IAI Journals.ASM.org

Structural Conservation, Variability, and Immunogenicity of the T6 Backbone Pilin of Serotype M6 *Streptococcus pyogenes*

Paul G. Young,[a,b] Nicole J. Moreland,[a,b] Jacelyn M. Loh,[b,c] Anita Bell,[d] Polly Atatoa Carr,[e] Thomas Proft,[b,c] Edward N. Baker[a,b]

School of Biological Sciences, University of Auckland, Auckland, New Zealand[a]; Maurice Wilkins Centre, University of Auckland, Auckland, New Zealand[b]; Department of Molecular Medicine and Pathology, University of Auckland, Auckland, New Zealand[c]; Population Health, Waikato District Health Board, Hamilton, New Zealand[d]; Waikato Clinical School, University of Auckland, Hamilton, New Zealand[e]

Group A streptococcus (GAS; *Streptococcus pyogenes*) is a Gram-positive human pathogen that causes a broad range of diseases ranging from acute pharyngitis to the poststreptococcal sequelae of acute rheumatic fever. GAS pili are highly diverse, long protein polymers that extend from the cell surface. They have multiple roles in infection and are promising candidates for vaccine development. This study describes the structure of the T6 backbone pilin (BP; Lancefield T-antigen) from the important M6 serotype. The structure reveals a modular arrangement of three tandem immunoglobulin-like domains, two with internal isopeptide bonds. The T6 pilin lysine, essential for polymerization, is located in a novel VAKS motif that is structurally homologous to the canonical YPKN pilin lysine in other three- and four-domain Gram-positive pilins. The T6 structure also highlights a conserved pilin core whose surface is decorated with highly variable loops and extensions. Comparison to other Gram-positive BPs shows that many of the largest variable extensions are found in conserved locations. Studies with sera from patients diagnosed with GAS-associated acute rheumatic fever showed that each of the three T6 domains, and the largest of the variable extensions (V8), are targeted by IgG during infection *in vivo*. Although the GAS BP show large variations in size and sequence, the modular nature of the pilus proteins revealed by the T6 structure may aid the future design of a pilus-based vaccine.

Streptococcus pyogenes (group A streptococcus [GAS]) is a highly adapted obligate human pathogen that readily infects and colonizes the mucous membranes of the upper respiratory tract or the epidermal layer of the skin. Although mostly mild and self-limiting, these infections can lead to severe invasive diseases, such as necrotizing fasciitis and streptococcal toxic shock syndrome (1). The poststreptococcal sequelae of glomerulonephritis and acute rheumatic fever (ARF) with subsequent rheumatic heart disease are associated with significant disease burden in the developing world (2). Although ARF is now rare in most high-income countries, the rates in Indigenous populations in selected developed countries such as Australia and New Zealand are among the highest in the world (3, 4).

The surface of *S. pyogenes* is decorated with numerous virulence factors to aid adhesion and colonization. Prominent among GAS virulence factors are pili. Pili (or fimbriae) are proteinaceous filaments that extend from the bacterial surface. They are formed from a single chain of covalently linked proteins (pilins), usually comprising a major or backbone pilin (BP), which is polymerized into the shaft (5–7), and two ancillary pilins that function as an adhesin at the tip of the pilus (AP-1) (7, 8) and an adaptor protein (AP-2) at the pilus base for covalent linkage to peptidoglycan of the cell wall (8–10).

GAS pili have been shown to be effective protective antigens in mouse immunization studies and are promising candidates for vaccine development against GAS disease (5). The genes encoding the individual pilus proteins are located within the FCT (fibronectin-binding, collagen-binding, Lancefield T antigen) region. The collagen-binding protein and the T antigen are associated with the pilus, whereas several fibronectin-binding proteins are encoded outside the pilus operon. There is a much greater degree of variability among GAS pili compared to other streptococcal species, which is reflected in their use in strain typing (T-serotyping, *emm*-sequence typing) (11, 12). To date, nine FCT types have been documented (13, 14). The only known structure of a *S. pyogenes* BP is from the FCT 2 strain SF370 (serotype M1/T1) (6). This BP (Spy0128) comprises two tandem immunoglobulin (Ig)-like domains, each with a CnaB-like fold with reverse IgG topology (IgG-rev) (15). Structures of BPs from other Gram-positive bacteria show a modular structure similar to that of Spy0128, but with either three domains, as in *Corynebacterium diphtheriae* SpaA (16), *Streptococcus agalactiae* GBS80 (17), and *Actinomyces oris* FimA (18) and FimP (19), or four domains, as in *Streptococcus pneumoniae* RrgB (20), *Streptococcus agalactiae* BP-2a (21) and *Bacillus cereus* BcpA (22, 23). The three-domain pilins all have a domain with a CnaA-type (DEv-IgG) fold inserted between two CnaB domains: a CnaA domain also has a reverse Ig-like topology but contains two extra β-strands compared to a CnaB domain (24). The four-domain pilins either have an additional CnaB domain inserted between the N-terminal CnaB domain and the CnaA domain giving four tandem domains as in BcpA, or a CnaB domain inserted into the CnaA domain so that it is lateral to the CnaA domain, as in RrgB and BP-2a (25).

The high degree of divergence between different GAS pilus types is mirrored in both sequence diversity and length. Pili with the closest sequence similarity to the two-domain BP of FCT 2

Received 3 March 2014 **Returned for modification** 12 April 2014
Accepted 20 April 2014
Published ahead of print 28 April 2014
Editor: A. Camilli
Address correspondence to Paul G. Young, p.young@auckland.ac.nz.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/IAI.01706-14.

doi:10.1128/IAI.01706-14

(Spy0128) include those from FCT types 3, 4, 7, and 8 (13, 26). They have an approximate size of 350 amino acids and retain greater sequence identity between FCT types (14). The two-domain BPs are unusual since they colocalize with a signal peptidase-like gene that is essential for their polymerization into fimbriae (27–29) and lack a conserved YPKN pilin motif that is present in most other Gram-positive BPs. The YPKN motif presents the lysine residue that is ligated to the threonine of the pilus sortase recognition motif (a variant of the LPXTG cell wall anchor motif) of the next BP subunit during pilus polymerization (30). In the two-domain GAS BP the pilin lysine is instead presented on an omega loop on the last strand of the N-terminal CnaB domain (6).

Sequence identity between BPs from FCT types outside this two-domain group (FCT types 1, 5, 6, and 9) rarely exceeds 25%. This makes it difficult to predict structure, although their genomic organization suggests that these GAS pili may be more similar to pili from other *Streptococcus* species (26). The open reading frames of these GAS BPs range from 537 to 720 amino acids, suggestive of either a three- or four-domain structure analogous to either *S. agalactiae* GBS80 or BP-2a. With the exception of FCT 1 they all share the recognizable pilin features of a sortase recognition motif and a YPKN pilin motif.

The FCT 1 BP is unique as, like the smaller two-domain BP, it lacks the canonical YPKN pilin motif. The pilin lysine is instead presented in a novel VAKS motif (31). Another distinguishing feature of FCT 1 is that the pilus operon has only two structural proteins, an adhesin (FctX) and the BP (T6) (5). Unlike other pili, which are anchored to the cell wall peptidoglycan by a specialized ancillary protein (AP-2), in the mature FCT 1 pilus the BP (T6) is directly attached to the peptidoglycan by the housekeeping sortase A (SrtA) (31).

This study describes the high-resolution crystal structure of T6, a representative FCT 1 backbone pilin from *S. pyogenes*. The protein structure reveals a modular arrangement of three tandem Ig-like domains, two of which contain internal isopeptide bonds. The pilin lysine of the novel VAKS motif is presented in a context structurally equivalent to the lysine in the classical YPKN pilin motif observed in other Gram-positive species. The T6 structure highlights the nature of structural variation in major pilins, in which a common scaffold is decorated with highly variable extensions that provide different surface features and may contribute to strain specific immune responses.

