Structural Basis of Pilus Anchoring by the Ancillary Pilin RrgC of *Streptococcus pneumoniae*

Md Munan Shaik†§¶, Amandine Maccagni†§¶, Guillaume Tourcier†§¶, Anne Marie Di Guilmi†§¶, and Andréa Dessen†§¶

From the †Institut de Biologie Structurale (IBS), Université Grenoble Alpes, 41 avenue des Martyrs, 38044 Grenoble, France, the §Commissariat à l’Energie Atomique (CEA), 38000 Grenoble, France, the ¶Centre National de la Recherche Scientifique (CNRS), UMR 5075, Grenoble, France, and the ¶Brazilian National Laboratory for Biosciences (LNBio), CNPEM, Campinas, 13083-100 São Paulo, Brazil

**Background:** RrgC is a key component of the pneumococcal pilus, a virulence factor that plays an important role in pathogenesis.

**Results:** RrgC folds into three independent domains and requires the housekeeping sortase for surface association.

**Conclusion:** The rod-like structure of RrgC suggests that it stably bridges peptidoglycan and pilus fiber.

**Significance:** A complete model of the pneumococcal pilus reveals a multidomain, flexible assembly.

Pili are surface-attached, fibrous virulence factors that play key roles in the pathogenesis process of a number of bacterial agents. *Streptococcus pneumoniae* is a causative agent of pneumonia and meningitis, and the appearance of drug-resistance organisms has made its treatment challenging, especially in developing countries. Pneumococcus-expressed pili are composed of three structural proteins: RrgB, which forms the polymerized backbone, RrgA, the tip-associated adhesin, and RrgC, which presumably associates the pilus with the bacterial cell wall. Despite the fact that the structures of both RrgA and RrgB were known previously, structural information for RrgC was still lacking, impeding the analysis of a complete model of pilus architecture. Here, we report the structure of RrgC to 1.85 Å and reveal that it is a three-domain molecule stabilized by two intradomain isopeptide bonds. RrgC does not depend on pilus-specific sortases to become attached to the cell wall; instead, it binds the preformed pilus to the peptidoglycan by employing the catalytic activity of SrtA. A comprehensive model of the type 1 pilus from *S. pneumoniae* is also presented.

*Streptococcus pneumoniae* (the pneumococcus) is a mucosal commensal as well as a major human pathogen that contributes outstandingly to mortality rates among young children and the elderly in the developing world (1, 2). *S. pneumoniae* is the most important bacterial cause of pneumonia and is also the causative agent of meningitis and septicemia; the elevated number of victims is directly related to the limited efficacy and availability of vaccines as well as the spread of antibiotic-resistant strains (2–4). A number of virulence factors, such as the polysaccharide capsule, pneumolysin, and autolysin, play important roles in infection (5), but the mechanism employed by the pneumococcus to adhere to host targets, a key step in infection initiation, is still not well understood.

Pili, flexible hair-like structures that are directly associated with the bacterial surface, have been identified in a number of Gram-positive bacteria, including *S. pneumoniae*, * Corynebacterium diphtheriae*, *Streptococcus pyogenes*, and *Bacillus anthracis*. These thin fibers have been shown to play important roles in host tissue colonization and pathogenesis by enhancing adherence to host cells, recognizing the extracellular matrix, and participating in biofilm formation (6–11). Notably, in *S. pneumoniae*, pili have been shown to manipulate the host inflammatory response, and pilin subunits have been reported to confer protection against a lethal bacterial challenge in animal models (7, 12).

In contrast to the well studied pili of Gram-negative organisms, which are formed through the noncovalent association of building blocks, pili in Gram-positive bacteria are assembled covalently through the action of specific transpeptidases (sortases) that recognize cell wall sorting signals such as LPXTG-like sequences. The quintessential sortase, sortase A from *Staphylococcus aureus*, was the first enzyme to be reported to catalyze the covalent association of virulence factors directly to the peptidoglycan (13); since then, sortases have been classified into four distinct classes (A, B, C, and D), with class C enzymes participating in covalent pilin association (14). The catalytic mechanism of class C sortases involves the nucleophilic attack of the epsilon-amino from a surface lysine residue of the secondary substrate (15).

In the pilus formation mechanism, sortases are responsible for the covalent association of one major pilin (forming the backbone) as well as two minor pilins (tip and base pilins) in most cases. The first pilus assembly model was proposed for *C. diphtheriae* (16) and later found to be applicable to pili from other Gram-positive organisms (17). In *S. pneumoniae*, pili are encoded by the *rlrA* pathogenicity islet, which carries genes for...
three pilin proteins (the major pilin RrgB and two minor pilins, RrgA and RrgC) as well as three sortases (SrtC-1, SrtC-2, and SrtC-3) (7, 18, 19). Cell fractionation, transmission electron microscopy, in vitro polymerization tests, and cell wall sorting signal domain swapping experiments have confirmed that covalently linked repeating units of RrgB form the S. pneumoniae pilus backbone whereas RrgA is present at the tip of the pilus (18, 20–25). Notably, RrgA was shown to be able to recognize extracellular matrix elements in vitro, a finding confirmed by its three-dimensional structure, which revealed the presence of an integrin collagen-recognition domain (9, 25, 26). These elements thus collectively suggest that RrgA plays the role of pilus adhesin. However, the role played by RrgC in pilus functionality is less well understood.

