NMICROBIOL-19082155B

Revised version (2020/03/06-2)

Structure of polymerized Type V pilin reveals assembly mechanism involving protease-mediated strand-exchange

Satoshi Shibata[1,†], Mikio Shoji[2,†], Kodai Okada[3], Hideyuki Matsunami[1], Melissa M. Matthews[1], Katsumi Imada[3,*], Koji Nakayama[2,*] and Matthias Wolf[1,*]

[1]Molecular Cryo-Electron Microscopy Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan.

[2]Department of Microbiology and Oral Infection, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan.

[3]Department of Macromolecular Science, Graduate School of Sciences, Osaka University, Osaka, Japan.

[†]These authors have contributed equally

*Corresponding authors: knak@nagasaki-u.ac.jp (KN), kimada@chem.sci.osaka-u.ac.jp (KI), and matthias.wolf@oist.jp (MW)

Bacterial adhesion is a general strategy for host-microbe and microbe-microbe interactions. Adhesive pili are essential for colonization, biofilm formation, virulence, and pathogenesis of many environmental and pathogenic bacteria[1,2]. Members of the class Bacteroidia have unique Type V pili, assembled by protease-mediated polymerization[3]. *Porphyromonas gingivalis* is the main contributor to periodontal disease and its Type V pili are a key factor for its virulence[4]. However, the structure of the polymerized pilus and its assembly mechanism are unknown. Here, we show structures of polymerized and monomeric states of FimA stalk pilin from *P. gingivalis*, determined by cryo-EM and crystallography. The atomic model of assembled FimA shows that the C-terminal strand of a donor subunit is inserted into a groove in the β-sheet of an acceptor subunit after N-terminal cleavage by the protease, RgpB. The C-terminus of the donor strand is essential for polymerization. We propose a sequential polar assembly mechanism for Type V pili at the cell surface, involving protease-mediated strand-exchange, employed by various Gram-negative species belonging to the class Bacteroidia. Our results reveal functional surfaces related to pathogenic properties of polymerized FimA. These insights may facilitate development of anti-bacterial drugs.

Bacterial pili are classified as chaperone-usher, curli, Type IV, sortase-dependent, the recently discovered Type V, and as yet unclassified e-pili[1,2,5–7]. Type V pili are unique to Gram-negative bacteria belonging to the class Bacteroidia in the phylum Bacteroidetes, and are composed of anchor, stalk, accessory (or adaptor), and tip pilins[2]. These pilins are structurally conserved and have distinct N- and C-terminal domains (NTD and CTD)[3]. Each domain has a transthyretin-like fold, consisting of seven β-strands arranged in two β-sheets. Based on the conformation of the C-terminal strand in crystal structures, pilins have been classified as either “closed-form” or “open-form”. The C-terminal strand of open-form pilins extends into the solvent and is not stabilized by crystal contacts[3]. In analogy to donor-strand

exchange in chaperone-assisted pilus assembly[2,8], we refer to the monomeric "closed-form"[3] of FimA, in which the C-terminal strand constitutes part of the CTD β-sheet, as "self-complementing", and distinguish the polymerized form as "donor-strand exchanged".

Porphyromonas gingivalis belongs to the class Bacteroidia and is a major oral pathogen, associated not only with development and progression of chronic periodontitis in adults[9], but also several systemic diseases, such as rheumatoid arthritis, cardiovascular disease, and pancreatic cancer[10–12]. Recently, *P. gingivalis* has also been implicated in formation of β-amyloid-like deposits in the brain[13]. *P. gingivalis* possesses Type V pili that are either of the long, major Fim pilus type (0.3-1.6 μm) or the short, minor Mfa type (80-120 nm)[14,15], both of which are key virulence factors for host colonization and evasion of the innate immune system[16,17]. Both types of pili serve to attach cells to oral epithelia via extracellular matrix proteins[16,17], as well as in co-aggregation of *P. gingivalis* with other pathogens in the oral biofilm[18].

Fim and Mfa pilins are expressed from operons with similar structures, composed of genes encoding stalk pilins (major: FimA or minor: Mfa1), anchor pilins (major: FimB or minor: Mfa2), accessory pilins (major: FimC, FimD, and FimE or minor: Mfa3, Mfa4 and Mfa5), and regulatory proteins[18–21]. FimA pilins from various *P. gingivalis* strains are further classified into five subtypes, one of which is divided into two sets (FimA1-FimA5 and FimA1b) that differ in immunogenicity[22–25]. FimA of *P. gingivalis* ATCC 33277 (FimA1) is the best studied pilin and shares 50-85% sequence identity with other FimA subtypes[3].

The Lol pathway delivers lipoproteins to the inner leaflet of the outer membrane of Gram-negative bacteria[26]. During Fim pilus biogenesis, pilins are synthesized in the cytoplasm as lipoprotein precursors containing a signal peptide and a conserved cysteine residue that serves as the lipidation site. Pilins are then exported to the periplasm by the Sec system[27], where the signal peptide is cleaved by type II signal peptidase, and the cysteine is conjugated with lipids. Lipidated pilins are presumably

transported to the outer membrane surface via lipoprotein-sorting machinery[28]. However, the precise mechanism, how subunits are exported to the cell surface, is not fully understood[29]. At the cell surface, an arginine- or lysine-specific protease, R-gingipain (Rgp) or K-gingipain (Kgp), removes an N-terminal portion of the pilin that includes the conserved arginine or lysine cleavage residue and the lipidated cysteine. In this way, pilins mature and assemble into the pilus filament[28,30,31]. Previous experiments have shown that the C-terminal strand of one pilin interacts with residues in the hydrophobic groove of a neighboring pilin in mature pilus filaments[3]. These insights and recent crystal structures led to the hypothesis that Type V pili are assembled by a protease-mediated strand-exchange mechanism[2,3].

We cloned and overexpressed recombinant FimA1 from the Type V pilus of *P. gingivalis* strain ATCC 33277 (rFimA) containing residues 20-383 in *Escherichia coli*, and purified and crystalized rFimA in monomeric form at 2.1 Å resolution (Fig. 1)[3,28,30]. There were two nearly identical molecules in the asymmetric unit, except for loop A1-B1 (residues 39-57) and a loop segment of residues 365-370. The structure comprises two domains, an N-terminal domain (NTD) and a C-terminal domain (CTD), consistent with the crystal structure of the homologous FimA4 (PDB 4q98). Each domain contains two β-sheets with seven β-strands in the NTD and in the CTD, arranged in a characteristic transthyretin-like fold. C-terminal residues 380-383 were disordered (Fig. 1b, dashed red line). The C-terminal donor strand A1′-A2′ (Fig. 1b, red) passes under loop A1-B1 (residues 39-57), which contains a conserved protease cleavage site at Arg46 (Fig. 1b). Cleavage of the loop by RgpB removes the safety pin and primes the structure for release of the donor strand. Strand A1’ is located at the edge of the CTD β-sheet. It is held in place between the edge of the CTD β-sheet and a short adjacent loop in anti-parallel β-strand conformation (residues 322-326) (Fig. 1b). The anchor strand A1 (Fig. 1b), however, is fully integrated into its β-sheet, suggesting greater stability at this position. FimA1 shows 79% amino acid sequence identity with FimA2 (Extended Data Fig. 1a, aligned with Clustal W[32]). *P. gingivalis* strains with

FimA2 are most frequently found in severe periodontitis patients, whereas those with FimA1 are most prevalent in *P. gingivalis*-positive, healthy adults. We also determined the crystal structure of FimA2 (residues 31-384) at 1.6 Å resolution and compared it with our structures of FimA1 and previously crystalized FimA4[3]. Crystal structures of FimA1, FimA2, and FimA4 were very similar overall. However, loop regions show significant structural differences (Extended Data Fig. 1 b-d).

Previous studies showed that recombinant Type V pilins can polymerize *in vitro*[31,33]. We found that rFimA monomers self-assemble into long filaments after *in vitro* treatment with the arginine-specific protease, RgpB. Negative stain transmission electron microscopy (TEM) of purified rFimA filaments showed a twisted ribbon structure similar to Fim pili isolated from Mfa pilus-defective *mfa1* mutant cells (Fig. 2a). We estimated the maximum observed assembly rate of rFimA *in vitro* as ~2.5 subunits/min (Extended Data Fig. 2 a-d and source data), with polymerization starting ~5 min after protease cleavage (Extended Data Fig. 2e). The longest filaments reached ~1.2 µm after 1 h. These values are similar to *in vitro* growth rates reported for Type 1 pili in *E. coli*[34]. Two-dimensional (2D) class-average cryo-EM images of the rFimA filament and the Fim pilus, purified from the *mfa1* mutant in amorphous ice, show that their dimensions and secondary structural features are very similar (Fig. 2a; approximate widths, 39 Å and 42 Å, respectively).

We performed single-particle cryo-EM analysis and used homogeneous *in vitro*-assembled homopolymeric rFimA filaments, because the *mfa1* mutant also expresses accessory pilins such as FimC, FimD, and FimE[35]. The resulting unsymmetrized cryo-EM electron potential map of the rFimA filament was determined at 3.6 Å resolution (Extended Data Fig. 3 and methods), enabling us to build an atomic model containing three pilin subunits (Fig. 2b and 2c). The fitted atomic model depicts the geometry of the rFimA filament. The rFimA filament formed a right-handed helix (twist 71.0° or 5.1 subunits/turn, rise 66.7 Å or pitch 339 Å/turn). Compared to an additional 3D-reconstruction of the Fim pilus from the *mfa1* mutant, helical parameters

of the homopolymeric rFimA filament were almost identical (Extended Data Fig. 4). The electron potential map clearly revealed features of the aromatic Phe48 and its neighboring Ala47 at the new N-terminus of the NTD (Fig. 2d), confirming that the N-terminus was cleaved at Arg46 by RgpB. The reconstructed electron potential corresponding to the donor strand A1′-A2′ extended into the acceptor pilin subunit, and the C-terminal Trp383 from the donor pilin subunit was found in the CTD of the acceptor pilin subunit (Fig. 2e). Comparison of the donor-strand exchanged structure of rFimA with its self-complementing structure shows that the donor strand A1′-A2′ swings out of the CTD at the loop located at residues 356-359 (Fig. 2f).

