



HAL
open science

Structure of the calcium dependent type 2 secretion pseudopilus

Aracelys Lopez-Castilla, Jenny-Lee Thomassin, Benjamin Bardiaux, Weili Zheng, Mangayarkarasi Nivaskumar, Xiong Yu, Michael Nilges, Edward Egelman, Nadia Izadi-Pruneyre, Olivera Francetic

► **To cite this version:**

Aracelys Lopez-Castilla, Jenny-Lee Thomassin, Benjamin Bardiaux, Weili Zheng, Mangayarkarasi Nivaskumar, et al.. Structure of the calcium dependent type 2 secretion pseudopilus. *Nature Microbiology*, 2017, 10.1038/s41564-017-0041-2 . pasteur-02283219

HAL Id: pasteur-02283219

<https://pasteur.hal.science/pasteur-02283219>

Submitted on 10 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

1 **Structure of the calcium dependent type 2 secretion pseudopilus**

2

3 Aracelys Lopez-Castilla^{1,5}, Jenny-Lee Thomassin^{2,5}, Benjamin Bardiaux^{3,5}, Weili Zheng⁴,
4 Mangayarkarasi Nivaskumar², Xiong Yu⁴, Michael Nilges³, Edward H. Egelman⁴, Nadia
5 Izadi-Pruneyre^{1*}, Olivera Francetic^{2*}

6

7 ¹NMR of Biomolecules Unit, Department of Structural Biology and Chemistry, Institut
8 Pasteur, CNRS UMR3528, 28 rue du Dr Roux, 75724 Paris, France

9 ²Biochemistry of Macromolecular Interactions Unit, Department of Structural Biology and
10 Chemistry, Institut Pasteur, CNRS ERL6002, 28 rue du Dr Roux, 75724 Paris, France.

11 ³Structural Bioinformatics Unit, Department of Structural Biology and Chemistry, CNRS
12 UMR3528, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris, France

13 ⁴Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville,
14 VA22908, USA

15 ⁵Equal contribution.

16 *Corresponding authors.

17

18 **Keywords:** type 2 secretion system, solution NMR, cryoEM, calcium, protein folding, type 4
19 filaments

20

21 **Abstract**

22 Many Gram-negative bacteria use type 2 secretion systems (T2SS) to secrete proteins
23 involved in virulence and adaptation. Transport of folded proteins *via* T2SS nanomachines
24 requires the assembly of inner membrane-anchored fibers called pseudopili. Although
25 efficient pseudopilus assembly is essential for protein secretion, structure-based functional
26 analyses are required to unravel the mechanistic link between these processes. Here, we report
27 an atomic model for a T2SS pseudopilus from *Klebsiella oxytoca*, obtained by fitting the
28 NMR structure of its calcium-bound subunit PulG into the $\sim 5 \text{ \AA}$ resolution cryo-electron
29 microscopy (cryoEM) reconstruction of assembled fibers. This structure reveals the
30 comprehensive network of inter-subunit contacts and unexpected features, including a
31 disordered central region of the PulG helical stem, and highly flexible C-terminal residues on
32 the fiber surface. NMR, mutagenesis and functional analyses highlight the key role of calcium
33 in PulG folding and stability. Fiber disassembly in the absence of calcium provides a basis for
34 pseudopilus length control, essential for protein secretion, and supports the Archimedes'
35 screw model for T2S mechanism.

36

37 **Main text**

38 Microorganisms profoundly transform their environment by colonizing biotic and abiotic
39 surfaces and assimilating nutrients from organic and mineral sources. In many prokaryotes,
40 nanomachines of the type 4 filament (Tff) superfamily¹ mediate adaptive functions by
41 promoting cell adhesion, motility and macromolecular transport. The Tff assembly systems,
42 including archaeal flagella², bacterial type 4 pili (T4P) and T2SSs³, use a conserved
43 mechanism to polymerise fibres from protein subunits initially localised in the plasma
44 membrane. In Gram-negative bacteria T2SSs promote assembly of periplasmic fibres called

45 pseudopili to drive secretion of specific folded proteins, including macromolecule hydrolases,
46 toxins, adhesins or cytochromes with key roles in virulence and niche survival⁴.

47 T2SSs are large trans-envelope nanomachines that coordinate the activity of a
48 cytoplasmic ATPase and inner membrane assembly sub-complex to polymerize pseudopilin
49 subunits into helical fibres (Fig. 1a). Pseudopili are composed mainly of major pseudopilin,
50 called PulG in the prototypical *Klebsiella oxytoca* pullulanase T2SS⁵. Four minor
51 pseudopilins PulH, I, J and K, likely cap the pseudopilus tip and promote efficient initiation
52 of fibre assembly^{6,7}. Pseudopilus assembly is coupled to secretion of folded proteins from the
53 periplasm across the dedicated outer membrane channel to the extracellular milieu⁸, however
54 the mechanism of this coupling remains unknown.

55 Structural information is essential to understand the secretion process at molecular and
56 atomic levels. The first crystal structure of the major pseudopilin PulG from *K. oxytoca* T2SS
57 showed a typical α - β pilin fold in a crystallographic dimer where the C-terminal β strand was
58 swapped between two monomers⁹. Subsequent crystallographic studies of PulG homologues
59 from *Vibrio cholerae*, *V. vulnificus* and enterohaemorrhagic *Escherichia coli* revealed local
60 structural differences at the C-terminus including a long loop harbouring a calcium atom¹⁰.
61 Substitutions of two calcium-coordinating residues in the *V. cholerae* pseudopilin EpsG
62 abolished protein secretion¹⁰, suggesting an essential role for calcium in T2SS function.

63 To gain insight into the native pseudopilus structure and conformational dynamics
64 during protein secretion, we solved the structure of the periplasmic PulG domain (PulGp)
65 using NMR. We combined these data with a cryoEM map of purified PulG pili to generate
66 T2SS pseudopilus atomic model. Furthermore, we characterized the role of calcium in PulG
67 folding, stability and pseudopilus assembly *in vivo* and *in vitro*. Our structural and functional
68 data suggests a role of calcium in pseudopilus folding, assembly and length control, with
69 general implications for T2S mechanism.

70

71 **Results**

72 **Calcium is required for PulG monomer and polymer stability**

73 The calcium-binding site identified in the periplasmic domain of major pseudopilins from
74 *Vibrio* and *E. coli* is highly conserved in its homologues, including PulG¹⁰. To test the effect
75 of calcium depletion on PulG pilus assembly, an *E. coli* K-12 strain expressing all the *pul*
76 genes encoding the *K. oxytoca* T2SS was grown in the presence of increasing concentrations
77 of EGTA (Fig. 1b, WT). PulG cellular levels were dramatically reduced at EGTA
78 concentrations of 1 mM or higher and pseudopilus assembly was abolished (Fig. 1b). As the
79 *pulM* gene is located downstream of *pulG* in the same operon and encodes a T2SS component
80 with the same topology as PulG (Fig. 1a), PulM levels were used as a control for *pul* gene
81 expression. As expected, PulM was unaffected by the addition of EGTA. In the presence of 2
82 mM EGTA, PulG detection was rescued upon addition of Ca²⁺, but not of Mg²⁺ or K⁺ (Fig.
83 1c). The PulG-FLAG variant was as sensitive to EGTA as PulG^{WT}, excluding the possibility
84 that the lack of PulG detection was due to disruption of crucial epitopes recognized by anti-
85 PulG antibodies (Fig. 1d). Addition of Ca²⁺ to the media restored piliation (Supplementary
86 Fig. 1). Together, these data indicate that maintaining *in vivo* cellular levels of PulG depends
87 on the presence of calcium and suggest that calcium depletion leads to PulG degradation.

88 Calcium depletion also affected the stability of PulG pseudopilus fibres *in vitro* (Fig.
89 1e). Upon incubation of purified pseudopili in buffer or buffer supplemented with Ca²⁺ for 1.5
90 or 18 hours, intact fibres were recovered in the pellet fraction following ultracentrifugation. In
91 contrast, PulG was found in the supernatant fraction of pseudopili incubated in the presence
92 of EGTA, indicating pseudopilus disassembly. Partial disassembly was observed for pili
93 incubated for 18 hours in buffer, suggesting an exchange and gradual release of calcium. The
94 disassembly was more dramatic after a long incubation with EGTA, whereas pseudopili

95 remained intact in buffer supplemented with Ca^{2+} . Together these data demonstrate that
96 calcium is required for PulG stability *in vivo* and for the pseudopilus stability *in vitro*.

97

98 **Essential role of calcium in PulG folding and stability**

99 To understand the effect of calcium on PulG stability, solution NMR was used to analyse the
100 structural features of the protein with and without calcium. For this study, we used the soluble
101 periplasmic domain of PulG (residues 25-134), designated PulGp. The NMR spectral
102 signature (^{15}N - ^1H HSQC) of PulGp purified from the periplasm was unchanged after the
103 addition of 1 mM calcium (Fig. 2a). In contrast, significant spectral changes were observed in
104 the presence of 20 mM EGTA (PulGp-EGTA) (Fig. 2a). Addition of 1 mM calcium to the
105 PulGp-EGTA sample, after EGTA was removed, restored the initial PulGp spectrum
106 (Supplementary Fig. 2). These observations indicate that PulGp purified from the periplasm is
107 calcium-bound and that this native folding is modified following calcium chelation by EGTA.