Group A Streptococcus strains carrying a knock-out of a close homolog of RrgC, spv0130, can still form pilin, but these are not attached to the cell wall and are released into the medium (27). In addition, deletion of the gene coding for the minor base pilin GBSS2 from Streptococcus agalactiae significantly reduces bacterial adhesion to pulmonary epithelial cells, an effect confirmed by the ability of anti-GBSS2 antibodies to inhibit epithelial cell adhesion in vitro (28). These observations, coupled to the fact that RrgC was shown to be located at the base of the pilus fiber, suggested that it could play the role of pilus anchor to the cell wall (25). Despite these advances, the three-dimensional fold of RrgC, as well as its mode of attachment to the pilus base and/or the cell wall, remained unknown, impeding a more detailed comprehension of streptococcal pilus architecture.

In this work we report the crystal structure of RrgC to 1.85 Å resolution and show that it displays three independent domains, two of which are stabilized by isopeptide bonds. RrgC depends mostly on the housekeeping SrtA, and not any of the expected pilus related SrtC enzymes, to link the pilus to the cell wall (25). Despite these advances, the three-dimensional structure of RrgC, as well as its mode of attachment to the pilus, remained unknown, impeding a more detailed comprehension of streptococcal pilus architecture.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Plasmids—S. pneumoniae strains were grown at 30 °C or 37 °C in CY (29) or Todd Hewitt medium (TH; BD Biosciences). Blood agar plates were made from Columbia agar containing 5% defibrinated horse blood (Difco). For transformation of S. pneumoniae, 500 ng of plasmid DNA was added to competent cells (diluted in 1/10 CY medium containing 0.015% albumin) followed by a phenotypic expression period of 30 min at 30 °C, 2 h at 37 °C, and overnight growth on Columbia blood agar plates containing tetracycline (final concentration 2.5 µg/ml).

The pLIM25 plasmid was used to express RrgC, Srt-C1, and Srt-C3 in pneumococcal strains (30). Protein expression was under the control of the zinc-inducible PczcD promoter and required addition of 0.15 mM ZnCl2 until A600 = 0.3–0.4, resuspended in fresh CY medium, and transferred to microscope slides. Slides were observed using an Olympus BX61 optical microscope equipped with a UPFLN 100∗ O-2PH/1.3 objective and a QImaging Retiga-SRV 1394 cooled charge-coupled device camera. Image acquisition was performed using the Volocity software package and processed with Adobe Photoshop 6.0.

Pneumococcal Subcellular Fractionation—The presence of RrgC in different cell compartments of the R6 wild-type and R6ΔsrtA strains was analyzed by cell fractionation. One-tenth of a 100-ml culture in exponential growth phase was centrifuged (15 min at 3000 × g), and the pellet was resuspended in 1 ml of PBS containing 100 µg/ml lysozyme and 50 units/ml mutanolysin and incubated for 2 h at 37 °C, yielding the total cell lysate extract. The remaining 90 ml were centrifuged (15 min at 3000 × g), and the pellet was resuspended in 9 ml of PBS containing 100 µg/ml lysozyme, 50 units/ml mutanolysin, and 30% sucrose and incubated for 2 h at 37 °C. After centrifugation, the supernatant containing the cell wall was collected, and protein quantification was performed by the BCA method. The pellet protoplasts were resuspended in 9 ml of 10 mM HEPES, 10 mM KCl (pH 7.4), lysed by three cycles of freeze/thawing, and ultracentrifuged for 45 min at 190,000 × g. The resulting membrane pellet was resuspended in 9 ml of buffer containing 10 mM HEPES, 10 mM KCl (pH 7.4). Total extracts, cell wall (20–40 µg), and membrane fractions were loaded on SDS-PAGE, transferred onto a nitrocellulose membrane, and analyzed by Western blotting using rabbit anti-pneumococcal RrgC antibodies (1/5000 dilution), horseradish peroxidase-conjugated anti-rabbit antibodies (1/5000 dilution; Sigma-Aldrich), and an ECL solution (Pierce) as detection reagents.

Generation of RrgC Mutants—The RrgC-expressing clone employed in Ref. 23 was used for mutagenesis experiments. Mutants were designed based on a surface entropy reduction (SER) method through the employment of the SER-prediction (SERp) server which suggested the modification of Glu-179, Lys-180, and Glu-181 to Ala, and were generated by using the QuikChange mutagenesis kit (Stratagene). Oligonucleotides used to introduce the mutation were: forward, 5’-GTA GCA AGA GAT GTT TCT GCG GCA GCG GTT CCC TTG ATT GGA GAA TAC and reverse, 5’-GTA TTC TCC AAT CAA GGG AAC CGC TGC CGC AGA AAC ATC TCT TGC TAC TGA TAC. The insertion of the required mutations in the resulting plasmid (mutpLIM01-RrgC) was verified by sequencing.

