Supporting Information

Kang et al. 10.1073/pnas.0906826106

SI Materials and Methods

Construction of Plasmid pAG153 Encoding SpaA and SrtA. Primers SpaA-498–5 (aaaagatcttgagttcgattggctttttttc) and SpaA-KpnI-3 (aaaggta ccgttctttttcttgttgtcgatc) were used to PCR-amplify the promoter and coding sequence of *spaA* (devoid of the sequence encoding the last 39 C-terminal residues) from plasmid pHTT11 DNA while appending BglII and KpnI sites for cloning purposes (1). Primers Kpn-MH6–5 (aaaggtaccatgcatcaccatcaccatcacgcc ggatttgaactgccac) and $SrtA_{\Delta 13}$-Bg-3 (cgcagatctttaggtgatgatgatg-gcgatg) were used to amplify a DNA fragment containing the coding sequence for 6 histidine residues and the last 39 residues of *spaA* fused to the coding sequence of *srtA* lacking the coding sequences for the last 13 residues, while appending KpnI and BglII sites. Both DNA fragments were digested with BglII and KpnI restriction enzymes and ligated with the cleaved BglII site of the *E. coli*/*Corynebacterium* shuttle vector pCGL0243 (1) to generate pAG153. This recombinant plasmid was transformed into a *C. diphtheriae* strain that lacks *spaA* and *srtA*.

Isolation of SpaA Pili. Overnight cultures of *C. diphtheriae* harboring pAG153 were diluted 1:10 in a defined medium supplemented with 50 μg/mL kanamycin. The supernatant of the overnight culture was collected by centrifugation and saturated with $(NH_4)_2SO_4$ to 30% and then 75% for 1 h at 4 °C. Precipitated proteins were solubilized in 15 mL of EQ buffer (150 mM NaCl and 50 mM Tris-HCl, pH 7.5) and dialyzed twice in EQ buffer for 24 h at 4 °C. Dialyzed material was centrifuged to remove insoluble proteins, and soluble His_6-tagged SpaA pilus proteins were purified using nickel-affinity chromatography as previously described (1).

Proteolytic Digestion and Mass Spectral Analyses. Purified SpaA pili and recombinant SpaA protein were digested and analyzed in a similar way to that previously described (2). SDS-PAGE gel bands containing ≈10 μg of SpaA or SpaA pili were diced and washed with 50% acetonitrile and 25 mM NH_4HCO_3 to remove the gel stain. The gel pieces were incubated with 100% acetonitrile and dried under vacuum before adding proteases dissolved in 25 mM NH_4HCO_3 and 10% acetonitrile. The gel pieces were incubated with trypsin (Promega) for 4 h followed by AspN endopeptidase (Roche) for 16 h at 37 °C. For SpaA pili, 1.5 μg of trypsin and 2 μg of AspN were used. For recombinant SpaA, 500 ng of trypsin and 200 ng of AspN were used. The digested samples were analyzed using a Q-STAR XL Hybrid MS/MS system (Applied Biosystems) and on a 0.3 × 150-mm C_{18} column with 0.1% aqueous formic acid (A) and 0.1% formic acid in acetonitrile (B), in which a linear gradient of 10–35% B was applied to separate the peptides. Data within the *m*/*z* range 300–1,600 were acquired, and 3 product ion scans were performed on the 3 most abundant multiply charged precursors. Initial searches against SpaA sequence using Mascot search engine version 2.0.05 (Matrix Science) identified linear peptides only. Unmatched peptides were manually searched to identify peptides crosslinked by the isopeptide bonds, either intramolecular or intermolecular.

1. Ton-That H, Schneewind O (2003) Assembly of pili on the surface of *Corynebacterium diphtheriae*. *Mol Microbiol* 50:1429–1438.

2. Kang HJ, Coulibaly F, Clow F, Proft T, Baker EN (2007) Stabilizing isopeptide bonds revealed in Gram-positive bacterial pilus structure. *Science* 318:1625–1628.

