

# New insights into the role of the glutamic acid of the E-box motif in group B *Streptococcus* pilus 2a assembly

Roberta Cozzi,\* Annalisa Nuccitelli,\* Mariapina D'Onofrio,<sup>†</sup> Francesca Necchi,\* Roberto Rosini,\* Francesca Zerbini,\* Massimiliano Biagini,\* Nathalie Norais,\* Christian Beier,<sup>‡</sup> John L. Telford,\* Guido Grandi,\* Michael Assfalg,<sup>†</sup> Martin Zacharias,<sup>‡</sup> Domenico Maione,\* and C. Daniela Rinaudo\*<sup>1</sup>

\*Novartis Vaccines and Diagnostics, Siena, Italy; <sup>†</sup>Nuclear Magnetic Resonance Laboratory, Department of Biotechnology, University of Verona, Verona, Italy; and <sup>‡</sup>Physics Department, Technical University Munich, Munich, Germany

**ABSTRACT** Group B *Streptococcus* pili are covalently linked structures assembled *via* a sortase-catalyzed transpeptidation mechanism involving specific residues and motifs. A sequence element containing a conserved glutamic acid, called the E-box, has been described to be involved in pilus formation. Although it is known that the glutamic acid is involved in stabilizing the internal isopeptide bonds, its role in pilus assembly still needs to be investigated. Using site-specific mutagenesis and complementation studies of knockout strains, we found that the E-box glutamic residue of the backbone and the major ancillary proteins is essential for pilus protein polymerization. NMR analysis revealed that the mutation of this residue seriously affected the folding of the protein. By contrast, the mutation of the lysine involved in the same isopeptide bond did not engender a structural destabilization, and the native fold was preserved. Moreover, molecular dynamics simulations on the E-box-containing domain of the backbone protein showed that the E-box glutamic acid is necessary to maintain the appropriate dryness of the domain core and that its mutation favors an unfolded state. The data provide the first direct evidence that the E-box has an additional and key role in maintaining the correct protein fold independently of isopeptide bond formation.—Cozzi, R., Nuccitelli, A., D'Onofrio, M., Necchi, F., Rosini, R., Zerbini, F., Biagini, M., Norais, N., Beier, C., Telford, J. L., Grandi, G., Assfalg, M., Zacharias, M., Maione, D., Rinaudo, C. D. New insights into the role of the glutamic acid of the E-box motif in

group B *Streptococcus* pilus 2a assembly. *FASEB J.* 26, 2008–2018 (2012). [www.fasebj.org](http://www.fasebj.org)

*Key Words:* isopeptide bonds • molecular dynamic simulations • NMR spectroscopy • backbone protein

THE GRAM-POSITIVE HUMAN pathogen *Streptococcus agalactiae* (also known as group B *Streptococcus* or GBS) is the leading cause of neonatal bacterial sepsis, pneumonia, and meningitis (1). A wide range of surface-exposed proteins have been shown to be involved in GBS virulence and pathogenesis, including cell surface pili (2). These long filamentous structures have aroused great interest also because of their importance as vaccine candidates (3, 4). Gram-positive pili are covalently linked structures assembled *via* a transpeptidation mechanism by specific class C sortases (5). These enzymes recognize specific sequence motifs in the pilin subunits and elongate the pilus shaft by progressive addition of subunits joined by intermolecular isopeptide bonds between a specific lysine residue of one subunit and the conserved threonine residue of the LPxTG sorting motif of the next subunit. Finally, the assembled pilus is anchored to the cell wall peptidoglycan by the housekeeping gene sortase A (6–9). The genes encoding pilin subunits are located within pathogenicity islands, which are known as pilus islands (PIs).

In GBS, 3 different genomic islands have been identified, named PI-1, PI-2a, and PI-2b (10). Each PI carries genes coding for 3 structural proteins containing the conserved C-terminal LPxTG motif typical of the associated cell wall-anchored proteins, and 2 other

Abbreviations: AP, ancillary protein; BP, backbone protein; Cna domain, collagen adhesion domain; CWSS, cell-wall sorting signal; D, domain; GBS, group B *Streptococcus*; HMW, high molecular weight; HSQC, heteronuclear single quantum coherence; IgG, immunoglobulin G; KO, knockout; MALDI-TOF, matrix-assisted laser desorption/ionization-time of flight; MD, molecular dynamics; MS, mass spectrometry; PDB, protein data bank; PI, pilus island; PIPE, polymerase incomplete primer extension; RMSD, root mean square deviation; RMSF, root mean square fluctuation; SrtC, sortase C; WT, wild type.

<sup>1</sup> Correspondence: Novartis Vaccines and Diagnostics, via Fiorentina, 1 53100 Siena, Italy. E-mail: [daniela.rinaudo@novartis.com](mailto:daniela.rinaudo@novartis.com)

doi: 10.1096/fj.11-196378

This article includes supplemental data. Please visit <http://www.fasebj.org> to obtain this information.

genes coding for sortase C enzymes (SrtC1 and SrtC2). The major pilin protein, known also as the backbone protein (BP), is distributed uniformly along the pilus structure; the ancillary proteins (APs), AP1 and AP2, are currently understood to be located at the pilus tip (AP1) and to be involved in pilus anchoring to the cell wall (AP2) (8, 9). BP and sortase C are necessary for pilus formation, whereas AP1 and AP2 are dispensable (10). Despite low sequence similarities, the pilin subunits of gram-positive bacteria show very similar tertiary structures comprising immunoglobulin G (IgG)-like domains of shared evolutionary origin. Each pilin subunit is stabilized by intramolecular isopeptide bonds, and all contain sequence elements and/or residues that are essential for pilus assembly, and which are conserved among pilin subunits in different bacteria (11). Such motifs include the pilin motif (consensus Wxxx-VxVYPK) that contain a specific lysine residue (K or asparagine), the side chain of which can participate in sortase-catalyzed amide bond formation by reaction with the C terminus of the next subunit molecule during polymerization (12, 13). In addition, pilin subunits have a cell-wall sorting signal (CWSS) containing the sortase recognition site LPxTG motif, and finally the E-box motif (consensus YxLxETxAPxGY), as assigned for the first time in the major pilin subunit SpaA of *Corynebacterium diphtheriae* (12) and subsequently in other bacterial pilins (14). The E-box contains a conserved glutamic acid residue, which in *C. diphtheriae* SpaA (Glu-446) has been demonstrated to be essential for the incorporation of the minor pilins SpaB and SpaC (12). Intriguingly, in SpaA, this glutamate is the catalytic residue that mediates the formation of the Lys-363–Asn-462 intramolecular isopeptide bond (15), similar to the role assigned to Glu-258 in *Streptococcus pyogenes* Spy0128, in which this residue was shown to be essential for the corresponding intramolecular reaction to occur (16). Moreover, several X-ray crystal structures of backbone pilins have shown that the E-box domain is involved in the formation of such isopeptide bonds and that these linkages confer higher stability to the monomeric subunit (16–19). Recently, we solved the X-ray crystal structure of the shaft-forming BP of *S. agalactiae* pilus 2a (BP-2a; ref. 20). The 3-dimensional structure revealed an IgG-like fold domain organization, comprising 4 structural units, designated D1–D4. The domains D2, D3, and D4 are each stabilized by an intramolecular Lys-Asn isopeptide bond, located in a largely hydrophobic pocket, comprising several aromatic residues, including a bond-catalyzing aspartyl or glutamyl residue (20). However, the role of intramolecular isopeptide bonds and of the E-box motif in pilus assembly still needs to be clarified.

In the present study, we explore the role of the intramolecular isopeptide bond and E-box motif in the polymerization of GBS pilus type 2a subunits, using a structure-based mutagenesis approach, coupled with *in vivo* complementation studies. We use NMR spectroscopy to study the contribution of E-box residues to the folding of BP-2a and AP1–2a and assess the role of these

residues in pilus protein polymerization. Finally, molecular dynamics simulations are used to understand in detail the effects of the burial of the glutamic acid within the core of domain D4 of the pilin subunits.

## MATERIALS AND METHODS

### Comparative homology modeling

All molecular simulations were performed using Discovery Studio 2.5 software from Accelrys (San Diego, CA, USA). The amino acid sequences of AP1–2a (TIGR annotation SAL\_1487) and AP2–2a (SAL\_1482) were used to search against the Protein Data Bank (PDB) with the Basic Local Alignment Search Tool (BLAST) program. Ten models have been generated, and the model that shared the least root mean square deviation (RMSD) with respect to trace (C $\alpha$  atoms) of the crystal structure of the template was selected for further refinements and validations. The quality of the refined structures obtained was checked with the Profile-3D module.

