Supporting Information

Campos et al. 10.1073/pnas.1001703107

SI Materials and Methods

Search Strategy for Pilus Modeling. After random placement of the protomer parallel to the pilus axis, the minimization proceeded in a three-stage protocol. In all three stages, the symmetry of the pilus was enforced by using the noncrystallographic symmetry operators describing the symmetry deduced from scanning transmission EM measurements, by using the NCS STRICT option in CNS software. The first stage consisted of fast optimization of the geometry and the packing, with a simplified nonbonded interaction (repulsive van der Waals interaction only). The second stage was a refinement *in vacuo*, using adapted nonbonded parameters (a distance-dependent dielectric, a switching function between 2 and 9 Å, and a nonbonded cutoff of 10 Å). The third stage was a short refinement in explicit water, similar to the one used in NMR structure determination (1), without any conformational restraints. We used a water layer of 10-Å thickness, a nonbonded cutoff of 12 Å, and the TIP3P water model (2). After the second stage, the pilus structures were compared and clustered using an rmsd criterion. For this process, the pilins were not superimposed but only were placed on an equivalent position within the pilus (i.e., the distance from the pilus axis and the orientation of the protomer remained unchanged). In this way, the rmsd criterion includes differences in the subunit structure and in the distance from the pilus axis and in the orientation of the pilin.

Use of Conformational Restraints. In addition to the constraint on the overall geometry of the pilus, the following restraints were used:

(*i*) The initial structures were maintained in a flexible and adaptive way using log-harmonic distance restraints and automated weighting (3) on alpha carbons of residues 3–131.

(*ii*) In the first stage of modeling, a linear restraint potential (i.e., a force independent of the distance to the pilus axis) was used to maintain the protomers close to the pilus axis.

(*iii*) To improve packing of the charged side chains, an adaptive and ambiguous potential (log-harmonic with automated weighting) was used as an attractive interaction between a charged residue and all residues of opposite charge on any protomer in the pilus.

(*iv*) In some of the calculations, an ambiguous restraint was used to incorporate the experimentally validated salt bridge between residues Asp^{48} and Arg^{87}. This restraint was active in stages 1 and 2 of the minimization protocol.

(*v*) In some of the calculations, a linear restraint (a force independent of distance) was used to incorporate the cysteine crosslinking data. This restraint was active only in stages 1 and 2 of the minimization protocol.

Structure Validation. Structures of the pilus were validated by calculating the overall energy (see below) and with PROCHECK (4). No bad nonbonded contacts were identified.

Energy Evaluation. The energy of the total structure and the energetic contributions of single residues were evaluated using the generalized Born electrostatic options in CNS software modified for symmetric systems (5). The energetic contribution of a side chain was evaluated as the difference in interaction energy of a residue with the rest of the protomer and a pilus with the protomer placed at 100 Å from the pilus axis (no interaction between protomers). The surface difference was evaluated separately by explicitly modeling a short stretch of the pilus (the central protomer plus 10 protomers in each of the positive and negative x directions), also with CNS software. The nonpolar free energy difference was evaluated as the surface area difference multiplied by a constant [5 cal/(mol.$Å^2$)] (6).

1. Linge JP, Williams MA, Spronk CAEM, Bonvin AMJJ, Nilges M (2003) Refinement of protein structures in explicit solvent. *Proteins* 50:496–506.
2. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935.
3. Nilges M, et al. (2008) Accurate NMR structures through minimization of an extended hybrid energy. *Structure* 9:1305–1312.
4. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst* 26:283–291.
5. Moulinier L, Case DA, Simonson T (2003) Reintroducing electrostatics into protein X-ray structure refinement: Bulk solvent treated as a dielectric continuum. *Acta Crystallogr D Biol Crystallogr* 59:2094–2103.
6. Lafont V, Schaefer M, Stote RH, Altschuh D, Dejaegere A (2007) Protein-protein recognition and interaction hot spots in an antigen-antibody complex: Free energy decomposition identifies "efficient amino acids". *Proteins* 67:418–434.