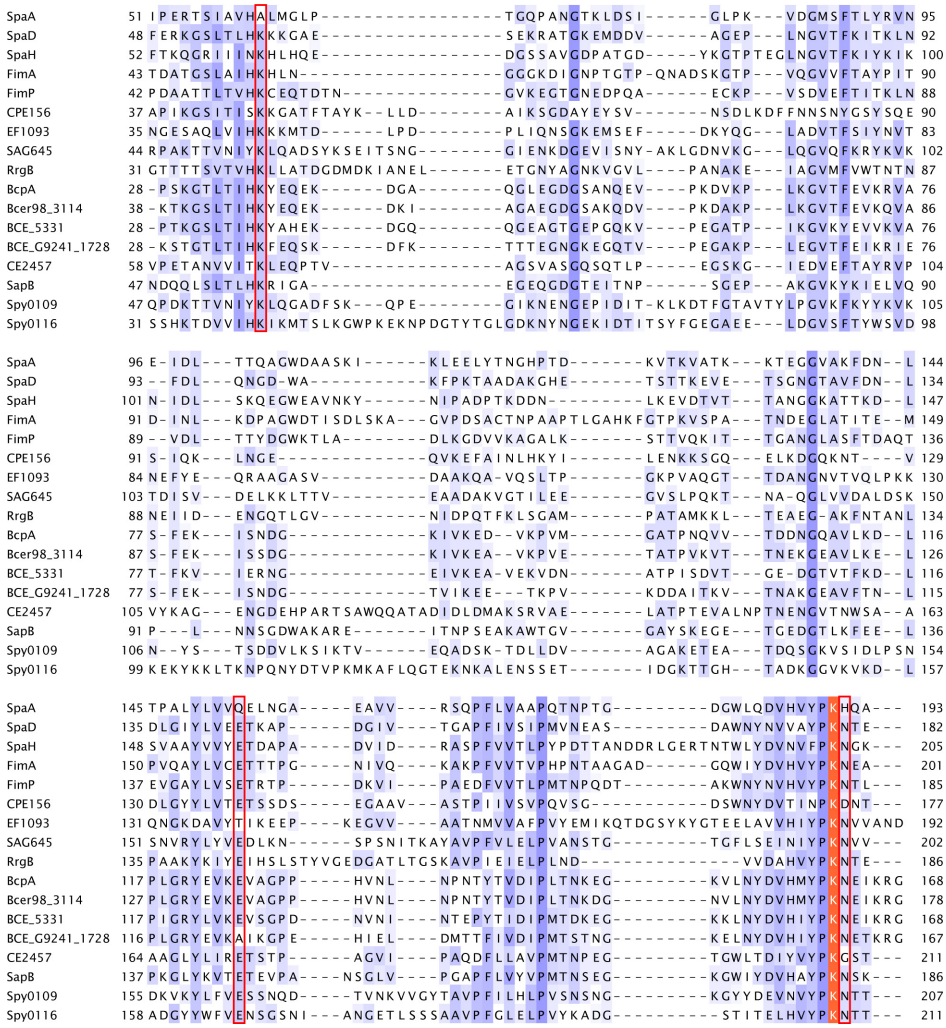

Fig. S1. Sequence alignment of N-terminal regions of Gram-positive bacterial major pilins. Residues that are aligned with the isopeptide bond forming residues of BcpA (K37 and Asn-163) [Budzik JM et al. (2008) Amide bonds assemble pili on the surface of bacilli. *Proc Natl Acad Sci USA* 105:10215–20] and predicted catalytic residue (E125 of BcpA) are outlined with red boxes. The pilin motif lysines are highlighted in orange. SpaA, SpaD, and SpaH from *C. diphtheriae*, FimA and FimP from *Actinomyces naeslundii*, CPE156 from *Clostridium perfringens*, EF1093 from *Enterococcus faecalis*, SAG645 from GBS, RrgB from *S. pneumoniae*, BcpA, Bcer98_3114, BCE_5331, and BCE_G9241_1728 from *B. cereus*, CE2457 from *Corynebacterium efficiens*, SapB from *Corynebacterium jeikeium*, Spy0109 from *S. pyogenes* strain M2, and Spy0116 from *S. pyogenes* strain M4.

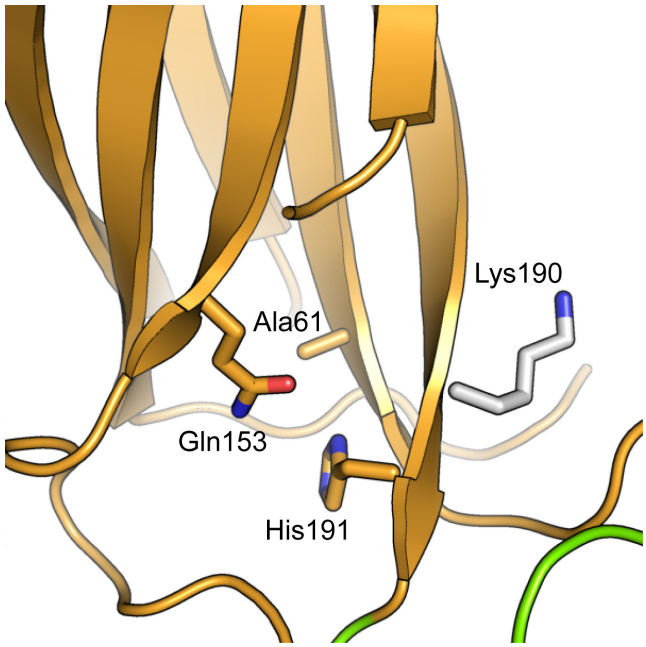

Fig. S2. Ribbon representation of SpaA N-domain. Residues of Ala-61, Gln-153, His-191, and Lys-190 of SpaA are shown in stick mode.