108 Further analyses comparing the ^{15}N - ^1H HSQC spectrum of native PulGp with that of
109 PulGp-EGTA highlight the structural role of calcium. The ^{15}N - ^1H HSQC spectrum of native
110 PulGp displayed good signal dispersion, typical of a well-structured protein. The number of
111 peaks corresponded to those expected based on the protein sequence, indicating that only one
112 conformer exists in the native condition. After calcium chelation by EGTA, the PulGp
113 spectrum showed reduced signal dispersion and an increased number of peaks. The additional
114 peaks appeared mainly between 7.7 and 8.7 ppm (Fig. 2a), indicating the presence of
115 disordered regions in slow exchange on the chemical shift timescale between different
116 conformations. To further characterize PulGp behaviour in the calcium-free state, its
117 backbone resonances were assigned and used to estimate their secondary structure contents.
118 Two conformers were identified (Fig. 2a). The first one with its backbone resonances almost
119 entirely assigned (at 94%) displayed a slight decrease of secondary structure content in favour

120 of random coil (Supplementary Fig. 3a). The second conformer (assigned at 32%) is mostly
121 disordered (Supplementary Fig. 3b). The propensity of PulGp to become partially unfolded or
122 disordered in the absence of calcium shows the importance of this metal in protein folding
123 and structural stability. In support of this conclusion, PulGp was thermostatically more stable
124 in the presence of calcium, since its T_m increased by 17.2 °C, from 34°C in the calcium-free-
125 to 51.8°C in the calcium-bound state (Supplementary Fig. 3c).

126

127 **PulG calcium-binding site**

128 To identify PulG residues affected by calcium, we analysed the chemical shift perturbation
129 (CSP) of its amide resonances between the calcium-bound and -free states. C-terminal
130 residues S113, L114, G115, D117, V119, D124, D125, I126 and G127 displayed the largest
131 CSP (> 0.2 ppm, Fig. 2b), indicating that their chemical environment was modified either by
132 their interaction with calcium or by conformational changes that were induced by the
133 presence of calcium. Residues E121 and S122 in this region were also significantly perturbed,
134 although their signals could not be unambiguously assigned in the calcium-free state.

135 To determine the effect of calcium binding *in vivo*, several PulG residues with high
136 (113, 114, 117, 119, 124, 125), intermediate (131, 132), low (128, 129, 130, 133, 134) and
137 undetermined (122) CSP values were individually substituted by an alanine. We analysed
138 bacterial and pili fractions of strains containing PulG or PulG variants grown under pilus-
139 inducing conditions (Fig. 2c). PulG levels were significantly affected by alanine substitutions
140 of residues 113, 117, 122, 124, 125, 129 indicating their requirement for PulG stability *in vivo*
141 and suggesting either direct involvement in calcium coordination or in intra-molecular
142 contacts essential for PulG folding and stability.

143

144 **Solution structure of PulGp in the calcium-bound state**

145 Given the important role of calcium for native PulG folding in the periplasm, we solved the
146 structure of PulGp in its calcium bound state. For this structure calculation, 2239 distance
147 restraints, 44 hydrogen bonds and 156 dihedral angles were used. The details of restraints and
148 structural characteristics of the family of 15 conformers representing the solution structure of
149 PulGp are summarized in Supplementary Table 1. The structures display good convergence
150 with a mean pairwise root mean square deviation (RMSD) of 0.5 and 1 Å for the backbone
151 and the heavy atoms of ordered regions, respectively (Fig. 3a).

152 PulGp is a monomer in solution as indicated by its overall correlation time (7.8 ns). Its
153 structure includes an N-terminal α -helix (α 1, residues 29-54), a long $\alpha\beta$ -loop (55-98)
154 including a second α -helix (α 2, residues 60-68), two short antiparallel β -strands (β 1: 99-101
155 and β 2: 110-112) and a short C-terminal helix (α 3: 128 to 131) (Fig. 3b). Near the C-
156 terminus, a long β 2- α 3 loop, well defined and convergent in the ensemble of conformers (Fig.
157 3a), harbours the calcium-binding site (Fig. 3c). It contains residues that exhibited the highest
158 CSP values (Fig. 2b). In addition, other residues sensed the effect of calcium binding, as
159 indicated by low but significant CSP values (Fig. 2b, 3c). They are located in the N-terminal
160 α -helix (K30, D32, R33, Q34, V37), the C-terminal part of the $\alpha\beta$ -loop and its junction with
161 the β 1 strand (D92, W94, S96, D97, Y98, Q99) and near the C-terminus (I131 and G132)
162 (Fig. 3c). Although residue W129 is located outside the calcium-binding loop (Fig. 3c), its
163 substitution by Ala abolished PulG detection (Fig. 2c). The proximity of W129 to H106 in the
164 PulG structure (Supplementary Fig. 4a) suggested that stacking interactions between their
165 aromatic rings flanking the β 2- α 3 loop stabilize the C-terminal region. Introducing Cys
166 substitutions in both residues led to intra-molecular disulphide bond formation in variant
167 PulG^{H106C,W129C} allowing for pseudopilus assembly *in vivo* (Supplementary Fig. 4b).

168

169 **Structure of the PulG pseudopilus**

170 Recent advances in cryo-EM have provided an unprecedented insight into bacterial
171 filaments¹¹. We used cryo-EM to analyse purified pseudopili (Fig. 4a) formed by native PulG
172 (PulG^{WT}) and by the PulG^{H106C,W129C} variant (PulG^{CC}) containing an intra-molecular
173 disulphide bond flanking the C-terminal Ca²⁺-binding loop (Supplementary Fig. 4). The
174 analysis of PulG^{WT} and PulG^{CC} fibres using the Iterative Helical Real Space Reconstruction
175 (IHRSR) method¹² led to reconstructions that converged to highly similar helical parameters:
176 a rise of 10.2 Å for both variants and twist angles of 83.5° and 83.2° for PulG^{WT} and PulG^{CC},
177 respectively. The high degree of flexibility of PulG^{WT} filaments, observed previously by
178 negative staining EM¹³, limits the resolution of the cryo-EM reconstruction. Although PulGp
179 domain density is well defined in the cryo-EM map of PulG^{WT} pili, with an estimated
180 resolution of ~7 Å, the rod-shaped density of the N-terminal α-helix fades from residue G26
181 towards the N-terminus (Supplementary Fig. 5b), precluding the construction of an atomic
182 model of this region. However, the cryo-EM map obtained for PulG^{CC} has a higher resolution
183 (~5 Å overall) (Fig. 4b-d) and the density for PulGp is extremely well-defined. The thinner
184 density of the N-terminal α-helix at the junction with the PulGp could be connected to the
185 region with characteristic helical-shaped density corresponding to the N-terminal PulG α
186 helix (Fig. 4e) and allowed for unambiguous identification of individual PulG subunits.

187 The PulG^{CC} pilus atomic model was built by refining, in the cryo-EM density map, the
188 PulGp NMR structure connected to a model of the trans-membrane segment (TMS) by an
189 extended (non-helical) linker (Fig. 4f). During refinement, the structure of PulGp was
190 preserved, with a backbone RMSD of 2.1 Å between initial and final models for residues 27
191 to 130. The region surrounding the conserved P22 (from L19 to G26) has no regular
192 secondary structure and residues 1-21 of the helical TMS fit nicely in the N-terminal density
193 (Fig 4e). In the final atomic model, the PulG^{CC} pilus has a diameter of around 70 Å (Fig. 4d)
194 and contacts between PulG subunits P and P+1, P+3 and P+4 are observed along the right-

195 handed 1-start helix (Fig. 5a). The globular domains form an outer ring whereas the TMSs are
196 packed at the pilus core and the calcium-binding loop is on the pilus surface (Fig. 5b).
197 Residues conserved in the major pseudopilin family are located in the core of the pilus while
198 surface-exposed residues are mostly non-conserved (Supplementary Fig. 6a). The inner side
199 of the PulG subunit facing the core of the pilus shows a stretch of conserved, mostly
200 hydrophobic residues (Supplementary Fig. 6b).

201 Both the PulG^{WT} and PulG^{CC} pili displayed similar subunit organization, with PulG
202 subunits forming contacts *via* the PulGp domains, the hydrophobic domains, as well as
203 between PulGp and hydrophobic domains (Fig. 5 and Supplementary Fig. 7). The cryo-EM
204 reconstruction of PulG^{WT} and PulG^{CC} supports the proximity of residues 10_{P+1} and 16_P,
205 consistent with efficient cross-linking of PulG variants with Cys at these positions (C β -C β
206 distance of 6.9 Å), but not with Cys substitutions at positions 7 and 13 (distance of 11.3 Å)
207 used as a control¹⁴ (Fig. 5d,e).

208 The atomic model of the PulG^{CC} pilus fits well into the ~7 Å cryo-EM map of PulG^{WT}
209 (cross-correlation coefficient of 0.76 after rigid docking) where the unstructured L19-G26
210 region sits in the central thin density and the helical N-terminal portion of the TMS is
211 positioned in the rod-like density at the core of the map (Supplementary Fig. 5b). Inter-
212 subunit contacts between conserved residues of TMSs and PulGp predicted by co-
213 evolutionary analysis are also observed in the structure, further supporting the PulG^{CC} atomic
214 model (Supplementary Fig. 6e). Therefore, with the exception of the disulphide bond, the
215 atomic model of PulG^{CC} is a good model for the native PulG pilus structure.

216

217 **Discussion**

218

219 We report the structures of the PulG monomer and homo-polymer that constitute the
220 core of the pseudopilus, an essential T2SS sub-complex. We demonstrate the key role of
221 calcium in pseudopilus biogenesis, folding and stability, explaining the previously described
222 role of calcium in protein secretion¹⁰.

223 The structure of the periplasmic PulG monomer was identical to that of the calcium-
224 bound state formed *in vitro*, showing that calcium induces native PulG folding and is required
225 for its stability. Calcium depletion destabilized PulG *in vitro*, yielding two different non-
226 native conformations, which coexist and exchange with a rate slower than a millisecond.
227 Remarkably, the addition of calcium to these calcium-free conformations restored the native
228 fold. Calcium-dependent PulG folding may explain differences between the NMR and X-ray
229 structures solved respectively with and without calcium. Under NMR experimental conditions
230 PulGp is a monomer, whereas it is observed as a domain swapped PulGp dimer in the crystal
231 ⁹. The major differences between the PulGp NMR and crystal structure are found along the C-
232 terminus including the calcium-binding region (RMSD of 13 Å for residues 124-133, versus 3
233 Å for the rest of the protein). In the NMR structure, the PulGp C-terminus is in a helical
234 conformation, whereas in the X-ray structure two β strands are present (Fig. 3 and
235 Supplementary Fig. 8). Strikingly, EGTA treatment resulted in the formation of a shorter loop
236 followed by a β3 strand in the more folded PulGp conformer (Supplementary Fig. 3a)
237 reminiscent of the X-ray structure.

