Type IV Pili in Gram-Positive Bacteria

Stephen Melville, Lisa Craig

Department of Biological Sciences, Virginia Tech, Blacksburg, Virginia, USA; Molecular Biology and Biochemistry Department, Simon Fraser University, Burnaby, British Columbia, Canada

SUMMARY
Type IV pili (T4P) are surface-exposed fibers that mediate many functions in bacteria, including locomotion, adherence to host cells, DNA uptake (competence), and protein secretion and that can act as nanowires carrying electric current. T4P are composed of a polymerized protein, pilin, and their assembly apparatuses share protein homologs with type II secretion systems in eubacteria and the flagella of archaea. T4P are found throughout Gram-negative bacterial families and have been studied most extensively in certain model Gram-negative species. Recently, it was discovered that T4P systems are also widespread among Gram-positive species, in particular the clostridia. Since Gram-positive and Gram-negative bacteria have many differences in cell wall architecture and other features, it is remarkable how similar the T4P core proteins are between these organisms, yet there are many key and interesting differences to be found as well. In this review, we compare the two T4P systems and identify and discuss the features they have in common and where they differ to provide a very broad-based view of T4P systems across all eubacterial species.

INTRODUCTION
Type IV pili (T4P) are polymers of the major pilin protein that are displayed on the surfaces of many Gram-negative bacteria and were recently identified in a select number of Gram-positive species (1–3). In Gram-negative bacteria, T4P mediate specialized forms of locomotion, called gliding motility and twitching motility, which are independent of the flagella (3–5). In addition to motility, T4P carry out a multitude of functions, including adherence to eukaryotic cells, microcolony formation, DNA uptake, and protein secretion, and some even function as nanowires carrying electric current (6–9). T4P are polymerized from an inner membrane-bound pool of pilin monomers and are held together in the core of the pilus filament via hydrophobic interactions among the conserved N-terminal segments of the proteins (10–13). Beyond this common N-terminal segment, pilins can vary substantially in size and sequence, producing pili with unique surface features, interaction partners, and functions (1, 3).

The canonical T4P system is closely related to several other systems in bacteria and archaea. The type II secretion (T2S) sys-
TEM in Gram-negative bacteria, the competence systems (Com) for DNA uptake in Gram-positive bacteria (14), and archaeal flagella (recently renamed archeaella [15]) all have structural homologs of proteins found in T4P systems, including the major building blocks of the filament (16, 17) (Fig. 1). The central proteins in all these systems in Gram-negative bacteria are (i) the major pilin subunit (including the archeaellar flagellin), (ii) a prepilin peptidase that processes pilins to a mature form, (iii) an assembly ATPase, (iv) an inner membrane core protein, and (v) an outer membrane secretin channel (Fig. 1). Conserved sequences within these proteins can be used to identify T4P and related systems in any organism for which the genome sequence has been determined.

T4P were first described for Pseudomonas aeruginosa, in the 1970s (18, 19), and subsequently were identified on a vast number of Gram-positive bacteria, including many human pathogens. They were initially considered to be exclusive to Gram-negative bacteria, in part based on the architecture of the assembly apparatus, which spans both bacterial membranes and the intervening periplasm. Filaments resembling T4P were reported for Streptococcus sanguinis and Clostridium difficile (20, 21), two Gram-positive pathogenic bacteria, but T4P-dependent twitching or gliding motility was not described. The availability of multiple-strain genome sequences for Clostridium perfringens led to the identification of T4P gene products in that species and a realization that all of the clostridial species sequenced up to that time (2006) contained likely T4P-associated gene products (22). Since then, hundreds of additional Gram-positive bacterial genomes have been sequenced and it is clear T4P genes are nearly universal in clostridia and also occur sporadically in other families.

The goal of this review is to assimilate available data for Gram-positive T4P and compare them with T4P in Gram-negative bacteria to define commonalities and differences in order to understand the evolution of these important molecular machines, particularly as they relate to pathogenesis. We also note many similarities between Gram-positive T4P and Gram-negative T2S systems. This is of profound interest because virtually every Gram-positive human pathogen secretes toxins as part of the infection process, yet the secretion mechanisms are not well defined. Determining if a T2S-like mechanism is utilized for toxin secretion in pathogens that have T4P will be valuable for understanding virulence mechanisms. Although the Com systems are related to T4P and are widespread in Gram-positive bacteria, they are not covered in this review; instead, the reader is directed to other reviews of the Com systems (14, 23). Tad operons (for tight adherence), believed to encode a subset of T4P, are also present in many Gram-positive species (24) but have not been studied to any extent and are not discussed extensively here. We focus on the T4P systems in two Clostridium species, C. perfringens and C. difficile, chosen because they are the best-characterized examples so far of T4P in Gram-positive bacteria. These important human and animal pathogens are only very distantly related within the extremely diverse Clostridiales group (25); thus, the T4P features they have in common may be shared by many clostridia. Many of the observations and comparisons presented here are based on genomic sequences; new genome sequences coming into the public domain on nearly a daily basis are likely to reveal new information and, possibly, exceptions to the general trends we discuss in this review. We show here that the clostridial pilus systems share features with the type IVa and IVb pilus systems and the T2S systems, consistent with them being the precursors from which these various systems evolved. Because there is a vast array of research related to T4P and related systems, we apologize to investigators whose work is not cited in this review.

**OVERVIEW OF T4P IN GRAM-NEGATIVE BACTERIA**

The major components that make up the Gram-negative T4P and T2S systems and the Gram-positive T4P systems are illustrated in Fig. 1 and listed in Table 1. Unfortunately, the nomenclature for homologous proteins varies from species to species. To minimize
TABLE 1 Nomenclature for related filament systems in Gram-negative and Gram-positive bacteria

<table>
<thead>
<tr>
<th>Protein category</th>
<th>T4P&lt;sup&gt;a&lt;/sup&gt;</th>
<th>T2S&lt;sup&gt;b&lt;/sup&gt;</th>
<th>TAD pilin&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Gram-positive T4P&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major pilins/pseudopilins</td>
<td>PilA</td>
<td>GspG</td>
<td>Flp</td>
<td>PilA</td>
</tr>
<tr>
<td>Minor pilins/pseudopilins</td>
<td>FimU, PilV, PilW, PilX, PilE</td>
<td>GspH, GspI, GspJ, GspK</td>
<td>TadE, TadF, TadG</td>
<td>Various proteins</td>
</tr>
<tr>
<td>Assembly ATPase</td>
<td>PilB</td>
<td>GspE</td>
<td>TadA</td>
<td>PilB</td>
</tr>
<tr>
<td>Inner membrane core protein</td>
<td>PilC</td>
<td>GspF</td>
<td>TadB, TadC</td>
<td>PilC</td>
</tr>
<tr>
<td>Prepilin peptidase</td>
<td>PilD</td>
<td>GspO</td>
<td>TadV</td>
<td>PilD</td>
</tr>
<tr>
<td>Outer membrane secretin</td>
<td>PilQ</td>
<td>GspD</td>
<td>RcpA</td>
<td>PilT</td>
</tr>
<tr>
<td>Retraction ATPase</td>
<td>PilT, PilU</td>
<td></td>
<td></td>
<td>PilT</td>
</tr>
<tr>
<td>Inner membrane accessory protein</td>
<td>PilM, PilN, PilO</td>
<td>GspL, GspM</td>
<td>TadG, TadZ</td>
<td>PilM, PilN, PilO</td>
</tr>
</tbody>
</table>

<sup>a</sup> *Pseudomonas aeruginosa* nomenclature.
<sup>b</sup> General secretion pathway nomenclature (170).
<sup>c</sup> *Pseudomonas aeruginosa* nomenclature.
<sup>d</sup> *Clostridium perfringens* nomenclature.

complexity, we designated the homologs of T4P proteins found in the clostridia in the same fashion as that for *P. aeruginosa*, i.e., PilA for the major pilin, PilB for the assembly ATPase, PilC for the inner membrane core protein, PilD for the prepilin peptidase, and PilT for the retraction ATPase (22). We encourage others who identify and characterize these proteins in other Gram-positive species to follow this practice. For this review, proteins are referred to by their descriptive generic names wherever possible, based on their sequence homology to proteins whose functions are well established.