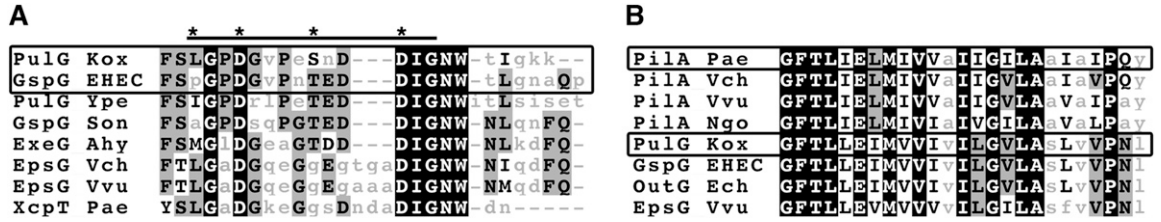


Fig. S1. (*A*) Protein sequence alignment of the C-terminal residues of major pilins from type II secretion systems (T2SS). Sequences of PulG and of GspG from enterohemorrhagic *Escherichia coli*, used as a template for the PulG modeling, are boxed. The line indicates the Ca^{2+}-binding loop region, and residues involved in the coordination of Ca^{2+} in the first two proteins are indicated by an asterisk. (*B*) N-terminal sequence alignment of four type IV (T4P) major pilins and four T2SS major pseudopilins. The sequences of PulG and PilA, used as a template for modeling the missing transmembrane segment, are boxed. Identical residues are written in white letters on a black background; conserved residues are written in black letters on a gray background; similar residues are written in black letters on a white background; and different residues are written in gray lowercase letters on a white background. Pseudopilin gene identification (gi) numbers: *EHEC*, enterohemorrhagic *Escherichia coli* (gi 254839309); *Kox*, *Klebsiella oxytoca* (gi 131595); *Ype*, *Yersinia pestis* (gi 167466636); Son, *Shewanella oneidensis* (gi 24371769); Ahy, *Aeromonas hydrophyla* (gi 117620343); *Vch*, *Vibrio cholerae* (gi 153802200); *Vvu*, *Vibrio vulnificus* (gi 254839341;); *Pae*, *Pseudomonas aeruginosa* (gi 15598297); *Ech*, *Erwinia chrysanthemi* (gi 151474). Type IV pilin gi numbers: *Pae*, *Pseudomonas aeruginosa* (gi 15599721); *Vch*, *Vibrio cholerae* (gi 229369678); *Vvu*, *Vibrio vulnificus* (gi 213876590); *Ngo*, *Neisseria gonorrheae* (gi 1169708).