238 When comparing available structures of major pseudopilin homologues the
239 conformation of the C-terminal part, either as α-helix or β-strand, correlates with the presence
240 of calcium^{10,15}. The orientation of the β1-β2 loop in these structures is similarly influenced by
241 the presence of calcium (Supplementary Fig. 8). Interestingly, the NMR structure of native
242 PulGp is more similar to the X-ray structures of EHEC and *V. cholerae* pseudopilins obtained
243 with calcium (DALI server Z score of 14.4 and 13.4 over 107 aligned residues) than to its

244 own X-ray structure (Z score of 11.8). While the structural effect of calcium is reversible,
245 calcium depletion and Ala substitutions of calcium-coordinating residues *in vivo* led to PulG
246 degradation and abolished T2SS function. These data support an important regulatory role of
247 calcium levels, which are known to influence expression or activity of many bacterial
248 virulence-related secretion systems¹⁶⁻¹⁸ and adherence factors^{19,20}, during adaptation to
249 specific niches or lifestyles²¹.

250 The structure of the calcium-bound form of PulGp fits well into the cryoEM map of
251 the PulG pseudopilus, and all residues can be assigned a density, except for the four C-
252 terminal residues, suggesting high flexibility. The helical symmetry of fibres formed by both
253 PulG^{WT} and PulG^{CC} is very close to previously reported values from STEM⁹ and negative
254 stain EM¹³ (Fig. 4b,c and Supplementary Fig. 9f). Compared to our previous models¹⁴, we
255 observe a major structural change involving parts of the N-terminal domain (Supplementary
256 Fig. 9a-c). The cryoEM structure shows that the central region of the pilus stem (L19-G26) is
257 non-helical and shifted downwards by ~7 Å along the pilus axis with respect to PulGp.
258 Several algorithms predict low helical propensity for this region (Supplementary Fig. 6d). A
259 similar feature of the α -helical stem was previously observed between residues 15 to 23 in the
260 major subunit PilE of meningococcal T4P²², despite having different subunit packing, and in
261 an EHEC pseudopilin crystal structure¹⁰ (Supplementary Fig. 9d,e). The X-ray structure of the
262 *Neisseria gonorrhoeae* pilin²³, which was the basis of PulG model used for our previous
263 pseudopilus model¹⁴ shows a long continuous helix, which is also present in the NMR
264 structure of the *Geobacter* pilin in solution²⁴. These structures display the all-helical
265 conformation that is likely present in the membrane prior to assembly, as also suggested by
266 PulG molecular dynamics simulations²⁵. Direct observation of this non-helical, extended
267 conformation in the pseudopilus suggests a structural modification, potentially induced during

268 the membrane extraction step of fibre assembly²⁵, whose mechanism and function remain to
269 be explored.

270 The pseudopilus cryoEM reconstruction provides the comprehensive description of
271 potential inter-protomer contacts. This structure is consistent with highly conserved
272 pseudopilus salt-bridges involving negatively charged residues (44 and 48) of protomer P and
273 positively charged residues (87 and 88) of protomer P+1, extensively validated by
274 mutagenesis¹⁴ (Fig. 5c and Supplementary Fig.6e) and essential for PulG dimer formation in
275 the membrane¹³, pseudopilus assembly and protein secretion¹⁴. Other long-range contacts
276 specifically involved in pseudopilus stability are consistent with the cryoEM structure,
277 including K30_{P+3} and D53_P¹³.

278 The high flexibility of T4P-like fibres, crucial for assembly and function, limits their
279 resolution by cryoEM. Here the PulG^{CC} variant allowed for increased resolution over the wild
280 type. Although the overall fold of PulG^{CC} subunit is conserved, the dynamics of its C-terminal
281 region constrained by the disulphide bond seems to have a global destabilizing effect since
282 PulG^{CC} fibres are less stable and, compared to the native fibres, disassemble more readily due
283 to calcium release (Supplementary Fig. 10).

284 By analogy to the closely related T4P, which extend and retract through the action of
285 two antagonistic ATPases, it has been proposed that the pseudopilus promotes protein
286 secretion in a piston-like fashion^{26,27}. However, T2SSs only have one assembly ATPase,
287 raising a question of the pseudopilus disassembly mechanism. In the alternative Archimedes'
288 screw model, pseudopilus assembly at the base is coupled to its disassembly and degradation
289 at the fibre tip¹³. T2SS overproduction and plate culture promote pseudopilus extension
290 beyond the bacterial surface^{5,28,29}. However, under liquid culture conditions that promote
291 exoprotein secretion, pseudopili are not detected on the cell surface suggesting a length
292 control mechanism. Fibre disassembly in the absence of calcium shown here provides a

293 mechanistic basis for pseudopilus length control. While continuous addition of pseudopilin
294 subunits to the base of the fibre in the membrane drives protein secretion¹³, pseudopilins at
295 the tip would be the first to encounter the extracellular environment, where calcium levels
296 could be locally lower than in the periplasm³⁰. These conditions combined with dynamic
297 interactions with T2SS components or substrates would favour subunit dissociation and
298 proteolysis³¹. The pseudopilus structure determined in this work, combined with NMR and
299 other tools to study interactions with T2SS components and substrates provides an ideal
300 framework to elucidate the protein secretion mechanism.

301

302 **Methods**

303

304 **Bacterial strains, plasmids and molecular biology techniques**

305 *E. coli* strains and plasmids used in this study are listed in Supplementary Table 2. Bacteria
306 were grown on LB media at 30°C or 37°C on LB medium supplemented, as required, with
307 100 µg.ml⁻¹ ampicillin (Ap) or 25 µg.ml⁻¹ chloramphenicol (Cm). DNA manipulations and
308 transformation were performed as described³². Modified Quick-Change method and primers
309 listed in Supplementary Table 2 were used for site-directed mutagenesis. All constructs were
310 verified by DNA sequencing (Eurofins and GATC).

311

312 **Pilus assembly and stability assays**

313 Pseudopilus assembly assays were performed as described previously. Briefly, *E. coli* strain
314 PAP7460 containing indicated plasmids were grown at 30°C for 48 hr on LB agar containing
315 0.2% D-maltose, Ap and Cm¹⁴. Proteins were analysed by denaturing polyacrylamide gel
316 electrophoresis (PAGE) on 10% Tris-Tricine gels³³, transferred onto nitrocellulose
317 membranes (ECL, Amersham) and processed for immunoblot. Custom-made polyclonal anti-

318 PulG (Genscript) and anti-PulM antisera were used as described³⁴. Monoclonal anti-FLAG
319 M2 antibodies (Sigma F-3165) were used at a working concentration of 0.8 $\mu\text{g}\cdot\text{ml}^{-1}$.
320 Secondary goat anti-rabbit or anti-mouse secondary antibodies coupled to HRP (GE) were
321 diluted to 1:20000. Blots were revealed by ECL2 (Thermo) and fluorescence signal was
322 recorded on Typhoon FLA7000 phosphorimager (GE). Images were processed using Adobe
323 Photoshop.

324 Pseudopili were purified from *E. coli* strains PAP7460 containing appropriate
325 plasmids (Supplementary Table 2) as described⁹. Pili were stored in HS buffer (50 mM
326 HEPES pH 7.2, 50 mM NaCl) at 4°C and used for cryoEM or stability assays. For pilus
327 stability experiments, HS buffer was supplemented with 5 mM CaCl_2 or 5 mM EGTA and
328 incubated for 1.5-18 hrs. The samples were subjected to ultracentrifugation for 30 min at 53
329 krpm in TLA-55 Beckmann rotor. Supernatant and pellet fractions were boiled in protein
330 sample buffer for 5 min, and analysed by PAGE on 10% polyacrylamide Tris-Tricine gels³³.
331 Proteins were visualized upon staining with Coomassie blue R.

332

333 **Expression and purification of isotope-labelled PulGp domain**

334 The PulG hydrophobic segment (residues 1-24) in pMS153, containing a cleavable PelB
335 signal peptide followed by a hexa-His tag fused to mature PulG was replaced by a TEV
336 cleavage site to generate plasmid pMS155, allowing us to produce the soluble PulG domain
337 (residues 25 through 134, hereafter called PulGp) in the periplasm. A C-terminal His₆-tagged
338 version of PulGp described previously⁹ was used for structure determination. Protein
339 expression, labelling and purification were performed as previously described³⁵. The
340 concentration of purified PulGp was determined from its absorbance at 280 nm and a
341 calculated extinction coefficient of 18 450 $\text{M}^{-1} \text{cm}^{-1}$.

342

343 **NMR spectroscopy**

344 The NMR samples were maintained in 50 mM HEPES buffer, pH 7, 50 mM NaCl, 10% D₂O,
345 supplemented with 1 mM CaCl₂ or 20 mM EGTA as required. All NMR experiments used for
346 structure calculation were recorded in the presence of 1 mM CaCl₂, to ensure the calcium
347 bound state of the protein. PulGp concentrations were between 0.5 and 0.6 mM. NMR
348 experiments were acquired at 298 K on a Varian spectrometer operating at 600 MHz (¹H
349 frequency) and equipped with a triple resonance cryogenically cooled ¹H {¹³C/¹⁵N} PFG
350 probe. 2, 2-dimethyl-2-silapentane-5-sulfonate (DSS) signal was taken as 0 ppm for
351 referencing ¹H chemical shifts and ¹⁵N and ¹³C chemical shifts were indirectly referenced to
352 DSS³⁶. The pulse sequences were employed as implemented in the Agilent BioPack
353 (*www.chem.agilent.com*). NMR Pipe³⁷ and CcpNmr Analysis³⁸ were used for NMR data
354 processing and analysis, respectively. 3D ¹³C and ¹⁵N NOESY-HSQC experiments with
355 mixing times of 120 ms were recorded on a Bruker AVANCE III 950 MHz equipped with a
356 TCI probe. In this case, the data was acquired and processed with TOPSPIN 3.1 (*Bruker,*
357 *Germany*). Resonance assignments were obtained from the analysis of standard NMR
358 experiments as previously described³⁵.