MATERIALS AND METHODS

Cloning, expression, and purification of T6. The *tee6* gene comprising the extracellular region of the protein ($T6_{35\text{-}520}$) was PCR amplified from *S. pyogenes* strain MGAS10394 (ATCC BAA-946) genomic DNA using the gene-specific primers tee6 Fwd and tee6 Rev (see Table S1 in the supplemental material). The amplified fragments were digested with the appropriate restriction enzymes (BamHI and HindIII) and cloned into the vector pProEXHTa (Invitrogen). For recombinant T6 protein expression, *Escherichia coli* BL21(λDE3)/pRIL cells were transformed with pProExHTa:tee6 and grown in Luria-Bertani medium supplemented with the required antibiotics at 37°C until the optical density at 600 nm reached 0.6. The cultures were induced with 0.3 mM IPTG (isopropyl-β-D-thiogalactopyranoside) at 37°C for 3 h, and the cells were harvested by centrifugation. Cell pellets were resuspended in lysis buffer (50 mM Tris-HCl [pH 8.0], 500 mM NaCl, 2% [vol/vol] glycerol) supplemented with 10 mM imidazole and Complete protease inhibitor cocktail EDTA-free minitablets (Roche) and stored at −20°C.

Cells were lysed using a cell disruptor (Constant Cell Disruption Systems) at 18 kpsi. Insoluble matter was sedimented by centrifugation (30,000 × *g*, 4°C, 30 min), and the soluble phase was loaded onto a HiTrap chelating 5-ml column (GE Healthcare). Bound protein was washed with wash buffer (lysis buffer plus 20 mM imidazole) and eluted in a gradient with elution buffer (50 mM Tris-HCl [pH 8.0], 150 mM NaCl, 2% [vol/vol] glycerol, 500 mM imidazole). The His_6 affinity tag was cleaved from T6 recombinant protein with a 1:50 ratio of rTEV-His_6 and concurrently dialyzed against 25 mM Tris-HCl [pH 8.0]–100 mM NaCl at 4°C for 16 h. T6 was separated from the rTEV-His_6 protease and uncleaved protein by immobilized metal affinity chromatography (IMAC). The unbound protein containing T6 was concentrated using a 30-kDa molecular-mass cutoff protein concentrator (VivaScience) and purified by size-exclusion chromatography on a Superdex S200 10/300 column (GE Healthcare) in crystallization buffer (10 mM Tris-HCl [pH 8.0], 100 mM NaCl). T6 eluted in a single peak that corresponds to a monomer of ~50 kDa and was ca. 99% pure, as indicated by SDS-PAGE.

Crystallization. Vapor diffusion crystallization trials were carried out at 18°C using a Cartesian nanoliter dispensing robot (Genomic Systems) and a locally compiled crystallization screen (32). Initial T6 crystals were grown in 0.1-μl format and subsequently optimized in a hanging-drop vapor diffusion format. The crystals used for X-ray data collection were grown by mixing 1 μl of protein solution (20 mg/ml in 10 mM Tris-HCl [pH 8.0], 100 mM NaCl) with 1 μl of precipitant (2.0 to 2.2 M $NaKPO_4$ [pH 6.4], 2% PEG 400, 0.1 M imidazole [pH 6.2 to 6.5]) at 18°C.

Data collection and structure determination. T6 crystals were flash-cooled in liquid nitrogen without cryoprotectant. For phase determination, crystals were soaked in 1.6 M sodium potassium phosphate (pH 6.4), 2% PEG 400, 0.1 M imidazole (pH 6.5), 50 mM NaCl, and 5 mM Tris-HCl (pH 8.0) supplemented with 500 mM KI for 10 s prior to flash-cooling. X-ray diffraction data from the KI-soaked crystals were collected in-house (Micromax-007HF, Rigaku; MAR345DTB, MAR Research) at −163°C. All data sets were integrated using XDS (33), reindexed using POINTLESS (34), and scaled using SCALA (34). The structure of T6 was determined by single-wavelength anomalous diffraction phasing from the KI-soaked data, using Phenix Autosol, followed by autobuilding using Phenix (35, 36), which built 60% of the model. This partial model was used for molecular replacement into the higher-resolution native data. Native X-ray diffraction data were recorded at the Australian Synchrotron on the MX1 beamline (ADSC quantum 210r detector). The native structure was determined by molecular replacement using Phaser-MR (37). The T6 structure was subsequently refined using iterative cycles of manual building in COOT (38) and refinement with Buster (39) and REFMAC (40). The quality of the T6 model was inspected using the program MolProbity (41). Data collection and refinement statistics are shown in Table 1. All figures were generated using PyMOL. The T6 coordinates have been deposited in the worldwide Protein Data Bank (PDB) using PDB code 4P0D.

Truncation mutagenesis of T6. T6 truncation mutants were PCR-amplified from pProExHTa:tee6 using the gene-specific primers listed in Table S2 in the supplemental material. The amplified products were cloned into the KasI and SalI restriction sites of the pProExHTa expression vector. The T6-minus C-terminal extension clone was produced using inverse PCR site-directed mutagenesis (42). Briefly, a high-fidelity DNA polymerase (iProof; Bio-Rad) was used for the PCR amplification of the pBluescript:tee6 construct to produce a linearized PCR product with the C-terminal extension domain deleted. Template vector is removed by DpnI digestion, which digests only methylated parent DNA, and then recircularized by intramolecular ligation to produce a modified construct. Primers used for mutagenesis are listed in Table S2 in the supplemental material. All mutants were sequence verified and transformed into *E. coli* BL21(λDE3) cells. Protein expression and purification were performed as described for the wild-type T6 construct.

Cloning, expression, and purification of backbone pilins from M1, M2, M4, M28, and M49. The backbone pilin genes comprising the extracellular region of the protein were PCR amplified using gene-specific primers from genomic DNA of the *S. pyogenes* strains listed in Table S1 in the supplemental material. The amplified fragments were subcloned into

TABLE 1 Data collection, phasing, and refinement statistics[a]

Parameter	Native (4P0D)	KI
Data collection statistics		
Wavelength (Å)	0.95468	1.5418
No. of images	360	360
Oscillation angle (°)	1.0	1.0
Resolution range (Å)	24.48–1.90 (1.94–1.90)	52.0–2.04 (2.15–2.04)
Total no. of observations	665,673 (27,669)	558,639 (49,337)
Unique reflections	36,565 (1,592)	32,118 (4,018)
Redundancy	18.2 (17.4)	17.4 (12.7)
Space group	$P6_3$	$P6_3$
Unit cell dimensions		
a, *b*, *c* (Å)	104.2, 104.2, 84.2	104.2, 104.2, 84.2
α, β, γ (°)	90, 90, 120	90, 90, 120
Completeness (%)	89.1(57.9)	96.3 (83.0)
Avg $I/\sigma\langle I\rangle$	16.6 (3.8)	17.9 (1.6)
R_{merge}	0.148 (0.915)	0.104 (1.52)
CC(1/2)	0.998 (0.904)	0.37 (2.04)
Phasing statistics		
Anomalous completeness (%)		95.2 (77.9)
Anomalous redundancy (%)		8.9 (6.4)
No. of I sites		9
Figure of merit		0.41
Autobuilt residues		315
CC		0.72
R_{work} (%)		0.32
R_{free} (%)		0.41
Refinement statistics		
Resolution range (Å)	24.48–1.90 (1.94–1.90)	
Molecules/AU	1	
Solvent content (%)	52.0	
R_{work} (%)	20.9	
R_{free} (%)	25.0	
No. of protein atoms	3,517	
No. of water molecules	441	
RMSD from ideal geometry		
Bond length (Å)	0.008	
Bond angle (°)	1.26	
Mean B-factor (Å^2)		
Protein	20.0	
Water	28.9	
Ligands	27.7	
Ramachandran plot (%)[b]		
Most favored	97.4	
Outliers	0.22	

[a] Values for the outermost shell are indicated in parentheses. Mean(I) half-set correlation CC(1/2) as calculated by Scala. $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$; $R = \Sigma_{hkl}||F_o(hkl)| - |F_c(hkl)||/\Sigma_{khl}|F_o(hkl)|$. The R_{work} value is calculated using 95% of the data selected randomly and used in refinement. R_{free} is calculated from the remaining 5% of the data not used in refinement.

