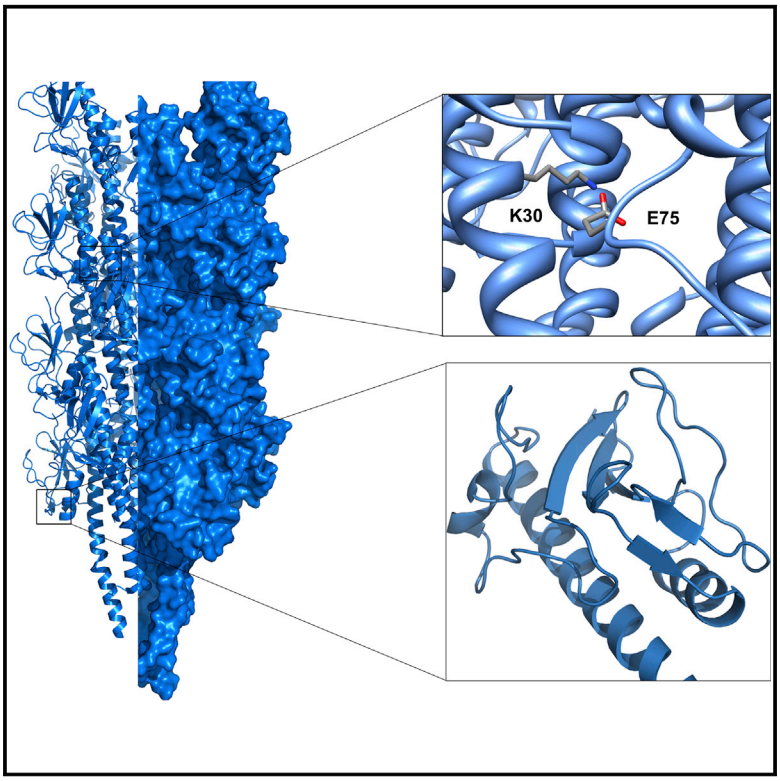
Article

Structure

Structural and Evolutionary Analyses Show Unique Stabilization Strategies in the Type IV Pili of *Clostridium difficile*

Graphical Abstract



Authors

Kurt H. Piepenbrink,
Grace A. Maldarelli, ...,
Michael S. Donnenberg, Eric J. Sundberg

Correspondence

esundberg@ihv.umaryland.edu

In Brief

Piepenbrink et al. describe the structure of PilA1, the major pilin of the type IV pili of *Clostridium difficile* from three distinct strains. The fold of PilA1 shows a resemblance to type IVb pilins as well as pseudopilins from type II secretion systems.

Highlights

- PilA1 is the major pilin of the type IV pili (T4P) in *C. difficile*
- *C. difficile* strains show distinct stabilization strategies in the fold of PilA1
- PilA1 shows structural similarity to the type IVb pili of Gram-negative bacteria
- Both PilA1 and PilJ are incorporated into T4P through an intersubunit salt bridge

Accession Numbers

4IXJ
4TSM
4OGM
4PE2

Piepenbrink et al., 2015, Structure *23*, 385–396
February 3, 2015
http://dx.doi.org/10.1016/j.str.2014.11.018

CellPress

Structure
Article

Structural and Evolutionary Analyses Show Unique Stabilization Strategies in the Type IV Pili of *Clostridium difficile*

Kurt H. Piepenbrink,[1] Grace A. Maldarelli,[2] Claudia F. Martinez de la Peña,[4] Tanis C. Dingle,[4] George L. Mulvey,[4] Amanda Lee,[4] Erik von Rosenvinge,[2] Glen D. Armstrong,[4] Michael S. Donnenberg,[2,3] and Eric J. Sundberg[1,2,3,*]
[1]Institute of Human Virology
[2]Department of Medicine
[3]Department of Microbiology and Immunology
University of Maryland School of Medicine, Baltimore, MD 21201, USA
[4]Department of Microbiology, Immunology and Infectious Disease, University of Calgary, Calgary, AB T2N 1N4, Canada
*Correspondence: esundberg@ihv.umaryland.edu
http://dx.doi.org/10.1016/j.str.2014.11.018

SUMMARY

Type IV pili are produced by many pathogenic Gram-negative bacteria and are important for processes as diverse as twitching motility, biofilm formation, cellular adhesion, and horizontal gene transfer. However, many Gram-positive species, including *Clostridium difficile*, also produce type IV pili. Here, we identify the major subunit of the type IV pili of *C. difficile*, PilA1, and describe multiple 3D structures of PilA1, demonstrating the diversity found in three strains of *C. difficile*. We also model the incorporation of both PilA1 and a minor pilin, PilJ, into the pilus fiber. Although PilA1 contains no cysteine residues, and therefore cannot form the disulfide bonds found in all Gram-negative type IV pilins, it adopts unique strategies to achieve a typical pilin fold. The structures of PilA1 and PilJ exhibit similarities with the type IVb pilins from Gram-negative bacteria that suggest that the type IV pili of *C. difficile* are involved in microcolony formation.

INTRODUCTION

Type IV pili are extracellular appendages that have been found in a variety of bacteria, both Gram-positive and Gram-negative, including *Pseudomonas aeruginosa*, *Vibrio cholerae*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Salmonella enterica* serovar Typhi, *Legionella pneumophila*, enteropathogenic and enterotoxigenic *Escherichia coli, Streptococcus pneumoniae, Clostridium perfringens*, and *Clostridium difficile* (Giron et al., 1991; Goulding et al., 2009; Laurenceau et al., 2013; Lee et al., 1994; Stone and Abu Kwaik, 1998; Strom and Lory, 1993; Taniguchi et al., 1995; Varga et al., 2006; Zhang et al., 2000). They are utilized in a variety of adhesive processes such as biofilm formation (O'Toole and Kolter, 1998), cellular adhesion (Rudel et al., 1995), twitching motility (Bradley, 1980; Henrichsen, 1983; Merz et al., 2000; Wall and Kaiser, 1999), and horizontal gene transfer (Seifert et al., 1988; Yoshida et al., 1999). In many cases, these properties are essential for virulence (Bieber et al., 1998; Herrington et al., 1988; Tacket et al., 1998). These fimbrial appendages are composed of many pilin proteins tightly packed in a helix to bury the hydrophobic amino terminus of each subunit in the pilus core (Craig et al., 2004). Typically, a single pilin protein, referred to as the major pilin, has been found to predominate and the other pilin proteins included as subunits are termed minor pilins. In type II secretion, the equivalent subunits are referred to as major and minor pseudopilins.

All type IV pilins share certain structural features, namely a prepilin leader sequence, an N-terminal α helix, which is hydrophobic for the first ~60 amino acids, and a central β sheet. The pili they form, however, diverge based on the details of their structures, particularly in the $\alpha\beta$ loop and at the C terminus. Nearly all pilins from Gram-negative bacteria have a disulfide bond that bounds a portion of the C terminus referred to as the D region. Based on the length of the pre-pilin leader peptide and sequence similarity of the N-terminal α helix, among other factors, type IV pili have been divided into two classes, type IVa and type IVb pilins. type IVa pilins occur in a wide range of Gram-negative bacteria while type IVb pilins have been found only in a subset capable of colonizing the intestine and are larger, particularly in the D region (Craig et al., 2004).

Type IV pili are also produced by several Gram-positive species (Laurenceau et al., 2013; Melville and Craig, 2013), although they are only sparsely characterized as yet. Genes encoding type IV pilins and pilus biogenesis proteins have been found in all of the Clostridiales and type IV pili have been observed in several *Clostridium* species, including *C. perfringens*, where they are required for gliding motility (Varga et al., 2006), *Ruminococcus albus* (Rakotoarivonina et al., 2002), and *C. difficile* (Borriello et al., 1988; Goulding et al., 2009; Maldarelli et al., 2014; Piepenbrink et al., 2014). Because the *C. difficile* type IV pili represent an opportunity to better understand both the diversity of type IV pili and the survival strategy of *C. difficile*, we have sought to characterize the type IV pili from this Gram-positive species.

Previously, we reported the structure of PilJ, the first structure of a type IV pilin from a Gram-positive organism (Piepenbrink et al., 2014). We found that PilJ has a unique fold compared

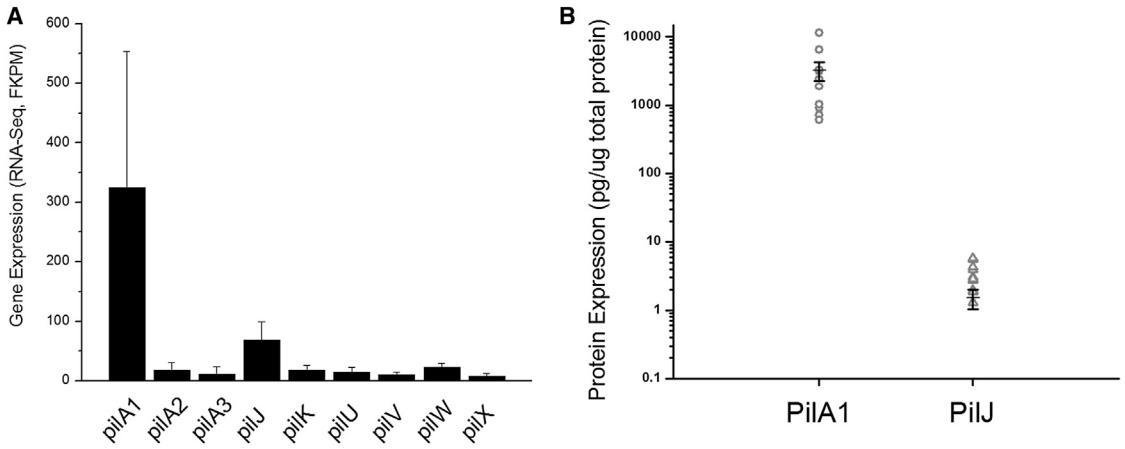


Figure 1. Expression of PilA1 and PilJ

(A) mRNA expression of putative pilin genes in *C. difficile*. Average mRNA levels are shown for each protein, with error bars indicating the standard deviation. (B) Protein expression of PilA1 and PilJ in *C. difficile*. Expression levels from multiple independent experiments are plotted on a logarithmic scale with the error bars indicating the standard error.

with pilin structures characterized previously, containing two pilin-like domains joined by a Cys-Cys-Cys-His-coordinated zinc atom. While immunogold staining showed that PilJ is incorporated into the type IV pilus of *C. difficile*, models of pilus fragments composed entirely of PilJ were significantly wider than the observed *C. difficile* type IV pili, indicating that PilJ is a minor pilin.

