 Å, consistent with the 40 Å radius for TCP filaments. This yields a Bessel order of J_3 or a *three-start helix* (Figure 3A). For negatively stained TCP, the layer line peak is further from the meridian at R = 1/40 Å, giving a radius of 27 Å for Bessel order J_3 (Figure 3B). Rather than being an overlapping Bessel order, we believe that this layer line is

a secondary J_3 layer line, which results from negative stain pooled inside the fiber. This would highlight features of the three-start helix that are not at the maximum radius. Thus, the 1/45 Å layer line indicates a three-start helix with 45 Å pitch for TCP. To derive TCP models having a dominant three-start helix assembly with a pitch of ~45 Å and a diameter of ~80 Å, ~23 million fiber models were computationally generated and evaluated as described in the Supplemental Materials (at http://www.molecule.org/cgi/content/full/11/5/1139/DC1).

Acknowledgments

We thank Carey Kassmann and Anchi Cheng for technical advice, and the staff at beamlines 9-2 and 11-1 at SSRL for assistance in data collection. We thank Atsushi Yamagata, Michael DiDonato, Ronald Brudler, Edward Egelman, Ronald Milligan, and Thomas Kirn for comments. This work was supported by NIH grants AI22160 (J.A.T.), AI31535 (M.Y.), and GM59721 (K.T.F.), a fellowship from The Canadian Institutes of Health Research (L.C.), and by The American Heart Association, Bristol-Myers Squibb, and The Burroughs Wellcome Fund (M.Y).

Received: December 18, 2002
Revised: February 24, 2003
Accepted: March 6, 2003
Published: May 22, 2003

References

Blank, T.E., Zhong, H., Bell, A.L., Whittam, T.S., and Donnenberg, M.S. (2000). Molecular variation among type IV pilin (bfpA) genes from diverse enteropathogenic *Escherichia coli* strains. Infect. Immun. *68*, 7028–7038.

Boyd, E.F., and Waldor, M.K. (2002). Evolutionary and functional analyses of variants of the toxin-coregulated pilus protein TcpA from toxigenic *Vibrio cholerae* non-O1/non-O139 serogroup isolates. Microbiology *148*, 1655–1666.

Bradley, D.E. (1980). A function of *Pseudomonas aeruginosa* PAO polar pili: twitching motility. Can. J. Microbiol. *26*, 146–154.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Chattopadhyaya, R., and Ghose, A.C. (2002). Model of *Vibrio cholerae* toxin coregulated pilin capable of filament formation. Protein Eng. *15*, 297–304.

Chiang, S.L., Taylor, R.K., Koomey, M., and Mekalanos, J.J. (1995). Single amino acid substitutions in the N-terminus of *Vibrio cholerae* TcpA affect colonization, autoagglutination, and serum resistance. Mol. Microbiol. *17*, 1133–1142.

Crowther, R.A., Henderson, R., and Smith, J.M. (1996). MRC image processing programs. J. Struct. Biol. *116*, 9–16.

de la Fortelle, E., and Bricogne, G. (1999). Maximum likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. Methods Enzymol. *276*, 472–493.

Elleman, T.C., and Peterson, J.E. (1987). Expression of multiple types of N-methyl Phe pili in Pseudomonas aeruginosa. Mol. Microbiol. *1*, 377–380.

Folkhard, W., Marvin, D.A., Watts, T.H., and Paranchych, W. (1981). Structure of polar pili from *Pseudomonas aeruginosa* strains K and O. J. Mol. Biol. *149*, 79–93.

Forest, K.T., Dunham, S.A., Koomey, M., and Tainer, J.A. (1999). Crystallographic structure reveals phosphorylated pilin from *Neisseria*: phosphoserine sites modify type IV pilus surface chemistry and fibre morphology. Mol. Microbiol. *31*, 743–752.

Giron, J.A., Ho, A.S., and Schoolnik, G.K. (1991). An inducible bundle-forming pilus of enteropathogenic *Escherichia coli*. Science *254*, 710–713.

Giron, J.A., Gomez-Duarte, O.G., Jarvis, K.G., and Kaper, J.B. (1997). Longus pilus on enterotoxigenic *Escherichia coli* and its relatedness to other type-4 pili—a minireview. Gene *192*, 39–43.

Hazes, B., Sastry, P.A., Hayakawa, K., Read, R.J., and Irvin, R.T. (2000). Crystal structure of *Pseudomonas aeruginosa* PAK pilin suggests a main-chain-dominated mode of receptor binding. J. Mol. Biol. *299*, 1005–1017.