[b] Determined using MolProbity.

the expression vectors pProEXHTb (M2 and M4) or pET32a3C (M1, M28, and M49) and expressed in *E. coli* BL21(λDE3)/pLysS cells. In each case, expression was induced with 0.1 mM IPTG for 4 h at 30°C, and the recombinant protein was then purified by Ni^{2+} affinity chromatography by using Profinity IMAC Ni-charged resin (Bio-Rad). For proteins expressed in pET32a3C, the thioredoxin fusion protein was cleaved using 3C protease and subjected to Ni^{2+}affinity chromatography to separate the recombinant pilin from the cleaved thioredoxin tag. For proteins expressed in pProEXHTb the His_6 affinity tag was not removed. Further purification was performed using size exclusion chromatography on a Superdex S200 10/300 column (GE Healthcare) in phosphate-buffered saline (PBS).

T6 immunogenicity. Blood samples from patients diagnosed with acute rheumatic fever (ARF) and a preceding GAS infection confirmed by an elevated anti-streptolysin-O titer were collected at Waikato Hospital, Hamilton, New Zealand, as part of an ongoing study to examine GAS disease. Participants had provided written informed consent and the research protocols were approved by the Central Regional Ethics Committee (CEN/12/06/017). Serum was prepared by standard clotting methods and stored in aliquots at −80°C until use. For Western blots, pilus proteins (100 ng per lane) were separated on SDS–12% PAGE gels and electrotransferred using the wet transfer technique to polyvinylidene difluoride membranes (GE Healthcare). Membranes were blotted with patient serum diluted 1:500 in PBS (pH 7.2) supplemented with 5% skim milk, followed by an anti-human IgG HRP conjugate (Santa Cruz). For enzyme-linked immunosorbent assay (ELISA), Maxisorb immunoplates (Nunc) were coated with the relevant T6 antigen (5 μg/ml) in PBS (pH 7.2) and blocked with 5% skim milk in PBS-T (PBS with 0.1% Tween 20). Blocked wells were incubated with patient serum (1:200 in PBST with 5% skim milk) at room temperature for 1 h. The plates were washed with PBS-T and incubated with an anti-human antibody–horseradish peroxidase conjugate for detection.

RESULTS

Structure determination. The T6 backbone pilin from *Streptococcus pyogenes* strain MGAS10394 (serotype M6) is expressed as a 549-residue preprotein that includes an N-terminal signal peptide and a classical C-terminal cell wall anchoring sortase recognition motif (LPSTG). For structural analysis, a construct encompassing the mature protein (residues 35 to 520) was expressed in *E. coli*, purified, and crystallized. This construct lacks the N-terminal signal peptide and ends at the glycine residue of the LPSTG sortase recognition motif. T6 was crystallized in space group $P6_3$, and the structure was solved by single wavelength anomalous dispersion methods from KI-soaked crystals. An initial model was built which was further refined against native data to 1.90-Å resolution ($R = 21.4\%$, $R_{free} = 25.8$, see Table 1 for full details). There is one molecule per asymmetric unit with only one external loop (residues 158 to 163) and the C-terminal residues (residues 515 to 520) devoid of electron density, and three short mobile sections with uninterpretable electron density (residues 64 to 65, 335 to 336, and 435 to 436). The first four N-terminal residues (GSLS) are vector derived, with only the N-terminal glycine not modeled.

Structure of T6. T6 is folded into three tandem Ig-like domains arranged linearly to form an elongated structure, 115 Å in length. It has a conserved modular structure typical of Gram-positive backbone pilins (BPs), with an N-terminal CnaB domain (N domain, residues 35 to 178), a middle CnaA domain (M domain, residues 179 to 349), and a C-terminal CnaB domain (C domain, residues 350 to 520) (Fig. 1). This three-domain arrangement closely resembles the BPs from *C. diphtheriae* (SpaA), *A. oris* (FimP), and *S. agalactiae* (GBS80) (16, 17, 19). Each T6 domain, however, has insertions into the basic fold. These variant loops and extensions differ from those in three domain BPs from other organisms.

The most distinctive feature of the CnaB-type N domain is the presence of a 40-residue insertion (variable region 2 [V2]) that precedes β-strand C (Fig. 1). V2 forms a β-strand followed by an α-helix that packs against the top of the M domain, such that the β-sandwich of the N domain is offset from but still parallel to the principal axis formed by the M and C domains. A cleft is formed between the α-helix and β-strands F and G from the main body of the domain. A second point of variation involves the 21-residue loop (V3) between β-strands F and G. This loop sits at the top of, and partially occluding, the cleft and is partially disordered with uninterpretable electron density between residues 158 and 163. At the bottom of the cleft, the side chain of Lys175 protrudes into the solvent. Lys175 is the lysine residue in the novel VAKS pilin motif, which is conserved in *S. pyogenes* T6 strains (31). Lys175 is essential for sortase-mediated pilin polymerization, forming an inter-