We mapped non-conserved regions, functional binding sites, and epitopes onto the assembled FimA atomic model (Extended Data Fig. 1e and 1f). FimA binds with salivary molecules (salivary protein, statherin, and a proline-rich protein, PRP-1) and is recognized by three pattern-recognition receptors (CD14, CD11b, and CD18)[16,36]. The salivary protein and statherin binding regions are distributed on well-conserved surfaces, consistent with saliva-binding capability of FimA1 and FimA2[24] (Extended Data Fig. 1). The PRP1 binding site is not conserved. Epitopes recognizing CD14, CD11b, and CD18 map to the CTD of the FimA subunit (Extended Data Fig. 1). *P. gingivalis* strains expressing FimA2 induce more severe inflammation than those expressing FimA1[37–39].

In the donor-strand exchanged structure (Fig. 2), the A1′-A2′ strand of the donor pilin is relocated to the hydrophobic groove of the acceptor pilin. Removal of the N-terminal β-strand A1 and the A1′-A2′ strand from our atomic model exposes grooves traversing both β-sheets of NTD and CTD (termed NTD groove and CTD groove, respectively). Surface properties of these grooves are illustrated in a hydrophobicity map (Fig. 3a). Our structure of the assembled donor-strand exchanged-form pilin confirms that the C-terminal strand A1′-A2′ rests in the grooves of the acceptor pilin, restoring the full β-sheet (Fig. 3a). The A1′ strand (residues 363-372) is located in the NTD groove (Fig. 3b), and A2′ strand (residues 377-383) is inserted into the CTD

groove (Fig. 3c). Thus, the extended strand of the donor pilin is oriented by hydrophobic interactions in the grooves of the acceptor pilin and held in place by reintegration into its β-sheets via inter-strand hydrogen bonds.

The high thermal stability of the assembled pilus[40] (ladder pattern in Fig. 3d and 3f after incubation for 10 min in SDS buffer at 80°C) may be explained by increased interactions after strand exchange, as well as the entropic cost of exposing the hydrophobic grooves by strand removal. Donor strand A1′ is now found in the NTD groove, replacing the A1 strand (Extended Data Fig. 5a). The donor-strand exchanged-form A1′ β-strand is longer than A1 in the self-complementing form and has additional side-chain interactions with the NTD. In addition to the β-sheet interactions and the entropically favorable burial of the hydrophobic groove, which are already present in the A1′-CTD pair of the monomeric, self-complementing form, interactions in A2′-CTD of the donor-strand exchanged form include additional hydrogen-bonding between Trp383 and Gln379 of A2′ and the main chain atoms of the CTD. Interestingly, the hydrophilic Gln379 is located at the interior hydrophobic face of A2′ (Extended Data Fig. 5a). However, the Q379A mutation did not affect pilus formation (Extended Data Fig. 5b and 5c). In our atomic model, the side chain of Asn366 of A1′ is within hydrogen-bonding distance of both the main chain carbonyl oxygen and the side chain carboxy oxygen of Glu83 on the NTD. The geometry of Gln368 suggests a similar hydrogen bond with NTD-Thr80.

Interactions outside the hydrophobic cleft are also observed in the pilin model. At the interface between subunits, Tyr186 of the donor subunit is within hydrogen bonding distance of the main chain carbonyl oxygen of acceptor subunit Gln70. Similarly, the side chain amide nitrogen of Gln70 pairs with the backbone carbonyl oxygen of donor subunit Ala185 (Extended Data Fig. 5e and 5f). In their crystallographic unit cells, FimA4 or FimA1 pilins from *Bacteroides eggerthii* were stacked head-to-tail, forming pilus-like filaments, despite still being in an immature

188 conformation[3]. This suggests that the subunit head-tail interface contributes to relative
189 subunit alignment and stability of the pilus.

190 We tested the effects of C-terminal mutants. Deletion of the three C-terminal
191 residues of FimA1 prevents Fim pilus assembly[3]. Trp383 at the C-terminus of FimA1 is
192 conserved among FimA subtypes (Extended Data Fig. 1a). We deleted C-terminal
193 residues Trp383, Thr382-Trp383, or Ala381-Trp383 by replacement with a stop codon
194 and found that each of these mutations prevented FimA pilus formation on the cell
195 surface (Fig. 3d and 3e). rFimA1(Δ383) also failed to polymerize *in vitro* (Extended
196 Data Fig. 5d). This indicates the importance of the C-terminal Trp383 for FimA
197 polymerization. Furthermore, we assessed whether Trp383 is specifically required at
198 this position: Substitutions of Trp383 with aromatic (Phe, Tyr) or hydrophobic (Leu)
199 residues were tolerated, but substitutions with His, Ala, or Gly reduced polymerization
200 significantly (Fig. 3f and 3g). The C-terminus is essential for polymerization. When the
201 donor strand A1' is folded back (self-complementing form), it is unlikely that an
202 incoming strand A2' from another donor pilin would be capable of peeling the donor
203 strand from its β-sheet. Therefore, release of the donor strand A1' following cleavage
204 by RgpB should precede binding of the C-terminus of the next pilin. Indeed, related
205 Type V pilins from other species (BfrFim1l and BthFim1A) have been crystallized with
206 their donor strands released from the CTD while the anchor strand, A1, was still
207 present[3], suggesting that the donor strand can flip out from the CTD easily.

208 We noticed that FimA(Δ383) accumulated in cleaved form in the cell fraction
209 (bands near 40 kDa molecular weight in Fig. 3d iii-iv). Indeed, a dot blot assay using
210 wild-type and mutant cells (Fig. 3h) detected the FimA(Δ383) protein at the cell surface.
211 Both observations demonstrated that the FimA(Δ383) protein remained at the cell
212 surface after RgpB cleavage. Furthermore, to test whether the anchor strand is released
213 from the NTD after RgpB cleavage, we created cross-linking mutants by introducing
214 cysteine pairs in the FimA(Δ383) mutant after substituting all intrinsic cysteines with
215 alanines (Fig. 3i). Cross-linking these mutants with dithio-bis-maleimidoethane

(DTME) resulted in apparent increased molecular weight of the FimA band, which was prevented by addition of β-mercaptoethanol (βMe) (Fig. 3j). Furthermore, immuno-TEM of the rFimA filament with anti-His-tag antibody showed gold particles localized at one end of the rFimA filament (Extended Data Fig. 6), suggesting that the His-tagged N-terminal strand was still present in the FimA subunit at the base of the FimA filament. Together, these observations strongly suggest that the anchor strand remains attached to the NTD of FimA after RgpB cleavage.

Pilin polymerization relies on sequential elongation from the cell membrane until termination. Simultaneous dissociation of donor and anchor strands would detach the subunit from the cell surface, releasing it into the extracellular environment. Hence, we propose that Type V pili grow from the base (Fig. 4). In *P. gingivalis* strain ATCC 33277, which does not express anchor pilin, pili are attached to the cell surface although they are easily removed by vortexing, consistent with our finding that pili remain anchored to the cell membrane during assembly. Pilus assembly is prevented by deletion of Trp383 at the C-terminal end (Fig. 3 d-g, Extended Data 5b and 5d). Furthermore, peptides corresponding to the C-terminal region inhibit assembly of the minor Mfa pilus[41]. The fact that binding of these C-terminal residues is essential for polymerization suggests that addition of donor strand A2′ to the empty CTD groove of the acceptor pilin precedes replacement of the anchor strand (Fig. 4 III). After A2’ fills the CTD groove, A1’ may be positioned at the NTD-CTD interface to help trigger release of the anchor strand of the acceptor pilin from its NTD, thereby detaching the acceptor subunit from the cell membrane (Fig. 4 IV). As each stalk pilin subunit is added to the assembled pilus (Fig. 4 V), the initial binding of the A2′ strand brings the A1′ strand into close proximity to the CTD groove. The length of Type V pili is not strictly regulated and there is a wide distribution[3]. As long as an ample supply of stalk pilin subunits exists on the cell membrane, pili should be able to continue growing. Binding of a membrane-anchored stalk-pilin multimer to an anchor pilin that does not contain the protease cleavage site, will terminate elongation (Fig. 4 VI-VII). Therefore,

the final pilus number and the distribution of pilus lengths may be governed by expression level and relative abundance of each gene product. Pilus polymerization at the cell surface may start from a tip pilin, which lacks the donor strand. However, the mechanism is not clear. It is unlikely that a tip pilin would be incorporated into a pilus at the filament tip after pilus polymerization. The stalk pilin FimA alone is sufficient for pilus assembly in vivo[40]. Hence, initiation may not require a tip pilin,[40], which would result in a mixed population of tipped and untipped pili.

Pilin polymerization of Type V pili may be similar to that of chaperone-usher pili with respect to their use of donor-strand exchange[2,6]. Our hypothesis about the initiating role played by the C-terminal residues in Type V pili is reminiscent of the so-called "P5" residue in chaperone-usher pilus assembly, which binds an accessible hydrophobic pocket in the acceptor groove before displacing a complementing strand donated by the chaperone[8]. However, there are several differences between Type V pili and chaperone-usher pili. Their overall pilin domain structure is different and donor strands of chaperone-usher pilins are N-terminal, while those of Type V pilins are C-terminal. Finally, chaperone-usher pili require chaperone proteins for assembly, while Type V pili appear to be self-assembled with the aid of arginine-specific proteases.

Methods

Media and growth condition. *P. gingivalis* ATCC 33277 cells were grown under anaerobic conditions (80% N_2, 10% CO_2, 10% H_2) at 37°C in enriched brain–heart infusion (BHI) broth or on enriched tryptosoya (TS) agar plates supplemented with 5 μg/mL hemin (Sigma) and 0.5 μg/mL menadione (Sigma). Luria–Bertani (LB) broth and LB agar plates were used for growth of *E. coli* strains. Antibiotics were used at the following concentrations: ampicillin (50 or 100 μg/mL for *E. coli*) and erythromycin (Em; 10 μg/mL for *P. gingivalis*).

Purification of recombinant FimA and RgpB protease. The genes, *fimA1* and *fimA2*, of *P. gingivalis* strains ATCC 33277 and TDC60, respectively, were amplified with primer pairs F and R, and then cloned into plasmid pET15b. *E.coli* strain BL21 (DE3) was transformed with this plasmid. Primer sequences are available as supplementary information. Recombinant FimA protein with an N-terminal His-tag was purified by affinity chromatography using Talon resin (Clontech, Takara Bio, Japan) or a His Trap^{TM} HP column (GE Healthcare). His-tagged RgpB protease was expressed in *P. gingivalis* strain 662i6H[42] and purified accordingly.

