Published in final edited form as:

J Mol Biol. 2011 April 15; 407(5): 731–743. doi:10.1016/j.jmb.2011.02.024.

A Model for Group B Streptococcus Pilus Type 1: The Structure of a 35 kDa C-terminal Fragment of the Major Pilin GBS80

Vengadesan Krishnan¹, Ma Xin², Dwivedi Prabhat², Ton-That Hung², and Narayana V.L Sthanam^{1,*}

¹Center for Biophysical Sciences and Engineering, School of Optometry, University of Alabama at Birmingham, Birmingham AL 35294

²University of Texas Health Science Center, Houston, TX 77030

Abstract

The Gram-positive pathogen *Streptococcus agalactiae*, known as group B streptococcus (GBS), is the leading cause of bacterial septicemia, pneumonia, and meningitis among neonates. GBS assemble two types of pili, PI-1 and PI-2, on their surface to adhere to host cells and initiate colonization for pathogenesis. The GBS PI-1 pilus is made of one major pilin, GBS80, which forms the pilus shaft, and two secondary pilins (GBS104 and GBS52), which are incorporated into the pilus at various places. We report here the crystal structure of the 35 kDa C-terminal fragment from GBS80, which is composed of two IgG-like domains (N2-N3). The structure was solved by the single-wavelength anomalous dispersion (SAD) method using sodium iodide (NaI) soaked crystals and diffraction data collected at the home source. The N2 domain exhibits a cnaA/Dev-IgG fold with two calcium binding sites, while the N3 domain displays a cnaB/IgG-rev fold. We have built a model for full-length GBS80 (N1, N2 and N3) with the help of available homologous major pilin structures, and we propose a model for the GBS PI-1 pilus shaft. The N2 and N3 domains are arranged in tandem along the pilus shaft, whereas the respective N1 domain is tilted by approximately 20° away from the pilus axis. We have also identified a pilin-like motif in the minor pilin GBS52, which might aid its incorporation at the pilus base.

Introduction

Bacterial pathogens assemble hair-like, extended surface organelles referred as pili that play pivotal role in colonization and pathogenesis by mediating adhesion to the host tissues. The presence of pili in Gram-positive bacteria is a relatively recent discovery, and knowledge of this assembly mechanism is still evolving¹⁻⁴ compared with the well-investigated Gram-negative pilus assembly (for review⁵⁻⁷). A typical Gram-positive pilus is composed of multiple copies of two or more distinct protein subunits referred to as pilins. According to the well-characterized Gram-positive *Corynebacterium diphtheriae* pilus model, its SpaABC pilus is made of three pilins, SpaA, SpaB and SpaC.¹ The major pilin SpaA is distributed

© 2011 Elsevier Ltd. All rights reserved.

*Corresponding Author: Narayana Sthanam, PhD Center for Biophysical Science and Engineering University of Alabama at Birmingham Birmingham, AL 35294 USA Phone: (+1) 205 969 3474 Fax: (+1) 205 975 0538 narayana@uab.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCESSION NUMBERS

Coordinates and structural factors have been deposited in the Protein Data Bank with accession numbers **3PF2**, **3PG2**

along the pilus shaft and in contrast, the minor pilin SpaB is observed at regular intervals, and SpaC is positioned at the tip of the shaft. Many of the unique features observed in the corynebacterial pilus assembly model are also observed in several other Gram-positive bacteria, such as group A streptococcus (GAS, *Streptococcus pyogenes*), group B streptococcus (GBS, *Streptococcus agalactiae*), *Streptococcus pneumoniae*, *Enterococcus faecalis*, and *Actinomyces*. Two classes of cysteine transpeptidases, pilus-specific sortase and house-keeping sortase, have been shown to assist in pili assembly. The pilus-specific sortase builds the pilus shaft by linking major or backbone pilins via inter-molecular covalent amide bonds in a head-to-tail fashion. The house-keeping sortase anchors the assembled pilus polymer onto the cell wall by linking it to the peptidoglycan cross bridge. However, an understanding of the ancillary/minor pilin incorporation as well as the final pilus assembly is still evolving.⁸

Gram-positive pilins contain a signal peptide at the N-terminus approximately 30 residues long and a sorting signal at the C-terminus, similar to those observed for MSCRAMMs (microbial surface component recognizing adhesive matrix molecules)⁹ (for example, Fig. 1a and 1b represent the arrangement in GBS major and minor pilins), and these motifs are sufficient and essential for cell wall anchoring.¹⁰ The C-terminal sorting signal contains a LPXTG-like motif, followed by a hydrophobic domain and positively charged tail.^{10,11} In addition to the conserved C-terminal motifs, the major pilins of several Gram-positive pathogens also contain a hydrophobic (WxxxVxVYPK) pilin motif, and the Lys residue from this motif was shown to be essential for major pilin polymerization.^{1,12}

Streptococcus agalactiae, known as group B streptococcus (GBS), is one of the major causes of neonatal morbidity and mortality.¹³ Either vertical transmission to the fetus or exposure during passage through the birth canal can lead to GBS infections in neonates, which include septicemia, pneumonia and meningitis. Prophylaxis during child birth has reduced GBS-related neonatal mortality in the past thirty years; however, the pathogen remains a significant threat to neonates. In addition, GBS is also responsible for invasive diseases in non-pregnant, elderly or immune-compromised adults, and their mortality rates have exceeded those seen in neonates.¹⁴⁻¹⁷ As there is a necessity for a universal vaccine against multiple GBS strains,¹⁸ given increased antibiotic resistance and the need for alternative targets in the development of anti-infective drugs,^{19,20} the GBS pilins have become important vaccine candidates.²¹

Genomic analysis of GBS strains (2603V/R, NEM316, and A909) revealed the presence of two similar pilus islands (PIs).²⁻⁴ The Pilus Island 1 (PI-1) from strain 2603V/R encodes three pilins, GBS80 (SAG0645), GBS52 (SAG0646) and GBS104 (SAG0649); two pilus specific sortases, SAG0647 and SAG0648; and a housekeeping sortase, SAG0650. The pilins GBS80, GBS52 and GBS104 are homologs of the *C. diphtheriae* pilins SpaA, SpaB and SpaC, respectively. According to the present model of pilus assembly in GBS,²² either pilus-specific sortase (SAG0647 or SAG0648) is capable of polymerizing the major pilin GBS80. This mechanism occurs via linking the Thr residue in the C-terminal IPNTG sorting motif of one molecule to the side chain of a conserved Lys within the pilin YPKN motif of another GBS80 molecule. While GBS80 forms the pilus shaft, incorporation of the adhesive minor pilins, GBS52 and GBS104, requires specific sortases, SAG647 and SAG0648, respectively. While the minor pilin GBS104 is incorporated at the pilus tip,^{4,8,23} the location of GBS52 is still not clear, as it is not detected by imaging techniques. However, immunoblotting data suggests it is associated with the polymerized pilus shaft,²² and labeled GBS1474 (strain NEM316), an ortholog of GBS52, was observed primarily at the base of the pilus and randomly along the pilus shaft.³