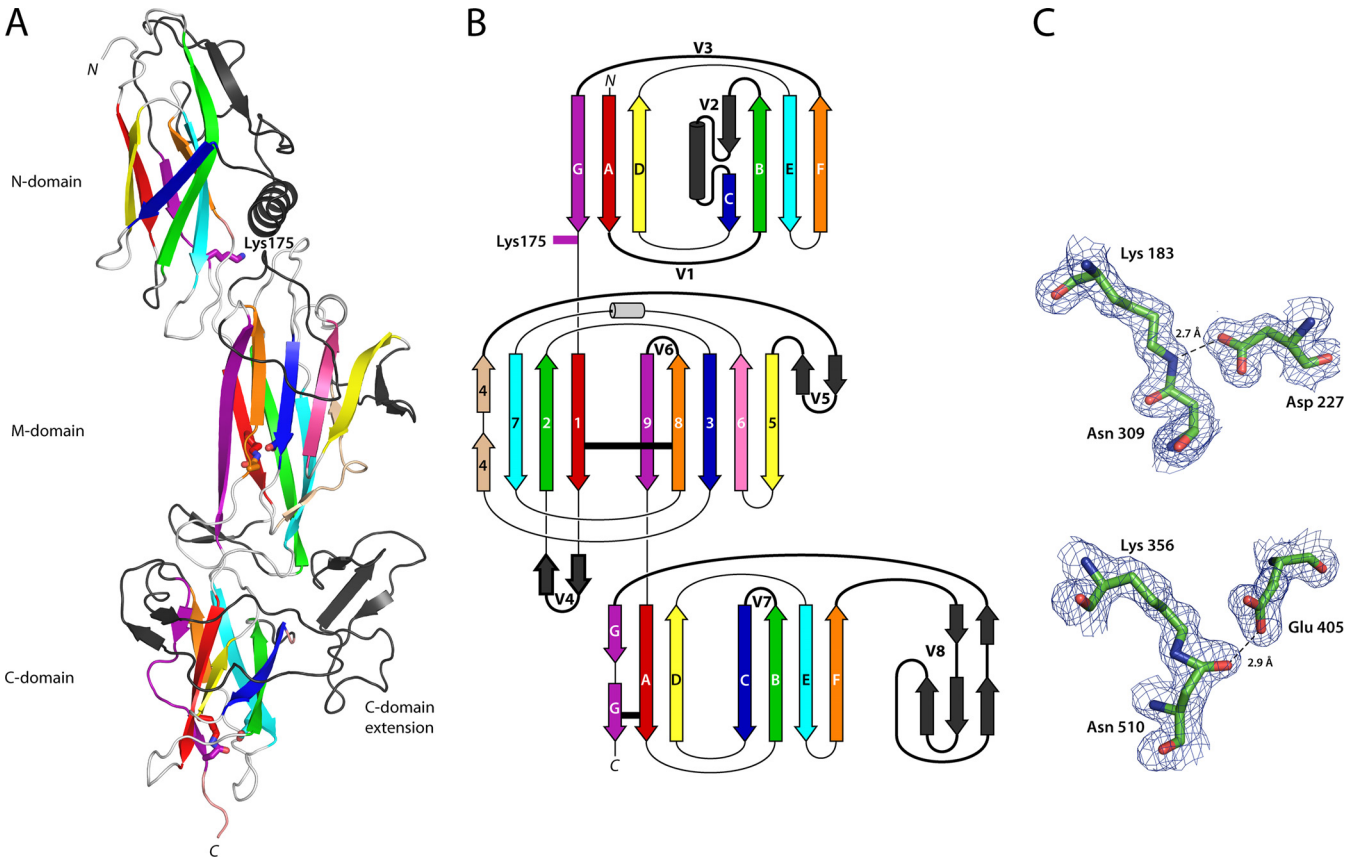

FIG 1 Structure of the T6 backbone pilin. (A) Ribbon diagram depicting the three domains of T6 colored as in the topology diagram (B) from red (N terminus) to purple (C terminus). Each domain is decorated with extended loops or variable regions (colored black, V1 to V8), the largest in the C domain is termed the C-domain extension. Residues involved in isopeptide bond formation and the N-domain pilin lysine (Lys175) are shown in stick form, colored the same as the β-strand from which they originate. (B) Topology diagram color-coded the same as the ribbon diagram. CnaB domain β-strands are labeled A to G, while the CnaA domain β-strands are labeled 1 to 9. The variable loop regions connecting β-strands are shown as black, labeled V1 to V8. Isopeptide bond position is depicted as horizontal black lines. (C) Residues involved in isopeptide bond formation in the M domain (top) and C domain (bottom) are shown in stick form, colored by atom type, in electron density from a 2Fo-Fc map contoured at 0.38 eÅ^{-3} (1.2σ). Hydrogen bonds are shown as broken lines.

molecular covalent linkage with the threonine from the LPSTG sorting motif of the preceding pilin (31). The crystal structure shows clearly that the VAKS motif is structurally homologous with the YPKN pilin motif first described in *C. diphtheriae* (30), a finding consistent with a common polymerization mechanism. Like Lys190 in SpaA (16), T6 Lys175 sits close to the C-terminal end of β-strand G, the last strand of the N domain before it passes into the M domain. The substitution of a serine for the asparagine of the YPKN motif is indicative of the fact that the T6 N-terminal domain does not contain an intramolecular isopeptide bond; it is the Asn of the YPKN motif that participates in the formation of such bonds in other pilin domains.

As with the majority of three-domain pilins, the M domain of T6 has a CnaA-type fold, with a core fold comprising nine β-strands arranged into two antiparallel β-sheets that form a β-sandwich. This is decorated with two β-hairpins and an α-helix. Beta-hairpin 1 (V4; between β-strands 1 and 2) packs against the C domain, and beta-hairpin 2 (V5; between β-strands 4 and 5) packs against β-strand 5 to form an extended β-sheet (Fig. 1). A prominent loop (V6; between β-strands 8 and 9) folds down over the surface to mask the face of this extended 7-stranded β-sheet, while the small single-turn helix sits at the top of β-strands 6 and 7, positioned over the top of the β-sandwich. A conserved helix is found in a similar location in all pilin M domains and in this position it "caps" the β-sandwich (Fig. 2).

The C domain, like the N domain, is a CnaB-type domain. The distinctive feature of the C domain is a large insertion (V8) of 72 residues between β-strands F and G. This major insertion has an unusual topology in which it "wraps around" the C domain from β-strand F and forms a small three-stranded antiparallel β-sheet that sits at the interface between the M and C domains on the opposite side of the molecule. The structure then runs back around the same face of the C domain in an antiparallel direction ending at β-strand G, which forms the last strand of the C domain leading to the C-terminal LPSTG cell wall-anchoring motif (Fig. 1).

Stabilizing isopeptide bonds. Intermolecular isopeptide bonds are a common stabilizing feature of Gram-positive pilins. First described in the *S. pyogenes* BP Spy0128 (6), these intramolecular cross-links have subsequently been found in most Gram-positive pilin domains and in some of the repeated "stalk" domains of other Gram-positive multidomain cell surface adhesins. In T6 the M and C domains have an isopeptide bond, as evident from the clearly defined continuous electron density linking the

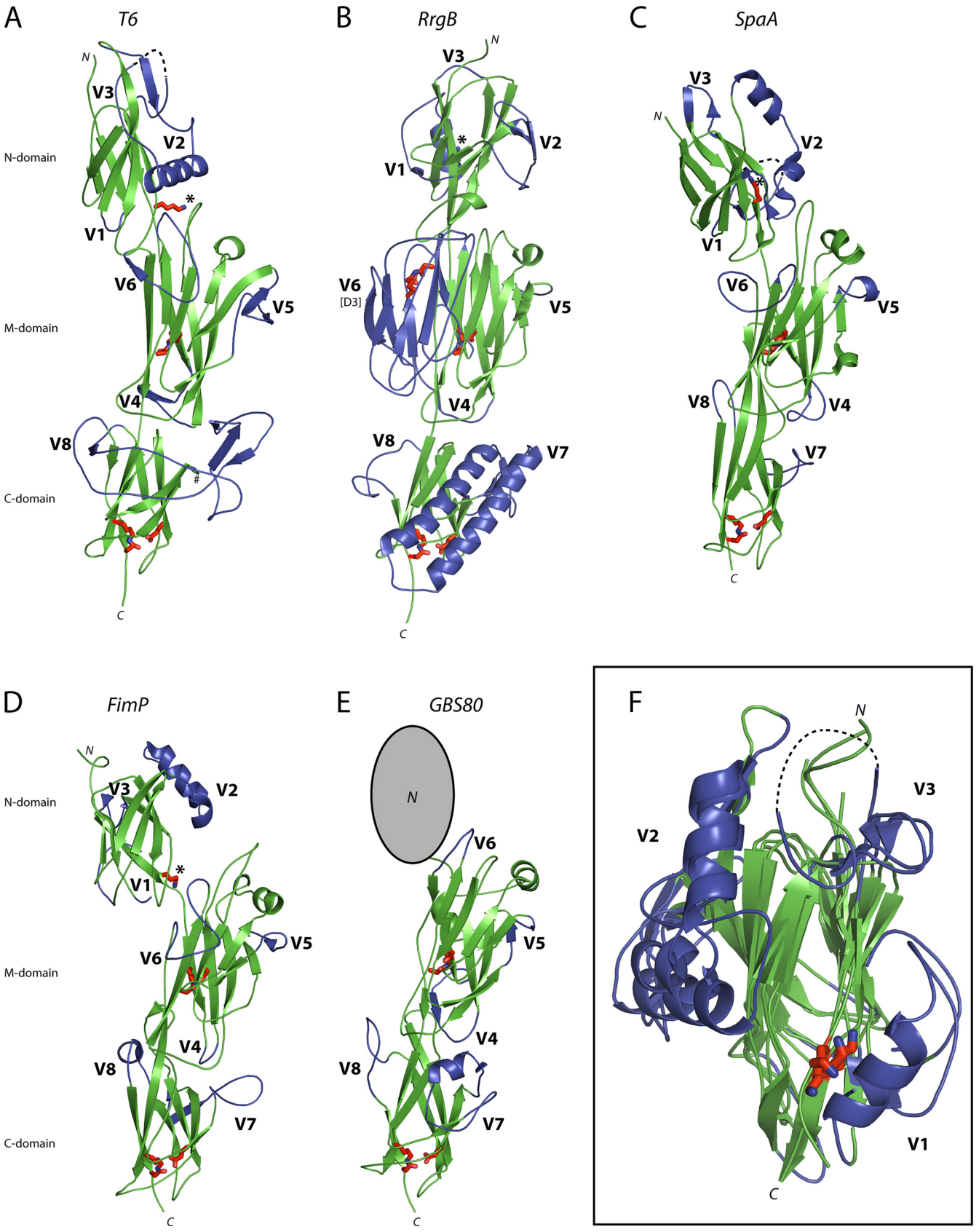

FIG 2 Comparison of Gram-positive backbone pilins. Ribbon diagrams depicting the backbone pilins from T6 (*S. pyogenes*) (A), RrgB (*S. pneumoniae*) (B), SpaA (*C. diphtheriae*) (C), FimP (*A. oris*) (D), and GBS80 (*S. agalactiae*) (E). Each pilin contains three common Ig-like domains, an N-terminal CnaB domain (missing in GBS80), a CnaA M domain, and a CnaB C domain. This three-domain core (Green) is decorated at conserved positions in the Ig-like folds with extensions or variable regions (V1 to V8, blue). Residues involved in isopeptide bond formation and the N-domain pilin lysine (*) are shown in stick form, colored red. (F) A structural alignment of all pilin N domains highlights the conserved positioning of the pilin lysine (red) at the base of a cleft, which is flanked by variable regions V1, V2, and V3 (blue). N, N terminus; C, C terminus.