359

360 **PulGp calcium binding monitored by NMR**

361 ¹H-¹⁵N HSQC experiments (1024 x 128 data points) were acquired on 0.5 mM PulGp in 50
362 mM HEPES buffer, pH 7, 50 mM NaCl, supplemented by 1 mM CaCl₂ or 20 mM EGTA.
363 Chemical shift perturbations (CSP) of PulGp backbone amide cross-peaks were quantified by
364 using the equation $CSP = [\Delta\delta H^2 + (\Delta\delta N * 0.159)^2]^{1/2}$, where $\Delta\delta H$ and $\Delta\delta N$ are the observed
365 ¹H and ¹⁵N chemical shift changes between the two experimental conditions. For the calcium-
366 free state (in the presence of EGTA), only the conformer for which 94% of backbone
367 resonances were assigned was considered. The secondary structure content of PulGp

368 conformers in the calcium-free and bound states were estimated based on the backbone
369 chemical shifts with the TALOS-N software³⁹.

370

371 **NMR structure calculation**

372 We determined the structure of PulGp in the calcium bound state by performing several
373 cycles of calculation with ARIA 2.3 software⁴⁰ coupled to CNS 1.2 software⁴¹, making use of
374 the standard torsion angle/Cartesian simulating annealing protocol. Each cycle consisted of
375 automatic NOESY assignment and structure calculations with 9 iterations with default
376 parameters. In the last iteration 200 structures were calculated and further refined in an
377 explicit water box⁴². The 15 lowest energy structures exhibiting no NOE restraint violations >
378 0.5 Å and no dihedral angle violations > 5° were selected as the final ensemble.

379 Distance restraints were derived from 3D ¹³C and ¹⁵N NOESY-HSQC experiments, by
380 making use of the almost complete PulGp resonances assignment previously reported³⁵.
381 Chemical shift tolerances were set to 0.045 for protons and 0.4 ppm for the heteroatoms. Phi
382 and psi dihedral angles were predicted with TALOS-N³⁹ and predictions classified as “strong”
383 were incorporated as dihedral angle restraints. Hydrogen bonds predicted from H/D exchange
384 experiments and supported by characteristic secondary structure NOE patterns were included
385 as restraints in the last cycles of calculations. Hydrogen bonds typical of antiparallel β-sheet
386 were incorporated for pairing residues 99-101 to residues 110-112 as well as for residue 33 to
387 54 of α-helix. Spin diffusion correction was applied as implemented in ARIA by using the τc
388 of PulGp (7.8 ns) obtained from ¹⁵N spin-relaxation data.

389 The calcium atom was coordinated by the side chain oxygen of residues D117, S122
390 and D125, and backbone carbonyl oxygen of residues L114 and V119. The coordination
391 scheme was set up as described for calcium-bound X-ray structures of homologous major

392 pseudopilins¹⁰ and supported by the PulGp-calcium/EGTA CSP results. The geometry of the
393 calcium coordination was maintained by imposing fixed distances between the calcium ion
394 and the above-mentioned oxygen atoms. Proline residues 75 and 103 were defined in *cis*
395 configuration based on their backbone chemical shifts and predictions using Promega
396 software⁴³. The structure ensemble was visualized and inspected with PyMOL⁴⁴ and the
397 quality was evaluated with PROCHECK-NMR⁴⁵ and the PSVS server⁴⁶. The structure atomic
398 coordinates and restraints used in the calculation were deposited in the Protein Data Bank
399 (PDB id: 5O2Y). We estimated the global rotation correlation time (τ_c) of the protein after the
400 measurement of ¹⁵N relaxation parameters assuming an isotropic tumbling model and using
401 the program Tensor2⁴⁷ as described previously⁴⁸.

402

403 **Circular dichroism (CD)**

404 Far-UV CD spectra were recorded on an Aviv 215 spectropolarimeter equipped by a Peltier
405 thermostated cell holder (*Aviv Biomedical*) and in 1 mm cell. Protein sample were at 50 μ M in
406 50 mM Hepes, pH 7, 50 mM NaCl, supplemented by 1 mM CaCl₂ or 1 mM EGTA. Thermal
407 denaturation of PulG was performed by heating the samples from 15 to 98°C, with an
408 increment of 1°C and an average time of 2 s per step. The PulGp melting temperature (T_m) in
409 each buffer condition was determined from the denaturation curves obtained by measuring
410 ellipticity at 222 nm.

411

412 **CryoEM data collection and image processing**

413 The PulG sample (3 μ L, double mutant H106C-W129C) was applied to glow-discharged lacey
414 carbon grids (TED PELLA, Inc., 300 mesh). Then the grids were plunged frozen using a
415 Vitrobot Mark IV (FEI, Inc.), and subsequently imaged in a Titan Krios at 300keV with a
416 Falcon II direct electron detector (pixel size 1.05 Å /pixel). A total of 1,819 images, each of

417 which was from a total exposure of 2 seconds dose-fractionated into seven chunks, were
418 collected and motion corrected using MotionCorr⁴⁹. Program CTFFIND3⁵⁰ was used for
419 determining the defocus and astigmatism. Images with poor CTF estimation as well as
420 defocus > 3 μ m were discarded. The SPIDER software package⁵¹ was used for most other
421 operations. The CTF was corrected by multiplying the images from the first two chunks
422 (containing a dose of ~ 20 electrons/ \AA^2) with the theoretical CTF, which is a Wiener filter in
423 the limit of a very poor signal-to-noise ratio (SNR). This both corrects the phases, which need
424 to be flipped, and improves the SNR. The e2helixboxer routine within EMAN2⁵² was used for
425 boxing the filaments from the images. A total of 85,619 overlapping segments (384 px long),
426 with a shift of 14 px between adjacent segments (~96% overlap), were used for the IHRSR¹²
427 reconstruction. The resolution of the reconstruction was determined by the FSC between two
428 independent half maps, generated from two non-overlapping data sets, which was ~5 \AA at
429 FSC=0.143.

430

431 **Model building and refinement**

432 The PulGp NMR structure was docked in the cryo-EM map of PulG^{H106CW129C} by rigid-body
433 docking with the program Situs⁵³. Next, we generated a homology model of PulG N-terminal
434 α -helical residues (T2-V20) with the software Modeller⁵⁴ based on the structure of the
435 *Pseudomonas* PAK pilin (PDB 1OQW) as template. This model of the PulG N-terminal helix
436 was rigidly docked in the remaining density of the cryo-EM map (not occupied by PulGp)
437 with Situs⁵³ (residues 25-26 were not included since they are not ordered in the NMR
438 ensemble). A full-length model of PulG was then constructed with Modeller⁵⁴ by connecting
439 the docked PulGp and N-terminal helix with an extended linker (V21-G26) and adding F1.
440 Side-chains of H106 and W129 were replaced by a Cysteine side-chain and a methyl group
441 was added to the N-terminal Phe (N-methylphenylalanine). Density for the four C-terminal

442 residues is absent in the cryo-EM map and these residues were excluded. Next, the full-length
443 model of PulG^{CC} was refined in the density map with CNS⁵⁵. Energy minimization followed
444 by torsion-angle MD simulated-annealing in the presence of stereo-chemical restraints was
445 performed with the real-space objective function calculated by RSRef⁵⁶ in CNS. Helical
446 symmetry of the pilus was imposed with strict NCS constraints using the symmetry
447 parameters of the experimental cryoEM map. Refinement with RSRef was repeated 1000
448 times with different random seeds and the lowest energy model was selected for further
449 refinement. Next, several cycles of geometry optimization and full-atom real-space
450 refinement with PHENIX⁵⁷ followed by manual inspection in Coot⁵⁸ were performed to
451 improve the fit to the experimental cryo-EM map and remove severe steric clashes and
452 outliers. The final model of PulG pilus was validated with Molprobity⁵⁹ and refinement
453 statistics are given in Supplementary Table 3. Figures were generated with UCSF Chimera⁶⁰
454 and PyMOL⁴⁴.

455

456 **Analysis of residue conservation, co-evolutionary analysis and helical propensity**

457 Residue conservation of PulG sequence was obtained with the ConSurf server⁶¹ (300
458 homologous sequences asked with between 35 and 95% identity). Co-evolutionary contact
459 predictions for PulG were obtained with the Gremlin tool⁶². The ratio between the number of
460 sequences used by Gremlin and the length of the query was 14.9. Helical propensity of PulG
461 residues F1-R56 was computed with the PSIPRED⁶³ and Jpred4⁶⁴ servers; the alpha-helical
462 scale of Chou & Fasman⁶⁵ was used to create the profile shown in Supplementary Fig. 6d.

463

464 **Data availability**

465 The data that support the findings of this study are available from the corresponding authors
466 upon request. The PulGp structure is available in the Protein Data Bank (accession code

467 5O2Y). The PulG pseudopilus cryo-EM map and model have been deposited to EMDB and
468 PDB, with codes EMD-8812 and 5WDA respectively.