Another of the type IV pilin proteins of *C. difficile*, PilA1, is similar in size (20 kDa, 164 residues for the mature protein) to the majority of previously characterized pilins. Due to its position in the primary type IV pilus operon in the *C. difficile* genome, as well as its high degree of sequence diversity among different *C. difficile* genomes, we and others have predicted that it was the major pilin for some, if not all, of the type IV pili expressed by *C. difficile* (Maldarelli et al., 2014; Melville and Craig, 2013; Piepenbrink et al., 2014). Here, we show that PilA1 is the major pilin for all observable *C. difficile* type IV pili. We also present the X-ray crystal structures of PilA1 pilins from three distinct strains, revealing how this protein has found more than one unique structural solution to adopting the common type IV pilin fold. With multiple independent PilA1 structures, we show that the genetic variations in *pilA1* map to surface-exposed regions in a modeled *C. difficile* type IVb pilus.

RESULTS

PilA1 Is Required for and Composes the Majority of *C. difficile* Type IV Pili

We reported previously that *C. difficile* produces type IV pili that incorporate both PilA1 and PilJ (Piepenbrink et al., 2014). Immunogold labeling revealed that several pilus fibers stained with immunogold particles corresponding to both pilins. Models of pili based on our PilJ crystal structure, however, produced pilus widths that were inconsistent with those observed by electron microscopy. Based on its size and position in the primary cluster of type IV pilus biogenesis genes in the *C. difficile* genome, we hypothesized that PilA1 was the major pilin, composing the majority of these pili (Maldarelli et al., 2014).

By definition, the major pilin should predominate in the pili. To test the hypothesis that PilA1 is the major pilin, we used RNA-Seq to measure the mRNA levels of all *C. difficile* pilin genes. As shown in Figure 1A, *pilA1* is the most highly expressed gene, several-fold more highly expressed than the next highest transcript, *pilJ*. In addition, we used quantitative immunoblotting to measure the amount of PilA1 and PilJ protein in whole cells. These measurements, shown in Figure 1B, indicate that when grown on blood agar plates, *C. difficile* expresses 3.23 ± 1.02 ng of PilA1 per microgram of total protein and 1.53 ± 0.485 pg of PilJ per microgram. This corresponds to a ratio of ~2000 molecules of PilA1 per molecule of PilJ.

To confirm these findings, we used ClosTron mutagenesis to engineer a variant strain of *C. difficile* R20291 deficient in PilA1. As shown in Figure 2A, numerous pili from the wild-type strain are stained with gold nanoparticles for both pilins. Conversely, the Δ*pilA1* strain produces no visible pili under conditions known to induce piliation in the wild-type strain. Quantification of the immunogold staining demonstrates the absence of both PilA1 and PilJ on the surface of *C. difficile* in the Δ*pilA1* mutant strain. The reintroduction of *pilA1* through a complementation plasmid (Dingle et al., 2011; Heap et al., 2009) restores piliation and, importantly, these pili incorporate both PilA1 and PilJ, as shown in Figure 2B.

Based on the sum of these factors, (1) PilA1 is expressed at a much higher level than PilJ, (2) the loss of PilA1 eliminates visible piliation, and (3) the restoration of PilA1 reincorporates PilJ into T4P, we assign PilA1 as the major pilin of all of the *C. difficile* type IV pili that we have observed. Whether *C. difficile* can produce T4P with a different composition under other conditions remains to be seen.

The Structure of PilA1 Shows Remarkable Conservation with Type IVb Gram-Negative Pilins

We determined the 1.9 Å resolution X-ray crystal structure of soluble PilA1 from *C. difficile* R20291 (ribotype 027) fused to the C terminus of maltose-binding protein (Table 1). PilA1 exhibits the typical pilin fold: an N-terminal α helix, followed by a loop

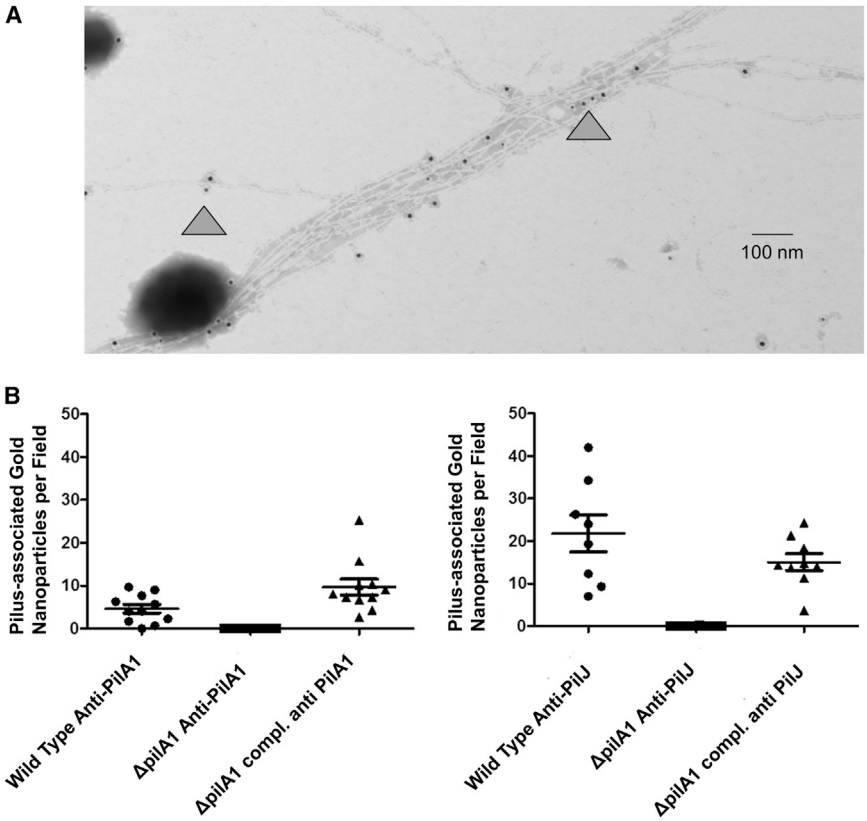


Figure 2. Type IV Pilus Expression in *Clostridium difficile*

(A) Immunogold electron microscopy of *C. difficile*. Electron micrographs were stained with anti-PilA1 and anti-PilJ antibodies and immunogold-labeled secondary antibodies with particle sizes 10 nm (PilA1) and 15 nm (PilJ). Colocalized 10- and 15-nm particles are indicated with triangles.

(B) Quantification of immunogold labeling. The average number of particles per field is shown as a horizontal line. The error bars show one standard deviation.

with some α-helical character (the α2 helix) and then by a four-stranded antiparallel β sheet (Figure 3). It contains an additional α helix (which we denote the α3 helix) at the C terminus and, more interestingly, a two-stranded antiparallel β sheet between the α3 helix and the C-terminal strand of the central β sheet (which we denote the B2 sheet).

The presence of a second β sheet is not unprecedented in the structures of type IV pilins, as the PAK pilin of *P. aeruginosa* contains a two-strand parallel β sheet in its αβ loop and the GC pilin of *N. gonorrhoeae* contains a β hairpin in addition to the central β sheet (Craig et al., 2003, 2006). However, these β structures are parallel to the central β sheet with respect to the N-terminal α helix. In contrast, the B2 sheet of PilA1 rests below the central β sheet at a 90° angle to it, in an analogous position to the disulfide bond found in the vast majority of type IV pilins from Gram-negative strains (including *N. gonorrhoeae* and *P. aeruginosa*) (Figure 3). Notably, the strands of the B2 sheet are placed at the midpoints of the β2-β4 and β4-β3 insertions in the central β sheet (Figure 3B). We propose that the B2 sheet of PilA1 serves primarily to stabilize the C terminus of the protein, in much the same way that the disulfide bond does in most Gram-negative pilins.

A superimposition of PilA1 with the two domains of PilJ (Figure S1 available online) reveals considerable conservation in the N-terminal half; the α1-C helices, αβ loops, and the first two strands of each β sheet overlay well (2.5 Å root-mean-square deviation [rmsd]). However, the C-terminal portions of PilA1 show no obvious similarity to those of either of the PilJ domains.