side chains of Lys183 and Asn309 in the M domain and the side chains of Lys356 and Asn510 in the C domain (Fig. 1C). As in other pilins, the T6 isopeptide bonds are buried in the hydrophobic core of the β-sandwich, stacked against aromatic residues. In T6 both of the isopeptide bonds have a *trans* configuration: in the M domain the isopeptide NH moiety hydrogen bonds with Asp227, and in the C domain the isopeptide O hydrogen bonds with Glu405. These configurations are identical to those observed in BPs from other streptococci (*S. pneumoniae* [RrgB] and *S. agalactiae* [BP-2a]) (20, 21). No isopeptide bond is present in the N domain since this domain lacks the residues necessary for bond formation; the typical triad of Lys, Asn, and Glu is substituted by Ile46, Ser176, and Leu131, respectively. A similar phenomenon is observed in the N domain of *C. diphtheriae* SpaA, which also has no capacity to form an isopeptide bond (16).

Conservation and variation in the Gram-positive backbone pilins. The structures currently available for backbone pilins from Gram-positive bacteria show that they have a recurring modular structure composed of two, three, or four Ig-like domains. More detailed comparisons show that the N-terminal domains of the three- and four-domain pilins are flexibly joined to the subsequent domains; they can adopt different orientations (Fig. 2) and can be removed relatively easily by proteolysis. In several cases, BPs could be crystallized only after proteolytic removal of their N-terminal domains (25). In contrast, there are substantial contacts between the M and C domains, which appear to be arranged as a single rigid entity. When the T6 structure is superimposed on to the structures of other BPs, its M and C domains, as a single unit, fit well to the M and C domains of both three-domain and four-domain BPs, despite a very limited sequence identity of ~20% (see Fig. S1 in the supplemental material). The root mean square differences (RMSDs) between the M and C domains of T6 and those of SpaA and GBS80 are 2.36 Å (over 201 Cα atoms) and 2.33 Å (over 207 Cα atoms), respectively, and RMSDs with RrgB and BP-2a are 2.19 Å (over 215 Cα atoms) and 2.12 Å (over 218 Cα atoms), respectively. Thus, although overall T6 appears more like the three-domain pilins of *S. agalactiae* and *C. diphtheriae*, it is just as closely related in its core structure to the four-domain pilins.

The structural alignments highlight a common core harboring decorations in the form of extended or variable regions (V1 to V8) that are positioned at conserved points in the CnaA and CnaB folds. In the T6 M domain the most substantial decoration (V6) is between β-strands 8 and 9, where a loop folds across the face of β-sheet 4. A similar loop is also observed in SpaA and FimP, and it is at this point that the additional domain (D3) of the four-domain pilins of BP-2a and RrgB is inserted (Fig. 2B). The major C-domain extension in T6 (V8) is inserted between β-strands F and G, where the pilins GBS80, RrgB, BP-2a, FimP, and SpaA all have extended loops, albeit on a much smaller scale. BP-2a, RrgB, and to a lesser extent GBS80, SpaA and FimP, also have a substantial insertion (V7) positioned between β-strands B and C. T6 lacks an insertion at this position, but the majority of the V8 extension in T6 is positioned on the same side of the protein as the V7 insertion of RrgB and BP-2a (Fig. 2).

Structural alignment of all N-terminal BP domains also reveals conserved points of decoration on the core CnaB fold (Fig. 2F). The major insertion on the T6 N domain (V2) is between β-strands B and C where an α-helix packs against the M domain. All pilin N domains have major decorations at this point in the CnaB fold, and in T6, SpaA, and FimP these are α-helices. The loop between β-strands A and B (V1) is another point at which the pilins are often heavily decorated. Whereas T6 only has a short loop, SpaA and FimP have disordered extensions and RrgB has an α-helix that sits adjacent to the pilin lysine. The third major point of variation (V3) is in the loop between β-strands F and G, adjacent to a surface cleft. In RrgB this cleft has been shown to accommodate the LPXTG motif of the preceding pilin in the pilus shaft (43). At the base of this cleft is the pilin lysine that forms the sortase-catalyzed intermolecular isopeptide bond with the LPXTG threonine residue of the next pilin (Fig. 2B). The position of this pilin lysine is highly conserved; in structural alignments the T6 Lys175 Cα atom is 0.5, 0.9, and 1.9 Å from the pilin lysine Cα atom in RrgB, SpaA, and FimP, respectively (Fig. 2F).

Expression and immunogenicity of T6 during infection. The *in vivo* expression of GAS pili has previously been confirmed indirectly by demonstrating that human sera from patients with GAS pharyngitis recognize recombinant BP proteins (44). When BPs from four different GAS strains (M1, M3, M6, and M12) were screened as part of a GAS protein array, 76 of 100 (76%) of the sera reacted with at least one BP. Similarly, in our study sera from patients diagnosed with ARF were screened for the presence of BP-specific antibodies. Western blots with sera against BPs from the six major FCT types (FCT 1, M6; FCT 2, M1; FCT 3, M49; FCT 4, M28; FCT 5, M4; and FCT 6, M2) identified by Falugi et al. (13) showed that IgG in five of the six sera (83%) reacted with T6. Serum from two patients reacted exclusively with T6, but in the three other patients multiple BPs were detected. Example blots for two patients are shown in Fig. 3A. Reactivity to more than one BP was also observed by Manetti et al. (44) and suggests either that cross-reactivity exists among homologous BP domains, that patients have experienced repeated infections with GAS strains carrying different BPs, or both.

The specificity of three sera with high T6 reactivity was further examined using a panel of recombinant T6 truncation constructs. It has previously been shown that the separate domains of the four-domain group B streptococcus pilin BP-2a elicit variable levels of protection in a murine challenge model (21). The D3 domain stimulated the highest titers of opsonophagocytic antibodies in mice, and flow cytometry with serum from immunized mice showed the D3 and D4 domains were accessible to antibodies, while antibody reactivity to the D1 and D2 domains was weak. This raised the question of whether certain domains of GAS BPs are also preferentially targeted by antibodies *in vivo*. Five different T6 truncation mutants, with various combinations of the N, M, and C domains, and the C-domain V8 extension, were expressed and purified to test in serum ELISAs alongside the full-length protein. The highest reactivity for all three serum samples was observed with the full-length T6 protein (N-M-C-ext), as expected (Fig. 3B). Removing the C-domain extension (V8) from both the full-length protein (N-M-C) and the M/C domain (M-C) reduced IgG reactivity for all three sera to various degrees, demonstrating that the V8 insertion is immunogenic. Detection of each of the domain constructs (M, M-C, and N-M) by the three patient sera shows all of the individual T6 domains are targeted by antibodies during infection. This differs from the observations for group B BP-2a in mice, where only two of the four domains showed strong reactivity (21), and suggests that human antibodies bind along the length of the T6 pilus shaft.