Overexpression and Purification of RrgC—Escherichia coli BL21(DE3) cells (Novagen) were transformed with the mutant pLIM01-RrgC plasmid and grown in LB medium supplemented with ampicillin (100 µg/ml) at 37 °C to an absorbance (A600) of 0.6 absorbance units. Protein expression was triggered by the addition of 0.5 mM isopropyl-β-D-thiogalactopyranoside (Inalco), and cells were incubated for 4 h at 25 °C, with shaking. Cultures were centrifuged and resuspended in buffer A (30 mM Tris-HCl (pH 8.0), 150 mM NaCl) and then disrupted by sonication. The lysate was centrifuged to remove cell debris (18,000 × g for 25 min) and loaded onto a 5-ml Ni2+TM column (GE Healthcare). After extensive washing using buffer A supplemented with 20 mM imidazole, RrgC was
Crystal Structure of the Pilus-association Protein RrgC

TABLE 1

Data collection and refinement statistics

| Dataset               | Native          | Selenomethionine |
|-----------------------|-----------------|------------------|
| **Data collection**   |                 |                  |
| Wavelength (Å)        | 1.0045          | 0.97923          |
| Space group           | P2₁             | P2₁              |
| **Cell dimensions**   |                 |                  |
| a, b, c (Å)           | a = 43.58, b = 65.76, c = 60.42 | a = 46.340, b = 65.85, c = 61.16 |
| Angles (°)            | 90, 110.97, 90  | 90, 112.47, 90   |
| Resolution (Å)        | 42.82–1.85 (1.95–1.85) | 65.85–2.10 (2.21–2.10) |
| Rmerge or Rsym        | 0.149 (0.502)   | 0.114 (0.581)    |
| <T> | 7.1 (3.0)       | 11.6 (3.8)       |
| Completeness (%)      | 97.1 (96.1)     | 98.0 (97.6)      |
| Anomalous completeness (%) |             | 95.7 (95.7)      |
| Redundancy            | 6.9 (7.1)       | 11.0 (10.5)      |

**Refinement statistics**

|                  | Native          | Selenomethionine |
|------------------|-----------------|------------------|
| No. reflections  | 26,430 (2,595)  |                  |
| Rmerge/Rwork     | 18.82/23.69     |                  |
| No. atoms        | 2,884           |                  |
| Protein          | 2,647           |                  |
| Water molecules  | 237             |                  |
| r.m.s.d.         |                 |                  |
| Bond lengths (Å) | 0.007           |                  |
| Bond angles (°)  | 1.17            |                  |
| Ramachandran plot (%) | 98.6          | 1.4              |
| Favored          |                 |                  |
| Generously allowed |               |                  |

Native data were collected (without additional cryoprotection) at beamline ID14-4 of the European Synchrotron Radiation Facility (ESRF) in Grenoble, France. Crystals belonged to space group P2₁ and had one molecule in the asymmetric unit, with Vₘ = 2.05 Å³/Da, corresponding to an approximate solvent content of 40.15%. Crystals of the Se-Met derivative were grown in the same condition, and a single wavelength anomalous dispersion (SAD) experiment was performed on the selenium edge on beamline ID29 of the ESRF (2.1 Å resolution, P2₁). All datasets were indexed and integrated with Mosflm (31) and merged and scaled with Scala (32), available in the CCP4 crystallographic package (33). The structure was solved using the SAD dataset; six selenium positions were identified by Autosol, and further model building, including phase extension to the higher resolution native dataset (1.85 Å), was performed with Autobuild (34). Model building was further extended with Buccaneer (35) and manual rebuilding with Coot (36). Refinement was carried out using the Phenix package (34) and Refmac (37). Solvent molecules were added with the automated procedure in Phenix. Geometrical parameters of the model were verified by Procheck (38). Data collection and refinement statistics are presented in Table 1.