The T4P filament is composed of thousands of copies of the major pilin polymerized from a pool of subunits in the cytoplasmic membranes of Gram-negative bacteria (Fig. 1). Pilins are synthesized in a prepilin form with a signal peptide that is rich in basic sequence on the cytoplasmic side of the membrane is removed by a dedicated membrane-bound prepilin peptidase, called PilD in many species (27–29). Only the processed pilin is competent for polymerization (30). The N-terminal ~50 residues of the mature pilin form an extended α-helix (31–33). The N-terminal hydrophobic half of this α-helix anchors the C-terminal globular domain in the inner membrane prior to pilus assembly and forms a helical array in the filament core in the assembled pilus filament, with the globular domain forming the outer shell of the filament (11, 12). Minor pilins share N-terminal sequence homology with the major pilin and act to optimize filament assembly or retraction; some have been found to be incorporated into the pilus itself and to influence pilus functions (34–44).

The mechanism by which pilin polymerization occurs is not well understood but involves, at a minimum, an assembly ATPase and an inner membrane core protein (45–51). One model that has been proposed is that upon binding ATP, the assembly ATPase undergoes a conformational change that is transmitted via the core protein to the base of the pilus, wherein the core protein functions as a lever that pushes the pilus out of the membrane a sufficient distance to allow a new pilin monomer to diffuse from the inner membrane into a gap at the base of the pilus filament (11) (Fig. 2). After hydrolysis of the ATP to ADP and inorganic phosphate (P<sub>i</sub>), the complex returns to its resting state and the cycle is repeated, iteratively adding one subunit at a time to the base of the filament. Pilus assembly requires three or more additional inner membrane accessory proteins that recruit the amphi-tropic ATPase, stabilize the inner membrane assembly platform, and interact with T4P components in the outer membrane (45, 46, 52–59). For pilus retraction, the assembly process is presumably reversed and involves a retraction ATPase, i.e., PilT or PilU (Fig. 1) (60–63), although this motor enzyme is not present in all T4P systems (2). The mechanism by which the retraction ATPase induces pilus disassembly is not understood, but the force generated by this process has been measured as >100 pN for a single PilT hexamer, making it the strongest biological motor that has been examined thus far (64).

In Gram-negative bacteria, the pilus must pass through the outer membrane barrier. This is accomplished via the secretin protein, which forms an oligomeric gated channel in the outer membrane (Fig. 1) (55, 65–69). The periplasmic domains of the inner membrane accessory proteins PilN and PilO contact the secretin channel via PilP, forming a periplasmic conduit through which the pilus grows (not shown in Fig. 1) (52, 55). Gram-positive bacteria have different requirements because they lack an outer membrane and have a thick peptidoglycan layer. The pilus (sometimes called fimbriae [70]) that are most commonly associated with Gram-positive bacteria are assembled by sortases, which polymerize pilin subunits at the base of the filament and then covalently attach the filament to either the peptidoglycan or the cytoplasmic membrane (reviewed in reference 71).

**GRAM-POSITIVE SPECIES HAVE T4P**

A few T4P- and T2S-related genes were found by chance in Gram-positive species, mostly clostridia, due to proximity to an unrelated gene of interest. The first full genome sequence obtained for a *Clostridium* species was that of *Clostridium acetobutylicum* ATCC 824 (72), a nonpathogenic bacterium used in solvent production. The T4P-related genes found in that strain were initially annotated as encoding components of the general secretory pathway (72), of which T2S represents the terminal branch, but were later described as a fimbrilin protein exporter system (73). Once
sequences became available for 3 strains of C. perfringens (74, 75), a Gram-positive pathogen, T4P-like genes were identified that were highly conserved among the strains (22). Immunofluorescence studies demonstrated the presence of two putative pilin proteins on the surfaces of the bacteria, and short (200 to 300 nm) pili were visible on the surface by use of field-emission scanning electron microscopy (22). Intriguingly, these bacteria exhibit a distinctive type of gliding motility on agar surfaces, characterized by the alignment of individual bacteria in an end-to-end conformation to form chains that moved away from the central colony (22) (see Video S1 in the supplemental material). Because C. perfringens lacks flagellar and chemotaxis genes, it seemed likely that the T4P-like gene products were responsible for this motility. Mutations in the pilC1 (previously named pilC; see T4P Operons in Clostridia) and pilT genes, encoding a putative inner membrane core protein and a retraction ATPase, respectively, inhibited gliding motility and reduced expression of the surface pili (22). These results suggested that pilC1 and pilT are necessary for pilin polymerization and gliding motility, marking the first demonstration of T4P-dependent motility in a Gram-positive bacterium.

Analysis of the 7 other clostridial genome sequences available at that time, including that of the opportunistic enteric pathogen C. difficile, indicated that each carried T4P-associated genes (22). Clostridium beijerinckii, a species closely related to C. perfringens, has a similar arrangement of T4P-associated genes but also has flagella. In the liquid film on the surface of 0.5% agar plates, C. beijerinckii is highly motile via flagellum-mediated swimming, but on 4% agar plates, which have less water on the surface, the bacterium demonstrates gliding motility similar to that seen with the nonflagellated C. perfringens, suggesting that an increased viscosity of the surrounding medium forces a switch from swimming to gliding motility (22).

Most, if Not All, Species of Clostridia Sequenced to Date Have T4P-Associated Gene Products

The gene sequences for the key assembly proteins found in C. perfringens, PilA, PilB, PilC, and PilD (Table 1), were used to identify homologs present in other clostridia. The genome sequences of ~800 species, spread over the >800 species in the Clostridia class, were analyzed using the BLAST program (http://www.ncbi.nlm.nih.gov/), and homologs of pilB, pilC, and pilD were identified in each of these (data not shown). Among these were two clostridial species in genera of particular interest: Epulopiscium, comprising multiple-spore-forming bacteria large enough to be seen with the naked eye (76); and Heliobacterium, comprising the only photosynthetic Gram-positive bacteria (77). A useful program, PilFind, was developed by M. Pohlschroeder and colleagues to detect candidate pili in many Gram-positive species (24). PilFind takes advantage of conserved features of the type IV pilin amino acid sequence: a signal peptide that typically ends in Gly(-1), followed by a hydrophobic stretch of ~25 amino acids with a glutamate at position 5 of the mature protein. By searching for the canonical peptide-terminal motif (A/G)X3(D/E), PilFind identified candidate pilin genes in Gram-positive genomes, including several in C. perfringens and C. difficile (Table 2). Some of these gene assignments are questionable, as they are not associated with other pili assembly genes or their protein sequences bear little resemblance to those of type IV pilins. Others, such as CPE1839 and CPE1841, appear to be part of an uncharacterized C. perfringens T4P operon.