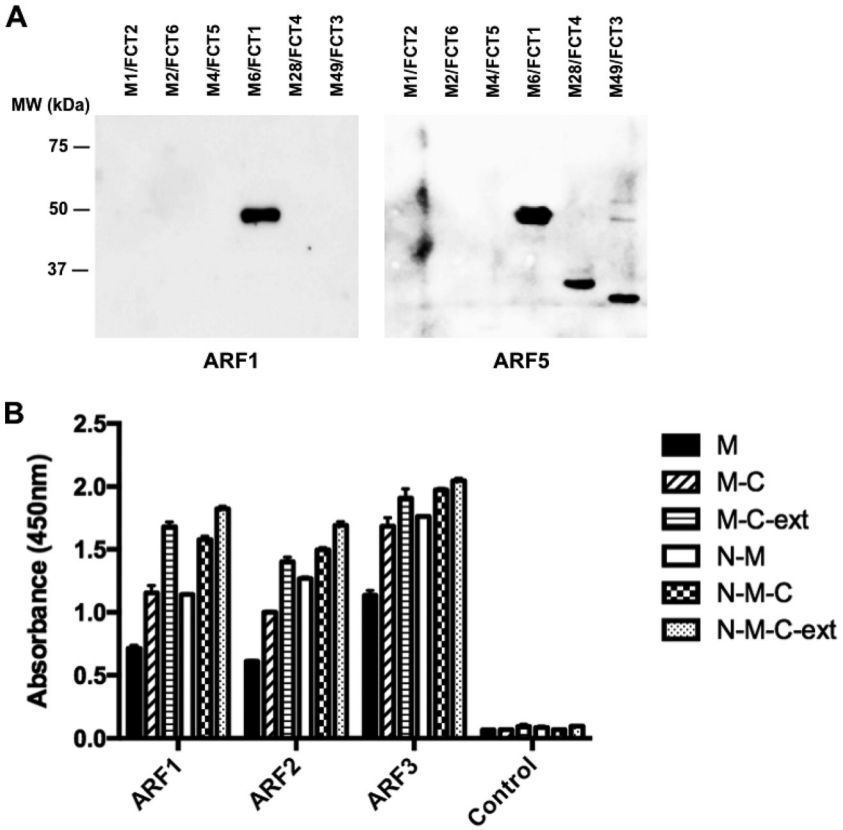

FIG 3 Immunogenicity of the T6 backbone pilin in a defined set of patient sera. (A) Western immunoblots with recombinant purified backbone pilin proteins. Blots were incubated with sera from patients recently diagnosed with acute rheumatic fever. Blots are shown for two patients (ARF1 and ARF5). (B) T6 domain specificity in patient sera as measured by ELISA. Plates were coated with recombinant T6-derived domains and incubated with sera from three individuals (ARF1, ARF2, and ARF3) that were highly reactive with full-length T6. Serum from an individual without anti-T6 antibodies is included as a control.

DISCUSSION

Unlike Gram-negative bacteria, in which pili have long been characterized, the pili of Gram-positive bacteria have been a relatively recent discovery (5, 45). The first structure of a Gram-positive pilin protein, Spy0128, the BP protein from M1 GAS, revealed a two-domain architecture and distinctive intramolecular isopeptide bonds that provide strength and stability (6). Spy0128 is one of the smallest BPs (32.5 kDa), and with its two Ig-like domains differs from subsequently characterized Gram-positive BP structures from *S. agalactiae*, *C. diphtheriae*, *A. oris*, *B. cereus*, and *S. pneumoniae*, which all contain three or four Ig-like domains.

The structure of T6 BP, from the important FCT 1 type M6 strain, is the second structure of a GAS BP. It reveals a domain organization distinctly different from the M1 BP Spy0128, but similar to the three-domain and four-domain BPs from other organisms, with a CnaA domain sandwiched between two CnaB domains. This finding further highlights the recurring, modular structure of Gram-positive pili. The FCT 1 T6 pilus, like the two-domain GAS BP encoded in FCT types 2, 3, and 4, has been shown to lack the canonical YPKN pilin motif that contains the essential pilin lysine involved in BP polymerization (31). In Spy0128, and presumably all other two-domain pilins, the essential pilin lysine is presented on an omega loop on the last β-strand of the N-terminal CnaB domain (6). The T6 structure, however, shows that the novel VAKS motif noted previously (31) is actually just a variant of the YPKN motif of most other pilins. Its position, near the base of the N-terminal CnaB domain with the pilin lysine (Lys 175) located at the end of a groove, exposed to the solvent, matches exactly to that of the YPKN motif (Fig. 3). It differs in that the lysine is not followed by an asparagine, but we now see that this results from the fact that the T6 N domain does not contain an isopeptide bond and the Asn is thus not required. Sequence alignments of FCT 1-type BPs predict an extended motif of YL(Y/F)G(Q/E)X(A/S)V**A**K**S**, where Lys175 (indicated in boldface) is preceded by residues that line the groove (underlined). This extended VAKS motif has some similarities to the canonical WXXXVXV YP**K**N pilin motif, but with notable differences in the residues that line the cleft. In FCT 1-type pilins these residues appear more to be polar than the corresponding residues in other BPs and may reflect differences in the binding mode of the polymerizing sortase.

Although the BPs that form the shaft of Gram-positive pili share a modular construction based on tandem Ig-like domains with similar core folds, they all harbor highly diverged decorations. Analysis of these decorations in BP structures shows a conserved positioning of many of the largest insertions. A common feature of the three Ig-like domains in T6 is that they all have variant loops situated between the last two β-strands, the largest being V8 in the C domain. This extension wraps around one face of the C domain and positions a small β-sheet on the opposite side of the CnaB domain, effectively masking most of the core CnaB fold. As shown in the surface representation in Fig. 4, the series of variant loops in T6 (blue) cover large parts of the structurally conserved Ig-like core (green). The variant

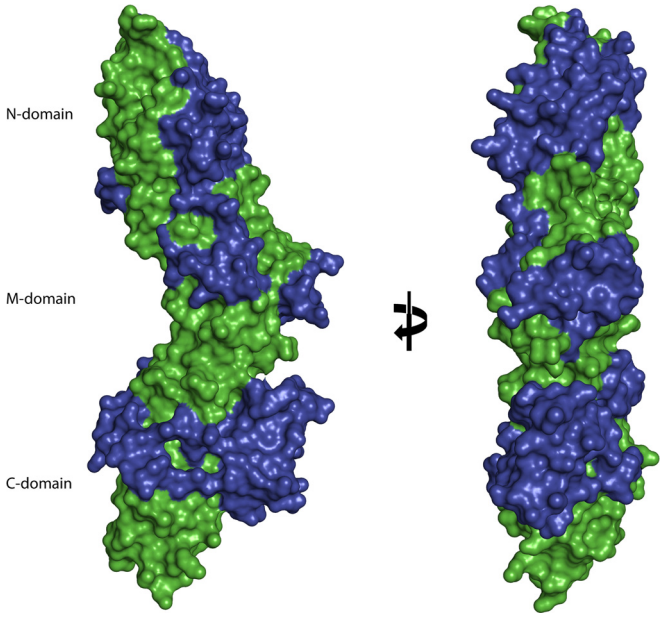

FIG 4 Surface representation of the T6 backbone pilin. A surface representation of the T6 structure highlights the extent that the variable regions or extensions (blue) mask the core Ig-like domains (green). Views are at 0° and 90° rotation around the *y* axis.

loops in GAS pilins likely contribute to the specificity of T sera and the capacity of the sera to recognize specific BP proteins for identification of T-type (13). Similarly, the generation of strain-specific antibody responses to BP in humans is probably driven by these variant loops. IgGs specific for highly variable loops are more likely to be strain specific, whereas IgGs that bind the conserved core of the BP may cross-react with structurally similar epitopes on BPs from different strains. In the present study, the V8 extension was shown to be immunogenic in individuals with GAS-associated ARF. The distinct IgG epitopes of BPs are yet to be mapped in detail, but the immunogenicity of flexible loops and highly variable regions are well documented for many other surface-exposed bacterial proteins (46–48).

Murine challenge studies with GAS pilus components have shown that the BP proteins are effective protective antigens and promising candidates for vaccine development (5). However, developing a vaccine against GAS pili is complicated by high sequence variability, with previous estimates suggesting that a vaccine containing a combination of at least 12 BP variants would be required for 90% protection against common GAS strains (13). Although GAS BPs show large variations in size and sequence, the modular nature of the pilus proteins revealed by the current T6 structure and that of M1 Spy0128 (6) may aid future vaccine design. Such structure-based antigen design approaches have proved successful for serogroup B meningococcus and group B streptococcus (49, 50). Careful examination of GAS BP structures from common serotypes, alongside studies of the human immune response to BP domains, may reduce the number of BP variants needed to provide protection against widely circulating GAS strains.