### Molecular dynamics (MD) simulations

All MD simulations and trajectory analysis were performed with the Gromacs 4.0.5 simulation package (21) using the AMBER99SB-ILDN force field (22) with explicit water (TIP3P; ref. 23). The wild-type (WT) crystal structure of the BP-2a domain D4 from *S. agalactiae* (20) served as starting structure for MD simulations at 310 K and a pressure of 1 bar. A covalent isopeptide bond (between K463 and N636) was added to the topology, according to the force-field parameters of an amide peptide bond. Mutants K463A and E589A were generated by truncating the side chains up to the C- $\beta$  atom. None of the mutations contained an isopeptide bond between K463 and N636. Missing solvent exposed flexible C-terminal residues (up to residue G648) were added to each structure and relaxed using MD simulations (in vacuum), keeping the rest of the structure fixed. Each of the 3 structures was subsequently placed in a cubic box (90 $\times$ 90 $\times$ 90 Å) containing  $\sim$ 13,000 water molecules. The distance between the protein and the edge of the box was set to 1.2 nm; an appropriate number of Cl $^-$  ions were placed in each system to keep electroneutrality. After stepwise heating of the systems to 310 K, production runs were performed for up to 50 ns with a time step of 2 fs in the NPT ensemble at 310 K and 1 bar. Temperature and pressure were controlled by Nosé-Hoover (coupling constant  $t_c = 2.524$ , 25) and Parrinello-Rahman ( $t_p = 5.0$  ps; refs. 26, 27) schemes, respectively.

The RMSD and RMS fluctuation (RMSF) measurements have been performed on the backbone of the proteins. The water molecule monitoring was performed by counting the number of solvent contacts within 4 Å of any atom of the residue phenylalanine 477 (F477). The F477 residue is located in the pocket close to the isopeptide bond. Using this setting, we considered all of the water molecules that moved from the bulk into the core of the domain. Figures of the molecular structures were generated with VMD (28) and Discovery Studio 3.0 (Accelrys).

### Bacterial strains, media, and growth conditions

*S. agalactiae* strains were grown at 37°C in 5% CO $_2$  in Todd Hewitt Broth (Difco Laboratories, Detroit, MI, USA) or trypticase soy agar supplemented with 5% sheep blood.

## Complementation vectors and site-directed mutagenesis

GBS-knockout (KO) mutant strains for each of the 3 structural proteins of pilus type 2a (515 $\Delta$ BP-2a, 515 $\Delta$ AP1-2a, and 515 $\Delta$ AP2-2a) and complementation vectors (pAM\_BP-2a, pAM\_API-2a, and pAM\_AP2-2a) were generated, as described previously (10). Site-directed mutagenesis was performed by the polymerase incomplete primer extension (PIPE) method (29), as previously reported (30). Mutations were verified by DNA sequencing. All complementation plasmids expressing mutated forms of the pilus subunits were electroporated into the KO strains. Complementation was confirmed by checking protein expression by Western blot analysis. For all complementation experiments, the empty pAM401/gbs80P+T plasmid, introduced into WT GBS strain 515 and mutant strains, was used as a negative control.

## Western blot analysis

GBS total cell extracts were carried out as described previously (10). Briefly, GBS strains were grown to midexponential phase. Cell pellets were resuspended in 50 mM Tris-HCl containing 400 U of mutanolysin (Sigma-Aldrich, St. Louis, MO, USA) and protease inhibitors. The mixtures were then incubated at 37°C for 1 h, and cells were lysed by 3 cycles of freeze-thawing. Cellular debris was removed by centrifugation, and protein concentration was determined using BCA protein assay (Pierce, Rockford, IL, USA). Total protein extracts (20  $\mu$ g) were resolved on 3–8% NuPAGE precast gels (Invitrogen, Carlsbad, CA, USA) by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose. Membranes were probed with mouse antiserum directed against BP, AP1, and AP2 pilus proteins (1:1000 dilution), followed by a rabbit anti-mouse horseradish peroxidase-conjugated secondary antibody (Dako, Glostrup, Denmark). Bands were then visualized using an Opti-4CN substrate kit (Bio-Rad, Hercules, CA, USA). Antisera specific for the recombinant proteins were produced by immunizing CD1 mice with the purified recombinant proteins, as previously reported (3).

## Cloning, expression, and purification of recombinant proteins

PCR fragments encoding domains D3 and D4 of BP-2a (SAL\_1486) were cloned into pENTR/TEV/D-TOPO vector (Invitrogen) and then subcloned into pET54 DEST or pET59 DEST vectors (Novagen, San Diego, CA, USA) by Gateway LR reaction to generate His- or TRX-His-tagged proteins. PCR fragment encoding domain D4 of AP1-2a and the full-length AP2-2a were cloned in the pET-15 TEV vector modified in house for PIPE cloning. The recombinant mutants BP-D3<sub>E416A</sub>, BP-D4<sub>E589A</sub>, BP-D4<sub>K463A</sub>, AP1-D4<sub>K472A</sub>, AP1-D4<sub>E801A</sub>, and AP2<sub>E216A</sub> were generated by site-directed mutagenesis using as template the WT genes and the KAPA Hi-FI polymerase (KAPA Biosystems, Cambridge, MA, USA). Proteins were expressed in *Escherichia coli* BL21 (DE3; Novagen) cells as His- or TRX-tagged fusion proteins and purified by affinity chromatography and gel filtration. The cells were grown in Luria-Bertani medium, or in M9 minimal medium [containing 1 mg/ml of (<sup>15</sup>NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>] for the expression of <sup>15</sup>N-labeled samples, at 37°C until OD<sub>600</sub> ~ 0.7, and then induced with 1 mM isopropyl- $\beta$ -D-thio-galactoside (IPTG) overnight at 25°C. The soluble proteins were extracted by sonication in 50 mM Tris-HCl (pH 7.5), 250 mM NaCl, 10 mM imidazole, lysozyme, and DNase, and they were purified by a FF-Crude His-Trap HP nickel-chelating column (Amersham Biosciences, Piscataway, NJ, USA). The recombinant proteins,

eluted with 300 mM imidazole, were concentrated by ultrafiltration to 10 mg/ml, and the buffer was exchanged using a PD-10 desalting column (Amersham Biosciences) equilibrated with TEV cleavage buffer [50 mM Tris-HCl (pH 8), 1 mM DTT, and 0.5 mM EDTA]. The His and TRX-His tags were cleaved with AcTEV protease (Invitrogen) and then removed by a subtractive immobilized metal ion affinity chromatography (IMAC) purification step. The proteins were concentrated and loaded onto HiLoad 26/60 Superdex 75 (Amersham Biosciences) preequilibrated in 50 mM Tris-HCl (pH 7.5), 100 mM NaCl. The protein abundance in pure fractions was quantified with the BCA assay (Pierce).

For NMR experiments, the protein's buffers were exchanged using a PD-10 desalting column (Amersham Biosciences) equilibrated with 50 mM phosphate buffer, pH = 6.5, and finally concentrated by ultrafiltration.

## NMR spectroscopy analysis

<sup>1</sup>H-<sup>15</sup>N heteronuclear single quantum coherence (HSQC) spectra were recorded at 25°C on a Bruker Avance III spectrometer (Bruker Daltonics, Bremen, Germany) operating at 600.13-MHz proton Larmor frequency, equipped with a triple-resonance TCI cryoprobe incorporating a z-axis gradient. A standard <sup>1</sup>H-<sup>15</sup>N HSQC pulse sequence was used, with pulsed field gradients for suppression of the solvent signal and spectral artifacts. All the <sup>1</sup>H-<sup>15</sup>N HSQC spectra on different protein samples were performed in the presence of 7% D<sub>2</sub>O for NMR spectrometer frequency lock, in the same conditions, using a spectral width of 9515.385 Hz and 2048 complex points in the <sup>1</sup>H dimension, and a spectral width of 2432.718 Hz and 256 complex points in the <sup>15</sup>N dimension. Processing of all the spectra has been obtained with Topspin 2.1 (Bruker Daltonics).