PilA1 and PilJ Show Significantly Different Levels of Conservation

A comparison of the sequences of PilA1 and PilJ in the genomes of a large collection of *C. difficile* strains confirms our prior observation that PilA1 is much less conserved than any of the putative minor pilins (Maldarelli et al., 2014), including PilJ. Here, we compared all of the publically available genomes of *C. difficile*, encompassing 238 strains, and found that the average variation is considerably greater in PilA1 than in PilJ (Figure 4). Furthermore, using the larger data set available with newly sequenced genomes, we confirm our prior observation that residues of PilA1 are subject to diversifying selection ($p < 0.01$) (Maldarelli et al., 2014). The evidence for diversifying selection in PilJ is less clear, with a p value between 0.05 and 0.01.

Figure 4 shows the regions of conservation for PilA1 and PilJ. In both cases, the areas expected to be involved in pilus formation are conserved. In PilA1, the regions of largest variability are the loops between the strands of the β sheet, which vary in both composition and length between the various strains of *C. difficile*. The sequences of PilJ show no insertions and only one deletion (in strain CD160) and have a much higher level of overall conservation (97% identity over 238 genomes as opposed to 90% for PilA1). As noted previously (Piepenbrink et al., 2014), the polymorphisms are primarily limited to one section of the αβ loop and one face of the C-terminal domain.

The structure of PilA1 is a substantial addition to our knowledge of type IV pilin structures because there is such great divergence between the sequences of type IV pilins from Gram-positive and Gram-negative strains, despite the conservation of the overall fold. A recent effort to model the structure of PilA1 based on the structure of the GC pilin of *N. gonorrhoeae* correctly predicted the structure of the αβ loop but was less successful in modeling the C terminus, hindered by the low (17%) sequence identity between the two proteins (Melville and Craig, 2013). Homologs of PilA1 are found in many other Clostridiales but no homologs of PilJ are known (Figure S2).

Table 1. Data Collection and Refinement Parameters

Crystallographic Parameters	MBP-PilA1 R202091	MBP-PilA1 NAP08	MBP-PilA1 CD160
Resolution range (Å)	38.58–1.899 (1.966–1.899)	59.31–2.23 (2.313–2.234)	36.59–1.724 (1.785–1.724)
Space group	C 1 2 1	P 21 21 21	I 1 2 1
Unit cell	143.951, 79.366, 164.746, 90, 90.23, 90	66.224, 74.71, 97.517, 90, 90, 90	62.484, 70.468, 117.153, 90, 90.24, 90
Total reflections	998,770 (98,860)	47,582 (4,141)	97,145 (5,144)
Unique reflections	139,534 (14,028)	23,808 (2,087)	51,185 (784)
Multiplicity	7.1 (7.0)	2.0 (2.0)	6.7 (6.6)
Completeness (%)	95.19 (92.48)	98.73 (87.51)	95.27 (69.38)
Mean $I/\sigma(I)$	8.74 (1.14)	14.14 (2.00)	11.99 (7.25)
Wilson B factor	22.94	26.63	27.53
R_{merge}	0.1776 (1.844)	0.04727 (0.4165)	0.09605 (0.1519)
R_{meas}	0.1914	0.06684	0.1043
R_{work}	0.1969 (0.2887)	0.2011 (0.2608)	0.1968 (0.3684)
R_{free}	0.2326 (0.3136)	0.2476 (0.3051)	0.2218 (0.4014)
Ramachandran favored (%)	98	97	99
Ramachandran outliers (%)	0	0.4	0
Average B factor	25.60	29.10	33.30
Macromolecules	25.20	29.20	32.80
Ligands	29.40	16.70	27.40
Solvent	29.20	29.10	38.60

Structural Divergence of PilA1 between *C. difficile* Strains

Because of the unusual degree of sequence divergence in PilA1, at least in comparison with other *C. difficile* T4P genes, we solved the structure of PilA1 from two additional strains of *C. difficile,* NAP08 (ribotype 078) and a recently discovered divergent strain of *C. difficile* termed CD160. A sequence alignment of PilA1 from R20291, NAP08, and CD160 is shown in Figure 5A with the three regions of greatest divergence highlighted. Each of these stretches of amino acids is located within an exposed loop in the R20291 structure, leading us to hypothesize that one side of PilA1, which likely faces in toward the core of the formed pilus, would be conserved between PilA1 from different strains of *C. difficile,* while the region that likely faces out into solution would contain all significant structural differences.

Using the same crystallographic methods, we solved the structure of PilA1 from NAP08 to a resolution of 2.2 Å (Table 1). As expected, the overall fold is quite similar between the two structures, with the divergent regions restricted to the face of the protein we expect to be exposed in the type IV pilus, namely the αβ loop, the loop between the second strand of the central β sheet and the first strand of the B2 sheet, and the loop between the α3 helix and the second strand of the B2 sheet (Figure 5C). These structural differences are not expected to have a significant effect on function but may well contribute to immune escape as antibodies against these regions may not be cross-protective against different strains of *C. difficile*.

We also solved the structure of PilA1 from CD160 to 1.7 Å resolution (Table 1). An overlay of the R20291, NAP08, and CD160 structures is shown in Figure 5B. The overall fold of PilA1 CD160 is similar to those of R202091 and NAP08 but there are substantial differences in the loops that we expect to be exposed in the assembled pilus. In comparison with the R20291 structure, the αβ loop of the CD160 structure is four residues shorter and slightly more helical, and the loop between the second strand of the central β sheet and the first strand of the B2 sheet is two residues longer and more exposed. However, the greatest difference is in the C terminus, where the region that forms the second strand of the B2 sheet in the R20291 structure stretches farther away from the core of the protein, eliminating any possibility for direct interactions with the stretch that forms the first strand of the B2 sheet in the R20291 and NAP08 structures (Figure 5A).

This initially puzzling finding is explained by the presence of a network of water molecules between the two strands. As shown in Figure 5D, five crystallographic water molecules lie between these two strands and make hydrogen bonds to the backbone of one or both strands. Of these five waters, two in particular are highly ordered, with B factors of 27.27 and 27.38 (the average for the structure is 33.30 overall and 38.60 for solvent atoms). In addition, they each form one direct hydrogen bond to each strand, creating water-mediated hydrogen bonds between the two strands in place of the direct backbone-backbone hydrogen bonds of the B2 sheet found in the R202091 and NAP08 PilA1 structures (Figure S4).

The structural divergence of PilA1 from R20291, NAP08, and CD160, taken together with the structure of FimA from *Dichelobacter nodosus* (Hartung et al., 2011), show that multiple energetic strategies exist for stabilizing the C termini of pilins in the absence of disulfide bonds. We expect that as more structures of pilins from Gram-positive species are characterized, the range of known pilin C-terminal structural motifs will continue to expand.

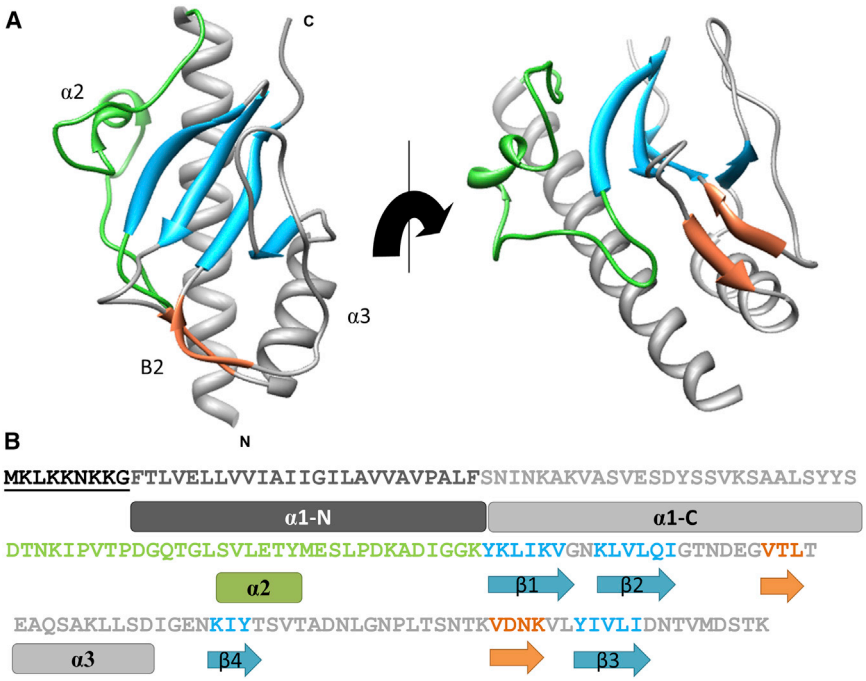


Figure 3. PilA1 3D Structure

(A) 3D structure of PilA1 from *C. difficile* R20291. The αβ loop is shown in green and the central beta sheet in blue. The B2 sheet is depicted in orange. (B) Schematic diagram of PilA1 secondary structure. The pre-pilin leader sequence is underlined and the α-1N helix is colored dark gray. The head group is colored gray save for the αβ loop and the two beta sheets, which are colored to match those in (A).