In conclusion, the structure of the BP protein T6 highlights a structurally conserved core for pilus proteins. The three conserved Ig-like domains are decorated with variant loops that likely contribute to strain-specific immune responses in the human host. There are currently no vaccines available for GAS, and the design of a pilus-based vaccine would be facilitated by the availability of further BP structures.

ACKNOWLEDGMENTS

This study was funded by the Maurice Wilkins Centre and the Health Research Council of New Zealand. N.J.M. is a New Zealand Heart Foundation Research Fellow.

REFERENCES

1. **Cunningham MW.** 2008. Pathogenesis of group A streptococcal infections and their sequelae. Adv. Exp. Med. Biol. **609:**29–42. http://dx.doi.org/10.1007/978-0-387-73960-1_3.
2. **Carapetis JR, McDonald M, Wilson NJ.** 2005. Acute rheumatic fever. Lancet **366:**155–168. http://dx.doi.org/10.1016/S0140-6736(05)66874-2.
3. **Jaine R, Baker M, Venugopal K.** 2008. Epidemiology of acute rheumatic fever in New Zealand 1996-2005. J. Paediatr. Child Health **44:**564–571. http://dx.doi.org/10.1111/j.1440-1754.2008.01384.x.
4. **Maguire GP, Carapetis JR, Walsh WF, Brown ADH.** 2012. The future of acute rheumatic fever and rheumatic heart disease in Australia. Med. J. Australia **197:**133–134. http://dx.doi.org/10.5694/mja12.10980.
5. **Mora M, Bensi G, Capo S, Falugi F, Zingaretti C, Manetti AG, Maggi T, Taddei AR, Grandi G, Telford JL.** 2005. Group A *Streptococcus* produce pilus-like structures containing protective antigens and Lancefield T antigens. Proc. Natl. Acad. Sci. U. S. A. **102:**15641–15646. http://dx.doi.org/10.1073/pnas.0507808102.
6. **Kang HJ, Coulibaly F, Clow F, Proft T, Baker EN.** 2007. Stabilizing isopeptide bonds revealed in gram-positive bacterial pilus structure. Science **318:**1625–1628. http://dx.doi.org/10.1126/science.1145806.
7. **Quigley BR, Zahner D, Hatkoff M, Thanassi DG, Scott JR.** 2009. Linkage of T3 and Cpa pilins in the *Streptococcus pyogenes* M3 pilus. Mol. Microbiol. **72:**1379–1394. http://dx.doi.org/10.1111/j.1365-2958.2009.06727.x.
8. **Smith WD, Pointon JA, Abbot E, Kang HJ, Baker EN, Hirst BH, Wilson JA, Banfield MJ, Kehoe MA.** 2010. Roles of minor pilin subunits Spy0125 and Spy0130 in the serotype M1 *Streptococcus pyogenes* strain SF370. J. Bacteriol. **192:**4651–4659. http://dx.doi.org/10.1128/JB.00071-10.
9. **Linke C, Young PG, Kang HJ, Bunker RD, Middleditch MJ, Caradoc-Davies TT, Proft T, Baker EN.** 2010. Crystal structure of the minor pilin FctB reveals determinants of group A streptococcal pilus anchoring. J. Biol. Chem. **285:**20381–20389. http://dx.doi.org/10.1074/jbc.M109.089680.
10. **Hendrickx AP, Budzik JM, Oh SY, Schneewind O.** 2011. Architects at the bacterial surface: sortases and the assembly of pili with isopeptide bonds. Nat. Rev. Microbiol. **9:**166–176. http://dx.doi.org/10.1038/nrmicro2520.
11. **Koller T, Manetti AGO, Kreikemeyer B, Lembke C, Margarit I, Grandi G, Podbielski A.** 2010. Typing of the pilus-protein-encoding FCT region and biofilm formation as novel parameters in epidemiological investigations of *Streptococcus pyogenes* isolates from various infection sites. J. Med. Microbiol. **59:**442–452. http://dx.doi.org/10.1099/jmm.0.013581-0.
12. **Luca-Harari B, Darenberg J, Neal S, Siljander T, Strakova L, Tanna A, Creti R, Ekelund K, Koliou M, Tassios PT, van der Linden M, Straut M, Vuopio-Varkila J, Bouvet A, Efstratiou A, Schalen C, Henriques-Normark B, Jasir A, Grp S-ES.** 2009. Clinical and microbiological characteristics of severe *Streptococcus pyogenes* disease in Europe. J. Clin. Microbiol. **47:**1155–1165. http://dx.doi.org/10.1128/JCM.02155-08.
13. **Falugi F, Zingaretti C, Pinto V, Mariani M, Amodeo L, Manetti AGO, Capo S, Musser JM, Orefici G, Margarit I, Telford JL, Grandi G, Mora M.** 2008. Sequence variation in group A streptococcus pili and association of pilus backbone types with Lancefield T serotypes. J. Infect. Dis. **198:**1834–1841. http://dx.doi.org/10.1086/593176.
14. **Kratovac Z, Manoharan A, Luo F, Lizano S, Bessen DE.** 2007. Population genetics and linkage analysis of loci within the FCT region of *Streptococcus pyogenes*. J. Bacteriol. **189:**1299–1310. http://dx.doi.org/10.1128/JB.01301-06.
15. **Deivanayagam CCS, Rich RL, Carson M, Owens RT, Danthuluri S, Bice T, Hook M, Narayana SVL.** 2000. Novel fold and assembly of the repetitive B region of the *Staphylococcus aureus* collagen-binding surface protein. Struct. Fold Des. **8:**67–78. http://dx.doi.org/10.1016/S0969-2126(00)00081-2.
16. **Kang HJ, Paterson NG, Gaspar AH, Ton-That H, Baker EN.** 2009. The *Corynebacterium diphtheriae* shaft pilin SpaA is built of tandem Ig-