## In-gel digestion and matrix-assisted laser desorption/ionization-time of flight (MALDI TOF)/TOF mass spectrometric analysis

Coomassie blue G-250 stained bands were excised from the gel, destained and in-gel digested in 5 mM NH<sub>4</sub>HCO<sub>3</sub> with 12  $\mu$ g/ml of modified Lys-C protease (Roche Diagnostics, Mannheim, Germany) or trypsin (Promega, Madison, WI, USA) overnight at 37°C. Guanidination of the C-term lysine residues was performed using the ProteoMass Guanidination Kit (Sigma-Aldrich), following the manufacturer's instructions. Modified and unmodified peptide solutions were directly spotted onto a Prespotted AnchorChip MALDI target (Bruker Daltonics). Mass spectra were acquired with a MALDI-TOF/TOF mass spectrometer UltraFlex (Bruker Daltonics). Ions generated by laser desorption at 337 nm (N<sub>2</sub> laser) were recorded with an acceleration voltage of 25 kV in the reflector mode. In general, ~200 single spectra were accumulated for improving the signal/noise ratio and analyzed by FlexAnalysis 2.4 (Bruker Daltonics). Peptide identification was performed using BioTools and SequenceEditor 3.0 (Bruker Daltonics).

## RESULTS

### Intramolecular isopeptide bonds are dispensable for pilus protein polymerization

In the 3-dimensional structure of the shaft-forming BP-2a (20), the residues involved in isopeptide bond formation are Lys199-Asn325 for D2 domain, Lys355-

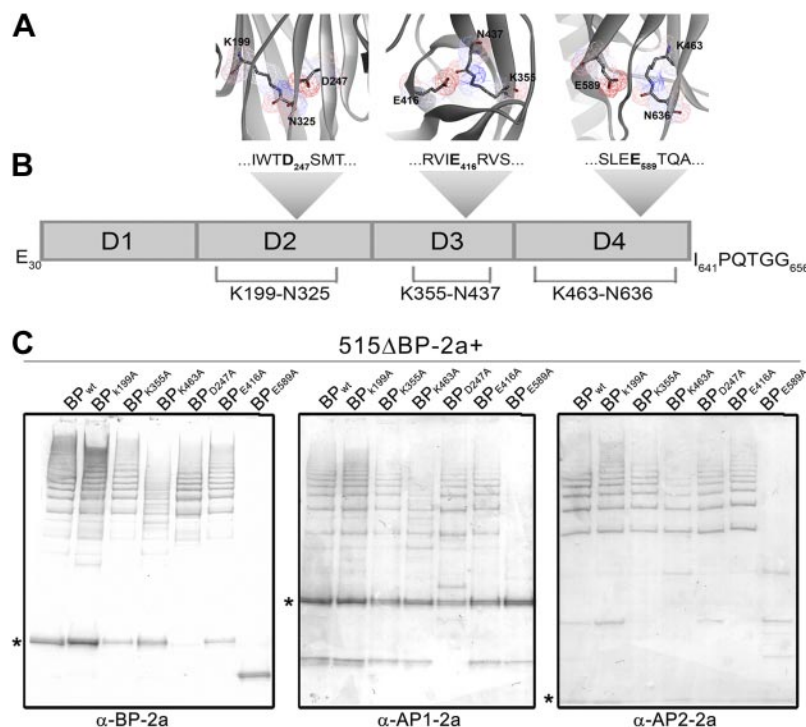
Asn437 for D3 domain, and Lys463-Asn636 for D4 domain. In addition, 3 essential residues (Asp247, Glu416, and Glu589), one for each isopeptide bond, have been identified as catalyzing and stabilizing amino acids (ref. 20 and **Fig. 1A, B**). To examine the contribution of these linkages in pilus 2a assembly, we mutated the Lys residues involved in the isopeptide bonds (K199, K355, and K463) into alanine and generated 3 different complementation plasmids, pAM\_BP<sub>K199A</sub>, pAM\_BP<sub>K355A</sub>, and pAM\_BP<sub>K463A</sub>, expressing mutated forms of the BP. As previously observed by mass spectrometry (MS) analysis, these specific mutations were able to abrogate the formation of isopeptide bonds in a recombinant BP-2a expressed in *E. coli* (20), confirming that these residues are absolutely necessary for isopeptide bond formation. Thus, the complementation vectors were used to transform the GBS-KO strain 515 lacking the main subunit (515ΔBP), and the effects of each mutation were assessed by immunoblotting analysis performed with antibodies raised against pilin subunits. The individual substitution of K199, K355, or K463 alone did not abrogate pilus protein polymerization, as the BP could still polymerize in high-molecular weight (HMW) structures, and the APs (AP1 and AP2) could be incorporated into pili (Fig. 1C).

### The glutamic acid 589 (E589) in the E-box motif of BP-2a is required for protein folding and pilus assembly

To investigate the role in pilus polymerization of the stabilizing residues of each isopeptide bond (D247, E416, and E589), we generated 3 new complementation vectors, pAM\_BP<sub>D247A</sub>, pAM\_BP<sub>E416A</sub>, and pAM\_

BP<sub>E589A</sub>, and transformed the KO strain 515ΔBP with these plasmids. Total protein extracts obtained from these complemented strains were analyzed by Western blot analysis using antisera raised against pilus proteins. The D247A and E416A substitutions did not affect pilus protein polymerization; in contrast, the mutation E589A completely abolished the formation of HMW structures (Fig. 1C), indicating that this isopeptide bond-catalyzing residue is pivotal for pilus 2a polymerization.

The conserved E589 is also the putative residue of the E-box motif of BP-2a (Supplemental Fig. S1). To explore in more detail the role of the E-box motif in BP stability and assembly, we used NMR spectroscopy. To this end, we used the single domains instead of the full-length protein because its large size is not compatible with NMR analysis. We have previously demonstrated that the single BP2a domains can fold independently and form the same isopeptide bonds found in the full-length protein (20). <sup>1</sup>H-<sup>15</sup>N-HSQC NMR spectra in principle contain signals from all HN groups of the protein, including backbone amide groups, as well as a number of side chains, therefore providing amino acid-specific information. The position of HN signals (chemical shift) depends on the chemical environment of the observed nuclei, which is determined by the protein's spatial arrangement and is, therefore, related to both secondary and tertiary structure. An unfolded protein displays poor signal dispersion because all amide atoms are in similar chemical environments. In contrast, in a well-folded protein, all HN atoms are influenced by different electronic distributions, determining good signal separation (31). The <sup>1</sup>H, <sup>15</sup>N-HSQC spectrum is expected to be only marginally perturbed by single-point mutations if the protein fold is pre-