Table S1. Data collection, phasing, and refinement statistics

Parameters	Native SpaA	SeMet-SpaA
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell dimensions (Å, deg)	a = 34.76, b = 63.68, c = 198.11, $\alpha = \beta = \gamma = 90.0$	a = 34.78, b = 64.04, c = 199.19, $\alpha = \beta = \gamma = 90.0$
Data collection*		
Resolution range (Å)	40.00–1.60 (1.69–1.60)	40.00–1.80 (1.90–1.80)
Wavelength (Å)	0.773	0.980
Total reflections	745,330 (54,546)	545,315 (43,503)
No. of unique reflections	58,024 (7,383)	41,950 (5,546)
Redundancy	12.8 (7.4)	6.9 (4.1)
Completeness (%)	97.7 (87.1)	98.2 (89.7)
Mean $I/\sigma(I)$	13.9 (2.6)	23.8 (4.2)
R_{sym} (%)	11.4 (50.3)	6.1 (35.1)
Phasing statistics for SeMet-SpaA		
Resolution range (Å)	40.00–1.80	
Figure of Merit (acentric/centric)	0.45/0.24	
Phasing power (anomalous)	1.47	
Refinement statistics for Native SpaA		
Resolution range (Å)	40.00–1.60 (1.64–1.60)	
No. of reflections	54,856 (3,418)	
No. of reflections for R_{free} test set	2,898 (157)	
R_{work}/R_{free} (%)	19.3/22.0 (27.1/29.0)	
RMS deviations from standard bond lengths/angles (Å/°)	0.014/1.53	
Average B-factor (Å^2)	12.57	
Protein atoms	3,246	
Ions	1 Ca^{2+}, 1 I^-	
Water molecules	571	
Ramachandran plot by MOLPROBITY (1)		
Most favored/allowed (%)	98.6/100.0	
Generously allowed or outliers (%)	0.0	

*Items in parentheses are for the outermost shell of data.

1. Richardson JS, Arendall WB, Richardson DC (2003) New tools and data for improving structures, using all-atom contacts. *Methods Enzymol* 374:385–412.

Table S2. MS/MS of a peptide at *m/z* 996.8^{3+} containing Lys199-Asn321 isopeptide bond of SpaA

Observed *m/z**	Charge	Calculated *m/z*†	$\Delta_{obs\text{-}calc}$‡	Proposed structure	Ion type
189.08	+1	189.12	−0.04	AV	y_2
205.12	+1	205.10	0.02	W	$y_{1'}$
266.09	+1	266.12	−0.03	HQ	b_2
290.17	+1	290.17	0	TAV	y_3
299.11	+1	299.13	−0.02	*PSN*§	Internal
333.14	+1	333.16	−0.02	QW	$y_{2'}$
337.16	+1	337.16	0	HQA	b_3
398.21	+1	398.20	0.01	*PTAQ*§	Internal
404.19	+1	404.19	0	AQW	$y_{3'}$
450.28	+1	450.25	0.03	HQAL	b_4
537.31	+1	537.28	0.03	HQALS	b_5
568.28	+1	568.27	0.01	*PSNPTA*§	Internal
602.33	+1	602.30	0.03	PTAQW	$y_{5'}$
666.28	+1	666.32	−0.04	HQALSE	b_6
696.34	+1	696.33	0.01	*PSNPTAQ*§	Internal
716.31	+1	716.34	−0.03	NPTAQW	$y_{6'}$
810.41	+2	810.40	0.01	HQALSEPVKTAV and DNQ (-NH_3)¶	Parent-$y_{12'}$
846.46	+2	846.43	0.03	HQALSEPVKTAV and DNQA (-NH_3)¶	Parent-$y_{11'}$
900.51	+1	900.42	0.09	PSNPTAQW	$y_{8'}$
939.48	+2	939.47	0.01	HQALSEPVKTAV and DNQAW (-NH_3)¶	Parent-$y_{10'}$
989.09	+2	989.01	0.08	HQALSEPVKTAV and DNQAWV (-NH_3)¶	Parent-$y_{9'}$
1,045.66	+2	1045.55	0.11	HQALSEPVKTAV and DNQAWVL (-NH_3)¶	Parent-$y_{8'}$