The Structures of PilA1 Support a Model of Pilus Formation Similar to Type IVb Pili

Previously, we found that the best model for incorporating PilJ into a type IV pilus was based on the toxin-coregulated pilus (TCP) of *V. cholerae* (Piepenbrink et al., 2014). However, the unusual dual-pilin fold of PilJ complicated comparisons with other single-domain pilins. We used the DALI server (Holm and Rosenstrom, 2010) to find the closest structural analogs of PilA1 and the results of this search are presented in Table S1. The highest scoring matches were to TcpA of *V. cholerae* El Tor (Z = 7.8) and CofA of *E. coli* (Z = 7.7), while numerous pilins and pseudopilins were also found to have significant structural similarity including PilS from *Salmonella* Typhi, type 2 secretion pseudopilins GspG and EspG, and TcpA from the classical strain of *V. cholerae*.

A structural alignment of PilA1 with TcpA from the classical strain of *V. cholerae* is shown in Figure 6A. Although there are substantial differences in the folds of the two proteins, there is an obvious similarity (3.0 Å rmsd), particularly in the αβ loop and the C-terminal accessory helix (the α3 helix of TcpA and PilA1). Based on these similarities, we modeled full-length PilA1 as previously described (Piepenbrink et al., 2014) and subsequently created a model of PilA1 pilus assembly based on the high-resolution electron micrographs of the TCP (Li et al., 2008), as shown in Figure 6D. When the PilA1 pilus model was superimposed onto the TCP electron density, the correlation coefficient was 0.90.

Because the structure of PilJ is also known and is incorporated into the same type IV pili as PilA1 (Piepenbrink et al., 2014), we can also judge any model of PilA1 pilus formation by its ability to accommodate PilJ subunits. As shown in Figure 6D, we were able to generate models of pili composed primarily of PilA1 but with PilJ incorporated either sporadically (ii) or in stretches (iii). This was expected as a superimposition of PilA1 and PilJ (Figure 6C) shows that the two are structurally similar (rmsd = 3.65 Å) in the regions expected to contact other pilus subunits, namely the α1 helix, the αβ loop, and the helical portion of the C terminus.

In addition, we noted that our model resulted in the formation of an inter-pilin salt bridge between residues lysine 30 and glutamate 75 of PilA1 (Figure 6B). This salt bridge was very similar in position to one shown to be important in the TCP between residues R26 and E83 (and is also found in CofA in residues R31 and E93) (Li et al., 2008). Although there is no detectable sequence similarity between PilA1 and PilJ outside of the N terminus, PilJ contains a lysine and a glutamate in nearly identical positions when the two proteins are superimposed (Figure 6C), namely K30 and E76, indicating a remarkable conservation of functional residues despite complete divergence in overall sequence. In 238 genomes of *C. difficile*, the only polymorphisms found at these positions were E75D substitutions in 2% of the PilA1 sequences. To investigate the importance of this potential interaction in pilus formation by *C. difficile*, we created PilA1 complementation plasmids with point mutations reversing the charges at positions 30 (K30E) and 75 (E75K). Using these plasmids to complement the Δ*pilA1* mutant strain of *C. difficile*, we found that unlike wild-type *pilA1*, the K30E and E75K mutants of *pilA1* were completely unable to produce type IV pili (Figure 7). The double mutant *pilA1* K30E E75K was also unable to produce type IV pili, which may stem from an imperfect recreation of the salt bridge but could also be due to an inability to incorporate one or more minor pilins as these proteins would not be able to form salt bridges with PilA1 K30E E75K.

DISCUSSION

The structures of type IV pilins from Gram-negative bacteria elucidated to date are remarkable in the commonality of their folds despite their vast divergence in sequence. The most logical explanation for this structural conservation is a preservation of function; not that all type IV pilins share any one function, but all share the requirement that they must be able to polymerize into pili and all of the functions ascribed to type IV pili are related in one way or another to adhesion, be it to host cells, bacterial cells (microcolony formation), DNA (horizontal transfer), or abiotic surfaces (biofilm formation). Indeed, the division between type IVa and type IVb pilins is a functional one but is represented

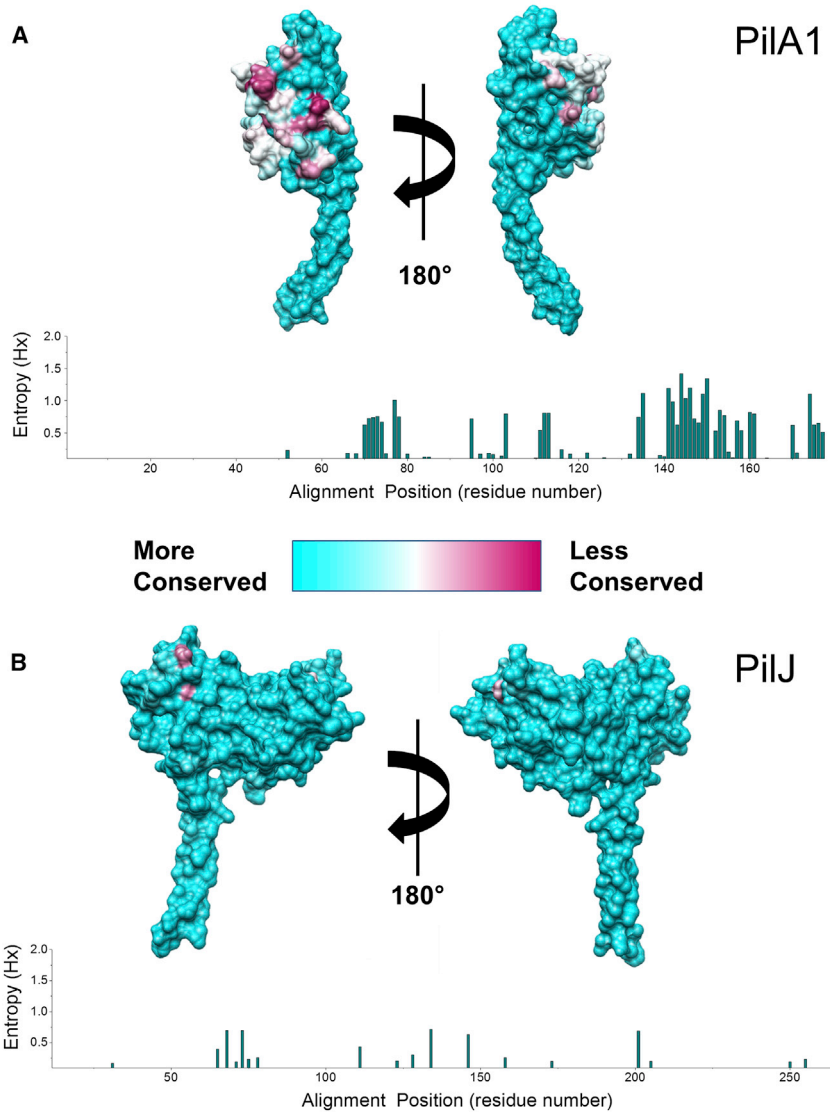


Figure 4. Genetic Analysis of *C. difficile* Type IV Pilins

Conservation heat maps of the PilA1 (A) and PilJ 3D (B) structures showing regions of conservation in cyan and variable regions in maroon. Below, Shannon entropy plots (BioEdit) show the degree of sequence variability at each position.

in structural differences, particularly in the larger C-terminal loops found in type IVb pilins (Giltner et al., 2012).

The structural similarity between PilA1 and pseudopilins is also noteworthy as pseudopilins from Gram-negative bacteria also lack a C-terminal disulfide bond and the overall similarity between type II secretion and type IV pili implies some common ancestral pilus-like structure. The αβ loop of PilA1 resembles the variable loop of major pseudopilins in that it is largely disordered but contains some helical character in the center (Alphonse et al., 2010; Kohler et al., 2004; Korotkov et al., 2009). The placement of the calcium binding site in major pseudopilins is also similar to that of the B2 sheet of PilA1, both sequentially (it occurs in a loop between two strands of the β sheet) and topographically (it rests below the β sheet on an axis paralleling the α1-N helix) (Giltner et al., 2012). Taken together, these data suggest that the ancestral pilin-like proteins used a variety of means to stabilize their C terminus and that the disulfide bond became nearly ubiquitous in type IV pilins from Gram-negative strains as a consequence of the formation of a periplasmic space. However, this hypothesis does not explain the presence of calcium-binding motifs rather than disulfide bonds in the pseudopilins of Gram-negative type II secretion systems.

Like other pilins identified in Gram-positive strains (Imam et al., 2011), PilA1 contains no cysteines, meaning that it lacks the ability to form the aforementioned disulfide bond. The only other known structure of a type IV pilin without a C-terminal disulfide bond is that of FimA from *D. nodosus*, which instead contains a disulfide bond between the αβ loop and the third strand of its β sheet, and is proposed to use a combination of hydrogen bonds and hydrophobic interactions to stabilize its C terminus (Hartung et al., 2011). However, unlike FimA, which is a type IVa pilin and has a continuous β sheet with relatively short sequences between the strands, PilA1 has a discontinuous β sheet. The C-terminal β strand is not the final strand; instead the α3 helix, the final (fourth) β strand, and the B2 sheet are placed sequentially between the second and third strands, analogous to what is seen in the type IVb pilins, including TcpA from *V. cholerae*, CofA from enterotoxigenic *E. coli* (ETEC), and PilS from *Salmonella typhi* (Balakrishna et al., 2009; Craig et al., 2003; Kolappan et al., 2012). This leaves a 34-residue loop between the second and fourth strands of the central β sheet.