469

470

471 **References**

- 472 1. Berry, J.L. & Pelicic, V. Exceptionally widespread nanomachines composed of type
473 IV pilins: the prokaryotic Swiss Army knives. *FEMS Microbiol. Rev.* **39**, 134-54
474 (2015).
- 475 2. Jarrell, K.F. & Albers, S.V. The archaellum: an old motility structure with a new
476 name. *Trends Microbiol.* **20**, 307-12 (2012).
- 477 3. Thomassin, J.L., Santos Moreno, J., Guilvout, I., Tran Van Nhieu, G. & Francetic, O.
478 The trans-envelope architecture and function of the type 2 secretion system: new
479 insights raising new questions. *Mol. Microbiol.* **105**, 211-226 (2017).
- 480 4. Cianciotto, N.P. & White, R.C. Expanding role of type II secretion in bacterial
481 pathogenesis and beyond. *Infect. Immun.* **85**, pii: e00014-17 (2017).
- 482 5. Sauvonnet, N., Vignon, G., Pugsley, A.P. & Gounon, P. Pilus formation and protein
483 secretion by the same machinery in *Escherichia coli*. *EMBO J.* **19**, 2221-8 (2000).
- 484 6. Korotkov, K.V. & Hol, W.G. Structure of the GspK-GspI-GspJ complex from the
485 enterotoxigenic *Escherichia coli* type 2 secretion system. *Nat. Struct. Mol. Biol.* **15**,
486 462-8 (2008).
- 487 7. Cisneros, D.A., Bond, P.J., Pugsley, A.P., Campos, M. & Francetic, O. Minor
488 pseudopilin self-assembly primes type II secretion pseudopilus elongation. *EMBO J.*
489 **31**, 1041-53 (2012).
- 490 8. Korotkov, K.V., Sandkvist, M. & Hol, W.G. The type II secretion system: biogenesis,
491 molecular architecture and mechanism. *Nat. Rev. Microbiol.* **10**, 336-51 (2012).
- 492 9. Kohler, R. *et al.* Structure and assembly of the pseudopilin PulG. *Mol. Microbiol.* **54**,
493 647-64 (2004).
- 494 10. Korotkov, K.V. *et al.* Calcium is essential for the major pseudopilin in the type 2
495 secretion system. *J. Biol. Chem.* **284**, 25466-70 (2009).
- 496 11. Egelman, E.H. Cryo-EM of bacterial pili and archaeal flagellar filaments. *Curr. Op.*
497 *Struct. Biol.* **46**, 31-37 (2017).
- 498 12. Egelman, E.H. A robust algorithm for the reconstruction of helical filaments using
499 single-particle methods. *Ultramicroscopy* **85**, 225-34 (2000).
- 500 13. Nivaskumar, M. *et al.* Distinct docking and stabilization steps of the pseudopilus
501 conformational transition path suggest rotational assembly of type IV pilus-like fibers.
502 *Structure* **22**, 685-96 (2014).
- 503 14. Campos, M., Nilges, M., Cisneros, D.A. & Francetic, O. Detailed structural and
504 assembly model of the type II secretion pilus from sparse data. *Proc. Natl. Acad. Sci.*
505 *USA* **107**, 13081-6 (2010).
- 506 15. Alphonse, S. *et al.* Structure of the *Pseudomonas aeruginosa* XcpT pseudopilin, a
507 major component of the type II secretion system. *J. Struct. Biol.* **169**, 75-80 (2010).
- 508 16. Williams, A.W. & Straley, S.C. YopD of *Yersinia pestis* plays a role in negative
509 regulation of the low-calcium response in addition to its role in translocation of Yops.
510 *J. Bacteriol.* **180**, 350-8 (1998).

- 511 17. Lu, D. *et al.* Structural insights into the T6SS effector protein Tse3 and the Tse3-Tsi3
512 complex from *Pseudomonas aeruginosa* reveal a calcium-dependent membrane-
513 binding mechanism. *Mol. Microbiol.* **92**, 1092-112 (2014).
- 514 18. O'Brien, D.P. *et al.* Structural models of intrinsically disordered and calcium-bound
515 folded states of a protein adapted for secretion. *Sci Rep* **5**, 14223 (2015).
- 516 19. Bilecen, K. & Yildiz, F.H. Identification of a calcium-controlled negative regulatory
517 system affecting *Vibrio cholerae* biofilm formation. *Environ. Microbiol.* **11**, 2015-29
518 (2009).
- 519 20. Orans, J. *et al.* Crystal structure analysis reveals *Pseudomonas* PilY1 as an essential
520 calcium-dependent regulator of bacterial surface motility. *Proc. Natl. Acad. Sci. USA*
521 **107**, 1065-70 (2010).
- 522 21. Broder, U.N., Jaeger, T. & Jenal, U. LadS is a calcium-responsive kinase that induces
523 acute-to-chronic virulence switch in *Pseudomonas aeruginosa*. *Nat. Microbiol.* **2**,
524 16184 (2016).
- 525 22. Kolappan, S. *et al.* Structure of the *Neisseria meningitidis* Type IV pilus. *Nat.*
526 *Commun.* **7**, 13015 (2016).
- 527 23. Craig, L. *et al.* Type IV pilus structure by cryo-electron microscopy and
528 crystallography: implications for pilus assembly and functions. *Molecular Cell* **23**,
529 651-62 (2006).
- 530 24. Reardon, P.N. & Mueller, K.T. Structure of the type IVa major pilin from the
531 electrically conductive bacterial nanowires of *Geobacter sulfurreducens*. *J. Biol.*
532 *Chem.* **288**, 29260-6 (2013).
- 533 25. Santos-Moreno, J. *et al.* Polar N-terminal residues conserved in Type 2 secretion
534 pseudopilins determine subunit targeting and membrane extraction steps during fibre
535 assembly. *J. Mol. Biol.* **429**, 1746-1765 (2017).
- 536 26. Shevchik, V.E., Robert-Baudouy, J. & Condemine, G. Specific interaction between
537 OutD, an *Erwinia chrysanthemi* outer membrane protein of the general secretory
538 pathway, and secreted proteins. *EMBO J.* **16**, 3007-16 (1997).
- 539 27. Sandkvist, M. Biology of type II secretion. *Mol. Microbiol.* **40**, 271-83 (2001).
- 540 28. Durand, E. *et al.* Type II protein secretion in *Pseudomonas aeruginosa*: the
541 pseudopilus is a multifibrillar and adhesive structure. *J Bacteriol.* **185**, 2749-58
542 (2003).
- 543 29. Vignon, G. *et al.* Type IV-like pili formed by the type II secretion: specificity,
544 composition, bundling, polar localization, and surface presentation of peptides. *J.*
545 *Bacteriol.* **185**, 3416-28 (2003).
- 546 30. Jones, H.E., Holland, I.B. & Campbell, A.K. Direct measurement of free Ca⁽²⁺⁾ shows
547 different regulation of Ca⁽²⁺⁾ between the periplasm and the cytosol of *Escherichia*
548 *coli*. *Cell Calcium* **32**, 183-92 (2002).
- 549 31. East, A. *et al.* Structural basis of pullulanase membrane binding and secretion revealed
550 by X-ray crystallography, molecular dynamics and biochemical analysis. *Structure* **24**,
551 92-104 (2016).
- 552 32. Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular cloning : a laboratory manual*,
553 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989).
- 554 33. Schagger, H. Tricine-SDS-PAGE. *Nat. Protoc.* **1**, 16-22 (2006).
- 555 34. Nivaskumar, M. *et al.* Pseudopilin residue E5 is essential for recruitment by the type 2
556 secretion system assembly platform. *Mol. Microbiol.* **101**, 924-41 (2016).
- 557 35. Lopez-Castilla, A. *et al.* ¹H, ¹⁵N and ¹³C resonance assignments and secondary
558 structure of PulG, the major pseudopilin from *Klebsiella oxytoca* type 2 secretion
559 system. *Biomol. NMR Assign.* doi: 10.1007/s12104-017-9738-7 (2017).

- 560 36. Wishart, D.S. *et al.* ^1H , ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR.
561 *J. Biomol. NMR* **6**, 135-40 (1995).
- 562 37. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on
563 UNIX pipes. *J. Biomol. NMR* **6**, 277-93 (1995).
- 564 38. Vranken, W.F. *et al.* The CCPN data model for NMR spectroscopy: development of a
565 software pipeline. *Proteins* **59**, 687-96 (2005).
- 566 39. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from
567 NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227-41
568 (2013).
- 569 40. Rieping, W. *et al.* ARIA2: automated NOE assignment and data integration in NMR
570 structure calculation. *Bioinformatics* **23**, 381-2 (2007).
- 571 41. Brunger, A.T. *et al.* Crystallography & NMR system: A new software suite for
572 macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**,
573 905-21 (1998).
- 574 42. Linge, J.P., Williams, M.A., Spronk, C.A., Bonvin, A.M. & Nilges, M. Refinement of
575 protein structures in explicit solvent. *Proteins* **50**, 496-506 (2003).
- 576 43. Shen, Y. & Bax, A. Prediction of Xaa-Pro peptide bond conformation from sequence
577 and chemical shifts. *J. Biomol. NMR* **46**, 199-204 (2010).
- 578 44. Schrödinger, L. The PyMOL Molecular Graphics System, Version 1.8. (2015).
- 579 45. Laskowski, R.A., Rullmann, J.A., MacArthur, M.W., Kaptein, R. & Thornton, J.M.
580 AQUA and PROCHECK-NMR: programs for checking the quality of protein
581 structures solved by NMR. *J. Biomol. NMR* **8**, 477-86 (1996).
- 582 46. Bhattacharya, A., Tejero, R. & Montelione, G.T. Evaluating protein structures
583 determined by structural genomics consortia. *Proteins* **66**, 778-95 (2007).
- 584 47. Dosset, P., Hus, J.C., Blackledge, M. & Marion, D. Efficient analysis of
585 macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol.*
586 *NMR* **16**, 23-8 (2000).
- 587 48. Wolff, N. *et al.* Comparative analysis of structural and dynamic properties of the
588 loaded and unloaded hemophore HasA: functional implications. *J. Mol. Biol.* **376**,
589 517-25 (2008).
- 590 49. Li, X. *et al.* Electron counting and beam-induced motion correction enable near-
591 atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584-90 (2013).
- 592 50. Mindell, J.A. & Grigorieff, N. Accurate determination of local defocus and specimen
593 tilt in electron microscopy. *J. Struct. Biol.* **142**, 334-347 (2003).
- 594 51. Frank, J. *et al.* SPIDER and WEB: Processing and visualization of images in 3D
595 electron microscopy and related fields. *J. Struct. Biol.* **116**, 190-199 (1996).
- 596 52. Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy.
597 *J. Struct. Biol.* **157**, 38-46 (2007).
- 598 53. Wriggers, W. Conventions and workflows for using Situs. *Acta Crystallogr. D Biol.*
599 *Crystallogr.* **68**, 344-51 (2012).
- 600 54. Sali, A. & Blundell, T.L. Comparative protein modelling by satisfaction of spatial
601 restraints. *J. Mol. Biol.* **234**, 779-815 (1993).
- 602 55. Brunger, A.T. Version 1.2 of the Crystallography and NMR system. *Nat. Protoc.* **2**,
603 2728-33 (2007).
- 604 56. Chapman, M.S., Trzynka, A. & Chapman, B.K. Atomic modeling of cryo-electron
605 microscopy reconstructions--joint refinement of model and imaging parameters. *J.*
606 *Struct. Biol.* **182**, 10-21 (2013).
- 607 57. Adams, P.D. *et al.* PHENIX: a comprehensive Python-based system for
608 macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213-21
609 (2010).