The GBS pili are virulence factors and have multiple functions including host cell adhesion²⁴ and knowledge of their structural assembly is essential for inhibitor design efforts. Toward this goal, we have previously reported the crystal structure of the minor pilin GBS52²³ and crystallization of two GBS sortases that are responsible for pilus assembly and anchoring.²⁵ Available full-length and partial crystal structures of major pilins in Gram-positive bacteria include Spy0128 of *S. pyogenes*,²⁶ SpaA of *C. diphtheriae*,²⁷ BcpA of *B. cereus*,²⁸ and RrgB of *S. pneumoniae*.²⁹ Certain unique structural features distinguish these major pilins from each other, which probably dictate the diversity in their cellular targets and the resulting infections. In this publication, we report the crystal structure of a 35 kDa C-terminal fragment of GBS80 in two different space groups using 1.7 and 1.8 Å resolutions, respectively that were collected using an in-house X-ray source. Halide quick-soaking and single-wavelength anomalous dispersion (SAD) methods were used to generate the initial phases and determine the structure. Homology modeling helped us generate the full length GBS80 structure and a model for GBS pilus type I assembly. We identified a pilin-like motif (Fig. 1b) in the minor pilin GBS52 crystal structure, which aided us in explaining its possible incorporation at the pilus base. We also present a rationale for our inability to crystallize the full-length GBS80, while structures for some full-length major pilins from other Gram-positive bacteria are available.

Results

Overall structure of the pilin GBS80 35 kDa fragment

The recombinant protein GBS80 (residues 36-518; rGBS80₃₆₋₅₁₈) containing an N-terminal His-tag, though lacking the N-terminal signal peptide and C-terminal sorting motif, was expressed in *E. coli* and purified. The rGBS80₃₆₋₅₁₈ was sensitive to proteases and failed to yield crystals. Inclusion of common protease inhibitors during the purification and crystallization did not prevent degradation. Therefore, rGBS80₃₆₋₅₁₈ was subjected to limited proteolysis by α -chymotrypsin to obtain a stable, 35-kDa crystallizable fragment. N-terminal sequencing identified residue Asn200 in the pilin motif (YPKN) as the starting residue of the crystallized fragment. Secondary structure prediction and in-house 3D model building studies suggested that the 35-kDa fragment (GBS80₂₀₁₋₅₁₈), which consists of 318 amino acids, possibly includes two immunoglobulin-like (IgG-like) domains. Inclusion of 10 mM CaCl₂ as an additive during rGBS80₃₆₋₅₁₈ purification also resulted in the same 35-kDa fragment. Interestingly, both the α -chymotrypsin and CaCl₂-treated fragments were crystallized under the same conditions, though in two different space groups (Table 1). The crystals resulting from the α -chymotrypsin treatment, referred to as tGBS80₂₀₀₋₅₁₈ below, belong to the P2₁ space group, with two molecules in the asymmetric unit, which are related by a pseudo-translation along the x-axis. The crystals grown from the CaCl₂-treated GBS80, referred to as CaGBS80₂₀₀₋₅₁₈ crystals, belong to the C2 space group and contain one molecule in the asymmetric unit. However, both crystal forms diffracted to 1.7 and 1.8 Å resolutions, using an in-house X-ray source (Table 1). The structure of tGBS80₂₀₀₋₅₁₈ was solved by the single-wavelength anomalous dispersion (SAD) method using sodium iodide (NaI)-soaked crystals, prepared by the halide quick-soaking technique as reported earlier.^{30,31} Figure 2a illustrates the positions of the identified halide ions in the anomalous difference map and the backbone trace of GBS52N2N3. The tertiary structures of the three GBS80₂₀₀₋₅₁₈ crystal forms (tGBS80₂₀₀₋₅₁₈, CaGBS80₂₀₀₋₅₁₈, and NaI derivative crystals) were similar with a root mean square deviation (rmsd) of less than 1 Å for all aligned C α atoms. Therefore, the CaGBS80₂₀₀₋₅₁₈ structure has been used for the structural analysis presented in the following sections. A well-defined electron density was apparent for residues 203 to 518, except for the Phe299 side chain. The overall B-factor is 18.8 Å², and the R_{factor} and R_{free} are 20.3 and 23.8%, respectively.

CaGBS80₂₀₀₋₅₁₈ is folded into two IgG-like domains with two intra-domain isopeptide bonds and two metal-binding sites, and it has an elongated rod shape with dimensions that are approximately $85 \times 48 \times 38 \text{ \AA}$ (Fig. 2b). We have named the two C-terminal domains N2 and N3 (following the MSCRAMMs and GBS52 conventions) (Fig. 1a and 1b). Flexibility between the two domains is restricted by a short linker and large interface with a buried surface area of $1642 \text{ \AA}^2$, contributed by approximately 53 contacts (distance $< 4 \text{ \AA}$) including a salt bridge. In addition, inter-domain flexibility is further restricted by the presence of a small, three-stranded (I1, I2, and I3) anti-parallel β -sheet.

Structure of the N2 domain

The core of the N2 domain (residues 203-380) exhibits a DEv-IgG fold/cnaA-type, previously observed in the *S. aureus* collagen-binding CNA A-region N2 domain.³² Similar to the ubiquitous IgG-C fold, the DEv-IgG fold of the N2 domain is composed of two β -sheets, where four anti-parallel strands (A, B, E, D) form the β -sheet I. However, the second β -sheet is composed of five strands, the three conserved strands (C, F, G) of the IgG-C fold and additional D' and D'' strands, situated between β strands E and D. These two β -sheets are parallel, with a large hydrophobic interior, and are connected by an inter-sheet isopeptide bond. The loops between D and D' as well as the D'' and E strands are long and extend toward the outer surface of β -sheet I, each hosting two short β -strands and two α -helices, respectively. One of the helices (H2) present in the D''E loop is parallel to strand A and partially covers one of the two metal-binding sites observed in this domain.

Structure of the N3 domain

The N3 domain (residues 381-518) exhibits an IgG-rev fold/cnaB-type, first observed in collagen-binding *S. aureus* CNA B-region repeats³³ and made of two β -sheets similar to those in the IgG-C fold, though the order of the β strands is different (Fig. 2c). Three connecting loops, B-C, D-E, and F-G, are extended toward the N2 domain and stabilize the domain interface. The BC loop is longer than the other two (DE and FG), containing a helix and a short β strand. It covers the N2-N3 domain interface and the interface between sheets I and II of the N3 domain, while its helix is inserted in the N2-N3 domain interface cleft. The BC loop of the N3 domain and the D''E loop of the N2 domain are placed on opposite sides giving a symmetrical shape to the GBS80N2N3 molecule.