like modules with stabilizing isopeptide and disulfide bonds. Proc. Natl. Acad. Sci. U. S. A. **106:**16967–16971. http://dx.doi.org/10.1073/pnas.0906826106.
17. **Vengadesan K, Ma X, Dwivedi P, Hung TT, Narayana SVL.** 2011. A model for group B streptococcus pilus type 1: the structure of a 35-kDa C-terminal fragment of the major pilin GBS80. J. Mol. Biol. **407:**731–743. http://dx.doi.org/10.1016/j.jmb.2011.02.024.
18. **Mishra A, Devarajan B, Reardon ME, Dwivedi P, Krishnan V, Cisar JO, Das A, Narayana SVL, Ton-That H.** 2011. Two autonomous structural modules in the fimbrial shaft adhesin FimA mediate *Actinomyces* interactions with streptococci and host cells during oral biofilm development. Mol. Microbiol. **81:**1205–1220. http://dx.doi.org/10.1111/j.1365-2958.2011.07745.x.
19. **Persson K, Esberg A, Claesson R, Stromberg N.** 2012. The pilin protein FimP from *Actinomyces oris*: crystal structure and sequence analyses. PLoS One **7:**1–10. http://dx.doi.org/10.1371/journal.pone.0048364.
20. **Paterson NG, Baker EN.** 2011. Structure of the full-length major pilin from *Streptococcus pneumoniae*: implications for isopeptide bond formation in gram-positive bacterial pili. PLoS One **6:**1–8. http://dx.doi.org/10.1371/journal.pone.0022095.
21. **Nuccitelli A, Cozzi R, Gourlay LJ, Donnarumma D, Necchi F, Norais N, Telford JL, Rappuoli R, Bolognesi M, Maione D, Grandi G, Rinaudo CD.** 2011. Structure-based approach to rationally design a chimeric protein for an effective vaccine against group B streptococcus infections. Proc. Natl. Acad. Sci. U. S. A. **108:**10278–10283. http://dx.doi.org/10.1073/pnas.1106590108.
22. **Budzik JM, Poor CB, Faull KF, Whitelegge JP, He C, Schneewind O.** 2010. Intramolecular amide bonds stabilize pili on the surface of bacilli. Proc. Natl. Acad. Sci. U. S. A. **107:**5260–5260. http://dx.doi.org/10.1073/pnas.1000441107.
23. **Hendrickx APA, Poor CB, Jureller JE, Budzik JM, He C, Schneewind O.** 2012. Isopeptide bonds of the major pilin protein BcpA influence pilus structure and bundle formation on the surface of *Bacillus cereus*. Mol. Microbiol. **85:**152–163. http://dx.doi.org/10.1111/j.1365-2958.2012.08098.x.
24. **Symersky J, Patti JM, Carson M, HousePompeo K, Teale M, Moore D, Jin L, Schneider A, DeLucas LJ, Hook M, Narayana SVL.** 1997. Structure of the collagen-binding domain from a *Staphylococcus aureus* adhesin. Nat. Struct. Biol. **4:**833–838. http://dx.doi.org/10.1038/nsb1097-833.
25. **Kang HJ, Baker EN.** 2012. Structure and assembly of Gram-positive bacterial pili: unique covalent polymers. Curr. Opin. Struc Biol. **22:**200–207. http://dx.doi.org/10.1016/j.sbi.2012.01.009.
26. **Kreikemeyer B, Gamez G, Margarit I, Giard JC, Hammerschmidt S, Hartke A, Podbielski A.** 2011. Genomic organization, structure, regulation and pathogenic role of pilus constituents in major pathogenic streptococci and enterococci. Int. J. Med. Microbiol. **301:**240–251. http://dx.doi.org/10.1016/j.ijmm.2010.09.003.
27. **Nakata M, Koller T, Moritz K, Ribardo D, Jonas L, McIver KS, Sumitomo T, Terao Y, Kawabata S, Podbielski A, Kreikemeyer B.** 2009. Mode of expression and functional characterization of FCT-3 pilus region-encoded proteins in *Streptococcus pyogenes* serotype M49. Infect. Immun. **77:**32–44. http://dx.doi.org/10.1128/IAI.00772-08.
28. **Zahner D, Scott JR.** 2008. SipA is required for pilus formation in *Streptococcus pyogenes* serotype M3. J. Bacteriol. **190:**527–535. http://dx.doi.org/10.1128/JB.01520-07.
29. **Young PG, Kang HJ, Baker EN.** 2013. An arm-swapped dimer of the *Streptococcus pyogenes* pilin specific assembly factor SipA. J. Struct. Biol. **183:**99–104. http://dx.doi.org/10.1016/j.jsb.2013.05.021.
30. **Ton-That H, Schneewind O.** 2003. Assembly of pili on the surface of *Corynebacterium diphtheriae*. Mol. Microbiol. **50:**1429–1438. http://dx.doi.org/10.1046/j.1365-2958.2003.03782.x.
31. **Nakata M, Kimura KR, Sumitomo T, Wada S, Sugauchi A, Oiki E, Higashino M, Kreikemeyer B, Podbielski A, Okahashi N, Hamada S, Isoda R, Terao Y, Kawabata S.** 2011. Assembly mechanism of FCT region type 1 pili in serotype M6 *Streptococcus pyogenes*. J. Biol. Chem. **286:**37566–37577. http://dx.doi.org/10.1074/jbc.M111.239780.
32. **Moreland N, Ashton R, Baker HM, Ivanovic I, Patterson S, Arcus VL, Baker EN, Lott JS.** 2005. A flexible and economical medium-throughput strategy for protein production and crystallization. Acta Crystallogr. D **61:**1378–1385. http://dx.doi.org/10.1107/S0907444905023590.
33. **Kabsch W.** 1993. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Crystallogr. **26:**795–800. http://dx.doi.org/10.1107/S0021889893005588.
34. **Evans P.** 2006. Scaling and assessment of data quality. Acta Crystallogr. D **62:**72–82. http://dx.doi.org/10.1107/S0907444905036693.
35. **Perrakis A, Morris R, Lamzin VS.** 1999. Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. **6:**458–463. http://dx.doi.org/10.1038/8263.
36. **Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC.** 2002. PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D **58:**1948–1954. http://dx.doi.org/10.1107/S0907444902016657.
37. **Mccoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ.** 2007. Phaser crystallographic software. J. Appl. Crystallogr. **40:**658–674. http://dx.doi.org/10.1107/S0021889807021206.
38. **Emsley P, Lohkamp B, Scott WG, Cowtan K.** 2010. Features and development of Coot. Acta Crystallogr. D **66:**486–501. http://dx.doi.org/10.1107/S0907444910007493.
39. **Smart OS, Womack TO, Flensburg C, Keller P, Paciorek W, Sharff A, Vonrhein C, Bricogne G.** 2012. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. Acta Crystallogr. D **68:**368–380. http://dx.doi.org/10.1107/S0907444911056058.
40. **Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA.** 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D **67:**355–367. http://dx.doi.org/10.1107/S0907444911001314.
41. **Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, Snoeyink J, Richardson JS, Richardson DC.** 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. **35:**W375–W383. http://dx.doi.org/10.1093/nar/gkm216.
42. **Ochman H, Gerber AS, Hartl DL.** 1988. Genetic applications of an inverse polymerase chain reaction. Genetics **120:**621–623.
43. **El Mortaji L, Contreras-Martel C, Moschioni M, Ferlenghi I, Manzano C, Vernet T, Dessen A, Di Guilmi AM.** 2012. The full-length *Streptococcus pneumoniae* major pilin RrgB crystallizes in a fibre-like structure, which presents the D1 isopeptide bond and provides details on the mechanism of pilus polymerization. Biochem. J. **441:**833–841. http://dx.doi.org/10.1042/BJ20111397.
44. **Manetti AG, Zingaretti C, Falugi F, Capo S, Bombaci M, Bagnoli F, Gambellini G, Bensi G, Mora M, Edwards AM, Musser JM, Graviss EA, Telford JL, Grandi G, Margarit I.** 2007. *Streptococcus pyogenes* pili promote pharyngeal cell adhesion and biofilm formation. Mol. Microbiol. **64:**968–983. http://dx.doi.org/10.1111/j.1365-2958.2007.05704.x.
45. **Proft T, Baker EN.** 2009. Pili in Gram-negative and Gram-positive bacteria: structure, assembly, and their role in disease. Cell Mol. Life Sci. **66:**613–635. http://dx.doi.org/10.1007/s00018-008-8477-4.
46. **Jones GJ, Gordon SV, Hewinson RG, Vordermeier HM.** 2010. Screening of predicted secreted antigens from *Mycobacterium bovis* reveals the immunodominance of the ESAT-6 protein family. Infect. Immun. **78:**1326–1332. http://dx.doi.org/10.1128/IAI.01246-09.
47. **Kvam AI, Mavenyengwa RT, Radtke A, Maeland JA.** 2011. *Streptococcus agalactiae* alpha-like protein 1 possesses both cross-reacting and Alp1-specific epitopes. Clin. Vaccine Immunol. **18:**1365–1370. http://dx.doi.org/10.1128/CVI.05005-11.
48. **Evans NJ, Harrison OB, Clow K, Derrick JP, Feavers IM, Maiden MCJ.** 2010. Variation and molecular evolution of HmbR, the *Neisseria meningitidis* haemoglobin receptor. Microbiol. SGM **156:**1384–1393. http://dx.doi.org/10.1099/mic.0.036475-0.
49. **Dormitzer PR, Grandi G, Rappuoli R.** 2012. Structural vaccinology starts to deliver. Nat. Rev. Microbiol. **10:**807–813. http://dx.doi.org/10.1038/nrmicro2893.
50. **Nuccitelli A, Rinaudo CD, Brogioni B, Cozzi R, Ferrer-Navarro M, Yero D, Telford JL, Grandi G, Daura X, Zacharias M, Maione D.** 2013. Understanding the molecular determinants driving the immunological specificity of the protective pilus 2a backbone protein of group B *Streptococcus*. PLoS Comput. Biol. **9:**1–11. http://dx.doi.org/10.1371/journal.pcbi.1003115.