The obvious question arises: how then does PilA1 stabilize its C terminus? The near-universal frequency of the C-terminal disulfide bond in type IV pilins from Gram-negative species has long been considered an indication of its importance (Giltner et al., 2012). However, its absence in Gram-positive type IV pilins shows that FimA is part of a larger pattern of type IV pilins in which noncovalent interactions are used in place of a covalent disulfide bond at the C terminus of the protein.

One possible explanation can be found by comparing the C terminus of PilA1 with other type IV pilins. The C termini are the least conserved regions of type IV pilins and we expect that they will vary considerably in terms of the strength of the interactions needed to anchor C-terminal loops to the core of the protein. Figure S5 shows the C-terminal structures of various type IV pilins and pseudopilins. The disulfide bonds of PilE and PilA (PAK) anchor C-terminal loops to their β sheets, while the

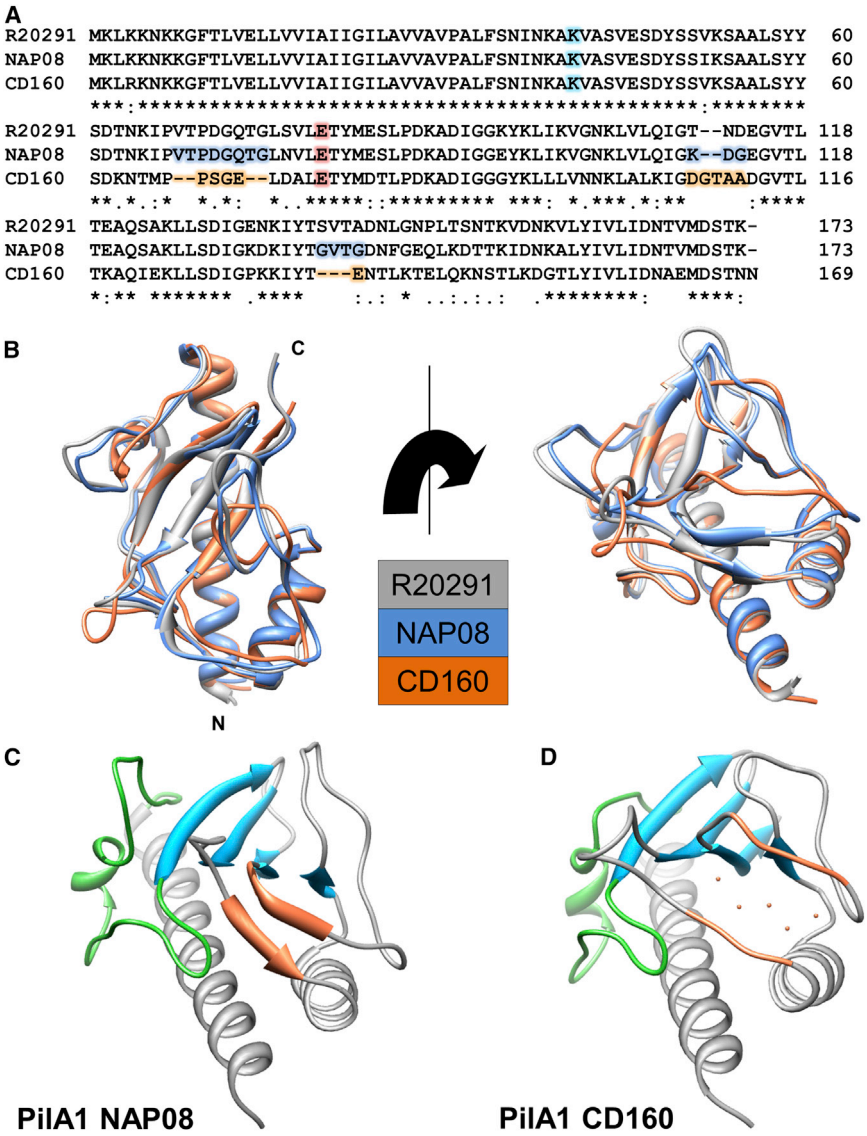


Figure 5. Variability in PilA1 Sequence and Structure

(A) Sequence alignment of PilA from *C. difficile* R20291, NAP08, and CD160. Lysine 30 and glutamate 75 are highlighted in blue and red, respectively. Three variable loops are also highlighted.

(B) Overlay of the 3D structures of PilA1 from *C. difficile* R202091 (gray), NAP08 (blue), and CD160 (orange).

(C and D) 3D structure of PilA1 NAP08 (C) and PilA1 CD160 (D). The αβ loop is shown in green and the central beta sheet in blue. The strands that make up the B2 sheet in the R20291 and NAP08 structures are depicted in orange, as are the water molecules forming water-mediated hydrogen bonds between the two strands in CD160.

disulfide bonds of TcpA, PilS, and CofA bridge two C-terminal α helices found in loops between the strands of the β sheet (Giltner et al., 2012). The PilA1 fold contains the discontinuous β sheet found in type IVb pilins but contains only one C-terminal α helix, making PilA1 more compact than the Gram-negative pilins that it most closely resembles. This more compact fold leaves only two long insertions between the strands of the central β sheet, each of which contains one strand of the B2 sheet, anchoring the extended β4-β3 loop to the core of the protein.

That same portion of the β4-β3 loop in PilA1 CD160 extends outward, dissolving the B2 sheet and creating a network of water-mediated hydrogen bonds in its place. This alternative conformation alters the surface of PilA1 significantly, which is likely to affect both function and immune recognition, although all three forms of PilA1 are stabilized by hydrogen bonds between the β2-β4 and β4-β3 loops.

The structure of the *C. difficile* minor pilin, PilJ, demonstrated how far a pilin can deviate from these established norms while maintaining the ability to incorporate into a type IV pilus (Piepenbrink et al., 2014). However, the structure of the major pilin of *C. difficile*, PilA1, allows for a better comparison with the structures of major pilins from other species. All of the structures of PilA1 presented here show an overall shape similar to that of PilS (*S. typhi*), TcpA (*V. cholerae*), and CofA (ETEC), including a helical region in the C terminus that superimposes onto the α3 helix of TcpA. Despite the overall differences in fold, PilJ also has a helical region that superimposes onto the same α3 region. In addition, TcpA, CofA, PilJ, and PilA1 share a pair of residues in similar positions that we expect form a salt bridge in all three pili.

The type IV pili of Gram-positive species are as yet only sparsely characterized, but a recent review noted the similarities in genetic organization between clostridial pilus biogenesis machinery and the type IVb pilus operons (Melville and Craig, 2013). Here, we show that the structures of two type IV pilins from *C. difficile* show a remarkable resemblance to type IVb pilins regarding the width of their head groups, the relative positions of helical regions, and the configuration of the central β sheet of the major pilin, PilA1.

While type IVb pili have been implicated in a variety of functions, the most common is microcolony formation, which has been demonstrated to occur through type IV pilus bundling in enteropathogenic *E. coli* and *S. typhi* (Bieber et al., 1998; Raza et al., 2011). If we examine those type IV pilus systems most closely related to the *C. difficile* pili by the structures of their major pilins, we find that the type IV pili of *V. cholerae*, *S. typhi*, and ETEC have all been implicated in microcolony formation (Clavijo et al., 2010; Tam et al., 2006). As *C. difficile* is known to colonize the human intestine and has been observed to form microcolonies in a hamster model (Buckley et al., 2011), which appear to be mediated by flagella-like appendages resembling

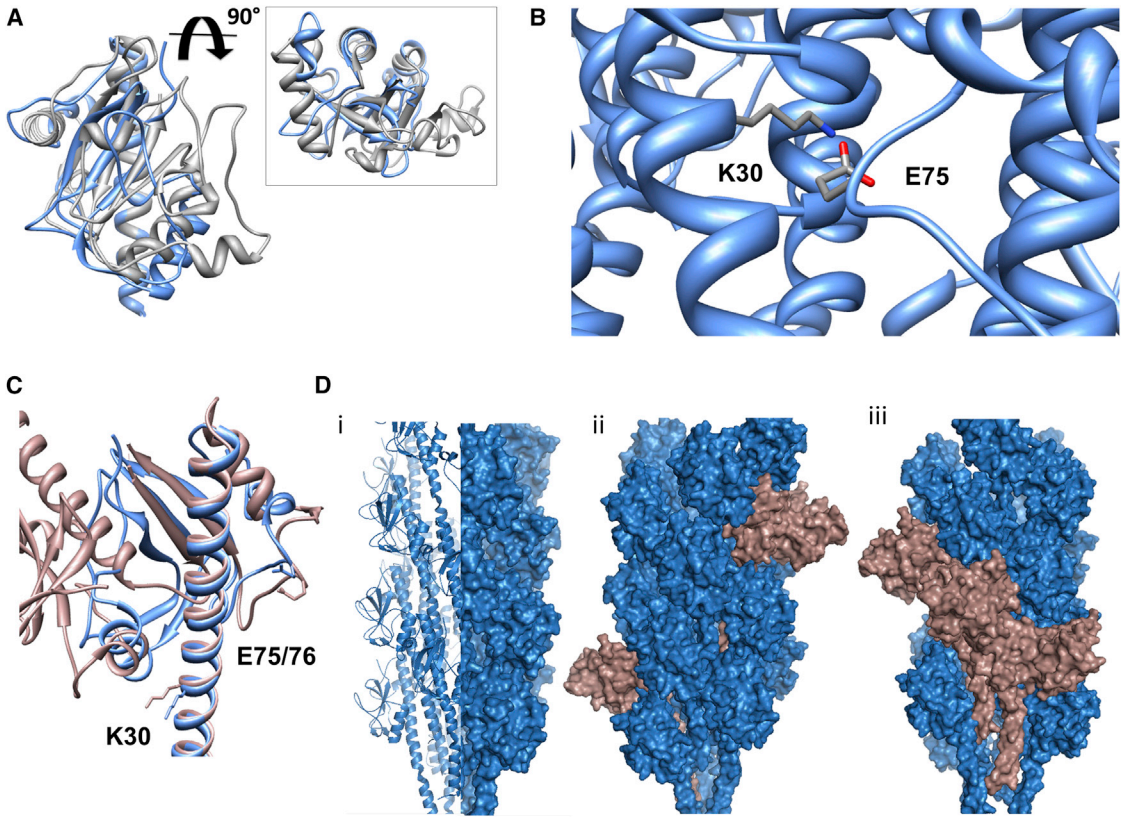


Figure 6. Incorporation of PilA1 and PilJ into T4P

(A) Superimposition of PilA1 and TcpA from *V. cholerae*. PilA1 is shown in blue, TcpA in gray.