Purification, crystallization, and structure determination of Type I and Type II FimA of *P. gingivalis*. His-FimA1 from strain ATCC 33277 and His-FimA2 (residues 20-390) from strain TDC60 were expressed in *E. coli* and crudely purified by affinity chromatography, as described above. Each protein solution was mixed with thrombin to remove the His-tag, and dialyzed in 1 L of buffer A (500 mM NaCl in 50 mM Tris-HCl, pH 8.0) at 4°C for 12 h. The protein solution was loaded onto a His Trap^{TM} HP column (GE Healthcare) to remove uncleaved proteins, and further purified with a High Load 26/60 Superdex 200 gel filtration column in 10 mM Tris-HCl and 100 mM NaCl (pH 8.0). Peak fractions were collected and concentrated.

Crystallization was performed by the sitting-drop, vapor-diffusion method. Crystallization drops were prepared by mixing the FimA1 or FimA2 solution with an equal volume of reservoir solution. Initial screening was carried out using screening kits (Wizard I and II, Cryo I and II (Emerald Biostructures) and Crystal Screen I and II (Hampton Research)), and then conditions were optimized. FimA1 crystals used for X-ray data collection were obtained at 20°C from a drop prepared by mixing 1.5 μL protein solution (10 mg/mL) containing 20 mM Tris-HCl pH 8.0 / 150 mM NaCl with an equal volume of reservoir solution containing 0.1 M sodium citrate pH 5.6, 1.8 M $(NH_4)_2$ SO_4 and 0.2 M sodium-potassium tartrate. The FimA1 crystal belongs to an orthorhombic space group of $P2_12_12_1$ with unit cell dimensions of a = 35.8 Å, b = 85.5

Å and c = 242.6 Å. FimA2 crystals suitable for X-ray data collection were obtained at 20°C from a drop prepared by mixing 1.5 μL protein solution (47 mg/mL) containing 20 mM Tris-HCl pH 8.0 and 150 mM NaCl with an equivalent volume of reservoir solution containing 0.1 M HEPES pH 7.5, 2.2 M $(NH_4)_2 SO_4$ and 1.5 % (w/v) PEG400. FimA2 was crystallized in an orthorhombic space group of $P2_12_12_1$ with unit cell dimensions of a = 58.9 Å, b = 86.0 Å and c = 102.1 Å.

X-ray diffraction data were collected at synchrotron beamline BL32XU and BL41XU in SPring-8 (Harima, Japan). Crystals were soaked in a solution containing 90% (v/v) reservoir solution and 10% (v/v) glycerol for a few seconds, frozen in liquid nitrogen, and mounted under nitrogen gas flow at 100 K for X-ray data collection at a wavelength of 1.000 Å. Diffraction data were processed with MOSFLM[43] and scaled with AIMLESS[44]. Initial FimA2 crystal data were calculated by molecular-replacement (MR) with Phaser[45] using the FimA4 structure (PDB ID: 4q98) as a search model. The atomic model was constructed with Coot[46] and refined with Phenix.refine[47] to 1.6 Å. Eleven N-terminal residues, 6 C-terminal residues, and 11 residues (N43-D53) in loop A1-B1 were invisible in the electron density map. The Ramachandran plot statistics indicated that 97.4% and 2.6% residues were in the most favorable and allowed regions, respectively. The initial phase of the FimA1 crystal was obtained by MR using the refined FimA2 structure as a search model. The atomic model of FimA1 was constructed with Coot and refined with Phenix.refine to 2.1 Å. During the refinement process, iterative manual modifications were performed. The Ramachandran plot statistics indicated that 97.7% and 2.2% residues were in the most favorable and allowed regions, respectively. Crystallographic statistics can be found in Extended Data Fig. 7a.

***In vitro* assembly of FimA1 pili.** rFimA1 monomers at 1.5 mg/mL concentration were mixed with 1 mg/mL RgpB and incubated at 37℃ for 4 h. Protease inhibitors N-α-p-tosyl-L-lysine-chloromethyl ketone hydrochloride (TLCK) and leupeptin were added to the solution at 0.1 mM concentration to stop assembly. Assembled rFimA1 filaments were precipitated with 4% (w/vol) polyethylene glycol (PEG-6000), 0.2 M NaCl at 4ºC and collected by centrifugation at 21,000*g* for 30 min at 4ºC. Pellets were

resuspended with 20 mM Tris-HCl pH 7.5 and dialyzed against 20 mM Tris-HCl pH 7.5. Assembly rate was estimated by length measurement of purified filaments imaged by negative stain TEM after 30 min incubation with RgpB, and after 60 min incubation (Extended Data Fig. 2 b-d). 100 filaments were picked from randomly imaged areas. Experiments were carried out in triplicate. Filaments shorter than 100 nm could not be precipitated.

Preparation of Fim pili from the Mfa pilus-defective mutant. Fim pili from *P. gingivalis* KDP225 cells (*mfa1*::*cepA*)[3] were prepared by vesicle-depleted supernatant fractionation, as described previously[48]. Briefly, colonies on a TS agar plate were suspended in phosphate buffered saline, vortexed vigorously, and centrifuged at 5,900*g* for 15 min at 4°C. The supernatant contained soluble extracellular proteins and outer membrane vesicles. The latter were removed by centrifugation at 88,000*g* for 60 min at 4°C and the supernatant, containing Fim pili was used for further analysis.

Immunoelectron microscopy. An rFimA filament sample was blocked against nonspecific binding with PBS containing 0.2% BSA for 30 min at 4°C, and was then incubated with primary anti-His-tag antibody (MBL, Aichi, Japan) at a 1:100 dilution in PBS containing 0.2% BSA for 1 h. Filaments were precipitated with 4% (w/vol) polyethylene glycol (PEG-6000), 0.2 M NaCl at 4ºC and collected by centrifugation at 100,000*g* for 30 min at 4ºC. Filaments were resuspended with PBS containing 0.2% BSA and then incubated with secondary goat anti-mouse IgG antibody with 10 nm gold conjugates (Sigma-Aldrich) at 1:20 dilution for 1 h. A 1 µL sample was applied to a carbon-coated EM grid, washed twice with PBS and stained with 1% uranyl acetate.

Construction of *P. gingivalis* strains expressing FimA variants with C-terminal deletions. A plasmid p*mfa1mfa2* containing the protein expression system induced by anhydrotetracycline (aTC), p*mfa1mfa2*::*ermF* with *tetR* and aTC-inducible promoter,

was constructed and integrated into the *mfa1 maf2* gene regions by homologous recombination, resulting in an Mfa pilus-deficient *P. gingivalis* strain. We created four different *P. gingivalis fimA mfa1 mfa2* strains expressing FimA^{+}, FimA(Δ383), FimA(Δ382-383), FimA(Δ381-383). PCR products were obtained with primer pairs of FimAFw/FimARv, FimAFw/FimA383Rv, FimAFw/FimA382Rv, and FimAFw/FimA381Rv, respectively. Primer sequences are available as supplementary information. PCR products were cloned into the pUC118 vector (Takara, Japan) and sequenced. The NcoI and SalI DNA fragments from cloned vectors were inserted into the NcoI and SalI region of the p*mfa1mfa2* vector, resulting in p*mfa1mfa2*-FimA^{+}, p*mfa1mfa2*-FimA(Δ383), p*mfa1mfa2*-FimA(Δ382-383), p*mfa1mfa2*-FimA(Δ381-383). These plasmids were linearized by SphI digestion and introduced into KDP304 cells (*fimA*::*tetQ*)[49] by electroporation. Cells were spread on TS agar containing 10 μg/mL Em and incubated under anaerobic conditions for 7 days. To express various FimA proteins, a preculture in BHI broth was used at a 50-fold dilution to inoculate a new BHI broth containing aTC (100 ng/mL). Cell pellets were resuspended with PBS and mixed with SDS sample buffer containing β-mercaptoethanol. Samples were incubated at 80°C for 10 min. FimA protein was detected by Western blot analysis with anti-FimA antibody.

Construction of *P. gingivalis* strains with FimA Trp383 substitutions. To generate mutated proteins of FimA at Trp383, amplified PCR products with primer pairs confimAFwNcoI/RvFimAW383F, confimAFwNcoI/RvFimAW383Y, confimAFwNcoI/RvFimAW383H, confimAFwNcoI/RvFimAW383A, confimAFwNcoI/RvFimAW383G, confimAFwNcoI/RvFimAW383L, using the ATCC 33277 genome as template DNA, were cloned into pUC118 (Takara, Kusatsu, Japan). Primer sequences are available as supplementary information. Cloned DNAs were verified by DNA sequencing analysis. The lower NcoI-SalI fragment from those cloned plasmids was inserted in corresponding sites of plasmid p*mfa1mfa2*::*ermF* with *tetR* and aTC-inducible promoter. Then, each recombinant plasmid was digested with SphI and

introduced into the *fimA*::*tetQ* (KDP304) strain by electroporation. Transformants were selected on blood agar plates containing 10 μg/mL Em.

Construction of FimA(Δ383 Cys-null with K40C K81C, or with K40C K167C). Full-length mutated genes were synthesized (Integrated DNA technology). DNA sequences are available as supplementary information. Synthetic DNAs were cloned into pUC118. Cloned DNAs were verified by DNA sequencing analysis. Each lower NcoI-SalI fragment from those cloned plasmids was inserted into corresponding sites of plasmid p*mfa1mfa2*::*ermF* with *tetR* and aTC-inducible promoter. Then, each recombinant plasmid was digested with SphI and introduced into the *fimA*::*tetQ* (KDP304) strain by electroporation. Transformants were selected on blood agar plates containing 10 μg/mL Em.