*Monoisoptic masses of observed ions.
†Calculated ions. Monoisotopic masses were calculated using the Fragment Ion Calculator (http://db.systemsbiology.net:8080/proteomicsToolkit/FragIonServlet.html).
‡Difference between observed ion mass and calculated ion mass.
§Internal ions are shown in italics.
¶Loss of 17 Da from losing NH_3 is shown in parentheses.

Table S3. MS/MS of a peptide at *m/z* 751.7^{3+} containing Lys363-Asn482 isopeptide bond of SpaA.

Observed *m/z**	Charge	Calculated *m/z*†	$\Delta_{obs-calc}$‡	Proposed structure	Ion type
147.12	+1	147.08	0.04	GA	y_2
147.12	+1	147.11	0.01	K	$y_{1'}$
155.07	+1	155.08	−0.01	*PG*§	Internal
177.11	+1	177.10	0.01	FG	b_2
205.12	+1	205.10	0.02	FG	a_2
244.12	+1	244.13	−0.01	PGA	y_3
244.12	+1	244.15	−0.03	NK (-NH_3)¶	$y_{2'}$ (-NH_3)
333.10	+1	333.15	−0.05	FGQ	b_3
345.15	+1	345.18	−0.03	TPGA	y_4
418.25	+1	418.25	0	FGQI	a_4
446.26	+1	446.24	0.02	FGQI	b_4
472.27	+1	472.27	0	IDNK (-NH_3)¶	$y_{4'}$ (-NH_3)
516.26	+1	516.24	0.02	GNTPGA	y_6
547.33	+1	547.29	0.04	FGQIT	b_5
631.26	+1	631.27	−0.01	DGNTPGA	y_7

*Monoisoptic masses of observed ions.
†Calculated ions. Monoisotopic masses were calculated using the Fragment Ion Calculator (http://db.systemsbiology.net:8080/proteomicsToolkit/FragIonServlet.html).
‡Difference between observed ion mass and calculated ion mass.
§Internal ions are shown in italics.
¶Loss of 17 Da from losing NH_3 is shown in parentheses.

Table S4. Daughter ions produced during MS/MS of a peptide (*m/z* 585.7^{3+}) that was generated from trypsin and AspN digestion of SpaA pili (this peptide contains the intersubunit isopeptide bond between SpaA Lys190 and Thr494)

Observed *m/z**	Charge	Calculated *m/z*†	$\Delta_{obs-calc}$‡	Proposed structure	Ion type
106.05	+1	106.05	0	S	y_1
187.15	+1	187.11	0.04	DV	a_2
211.14	+1	211.16	−0.02	*PL*§	Internal
215.11	+1	215.10	0.01	DV	b_2
215.11	+1	215.14	−0.03	EL	$a_{2'}$
219.13	+1	219.13	0	LS	y_2
243.13	+1	243.13	0	EL	$b_{2'}$
312.22	+1	312.20	0.02	*PLT*§–H_2O	Internal
340.22	+1	340.19	0.03	ELP	$b_{3'}$
352.15	+1	352.16	−0.01	DVH	b_3
418.22	+1	418.23	−0.01	QALS	y_4
451.21	+1	451.23	−0.02	DVHV	b_4
472.57	+3	472.58	−0.01	DVHVYPKHQ and PLT (- H_2O)¶	Parent-y_3-$b_{2'}$
496.26	+3	496.27	−0.01	DVHVYPKHQA and PLT (- H_2O)¶	Parent-y_2-$b_{2'}$
533.97	+3	534.96	0.01	DVHVYPKHQAL and PLT (- H_2O)¶	Parent-y_1-$b_{2'}$
555.32	+1	555.29	0.03	HQALS	y_5
568.99	+3	568.97	0.02	DVHVYPKHQALS and PLT (- H_2O)¶	Parent-$b_{2'}$
614.33	+1	614.29	0.04	DVHVY	b_5

*Monoisoptic masses of observed ions.
†Calculated ions. Monoisotopic masses were calculated using the Fragment Ion Calculator (http://db.systemsbiology.net:8080/proteomicsToolkit/FragIonServlet.html).
‡Difference between observed ion mass and calculated ion mass.
§Internal ions are shown in italics.
¶Loss of 18 Da from losing H_2O is shown in parentheses