T4P in Other Gram-Positive Genera

Streptococcus sanguinis has both a T4P operon and a com operon, making it the only known species in the genus Streptococcus with both systems (78). Xu et al. suggested the T4P genes were acquired by horizontal gene transfer, perhaps from a clostridial species (78). Imam et al. (24) used the PilFind program to identify a T4P operon in a Bacillus species, NRRL_B-14911, that also possesses com and tad operons. To our knowledge, this is the only Gram-positive organism reported to have all three types of pilin-related complexes. It seems that T4P operons are ubiquitous in clostridia but rare in bacilli, whereas com operons are quite common in the bacilli but have not been found in any clostridial species thus far.

FIG 2. Model for T4P assembly. (A) Side view of the inner membrane assembly platform. In the pilus assembly cycle shown, a single pilin subunit docks into a gap in the growing pilus filament, attracted by complementarity between its negatively charged Glu (-) residue and the positively charged main chain amine on the N-terminal residue (+) of the terminal pilin subunit in the growing filament. ATP is hydrolyzed by the assembly ATPase, inducing a conformational change in the inner membrane core protein (IMCP) that extrudes the filament outward a short distance, opening up a new gap ~120° around the base of the filament for a new pilin subunit to dock. Thus, subunits are added iteratively at 3 sites around the base of the growing filament, corresponding to each of the three helical strands in the T4P, shown in red, blue, and yellow. Each subunit is staggered by an axial distance of 8 to 10 Å along the length of the filament. Only one of three predicted inner membrane core proteins is shown. IM, inner membrane. (B) Top view of the assembly apparatus, looking down on the growing filament. The red subunit is added first, followed by the blue, then the yellow, etc. (Modified from reference 11 with permission from Elsevier.)
Most of the clostridial genomes we have analyzed thus far possess multiple gene clusters encoding T4P components (22). Both *C. perfringens* and *C. difficile* have a T4P operon, referred to here as the primary operon, that encodes all the proteins required to produce a pilus, plus additional gene clusters that encode single pilin proteins and/or pilus assembly components (Fig. 3). Because some of the T4P genes on the primary operons of *C. perfringens* and *C. difficile* were discovered after their homologs on the secondary operons, they are given a designation of “2,” whereas those genes that were identified initially are given a designation of “1.” The primary T4P operon encodes one or more major pilins (PilA; *CD3513* in *C. difficile*), a prepilin peptidase (PilD), an assembly ATPase (PilB), an inner membrane core protein (PilC), inner membrane accessory proteins (PilM and PilO; and PilN in *C. perfringens*), three putative minor pilins, and, in the case of *C. difficile*, a retraction ATPase. These primary operons resemble those encoding type IVb pili in *Vibrio cholerae*, enterotoxigenic *Escherichia coli* (ETEC), enteropathogenic *E. coli* (EPEC), and *Citrobacter rodentium*, where all pilus assembly proteins are encoded on a single operon (79–84). Such an arrangement is also seen for the genes encoding the T2S system (85). In contrast, genes encoding type IVa pilus assembly proteins are more numerous and are distributed in small clusters throughout the genome (2).

In addition to the apparently complete primary T4P operon, both *C. perfringens* and *C. difficile* possess two or three secondary gene clusters that encode a single pilin and/or one or more pilus assembly components (Fig. 3). Both organisms have a cluster of three syntenic genes, encoding a pilin, an assembly ATPase (PilB1), and an inner membrane core protein (PilC1). *C. perfringens* has additional gene clusters that encode single pilins and/or one or more pilus assembly components (Fig. 3). Because some of the T4P genes on the primary operons of *C. perfringens* and *C. difficile* were discovered after their homologs on the secondary operons, they are given a designation of “2,” whereas those genes that were identified initially are given a designation of “1.” The primary T4P operon encodes one or more major pilins (PilA; *CD3513* in *C. difficile*), a prepilin peptidase (PilD), an assembly ATPase (PilB), an inner membrane core protein (PilC), inner membrane accessory proteins (PilM and PilO; and PilN in *C. perfringens*), three putative minor pilins, and, in the case of *C. difficile*, a retraction ATPase. These primary operons resemble those encoding type IVb pili in *Vibrio cholerae*, enterotoxigenic *Escherichia coli* (ETEC), enteropathogenic *E. coli* (EPEC), and *Citrobacter rodentium*, where all pilus assembly proteins are encoded on a single operon (79–84). Such an arrangement is also seen for the genes encoding the T2S system (85). In contrast, genes encoding type IVa pilus assembly proteins are more numerous and are distributed in small clusters throughout the genome (2).

### TABLE 2 Pilins and other prepilin peptidase-dependent proteins found in *C. perfringens* and *C. difficile*

<table>
<thead>
<tr>
<th>Strain</th>
<th>Pilins</th>
<th>Putative minor pilins</th>
<th>Other putative PilD-dependent proteins&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. perfringens</em></td>
<td>PilA1/CPE2288 (12.6)</td>
<td>CPE2280 (38.1)</td>
<td>CPE0130 (3.1)</td>
</tr>
<tr>
<td></td>
<td>PilA2/CPE2284 (18.1)</td>
<td>CPE2279 (78.0)</td>
<td>CPE1839 (153)</td>
</tr>
<tr>
<td></td>
<td>PilA3/CPE2278 (20.0)</td>
<td>CPE2278/PilA3 (20.0)</td>
<td>CPE1841 (18.8)</td>
</tr>
<tr>
<td></td>
<td>PilA4/CPE1842 (15.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>C. difficile</em></td>
<td>CD630_3513 (17.2)</td>
<td>CD630_3508 (20.6)</td>
<td>CD630_1242 (13.6)</td>
</tr>
<tr>
<td></td>
<td>CD630_3294 (11.6)</td>
<td>CD630_3507 (19.8)</td>
<td>CD630_1245 (19.3)</td>
</tr>
<tr>
<td></td>
<td>CD630_2305 (17.2)</td>
<td>CD630_3506 (57.7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CD630_0755 (29.1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Proteins identified by PilFind software (http://signalfind.org/pilfind.html) as having a pilin-like N-terminal sequence.

**FIG 3** T4P gene clusters in *C. perfringens* and *C. difficile*. The primary T4P operon is shown at the top for *C. perfringens* strain 13 (A) and *C. difficile* strain 630 (B), and the secondary gene clusters are shown below. Protein annotations are based on sequence homology with proteins of known function, gene organization, and functional analysis. Genes are colored the same as their protein products shown in Fig. 1 and 2.
type IV pilin genes. As PilA2 is the major pilin and has a hydrophobic N-terminal residue, it is the polymerization domain, the presence of this implies evolutionary selection, the exact nature of which is unknown. In some obligate human pathogens, such as the pathogenic Neisseria species, pilin sequence heterogeneity occurs via a dedicated homologous recombination system that helps these organisms to resist antibody-mediated host immune responses (94). However, rather than being an obligate pathogen, C. perfringens is one of the most ubiquitous bacteria found in nature (95), and tissue infections of living animals likely represent a minor niche, suggesting that immune evasion does not exert a strong selective pressure. What seems more likely is that the PilA2 pilus, which has been shown to be important for attachment to host cells (see Functions of T4P Systems in Gram-positive Bacteria), might be highly variable to facilitate attachment to different surfaces, including but not exclusive to animal cells.