Cys-Cys cross-linking analysis of the anchor strand with NTD. To detect Cys-Cys cross bridges of mutated FimA(Δ383) protein, aTC-induced cell pellets were collected and washed once with PBS, and then gently suspended in PBS including 0, 0.1, 0.2, 0.4 mM dithio-bis-maleimidoethane (DTME) (Tokyo Chemical Industry, Tokyo, Japan), and then incubated at room temperature for 1 h. Cell pellets were then collected and resuspended in PBS with protease inhibitors and then mixed with SDS sample buffer with or without βME. All samples were incubated at 100°C for 10 min. *fimA*::*tetQ* TetR-FimA(Δ383 Cys-null with two residues substituted to Cys) were used as test strains. Preculture (0.1 mL) was added to 5 mL of new BHI broth with aTC (100 ng/mL) and was anaerobically grown overnight. 100 μL of overnight culture was spun down and the supernatant was discarded. The cell pellet was washed once with PBS and suspended in 100 μL PBS with and without DTME at final concentrations of 0.1, 0.2, 0.4 mM. Samples were stored at room temperature for 1 h. Then, samples were spun down at 20,400*g* at 4°C for 1 min. The cell pellet was dissolved with 60 μL of PBS including 0.1 mM TLCK and 0.1 mM leupeptin. 20 μL of 4X SDS sample buffer with or without βME were added. Samples were heat denatured at 100°C for 10 min. 10 μL of samples were applied in each well.

Construction of *P. gingivalis* FimA(Q379A). To generate mutated protein of FimA(Q379A), amplified PCR product with primer pairs of

fimAFwNcoI/fimAG379ARv using ATCC 33277 genome as template DNA was cloned into pUC118 (Takara, Kusatsu, Japan). Primer sequences are available as supplementary information. Cloned DNA was verified by DNA sequencing analysis. The lower NcoI-SalI fragment from the plasmid was inserted into corresponding sites of plasmid p*mfa1mfa2*::*ermF* with *tetR* and aTC-inducible promoter. Then, the recombinant plasmid was digested with SphI and introduced into *fimA*::*tetQ* (KDP304) by electroporation. Transformants were selected on blood agar plates containing 10 μg/mL Em.

Construction of recombinant FimA(Δ383) and FimA(Q379A) proteins in *E. coli.* To generate recombinant FimA(Δ383) and FimA(Q379A) proteins with His-tag in N-terminus, amplified PCR products with primer pairs pET15bFimAFw/pET15bFimAW383stopRv and pET15bFimAFw/pET15bFimAG379ARv using the ATCC 33277 genome as template DNA were cloned into pUC118 (Takara, Kusatsu, Japan). Primer sequences are available as supplementary information. Cloned DNAs were verified by DNA sequencing analysis. Each of the smaller NdeI-BamHI fragments from those cloned plasmids was inserted into corresponding sites of pET15b (Novagen). *E. coli* strain BL21 (DE3) was transformed with each plasmid.

Gel electrophoresis and immunoblot analysis. SDS-PAGE and immunoblot analysis were performed as previously described[3]. Anti-FimA[3] or anti-Hemin-binding-protein-35 (HBP35)[50] rabbit polyclonal primary antibodies was used for dot blots and Western blots. HRP-conjugated secondary antibody (Sigma) was used for detection.

Cryo-EM sample preparation and data collection. Three μL of sample solution was applied to holey carbon grids (Quantifoil 1.2/1.3 Cu 400-mesh) after plasma-cleaning with a Solarus (Gatan, Pleasanton, USA) for 30 s at 25°C in a 23% H_2, 77% O_2 gas mix. Grids were blotted for 3 s at 4°C and 100% relative humidity, then immediately plunge-frozen in liquid ethane with a Vitrobot Mark IV (Thermo Fisher Scientific, Hillsboro, USA (TFS)). Image data were collected on a Talos Arctica TEM (TFS)

operating at 200 kV in nanoprobe mode using parallel illumination and coma-free alignment, on a Falcon-3 direct electron detector (TFS) in counting mode at a calibrated magnification of 134,048× (corresponding to a pixel size of 1.12 Å at the specimen level) with defocus range from -1.5 to -2.5 µm. The microscope and camera were controlled using EPU software (TFS). Images were recorded as movies at a dose rate of 0.77 el/Å^2/sec at 40 frame/sec, 60-second exposure, with an accumulated total dose of 46 el/ Å^2. Cryo-EM statistics can be found in Extended Data Fig. 7b.

Cryo-EM image processing. Frames in 1,153 movies were aligned using motioncor2[51], dose-weighted, and summed. 831,459 FimA filament fragments were picked by relion_autopick in RELION 2.1[52] using a Gaussian blob as template. Particle coordinates were exported to the cisTEM[53] software package. All subsequent operations were carried out with cisTEM. Particles were boxed with dimension 384 × 384 pixel covering 3-4 FimA molecules within a circular mask, followed by 2D classification into 50 classes. 159,300 particles from good classes (members of class-average images in which the filament shape was recognizable) were retained. The initial 3D alignment model was created with Spider[54] by command “MO 3” as a cylinder (diameter 50 Å, length 210 Å), used as an initial template in the auto-refinement function of cisTEM with three 3D classes. 3D classes were qualitatively similar. The class with the highest resolution, containing 118,143 particles, was further split into 3 classes, followed by additional refinement. The best of these 3D classes refined to 3.6 Å resolution, including CTF refinement and likelihood blurring. No symmetry was applied. The final reconstruction contained 61,728 particles. The final figure-of-merit-weighted electron potential map was sharpened with a B-factor of -150 Å^{-1} at 3.6 Å resolution. Resolution was estimated using a Fourier shell correlation between two independently refined half sets as implemented in cisTEM (FSC threshold 0.143) (Extended Data Fig. 3 and 7b). Although the angular distribution of Euler angles showed a strong preference for certain orientations, these were clustered along an arch spanning the hemisphere of the angular

plot and all orientations outside this preferred region in the plot were represented by at least one particle image (Extended Data Fig. 3). Absolute hand of the map was validated by docking of the FimA1 crystal structure (see below). The cryo-EM reconstruction of the Fim pilus from the Mfa pilus-deficient mutant was created using the same procedure.

Cryo-EM model building and refinement. The crystal structure of FimA pilin was docked as a rigid body into the cryo-EM density map using *UCSF Chimera*[55]. The initial model (PDB 6JZJ) was manually adjusted, and C-terminal residues in the donor strand were manually built with inverted sequence using Coot[46]. The model containing three FimA molecules was refined using PHENIX_real_space refine[47]. The hydrophobicity map was colored according to the Kyte-Doolittle scale implemented in *UCSF Chimera*[56].

Data availability

The atomic coordinates of the crystal structures have been deposited into the RCSB Protein Data Bank (FimA1: PDB ID 6JZJ, FimA2: PDB ID 6JZK). The coordinates of the cryo-EM based model have been deposited under PDB ID 6KMF. The cryo-EM map has been deposited into the EM database under accession code EMD-0724. Uncropped gel or Western blot scans from Fig. 3d, 3f, 3h, 3j and Extended Data Fig. 2a, 2e, 5b are available as source data Fig 3, source data Extended Data Fig 2, and source data Extended Data Fig 5, respectively. Filament length data used to create Extended Data Fig. 2d is available as source data Extended data Fig 2d. Additional raw data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Allen, W. J., Phan, G. & Waksman, G. Pilus biogenesis at the outer membrane of Gram-negative bacterial pathogens. *Curr Opin Struct Biol* **22**, 500–506 (2012).

2. Hospenthal, M. K., Costa, T. R. D. & Waksman, G. A comprehensive guide to pilus biogenesis in Gram-negative bacteria. *Nat Rev Microbiol* **15**, 365–379 (2017).
3. Xu, Q. *et al.* A Distinct Type of Pilus from the Human Microbiome. *Cell* **165**, 690–703 (2016).
4. Lamont, R. J. & Jenkinson, H. F. Life below the gum line: pathogenic mechanisms of Porphyromonas gingivalis. *Microbiol. Mol. Biol. Rev.* **62**, 1244–63 (1998).
5. Wang, F. *et al.* Structure of Microbial Nanowires Reveals Stacked Hemes that Transport Electrons over Micrometers. *Cell* **177**, 361-369.e10 (2019).
6. Filman, D. J. *et al.* Cryo-EM reveals the structural basis of long-range electron transport in a cytochrome-based bacterial nanowire. *Commun. Biol.* **2**, 219 (2019).
7. Lukaszczyk, M., Pradhan, B. & Remaut, H. The Biosynthesis and Structures of Bacterial Pili. *Subcell. Biochem.* **92**, 369–413 (2019).
8. Remaut, H. *et al.* Donor-Strand Exchange in Chaperone-Assisted Pilus Assembly Proceeds through a Concerted β Strand Displacement Mechanism. *Mol. Cell* **22**, 831–842 (2006).
9. Holt, S. C. & Ebersole, J. L. Porphyromonas gingivalis, Treponema denticola, and Tannerella forsythia: the 'red complex', a prototype polybacterial pathogenic consortium in periodontitis. *Periodontol. 2000* **38**, 72–122 (2005).
10. Demmer, R. T. & Desvarieux, M. Periodontal infections and cardiovascular disease: the heart of the matter. *J Am Dent Assoc* **137 Suppl**, 14S-20S; quiz 38S (2006).
11. Michaud, D. S. *et al.* Plasma antibodies to oral bacteria and risk of pancreatic cancer in a large European prospective cohort study. *Gut* **62**, 1764–1770 (2013).
12. Leech, M. T. & Bartold, P. M. The association between rheumatoid arthritis and periodontitis. *Best Pr. Res Clin Rheumatol* **29**, 189–201 (2015).
13. Dominy, S. S. *et al.* Porphyromonas gingivalis in Alzheimer's disease brains: Evidence for disease causation and treatment with small-molecule inhibitors. *Sci Adv* **5**, eaau3333 (2019).
14. Yoshimura, F., Takahashi, K., Nodasaka, Y. & Suzuki, T. Purification and characterization of a novel type of fimbriae from the oral anaerobe Bacteroides