Herrington, D.A., Hall, R.H., Losonsky, G., Mekalanos, J.J., Taylor, R.K., and Levine, M.M. (1988). Toxin, toxin-coregulated pili, and the toxR regulon are essential for *Vibrio cholerae* pathogenesis in humans. J. Exp. Med. *168*, 1487–1492.

Holm, L., and Sander, C. (1995). Dali: a network tool for protein structure comparison. Trends Biochem. Sci. *20*, 478–480.

Irvin, R.T., Doig, P., Lee, K.K., Sastry, P.A., Paranchych, W., Todd, T., and Hodges, R.S. (1989). Characterization of the *Pseudomonas aeruginosa* pilus adhesin: confirmation that the pilin structural protein subunit contains a human epithelial cell-binding domain. Infect. Immun. *57*, 3720–3726.

Kaper, J.B., Morris, J.G., Jr., and Levine, M.M. (1995). Cholera. Clin. Microbiol. Rev. *8*, 48–86.

Karaolis, D.K., Somara, S., Maneval, D.R., Jr., Johnson, J.A., and Kaper, J.B. (1999). A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. Nature *399*, 375–379.

Keizer, D.W., Slupsky, C.M., Kalisiak, M., Campbell, A.P., Crump, M.P., Sastry, P.A., Hazes, B., Irvin, R.T., and Sykes, B.D. (2001). Structure of a pilin monomer from *Pseudomonas aeruginosa*: implications for the assembly of pili. J. Biol. Chem. *276*, 24186–24193.

Kirn, T.J., Lafferty, M.J., Sandoe, C.M., and Taylor, R.K. (2000). Delineation of pilin domains required for bacterial association into microcolonies and intestinal colonization by *Vibrio cholerae*. Mol. Microbiol. *35*, 896–910.

Laskowski, R.A., McArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.

Lee, K.K., Doig, P., Irvin, R.T., Paranchych, W., and Hodges, R.S. (1989). Mapping the surface regions of *Pseudomonas aeruginosa* PAK pilin: the importance of the C-terminal region for adherence to human buccal epithelial cells. Mol. Microbiol. *3*, 1493–1499.

Lee, K.K., Sheth, H.B., Wong, W.Y., Sherburne, R., Paranchych, W., Hodges, R.S., Lingwood, C.A., Krivan, H., and Irvin, R.T. (1994). The binding of *Pseudomonas aeruginosa* pili to glycosphingolipids is a tip-associated event involving the C-terminal region of the structural pilin subunit. Mol. Microbiol. *11*, 705–713.

Leslie, A.G.W. (1992). Joint CCP4 and EACMB Newsletter Protein Crystallography, 26 (Warrington, UK: Daresbury Laboratory).

Macdonald, D.L., Pasloske, B.L., and Paranchych, W. (1993). Mutations in the fifth-position glutamate in *Pseudomonas aeruginosa* pilin affect the transmethylation of the N-terminal phenylalanine. Can. J. Microbiol. *39*, 500–505.

Maier, B., Potter, L., So, M., Seifert, H.S., and Sheetz, M.P. (2002). Single pilus motor forces exceed 100 pN. Proc. Natl. Acad. Sci. USA *99*, 16012–16017.

McRee, D.E. (1999). XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. J. Struct. Biol. *125*, 156–165.

Merz, A.J., and So, M. (2000). Interactions of pathogenic *Neisseriae* with epithelial cell membranes. Annu. Rev. Cell Dev. Biol. *16*, 423–457.

Merz, A.J., So, M., and Sheetz, M.P. (2000). Pilus retraction powers bacterial twitching motility. Nature *407*, 98–102.

Mu, X.Q., Egelman, E.H., and Bullitt, E. (2002). Structure and function of Hib pili from *Haemophilus influenzae* type b. J. Bacteriol. *184*, 4868–4874.

Nassif, X., Beretti, J.L., Lowy, J., Stenberg, P., O'Gaora, P., Pfeifer, J., Normark, S., and So, M. (1994). Roles of pilin and PilC in adhesion of *Neisseria meningitidis* to human epithelial and endothelial cells. Proc. Natl. Acad. Sci. USA *91*, 3769–3773.

Navaza, J. (2001). Implementation of molecular replacement in AMoRe. Acta Crystallogr. D Biol. Crystallogr. *57*, 1367–1372.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Parge, H.E., Forest, K.T., Hickey, M.J., Christensen, D.A., Getzoff, E.D., and Tainer, J.A. (1995). Structure of the fibre-forming protein pilin at 2.6 A resolution. Nature *378*, 32–38.

Park, H.S., Wolfgang, M., van Putten, J.P., Dorward, D., Hayes, S.F., and Koomey, M. (2001). Structural alterations in a type IV pilus subunit protein result in concurrent defects in multicellular behaviour and adherence to host tissue. Mol. Microbiol. *42*, 293–307.