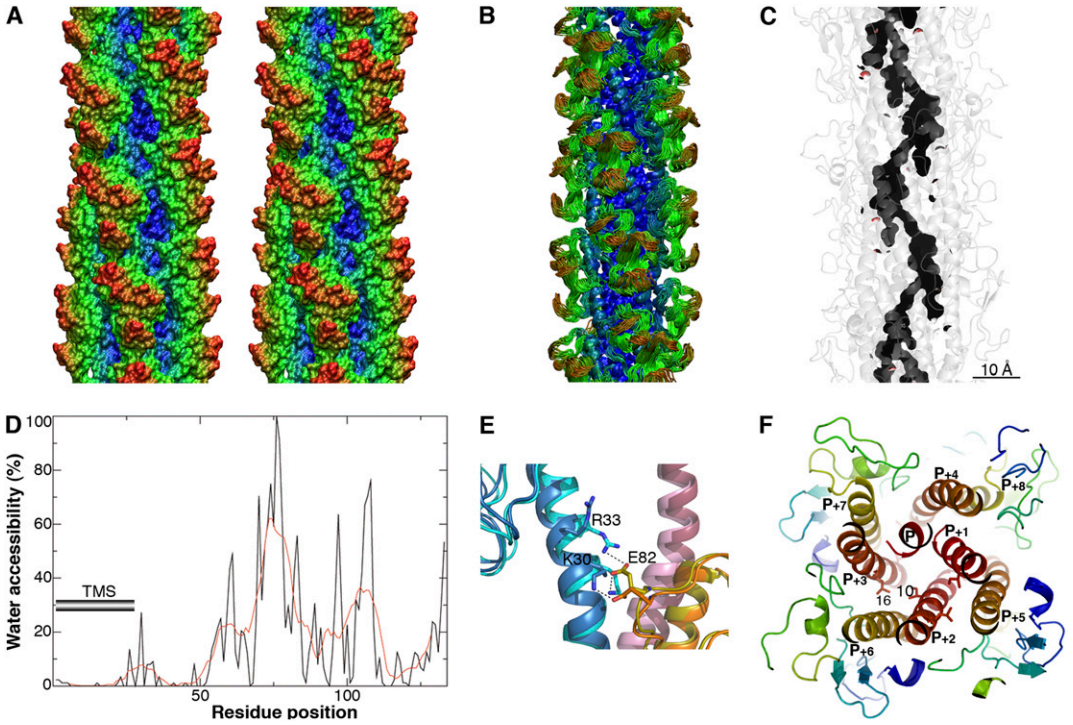


Fig. S2. Surface exposure, model robustness, and solvent accessibility of PulG residues in the pilus. (*A*) Stereo view of the pilus colored according to residue surface exposure with a blue-green-red gradient, red being the most exposed. (*B*) Superimposition of 20 models from the main cluster in ribbon view, colored as in *A*. (*C*) Internal pilus surface (solvent radius 1.4 Å) represented in black inside the ribbon view of the pilus (light gray). Water molecule does not fit in this pseudocavity. (*D*) Graphic representation of the solvent accessibility per residue (black line) or with a sliding 10-residue window average (red line). (*E*) P–P_{+3} interaction interface is shown in a ribbon view. Yellow (P), light blue (P_{+3}), and pink (P_{+4}) protomers show an alternative conformation of orange, blue, and purple protomers P, P_{+3}, and P_{+4}, respectively. (*F*) A 30-Å slice of pilus model in ribbon view colored as a rainbow, from red (N terminus) to blue (C terminus). Protomer numbers and residues 10 (Ile) and 16 (Leu) are indicated.

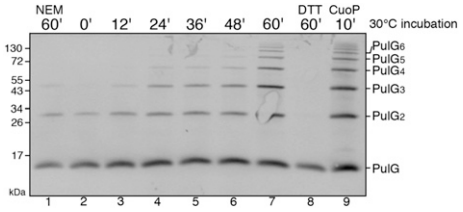
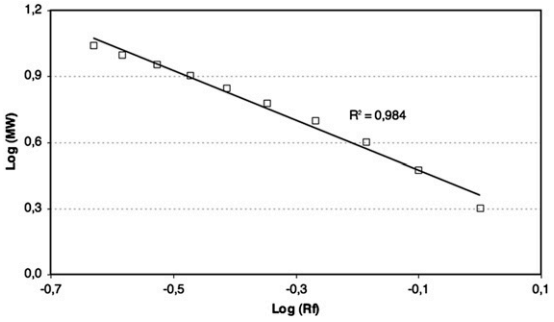


Fig. S3. Time course of PulG$^{I10C/L16C}$ pili cross-linking. After shearing, pili fractions were incubated at 30 °C with constant agitation for the indicated times, treated with 10 mM N-ethylmaleimide (NEM), 100 mM DTT, or 0.1 mM Cu-orthophenanthroline (CuoP) when appropriate. Samples were analyzed as described in Fig. 3.



Multimer	Log(Rf)	Log(MW)
11	-0,63	1,041
10	-0,584	1
9	-0,527	0,954
8	-0,473	0,903
7	-0,412	0,845
6	-0,347	0,778
5	-0,268	0,699
4	-0,185	0,602
3	-0,1	0,477
2	0	0,301

Fig. S4. (*Left*) Logarithm of the relative mobility (Rf) of PulG$^{I10C/L16C}$ cross-linked products on a 4–20% gradient SDS-polyacrylamide gel as a function of the logarithm of the number of cross-linked subunits. The linear relationship in the plot indicates that the relative mobility of each band corresponds to its assigned subunit stoichiometry. (*Right*) The coordinate values for each point are given in the table.