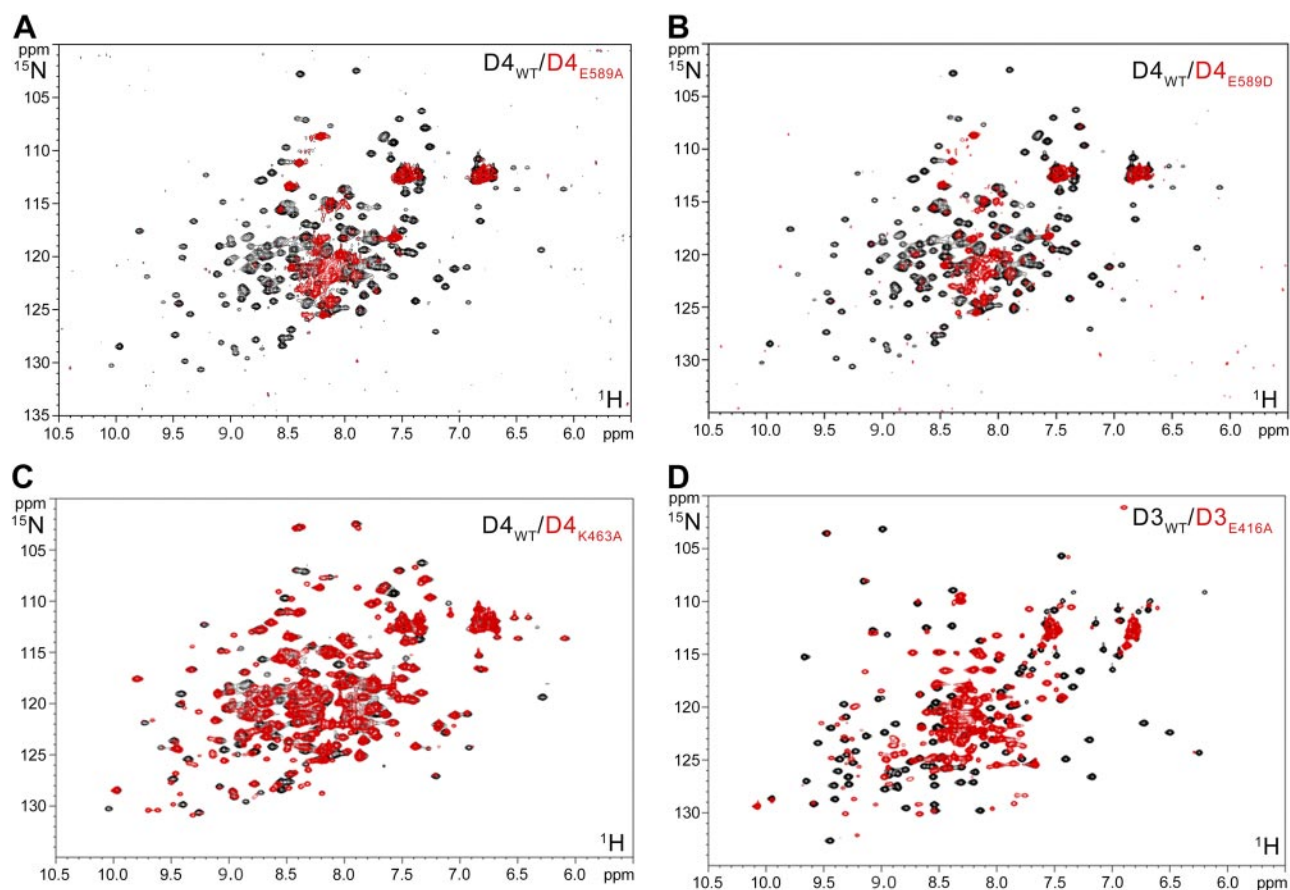


**Figure 1.** Isopeptide bonds are dispensable for pilus assembly. A) Structural details of the D2, D3, and D4 domains of BP-2a in the regions involved in isopeptide bond formation. B) Domain architecture of BP-2a. Position of the intramolecular isopeptide bonds is indicated. C) Immunoblotting analysis of total protein extracts from 515 mutant strain lacking the pilus 2a BP gene (515ΔBP2a) complemented by plasmids expressing BP-2a WT (BP<sub>wt</sub>) or BP carrying alanine substitutions of residues involved into isopeptide bond formation and stabilization. Nitrocellulose membranes were probed with antisera raised against the BP (α-BP-2a), AP1 (α-AP1-2a) and AP2 (α-AP2-2a).

served. The D4 domains of the WT BP (BP-D4<sub>WT</sub>) and of the mutated form BP-D4<sub>E589A</sub>, spanning from residue 459 to 649 and containing both the E-box motif and the LPxTG sorting signal, were produced as recombinant <sup>15</sup>N-labeled proteins and analyzed by NMR. The <sup>1</sup>H,<sup>15</sup>N-HSQC spectrum of the WT protein (Fig. 2A–C) showed a significant dispersion of peaks in both the proton (~4 ppm) and nitrogen (~28 ppm) frequency dimensions, while the spectrum of the E589A mutant is characterized by a much narrower chemical shift distribution (<sup>1</sup>H: ~1 ppm, <sup>15</sup>N: ~20 ppm), typical of a disordered protein (Fig. 2A). This result indicates that the E589A mutation caused the D4 domain to lose its native structure. To verify whether the replacement of the glutamic acid with a residue with similar physicochemical properties could preserve the structural role of E589, we generated a new D4 mutant (D4<sub>E589D</sub>), in which the E589 residue was replaced with an aspartic acid, and we analyzed this mutant by NMR (Fig. 2B). The recombinant <sup>15</sup>N-labeled D4<sub>E589D</sub> mutant showed a <sup>1</sup>H,<sup>15</sup>N-HSQC spectrum comparable to that of the E589A mutant, indicating that a glutamic acid in position 589 has an exclusive role on the stability and function of the protein and that removal of even a

single methylene group produces perturbation in the native structure.

To investigate whether the structural destabilization was related to the disruption of the isopeptide bond in the D4 domain after mutation of E589, we also analyzed the recombinant mutant domain D4, in which lysine 463, forming the isopeptide bond, is substituted by alanine (D4<sub>K463A</sub>). The <sup>1</sup>H,<sup>15</sup>N-HSQC spectrum of recombinant D4<sub>K463A</sub> showed that the native fold is preserved (Fig. 2C). Chemical shift perturbations observed for some peaks in the D4<sub>K463A</sub> mutant can be attributed to local changes in chemical environments introduced around the mutation site. Finally, the effect on domain folding of another stabilizing residue, E416, was explored analyzing by NMR the BP domain D3<sub>E416A</sub> in comparison with D3<sub>WT</sub> (Fig. 2D). It appears that the mutant protein is well structured; however, significant peak position changes are observed, suggesting an overall conformational rearrangement. Altogether, these observations confirm that E589 has an additional and key role in maintaining the correct protein fold, independently from intramolecular isopeptide bond formation. This specific function is exclusive for the E-box residue and



**Figure 2.** NMR spectroscopy assessment of the role of isopeptide bonds and E-box-motif in BP-2a folding. Comparison of NMR <sup>1</sup>H,<sup>15</sup>N-HSQC spectra of <sup>15</sup>N-BP-2a WT and mutants. *A*) Superimposition of HSQC spectra of D4<sub>WT</sub> (black) and D4<sub>E589A</sub> (red). *B*) Superimposition of HSQC spectra of D4<sub>WT</sub> (black) and D4<sub>E589D</sub> (red). *C*) Superimposition of HSQC spectra of D4<sub>WT</sub> (black) and D4<sub>K463A</sub> (red). *D*) Superimposition of HSQC spectra of D3<sub>WT</sub> (black) and D3<sub>E416A</sub> (red). All samples are in 50 mM phosphate buffer, pH 6.5. All the spectra were acquired at 25°C.

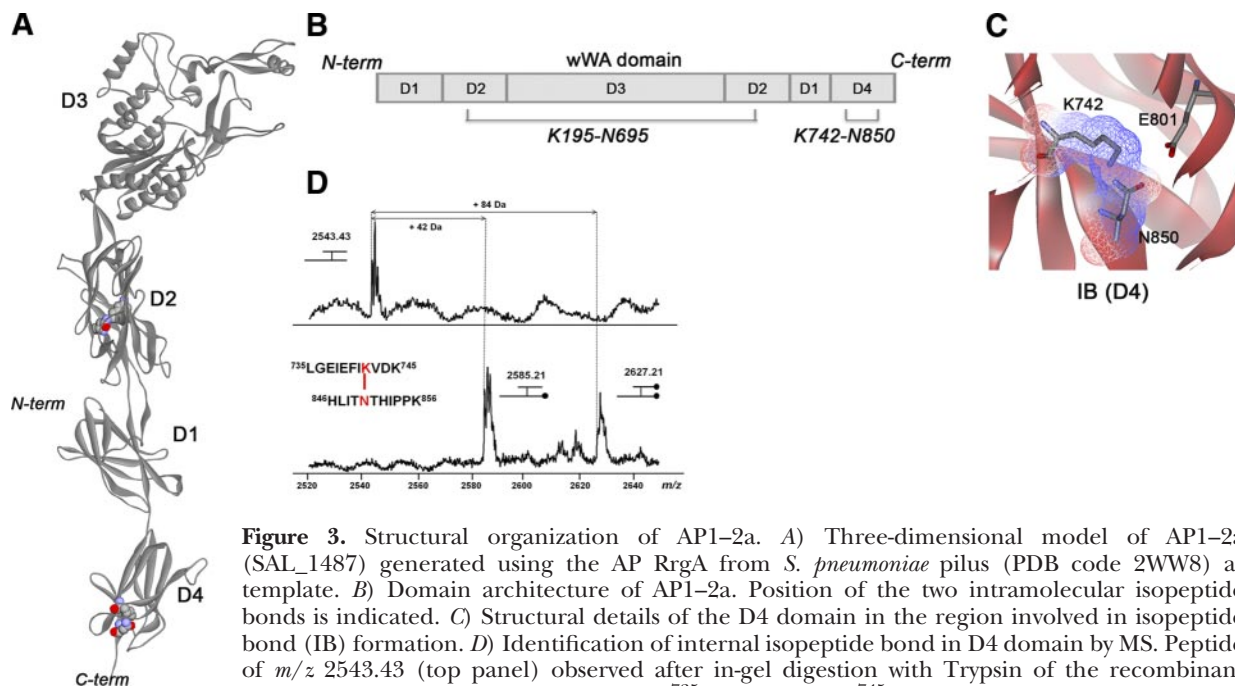
does not involve the other isopeptide bond-catalyzing amino acids in domains D2 and D3.