gingivalis. *J Bacteriol* **160**, 949–957 (1984).
15. Hamada, N., Sojar, H. T., Cho, M. I. & Genco, R. J. Isolation and characterization of a minor fimbria from Porphyromonas gingivalis. *Infect Immun* **64**, 4788–4794 (1996).
16. Hajishengallis, G., Shakhatreh, M.-A. K., Wang, M. & Liang, S. Complement Receptor 3 Blockade Promotes IL-12-Mediated Clearance of Porphyromonas gingivalis and Negates Its Virulence In Vivo. *J. Immunol.* **179**, 2359–2367 (2007).
17. Amano, A. Bacterial adhesins to host components in periodontitis. *Periodontol 2000* **52**, 12–37 (2010).
18. Park, Y. *et al.* Short fimbriae of Porphyromonas gingivalis and their role in coadhesion with Streptococcus gordonii. *Infect Immun* **73**, 3983–3989 (2005).
19. Sojar, H. T., Hamada, N. & Genco, R. J. Isolation and characterization of fimbriae from a sparsely fimbriated strain of Porphyromonas gingivalis. *Appl. Environ. Microbiol.* **63**, 2318–2323 (1997).
20. Hasegawa, Y. *et al.* Anchoring and length regulation of Porphyromonas gingivalis Mfa1 fimbriae by the downstream gene product Mfa2. *Microbiology* **155**, 3333–3347 (2009).
21. Nagano, K., Hasegawa, Y., Murakami, Y., Nishiyama, S. & Yoshimura, F. FimB Regulates FimA Fimbriation in Porphyromonas gingivalis. *J Dent Res* **89**, 903–908 (2010).
22. Hamada, S. *et al.* Molecular and immunological characterization of the fimbriae of Porphyromonas gingivalis. *Microbiol Immunol* **38**, 921–930 (1994).
23. Fujiwara, T., Nakagawa, I., Morishima, S., Takahashi, I. & Hamada, S. Inconsistency between the fimbrilin gene and the antigenicity of lipopolysaccharides in selected strains of Porphyromonas gingivalis. *FEMS Microbiol Lett* **124**, 333–341 (1994).
24. Nakagawa, I. *et al.* Distribution and molecular characterization of Porphyromonas gingivalis carrying a new type of fimA gene. *J Clin Microbiol* **38**, 1909–1914 (2000).
25. Nakagawa, I. *et al.* Identification of a new variant of fimA gene of Porphyromonas gingivalis and its distribution in adults and disabled populations with periodontitis. *J Periodontal Res* **37**, 425–432 (2002).

26. Okuda, S. & Tokuda, H. Lipoprotein Sorting in Bacteria. *Annu. Rev. Microbiol.* **65**, 239–259 (2011).
27. Korotkov, K. V *et al.* Calcium is essential for the major pseudopilin in the type 2 secretion system. *J Biol Chem* **284**, 25466–25470 (2009).
28. Shoji, M. *et al.* The major structural components of two cell surface filaments of Porphyromonas gingivalis are matured through lipoprotein precursors. *Mol Microbiol* **52**, 1513–1525 (2004).
29. Konovalova, A. & Silhavy, T. J. Outer membrane lipoprotein biogenesis: Lol is not the end. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**, (2015).
30. Nakayama, K., Yoshimura, F., Kadowaki, T. & Yamamoto, K. Involvement of arginine-specific cysteine proteinase (Arg-gingipain) in fimbriation of Porphyromonas gingivalis. *J Bacteriol* **178**, 2818–2824 (1996).
31. Shoji, M. *et al.* Recombinant Porphyromonas gingivalis FimA preproprotein expressed in Escherichia coli is lipidated and the mature or processed recombinant FimA protein forms a short filament in vitro. *Can J Microbiol* **56**, 959–967 (2010).
32. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
33. Lee, J. Y. *et al.* Maturation of the Mfa1 Fimbriae in the Oral Pathogen Porphyromonas gingivalis. *Front Cell Infect Microbiol* **8**, 137 (2018).
34. Nishiyama, M., Ishikawa, T., Rechsteiner, H. & Glockshuber, R. Reconstitution of pilus assembly reveals a bacterial outer membrane catalyst. *Science (80-.).* **320**, 376–379 (2008).
35. Nishiyama, S. -i. S. I. *et al.* Involvement of minor components associated with the FimA fimbriae of Porphyromonas gingivalis in adhesive functions. *Microbiology* **153**, 1916–1925 (2007).
36. Hajishengallis, G., Ratti, P. & Harokopakis, E. Peptide mapping of bacterial fimbrial epitopes interacting with pattern recognition receptors. *J Biol Chem* **280**, 38902–38913 (2005).
37. Nakano, K. *et al.* Comparison of inflammatory changes caused by

608 Porphyromonas gingivalis with distinct fimA genotypes in a mouse abscess
609 model. *Oral Microbiol. Immunol.* **19**, 205–209 (2004).
610 38. Sugano, N. *et al.* Differential cytokine induction by two types of Porphyromonas
611 gingivalis. *Oral Microbiol Immunol* **19**, 121–123 (2004).
612 39. Bodet, C., Chandad, F. & Grenier, D. Porphyromonas gingivalis-induced
613 inflammatory mediator profile in an ex vivo human whole blood model. *Clin.*
614 *Exp. Immunol.* **143**, 50–57 (2006).
615 40. Nagano, K. *et al.* Porphyromonas gingivalis FimA fimbriae: fimbrial assembly
616 by fimA alone in the fim gene cluster and differential antigenicity among fimA
617 genotypes. *PLoS One* **7**, e43722 (2012).
618 41. Alaei, S. R., Park, J. H., Walker, S. G. & Thanassi, D. G. Peptide-Based
619 Inhibitors of Fimbrial Biogenesis in Porphyromonas gingivalis. *Infect. Immun.* **87**,
620 e00750-18 (2019).
621 42. Veillard, F. *et al.* Purification and characterisation of recombinant His-tagged
622 RgpB gingipain from Porphymonas gingivalis. *Biol Chem* **396**, 377–384 (2015).
623 43. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W.
624 iMOSFLM: A new graphical interface for diffraction-image processing with
625 MOSFLM. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 271–281 (2011).
626 44. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta*
627 *Crystallogr D Biol Crystallogr* **67**, 235–242 (2011).
628 45. McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–
629 674 (2007).
630 46. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development
631 of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501 (2010).
632 47. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for
633 macromolecular structure solution. *Acta Crystallogr. Sect. D* **66**, 213–221 (2010).
634 48. Chen, T., Nakayama, K., Belliveau, L. & Duncan, M. J. Porphyromonas
635 gingivalis gingipains and adhesion to epithelial cells. *Infect. Immun.* **69**, 3048–56
636 (2001).
637 49. Kikuchi, Y. *et al.* Novel stationary-phase-upregulated protein of Porphyromonas
638 gingivalis influences production of superoxide dismutase, thiol peroxidase and
639 thioredoxin. *Microbiology* **151**, 841–853 (2005).
640 50. Shoji, M. *et al.* Characterization of hemin-binding protein 35 (HBP35) in

Porphyromonas gingivalis: its cellular distribution, thioredoxin activity and role in heme utilization. *BMC Microbiol.* **10**, 152 (2010).

51. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
52. Scheres, S. H. W. Chapter Six - Processing of Structurally Heterogeneous Cryo-EM Data in RELION. in *Methods in Enzymology* (ed. Crowther, R. A.) **579**, 125–157 (Academic Press, 2016).
53. Grant, T., Rohou, A. & Grigorieff, N. cisTEM, user-friendly software for single-particle image processing. *Elife* **7**, (2018).
54. Frank, J. *et al.* SPIDER and WEB: Processing and Visualization of Images in 3D Electron Microscopy and Related Fields. *J. Struct. Biol.* **116**, 190–199 (1996).
55. Pettersen, E. F. *et al.* UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J Comput Chem* **25**, 1605–1612 (2004).
56. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105–132 (1982).

Acknowledgements

The authors thank J. Potempa (University of Louisville) for kindly providing the His-tagged RgpB expressing *P. gingivalis* strain and S. Aizawa (Prefectural University of Hiroshima) for critically reading the manuscript. We thank Steven D. Aird for technical editing. X-ray diffraction data were collected at synchrotron beamlines BL32XU and BL41XU in SPring-8 (Harima, Japan) with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (Proposal 2014B1478, 2016A2539, 2017A2588 and 2018A2569). This work was supported by the Platform Project for Supporting Drug Discovery and Life Science Research (BINDS) from AMED, under grant number JP18am0101076 (to M.W.), by JSPS KAKENHI Grants JP17K17085 and JP19K10083 (to S.S.), by JSPS KAKENHI Grant JP16H05504 (to K.N., K.I. and M.S.), and by a JSPS KAKENHI Grant JP17K07318 (to H.M.). M.W. was supported by direct funding from Okinawa Institute of Science and Technology Graduate University.

Author contributions

S.S., M.S., K.I., K.N., and M.W. conceived and designed experiments. M.S., K.O., and S.S. performed molecular cloning and protein purification. M.S. created mutants. S.S., H.M., M.M., and M.W. carried out cryo-EM experiments and image processing. H.M. and M.M. built atomic models into cryo-EM maps. K.O. and K.I. performed crystallization experiments and analysis. S.S. wrote the first draft. All authors analyzed results and contributed to writing the paper.

Corresponding authors

Correspondence to Koji Nakayama, Katsumi Imada, or Matthias Wolf.

Competing Interests

The authors declare no competing interests.

Figure legends

Figure 1. Crystal structure of the monomeric, self-complementing form of FimA1 pilin. a, Schematic of full-length FimA1 protein architecture. The signal sequence enables export and membrane integration by the Sec system. The conserved cysteine serves as a lipidation site. The RgpB cleavage site is marked with a “scissors” symbol. The β-strands A1 (residues 32-37), A1′ (365-374), and A2′ (376-383) are depicted as arrows. **b**, The crystal structure of monomeric rFimA1 contains ordered density for Glu30-Gln379. Anchor strand A1 (blue) and donor strand A1′-A2′ (red) are in the self-complementing (closed) conformation, in which the C-terminal strand (red) is folded back and securely tucked under the N-terminal loop A1-B1 (blue). The dashed red line indicates the disordered C-terminal end of the A2′ polypeptide. The scissors icon and blue sphere mark the RgpB cleavage site at Arg46. Scale bar, 20 Å.