C. difficile

The seven putative pilins of C. difficile have fairly short signal peptides (5 to 12 residues for strain 630) (Fig. 5A). The mature proteins vary in length (from 108 to 262 residues) and are dissimilar in sequence beyond their N-terminal 25 residues, with the pilins encoded at loci CD3513 and CD2305, carried on separate operons, being related most closely overall (Fig. 5B). The N-terminal sequence is hydrophobic, with the exception of Glu5, but is not as well conserved among the seven pilins within a single C. difficile strain as it is in the C. perfringens pilus. The arrangement of the pilin genes on the C. difficile primary T4P operon suggests that CD3513 encodes the major pilin and CD3508, CD3507, and CD3506 encode minor pilins. Similar to the C. perfringens putative
minor pilin CPE2279, CD3506 is large for a pilin (499 amino acids) and lacks a Glu5 residue, suggestive of a GspK family protein. CD3506 also has an alanine, rather than phenylalanine, at position 1. Individual pilins are highly conserved among the three C. difficile strains (see Fig. S2 in the supplemental material) but do not resemble the C. perfringens pilins beyond their hydrophobic N termini (not shown).

Predicting Type IV Pilin Structures in Clostridia

The Gram-negative type IV pilins have a conserved architecture of an ~53-amino-acid N-terminal α-helical stalk, α1, and a globular C-terminal domain with a central 4- or 5-stranded antiparallel β-sheet (1). α1 anchors the pilin subunit in the inner membrane prior to pilus assembly, and it holds together the subunits in the pilus filament via hydrophobic interactions (11, 12). The regions flanking the β-sheet vary substantially in sequence, length, and structure among the pilins. These regions, i.e., the αβ-loop between α1 and the β-sheet and the D-region, delineated by the disulfide bond formed by two conserved cysteines, in large part define the pilus surface and at least some of its functions. Though similar in overall architecture, the type IVa and IVb pilins have distinct topologies: the D-region and C-terminal loop of the type IVa pilins are continuous and lie at the periphery of the globular domain, whereas the C-terminal segment of the type IVb pilins forms the central strand of the β-sheet (1). The clostridial type IV pilins lack the conserved cysteines but share the hydrophobic N-terminal segment, and their assembly proteins are homologous to those found in Gram-negative T4P systems. Thus, they are likely to share the canonical type IV pilin architecture. Structure predictions using the online servers Phyre2 (96) and Fugue (97) yielded high-confidence models for C. perfringens PilA1 and PilA2 and C. difficile pilins CD3513, CD3294, CD2305, and CD0755, based on sequences and predicted secondary structure homologies with full-length type IVa pilin structures from Neisseria gonorrhoeae.

FIG 4 Amino acid sequence similarity among C. perfringens pilins. (A) Amino acid sequences of the signal peptides and N-terminal 30 residues of the mature pilins from C. perfringens strain 13. The red background indicates amino acid sequence identity; the yellow background and bold letters indicate sequence similarity (nonbolded amino acids within each yellow column are not conserved). Residues are numbered above the sequences: negative numbers are used for the signal peptide, and positive numbers are used for the mature pilin sequence. (B) Schematic comparison of C. perfringens strain 13 pilins and putative pilins. Asterisks indicate putative minor pilins, among which CPE2279 is a GspK family member based on its large size and a tyrosine at position 5. The identities of the first and fifth amino acids of the mature protein are indicated. Gray bars represent the signal peptide, and yellow bars indicate amino acid sequence homology among the pilins. (C) Schematic comparison of pilins among multiple C. perfringens strains. Red bars indicate sequence identity for the majority of the residues among different strains, and blue bars indicate amino acid insertions/differences.
Pseudomonas aeruginosa, and Dichelobacter nodosus (Protein Data Bank [PDB] ID 2PIL, 1OQW, and 3SOK, respectively), as well as the major T2S pseudopilin PulG structure from Klebsiella pneumoniae (PDB ID 1T92). We were able to model only fragments of the remaining pilins, primarily for the N-terminal -helices. Despite the noted similarities between clostridial type IV pilins and the Gram-negative type IVb pilins, no models were obtained using type IVb pilins as templates, even though there are several such structures available in the Protein Data Bank.

Predictive models are shown for PilA1 and PilA2 from C. perfringens and the CD3513 pilin from C. difficile (Fig. 6), all of which are predicted to be major pilins based on their gene positions within the primary T4P operons and, in some cases, based on functional data. The αβ-loop and D regions are colored green and magenta, respectively, based on their respective positions in the template structures. C. perfringens PilA1 is a 117-amino-acid protein and is thus too short to possess a D region (Fig. 6A). High-quality models were obtained for C. perfringens PilA2 by using both N. gonorrhoeae GC pilin and D. nodosus FimA as templates. The FimA-based model is shown in Fig. 6B, because FimA, like the clostridial pilins, lacks the conserved cysteine pair that delineates the D-region in most Gram-negative type IV pilins yet retains the conserved D-region loop found in the type IVa pilins (33). The D-region loop in FimA is stabilized by hydrophobic interactions, and this may also be the case for some of the clostridial type IV pilins. The cysteines are not conserved in D. nodosus, C. perfringens, and C. difficile, most likely because these anaerobic bacteria live in a highly reducing environment that does not support disulfide bond formation. However, the cysteines are also lacking in the type IV pilins of two aerobic Gram-positive bacteria, S. sanguinis and Bacillus sp. NRRL_B-14911, suggesting that the absence of cysteines may be a characteristic of T4P in Gram-positive bacteria.

While the clostridial pilin models shown in Fig. 6 were derived from type IVa pilins, high-confidence models were also produced based on the T2S pseudopilin PulG from K. pneumoniae (98). This is despite the fact that the PulG crystal structure is for a truncated protein lacking the N-terminal 28-residue segment that bears the highest degree of sequence similarity to the Gram-positive pilins. Pseudopilins generally have smaller globular domains than type IV pilins. Most, like PulG, have a 3-stranded β-sheet and no disulfide-bonded cysteines. A characteristic feature of the major pseudopilins is a conserved loop–α-helix–loop structure in the αβ-loop region, between the N-terminal α-helix and the β-sheet, which contains a signature motif (99). In contrast, the type IV pilins display substantial sequence and structural variability in the αβ-loop. Although clostridial pilins can be modeled as pseudopilins, the pseudopilin αβ-loop signature motif is not present in the pilins we examined. Atomic structures of the clostridial pilins, once available, will likely aid in their classification.
and provide insights into their filament architecture and functions.