**RESULTS**

Expression, Crystallization, and Structure Solution—RrgC is a 393-residue protein that comprises a 21-residue signal peptide followed by three independent domains, a sortase recognition sequence (VPDTG) that deviates from the classical LPXTG motif, and a 25-amino acid C-terminal trans-membrane anchoring domain (Fig. 1). The predicted mature form of RrgC thus comprises residues 22–368, and a construct including this region was successfully cloned, expressed, and purified (see “Experimental Procedures”). However, extensive efforts to crystallize this form of RrgC were unsuccessful. We thus performed a surface entropy reduction (SER) analysis using the SERp server (39) to search for potential surface-exposed resi-
dyes that could be mutated to promote local stability. The top result from the server suggested the mutation of three residues, Glu-179, Lys-180, Glu-181 into alanines. This construct (which will be referred to simply as RrgC) was expressed and purified as the wild-type protein and was successfully crystallized in a form suitable for high resolution structural studies. RrgC crystallized in space group P2_1 with one molecule per asymmetric unit, and the structure was solved to a resolution of 1.85 Å by performing a SAD experiment on a selenomethionylated crystal at the ESRF in Grenoble. The excellent quality of the electron density map allowed for the clear tracing of the entire chain, with the exception of four N-terminal and nine C-terminal residues, as well as those corresponding to loop 344–346. It is of note that analysis of the crystal structure revealed that the mutated region lies within the β10/β11 loop, which contacts the β19/β20 and β24/β25 loops of neighboring symmetry mates within the crystal lattice, allowing for the formation of more stable crystal contacts and an improvement in the quality of the dataset.

RrgC Is a Three-domain Pilin—RrgC folds into a three domains, namely D1 (residues 25–147), D2 (residues 148–255), and D3 (residues 256–359; Fig. 1, lower panel). Its overall shape resembles a bent, narrow rod, and it displays approximate overall dimensions of 110 × 40 × 25 Å. The cores of all three domains display β-barrel folds; their comparison with the Multiprot server (40) reveals an r.m.s.d. value of 2 Å, indicating that the general fold of RrgC is in fact a repetition of a core domain to which minor secondary structure features were added. In D1, one major and three minor α-helices are exposed outside of the β-barrel, with helix α1 being positioned toward domain D2. An analysis of D1 using DALI (41) reveals that its fold is comparable with the N-terminal domains of the fimbrial protein FimP from Actinomyces oris (42) and GBS52 from S. agalactiae, with r.m.s.d. values in the range of 3.5 Å for all Cα. Notably, the distinctions among the three proteins lie primarily in the number of surrounding helices and linker positions, with the major difference encompassing the loop-helix-loop structure between β1 and β2, which, in all three structures, extend toward the adjacent domain, generating a completely distinct surface. Interestingly, despite the fact that GBS52 is a minor pilin protein, FimP forms the pilus backbone in A. oris and thus plays the role of major pilin, indicating that the same fold is used for multiple functions in pilus biogenesis.

D2 and D3 both fold into anti-parallel β-sheets (blue and brown in Fig. 1); the r.m.s.d. between the two domains is of 1.5 Å for all Cα atoms, indicating their high structural similarity. The resulting β-sandwiches are highly reminiscent of the B domain of the collagen-binding adhesin Cna (CnaB) from S. aureus (43). It is of note that D2 also displays structural similarities to the C-terminal domain of GBS52 (r.m.s.d. 1.6 Å). In DALI, both D2 and D3 demonstrate good fits to domains 1 and 3 of BcpA, the major pilin of Bacillus cereus (44) with r.m.s.d. values of 1.8 Å and 1.2 Å, respectively.

Topology comparisons between CnaB repeat subdomains and IgG-like domains, which participate in protein-protein interactions and have been shown to be involved in bacterial adhesion processes (28), have indicated that despite sharing the same overall fold, their respective β-strands display an inversed order. This led to the coining of the term “IgG-reverse,” or IgG-rev fold (43). The superposition of either D2 or D3 domains of RrgC onto the structure of an IgG constant domain (45) indicated that their β-strands map tail-to-head rather than head-to-tail, an arrangement that is similar to the CnaB fold.
described above. Thus, RrgC D2 and D3 domains also display the IgG-rev fold, which is also present in the minor basal pilin GBSS2 protein of *S. agalactiae* (28); in Fig. 2, these arrangements are depicted using the nomenclature adopted for the description of the GBSS2 structure. Despite the similarities in topology, it is of note that the F strand in both D2 and D3 of RrgC is displaced to the opposite sheet, resulting in a β-strand order of DAG and CBEF instead of DAGF and CBE as observed in the CnaB fold.

Despite the fact that the three domains are interconnected through potentially flexible linkers, a surface diagram (Fig. 3A) is suggestive of a more fixed structure that could play the role of “pedestal” or “connector.” The exposure of large charged patches, and most notably acidic regions on D1 that could be relevant for interactions with positively charged patches on partner molecule RrgB, is compatible with the bridging role played by RrgC between the peptidoglycan and the rest of the pilus structure. It is also of note that the three domains are stabilized with each other through direct salt bridges, suggesting that RrgC may have limited flexibility once bound to RrgB and/or the bacterial cell wall.