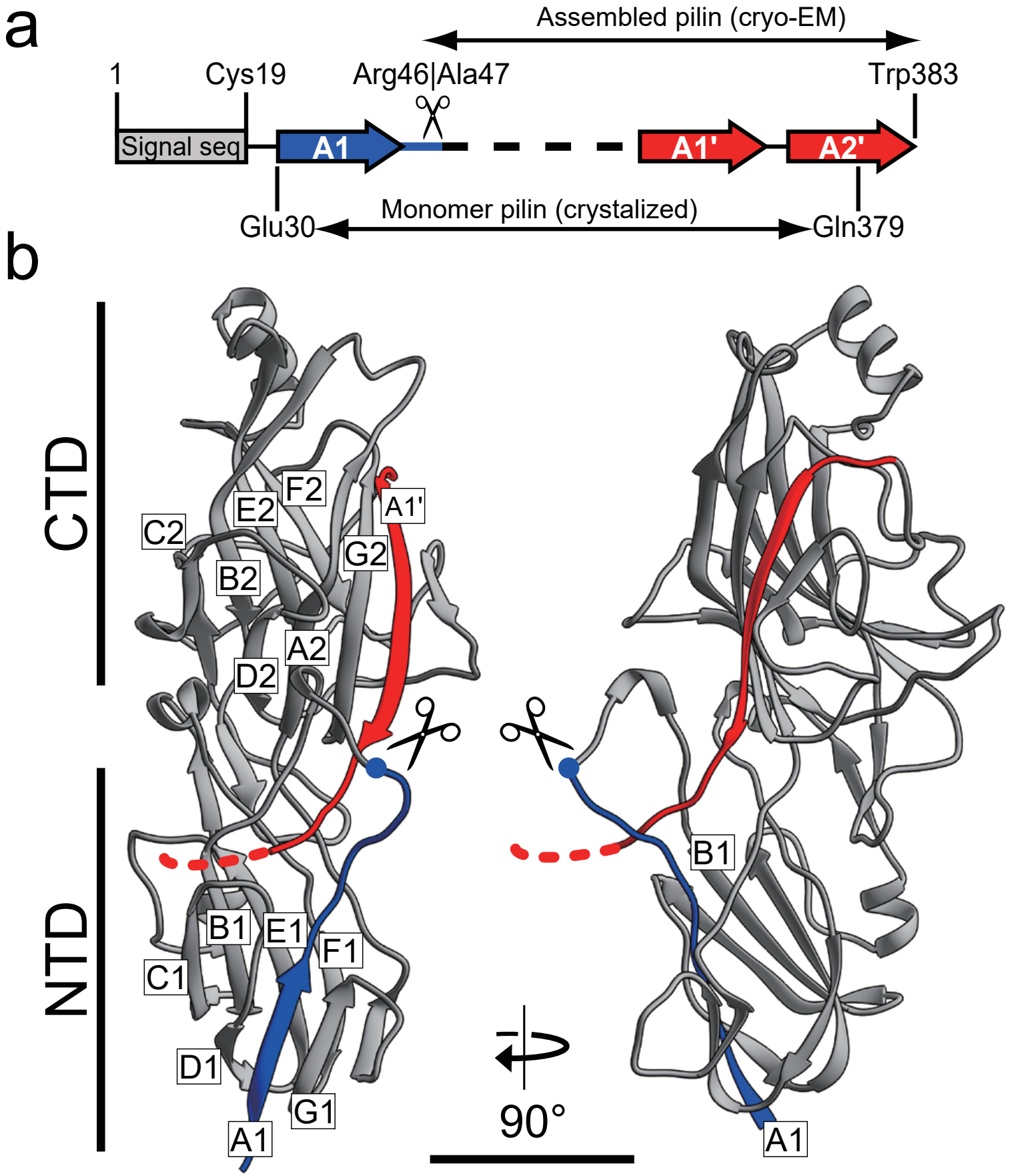
Figure 2. Cryo-EM structure of the polymerized, donor-strand-exchanged form of FimA1 after RgpB cleavage. a, Top row: Electron micrographs of negatively stained

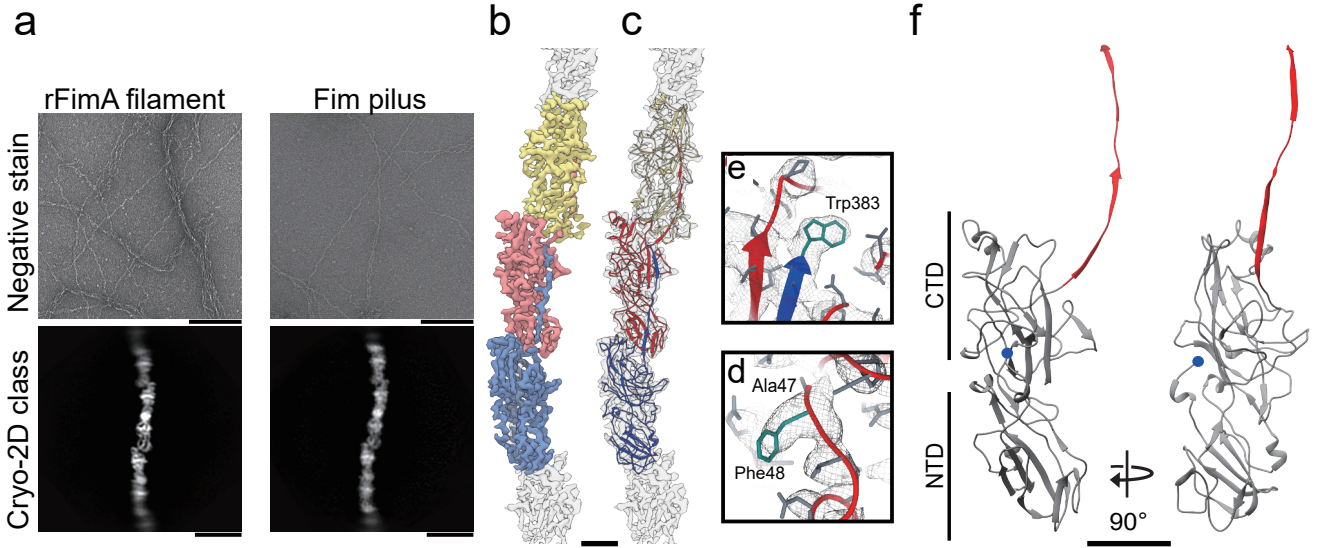
pili. Scale bar, 100 nm. Both results were reproduced 3 times. Bottom row: Reference-free 2D class averages from cryo-EM images of approximately 2,000 boxed, aligned filament segments. Scale bars, 100 Å. Left column: Self-assembled rFimA filament *in vitro*. Right column: A purified Fim pilus from Mfa pilus-defective mutant. **b**, An iso-electron potential surface representation of cryo-EM reconstruction (contoured at 5.5σ above mean) of an rFimA filament with individual pilin subunits highlighted in color. **c**, Same as B, including a fitted atomic model in ribbon representation. **d**, **e**, Ends of the FimA polypeptide chain. The reconstructed electron potential shows the characteristic shape of N-terminal Phe48 (**d**) and C-terminal residue Trp383 (**e**) in the rFimA filament. **f**, An individual rFimA subunit, showing donor strand A1′-A2′ in the donor-strand exchanged conformation. Scale bars (b, c, f), 20 Å.

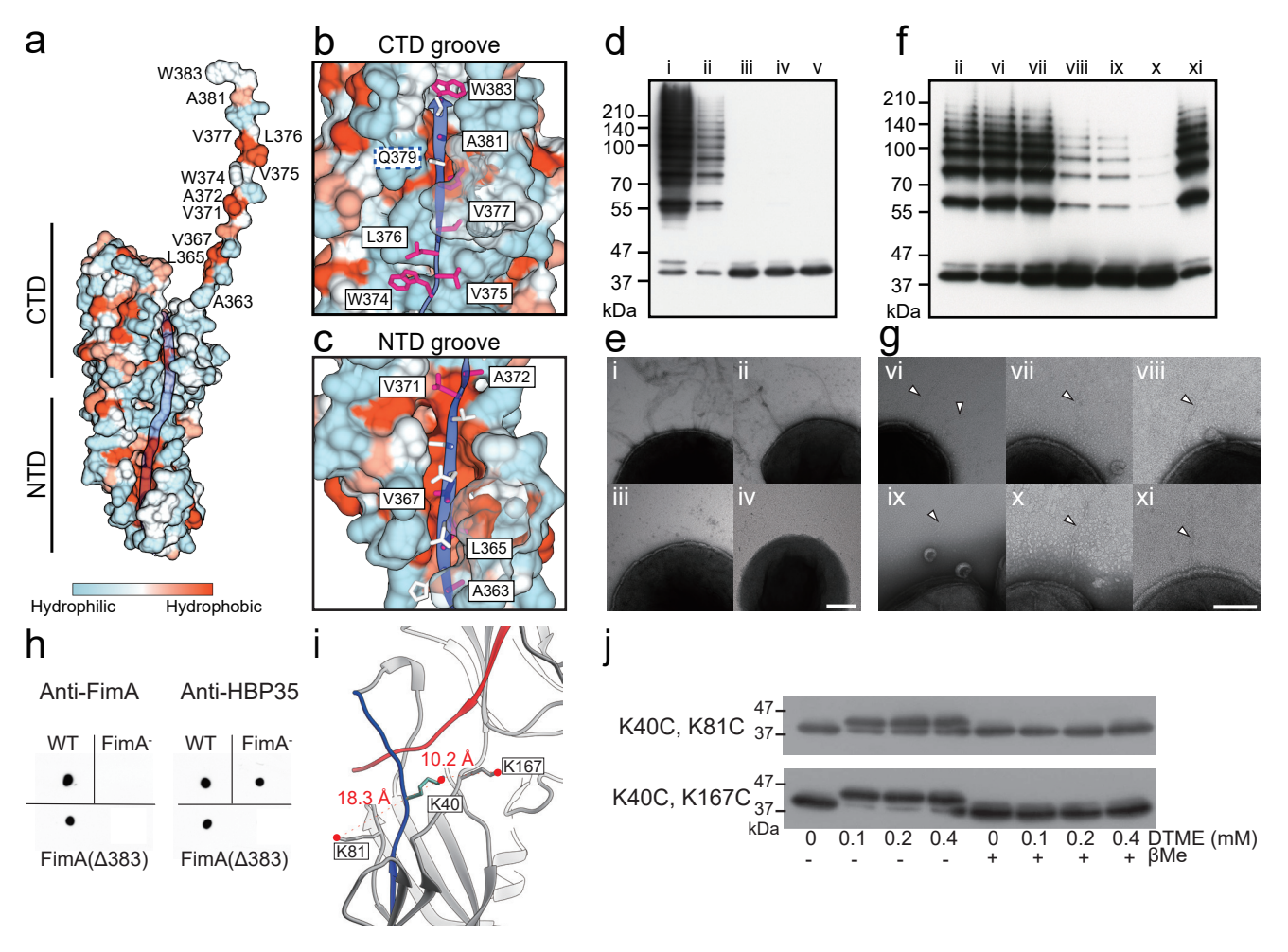
Figure 3. Integration of donor strand A1'-A2' into the acceptor pilin. a, A solvent-excluded surface generated from the atomic model of the donor-strand exchanged form FimA, colored by hydrophobicity. The bound donor strand from the next subunit is shown as a blue tube. **b**, A close-up of the NTD groove with strand A2′. **c**, A close-up of the CTD groove containing strand A1′. Hydrophobic residues and the hydrophilic Gln379 (blue dashed box) on the donor strand are labelled. **d**, Western blot of pilus-defective (*fimA mfa1 mfa2*) mutants expressing FimA with C-terminal deletions. Samples were heated at 80°C for 10 min. Lanes: (i), wild type (ATCC 33277); (ii), $FimA^+$; (iii), FimA(Δ383); (iv), FimA(Δ382-383); (v), FimA(Δ381-383). Primary antibodies against FimA. Ladder patterns indicate FimA polymerization. **e**, Electron micrographs of negatively stained pilus-defective mutants expressing FimA with a C-terminal deletion. Labels are the same as in (d). Scale bar, 200 nm. **f**, Western blot of pilus-defective (*fimA mfa1 mfa2*) mutants expressing FimA with C-terminal substitutions. Samples were heated at 80°C for 10 min. Lanes: (ii), $FimA^+$; (vi), FimA(W383F); (vii), FimA(W383Y); (viii), FimA(W383H); (ix), FimA(W383A); (x), FimA(W383G); (xi), FimA(W383L). **g**, Electron micrographs of substitution mutants from (f). Labels are the same as in (f). Scale bar, 200 nm. **h**, **i**, **j**, Anchor strand A1 is not released by RgpB cleavage *in vivo*. **h**, Immuno dot blot of wild-type (ATCC 33277) cells, Fim pilus-defective (*fimA*) mutant cells, and pilus-defective (*fimA mfa1 mfa2*) mutants expressing FimA(Δ383). Primary antibodies against FimA and Hemin binding protein 35 (HBP35), a known surface-attached protein. **i**, We substituted all intrinsic