**T4P ASSEMBLY COMPONENTS IN C. PERFRINGENS AND C. DIFFICILE**

**Prepilin Peptidases (PilD)**

*C. perfringens* possesses a single prepilin peptidase gene, *pilD*, encoding a homolog of the PilD prepilin peptidase (22), but *C. difficile* has two putative prepilin peptidase genes, *pilD1* (*CD3503*) and *pilD2* (*CD3504*), that are adjacent in the primary T4P cluster in strain 630 (Fig. 3). The three clostridial PilD proteins were aligned with each other and with PilD proteins from T4P systems and a T2S system from Gram-negative species (Fig. 7; see Fig. S3 in the supplemental material). Interestingly, *C. difficile* PilD2 is more similar in length and sequence to the single *C. perfringens* PilD homolog (41% identity for 254 residues) than to *C. difficile* PilD1 (25% for 207 residues), suggesting a common origin for their corresponding genes, with *C. difficile* pilD1 perhaps arising from duplication of pilD2 followed by genetic drift to its present form. Furthermore, *C. perfringens* PilD and *C. difficile* PilD2 align more closely with the prepilin peptidases from the type IVa pilus and T2S systems than with those of the type IVb pilus. *C. difficile* PilD1 appears to have an N-terminal truncation that is not seen in the other prepilin peptidases analyzed. The prepilin peptidases are polytopic membrane proteins belonging to the GXGD-type aspartyl protease family, which includespreflagellin peptidase and presenilin, a human enzyme that plays an important role in...
Alzheimer’s disease (27, 100–102). All three PilD proteins possess two conserved aspartates that are critical for peptidase activity, the second of which lies in the GXGD consensus sequence. In terms of substrate specificity, the clostridial prepilin peptidases likely do not discriminate between the lengths of the pilin signal peptides, as these vary substantially within *C. perfringens* (Fig. 4A), which appears to have only a single PilD protein. However, whereas most of the *C. perfringens* and *C. difficile* pilins have a highly positively charged sequence immediately preceding the target glycine at the /H11002/1 position of the signal peptide, the putative minor *C. difficile* pilins, CD3508, CD3507, and CD3506, which are encoded in the same operon as the two PilD proteins, have a glutamate at the /H11002/3 position. Thus, *C. difficile* PilD1 may recognize this negative charge near the cleavage site, with the more closely related *C. difficile* PilD2 and *C. perfringens* PilD having specificity for the more positively charged signal peptides found on the majority of the clostridial pilins. Interestingly, glutamates are also present in type IVb pilin signal peptides: at the /H11002/3 position in EPEC bundlin and at the /H11002, /H11002, /H11002, /H11002, /H11002, /H11002 positions in *V. cholerae* TcpA, along with an aspartate at the /H11002/9 position. The *C. perfringens* and *C. difficile* PilB2 proteins both possess an N-terminal ~60-residue extension present in the type IVa pilus and T2S assembly ATPases (40 to 50% identity) than to those of the type IVb enzymes. Secretion ATPases have a bilobed structure that binds ATP in its cleft via the Walker A and His boxes, and the tetracysteine motif found in other secretion ATPases (105, 106) (Fig. 8). The *C. perfringens* and *C. difficile* PilB2 proteins lack this segment, as do the type IVb pilus assembly ATPases and the T2S ATPases (Fig. 8). The N-terminal regions of assembly ATPases from *V. cholerae* T2S systems and type IVb pili interact with cytoplasmic domains of inner membrane accessory proteins (46, 111). Thus, differences in the N-terminal domains suggest that there are unique interaction partners for the two ATPases within each clostridial system.

**Assembly ATPases (PilB)**

There are two assembly ATPases, PilB1 and PilB2, in both *C. perfringens* (22) and *C. difficile*. These proteins belong to the GspE superfamily of secretion nucleoside triphosphatases (NTPases) (104). In each species, PilB2 is encoded on the primary T4P operon, along with all key pilus assembly components, whereas PilB1 is encoded on a secondary operon that also encodes a pilin and an inner membrane core protein (Fig. 3). Alignment of the four clostridial PilB proteins with T4P and T2S assembly ATPases shows significant sequence homology (40 to 50% identity), including the signature Walker A and Walker B boxes, the Asp and His boxes, and the tetracysteine motif found in other secretion ATPases (105, 106) (Fig. 8; see Fig. S4 in the supplemental material). All of the clostridial PilB sequences are more similar to those of the type IVa pilus and T2S assembly ATPases (40 to 50% identity) than to those of the type IVb enzymes. Secretion ATPases have a bilobed structure that binds ATP in its cleft via the Walker, Asp, and His boxes, and the tetracysteine motif found in other secretion ATPases (105, 106) (Fig. 8; see Fig. S4 in the supplemental material). All of the clostridial PilB sequences are more similar to those of the type IVa pilus and T2S assembly ATPases (40 to 50% identity) than to those of the type IVb enzymes. Secretion ATPases have a bilobed structure that binds ATP in its cleft via the Walker, Asp, and His boxes, and the tetracysteine motif found in other secretion ATPases (105, 106) (Fig. 8; see Fig. S4 in the supplemental material).
PilC proteins interact, either directly or indirectly, in both Gram-negative and Gram-positive bacteria.

**Inner Membrane Accessory Proteins**

The clostridial proteins PilM, PilN, and PilO are not closely related in sequence to their homologs in Gram-negative bacteria, but they share key features with each system. These putative membrane accessory proteins are encoded on the primary operons of C. perfringens and C. difficile (Fig. 3). C. perfringens PilM belongs to the HSP70 actin superfamily and is predicted by the membrane topology server TMHMM (http://www.cbs.dtu.dk/services/TMHMM/) to be a cytoplasmic protein, while PilN in PilO each have a single N-terminal transmembrane segment and a periplasmic C-terminal domain (Fig. 10). The predicted membrane topologies of the C. perfringens PilM, PilN, and PilO proteins resemble those of the Gram-negative T4P. The crystal structure of the cytoplasmic PilM protein from *Thermus thermophilus* was solved in complex with a peptide corresponding to the N-terminal cytoplasmic segment of PilN, implying that PilN anchors PilM to the inner membrane (56). PilM is an ATP binding protein that resembles the actin-like protein FtsA. The periplasmic domain of *P. aeruginosa* PilO has a ferredoxin-like fold, as shown by its crystal structure, and the corresponding region of PilN is thought to have a homologous structure, based on predicted secondary structure similarity with PilO (53). These periplasmic PilN/PilO domains interact to form a stable heterodimer. Furthermore, the PilM, PilN, and PilO proteins, which are encoded on a dedicated pilM-NOPQ operon in Gram-negative bacteria, form a complex that connects the inner membrane pilus assembly platform to the outer membrane secretin complex PilQ via the periplasmic lipoprotein PilP (52, 54, 55, 118–120). While *C. perfringens* lacks PilP and PilQ homologs, the presence of an apparent cytoplasmic membrane complex that extends into the region corresponding to cyto1. X-ray crystal structures for the cyto1 domains of each of the type IVa, type IVb, and T2S PilC homologs show a conserved 6-helix bundle predicted to interact with the cytoplasmic face of the inner membrane, exposing the conserved Glu and Gly residues in an otherwise hydrophobic groove on the face of the bundle that is directed away from the membrane (113–115). This face may interact with partner proteins in filament assembly, including the assembly ATPase and cytoplasmic domains of the accessory proteins. While most of cyto1 forms the 6-helix bundle in *V. cholerae* type IVb TcpE, *Thermus thermophilus* PilC (type IVa) and *V. cholerae* EpsF (T2S) have an additional ~60 residues at the N terminus that were not resolved in their crystal structures. This segment appears to also be present in the clostridial PilC2 protein and may represent a specialized cytoplasmic binding site in these proteins.