Table S1. Salt-bridge conservation in major T2SS pseudopilins

Salt bridge	Neg–pos (%)	Pos–neg (%)	Polar (%)	Nonpolar (%)	Same charge (%)
Asp^{48}–Arg^{87}	81.1	5.7	9.6	2.8	0.7
Glu^{44}–Arg^{88}	66.9	1.4	6.0	22.8	2.8

Salt-bridge conservation based on the alignment of the 281 nonredundant sequences annotated as major pseudopilin in the NCBI protein database (without taking into account the pseudopilin sequences annotated as "putative"). For the salt bridges Asp^{48}–Arg^{87} and Glu^{44}–Arg^{88}, percentages of pairs of residues at the corresponding positions are given. Neg–pos, negative charge at the first position and positive at the second position; Pos–neg, positive charge at the first position and negative charge at the second position; polar, at least one polar residue; nonpolar, at least one nonpolar residue; same charge, both residues have the same charge.

Table S2. *Escherichia coli* K-12 strains used in this study

Strain	Genotype	Source/reference
PAP105	Δ(*lac-pro*) pAPIP501 (F′ (*lacI*^{Q} Δ*lacZ*M15 *pro*+Tn*10*))(Tc^{R})	Laboratory collection
PAP7460	HS2019 (pAPIP501) *ΔmalE444 malG501* F′ (*lacI*^{Q} *ΔlacZM15 pro*+ Tn*10*)(Tc^{R})	(1)
PAP9304	PAP7460 pCHAP1216	Laboratory collection
PAP5299	*araD139 (argF-lac)U169 rpsL150 relA1 flb5301 deoC1ptsF25 thi pcnB::*Tn*10* (F′ *lacI*^{Q})	Laboratory collection
PAP5300	PAP5299 pCHAP8184	This study

1. Possot OM, Vignon G, Bomchil N, Ebel F, Pugsley AP (2000) Multiple interactions between pullulanase secreton components involved in stabilization and cytoplasmic membrane association of PulE. *J Bacteriol* 182:2142–2152.

Table S3. Plasmids carrying cloned *pul* genes used in this study

Plasmid	Vector/replicon /marker	Cloned genes	Source or reference
pSU19	p15A/Cm^{R}	—	(1)
pCHAP231	pBR322/ColE1/Ap^{R}	*pulS, pulAB pulCDEFGHIJKLMN*	(2)
pCHAP1216	pBR322/ColE1/Ap^{R}	pCHAP231 *pulB::kan1 ΔpulG1*	(3)
pCHAP8184	pBR322/ColE1/Ap^{R}	pCHAP231 *pelBsp-pulA ΔpulG1*	This study
pCHAP7303	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*	This study
pCHAP7304	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(E5A)	This study
pCHAP7308	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(P22A)	This study
pCHAP7312	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(K28E)	This study
pCHAP7354	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(I10C/L16C)	This study
pCHAP7357	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(E5D)	This study
pCHAP7358	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(E5Q)	This study
pCHAP7360	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(R78D)	This study
pCHAP7364	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(D48K)	This study
pCHAP7367	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(R87E)	This study
pCHAP7368	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(E82K)	This study
pCHAP7369	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(R56E)	This study
pCHAP7639	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(E29A/K51E)	This study
pCHAP7370	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(D117R)	This study
pCHAP7371	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(D124R)	This study
pCHAP7372	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(E29R)	This study
pCHAP7373	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(K51E)	This study
pCHAP7374	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(E44R)	This study
pCHAP7375	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(Y85F)	This study
pCHAP7354	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(I10C/L16C)	This study
pCHAP7611	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(I6C/I12C)	This study
pCHAP7612	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(I6C/L13C)	This study
pCHAP7613	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(M7C/I12C)	This study
pCHAP7614	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(M7C/L13C)	This study
pCHAP7615	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(V9C/V15C)	This study
pCHAP7616	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(V9C/L16C)	This study
pCHAP7617	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(I10C/V15C)	This study
pCHAP7618	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(I10C/A17C)	This study
pCHAP7619	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(V11C/L16C)	This study
pCHAP7620	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(V11C/A17C)	This study
pCHAP7621	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(D48R87/KD)	This study
pCHAP7624	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(I10C)	This study
pCHAP7641	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(R88D)	This study
pCHAP7642	pSU19/p15A/Cm^{R}	lacZp-FLAG::*pulG*(E44R/R88D)	This study

1. Bartolomé B, Jubete Y, Martinez E, de la Cruz F (1991) Construction and properties of a family of pACYC184-derived vectors compatible with pBR322 and its derivatives. *Gene* 102: 75–78.
2. d'Enfert C, Ryter A, Pugsley AP (1987) Cloning and expression in *Escherichia coli* of the *Klebsiella pneumoniae* genes for production, surface localization and secretion of the lipoprotein pullulanase. *EMBO J* 6:3531–3538.
3. Possot OM, Vignon G, Bomchil N, Ebel F, Pugsley AP (2000) Multiple interactions between pullulanase secreton components involved in stabilization and cytoplasmic membrane association of PulE. *J Bacteriol* 182:2142–2152.