The isopeptide bonds

Two stabilizing isopeptide bonds are identified in the structure of GBS80N2N3, one in the N2 and another in the N3 domain. The presence of continuous electron density between the side chains of respective Lys and Asn residues clearly confirms the existence of two isopeptide bonds (Fig. 3a). In the N2 domain, the isopeptide bond connects two β -sheets and is formed between Lys210 (strand A of β -sheet I) and Asn351 (on a loop leading to the N-terminal end of β -sheet II strand F). The conserved acidic residue Asp250 is found on strand C of sheet II within hydrogen bonding distance to the amide bond. The N2 domain isopeptide bond has *cis* configuration, allowing its amide hydrogen (NH) and carbonyl oxygen (O) to form bi-dentate hydrogen bonds with the Asp250 side chain carboxyl group (Fig. 3a). Interestingly, the isopeptide-forming Lys210 and its neighbors, Asp209 and Asp211, participate in metal-ion coordination, which probably restricts the environment of Lys210. Moreover, the isopeptide bond is sandwiched between the aromatic residues Trp230 and Phe252 on one side and Phe353 on the other side as well as surrounded by additional hydrophobic residues, Leu232, Val335, Pro376, and Leu256.

The N3 domain isopeptide bond connects two adjacent, parallel β -strands. The bond is formed between Lys387 and Asn515 in the first (A) and the last (G) strands of the three-stranded (DAG) β -sheet II with a *trans* configuration, and the conserved acidic Glu471

residue is anchored on strand E (Fig. 3b). Unlike the N2 domain isopeptide bond, a single hydrogen bond is observed between the Asn515 side chain O_{δ1} atom and the carboxyl group of Glu471. This isopeptide bond is surrounded by hydrophobic residues Leu396, Phe401, Phe447, Ile513, and Ala399.

Metal-binding sites

The difference electron density maps suggest the presence of two high-affinity metal-binding sites ($>10\sigma$ in the Fo-Fc map) in the GBS80 N2 domain, one near its isopeptide bond and the other in the FG loop. The metal ions have been confirmed as Ca²⁺ ions in the CaGBS80₂₀₀₋₅₁₈ crystals, based on the average metal-ligand bond lengths and the coordination environment. Interestingly, the same metal sites are also identified in the tGBS80₂₀₀₋₅₁₈ structure, and this recombinant protein was not treated with a CaCl₂ additive. The coordinating residues for the first metal ion come from strand A and the A-I1 loop. The metal ion is coordinated to seven oxygen atoms. Four are from the Asp209, Asp211, and Asp218 side chains, one is from the main chain carbonyl of the Lys210 that is involved in isopeptide formation, and two from water molecules, which are held in place by the carbonyls of Asp319 from helix (H2) and Arg374 from G-I3 loop (Fig. 4a). The seven out of eight ligands for the second metal ion come from the FG loop alone, which include the side chain oxygens of Asp358, Thr360, Asp362, and Asn366 as well as the main chain carbonyls of the Thr360 and Ala364 residues (Fig. 4b). The eighth coordinating residue is Glu244, and no water molecules are present in the coordination sphere.

Comparison with other Gram-positive pilin structures

We have used the Dali server³⁴ and identified RrgB (PDB ID: **2X9W**; Z-score 18.2), BcpA (PDB ID: **3KPT**; Z-score 17.2), SpaA (PDB ID: **3HTL**; Z-score 15.0), and SspB (PDB ID: **2WZA**; Z-score 11.7) as structural homologues of GBS80. However, the primary sequence identity with these homologues is comparatively low, with a maximum of 28% for RrgB. When a similar search was carried out for individual N2 and N3 domains, the collagen binding segment (PDB ID: **2Z1P**) of *E. faecalis* MSCRAMM ACE³⁵ and the minor pilin GBS52 (PDB ID: **3PHS**) of *S. agalactiae*²³ are identified as structural homologues to the N2 and N3 domains, respectively in addition to the above mentioned four pilin structures.

The comparison between the GBS80N2N3 structures and other available full length or C-terminal fragment structures of 3- to 4-domain backbone pilins (RrgB, SpaA, and BcpA) reveals the following four similarities (Fig. 5): 1) a conserved DEv-IgG fold/cnaA-type followed by an IgG-rev fold/cnaB-type at the C-terminal end, with significant variations in the loop lengths and conformations; 2) a large hydrophobic domain interface between the two domains with a short antiparallel β -sheet at the domain junction; 3) the presence of an isopeptide triad (Lys-Asp/Glu-Asn) at identical positions; and 4) two isoforms of isopeptide bonds are identified in the major pilin structures. The isopeptide bonds in the DEv-IgG fold/cnaA-type domains display a *cis* configuration in GBS80, SpaA, and BcpA as well as a *trans* configuration in RrgB; the IgG-rev fold/cnaB-type domains in GBS80, SpaA, and RrgB display a *trans* configuration for the isopeptide bond, while it is *cis* in BcpA.

In RrgB, the N3 domain is inserted in a loop connecting last two strands of the N2 domain, positioning it parallel to N2. Thus, the spatial arrangement of the N2 and N4 domains resemble the N2 and N3 domains of GBS80. The structural superposition of GBS80N2N3 onto RrgB N2 and N4 resulted in an rms deviation of 1.9 Å for 176 common C α atoms. The structural superposition of GBS80N2N3 onto BcpAN3N4 resulted in an rms deviation of 1.6 Å for 149 common C α atoms. The C-terminal domains of the RrgB and BcpA crystal structures also did not display metal-binding sites, but they do have the three stranded β -sheet (I1, I2, I3) at the domain interface similar to GBS80N2N3.

The structural superposition of GBS80N2N3 onto SpaAN2N3 resulted in an rms deviation of 1.9 Å for 156 common Cα atoms, and revealed similar orientations for the N2 and N3 domains (Fig. 5 and 6b). Similar to GBS80N2N3, the N2 domain of SpaA also contains a metal-binding site in its AB loop, but it is away from the core and toward the (N2N3) domain interface. The D'E loop of the N2 domain is shorter by 21 residues in SpaA. The FG loop of SpaA is positioned on the opposite side of the molecule to the D'E loop side. However, the corresponding loop in GBS80N2N3 contains the second metal site in the N2 domain and extends toward the N1-N2 domain interface. Similarly, the BC loop of the N3 domain is shorter in SpaA by 12 residues. The DE loop in the SpaA N3 domain is extended toward the N2 domain and composes the four-stranded interface β-sheet. Unlike GBS80N2N3, a disulfide bridge, which connects strands B and E of the N3 domain, is also present in SpaA.