It is not clear whether the ATPase and inner membrane core protein encoded on the secondary T4P gene cluster work with assembly components encoded on the primary T4P operon or use as yet unidentified components to make pili. Presumably, each ATPase-core protein pair operates within its own inner membrane assembly platform. In *P. aeruginosa*, a recombinant PilB protein fused to yellow fluorescent protein (PilB-YFP) was found to localize to the poles of the cells in a PilC-dependent manner, as this PilB-YFP fusion protein was diffusely distributed throughout the cytoplasm in a pilC mutant background (116). A similar situation was seen in *C. perfringens*, where a PilB1-YFP fusion protein was found to localize to the poles of the cell but polar localization was lost in a strain lacking PilC1 (117). This suggests that PilB and PilC proteins interact, either directly or indirectly, in both Gram-negative and Gram-positive bacteria.

**Inner Membrane Core Proteins**

Genes encoding the PilB secretion NTPases are typically paired with those encoding PilC, the inner membrane core protein, also referred to as the platform protein or the polytopic inner membrane protein. PilC proteins belong to the GspF superfamily of secretion proteins (112). Since the clostridia lack an outer membrane, we refer to their PilC homologs as “membrane core proteins” or just “core proteins.” The *C. perfringens* and *C. difficile* Cyto1 domains are conserved in sequence to their homologs in Gram-negative and Gram-positive bacteria.

**FIG 9** Alignment of PilC inner membrane core proteins. Proteins are listed based on the alignment shown in Fig. S5 in the supplemental material, with the clostridial sequences shown in gray. Absolutely conserved residues are indicated. These correspond to Glu82, Gly85, Pro115, and Pro135 in *C. perfringens* strain PAO1 PilC (GenBank accession number P22609), *N. gonorrhoeae* MS11 PilG (AAC43469), *C. perfringens* 13 PilC2 (NP_563201.1) and PilC1 (NP_562759), *C. difficile* 630 PilC2 (YP_001089811) and PilC1 (YP_001090032), *V. cholerae* N16961 TcpE (P0C6C9) and EpsF (NP_232358), *K. oxytoca* PulF (T2S) is directed away from the membrane (113–115). This face may interact with partner proteins in filament assembly, including the assembly ATPase and cytoplasmic domains of the accessory proteins. While most of cyto1 forms the 6-helix bundle in *V. cholerae* type IVb TcpE, *Thermus thermophilus* PilC (type IVa) and *V. cholerae* EpsF (T2S) have an additional ~60 residues at the N terminus that were not resolved in their crystal structures. This segment appears to also be present in the clostridial PilC2 protein and may represent a specialized cytoplasmic binding site in these proteins.
the periplasm of Gram-negative bacteria implies that the pilus assembly apparatus extends into the peptidoglycan layer of *C. perfringens*.

*C. difficile* PilM2, encoded on the *C. difficile* primary T4P operon, appears to be the product of a gene fusion between pilM and pilN: it is predicted to have a single transmembrane segment with an N-terminal cytoplasmic domain that is homologous to PilM from Gram-negative bacteria and a C-terminal periplasmic domain that is homologous to PilN (Fig. 10). A discrete pilN gene is not present in *C. difficile*. The type IVb pili and the T2S systems also have only two transmembrane accessory proteins, one with both a cytoplasmic domain and a periplasmic domain, analogous to *C. difficile* PilM2, and the other with only a periplasmic domain, analogous to *C. difficile* PilO (Fig. 10). The cytoplasmic domains of the EPEC type IVb protein BfpC and the *V. cholerae* T2S protein PilO are homologous in structure to *T. thermophilus* PilM, despite very limited sequence homology (56, 59, 121). Similarly, the periplasmic domains of *V. cholerae* PilO and *E. coli* PilM have ferredoxin folds (111, 121) like that shown for *P. aeruginosa* PilO (53). Thus, each filament system has a cluster of membrane accessory proteins with a single cytoplasmic domain and two periplasmic domains, but this is accomplished using three proteins for *C. perfringens* and type IVa pili and only two proteins for *C. difficile*, EPEC, and the T2S systems.

Genes encoding the inner membrane accessory proteins of the type IVb and T2S systems are clustered in the same operon as other pilus assembly genes, as they are in both *C. perfringens* and *C. difficile* (Fig. 3). Interestingly, the secondary operon of *C. difficile*, which encodes a single type IV pilin (CD3294), a membrane core protein (PilC1), and an assembly ATPase (PilB1), also has a pilM gene, CD3293 (Fig. 3B). Its gene product, which we call PilM1, lacks a transmembrane segment and is predicted to be cytoplasmic, similar to PilM in *C. perfringens* and the type IVa pili. The two genes downstream of pilM1, CD3292 and CD3291, encode gene products with predicted topologies matching those of *C. perfringens* PilN and PilO but without recognizable sequence similarity to these proteins. Thus, this secondary operon of *C. difficile* appears to encode a minimal but complete set of T4P assembly proteins.

The role of the clostridial PilM protein on the cytoplasmic side of the cell membrane likely mimics its role in Gram-negative bacteria, which is to interact with the inner membrane core protein and the assembly ATPase to form a functional pilus assembly platform (45, 46, 59, 111, 122–125). The roles of the PilM, PilN, and PilO proteins on the outer face of the cell membrane are less obvious in the absence of a secretin channel and outer membrane. These proteins likely interact with proteins that are unique to Gram-positive bacteria to form a channel through the thick peptidoglycan cell wall that allows pilus extension and surface display, and may also allow passage of exoproteins for release into the extracellular milieu.

**Retraction ATPases (PilT)**

In Gram-negative T4P systems, pili that have been shown to be retractable possess a second ATPase, usually called PilT, that somehow facilitates active disassembly of the pili. Some species, including *P. aeruginosa*, possess two retraction ATPases: PilT and PilU (105, 126). PilT homologs have some amino acid sequence similarity to assembly ATPases, including the Walker A and B boxes and Asp and His boxes, but have unique sequence motifs (AIRNLIRE and GMQTXXXXLXXLXXXXXI) in their C termini (127, 128). These signature PilT motifs are also present in the putative PilT proteins of *C. perfringens* and *C. difficile* (Fig. 11; see Fig. S6 in the supplemental material). *C. perfringens* and *C. difficile* PilT proteins are 56% identical in sequence (http://www.cbs.dtu.dk/services/TMHMM/). This similarity is much higher than that of other proteins in the pilus assembly machinery. They also share ~50% identity with *P. aeruginosa* PilT, suggesting that conservation of this protein sequence is somehow beneficial for bacterial survival. Retraction ATPases have not been identified for the T2S systems or for *V. cholerae* or ETEC type IVb pili, but the EPEC type IVb pili possess a putative retraction ATPase, BfpF, that is required for EPEC dissemination in cultured epithelial cells and for full virulence (129, 130). However, the amino acid sequence of BfpF is much less conserved, with a poorly matched AIRNLIRE sequence (SISSKIRS) and no GMQT sequence (see Fig. S6).