(B) Pilus fiber model of PilA1 based on the TCP of *V. cholerae*. A predicted salt bridge between lysine 30 and glutamate 75 is depicted.

(C) Superimposition of PilA1 and PilJ; PilA1 is shown in blue, PilJ in chocolate. Side chains from lysine 30 and glutamate 75 (PilA1)/76 (PilJ) are shown.

(D) Pilus fiber models showing (i) a pilus made up entirely of PilA1, (ii) a pilus with PilJ incorporated sporadically, and (iii) a pilus with PilJ incorporated in clusters.

type IV pili, it is likely that one function of type IV pili in *C. difficile* is to mediate this self-association in vivo.

The function of PilJ within the type IV pili of *C. difficile* also remains to be determined. Although hypotheses are somewhat speculative at this juncture, the fact that PilJ has so much more exposed surface area in our model of the pilus and its low incorporation rate suggest that it does not significantly aid in stabilizing the pilus but rather serves as an adhesin.

In conclusion, while the role of type IV pili in the colonization and pathogenic strategies of *C. difficile* remains to be determined, the fold of PilA1 shows the flexibility of the type IV pilin architecture and the variety of structural motifs that can be used to form similar folds. This flexibility allows for substantial variation in the sequence of the pilin head groups while maintaining an immediately recognizable common protein fold.

EXPERIMENTAL PROCEDURES

Reagents

Polyclonal antisera against PilA1 and PilJ used in this study were generated using recombinant proteins as described previously (Maldarelli et al., 2014; Piepenbrink et al., 2014).

Protein Expression and Purification

The codon-optimized sequences of PilA1 from strains R20291, NAP08, and CD160, starting with serine 26, were cloned into a maltose-binding fusion vector making use of previously described surface entropy reduction mutations (pMal E) (Moon et al., 2010). A C-terminal 6xHis tag was included for ease of purification. These clones were transformed into BL21 (DE3) pLysS cells and grown to saturation overnight with shaking at 37°C in LB media with 50 μg/ml ampicillin. These saturation cultures were then diluted into fresh LB-ampicillin and grown to an optical density (OD) of 0.4–0.6 at 37°C. These flasks were transferred to a refrigerated orbital shaker and cooled to 18°C before induction with 30 μM isopropyl β-D-1-thiogalactopyranoside. These flasks were allowed to grow overnight before being harvested by centrifugation at 7,500 *g* for 10 min. The cells were then lysed using lysozyme and the resulting lysate centrifuged again, this time at 20,000 *g* for 30 min. The supernatant was purified using a nickel-NTA column and the elution further purified through size-exclusion chromatography over a GE S200 Superdex column using an Åkta Purifier FPLC.

Structure Determination and Refinement

All three MBP-PilA1 constructs were initially screened at a concentration of 20 mg/ml in 20 mM Bis-Tris (pH 6.0), 100 mM NaCl, and 50 mM maltose. However, each crystallized in a different set of precipitating conditions and optimization produced different ideal concentrations of NaCl and maltose.

NAP08

MBP-PilA1 NAP08 initially crystallized in Hampton's Index screen condition D7 (25% PEG 3350, 100 mM Bis-Tris [pH 6.5]) at room temperature. The optimal buffer conditions were 20 mM Bis-Tris (pH 6.0), 100 mM NaCl, and the optimal precipitant conditions were 25% PEG 2000MME, 100 mM Bis-Tris (pH 6.5) (crystals were grown by hanging drop vapor diffusion at 4°C). Crystals grew overnight and were harvested and flash-cooled in the mother liquor

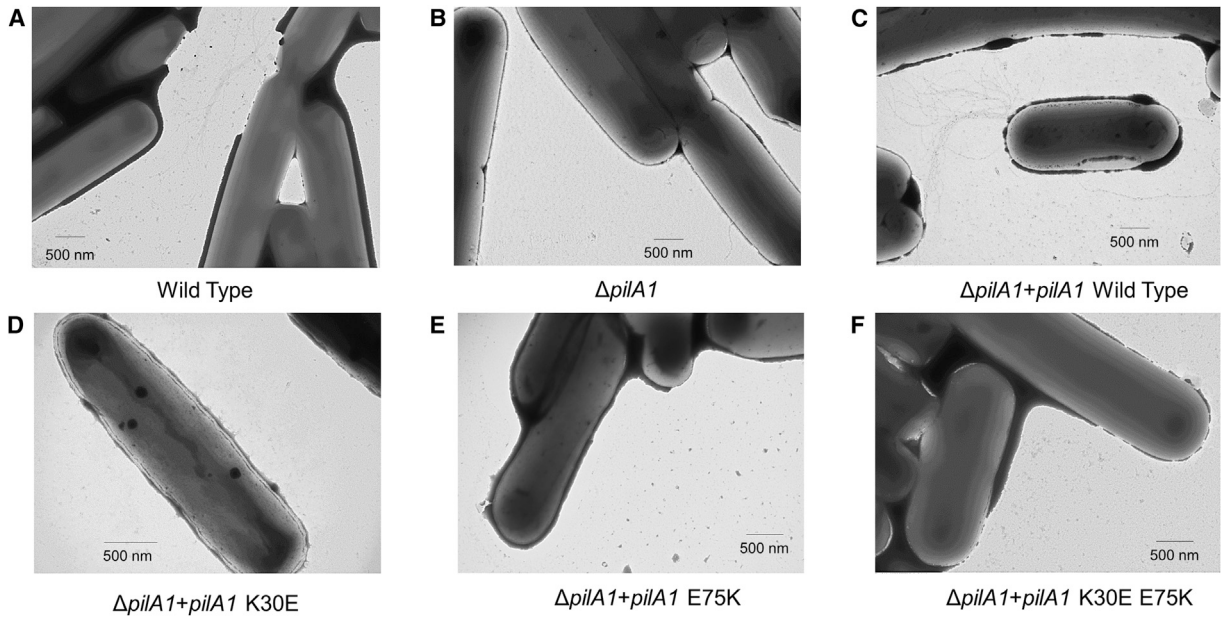


Figure 7. Pilation of PilA1 Mutant *C. difficile* Strains

Electron micrographs are shown of (A) *C. difficile* R20291 wild-type, (B) *C. difficile* R202091 with a gene interruption mutant of the *pilA1* gene, *C. difficile* Δ*pilA1*, (C) Δ*pilA1* complemented with *pilA1* wild-type, (D) Δ*pilA1* complemented with *pilA1* K30E, (E) Δ*pilA1* complemented with *pilA1* E75K, and (F) Δ*pilA1* complemented with *pilA1* K30E E75K.

supplemented with 25% glycerol. Data were collected at the National Light Source Beamline X25 in Brookhaven, NY.

CD160

MBPilA1 CD160 initially crystallized in Emerald Biosystems Wizard III/IV F12 (2.4 M sodium malonate). The optimal buffer conditions were 20 mM Bis-Tris (pH 6.0), 100 mM NaCl, 100 mM maltose, and optimization did not change the precipitant conditions (crystals were grown by sitting drop vapor diffusion at room temperature). Crystals grew overnight and were harvested and flash-cooled without the addition of any additional cryoprotectant. Data were collected remotely at Stanford Synchrotron Radiation Lightsource Beamline 12-2.

To ensure that the alternative conformation found in the CD160 structure is not a crystallographic artifact of the I 1 2 1 space group, which has close crystal contacts in this region of the protein, we also solved the structure in a different crystal form (space group, P 1 21 1), which has no crystal contacts in this region. The structures of CD160 PilA1 in these two crystal structures were found to be essentially identical and density for crystallographic waters was found in nearly identical positions in the space between the two B2 strands.

R20291

MBP-PilA1 R20291 Emerald Biosystems Wizard I/II B9 (20% PEG 8000, 100 mM HEPES [pH 7.5]). The optimal concentration was 30 mg/ml and the optimal buffer conditions 20 mM Bis-Tris (pH 6.0), 50 mM maltose; optimization did not change the precipitant conditions (crystals were grown by hanging drop vapor diffusion at room temperature).