### The conserved glutamic acid E801 in the putative E-box of the major AP (AP1) is essential for its incorporation into HMW polymers

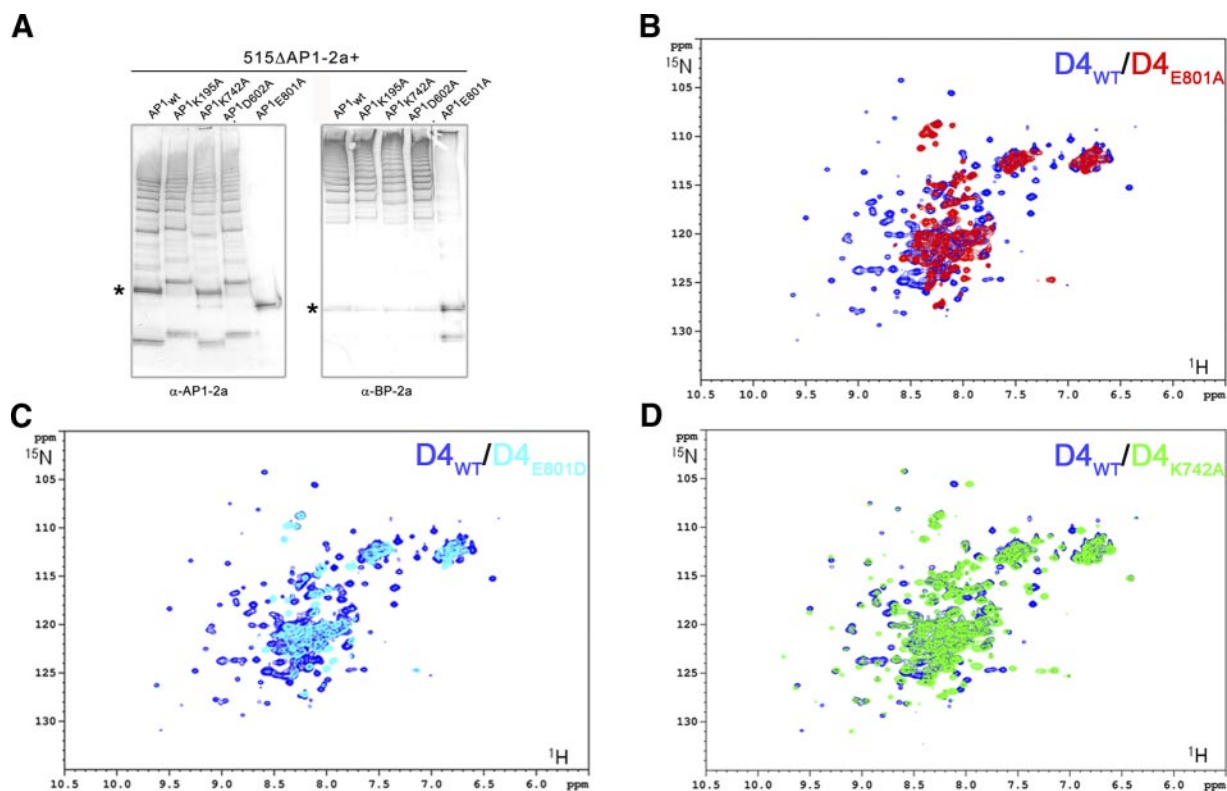
Sequence alignment analysis of the APs suggests that the major AP could contain a putative E-box-like motif, including the conserved glutamic acid E801 (Supplemental Fig. S1). In addition, intramolecular isopeptide bonds were also identified in both the APs of pilus 2a by comparative homology modeling and MS analysis. The structural model of AP1 (shown in Fig. 3A) was generated using the AP RrgA from *S. pneumoniae* pilus (PDB code 2WW8; ref. 32) as a template. Sequence alignment between GBS pilus 2a AP1 and RrgA showed a sequence identity of 38.9% and a sequence similarity of 57.1% (Supplemental Fig. S2). The structural model indicates a 4-domain organization: D1 domain carrying 5  $\beta$  strands (IgG domain); D2, 11  $\beta$  strands (CnaB-like domain). In the model, this arrangement is stabilized by an isopeptide bond between Lys195 and Asn695 that links the N-terminal part of the protein with the C terminus (Fig. 3B). D1 and D2 carry elements both from N- and C-terminal primary sequence. D3 is a von Willebrand factor type A (vWA)-like [domain] domain, found in many eukaryotic cell surface proteins involved in interactions with the extracellular matrix, with two extensions; D4 is made by 8  $\beta$  strands (IgG domain) with a second isopeptide bond involving Lys742 and Asn850, and stabilized by Glu801 (Fig. 3C), corresponding to the conserved glutamic acid residue identified in the putative E-box-like motif (Sup-

plemental Fig. S1). The presence of this intra-D4 domain isopeptide bond was verified by MS analysis by using a previously described protocol (20). The measured molecular mass of domain D4 ( $30,268.16 \pm 0.13$  Da) was consistent with the expected mass of the molecule carrying one isopeptide bond (data not shown). The specific involvement of Lys742-Asn850 in the bond formation was confirmed by in-gel trypsin digestion of recombinant D4 followed by peptide mass fingerprint analysis of the unmodified and *O*-methylisourea-modified peptides (Fig. 3D). As observed for the isopeptide bonds in the BP, the mutated AP1<sub>K195A</sub> or AP1<sub>K742A</sub> did not affect the incorporation of pilin subunits (AP1 and BP) into HMW pilus structures (Fig. 4A).

We also generated a plasmid, pAM\_AP1<sub>E801A</sub>, expressing a mutant form of AP1 with the glutamic acid residue of the putative E-box mutated into alanine. We transformed this vector in the KO strain lacking the WT AP1 ( $515\Delta$ AP1-2a), and the effects of this mutation were assessed by immunoblot analysis. We found that the glutamic acid within the E-box, as in the case of the BP, is absolutely necessary for AP1 incorporation into pilus polymers. In fact, only the monomeric species of AP1 could be detected by Western blot (Fig. 4A). In addition, to verify whether the same structural destabilization effect observed in the BP-D4<sub>E589A</sub> mutant was also present in the AP1<sub>E801A</sub> mutant, the recombinant D4 domains of the WT AP1 and mutants AP1<sub>E801A</sub>, AP1<sub>E801D</sub>, and AP1<sub>K742A</sub> were produced in *E. coli* as <sup>15</sup>N-labeled samples and analyzed by NMR. The <sup>1</sup>H, <sup>15</sup>N-HSQC spectrum of AP1<sub>WT</sub> is indicative of a folded domain, while the spectra of AP1<sub>E801A</sub> (Fig. 4B) and AP1<sub>E801D</sub> (Fig. 4C) show the presence of signals in a narrow range of the proton frequency dimension ( $\sim 1$



**Figure 3.** Structural organization of AP1-2a. *A*) Three-dimensional model of AP1-2a (SAL\_1487) generated using the AP RrgA from *S. pneumoniae* pilus (PDB code 2WW8) as template. *B*) Domain architecture of AP1-2a. Position of the two intramolecular isopeptide bonds is indicated. *C*) Structural details of the D4 domain in the region involved in isopeptide bond (IB) formation. *D*) Identification of internal isopeptide bond in D4 domain by MS. Peptide of  $m/z$  2543.43 (top panel) observed after in-gel digestion with Trypsin of the recombinant protein is in agreement with the peptide <sup>735</sup>LGEIEFIKVDK<sup>745</sup> (expected  $m/z$  1290.72) linked by and isopeptide bond to the peptide <sup>846</sup>HLITNTHIPPK<sup>856</sup> (expected  $m/z$  1270.72). After incubation with *O*-methylisourea, two molecular species with a mass increase of 42 Da (single reaction) and 84 Da (double reaction) were observed, confirming the cross-linked nature of the peptides.



**Figure 4.** Role of E-box-like motif of the major AP (AP1–2a) in pilus polymerization. *A*) Immunoblotting analysis of total protein extracts from 515 mutant strain lacking the pilus 2a AP1 gene (515ΔAP1–2a) complemented by plasmids expressing AP1–2a WT (AP1<sub>wt</sub>) or AP1 carrying alanine substitutions of residues involved into isopeptide bond formation. Nitrocellulose membranes were probed with antisera raised against AP1 (α-AP1–2a) and BP (α-BP–2a). *B*) Superimposition of <sup>1</sup>H,<sup>15</sup>N-HSQC spectra recorded on D4<sub>WT</sub> (blue) and D4<sub>E801A</sub> (red). *C*) Superimposition of <sup>1</sup>H,<sup>15</sup>N-HSQC spectra recorded on D4<sub>WT</sub> (blue) and D4<sub>E801D</sub> (cyan). *D*) Superimposition of <sup>1</sup>H,<sup>15</sup>N-HSQC spectra recorded on D4<sub>WT</sub> (blue) and D4<sub>K742A</sub> (green). All samples are in 50 mM phosphate buffer, pH 6.5. All the spectra were acquired at 25°C.

ppm), consistent with an unstructured polypeptide. It can be further observed that abrogation of the isopeptide bond established by K742, subsequent to the mutation in this position, does not impair proper tertiary structure formation (Fig. 4*D*). It can be concluded that the E-box glutamic acid is essential for protein folding but not as a consequence of a destabilization of the intramolecular isopeptide bond in the D4 domain.