Discussion

Bacterial pili are important determinants in mediating host-pathogen interactions. While the assembly and the role of pili are well-investigated for Gram-negative pathogens,^{5,6} it is an emerging story for Gram-positive microbes (see reviews^{7,22,36,37}). Recent structural studies on the Gram-positive pili components have revealed many structural features that are distinct from the Gram-negative pili. These distinctions specifically include the presence of inter- and intra-molecular isopeptide bonds, and involvement of sortases in forming covalent linkages between the individual pilins, as well as in anchoring the assembled pili onto the bacterial cell wall. To understand the specifics of pilus assembly and anchoring, we initiated structural studies on individual pili components and the corresponding *S. agalactiae* sortases.²⁵ Recently, we reported the structure of minor pilin GBS52,²³ and, in recent years, several Gram-positive pilin structures have also been detailed by others.^{26-29,38,39} Interestingly all of them exhibit a mixture of IgG-like domains, which were classified as DEv-IgG fold/cnaA-types and IgG-rev fold/cnaB-types in Gram-positive MSCRAMMs.^{32,33}

The crystal structure of GBS80N2N3, presented herein, reveals a DEv-IgG fold in its N2 domain and an IgG-rev fold in N3, similar to the recently reported structures for major pilins RrgB,²⁹ SpaA,²⁷ and BcpA.²⁸ The isopeptide bonds and the participating Lys and Asn residues are structurally conserved among them. The isopeptide bond in the N2 domain connects two β-sheets, while that in the N3 domain connects two adjacent parallel strands of the same sheet I. As first discussed in the Spy0128 pilin structure,²⁶ the intra-molecular isopeptide bonds are formed by a reaction between an ε-amino group of a Lys residue and a carboxamide group of an Asn residue. A conserved acidic residue (Glu/Asp) in close proximity to the partnering Lys and Asn residues is shown to be essential for the isopeptide bond formation.^{27,28,40} Mutation of the conserved acidic residue to any other residue abrogates the formation of the corresponding amide bonds.^{27,28,40} Crystallography and mass spectrometry tools have been used in the past few years to identify and study the role of isopeptide bonds in pilus structure and stability.⁴¹ Recently, single molecule force spectroscopy techniques have been used to show that isopeptide bonds block the mechanical extensibility of pili in *S. pyogenes*.⁴² Similarly, combined biochemical and biophysical approaches were used to identify intramolecular isopeptide bonds and their stabilization properties in *S. pneumoniae*.⁴³ The recent crystal structures from commensal bacteria, such as oral streptococci also revealed the presence of isopeptide bonds, for example in the crystal structure of C-terminal domains of SspB.⁴⁴ The roles of isopeptide bonds and two high affinity metal-binding sites of GBS80 in the GBS pilus type 1 assembly are yet to be determined by mutational analysis and other biophysical studies.

Based on primary sequence homology with other full length pilins, we suggest that the N1 domain of GBS80 may be similar to the N1 domain of *C. diphtheriae* SpaA, which displays an IgG-rev fold. Our modeling of full length GBS80 based on the SpaA structure suggests the possibility of an internal isopeptide bond in N1 domain between residues Lys54 and Asn200, where the later residue is next to Lys199 in the pilin motif and the required catalytic acidic residue for the amide bond formation is Glu159. The internal isopeptide bond in the N1 domain is also found in the 4-domain pilins RrgB and BcpA, though not in the 3-domain SpaA. Interestingly, the recombinant proteins GBS80, RrgB and BcpA could be crystallized only after limited proteolysis, which removes the respective N1 domains at the Lys residue of their pilin motif. In contrast, full length, recombinant *C. diphtheriae* SpaA was able to form crystals, and its N1 domain has no isopeptide bond.²⁷ The pilin motif (YPKN) of several major pilins contains a Lys residue for the inter-molecular isopeptide bond and an Asn residue for the intra-molecular isopeptide bond formation in the N-terminal (N1) domain, as observed in BcpA.²⁸ The pairwise sequence comparison of GBS80 to full length RrgB and BcpA reveals the presence of the conserved Lys and Glu residues, required for the isopeptide bond in the respective N1 domains. Moreover, the Lys residue that is involved in the N2 domain isopeptide bond is observed after the 10 residues following the pilin motif Lys and the conserved bulky residue (generally Trp or Phe) that is associated with N1 domain is 9 residues before it (FxxxIxYPKNxxxx PKxDKD). Budizk *et al.* (2009) have suggested that the N1 domain isopeptide bond is synthesized only after the BcpA pilin precursor has been polymerized via the Lys162-Thr522 bond. In the absence of the N1 domain amide bond, the pili remain partially sensitive to protease cleavage.²⁸ Hence, we suggest that in the absence of the N1 domain intra-molecular Lys-Asn isopeptide bond, GBS80 is not stable and is susceptible to proteolysis, which also may be the reason for the reported difficulties during the crystallization of full length RrgB and BcpA.^{28,29}

A pilus-like form is seen in the crystal packing of the spy0128 and SpaA crystal structures,^{26,27} in which the successive backbone pilin molecules are stacked head-to-tail and have N- and C-termini in close proximity, suitable for inter-molecular isopeptide bond formation. Although the GBS80N2N3 structure lacks the N1 domain, similar head-to-tail packing can be seen in the crystal packing of the P2₁ space group model (Fig. 6a). The C-terminal region of the N3 domain, facing the N-terminus of the N2 domain (Asn200) at a distance of approximately 16 Å, lacks seven residues that would be sufficient for the inter-molecular isopeptide bond (Lys199-Thr524) formation. Accommodation of the N1 domain model, placed in the same orientation and position as observed in the SpaA structure (Fig. 6b and c), does not interfere with the head-to-tail stacking due to a ~20° tilt of N1 (Fig. 6d). Thus, it appears that the N2 and N3 domains primarily contribute to the GBS type 1 pilus shaft, while the N1 domain is positioned to strengthen the pilus by constricting the major pilin links (Fig. 6d and e).

When we compare the available crystal structures, the minor/ancillary pilins can be grouped into two categories, base and tip pilins.^{23,38,39} The tip pilins, RrgA of *S. pneumoniae*³⁸ and GBS104 of *S. agalactiae*, which are typically larger (~90 kDa) than the base pilins, contain a vWFA (von Willebrand factor A-type) domain that is inserted between multiple IgG-like domains.⁴⁵ The base pilins can be further divided into the following two groups, according to size: one- and two-domain pilins. GBS52 of *S. agalactiae*²³ and likely RrgC of *S. pneumoniae* are composed of two IgG-rev domains, whereas FctB of *S. pyogenes* (GAS M49 strain)³⁹ and SpaB of *C. diphtheriae* consist of a single IgG-rev domain.