**Domains D2 and D3 Are Stabilized by Isopeptide Bonds**—Isopeptide amide bonds formed within pilin monomers have been shown to stabilize Gram-positive pili (46, 47). These bonds are autocatalytically generated; many examples exist of isopeptide bonds between the side chains of juxtaposed lysine and asparagine residues, with a resulting link that is stabilized by an acidic residue present in the vicinity (46), but an example of a Thr-
Gln bond has also been recently highlighted (48). The crystal structure of RrgC reveals the presence of two intramolecular isopeptide bonds, in domains D2 and D3 (Figs. 1, lower panel, and 3B). The isopeptide bond within domain D2 is formed between the side chains of Lys-155 and Asn-252 and is stabilized by Glu-222, whereas in domain D3 it is formed between Lys-264 and Asn-354 and is stabilized by Glu-322. Notably, in both domains, the isopeptide bonds connect the first and last β-strand in the β-barrel (which are located side-by-side due to the unusual IgG fold described above), in analogous fashion to a belt buckle, indicating that in RrgC these bonds assist in the stability of the core β-barrel. It is of note that these bonds had been previously shown to confer thermal stability and resistance to proteolysis to RrgC (22). This arrangement is distinct from that seen for RrgA, where isopeptide bonds stabilize the different halves of a β-sandwich (26). As in the case for other pilus-forming proteins, however, the Lys-Asn isopeptide bonds present in RrgC are located within hydrophobic cores and include Leu-163, Val-231, and Val-250 for D2, and Tyr-329, Leu-332, Val-352 for D3. Notably, isopeptide bonds are present in all pilin molecules whose structures have been solved to date and have been shown to play important roles in structure stabilization for major or minor pilin proteins (46) (Fig. 4).

Incorporation of RrgC into the Pneumococcal Pilus—Major pilins such as RrgB, which are responsible for the formation of the backbone of the pilus structure, carry a conserved region called a “pilin motif” (WXXXVXVPK); upon polymerization of the pilus fiber, the ε-amino moiety of the Lys side chain forms an isopeptide bond with the Thr of the LPXTG of a neighboring major pilin molecule, thus allowing formation of the chain. Minor pilins such as RrgC, however, do not display such sequences. These molecules, and most notably the base pilin, use an exposed Lys side chain to link to other pilus-related molecules (16, 49). This is the case for minor pilins SpaB from C. diphtheriae, FctB from a group A Streptococcus strain, and GBS52 from S. agalactiae, which become covalently incorporated into the pilus through exposed Lys residues located in flexible loop regions (28, 49–51). Despite the fact that the nucleophilic Lys residue is yet to be identified in RrgC, the comparison of its structure with that of GBS52 points clearly to a candidate residue. In GBS52, Lys-148, located in the flexible linker region between D1 and D2, lies within a pilin-like motif (IYPK) and has been suggested as being responsible for linking GBS52 to partner molecules (49). Likewise, Lys-142 of RrgC is located at precisely the same position and points into solvent,
much like its GBSS2 counterpart. Thus, it is conceivable that in RrgC, the linker region between D1 and D2 is involved in recognition of RrgB. The fact that in RrgC D1 is rotated from the main axis of the molecules suggests that this region may be exposed within the pilus-forming arrangement, which would facilitate binding of the LPXTG region of RrgB.

**RrgC Does Not Require a Pilus-related Sortase for Surface Attachment**—To gain further insight into the role played by RrgC in the pilus formation process, we set out to investigate the precise location of RrgC in the pneumococcus. We thus expressed full-length RrgC (harboring the signal peptide) in a GFP-expressing vector in the nonpiliated, noninfectious *S. pneumoniae* R6 strain and analyzed bacteria by immunofluorescence. Fig. 5a reveals that RrgC was homogeneously localized on the bacterial membrane, a result that is consistent with its predicted membrane-association topology.

We then set out to investigate the transpeptidation roles of SrtC-1 and SrtC-3, which had been previously linked to RrgC functionality within the pilus (19, 23, 52). To do so, we expressed different combinations of pilus-related proteins in *S. pneumoniae* R6 and R6ΔsrtA strains and tested for the presence of RrgC in total, cell wall, and membrane extracts (Fig. 5, B and C). Fig. 5B displays results of an experiment where the two R6 strains were transformed with plasmids overexpressing RrgC and SrtC-1. A Western blot developed with anti-RrgC antibodies shows that RrgC is present in all fractions. Analysis of the same samples from the *S. pneumoniae* R6ΔsrtA, however, shows that there is considerably less RrgC bound to the cell wall in the absence of SrtA, suggesting that this enzyme could play an important role in association of RrgC with the peptidoglycan. The thin band corresponding to RrgC potentially indicates background activity from SrtC-1, which was also overexpressed. This result becomes even clearer upon analysis of Fig. 5C, which depicts the result from an experiment where *S. pneumoniae* R6 and R6ΔsrtA strains were transformed with plasmids overexpressing RrgC and SrtC-3. In this case, despite the fact that in the wild-type strain RrgC was present in all fractions as expected, in R6ΔsrtA very little or no RrgC could be detected in the cell wall fraction. This not only discards the possibility that SrtC-3 could play an important role in RrgC association with the cell wall, but also underlines the importance of SrtA for this function. These experiments thus indicate that RrgC is covalently associated with the pneumococcal cell wall through the action of the housekeeping sortase, SrtA, with some background activity from SrtC-1. Despite the fact that these results diverge from those of LeMieux et al., who employed in vivo mutagenesis of pilus-related sortases and suggested that SrtA is dispensable for pilus cell wall attachment in *S. pneumoniae* (19), housekeeping sortases have been linked to the pilus association process in a number of organisms, such as *B. cereus*, *C. diphtheriae*, and *S. agalactiae* (53–55); thus, the results presented here are in agreement with those described for other Gram-positive species.