In Gram-negative bacteria, pilT mutants tend to be hyperpili-
ated, because the pili can assemble but not disassemble (4). However, in at least one Gram-negative bacterium, Francisella tularensis, disruption of the pilT gene results in loss of pilus assembly (51). Similarly, a C. perfringens pilT mutant is nonpiliated (22). While the mechanism is not understood at this point, it may rely on a regulatory pathway that terminates pilus assembly when genes in the assembly/retraction machinery are disrupted.

An unusual feature of the pilT gene in C. perfringens is its location in an apparent operon with the ftsA and ftsZ genes (Fig. 3A). In most clostridia, the pilT gene is located within the primary T4P operon, as is the case with C. difficile (Fig. 3B), or it is nearby. Reverse transcription-PCR (RT-PCR) experiments with C. perfringens indicated that there is a transcript that runs from the pilT gene into the ftsA gene (131). The synteny of the pilT/ftsA/ftsZ genes is maintained in the relatives of C. perfringens that fall into cluster 1 (132), such as Clostridium botulinum and Clostridium novyi (unpublished data). Why pilT expression is linked to genes involved in creating the division septum is unknown, but an intriguing observation is that T4P in C. perfringens tend to be located mostly at the poles (117; unpublished data). Since the division septum represents a new pole after division and separation occur, it may be necessary for the pilus assembly/retraction proteins to be located in this area prior to cell division. Coordination of pole formation with pilus assembly is well established for Gram-negative bacteria (133, 134). Cowles and Gitai showed that the MreB cytoskeleton is required for PilT localization to the poles in P. aeruginosa (135). This may be true in Gram-positive bacteria as well.

Secretins

Secretins are proteins that form a ring-shaped oligomeric gated channel in the outer membrane of Gram-negative bacteria through which the pilus grows (55, 136–141). Because the large majority of Gram-positive bacteria lack an outer membrane, it is not surprising that no secretin homologs have been found, except in species in the genera Desulfotomaculum and Pelomaculum (24), which are closely related to each other. The secretin homologs found in these bacteria are smaller (375 amino acids) than those seen in Gram-negative bacteria, such as the P. aeruginosa PilQ protein (714 amino acids), and they appear to comprise only the C-terminal half of the Gram-negative secretins that oligomerize to form a transmembrane β-barrel. Whether or not the Desulfotomaculum homolog actually functions as a secretin remains to be determined, but it is located in the main T4P operon (24), and thus it is likely to play a role in T4P functionality. In addition to lacking secretins, Gram-positive bacteria also lack homologs of PilF and PilP, which are lipoproteins found in the type IVa systems that are involved in assembly of the secretin in the Gram-negative outer membrane and in anchoring the secretin complex to the T4P inner membrane complex, respectively (52, 55, 119, 120, 142).

Sporomusa and Selenomonas organisms are members of the Firmicutes that have an outer membrane (143) and have characteristics of both Gram-negative and Gram-positive bacteria. For example, Sporomusa organisms are able to produce heat-stable endospores, a feature usually restricted to Gram-positive bacteria (144). Selenomonas sputigena has a T4P system with a 447-amino-acid secretin homolog that is comparable in size to the Desulfotomaculum secretin, with 17.8% identity, but may be related more closely in evolutionary history to P. aeruginosa PilQ, with 26.9% sequence identity.

FUNCTIONS OF T4P SYSTEMS IN GRAM-POSITIVE BACTERIA

In Gram-negative bacteria, T4P have been linked to many cellular functions, including twitching motility, gliding motility, DNA uptake, protein secretion, adherence to eukaryotic cells, biofilm formation, and acting as nanowires carrying electric current (1, 2, 145–148). Since the functions associated with T4P in Gram-positive bacteria are just beginning to be investigated, only a few examples are discussed here.

Gliding Motility

C. perfringens lacks flagella but can glide on the surfaces of agar plates by lining up in an end-to-end orientation to form long filaments that can be seen moving away from a colony (22) (see Video S1 in the supplemental material). Cells in the filaments also appear to align laterally, with filaments sliding along one another. Mutations in the retraction ATPase pilT gene and the core protein pilC gene result in nonmotile C. perfringens strains that have no regular end-to-end connectivity (22), but the exact nature of the cell-cell interactions is still not known.

Unlike C. perfringens, most clostridia possess flagella and can swim via flagellum-mediated motility. As noted above, C. beijerinckii swims in the liquid film on the surface on 0.5% agar plates, but on 4% agar, it moves via gliding motility in a manner similar to that seen with C. perfringens. For C. difficile, another flagellated species, bacteria can be seen swimming on 0.5% agar media (see Video S2 in the supplemental material) and gliding when the agar...
concentration is increased to 2% (see Video S3). Thus, this type of motility switch may be common among the clostridia. It is not clear whether the increase in viscosity induces a switch in protein expression or simply allows the pili greater access to the agar surface and to each other. Mutagenesis of T4P-associated genes in these species will determine if the gliding motility is in fact dependent on T4P.

Adherence to Host Cells
Given that many Gram-negative pathogens utilize T4P as adherence organelles, it would not be surprising if Gram-positive pathogens that have T4P do the same. Evidence in support of this comes from experiments in which the C. perfringens PilA2 gene from strain 13 was heterologously expressed in N. gonorrhoeae lacking its native pilin, PilE (149). While N. gonorrhoeae normally adheres to genitourinary tract epithelial cells, the strain expressing C. perfringens PilA2 was unable to adhere to such cells but instead bound to myoblasts (muscle cells) and fibroblasts (149), which are host cell types that C. perfringens would encounter in a gas gangrene infection (150). In addition, C. perfringens itself adheres to myoblasts in a PilT-dependent manner (149). Since C. perfringens pilT mutants are nonpiliated, these results suggest that the T4P are important for adherence.

In hamsters infected orally with C. difficile strain 630, bacteria were associated with crypt cells, and immunogold labeling of these bacteria with antibodies to the CD3507 pilin revealed a pattern of gold particles bound to distinct pilus-like filaments projecting from the cells (151). Whether or not CD3507 or these filaments are involved in adherence remains to be determined. We predict, based on sequence and gene location, that CD3507 is a minor pilin expressed under normal expression conditions and that these filaments mediate characteristic T4P functions, such as motility and adhesion. Empirical data support the classification of the C. perfringens primary operon as encoding a T4P rather than a T2S system: PilA1 and PilA2 are present on the surfaces of C. perfringens cells; their expression correlates with the presence of pilus-like filaments on the bacterial surface and with gliding motility (22); disruption of either the core membrane protein, PilC1, or the putative retraction ATPase, PilT, results in a loss of pili and gliding motility (22); and expression of PilA2 in N. gonorrhoeae correlates with binding to muscle cells (149). Furthermore, this system possesses a retraction ATPase, an enzyme that is not associated with T2S systems. Examination of Gram-positive bacterial genomes with predicted T4P systems shows PilT homologs are present in most or all of these species, suggesting that active pilus retraction occurs.