All three data sets were processed with XDS (Kabsch, 2010). Molecular replacement was carried out by Phaser (McCoy et al., 2007) using maltose-binding protein (MBP-PilA1 NAP08) or a sequential search of (1) maltose-binding protein and (2) PilA1 NAP08 (MBP-PilA1 CD160 and MBP-PilA1 R20291). Phenix and Coot were used for phasing, building, and refinement (Adams et al., 2002, 2010, 2011; Emsley and Cowtan, 2004). Multiple iterative rounds of model building and refinement resulted in an R_{work} factor of 20.1% and an R_{free} of 24.7% (NAP08), an R_{work} factor of 19.7% and R_{free} of 22.2% (CD160), and an R_{work} factor of 19.7% and R_{free} of 23.3% (R20291). The crystallographic parameters of the refined data are summarized in Table 1.

C. difficile Lysate Preparation

Cultures of *C. difficile* strain R20291 were grown on Columbia agar plates as above. Colonies were scraped off the plates, resuspended in PBS, and brought to an OD_{600} of 20. Bacteria were lysed as described previously (Lawley et al., 2009); in brief, resuspended cells were frozen at −20°C overnight, then incubated at 37°C for 40 min. Total protein present in each lysate was quantified with a commercial bichinoic acid assay kit (Thermo Scientific Pierce) according to the manufacturer's protocol.

In preparation for immunoblotting, lysates were diluted 1:1 in Laemmli buffer and boiled for 10 min. Lysates and a standard curve of purified untagged PilJ and PilA1 were separated on precast 4%–15% gradient Mini-PROTEAN TGX polyacrylamide gels (Bio-Rad) and transferred to polyvinylidene difluoride membranes. Blots were blocked for 1 hr with 5% nonfat dry milk, then incubated at 4°C overnight with polyclonal rabbit anti-PilJ at a 1:10,000 dilution. Blots were washed in PBS with Tween-20, incubated with IRDye 800CW donkey anti-rabbit IgG H+L (Li-Cor Biosciences) at a 1:20,000 dilution for 1 hr. Infrared signals were detected and quantified using the Odyssey imaging system (Li-Cor Biosciences).

Creation of *C. difficile* pilA1 Gene Interruption Mutant

C. difficile mutants were produced using the ClosTron system (Heap et al., 2010). Retargeted MTL007C-E2 vectors for the Cdi-*pilA1*-39a mutant were purchased from DNA2.0 Inc. The retargeted vector was transferred from *E. coli* CA434 to wild-type R20291 *C. difficile* by conjugation using cycloserine cefoxitin fructose agar (CCFA) plates containing 15 mg/ml thiamphenicol. Once colonies appeared on the plates (48–72 hr), they were restreaked onto CCFA containing 2.5 μg/ml erythromycin to select for bacteria in which the intron had been inserted. Insertions were confirmed by PCR using the EBS universal primer (5′ cga aat tag aaa ctt gcg ttc agt aaa c) and a reverse primer (*pilA1F*, 5′ ccc tgg cgg ccg cat agt ata aac agt gtt gct).

Creation of *C. difficile* pilA1, pilA1K30E, pilA1E75K, and pilA1K30E-E75K Complemented Strains

To complement the *pilA1* mutation, the *pilA1* gene and its promoter region were PCR amplified using the above indicated primer *pilA1F* and the *pilA1R*

(5′ ccc tgc tcg agc tat ttt gtt gag tcc atc ac) and cloned into the p84151 modular plasmid using the NotI and XhoI restriction enzymes (Heap et al., 2009). The vector was then transformed into *E. coli* CA434 and conjugated into Cdi-*pilA1*-39a, and positive conjugants were selected for on CCFA containing 15 μg/ml thiamphenicol. Complement strains were screened by PCR and Western immunoblotting to confirm the proper genotype and phenotype.

For the construction of the complementation plasmids *pilA1*K30E, *pilA1*E75K, and *pilA1*K30E-E75K, we used the Quick Change Mutagenesis system (Agilent Technologies) using the p84151-*pilA1* as a template and the primers *pilA1*K30EF (5′ cca gct tta ttt agt aat ata aac gaa gct aag gta gca agt gtt gag tct g), p*ilA1*k30ER (5′ cag act caa cac ttg cta cct tag ctt cgt tta tat tac taa ata aag ctg g), *pilAE*75K fF (5′ gtg ttt tag aga ctt ata tga aat ctc tgc ctg ata aag ctg), and *pilA*1E75KR (5′ cag ctt tat cag gca gag att tca tat aag tct cta aaa cac) to introduce the mutations in the desired codons. The resultant plasmids were confirmed by sequencing and transformed into CA434 for posterior conjugation with the Cdi-*pilA*1-39a mutant.

Immunogold Electron Microscopy

The immunogold electron microscopy assays where performed as described in Piepenbrink et al. (2014).

Genetic Analysis

Protein sequences of PilA1 and PilJ were translated from *C. difficile* genome sequences in the National Center for Biotechnology Information database, aligned with ClustalX (Jeanmougin et al., 1998), and analyzed using SAIS and Bioedit.

RNA-Seq

Cultures of *C. difficile* strain R20291 were grown on Columbia agar plates as described above. Within an anaerobic chamber (Coy Laboratories, Grass Lake, MI), colonies were removed and placed immediately into RNAprotect bacteria reagent (Qiagen) and stored at −80°C until used. After thawing, cells were pelleted and resuspended in 100 μl of TES buffer with 35,000 units of Ready-Lyse lysozyme solution (Epicentre Technologies, Madison, WI) and incubated with shaking at room temperature for 90 min. The RNeasy RNA purification kit (Qiagen) was then used according to the manufacturer's instructions. RNA quantity and integrity were measured with an Agilent bioanalyzer and all RNA integrity numbers exceeded 8. Total RNA samples were treated with DNase I (Invitrogen). The level of ribosomal RNA present in total RNA samples was reduced prior to library construction using the Ribo-Zero Gram Positive Bacteria rRNA Removal Kit (Epicentre Technologies, Madison, WI). Illumina RNA-Seq libraries were prepared with the TruSeq RNA Sample Prep kit (Illumina, San Diego, CA), omitting the poly-A selection steps. Adapters containing six-nucleotide indexes were ligated to the double-stranded cDNA. The DNA was purified between enzymatic reactions and the size selection of the library was performed with AMPure XT beads (Beckman Coulter Genomics, Danvers, MA). The libraries were sequenced on a 100 bp paired end run on the HiSeq 2000 (Illumina, San Diego, CA).

Pilus Modeling

Full-length PilA1 and PilJ were modeled using the structure of the full-length *P. aeruginosa* PAK pilin (Craig et al., 2003). The initial model of the PilA1 pilus was created by superimposition onto a model of the TCP of *V. cholerae* (Protein Data Bank ID 1OR9), which did not result in any clashes. The resulting model then underwent rigid-body minimization by UCSF Chimera (Pettersen et al., 2004) using the electron density from a high-resolution cryo-electron micrograph of the TCP (EMD-1954). The incorporation of PilJ was modeled by replacing PilA1 subunits with superimposed PilJ subunits without additional refinement.

ACCESSION NUMBERS

The X-ray crystal structure of PilJ is deposited in the Protein Data Bank as PDB ID 4IXJ; MBP-PilA1 R202091 is 4TSM, MBP-PilA1 NAP08 is 4OGM, and MBP-PilA1 CD160 is 4PE2.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and one table and can be found with this article online at http://dx.doi.org/10.1016/j.str.2014.11.018.

AUTHOR CONTRIBUTIONS

Structure determination and modeling were done by K.H.P. Genetic analysis was done by G.A.M. and K.H.P. Electron microscopy was done by C.F.M., G.L.M., and A.L. The Δ*pilA1* strain of *C. difficile* was engineered by T.C.D. Mutagenesis of *pilA1* complementation plasmids was done by C.F.M. Measurements of mRNA expression were made by E.v.R. and measurements of protein expression were made by G.A.M. All experiments were performed under the direction of G.D.A., M.S.D., or E.J.S.

ACKNOWLEDGMENTS

This work was supported by NIH grant R21 AI105881 (to M.S.D.). Additional work was supported by the Alberta Glycomics Centre. K.H.P. was supported in part by NIH training grant T32 AI095190 and by an NIH fellowship F32 AI 110045. C.M.P. was supported in part by Mexico's National Council for Science and Technology No. 130211 (CONACYT). We thank the staff at Brookhaven National Laboratory National Synchrotron Light Source beamline X25 and the staff at Stanford Synchrotron Radiation Lightsource Beamline 12-2 for technical assistance with X-ray data collection. We also thank Alison Scott for technical assistance in collecting and analyzing the RNA-Seq data.

Received: September 19, 2014
Revised: November 18, 2014
Accepted: November 22, 2014
Published: January 15, 2015

REFERENCES

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D Biol. Crystallogr. *58*, 1948–1954.

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Echols, N., Headd, J.J., Hung, L.W., Jain, S., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2011). The Phenix software for automated determination of macromolecular structures. Methods *55*, 94–106.

Alphonse, S., Durand, E., Douzi, B., Waegele, B., Darbon, H., Filloux, A., Voulhoux, R., and Bernard, C. (2010). Structure of the *Pseudomonas aeruginosa* XcpT pseudopilin, a major component of the type II secretion system. J. Struct. Biol. *169*, 75–80.

Balakrishna, A.M., Saxena, A.M., Mok, H.Y., and Swaminathan, K. (2009). Structural basis of typhoid: *Salmonella typhi* type IVb pilin (PilS) and cystic fibrosis transmembrane conductance regulator interaction. Proteins *77*, 253–261.