**FIGURE 6. Model of the pneumococcal pilus.** A–C, structures of the minor pilin RrgA (A), major pilin RrgB (B), and minor pilin RrgC (C). D, model of the complete pilus, including the three structural components. RrgC serves as a pedestal or anchor that associates the pilus to the cell wall through the action of SrtA and signals to the cell that pilus formation is complete. Note that the involvement of Lys-142 in RrgC association with RrgB is merely suggestive.
DISCUSSION

Adhesion to host cells is prerequisite for successful colonization by many pathogenic bacteria, and the host–bacterial attachment can be mediated by cell surface pili in the case of both Gram-negative and Gram-positive pathogens (15, 56). Pili from Gram-negative bacteria have been extensively studied; Type 1 and Type P pili, for example, have been shown to be formed through noncovalent interactions between adjacent subunits whose association depends on a complex β-strand swapping mechanisms (56). In addition, structural data for a number of pilins from Gram-negative organisms are available. One elegant study involves the cryo-EM structure of the Vibrio cholerae toxin co-regulated Type IV pilus associated with cryo-EM. The assembled pilus packs centralizes α-helices stabilized by hydrophobic interactions. The globular domain, located on the outside of the pilus, provides surface variation for the multiple functions of the virulence factor (57, 58). Despite the extensive effort that has been dedicated to the study of Gram-positive pili, structural information at this level of detail is still not available.

Pili from Gram-positive organisms are polymers formed by covalently associated subunits that are generally encoded on the same pathogenicity islet as their cognate sortases. In S. pneumoniae, the pilus backbone is formed by covalently associated RrgB units; microscopy efforts have indicated that the molecules are associated with a directionality, in which a nose-like protrusion defines the polarity of the fiber (25). The crystal structure of RrgB, fitted into a cryo-EM map of the pilus, confirms this polarity (24, 59). Notably, RrgA and RrgC were suggested as being located at the extremities of the fiber, with RrgC at the base and RrgA at the free end (25). The structure of RrgA revealed a four-domain flexible molecule carrying a von Willebrand factor (VWF) domain capable of recognizing extracellular matrix components (26), whereas RrgB displayed a three-domain fold (24, 59). Until recently, however, the structure of the third structural component of the pneumococcal pilus, RrgC, was still lacking, impeding a more complete understanding of the architecture of the fiber from S. pneumoniae.

Here, we report that RrgC folds into three independent domains linked by potentially flexible regions; it is smaller than both RrgA and RrgB. The contacts present between the domains could somewhat restrict interdomain flexibility, which is in line with that observed for RrgB, whose four domains also display limited structural malleability (24). Two of RrgC domains are IgG-like; their potential function could be to enhance the binding capacity of RrgC (and hence the pilus) to the cell wall or to eukaryotic targets. The minor pilin from S. agalactiae (GB52), for example, which is composed of two domains, employs the IgG-like N2 domain to bind to lung epithelial cells, an essential step in the infection process (28). These observations are in agreement with the fact that bacteria mimic the host immune system by incorporating eukaryotic immunoglobulin-like subdomain variants into their structures, to not only bind to partner molecules but also escape the host immune system elements. In addition, despite the fact that basal pilin proteins often harbor proline-rich C-terminal tails upstream from the sorting motif, which form PPII-type helices that assist with protein–protein interactions (60), RrgC only harbors two Pro residues in the vicinity of the VDPTG sorting motif (Pro-336, Pro-359). It is thus unlikely that these residues play key roles in interactions with chaperones or other protein complexes involved in pilus assembly, and it is more probable that RrgC contacts the pilus-forming machinery through its LPXTG-like motif and IgG-like domains.

Isopeptide bonds have now been identified in a number of pilin subunits from Gram-positive bacteria, including RrgA and RrgB, which display such bonds in almost all domains (24, 26, 59). In RrgC, the isopeptide bonds of domains D2 and D3 had been identified through sequence comparisons and mutagenesis studies; the latter indicated that both bonds play a key role in protein stability (22). Notably, in both RrgB and RrgC, the stabilization of the fold by intramolecular isopeptide bonds plays a significant role in the efficient recognition of the pilin subunits by their cognate sortases (22) underlining the key role played by these bonds for stability of the architecture of the entire pilus. Thus, the structuring of pilus-forming proteins as “beads on a string,” with each domain stabilized by a covalent isopeptide bond, is a common signature in extracellular Gram-positive fibers.