- 610 58. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of
611 Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486-501 (2010).
612 59. Davis, I.W. *et al.* MolProbity: all-atom contacts and structure validation for proteins
613 and nucleic acids. *Nucleic Acids Res.* **35**, W375-83 (2007).
614 60. Pettersen, E.F. *et al.* UCSF Chimera - a visualization system for exploratory research
615 and analysis. *J. Comput. Chem.* **25**, 1605-12 (2004).
616 61. Ashkenazy, H. *et al.* ConSurf 2016: an improved methodology to estimate and
617 visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344-
618 50 (2016).
619 62. Kamisetty, H., Ovchinnikov, S. & Baker, D. Assessing the utility of coevolution-
620 based residue-residue contact predictions in a sequence- and structure-rich era. *Proc.*
621 *Natl. Acad. Sci. USA* **110**, 15674-9 (2013).
622 63. Jones, D.T. Protein secondary structure prediction based on position-specific scoring
623 matrices. *J. Mol. Biol.* **292**, 195-202 (1999).
624 64. Drozdetskiy, A., Cole, C., Procter, J. & Barton, G.J. JPred4: a protein secondary
625 structure prediction server. *Nucleic Acids Res.* **43**, W389-94 (2015).
626 65. Chou, P.Y. & Fasman, G.D. Prediction of the secondary structure of proteins from
627 their amino acid sequence. *Adv. Enzymol. Relat. Areas Mol. Biol.* **47**, 45-148 (1978).
628
629

630 Correspondence and requests for materials should be addressed to: Olivera Francetic

631 (olivera.francetic@pasteur.fr) and Nadia Izadi-Pruneyre (nadia.izadi@pasteur.fr).

632

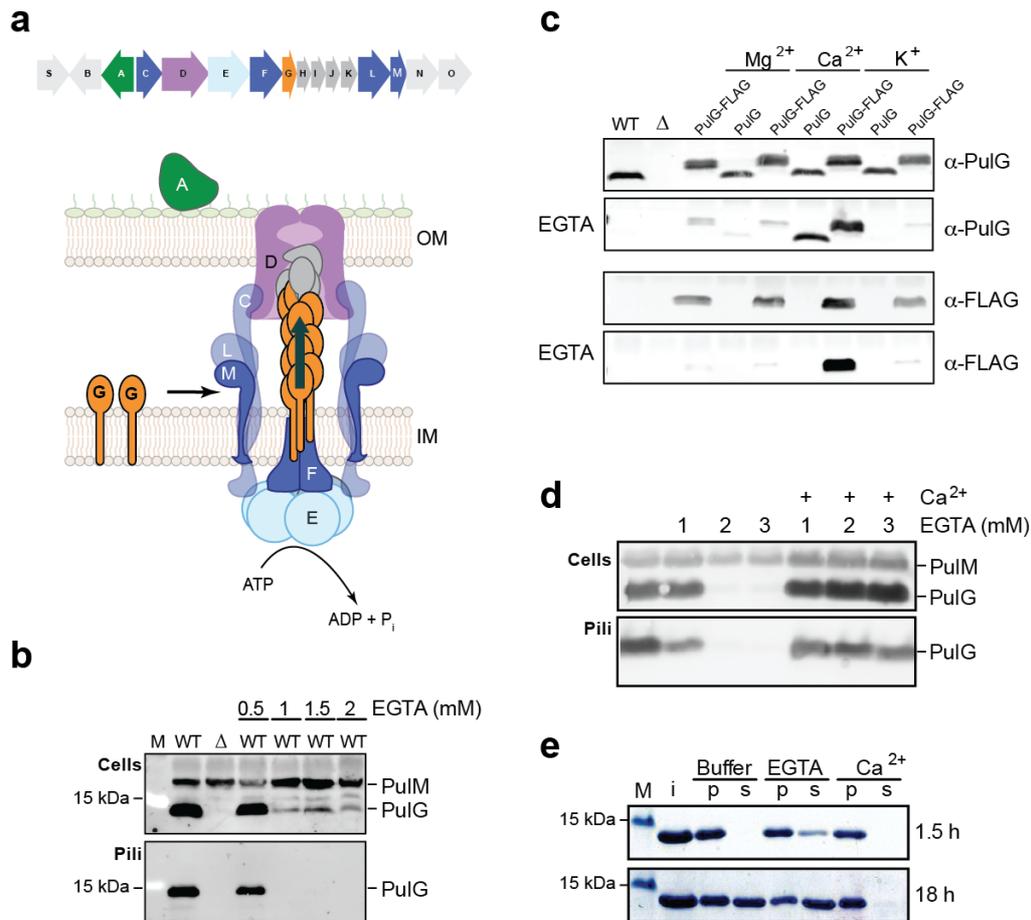
633 **Acknowledgements**

634 This work was funded by the Institut Pasteur, the Centre National de la Recherche
635 Scientifique (CNRS), the French Agence Nationale de la Recherche (ANR-14-CE09-0004),
636 the European Union FP7-IDEAS-ERC 294809 (to M. Nilges) and the NIH R35GM122510 (to
637 E.H.E.). We thank Léa Khoury for help in sample preparation, the Plateforme de Biophysique
638 Moléculaire of Institut Pasteur and Bruno Baron for assistance in CD experiments. We are
639 grateful to Muriel Delepierre and Daniel Ladant for support and interest in this work. We
640 acknowledge Nelly Morellet and financial support from the TGIR-RMN-THC Fr3050
641 CNRS.