The distinction between T4P and T2S systems is not an absolute one even among Gram-negative bacteria. Some T4P systems, like those of V. cholerae and D. nodosus, secrete protein substrates (7, 8), and some T2S systems have been shown to display pseudopilus filaments on the bacterial surface when the major pseudopilin is overexpressed (88, 160, 161). The distinction between T4P and T2S systems may not be meaningful for Gram-positive bacteria. The difference between displaying a long pilus on the bacterial surface versus a short pseudopilus in the periplasm is likely determined in part by interactions between the growing filament and the outer membrane secretin, which allows passage of T4P but not T2S pseudopilis. Given that most Gram-positive bacteria lack an outer membrane, all pilus filaments should, in theory, be displayed on the cell surface, provided that they can traverse the peptidoglycan layer. Some of these pili could function as pistons to extrude proteins that have been transported from the cytoplasm via the Sec secretion machinery in the cell membrane. Such a mechanism may explain how some of the many toxins and degradative enzymes are exported into the extracellular space by Gram-positive bacteria. A T4P/T2S system could provide a conduit in the peptidoglycan through which a pseudopilus could actively extrude larger folded proteins (Fig. 1). We predict that upon further investigation, a secretion function will be revealed for at least some of the Gram-positive T4P systems.

The T4P functions of C. perfringens are important for maximum biofilm formation, as mutants lacking the PilT and PilC1 proteins show significantly reduced numbers of bacteria in biofilms (153). Interestingly, immunofluorescence labeling of PilA1 and PilA2 in mature C. perfringens biofilms suggested that the pilin proteins are part of the extracellular matrix (153), a pattern seen in some Gram-negative bacterial biofilms (154).

EVCOLUTION OF TYPE IV PILI
The nature of a last universal common ancestor preceding eubacterial and archaeal evolution is not without controversy (155), but it seems clear that T4P-like structures must have been acquired early in organismal evolution, since they are widespread in eubacteria and archaea. Because many trees of life depict thermophilic clostridia as one of the most ancient types of eubacteria (e.g., see reference 156), and because T4P are ubiquitous in clostridia, it seems likely that a clostridium-like eubacterial organism is the source of T4P in all eubacteria. Our analysis shows that the clostridial T4P share commonalities with both the type IVa and IVb pilus systems and the T2S system but are not easily classified into a single system. These observations are consistent with clostridial T4P being precursors to all three systems.

How, then, did Gram-negative bacteria inherit T4P? One hypothesis regarding the origin of Gram-negative bacteria is that they are the result of an endosymbiotic event between an ancient clostridium and an ancient actinobacterium that produced a Gram-negative organism (157). If the ancient clostridium carried T4P genes as part of its genome, then this system would have been retained and evolved further in Gram-negative bacteria. The endosymbiotic origin of Gram-negative bacteria is certainly a controversial hypothesis (e.g., see reference 158), but it does explain the similarity of T4P systems throughout the Gram-positive and Gram-negative clades. Since the EPEC and ETEC type IVb pilus operons are located on virulence plasmids (84), it is also possible that the type IVb pilus components in enteric bacteria such as EPEC were acquired from a clostridial species via horizontal transfer of an intact T4P operon similar to the primary operons found in C. perfringens and C. difficile.
How Do Pili Pass through the Thick Peptidoglycan Layer of Gram-Positive Bacteria?

Regardless of the ultimate length and surface display of the pilus in Gram-positive organisms and whether or not they export proteins into the extracellular space, they must pass through the physical barrier of the peptidoglycan layer (Fig. 1). In *Bacillus subtilis*, the peptidoglycan layer is permeable only to globular proteins smaller than 25 kDa (162). Thus, it is unlikely that a solid pilus cylinder of 6 to 9 nm could pass through the cell wall without requiring significant local restructuring of the cell wall. Little is known about this aspect of the pilus assembly process for Gram-negative bacteria, which have a thinner but nonetheless substantial peptidoglycan barrier. FimV, a peptidoglycan binding protein in *P. aeruginosa*, has been shown to promote secretin complex formation and surface pilation (163). These functions are dependent on a peptidoglycan binding domain in FimV (163), suggesting that this protein provides a link between the peptidoglycan layer and the secretin to help guide the pilus through the periplasm and across the outer membrane. FimV homologs have not been identified in the clostridia; thus, they may use another system to allow passage of the pilus through the peptidoglycan layer. Peptidoglycan binding proteins have also been identified that participate in T2S system formation (for example, see reference 164). Most likely the PilM-PilN-PilO protein complex, which is conserved among the Gram-negative and Gram-positive T4P systems, plays a key role in ushering the pilus filament through the peptidoglycan cell wall, perhaps by linking the inner membrane assembly platform with a protein channel through this semipermeable barrier.

CONCLUSIONS AND FUTURE DIRECTIONS

It is clear from the evidence presented in this review that T4P are widespread among the clostridia and are also present in other Gram-positive species. We are only beginning to explore their structure, assembly mechanisms, and functions in this major class of bacteria and to understand the extent of their involvement in bacterial ecology and pathogenesis. The Gram-positive T4P represent an entirely new opportunity to study the mechanism and functionality of T4P and related filament systems. Since the clostridial T4P systems are among the simplest ones identified in bacteria thus far (2), as they require no outer membrane machinery, they should be excellent model systems for analysis of T4P assembly and retraction. In the past, genetic systems were limited for they should be excellent model systems for analysis of T4P assembly and use of fluorescent protein fusion tags (117) based on eukaryotic systems (167,168) have greatly accelerated progress and retraction. In the past, genetic systems were limited for they should be excellent model systems for analysis of T4P assembly and use of fluorescent protein fusion tags (117) based on eukaryotic systems (167,168) have greatly accelerated progress.

REFERENCES


65. Bose N, Taylor RK. 2005. Identification of a TcpC-TcpQ outer mem-


Downloaded from http://mmbr.asm.org on December 2, 2014 by The TMC Library


159. Reference deleted.


Stephen Melville received his Ph.D. from the University of California, Davis. He did postdoctoral training at the University of California, Los Angeles, in the laboratory of Robert Gunsalus, studying the global transcriptional regulatory protein FNR. He then did postdoctoral work in the laboratory of Abraham L. Sonenshein at Tufts University School of Medicine, where he began working with the Gram-positive pathogen *Clostridium perfringens*, which has been the focus of his research career ever since. After joining the University of Tennessee, Memphis (now The University of Tennessee Health Sciences Center), as an Assistant Professor in 1995, he moved to Virginia Tech as an Associate Professor in 2001, where he has remained. In recent years, his work has focused on the role of type IV pilus (T4P) in *Clostridium perfringens* gliding motility and pathogenesis. He is using the relatively simple T4P system in *C. perfringens* as a model to understand the basic mechanisms of T4P assembly in both Gram-positive and Gram-negative bacteria.

Lisa Craig received her Ph.D. from Simon Fraser University (SFU) in Burnaby, British Columbia, Canada, in 1998. She began her work on type IV pilus as a postdoctoral fellow in the laboratory of John Tainer at the Scripps Research Institute in La Jolla, CA. In 2005, she returned to SFU as a faculty member, where she is currently an Associate Professor in the Molecular Biology and Biochemistry Department. Dr. Craig has performed extensive structure-function studies on type IV pili from a number of Gram-negative bacterial pathogens, including *Vibrio cholerae*, *Neisseria gonorrhoeae*, and enterotoxigenic *Escherichia coli*. Her current research interests are focused on type IV pilus assembly and retraction and secretion mechanisms. Dr. Craig is funded by the Natural Sciences and Engineering Research Council and the Canadian Institutes of Health Research (CIHR). She is the recipient of a CIHR New Investigator Award and a Michael Smith Foundation for Health Research Scholar Award.