The incorporating mechanisms of the smaller pilin, which do not have a conventional pilin motif, into the pilus shaft are still emerging.⁸ In *C. diphtheriae*, a conserved Glu residue in the E-box (YxLxE⁺TxAPxGY) of the major pilin, SpaA, has previously been implicated in the incorporation of ancillary pilins, as its substitution by Ala or Arg disrupts incorporation

of the minor pilin, SpaB.¹² However, recent studies reveal that Glu446 in the E-box is the catalytic residue for the Lys363–Asn462 intra-molecular isopeptide bond,²⁷ suggesting that the E-box residues are required primarily for protein stability. Similarly, in the GBS80N2N3 structure, Glu471 in the E-box (LK $\underline{\text{E}}$ T) is associated with the isopeptide bond between Lys387 and Asn515 in the N3 domain, which is also the case with the other major pilins BcpA and RrgB. Recent biochemical and genetic studies on the corynebacterial minor pilin, SpaB,⁸ revealed that its Lys139 residue is required for the SpaB–SpaA linkage that is catalyzed by the pilus-specific sortase, similar to the mechanism that cross-links major pilins. Structural superposition of the Group A streptococcal single-domain minor pilin FctB and the N1 domain of its major pilin, Spy0128, reveals that the Lys110 of the former, located on the Ω loop in last the strand (G),³⁹ aligns with the pilin-motif Lys161 in the major pilin, also present in a similar loop. Using mass spectrometry and homology modeling, Linke *et al.* (2010) revealed that Lys110 in the minor pilin FctB is linked with the Thr residue in the LPXTG motif of the major pilus protein.³⁹

Interestingly, a similar pilin and pilin-like motif can be seen in GBS80 and GBS52, respectively upon their primary, secondary and tertiary structural superposition. The minor pilin, GBS52, is made of two IgG-like domains (Fig. 1b and 7a),²³ and we notice that the Lys148 residue is located at the C-terminal end of the last strand (G) in the N1 domain within a pilin-like motif (¹⁴⁵IYPK $\underline{\text{I}}$ ¹⁴⁹). On superposition, similarly folded (IgG-rev) and with an identical core, the GBS52 N1 and SpaA N1 domains (rmsd 1.9 Å for 96 Ca atoms) (Fig. 7a and b) have overlapping pilin and pilin-like motifs, respectively. In addition, the GBS80 model and GBS52 structure display relatively similar orientations between the respective N1 and N2 domains (Fig. 7a). Hence, the N1 domain in the GBS52 could be placed in a similar position and orientation with the N1 domain of the full length GBS80 model we built using the SpaA structure. Such superposition would align the N2 domain from GBS52 along the long axis of the pilus shaft, and bring Lys148 in GBS52 and Thr524 in GBS80 close enough for inter-molecular cross linking (Fig. 7c). Incidentally, the GBS52 has a striking sequence similarity (92%) to ancillary protein 2 in *S. pyogenes* (uniprotKB ID: **B8QYJ0**, **Q1JIZ8**-serotype M2, strain MGAS10270_Spy0110), where the above mentioned Lys residue in the pilin-like motif is also conserved in the primary sequence and possibly at the tertiary structural level. Such a high sequence similarity is uncommon between pilins from different organisms and may suggest some common features among the various Gram-positive pilus assemblies. GBS1474, an ortholog of GBS52 (strain NEM316) with a 25% sequence identity, also has a conserved pilin-like motif (INPK $\underline{\text{K}}$) and the required residues for an isopeptide bond in the N2 domain. The T-antigen-like fimbrial structural subunit protein fszE from *Streptococcus equi* subsp. *zooepidemicus* (strain MGCS10565, uniprotKB ID: **B4U0E6**) and *Streptococcus dysgalactiae* subsp. *equisimilis* (strain CGS_124, uniprotKB ID: **C5WE61**), which share 50% and 46% sequence identities with GBS52, respectively, also contain similar pilin-like motifs. Interestingly similar to the major pilins, in the two-domain minor pilins, the conserved Lys residue in the pilin-like motif is located approximately 10 residues away from the Lys residue that forms the isopeptide bond in the N2 domain.

The proline rich C-terminal end of GBS52 was shown to be rigid and extend further away from the N2 domain with a 13 residue separation between the sorting motif and the body of the molecule. A similar arrangement was seen in the FctB crystal structure and also identified in the minor pilins from GAS and *C. diphtheriae*.³⁹ Suggestions were made to the effect that the extended C-terminal tail participates as a linkage to the peptidoglycan, facilitating cell wall anchoring of the pili. However, unlike FctB in Group A streptococci and GBS52, where its role may be confined to the pilus base, SpaB in *C. diphtheriae* was shown to be irregularly interspersed along the pilus shaft and at the base.⁴⁶

The GBS80N2N3 structure and the structural model for the GBS type 1 pili we described above has enhanced our knowledge and understanding of the GBS pilus architecture in particular and Gram-positive pili assembly in general. In the future, we need to obtain detailed structural information that can facilitate our understanding of the inter-molecular isopeptide bond formation and ancillary pilin incorporation, characteristic features of Gram-positive pili.

Materials and Methods

Expression and Purification

The rGBS80₃₆₋₅₁₈ with an N-terminal His-tag was expressed in *E. coli* and purified by Ni-affinity (HiTrap Chelating column), gel filtration (S200 (26/60) Sephacryl column) and ion-exchange (HiTrap Q column) chromatographic techniques as described previously.³¹ The 35 kDa fragment containing the N2 and N3 domains was obtained via limited proteolysis using α -chymotrypsin.³¹ Briefly, the purified rGBS80₃₆₋₅₁₈ was incubated overnight using a 100:1 protein to protease ratio at room temperature in digestion buffer, 30 mM Tris-HCl pH 7.4, 100 mM NaCl. Protease inhibitor PMSF at a concentration of 0.5 mM was added to the digested GBS80, which was purified using size-exclusion chromatography.

Crystallization and Data collection

The GBS80 35 kDa fragment was concentrated to 10 mg/ml and crystallized using a vapor-diffusion method. Four initial conditions gave thin plate-like crystals after three weeks and were optimized to produce large crystals suitable for X-ray diffraction analysis. A droplet with 2 μ l of protein in buffer containing 20 mM Tris-HCl pH 7.0, 100 mM NaCl and 2 μ l reservoir solution (0.1 M MES pH 5.5, 25% PEG MME 2000) was equilibrated against 1.0 ml reservoir solution at 295K. The recombinant GBS80 was degraded to a similar 35 kDa fragment when purified with 10 mM CaCl₂ as an additive to all of the buffers. The CaCl₂-treated GBS80 fragment was also crystallized in the same condition (0.1 M MES pH 5.5, 25% PEG MME 2000), though the space group was different (Table 1).

Native diffraction data sets from the α -chymotrypsin and CaCl₂-treated GBS80 crystals were collected to 1.8 and 1.7 Å on an R-Axis IV imaging-plate detector mounted on an in-house RIGAKU® rotating-anode X-ray generator operating at 100 mA and 50 kV, using 20% (v/v) ethylene glycol (EG) as a cryoprotectant. The crystals of recombinant GBS80₂₀₁₋₅₁₈, obtained by treating GBS80 with α -chymotrypsin, belong to the monoclinic *P*2₁ space group with two molecules in the asymmetric unit, while the CaCl₂-treated GBS80 crystals belongs to the monoclinic *C*2 space group with one molecule in the asymmetric unit.