Thus, the RrgC structure reported here is the “missing link” that allows the construction of a model of the pneumococcal pilus in atomic detail. RrgA, the four-domain adhesin that binds elements of the eukaryotic extracellular matrix, is covalently associated with the backbone fiber through its YPRTG motif, exposing its von Willebrand factor region as far as possible from the bacterial surface to promote adhesion (26) (Fig. 6, cyan). The backbone is formed by the covalent association of RrgB monomers (Fig. 6, bright green), and its nucleophilic Lys-183 residue thus has a dual recognition potential, participating in the covalent association with RrgA–YPRTG and to the IPQTG motif of other RrgB molecules to generate the fiber with the adhesin at its tip (21, 24, 25, 59). Finally, RrgC (Fig. 6, blue) serves as a pedestal molecule that links the pilus directly to the peptidoglycan through the action of SrtA, and in doing so, “signals” that pilus synthesis is complete. Further studies aimed at identifying the precise signal that regulates pilus length and the timing of anchoring will be required to determine the mechanisms of pilus fiber assembly.

Acknowledgments—We thank J. Marquez and the HTX Laboratory team for access to and help with high throughput crystallization and the European Synchrotron Radiation Facility (ESRF) for access to beamlines.

REFERENCES
1. Cartwright, K. (2002) Pneumococcal disease in Western Europe: burden of disease, antibiotic resistance and management. Eur. J. Pediatr. 161, 188–195
2. Bronzwaer, S. L., Cars, O., Buchholz, U., Mölstad, S., Goetttsch, W., Veldhuizen, I. K., Kool, J. L., Sprenger, M. J., and Degener, J. E., and European Antimicrobial Resistance Surveillance System (2002) A European study on the relationship between antimicrobial use and antimicrobial resistance. Emerg. Infect. Dis. 8, 278–282
3. Greenwood, B. (1999) The epidemiology of pneumococcal infection in...
Crystal Structure of the Pilus-association Protein RrgC

children in the developing world. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **354**, 777–785

4. Levine, O. S., O’Brien, K. L., Knoll, M., Adegbola, R. A., Black, S., Cherian, T., Dagan, R., Goldblatt, D., Grange, A., Greenwood, B. M., Hennessy, T., Klugman, K. P., Madhi, S. A., Mulholland, K., Nohynk, H., Santosh, M., Saha, S. K., Socott, J. A., Sow, S., Whitney, C. G., and Cutts, F. (2006) Pneumococcal vaccination in developing countries. *Lancet* **367**, 1880–1882

5. AlonsoDeVelasco, E., Verheul, A. F., Verhoef, J., and Snippe, H. (1995) *Streptococcus pneumoniae*: virulence factors, pathogenesis, and vaccines. *Microbiol. Rev.* **59**, 591–603

6. Abbot, E. L., Smith, W. D., Siou, G. P., Chiriboga, C., Smith, R. J., Wilson, J. A., Hirsh, B. H., and Kehoe, M. A. (2007) Pili mediate specific adhesion of *Streptococcus pyogenes* to human tonsil and skin. *Cell. Microbiol.* **9**, 1822–1833

7. Barocchi, M. A., Ries, J., Zago, X., Hemsley, C., Albiger, B., Kanth, A., Dahlberg, S., Fernerbo, J., Moschioni, M., Massignani, V., Hultenby, K., Taddei, A. R., Beiter, K., Wartha, F., von Euler, A., Covacci, A., Holden, D. W., Normark, S., Rappuoli, R., and Henriques-Normark, B. (2006) A pneumococcal pili influences virulence and host inflammatory responses. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 2857–2862

8. Maisey, H. C., Hensler, M., Nizet, V., and Doran, K. S. (2007) Group B streptococcal pilus proteins contribute to adherence to and invasion of brain microvascular endothelial cells. *J. Bacteriol.* **189**, 1464–1467

9. Nelson, A. L., Ries, J., Bagnoli, F., Dahlberg, S., Fälker, S., Rounioja, S., Hultenby, K., Hilleringmann, M., Holden, D. W., Rappuoli, R., Normark, S., Barocchi, M. A., and Henriques-Normark, B. (2007) RrgA is a pilus-associated adhesin in *Streptococcus pneumoniae*. *Mol. Microbiol.* **66**, 329–340

10. Mandlik, A., Swierczynski, A., Das, A., and Ton-That, H. (2007) *Corynebacterium diphtheriae* employs specific minor pilins to target human pharyngeal epithelial cells. *Microb. Mol. Biol. Rev.* **61**, 111–124

11. Konto-Ghiorghi, Y., Mairey, E., Mallet, A., Duménil, G., Caliot, E., Trieu-Cuot, P., and Drancourt, M. (2009) Dual role for pili in adherence to epithelial cells and biofilm formation in *Streptococcus agalactiae*. *PLoS Pathog.* **5**, e1000422

12. Gianfaldoni, C., Censini, S., Hilleringmann, M., Moschioni, M., Facciotti, C., Paneagrau, W., Massignani, V., Covacci, A., Rappuoli, R., Barocchi, M. A., and Ruggiero, P. (2007) RrgC of *Streptococcus pneumoniae* subunits protect mice against lethal challenge. *Infect. Immun.* **75**, 1059–1062