Bieber, D., Ramer, S.W., Wu, C.Y., Murray, W.J., Tobe, T., Fernandez, R., and Schoolnik, G.K. (1998). Type IV pili, transient bacterial aggregates, and virulence of enteropathogenic *Escherichia coli*. Science *280*, 2114–2118.

Borriello, S.P., Davies, H.A., and Barclay, F.E. (1988). Detection of fimbriae amongst strains of *Clostridium difficile*. FEMS Microbiol. Lett. *49*, 65–67.

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

Buckley, A.M., Spencer, J., Candlish, D., Irvine, J.J., and Douce, G.R. (2011). Infection of hamsters with the UK *Clostridium difficile* ribotype 027 outbreak strain R20291. J. Med. Microbiol. *60*, 1174–1180.

Clavijo, A.P., Bai, J., and Gomez-Duarte, O.G. (2010). The Longus type IV pilus of enterotoxigenic *Escherichia coli* (ETEC) mediates bacterial self-aggregation and protection from antimicrobial agents. Microb. Pathog. *48*, 230–238.

Craig, L., Taylor, R.K., Pique, M.E., Adair, B.D., Arvai, A.S., Singh, M., Lloyd, S.J., Shin, D.S., Getzoff, E.D., Yeager, M., et al. (2003). Type IV pilin structure and assembly: X-ray and EM analyses of *Vibrio cholerae* toxin-coregulated pilus and *Pseudomonas aeruginosa* PAK pilin. Mol. Cell *11*, 1139–1150.

Craig, L., Pique, M.E., and Tainer, J.A. (2004). Type IV pilus structure and bacterial pathogenicity. Nat. Rev. Microbiol. *2*, 363–378.

Craig, L., Volkmann, N., Arvai, A.S., Pique, M.E., Yeager, M., Egelman, E.H., and Tainer, J.A. (2006). Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. Mol. Cell *23*, 651–662.

Dingle, T.C., Mulvey, G.L., and Armstrong, G.D. (2011). Mutagenic analysis of the *Clostridium difficile* flagellar proteins, FliC and FliD, and their contribution to virulence in hamsters. Infect. Immun. *79*, 4061–4067.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Giltner, C.L., Nguyen, Y., and Burrows, L.L. (2012). Type IV pilin proteins: versatile molecular modules. Microbiol. Mol. Biol. Rev. *76*, 740–772.

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

Goulding, D., Thompson, H., Emerson, J., Fairweather, N.F., Dougan, G., and Douce, G.R. (2009). Distinctive profiles of infection and pathology in hamsters infected with *Clostridium difficile* strains 630 and B1. Infect. Immun. *77*, 5478–5485.

Hartung, S., Arvai, A.S., Wood, T., Kolappan, S., Shin, D.S., Craig, L., and Tainer, J.A. (2011). Ultrahigh resolution and full-length pilin structures with insights for filament assembly, pathogenic functions, and vaccine potential. J. Biol. Chem. *286*, 44254–44265.

Heap, J.T., Pennington, O.J., Cartman, S.T., and Minton, N.P. (2009). A modular system for *Clostridium* shuttle plasmids. J. Microbiol. Methods *78*, 79–85.

Heap, J.T., Kuehne, S.A., Ehsaan, M., Cartman, S.T., Cooksley, C.M., Scott, J.C., and Minton, N.P. (2010). The ClosTron: mutagenesis in *Clostridium* refined and streamlined. J. Microbiol. Methods *80*, 49–55.

Henrichsen, J. (1983). Twitching motility. Annu. Rev. Microbiol. *37*, 81–93.

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 Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids Res. *38*, W545–W549.

Imam, S., Chen, Z., Roos, D.S., and Pohlschroder, M. (2011). Identification of surprisingly diverse type IV pili, across a broad range of gram-positive bacteria. PLoS One *6*, e28919.

Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., and Gibson, T.J. (1998). Multiple sequence alignment with Clustal X. Trends Biochem. Sci. *23*, 403–405.

Kabsch, W. (2010). Xds. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Kohler, R., Schafer, K., Muller, S., Vignon, G., Diederichs, K., Philippsen, A., Ringler, P., Pugsley, A.P., Engel, A., and Welte, W. (2004). Structure and assembly of the pseudopilin PulG. Mol. Microbiol. *54*, 647–664.

Kolappan, S., Roos, J., Yuen, A.S., Pierce, O.M., and Craig, L. (2012). Structural characterization of CFA/III and Longus type IVb pili from enterotoxigenic *Escherichia coli*. J. Bacteriol. *194*, 2725–2735.

Korotkov, K.V., Gray, M.D., Kreger, A., Turley, S., Sandkvist, M., and Hol, W.G. (2009). Calcium is essential for the major pseudopilin in the type 2 secretion system. J. Biol. Chem. *284*, 25466–25470.

Laurenceau, R., Pehau-Arnaudet, G., Baconnais, S., Gault, J., Malosse, C., Dujeancourt, A., Campo, N., Chamot-Rooke, J., Le Cam, E., Claverys, J.P., et al. (2013). A type IV pilus mediates DNA binding during natural transformation in *Streptococcus pneumoniae*. Plos Pathog *9*, e1003473.

Lawley, T.D., Croucher, N.J., Yu, L., Clare, S., Sebaihia, M., Goulding, D., Pickard, D.J., Parkhill, J., Choudhary, J., and Dougan, G. (2009). Proteomic and genomic characterization of highly infectious *Clostridium difficile* 630 spores. J. Bacteriol. *191*, 5377–5386.

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.

Li, J., Lim, M.S., Li, S., Brock, M., Pique, M.E., Woods, V.L., Jr., and Craig, L. (2008). *Vibrio cholerae* toxin-coregulated pilus structure analyzed by hydrogen/deuterium exchange mass spectrometry. Structure *16*, 137–148.

Maldarelli, G.A., De Masi, L., von Rosenvinge, E.C., Carter, M., and Donnenberg, M.S. (2014). Identification, immunogenicity, and cross-reactivity of type IV pilin and pilin-like proteins from *Clostridium difficile*. Pathog. Dis. *71*, 302–314.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

Melville, S., and Craig, L. (2013). Type IV pili in Gram-positive bacteria. Microbiol. Mol. Biol. Rev. *77*, 323–341.

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

Moon, A.F., Mueller, G.A., Zhong, X., and Pedersen, L.C. (2010). A synergistic approach to protein crystallization: combination of a fixed-arm carrier with surface entropy reduction. Protein Sci. *19*, 901–913.

O'Toole, G.A., and Kolter, R. (1998). Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. Mol. Microbiol. *30*, 295–304.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Piepenbrink, K.H., Maldarelli, G.A., de la Pena, C.F., Mulvey, G.L., Snyder, G.A., De Masi, L., von Rosenvinge, E.C., Gunther, S., Armstrong, G.D., Donnenberg, M.S., et al. (2014). Structure of *Clostridium difficile* PilJ exhibits unprecedented divergence from known type IV pilins. J. Biol. Chem. *289*, 4334–4345.

Rakotoarivonina, H., Jubelin, G., Hebraud, M., Gaillard-Martinie, B., Forano, E., and Mosoni, P. (2002). Adhesion to cellulose of the Gram-positive bacterium *Ruminococcus albus* involves type IV pili. Microbiology *148*, 1871–1880.

Raza, A., Sarwar, Y., Ali, A., Jamil, A., Haque, A., and Haque, A. (2011). Effect of biofilm formation on the excretion of *Salmonella enterica* serovar Typhi in feces. Int. J. Infect. Dis. *15*, e747–e752.

Rudel, T., Scheurerpflug, I., and Meyer, T.F. (1995). *Neisseria* PilC protein identified as type-4 pilus tip-located adhesin. Nature *373*, 357–359.

Seifert, H.S., Ajioka, R.S., Marchal, C., Sparling, P.F., and So, M. (1988). DNA transformation leads to pilin antigenic variation in *Neisseria gonorrhoeae*. Nature *336*, 392–395.

Stone, B.J., and Abu Kwaik, Y. (1998). Expression of multiple pili by *Legionella pneumophila*: identification and characterization of a type IV pilin gene and its role in adherence to mammalian and protozoan cells. Infect. Immun. *66*, 1768–1775.

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

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.

Tam, C.K., Morris, C., and Hackett, J. (2006). The *Salmonella enterica* serovar Typhi type IVB self-association pili are detached from the bacterial cell by the PilV minor pilus proteins. Infect. Immun. *74*, 5414–5418.

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.

Varga, J.J., Nguyen, V., O'Brien, D.K., Rodgers, K., Walker, R.A., and Melville, S.B. (2006). Type IV pili-dependent gliding motility in the Gram-positive pathogen *Clostridium perfringens* and other Clostridia. Mol. Microbiol. *62*, 680–694.

Wall, D., and Kaiser, D. (1999). Type IV pili and cell motility. Mol. Microbiol. *32*, 1–10.

Yoshida, T., Kim, S.R., and Komano, T. (1999). Twelve pil genes are required for biogenesis of the R64 thin pilus. J. Bacteriol. *181*, 2038–2043.

Zhang, X.L., Tsui, I.S.M., Yip, C.M.C., Fung, A.W.Y., Wong, D.K.H., Dai, X.Y., Yang, Y.H., 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.