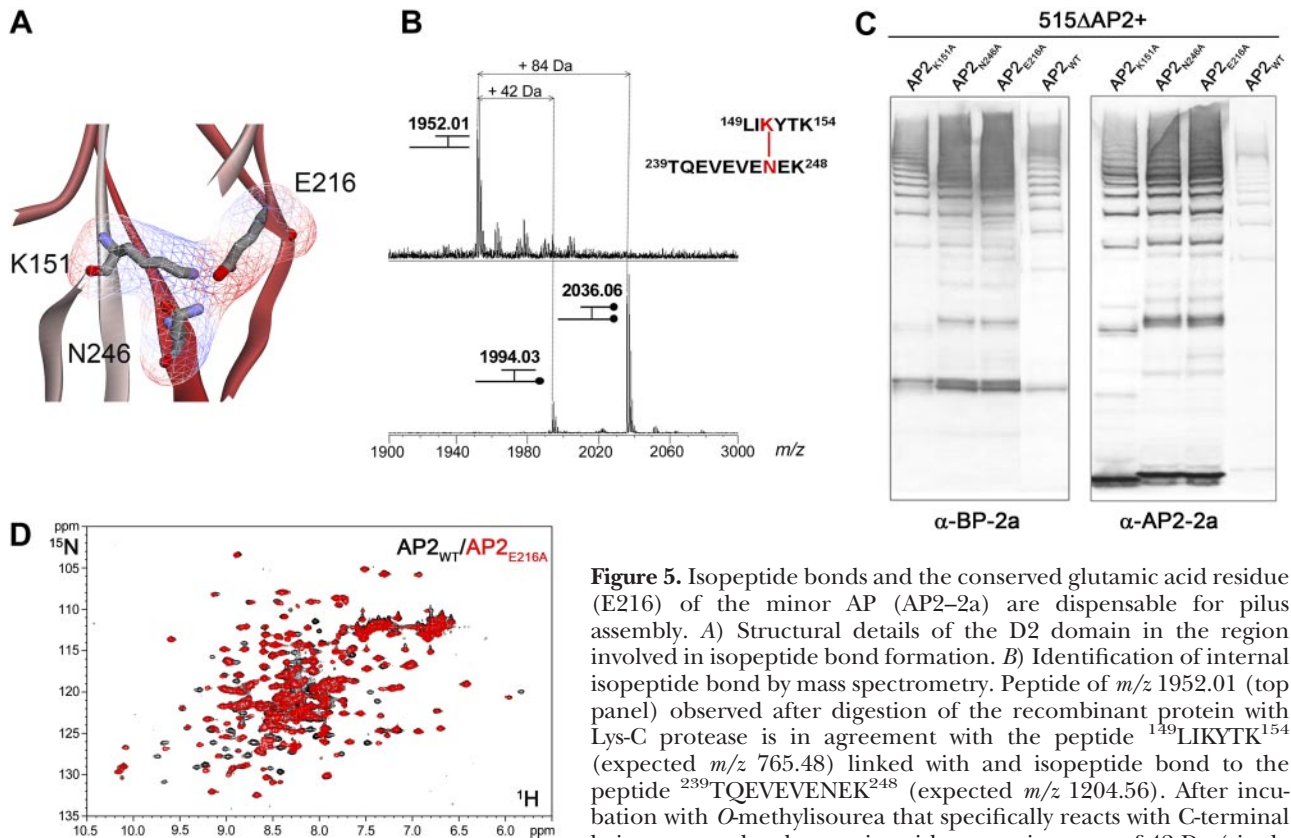
#### AP2 contains an isopeptide bond dispensable for its incorporation into pili and does not contain an E-box

The 3-dimensional model of the pilus anchor protein AP2 generated using the crystal structure of GBS52 (PDB code 2PZ4) as a template, the minor AP of PI-1 (33), revealed the presence of a K151-E216-N246 triad whose spatial arrangement is compatible with the formation of a Lys-Asn isopeptide bond in the C-terminal IgG-like fold domain (Fig. 5*A*). The presence of this intramolecular link was verified by MS analysis of the recombinant protein AP2–2a (Fig. 5*B*). No other cross-linked peptides were found in the MS spectra of recombinant AP2, confirming that this pilin subunit harbors only one intramolecular isopeptide bond. Therefore, we examined the potential contribution of this bond toward protein incorporation in the pilus by generating AP2 mutants, contain-

ing site-directed substitutions of Lys151, Asn246, and Glu216 into alanine. The KO 515ΔAP2 strain was transformed with plasmids harboring the 3 mutations (pAM<sub>AP2</sub><sub>K151A</sub>, pAM<sub>AP2</sub><sub>N246A</sub>, and pAM<sub>AP2</sub><sub>E216A</sub>), and the resultant strains were then examined for AP2 incorporation into pili. Total cell lysates were analyzed by Western blot using antisera raised against AP2–2a and BP–2a. No alanine substitution affected the incorporation of the minor ancillary subunit into the pilus (Fig. 5*C*). These data reveal that although sequence analysis suggests that the minor subunit harbors a putative E-box-like motif, the glutamic acid catalyzing the isopeptide bond formation is not required for AP2 incorporation. Moreover, the <sup>1</sup>H,<sup>15</sup>N-HSQC spectra of the recombinant full-length AP2 and the mutant AP2<sub>E216A</sub> show that the overall fold of the protein is not compromised by the mutation (Fig. 5*D*). The extensive overlap of peaks between the two spectra further indicates that only minor structural rearrangements occur on mutation.

#### Molecular simulations indicate a possible requirement of glutamic acid E589 for drying the hydrophobic core of D4 of BP-2a

The E589A mutation in the D4 domain of BP-2a causes a strong destabilization of the folded state, resulting in



**Figure 5.** Isopeptide bonds and the conserved glutamic acid residue (E216) of the minor AP (AP2-2a) are dispensable for pilus assembly. *A*) Structural details of the D2 domain in the region involved in isopeptide bond formation. *B*) Identification of internal isopeptide bond by mass spectrometry. Peptide of  $m/z$  1952.01 (top panel) observed after digestion of the recombinant protein with Lys-C protease is in agreement with the peptide  $^{149}\text{LIKTYK}^{154}$  (expected  $m/z$  765.48) linked with and isopeptide bond to the peptide  $^{239}\text{TQEVEVENEK}^{248}$  (expected  $m/z$  1204.56). After incubation with *O*-methylisourea that specifically reacts with C-terminal lysine, two molecular species with a mass increase of 42 Da (single reaction) and 84 Da (double reaction) were observed, confirming

the cross-linked nature of peptides. *C*) Immunoblotting analysis of total protein extracts from 515 mutant strain lacking the AP2 protein (515 $\Delta$ AP2) complemented by plasmids expressing AP2-2a WT (AP2<sub>wt</sub>) or AP2 carrying alanine substitutions of residues involved into isopeptide bond formation. Nitrocellulose membrane was probed with antiserum raised against the BP ( $\alpha$ -BP-2a) and AP2 ( $\alpha$ -AP2-2a). *D*) Superimposition of  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectra recorded on full-length AP2<sub>WT</sub> (black) and AP2<sub>E216A</sub> (red). The samples are in 50-mM phosphate buffer, pH 6.5. The spectra were acquired at 25°C.