Table S4. Mutagenic oligonucleotides used in this study

Name	Oligonucleotide sequence
G-L16C-3	5′ C CTC GGC GTG TGC GCA AGC CTC GTG GTG 3′
G-E5A-5	5′ GGA TTC ACC TTA CTG GCC ATC ATG GTG G 3′
G-E5Q-5	5′ GGA TTC ACC TTA CTG CAG ATC ATG GTG G 3′
G-E5D-5	5′ GGA TTC ACC TTA CTG GAT ATC ATG GTG G 3′
G-P22A-5	5′ C AGC CTC GTG GTG GCC AAC CTG ATG GGC AAC 3′
G-K28E-5	5′ C CTG ATG GGC AAC GAG GAG AAG GCC GAC C 3′
G-E29A-5	5′ CTG ATG GGC AAC AAG GCA AAG GCC GAT CGG CAA AAA GTG 3′
G-E29R-5	5′ G ATG GGC AAC AAG CGC AAG GCC GAT CGG CA 3′
G-E44R-5	5′ C GAT CTT GTA GCG CTC CGC GGC GCG CTG GAC 3′
G-D48K-5	5′ G CTC GAG GGC GCG CTT AAG ATG TAC AAA CTG G 3′
G-K51E	5′ G GAC ATG TAC GAG CTC GAC AAC AGC 3′
G-R56E-5	5′ G GAC AAC AGC GAG TAT CCC ACC AC 3′
G-R78D-5	5′ C GAG CCG CAT GCA GAC AAC TAT CCC GAA GG 3′
G-E82K-5	5′ C GAG CCG CAT GCG CGC AAC TAT CCC AAA GGC GGC TAT ATC C 3′
G-Y85F-5	5′ GC AAC TAT CCC GAA GGC GGA TTC ATC CGC CGT CTG CCG CAG 3′
G-R87E-3	5′ GTC CTG CGG CAG ACG TTC GAT ATA GCC GCC TTC G 3′
G-R88D-3	5′ TG CCC CAG GGA TCC TGC GG CAG ATC GCG GAT ATA GCC GCC T 3′
G-D117R-3	5′ GTC CGG CAC GCC GCG GGG CCC GAG GCT G 3′
G-D124-3	5′ CA GTT GCC GAT ATC GCG ATT GCT CTC CGG CA 3′
G-I6C-5	5′ C ACC TTA CTG GAA TGC ATG GTG GTG ATT G 3′
G-M7C-5	5′ CC TTA CTG GAG ATC TGC GTG GTG ATT GTT ATC 3′
G-I12C-3	5′ GAG GCT GGC GAG TAC TCC GAG GCA AAC AAT CAC CAC 3′
G-L13C-3	5′ GAG CAC GCC GCA TAT GAC AAT CAC C 3′
G-V9C-5	5′ G GAA ATC ATG GTG TGC ATT GTT ATC CTC 3′
G-I10C-5	5′ C ATG GTG GTG TGC GTT ATC CTC GGC GTG 3′
G-V11C-5	5′ G GTG GTT ATT TGC ATC CTC GGC GTG 3′
GV15C-3	5′ GCT AGC GAG ACA GCC GAG GAT AAC AAT C 3′
G-L16C-3	5′ C CTC GGC GTG TGC GCA AGC CTC GTG GTG 3′
G-A17C-3	5′ CAC CAC GAG GCT GCA GAG CAC GCC GAG G 3′
M13-Fw	5′ TGTAAAACGACGGCCAGT 3′

Substituted nucleotides are underlined.

Movie S1. Modeling run of the T2S pilus. Each frame represents a step in the modeling process with all of the implicit protomers represented by symmetry. Protomers are attracted toward the pilus axis, are arranged first in a left-handed helix, and then converge toward one of the right-handed helix models in the main cluster.

Movie S1