642

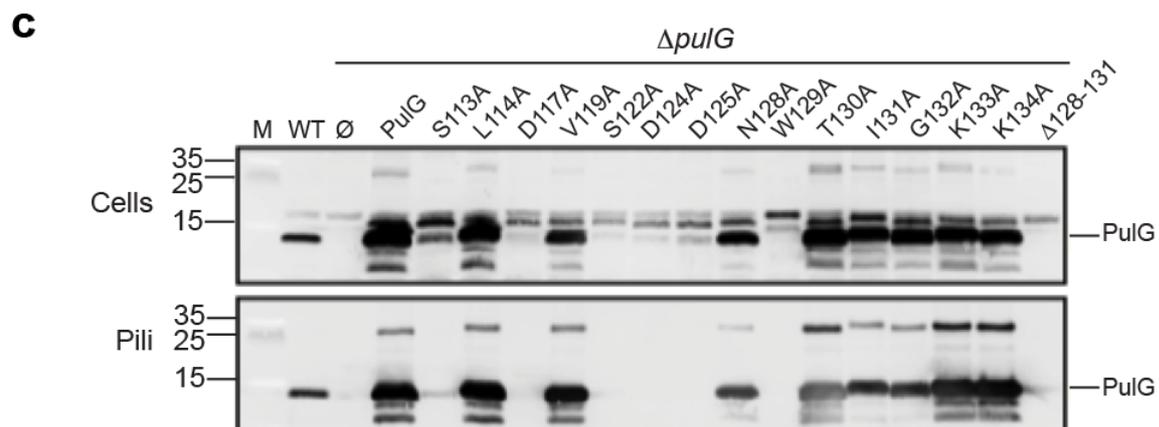
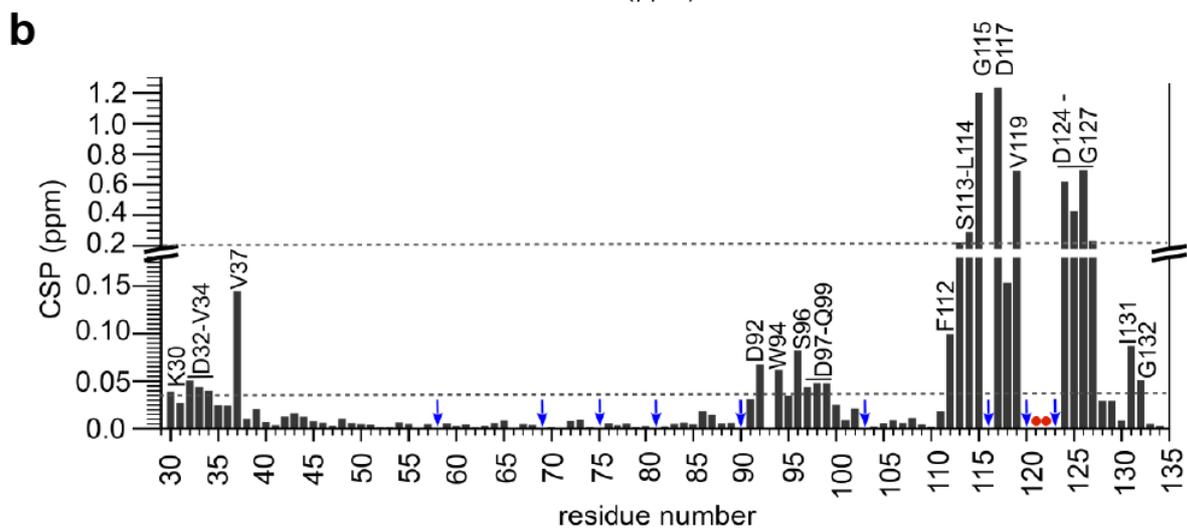
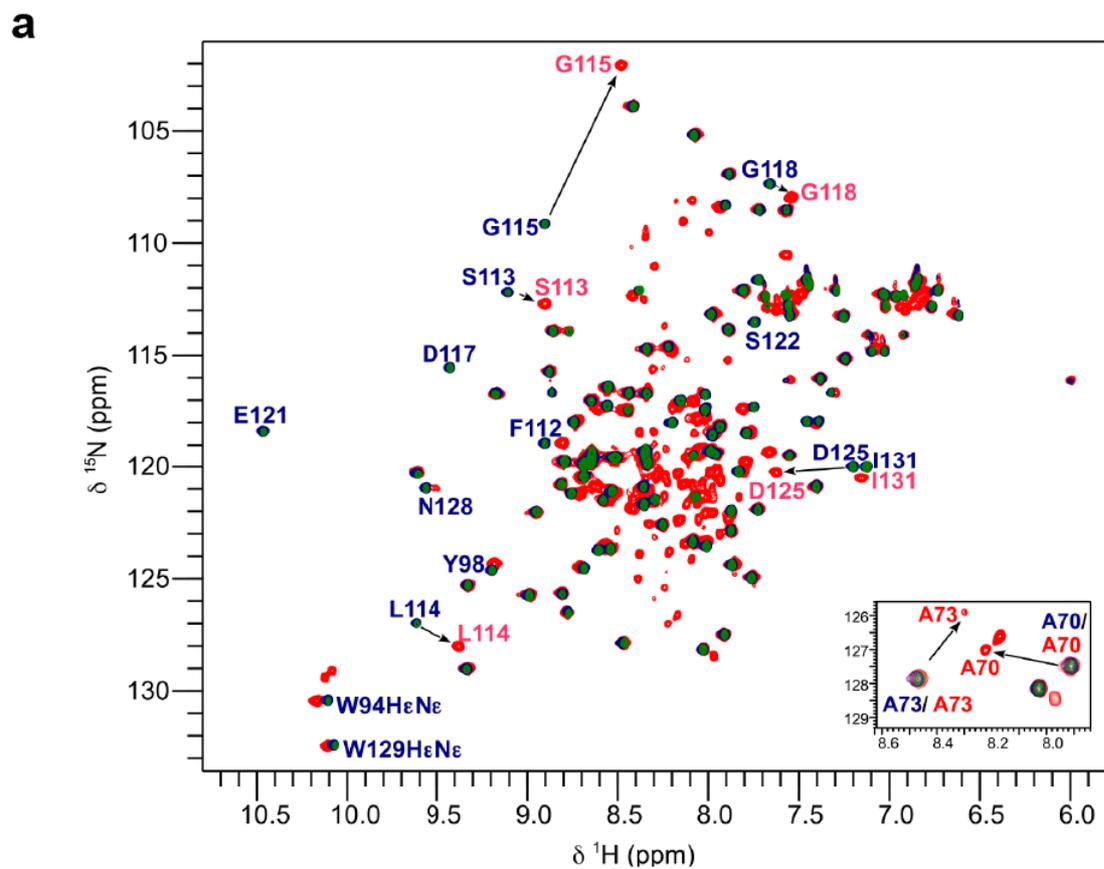
643 **Author contributions**

644 Conceived and designed the experiments: ALC, JLT, BB, M. Nilges, EHE, NIP and OF.
645 Performed the experiments: ALC, JLT, BB, WZ, XY, M. Nivaskumar. Analysed the data and
646 wrote the manuscript: ALC, JLT, BB, WZ, M. Nilges, EHE, NIP and OF.
647

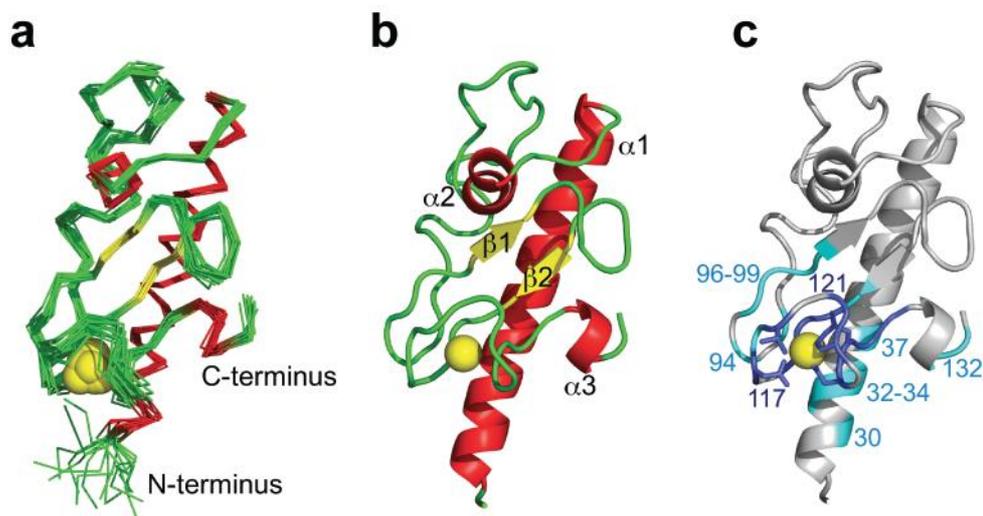


649

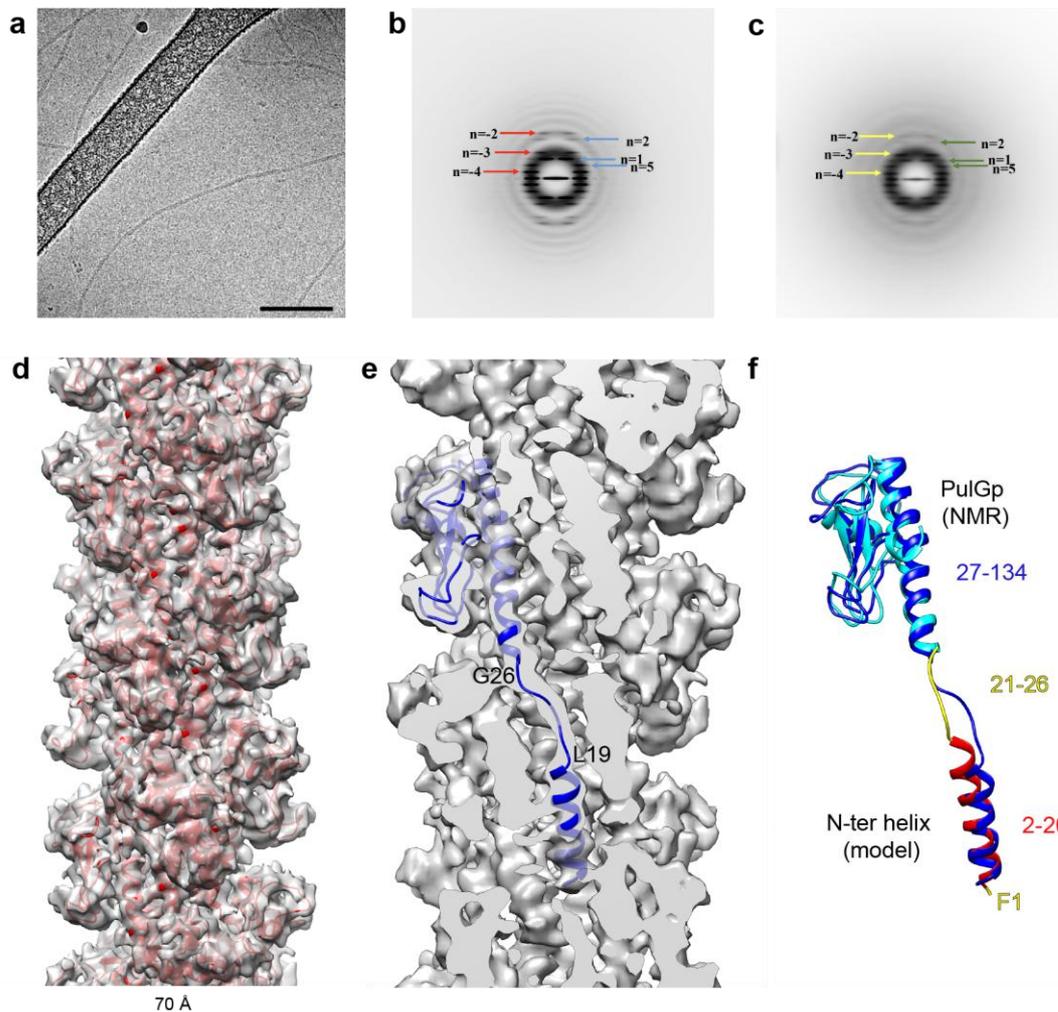
650 **Figure 1.** Calcium is required for PulG stability, pseudopilus assembly and stability. **a**, Top:
 651 cartoon of the gene organization in the *pul* cluster. Bottom: a schematic view of the assembled
 652 T2SS nanomachine, with individual components labelled with the corresponding single-letter
 653 and colour code. **b**, **c**, **d**, PAP7460 harbouring pCHAP8185 and pSU19 (WT), pCHAP8184
 654 and pSU19 ($\Delta pulG$; Δ), pCHAP8184 and pCHAP8658 (PulG), or pCHAP8184 and
 655 pCHAP5247 (PulG-FLAG) were grown under pilus-inducing conditions in media
 656 supplemented with EGTA and/or cations as indicated. Where appropriate, molecular weight
 657 markers (M) are indicated on the left in kDa. **b**, Effect of EGTA on T2SS function was
 658 monitored by probing cell extract and sheared pilus fractions with anti-PulG and anti-PulM
 659 antibodies. **c**, Effect of different cations on total PulG levels was monitored in bacterial cells
 660 using PulG (α -PulG) and FLAG (α -FLAG) antibodies in absence (top panels) or presence of
 661 EGTA (bottom panels). **d**, Ability of Ca²⁺ to restore PulG levels and pseudopilus formation in
 662 WT was assessed using anti-PulG and anti-PulM antibodies. **e**, Pseudopili present on the
 663 surface of PAP7460 cells harbouring pCHAP8184 and pCHAP8658 were mechanically
 664 sheared and isolated by ultracentrifugation (i). After purification, pseudopili were incubated
 665 in buffer alone or buffer supplemented with EGTA or Ca²⁺ for the indicated time and intact
 666 pseudopili (p) were separated from the supernatant containing broken or degraded pseudopili
 667 (s) by ultracentrifugation. All experiments were repeated three times with reproducible
 668 results.