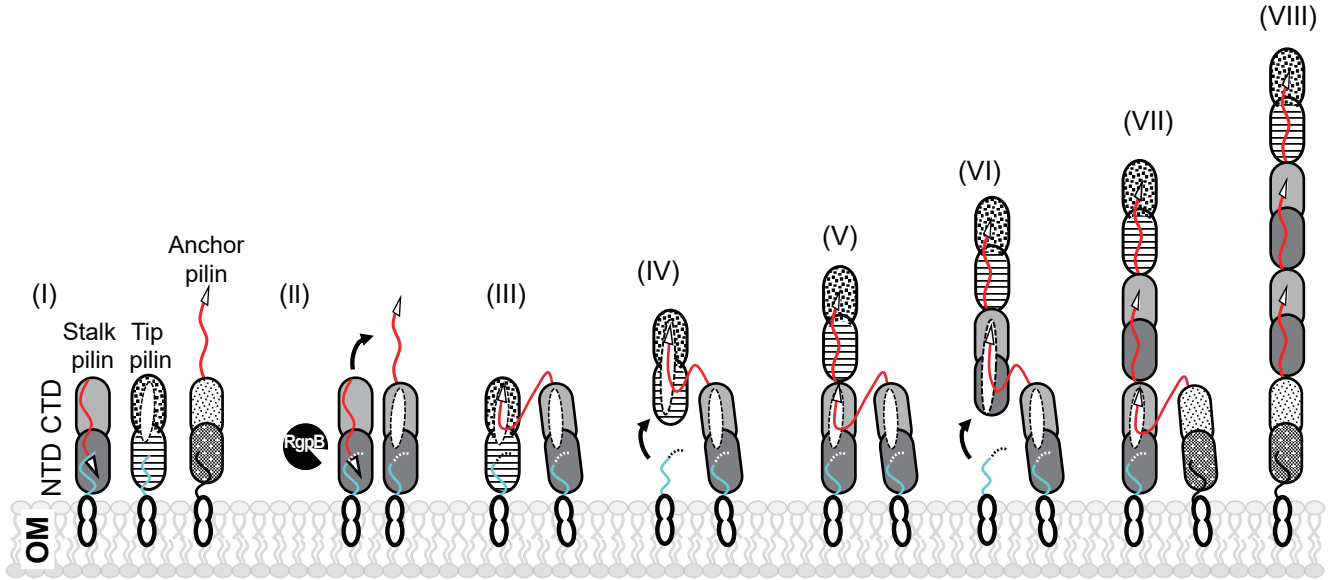
cysteines with alanines in FimA(Δ383) and created two mutants by introducing cysteine pairs in positions (K40, K81) and (K40, K167) for cross-linking the anchor strand (blue) with adjacent β-strands in the NTD. The cross linker dithio-bis-maleimidoethane (DTME, length 13.3 Å) could bridge the distance. **j**, Western blot of the two mutant strains in (i) subjected to increasing concentrations of DTME, with or without the reducing agent, β-mercaptoethanol (βMe). Samples were heated at 100°C for 10 min. Primary antibodies against FimA. Experiments for (d, f, h) were carried out in duplicate, and for (j) in triplicate. Each image in (e, g) is representative for 30 observations of cells, repeated twice.

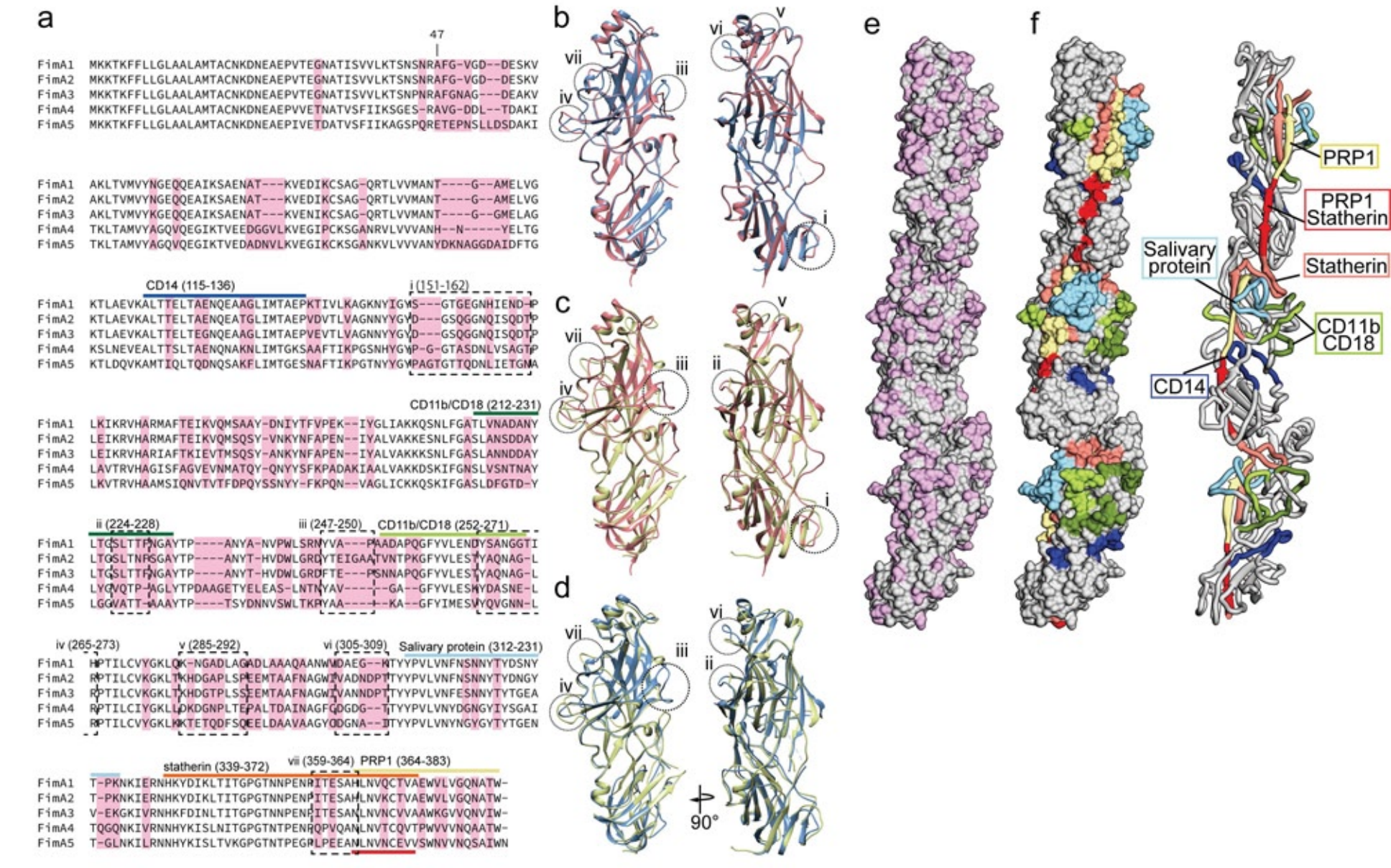
Figure 4. Model for Type V pilus assembly. (I) There are three types of pilin (anchor, stalk and tip). All are lipoproteins that are lipidated at a conserved cysteine after removal of the signal sequence during export. Lipid anchors at their N-terminal domains (NTD) attach the subunit to the outer membrane (OM) on the cell surface. We refer to strand A1 (Figure 1) as the anchor strand (black or cyan wavy line). The donor strand (red) of the stalk pilin is folded back in the self-complementing form and secured by the anchor strand (cyan) at the RgpB protease cleavage site. The anchor pilin does not have a protease cleavage site and its donor strand is extended. The tip pilin does not contain a donor strand (white ellipse indicates the unfilled groove in the β-sheet). (II) RgpB cleavage removes the safety pin and releases the donor strand of a stalk pilin, exposing a hydrophobic groove in the C-terminal domain (CTD). The pilin remains membrane-attached with its anchor strand. (III) An adjacent pilin in the same cleaved state binds with its donor strand to the unfilled groove of the receptor CTD of a tip or stalk pilin, releasing (IV) the anchor strand from the NTD and freeing the subunit from the cell membrane. (V) Sequential addition of stalk pilins extends the growing pilus. (VI-VIII) Random binding to a non-cleavable anchor pilin terminates polymerization. If assembly starts without a tip pilin, the end of the pilus will be untipped.