Structure determination, model building, and refinement

The 35-kDa fragment from GBS80 contains four Arg and forty-one Lys (14%) residues. Anticipating the potential for binding halide anions, we introduced anomalous scatterers by screening various halide anions. Initial phases were obtained via the single-wavelength anomalous dispersion (SAD) method using a Sodium Iodide (NaI) derivative. The NaI derivative crystals were prepared using a halide quick-soaking technique,³⁰ where the native GBS80₂₀₁₋₅₁₈ crystals were soaked in the reservoir solution containing 200 mM NaI for one minute. The SAD data were collected at a 1.5418 Å wavelength using an in-house source at our data collection facility (Table 1), and the data were processed with *D*TREK*.⁴⁷ Phase calculations and initial model building were carried out using *autoSHARP*.⁴⁸ Ten iodide sites (Fig. 2a) were used in the *autoSHARP* phase calculations (FOM 0.33, Phasing power 1.33), and the initial model consisted of 240 residues with R/R_{free} at 28.5/34.6%. Final model building and refinement were completed using *COOT*⁴⁹ and *REFMAC*⁵⁰ from the

CCP4 suite.⁵¹ The GBS80-NaI derivative structure was used in the refinement of the α -chymotrypsin- and CaCl_2 -treated GBS80 structures (Table 2). The quality of the final models was examined using *COOT*.

Acknowledgments

This work was supported by NIH grants (AI061381) to H. T-T and (AI073521) to S.V.L.N.

REFERENCES

1. Ton-That H, Schneewind O. Assembly of pili in Gram-positive bacteria. *Trends Microbiol.* 2004; 12:228–234. [PubMed: 15120142]
2. Lauer P, Rinaudo CD, Soriani M, Margarit I, Maione D, Rosini R, Taddei AR, Mora M, Rappuoli R, Grandi G, Telford JL. Genome analysis reveals pili in Group B *Streptococcus*. *Science.* 2005; 309:105. [PubMed: 15994549]
3. Dramsi S, Caliot E, Bonne I, Guadagnini S, Prevost MC, Kojadinovic M, Lalioui L, Poyart C, Trieu-Cuot P. Assembly and role of pili in group B streptococci. *Mol. Microbiol.* 2006; 60:1401–1413. [PubMed: 16796677]
4. Rosini R, Rinaudo CD, Soriani M, Lauer P, Mora M, Maione D, Taddei A, Santi I, Ghezzi C, Brettoni C, Buccato S, Margarit I, Grandi G, Telford JL. Identification of novel genomic islands coding for antigenic pilus-like structures in *Streptococcus agalactiae*. *Mol. Microbiol.* 2006; 61:126–141. [PubMed: 16824100]
5. Proft T, Baker EN. Pili in Gram-negative and Gram-positive bacteria - structure, assembly and their role in disease. *Cell. Mol. Life Sci.* 2009; 66:613–635. [PubMed: 18953686]
6. Sauer FG, Mulvey MA, Schilling JD, Martinez JJ, Hultgren SJ. Bacterial pili: molecular mechanisms of pathogenesis. *Curr. Opin. Microbiol.* 2000; 3:65–72. [PubMed: 10679419]
7. Kline KA, Dodson KW, Caparon MG, Hultgren SJ. A tale of two pili: assembly and function of pili in bacteria. *Trends Microbiol.* 2010; 18:224–232. [PubMed: 20378353]
8. Mandlik A, Das A, Ton-That H. The molecular switch that activates the cell wall anchoring step of pilus assembly in Gram-positive bacteria. *Proc. Natl. Acad. Sci. USA.* 2008; 105:14147–14152. [PubMed: 18779588]
9. Patti JM, Allen BL, McGavin MJ, Höök M. MSCRAMM-mediated adherence of microorganisms to host tissues. *Annu. Rev. Microbiol.* 1994; 48:585–617. [PubMed: 7826020]
10. Schneewind O, Model P, Fischetti VA. Sorting of protein A to the staphylococcal cell wall. *Cell.* 1992; 70:267–281. [PubMed: 1638631]
11. Fischetti VA, Pancholi V, Schneewind O. Conservation of a hexapeptide sequence in the anchor region of surface proteins from Gram-positive cocci. *Mol. Microbiol.* 1990; 4:1603–1605. [PubMed: 2287281]
12. Ton-That H, Marraffini LA, Schneewind O. Sortases and pilin elements involved in pilus assembly of *Corynebacterium diphtheriae*. *Mol. Microbiol.* 2004; 53:251–261. [PubMed: 15225319]
13. Dermer P, Lee C, Eggert J, Few B. A history of neonatal group B *Streptococcus* with its related morbidity and mortality rates in the United States. *J. Pediatr. Nurs.* 2004; 19:357–363. [PubMed: 15614260]
14. Bolanos M, Canas A, Santana OE, Perez-Arellano JL, de Miguel I, Martin-Sanchez AM. Invasive group B streptococcal disease in nonpregnant adults. *Eur. J. Clin. Microbiol. Infect. Dis.* 2001; 20:837–839. [PubMed: 11783707]
15. Edwards MS, Rench MA, Palazzi DL, Baker CJ. Group B streptococcal colonization and serotype-specific immunity in healthy elderly persons. *Clin. Infect. Dis.* 2005; 40:352–357. [PubMed: 15668856]
16. Schuchat A. Epidemiology of group B streptococcal disease in the United States: shifting paradigms. *Clin. Microbiol. Rev.* 1998; 11:497–513. [PubMed: 9665980]
17. Sendi P, Johansson L, Norrby-Teglund A. Invasive group B Streptococcal disease in non-pregnant adults : a review with emphasis on skin and soft-tissue infections. *Infection.* 2008; 36:100–111. [PubMed: 18193384]