670 **Figure 2.** The effect of calcium on PulGp folding and identification of calcium coordinating
671 residues. **a**, Superposed ^1H - ^{15}N HSQC spectra of 0.5 mM ^{15}N labelled-PulGp purified from
672 the bacterial periplasm (blue contours) in buffer (50 mM HEPES pH 7, 50 mM NaCl), buffer
673 supplemented with 1 mM of calcium (green contours) or buffer supplemented with 20 mM of
674 EGTA (red contours). Some residues exhibiting chemical shift variations are indicated by
675 sequence numbers and by one-letter amino acid code. Insert: zoom showing double peaks in
676 the presence of EGTA for residues A70 and A73. **b**, Histogram showing the chemical shift
677 perturbation (CSP) values of PulGp backbone amide signals (0.5 mM) in the presence of
678 calcium (1 mM) and EGTA (20 mM), as a function of residue numbers. Blue arrows indicate
679 proline residues or amide resonances that could not be unambiguously assigned in the calcium
680 bound state. Red dots indicate residues whose signals were perturbed upon EGTA addition
681 but could not be unambiguously assigned. Residues displaying significant levels of signal
682 perturbations (CSP>0.2 ppm or CSP>0.04 ppm, dashed lines) are labelled. **c**, Bacterial cell
683 and pilus fractions from strains harbouring T2SS (WT) or derivatives lacking *pulG* ($\Delta pulG$)
684 and empty vector (\emptyset) or vector encoding native PulG or its variants with indicated residue
685 substitutions. Equivalent of 0.05 OD_{600nm} of cell and pili fractions were analysed by western
686 blot using anti-PulG antibodies. Molecular mass of standard weight markers (M) is indicated
687 on the left (in kDa). Data shown are representative of three independent experiments.
688

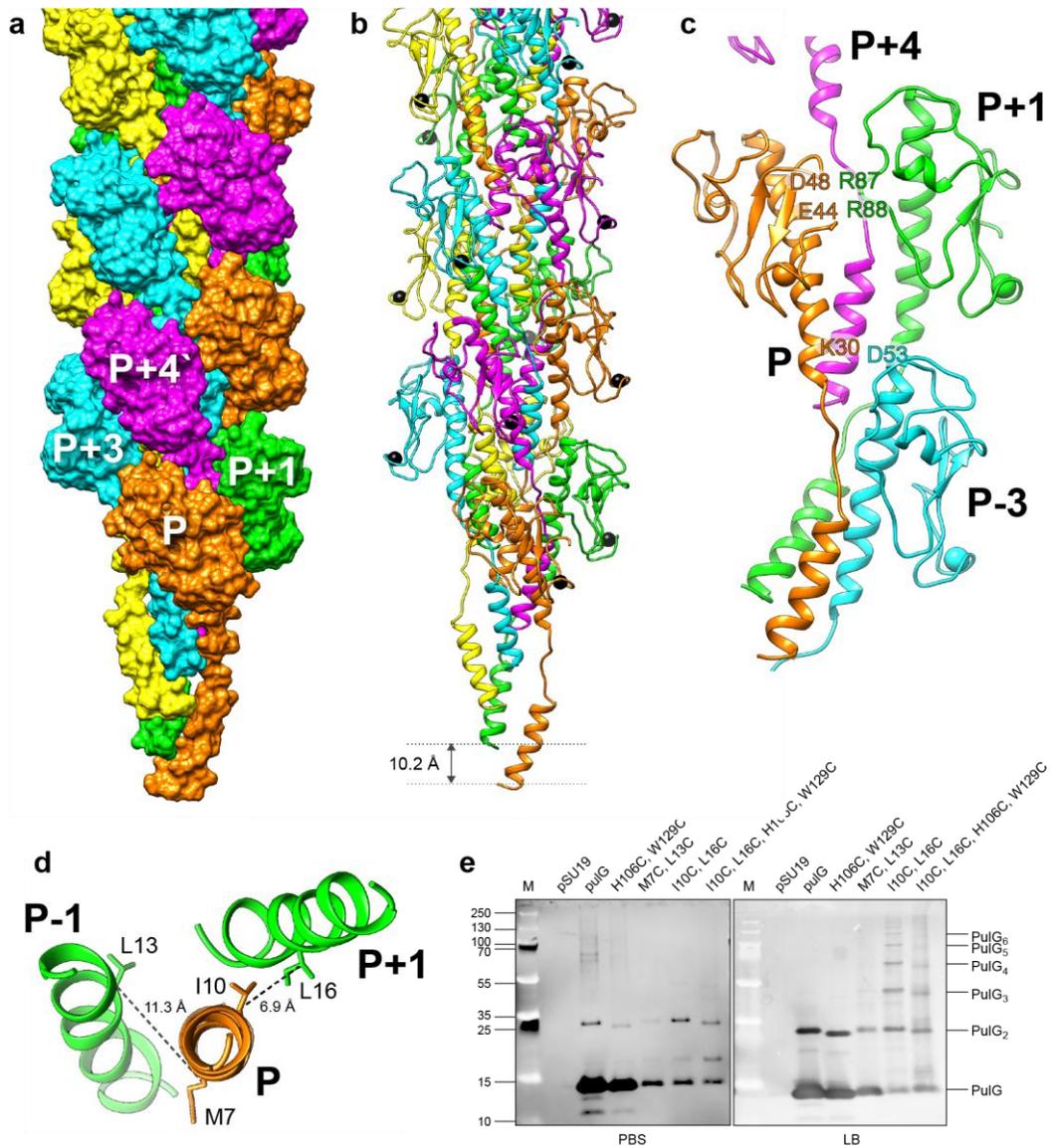


689 **Figure 3.** NMR structure of PulGp in the calcium bound state.
690 **a**, Superposition of the 15 best-calculated structures. **b**, Ribbon representation of the lowest-
691 energy PulGp conformer. Secondary structure elements are coloured in red (helices) and
692 yellow (β -strands) and calcium atoms are represented as yellow spheres. **c**, Residues that
693 exhibit significant CSP values (above the dashed line threshold in Fig. 2b) are coloured in
694 dark blue (CSP>0.2 ppm) and light blue (CSP>0.04 ppm) in PulGp structure. For clarity, not
695 all significantly affected residues are labelled. Side chains of some of the calcium-
696 coordinating residues are shown as sticks.
697
698



699
700
701
702
703
704
705
706
707
708
709
710
711
712

Figure 4. Morphology of *Klebsiella* T2SS pseudopilus PulG^{CC} filament by cryo-EM and reconstruction at ~5 Å resolution. **a**, Representative electron micrograph of *Klebsiella* T2SS pseudopilus PulG^{CC} filament in vitreous ice. Scale bar, 100 nm. Power spectra generated from 72,197 overlapping segments of pseudopilus PulG^{CC} (**b**) and 32,162 overlapping segments of pseudopilus PulG^{WT} (**c**). The layer lines with corresponding Bessel orders are indicated. **d**, An outside view of the cryo-EM reconstruction of PulG^{CC} pilus with the refined atomic model shown in ribbons (red). **e**, Cross-section of the cryo-EM reconstruction with a single PulG^{CC} monomer showing the elongated part (L19-G26) of the N-terminal region and the corresponding density. **f**, Superposition of the initial (cyan/yellow/red) and final (blue) model of a PulG monomer. The PulGp NMR structure (residues 27-134, cyan), excluding flexible M25 and G26, and a model of the N-terminal helix (residues 2-20, red) were docked in the reconstruction and joined by an extended linker (residues 21-26, yellow) before refinement.



713
714 **Figure 5.** Structure of the PulG^{CC} pilus

715 **a**, Surface view of the PulG^{CC} pilus structure where subunits P, P+1, P+2, P+3 and P+4 (along
716 the 1-start helix) are coloured in orange, green, yellow, blue and magenta, respectively. **b**,
717 Ribbon view of PulG^{CC} pilus structure where subunits are coloured as in **a** and calcium ions
718 are shown as black spheres. **c**, Arrangement of neighbouring subunits in the PulG^{CC} structure,
719 showing the location of residues involved in conserved interactions. Calcium ions are shown
720 as spheres. **d**, Top view of the segment F1-L19 of subunits P and P±1 and shortest distances
721 between PulG subunits in the pilus for M7Cβ-L13Cβ and I10Cβ-L16Cβ (dashed lines). **e**,
722 Specific cross-linking of Cys10-Cys16 in PulG^{WT} and PulG^{CC}. Pili were sheared from the cell
723 surface by vortexing and separated from bacteria harbouring plasmid pCHAP8184 (*ΔpulG*)
724 and either empty vector (∅) or its derivatives encoding PulG variants as indicated above each
725 lane (Supplementary Table 2). Pilus fractions were incubated with agitation in PBS buffer
726 (left) or LB (right) for 1 hr at 30°C. Pili samples were analysed by western blot using PulG
727 antibodies. M, molecular weight markers (in kDa) are indicated on the left. Migration of PulG
728 monomers and oligomers (n=2-6) are indicated on the right. A representative of three
729 independent experiments is shown.