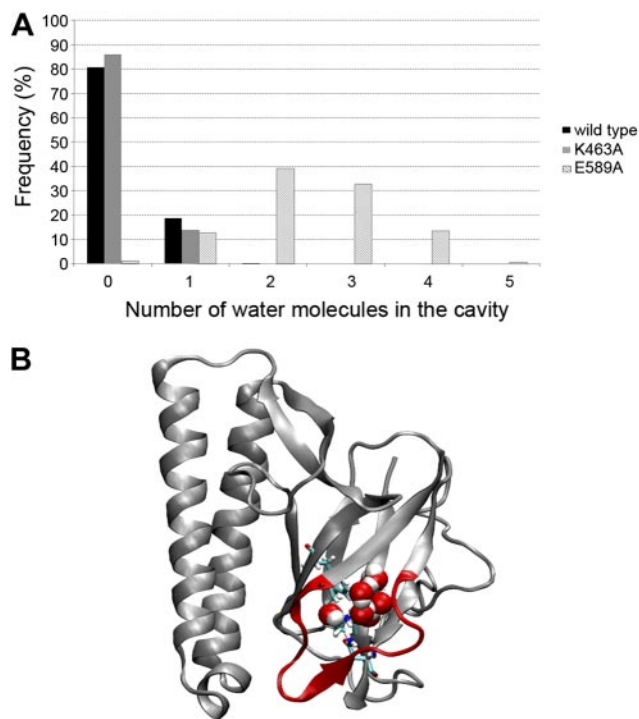
unfolding of the protein. Comparative MD simulations in explicit solvent were used to investigate the dynamics of the WT as well as the E589A and the K463A D4 domains. The aim was to identify possible molecular details of the destabilizing effect specific for the E589A mutation. The simulation time (50 ns) is too short to investigate complete unfolding of the D4 domain. However, even on this time scale, the analysis of conformational fluctuations and drift can give important insight into the molecular details that stabilize or destabilize the folded D4 state. The analysis of conformational fluctuations (RMSF) and drift (RMSD) of the MD simulations indicated a region spanning residues 590 to 600 (Supplemental Fig. S3A), which showed in case of the E589 mutation a much higher RMS deviation and RMS fluctuation compared with the WT and the other mutants (Supplemental Fig. S3B, C) and may indicate the onset of unfolding. The mutation of the glutamic acid into alanine, but not the disruption of the isopeptide bond, disturbed the proper folding of one loop related to the  $\beta$  barrel and made the core domain more accessible for solvent molecules. The core structure of the D4 domain contains a large hydrophobic area that surrounds the isopeptide bond and the conserved E589 completely (20). To investigate the solvation of this region, we recorded the water molecules

that entered during the MD simulations. The simulation of the E589A mutation resulted in the highest accumulation of water molecules in this pocket region compared to the WT and K463A mutations (Fig. 6A, B). These results, which perfectly correlate with those observed by NMR and mutagenesis experiments, revealed that the change in hydration of the hydrophobic core could be the first step toward the process of domain unfolding and that the glutamic acid 589 has a pivotal role in maintaining the proper conformational stability and solvation of the region around the isopeptide bond.

## DISCUSSION

Although in gram-positive bacteria, pilin subunits show low-sequence similarities, they have a quite similar structural architecture and assembly mechanism, in which these subunits are assembled in a modular fashion by covalent intermolecular linkages. The available 3-dimensional structures of the shaft-forming major pilins of *S. pyogenes* (16), *Bacillus cereus* (34), *C. diphtheriae* (15), *Streptococcus pneumoniae* (35), and the recently solved crystals of backbone subunits of Group B *Streptococcus pili* (20, 36) have revealed a common





**Figure 6.** Hydration of the hydrophobic core of BP-2a D4 domain. *A*) Observed number of water molecules within 4 Å of any atom of the F477 residue of WT and mutant BP-2a D4 domains during MD simulations. Residue F477 is located within the pocket region close to the isopeptide bond (formed in case of the WT protein). *B*) Cartoon representation of the E589A mutant (mobile loop highlighted in black) with the water molecules (CPK colors) entering the pocket close to residues F477, A589, and K463 (indicated as sticks) during the simulation.

IgG-like domain organization, in which single domains are stabilized by intramolecular isopeptide bonds. These linkages, which are covalent bonds between lysine and asparagine side chains, form autocatalytically during protein folding and confer proteolytic and thermal stability to pili (17, 18, 37). Neighboring glutamic or aspartic residues play an important role in catalyzing and stabilizing these bonds.

The *in vivo* mutagenesis experiments performed in GBS pilus 2a subunits and reported in this paper show that lysine residues involved in the intramolecular isopeptide bonds are dispensable for pilus protein polymerization, as well as the aspartic or glutamic acid stabilizing these linkages, with the exception of the glutamic acid in the E-box motifs. By sequence alignments among homologous pilin proteins, a putative E-box-like motif can be identified not only in the shaft-forming major subunits, as already reported (12), but also in the major APs, known also as tip pilins. In particular, although alanine substitution of Lys463 in the BP and Lys742 in the major AP (AP1) did not affect pilus assembly, the substitution of the isopeptide-bond-stabilizing glutamic acid residues into the E-box regions (BP-Glu589 or AP1-Glu801) completely abolished fiber formation. Therefore, the effect of the E-box glutamic acid for pilus polymerization is not related to the

stabilization of the internal isopeptide bonds. This finding is in agreement with previous observations showing that in the major pilin subunits, BcpA of *B. cereus* pili, SpaA of *C. diphtheriae*, and Spy128 of *S. pyogenes*, only the substitution of the E-box glutamic acid stabilizing the isopeptide bonds close to the LPXTG sorting signal was lethal for pili assembly (15, 17, 34, 38).

An NMR-based investigation has been conducted in order to address the role of the E-box motif in pilus stability and folding of GBS pilus 2a-forming proteins. D4 domains of BP and AP1 WT and E-box mutants were produced as recombinant <sup>15</sup>N-labeled proteins and analyzed by heteronuclear NMR experiments. The <sup>1</sup>H,<sup>15</sup>N-HSQC spectra of E-box mutants show that the E589A mutation causes unfolding of the D4 domain of BP-2a. The reason for this dramatic effect may be found by closer inspection of the crystallographic structure. Domain D4 can be viewed as constituted by two structurally distinct subdomains, one of which is all-helical, while the second displays several β strands. The latter subdomain appears as a relatively fluxional moiety, with a significant share of amino acid residues (~65% of total) not being part of well-defined secondary structure elements (Supplemental Fig. S4A). E589 belongs to the β-sheet core with the side chain pointing to its interior. Polar contacts are established between the carboxylate oxygen atoms and K598, L599, and an internal water molecule. The latter is, in turn, bound *via* hydrogen interactions to A597 and N636. Residues 597–599 belong to a nearby extended loop, while N636 is located in the C-terminal loop. It is, therefore, likely that this hydrophilic network stabilizes the overall structure by linking flexible regions to the protein core.

At variance with the case of BP-D4, mutation of the stabilizing residue E416 in BP-D3 does not produce disruption of the tertiary fold. Indeed, BP-D3 is a stable well-structured domain and the carboxylic group of E416 is not involved in polar interactions with other residues or water molecules. On the contrary, the side chain of E416 is located in a hydrophobic environment (Supplemental Fig. S4B), and its mutation into alanine, a more hydrophobic amino acid, does not abrogate preexisting interactions, rather it could favor hydrophobic core formation.

The above-described structural grounds for domain stability may, however, not be sufficient to provide a complete picture of mutation-induced unfolding due to the inherent complexity of the folding process. For example, because folding of globular proteins almost invariably involves the burial of hydrophobic side chains, it has been proposed that dehydration events play an essential role in the folding mechanism (39). The interaction of water with hydrophobic solutes has been deeply studied for many years (40, 41). In particular, solvation changes in limited protein regions have been connected to changes in protein structures (42) and binding affin-

ity with ligands (43). To explore the possibility that water molecules play a substantial role in the unfolding process of domain D4, we performed an explicit MD simulation. Although the simulation times were too short to directly investigate a possible unfolding of the D4 domain, it was possible to analyze a more rapid change in hydration. The latter analysis indicated that after single-point mutation of E589 into alanine, a more extended hydration of the protein core region occurs. In addition, increased local conformational mobility is observed for loop 590–600, in agreement with the loss of hydrophilic interactions fastening it to the domain core. Thus, it is possible that the disruption of interactions involving E589 renders the structure even more fluxional, allowing water molecules to penetrate into the hydrophobic core and energetically disfavor formation of the native fold.

In summary, the observation that the E-box mutation induces structural destabilization of the domain D4 of the BP suggests the impossibility for SrtC enzymes to interact with the LPxTG of substrate BP protein as the sorting signal is located close to the E-box motif, leading to the abrogation of pili polymerization *in vivo* when the E-box is mutated. Our work provides the first structural insight into the effects of E-box mutation in pilus protein polymerization. FJ

The authors are grateful to Henriette Molinari (NMR Laboratory, University of Verona, Verona, Italy) for valuable discussions and critical review of the manuscript, and Matthew Bottomley for editing the manuscript.