18. Johri AK, Paoletti LC, Glaser P, Dua M, Sharma PK, Grandi G, Rappuoli R. Group B *Streptococcus*: global incidence and vaccine development. *Nat. Rev. Microbiol.* 2006; 4:932–942. [PubMed: 17088932]
19. Heelan JS, Hasenbein ME, McAdam AJ. Resistance of group B *Streptococcus* to selected antibiotics, including erythromycin and clindamycin. *J. Clin. Microbiol.* 2004; 42:1263–1264. [PubMed: 15004089]
20. Lin FY, Azimi PH, Weisman LE, Philips JB III, Regan J, Clark P, Rhoads GG, Clemens J, Troendle J, Pratt E, Brenner RA, Gill V. Antibiotic susceptibility profiles for group B streptococci isolated from neonates, 1995–1998. *Clin. Infect. Dis.* 2000; 31:76–79. [PubMed: 10913400]
21. Margarit I, Rinaudo CD, Galeotti CL, Maione D, Ghezzi C, Buttazzoni E, Rosini R, Runci Y, Mora M, Buccato S, Pagani M, Tresoldi E, Berardi A, Creti R, Baker CJ, Telford JL, Grandi G. Preventing bacterial infections with pilus-based vaccines: the group B streptococcus paradigm. *J. Infect. Dis.* 2009; 199:108–115. [PubMed: 19086816]
22. Telford JL, Barocchi MA, Margarit I, Rappuoli R, Grandi G. Pili in Gram-positive pathogens. *Nat. Rev. Microbiol.* 2006; 4:509–519. [PubMed: 16778837]
23. Krishnan V, Gaspar AH, Ye N, Mandlik A, Ton-That H, Narayana SV. An IgG-like domain in the minor pilin GBS52 of *Streptococcus agalactiae* mediates lung epithelial cell adhesion. *Structure.* 2007; 15:893–903. [PubMed: 17697995]
24. Soriani M, Telford JL. Relevance of pili in pathogenic streptococci pathogenesis and vaccine development. *Future Microbiol.* 2010; 5:735–747. [PubMed: 20441546]
25. Khare B, Samal A, Vengadesan K, Rajashankar KR, Ma X, Huang IH, Ton-That H, Narayana SV. Preliminary crystallographic study of the *Streptococcus agalactiae* sortases, sortase A and sortase C1. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 2010; 66:1096–1100.
26. Kang HJ, Coulibaly F, Clow F, Proft T, Baker EN. Stabilizing isopeptide bonds revealed in Gram-positive bacterial pilus structure. *Science.* 2007; 318:1625–1628. [PubMed: 18063798]
27. Kang HJ, Paterson NG, Gaspar AH, Ton-That H, Baker EN. The *Corynebacterium diphtheriae* shaft pilin SpaA is built of tandem Ig-like modules with stabilizing isopeptide and disulfide bonds. *Proc. Natl. Acad. Sci. USA.* 2009; 106:16967–16971. [PubMed: 19805181]
28. Budzik JM, Poor CB, Faull KF, Whitelegge JP, He C, Schneewind O. Intramolecular amide bonds stabilize pili on the surface of bacilli. *Proc. Natl. Acad. Sci. USA.* 2009; 106:19992–19997. [PubMed: 19903875]
29. Spraggon G, Koesema E, Scarselli M, Malito E, Biagini M, Norais N, Emolo C, Barocchi MA, Giusti F, Hillergermann M, Rappuoli R, Lesley S, Covacci A, Maignani V, Ferlenghi I. Supramolecular organization of the repetitive backbone unit of the *Streptococcus pneumoniae* pilus. *PLoS ONE.* 2010; 5:e10919. [PubMed: 20559564]
30. Dauter Z, Dauter M, Rajashankar KR. Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. *Acta Crystallogr. D Biol. Crystallogr.* 2000; 56:232–237. [PubMed: 10666615]
31. Krishnan V, Ma X, Prabhat D, Ton-That H, Narayana SV. Purification, crystallization and halide phasing of a *Streptococcus agalactiae* backbone pilin GBS80 fragment. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 2010; 66:1666–1669.
32. Symersky J, Patti JM, Carson M, House-Pompeo K, Teale M, Moore D, Jin L, Schneider A, DeLucas LJ, Höök M, Narayana SV. Structure of the collagen-binding domain from a *Staphylococcus aureus* adhesin. *Nat. Struct. Biol.* 1997; 4:833–838. [PubMed: 9334749]
33. Deivanayagam CC, Rich RL, Carson M, Owens RT, Danthuluri S, Bice T, Höök M, Narayana SV. Novel fold and assembly of the repetitive B region of the *Staphylococcus aureus* collagen-binding surface protein. *Structure.* 2000; 8:67–78. [PubMed: 10673425]
34. Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 2010; 38(Suppl):W545–549. [PubMed: 20457744]
35. Liu Q, Ponnuraj K, Xu Y, Ganesh VK, Sillanpää J, Murray BE, Narayana SV, Höök M. The *Enterococcus faecalis* MSCRAMM ACE binds its ligand by the Collagen Hug model. *J. Biol. Chem.* 2007; 282:19629–19637. [PubMed: 17392280]

36. Mandlik A, Swierczynski A, Das A, Ton-That H. Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. *Trends Microbiol.* 2008; 16:33–40. [PubMed: 18083568]
37. Scott JR, Záhner D. Pili with strong attachments: Gram-positive bacteria do it differently. *Mol. Microbiol.* 2006; 62:320–330. [PubMed: 16978260]
38. Izoré T, Contreras-Martel C, El Mortaji L, Manzano C, Terrasse R, Vernet T, Di Guilmi AM, Dessen A. Structural basis of host cell recognition by the pilus adhesin from *Streptococcus pneumoniae*. *Structure.* 2010; 18:106–115. [PubMed: 20152157]
39. Linke C, Young PG, Kang HJ, Bunker RD, Middleditch MJ, Caradoc-Davies TT, Proft T, Baker EN. Crystal structure of the minor pilin FctB reveals determinants of Group A streptococcal pilus anchoring. *J. Biol. Chem.* 2010; 285:20381–20389. [PubMed: 20427291]
40. Kang HJ, Baker EN. Intramolecular isopeptide bonds give thermodynamic and proteolytic stability to the major pilin protein of *Streptococcus pyogenes*. *J. Biol. Chem.* 2009; 284:20729–20737. [PubMed: 19497855]
41. Kang HJ, Middleditch M, Proft T, Baker EN. Isopeptide bonds in bacterial pili and their characterization by X-ray crystallography and mass spectrometry. *Biopolymers.* 2009; 91:1126–1134. [PubMed: 19226623]
42. Alegre-Cebollada J, Badilla CL, Fernández JM. Isopeptide bonds block the mechanical extension of pili in pathogenic *Streptococcus pyogenes*. *J. Biol. Chem.* 2010; 285:11235–11242. [PubMed: 20139067]
43. El Mortaji L, Terrasse R, Dessen A, Vernet T, Di Guilmi AM. Stability and assembly of pilus subunits of *Streptococcus pneumoniae*. *J. Biol. Chem.* 2010; 285:12405–12415. [PubMed: 20147289]
44. Forsgren N, Lamont RJ, Persson K. Two intramolecular isopeptide bonds are identified in the crystal structure of the *Streptococcus gordonii* SspB C-terminal domain. *J. Mol. Biol.* 2010; 397:740–751. [PubMed: 20138058]
45. Konto-Ghiorgi Y, Mairey E, Mallet A, Dumenil G, Caliot E, Trieu-Cuot P, Dramsi S. Dual role for pilus in adherence to epithelial cells and biofilm formation in *Streptococcus agalactiae*. *PLoS Pathog.* 2009; 5:e1000422. [PubMed: 19424490]
46. Ton-That H, Schneewind O. Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol. Microbiol.* 2003; 50:1429–1438. [PubMed: 14622427]
47. Pflugrath JW. The finer things in X-ray diffraction data collection. *Acta Crystallogr. D Biol. Crystallogr.* 1999; 55:1718–1725. [PubMed: 10531521]
48. Vonrhein C, Blanc E, Roversi P, Bricogne G. Automated structure solution with autoSHARP. *Methods Mol. Biol.* 2007; 364:215–230. [PubMed: 17172768]
49. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 2004; 60:2126–2132. [PubMed: 15572765]
50. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 1997; 53:240–255. [PubMed: 15299926]
51. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 1994; 50:760–763. [PubMed: 15299374]

Figure 1. Schematic representation of the modules in the Group B Streptococcus (GBS) pilins

(a) Schematic diagram of the backbone pilin GBS80. It contains a signal peptide (S) at the N-terminus and a C-terminal cell wall sorting signal, consisting of a LPXTG motif (W, shown in red), membrane-spanning hydrophobic domain (M, magenta) and cytoplasmic positively charged tail (C, blue), as observed in MSCRAMMs and other Gram-positive pilins. The central portion contains three IgG-like domains referred to as N1, N2 and N3. The Pilin motif (YPKN) is present near the N1-N2 domain linker. The N3 domain contains an E-box (LKET).