13. Mazzanikian, S. K., Liu, G., Ton-That, H., and Schneewind, O. (1999) *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science* **285**, 760–763

14. Dramsi, S., Trieu-Cuot, P., and Bierne, H. (2005) Sorting sortases: a no-
the repetitive B region of the *Staphylococcus aureus* collagen-binding surface protein. *Structure* **8**, 67–78
44. Budzik, J. M., Poor, C. B., Faull, K. F., Whitelegge, J. P., He, C., and Schneewind, O. (2009) Intramolecular amide bonds stabilize pilus on the surface of bacilli. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 19992–19997
45. Prabakaran, P., Yu, B. K., Gan, J., Feng, Y., Dimitrov, D. S., and Ji, X. (2008) Structure of an isolated unglycosylated antibody CH2 domain. *Acta Crystallogr. D Biol. Crystallogr.* **64**, 1062–1067
46. Kang, H. J., and Baker, E. N. (2012) Structure and assembly of Gram-positive bacterial pili: unique covalent polymers. *Curr. Opin. Struct. Biol.* **22**, 200–207
47. Kang, H. J., Coulibaly, F., Clow, F., Proft, T., and Baker, E. N. (2007) Stabilizing isopeptide bonds revealed in Gram-positive bacterial pilus structure. *Science* **318**, 1625–1628
48. Kwon, H., Squire, C. J., Young, P. G., and Baker, E. N. (2014) Autocatalytically generated Thr–Gln ester bond cross-links stabilize the repetitive Ig-domain shaft of a bacterial cell surface adhesin. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1367–1372
49. Vengadesan, K., and Narayana, S. V. (2011) Structural biology of Gram-positive bacterial adhesins. *Protein Sci.* **20**, 759–772
50. Mandlik, A., Das, A., and Ton-That, H. (2008) The molecular switch that activates the cell wall anchoring step of pilus assembly in Gram-positive bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 14147–14152
51. Linke, C., Young, P. G., Kang, H. J., Bunker, R. D., Middleton, M. J., Caradoc-Davies, T. T., Proft, T., and Baker, E. N. (2010) Crystal structure of the minor pilin FctB reveals determinants of group A streptococcal pilus anchoring. *J. Biol. Chem.* **285**, 20381–20389
52. Neiers, F., Madhurantakam, C., Fälker, S., Manzano, C., Dessen, A., Normark, S., Henriques-Normark, B., and Achour, A. (2009) Two crystal structures of pneumococcal pilus sortase C provide novel insights into catalysis and substrate specificity. *J. Mol. Biol.* **393**, 704–716
53. Budzik, J. M., Marraffini, L. A., and Schneewind, O. (2007) Assembly of pili on the surface of *Bacillus cereus* vegetative cells. *Mol. Microbiol.* **66**, 495–510
54. Swaminathan, A., Mandlik, A., Świerczynski, A., Gaspar, A., Das, A., and Ton-That, H. (2007) Housekeeping sortase facilitates the cell wall anchoring of pilus polymers in *Corynebacterium diphtheriae*. *Mol. Microbiol.* **66**, 961–974
55. Nobbs, A. H., Rosini, R., Rinaudo, C. D., Maione, D., Grandi, G., and Telford, J. L. (2008) Sortase A utilizes an ancillary protein anchor for efficient cell wall anchoring of pili in *Streptococcus agalactiae*. *Infect. Immun.* **76**, 3550–3560
56. Allen, W. I., Phan, G., and Waksman, G. (2012) Pilus biogenesis a the outer membrane of Gram-negative bacterial pathogens. *Curr. Opin. Struct. Biol.* **22**, 500–506
57. Craig, L., Taylor, R. K., Pique, M. E., Adair, B. D., Arvai, A. S., Singh, M., Lloyd, S. I., Shin, D. S., Getzoff, E. D., Yeager, M., Forest, K. T., and Tainer, J. A. (2003) Type IV pilin structure and assembly: x-ray and EM analyses of *Vibrio cholerae* toxin-coregulated pilus and *Pseudomonas aeruginosa* PAK pilin. *Mol. Cell* **11**, 1139–1150
58. Lim, M. S., Ng, D., Zong, Z., Arvai, A. S., Taylor, R. K., Tainer, J. A., and Craig, L. (2010) *Vibrio cholerae* El Tor TcpA crystal structure and mechanism for pilus-mediated microcolony formation. *Mol. Microbiol.* **77**, 755–770
59. Paterson, N. G., and Baker, E. N. (2011) Structure of the full-length major pilin from *Streptococcus pneumoniae*: implications for isopeptide bond formation in Gram-positive bacterial pilus. *PLoS One* **6**, e22095
60. Kay, B. K., Williamson, M. P., and Sudol, M. (2000) The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* **14**, 231–241