## REFERENCES

- Johri, A. K., Paoletti, L. C., Glaser, P., Dua, M., Sharma, P. K., Grandi, G., and Rappuoli, R. (2006) Group B *Streptococcus*: global incidence and vaccine development. *Nat. Rev. Microbiol.* **4**, 932–942
- Lauer, P., Rinaudo, C. D., Soriani, M., Margarit, I., Maione, D., Rosini, R., Taddei, A. R., Mora, M., Rappuoli, R., Grandi, G., and Telford, J. L. (2005) Genome analysis reveals pili in group B *Streptococcus*. *Science* **309**, 105
- Maione, D., Margarit, I., Rinaudo, C. D., Massignani, V., Mora, M., Scarselli, M., Tettelin, H., Brettoni, C., Iacobini, E. T., Rosini, R., D'Agostino, N., Miorin, L., Buccato, S., Mariani, M., Galli, G., Nogarotto, R., Nardi Dei, V., Vegni, F., Fraser, C., Mancuso, G., Teti, G., Madoff, L. C., Paoletti, L. C., Rappuoli, R., Kasper, D. L., Telford, J. L., and Grandi, G. (2005) Identification of a universal group B *Streptococcus* vaccine by multiple genome screen. *Science* **309**, 148–150
- Margarit, I., Rinaudo, C. D., Galeotti, C. L., Maione, D., Ghezzi, C., Buttazzoni, E., Rosini, R., Runci, Y., Mora, M., Buccato, S., Pagani, M., Tresoldi, E., Berardi, A., Creti, R., Baker, C. J., Telford, J. L., and Grandi, G. (2009) Preventing bacterial infections with pilus-based vaccines: the group B *Streptococcus* paradigm. *J. Infect. Dis.* **199**, 108–115
- Proft, T., and Baker, E. N. (2009) Pili in gram-negative and gram-positive bacteria—structure, assembly, and their role in disease. *Cell. Mol. Life Sci.* **66**, 613–635
- Konto-Ghiorghi, Y., Mairey, E., Mallet, A., Dumenil, G., Caliot, E., Trieu-Cuot, P., and Dramsi, S. (2009) Dual role for pilus in adherence to epithelial cells and biofilm formation in *Streptococcus agalactiae*. *PLoS Pathog.* **5**, e1000422
- Ton-That, H., and Schneewind, O. (2004) Assembly of pili in Gram-positive bacteria. *Trends Microbiol.* **12**, 228–234
- 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
- Dramsi, S., Caliot, E., Bonne, I., Guadagnini, S., Prevost, M. C., Kojadinovic, M., Lalioui, L., Poyart, C., and Trieu-Cuot, P. (2006) Assembly and role of pili in group B *Streptococci*. *Mol. Microbiol.* **60**, 1401–1413
- Rosini, R., Rinaudo, C. D., Soriani, M., Lauer, P., Mora, M., Maione, D., Taddei, A., Santi, I., Ghezzi, C., Brettoni, C., Buccato, S., Margarit, I., Grandi, G., and Telford, J. L. (2006) Identification of novel genomic islands coding for antigenic pilus-like structures in *Streptococcus agalactiae*. *Mol. Microbiol.* **61**, 126–141
- Ton-That, H., and Schneewind, O. (2003) Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol. Microbiol.* **50**, 1429–1438
- Ton-That, H., Maraffini, L. A., and Schneewind, O. (2004) Sortases and pilin elements involved in pilus assembly of *Corynebacterium diphtheriae*. *Mol. Microbiol.* **53**, 251–261
- Yeates, T. O., and Clubb, R. T. (2007) Biochemistry. How some pili pull. *Science* **318**, 1558–1559
- Mandlik, A., Swierczynski, A., Das, A., and Ton-That, H. (2008) Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development. *Trends Microbiol.* **16**, 33–40
- Kang, H. J., Paterson, N. G., Gaspar, A. H., Ton-That, H., and Baker, E. N. (2009) The *Corynebacterium diphtheriae* shaft pilin SpaA is built of tandem Ig-like modules with stabilizing isopeptide and disulfide bonds. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 16967–16971
- 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
- Budzik, J. M., Poor, C. B., Faull, K. F., Whitelegge, J. P., He, C., and Schneewind, O. (2009) Intramolecular amide bonds stabilize pili on the surface of bacilli. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19992–19997
- Kang, H. J., and Baker, E. N. (2011) Intramolecular isopeptide bonds: protein crosslinks built for stress? *Trends Biochem. Sci.* **36**, 229–237
- Hendrickx, A. P., Budzik, J. M., Oh, S. Y., and Schneewind, O. (2011) Architects at the bacterial surface—sortases and the assembly of pili with isopeptide bonds. *Nat. Rev. Microbiol.* **9**, 166–176
- Nuccitelli, A., Cozzi, R., Gourlay, L. J., Donnarumma, D., Necchi, F., Norais, N., Telford, J. L., Rappuoli, R., Bolognesi, M., Maione, D., Grandi, G., and Rinaudo, C. D. (2011) Structure-based approach to rationally design a chimeric protein for an effective vaccine against group B *Streptococcus* infections. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 10278–10283
- Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. (2005) GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718
- Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950–1958
- Jorgensen, W., Chandrasekhar, J., Madura, J., Impey, R., and Klein, M. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926
- Nose, S. (1984) A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268
- Hoover, W. G. (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697
- Parrinello, M., and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190
- Nose, S., and Klein, M. (1983) Constant pressure molecular dynamics for molecular systems. *Mol. Phys.* **50**, 1055–1076
- Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38, 27–38
- Klock, H. E., and Lesley, S. A. (2009) The polymerase incomplete primer extension (PIPE) method applied to high-through-

- put cloning and site-directed mutagenesis. *Methods Mol. Biol.* **498**, 91–103
30. Cozzi, R., Malito, E., Nuccitelli, A., D'Onofrio, M., Martinelli, M., Ferlenghi, I., Grandi, G., Telford, J. L., Maione, D., and Rinaudo, C. D. (2011) Structure analysis and site-directed mutagenesis of defined key residues and motives for pilus-related sortase C1 in group B *Streptococcus*. *FASEB J.* **25**, 1874–1886
  31. Kwan, A. H., Mobli, M., Gooley, P. R., King, G. F., and Mackay, J. P. (2011) Macromolecular NMR spectroscopy for the non-spectroscopist. *FEBS J.* **278**, 687–703
  32. Izore, T., Contreras-Martel, C., El Mortaji, L., Manzano, C., Terrasse, R., Vernet, T., Di Guilmi, A. M., and Dessen, A. (2010) Structural basis of host cell recognition by the pilus adhesin from *Streptococcus pneumoniae*. *Structure* **18**, 106–115
  33. Krishnan, V., Gaspar, A. H., Ye, N., Mandlik, A., Ton-That, H., and Narayana, S. V. (2007) An IgG-like domain in the minor pilin GBS52 of *Streptococcus agalactiae* mediates lung epithelial cell adhesion. *Structure* **15**, 893–903
  34. Budzik, J. M., Marraffini, L. A., Souda, P., Whitelegge, J. P., Faull, K. F., and Schneewind, O. (2008) Amide bonds assemble pili on the surface of bacilli. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 10215–10220
  35. Spraggon, G., Koesema, E., Scarselli, M., Malito, E., Biagini, M., Norais, N., Emolo, C., Barocchi, M. A., Giusti, F., Hilleringmann, M., Rappuoli, R., Lesley, S., Covacci, A., Maignani, V., and Ferlenghi, I. (2010) Supramolecular organization of the repetitive backbone unit of the *Streptococcus pneumoniae* pilus. *PLoS. One* **5**, e10919
  36. Vengadesan, K., Ma, X., Dwivedi, P., Ton-That, H., and Narayana, S. V. (2011) A model for group B *Streptococcus* pilus type 1: the structure of a 35-kDa C-terminal fragment of the major pilin GBS80. *J. Mol. Biol.* **407**, 731–743
  37. Alegre-Cebollada, J., Badilla, C. L., and Fernandez, J. M. (2010) Isopeptide bonds block the mechanical extension of pili in pathogenic *Streptococcus pyogenes*. *J. Biol. Chem.* **285**, 11235–11242
  38. Kang, H. J., and Baker, E. N. (2009) Intramolecular isopeptide bonds give thermodynamic and proteolytic stability to the major pilin protein of *Streptococcus pyogenes*. *J. Biol. Chem.* **284**, 20729–20737
  39. Cheung, M. S., Garcia, A. E., and Onuchic, J. N. (2002) Protein folding mediated by solvation: water expulsion and formation of the hydrophobic core occur after the structural collapse. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 685–690
  40. Chandler, D. (2002) Hydrophobicity: two faces of water. *Nature* **417**, 491
  41. Ball, P. (2003) Chemical physics: How to keep dry in water. *Nature* **423**, 25–26
  42. Eggers, D. K., and Valentine, J. S. (2001) Crowding and hydration effects on protein conformation: a study with sol-gel encapsulated proteins. *J. Mol. Biol.* **314**, 911–922
  43. Young, T., Hua, L., Huang, X., Abel, R., Friesner, R., and Berne, B. J. (2010) Dewetting transitions in protein cavities. *Proteins* **78**, 1856–1869

Received for publication September 23, 2011.  
Accepted for publication January 10, 2012.