(b) Minor pilin GBS52, having two domains and a pilin-like motif that we identified in the C-terminal end of the N1 domain. The solid lines at the bottom indicate the intra-molecular isopeptide bonds.

Figure 2. Tertiary fold of GBS80N2N3

(a) Anomalous difference Fourier map contoured at the 5σ level for the GBS80N2N3 fragment showing ten iodide sites (magenta) used in the SAD phase calculation. (b) Ribbon representation of the 35 kDa C-terminal major fragment from GBS80, containing two IgG-like domains (N2 and N3). The strands from each domain are colored in rainbow style from red (N-terminus) to violet (C-terminus). The N2 domain displays a DEv-IgG fold/cnaA-type while the N3 domain exhibits an IgG-rev fold/cnaB-type. Both the N2 and N3 domains contain an isopeptide bond, shown as sticks. The two calcium binding sites in the N2 domain are shown in grey. The strands (I1, I2, and I3) of short β -sheet at domain interface and the three helices (H1, H2 and H3) are shown in pink and gray, respectively. (c) Topology diagram of IgG-C, IgG-rev and DEv-IgG folds, with their core β strands labeled A to G in a rainbow style (red to violet) as in Figure 2(b). The horizontal, solid straight line indicates approximate position of the isopeptide bond.

Figure 3. The intra-molecular isopeptide bonds in the N2 (a) and N3 (b) domains from GBS80N2N3

The isopeptide bond, its conserved acidic residue, and surrounding hydrophobic residues are shown using stick models. The dotted lines represent hydrogen bonds and the numbers listed are distances in angstroms.

Figure 4. The metal-binding sites (site 1 (a) & 2(b)) in the N2 domain of GBS80N2N3

The calcium ions are depicted using grey spheres. The residues involved in the metal coordination are shown with sticks and the metal coordination bonds are illustrated using broken lines. In site 1, the metal ion is coordinated by seven oxygen atoms from the side chains of Asp209, Asp211, and Asp218; the main chain of Lys210, which is involved in isopeptide formation; and two water molecules. In the second metal-binding site, the eight ligands are from the side chains of Glu244, Asp358, Thr360, Asp362, and Asn366 as well as the main chains of Thr360 and Ala364.

Figure 5. Comparison of the GBS80N2N3 structure with the equivalent domains from other similar backbone pilin structures RrgBN2N4, BcpAN3N4, and SpaAN2N3

The core β strands in each structure are labeled A to G in a rainbow style (red to violet) as in Figure 2(b). The GBS80N2N3 is kept in the same orientation as in Figure 2(b). The RrgBN2N4, BCPAN3N4, and SpaAN2N3 are superimposed on GBS80N2N3 and translated to show them in the similar orientation. The crystal structures of BcpA (PDB code [3KPT](#)) and RrgB (PDB code [2X9W](#)) lack the N1 domain, while the structure of full length SpaA (PDB code [3HTL](#)) is available.

Figure 6. The GBS pilus shaft model

(a) Crystal packing of the GBS80N2N3 N2 and N3 domains. (b) Superposition of the N2 and N3 domains from the GBS80N2N3 (cyan) structure onto the corresponding domains in SpaA (magenta). (c) Model for GBS80. The N1 domain is modeled and placed, based on the SpaA structure. (d) Schematic representation of pilus shaft formation via the backbone pilin GBS80. The two-sided arrow points the location of C terminus and pilin motif for inter-molecular isopeptide bond formation. (e) Corresponding cartoon representation of the GBS type 1 pilus shaft model. The two-sided arrow indicates the inter-molecular isopeptide bond and the dotted lines show the pilus shaft termini.

Figure 7. Minor pilin incorporation in the GBS pilus shaft model

(a) Superposition of the N1 domains from minor pilin GBS52 (green) onto the corresponding domain from major pilin SpaA (magenta). As the N2 domains of GBS52 and SpaA have different folds (IgG-rev and DEv-IgG), the N1 domain was used for superpositioning. (b) Superposition of the N1 domain from GBS52 (green) onto the SpaA N1 domain (magenta). The pilin-like motif (YPK in sticks) is enlarged to show their similarly positioned and oriented residues. (c) Model for incorporation of the minor pilins, GBS52 at the base and GBS104 at the tip. The two sided white arrow indicates the possible intermolecular isopeptide bond. The two sided black arrow indicates the possible covalent bond between the base pilin and peptidoglycan cross-bridge.

Table 1

Data collection and processing statistics. The numbers in parentheses correspond to the values in the highest resolution shell.

	tGBS80N2N3	CaGBS80N2N3	NaI-GBS80N2N3
Resolution range (Å)	40-1.8 (1.86-1.8)	40-1.7 (1.76-1.7)	40-2.1 (2.18-2.1)
Space group	<i>P2₁</i>	<i>C2</i>	<i>P2₁</i>
Unit-cell parameters (a, b, c in Å; β in °)	69.9, 60.3, 80.4; 101.6	130.5, 34.7, 74.6; 93.3	34.9, 59.7, 80.8; 101.6
Unique reflections	60630	37173	18718
Multiplicity	2.8 (2.7)	4.3 (4.1)	6.5 (6.4)
Mean $I/\sigma(I)$	9.6 (3.2)	26.7(6.5)	20.7(11.2)
Completeness (%)	99.6 (100)	99.8(100)	97.7 (95.5)
* R_{merge} (%)	6.3 (29.5)	2.8 (17.1)	5.4 (12.1)

* $R_{\text{merge}} = \sum_{hkl} \sum_j |I(hkl)_j - [I(hkl)]| / \sum_{hkl} \sum_i I(hkl)_i$, where $I(hkl)$ are the intensities of symmetry-related reflections and $[I(hkl)]$ is the average intensity over all observations

Table 2

Model refinement statistics.

	tGBS80N2N3	CaGBS80N2N3
Resolution range (Å)	40-1.8	40-1.7
No. of reflections used	60617	37170
R _{cryst} /R _{free} (%)	20.3/24.0	20.3/23.8
Average B value (Å ²) for protein/metal/solvent	17.3/13.7/27.4	17.8/13.1/25.8
Rmsd, bonds (Å)	0.011	0.009
Rmsd, angles (°)	1.25	1.22
No. of protein/metal atoms	4857/4	2457/2
No. of solvent atoms	880	398
Ramachandran plot Preferred/Allowed/outliers (%)	96.31/3.53/0.16	96.15/3.53/0.32