Pasloske, B.L., Scraba, D.G., and Paranchych, W. (1989). Assembly of mutant pilins in *Pseudomonas aeruginosa*: formation of pili composed of heterologous subunits. J. Bacteriol. *171*, 2142–2147.

Seifert, H., and So, M. (1984). Genetic mechanisms of bacterial antigenic variation. Microbiol. Rev. *52*, 327–336.

Shaw, C.E., and Taylor, R.K. (1990). Vibrio cholerae 0395 tcpA pilin gene sequence and comparison of predicted protein structural features to those of type 4 pilins. Infect. Immun. *58*, 3042–3049.

Sheldrick, G.M. (1990). Phase annealing in SHELX-90: direct methods for larger structures. Acta Crystallogr. A *46*, 467–473.

Strom, M.S., and Lory, S. (1991). Amino acid substitutions in pilin of *Pseudomonas aeruginosa*. Effect on leader peptide cleavage, amino-terminal methylation, and pilus assembly. J. Biol. Chem. *266*, 1656–1664.

Strom, M.S., and Lory, S. (1993). Structure-function and biogenesis of the type IV pili. Annu. Rev. Microbiol. *47*, 565–596.

Sun, D.X., Mekalanos, J.J., and Taylor, R.K. (1990). Antibodies directed against the toxin-coregulated pilus isolated from *Vibrio cholerae* provide protection in the infant mouse experimental cholera model. J. Infect. Dis. *161*, 1231–1236.

Sun, D., Seyer, J.M., Kovari, I., Siumrada, R.A., and Taylor, R.K. (1991). Localization of protective epitopes within the pilin subunit of the *Vibrio cholerae* toxin-coregulated pilus. Infect. Immun. *59*, 114–118.

Sun, D., Lafferty, M.J., Peek, J.A., and Taylor, R.K. (1997). Domains within the *Vibrio cholerae* toxin coregulated pilin subunit that mediate bacterial colonization. Gene *192*, 79–85.

Swanson, J. (1983). Gonococcal adherence: selected topics. Rev. Infect. Dis. Suppl. *5*, S678–S684.

Tacket, C.O., Taylor, R.K., Losonsky, G., Lim, Y., Nataro, J.P., Kaper, J.B., and Levine, M.M. (1998). Investigation of the roles of toxin-coregulated pili and mannose-sensitive hemagglutinin pili in the pathogenesis of *Vibrio cholerae* O139 infection. Infect. Immun. *66*, 692–695.

Taniguchi, T., Fujino, Y., Yamamoto, K., Miwatani, T., and Honda, T. (1995). Sequencing of the gene encoding the major pilin of pilus colonization factor antigen III (CFA/III) of human enterotoxigenic *Escherichia coli* and evidence that CFA/III is related to type IV pili. Infect. Immun. *63*, 724–728.

Taylor, R.K., Miller, V.L., Furlong, D.B., and Mekalanos, J.J. (1987). Use of phoA gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. Proc. Natl. Acad. Sci. USA *84*, 2833–2837.

Terwilliger, T.C., and Berendzen, J. (1999). Automated MAD and MIR structure solution. Acta Crystallogr. D *55*, 849–861.

Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L., and Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 Complexes with FK506 and Rapamycin. J. Mol. Biol. *229*, 105–124.

Virji, M., and Heckels, J.E. (1983). Antigenic cross-reactivity of *Neisseria* pili: investigations with type- and species-specific monoclonal antibodies. J. Gen. Microbiol. *129*, 2761–2768.

Waldor, M.K., and Mekalanos, J.J. (1996). Lysogenic conversion by a filamentous phage encoding cholera toxin. Science *272*, 1910–1914.

Watts, T.H., Kay, C.M., and Paranchych, W. (1983). Spectral properties of three quaternary arrangements of *Pseudomonas* pilin. Biochemistry *22*, 3640–3646.

Woods, D.E., Straus, D.C., Johanson, W.G., Jr., Berry, V.K., and Bass, J.A. (1980). Role of pili in adherence of *Pseudomonas aeruginosa* to mammalian buccal epithelial cells. Infect. Immun. *29*, 1146–1151.

Zhang, X.L., Tsui, I.S., Yip, C.M., Fung, A.W., Wong, D.K., Dai, X., Yang, Y., Hackett, J., and Morris, C. (2000). *Salmonella enterica* serovar typhi uses type IVB pili to enter human intestinal epithelial cells. Infect. Immun. *68*, 3067–3073.

Accession Numbers

The structures described in this paper have been deposited in the Protein Data Bank with accession numbers 1OQV (TcpA), 1OQW (PAK pilin), and 1OR9 (TCP model).