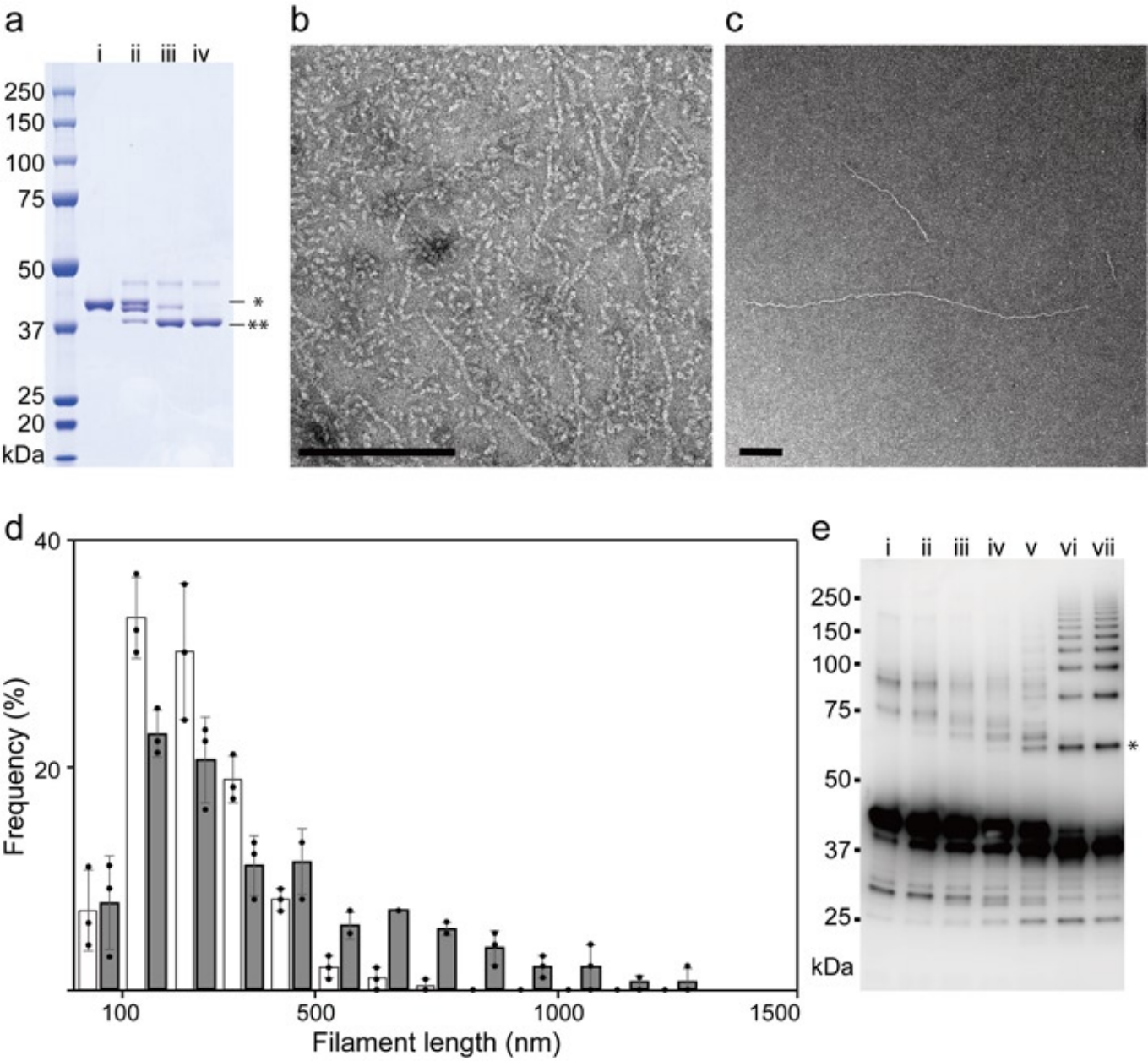

a
Assembled pilin (cryo-EM)
1
Cys19
Arg46|Ala47
Trp383
Signal seq
A1
A1'
A2'
Glu30
Monomer pilin (crystalized)
Gln379
b
CTD
NTD
F2
E2
A1'
C2
G2
B2
A2
D2
B1
E1
F1
C1
D1
G1
A1
90°
B1
A1

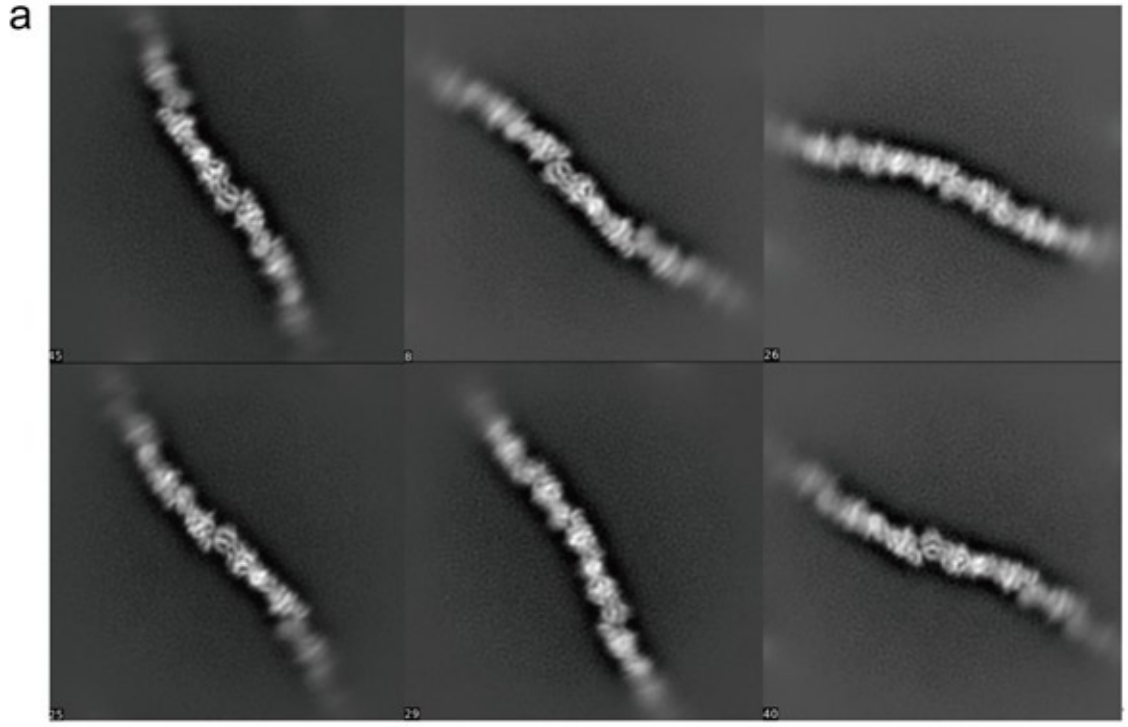

a
rFimA filament
Fim pilus
Negative stain
Cryo-2D class
b
c
e
Trp383
d
Ala47
Phe48
f
CTD
NTD
90°

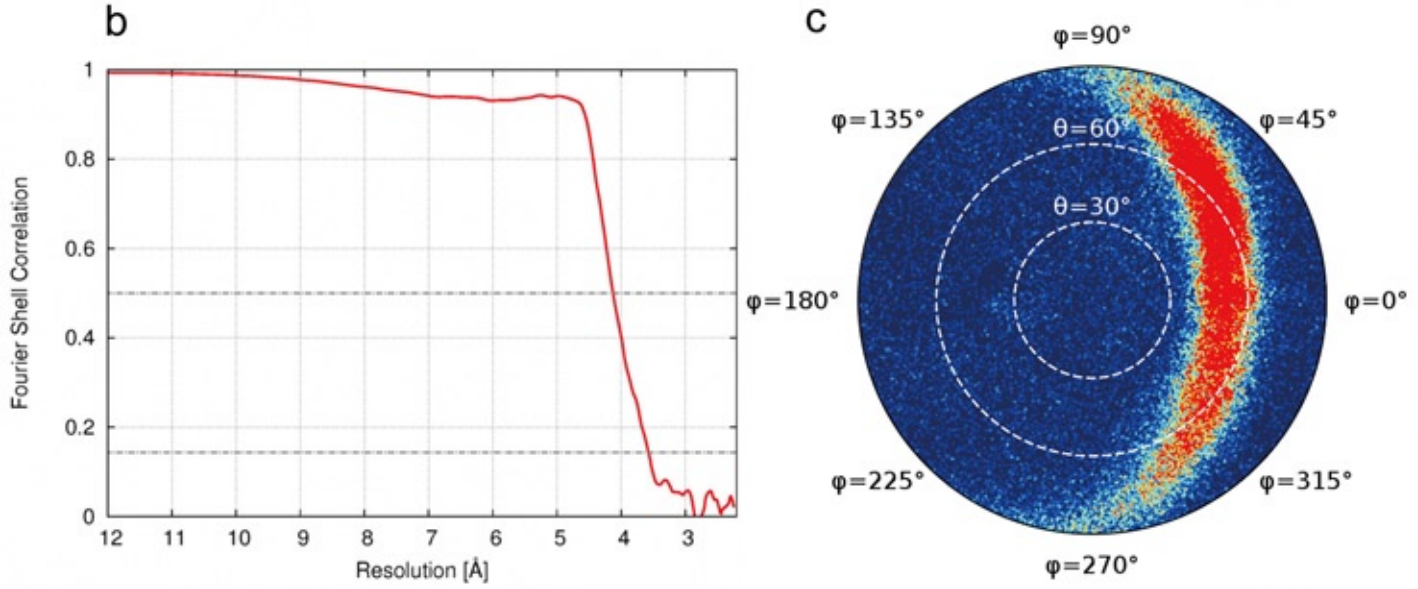

a
W383
A381
V377
L376
W374
A372
V375
V371
V367
L365
A363
CTD
NTD
Hydrophilic
Hydrophobic
b
CTD groove
W383
A381
Q379
V377
L376
W374
V375
c
NTD groove
V371
A372
V367
L365
A363
d
i ii iii iv v
210
140
100
70
55
47
37
kDa
f
ii vi vii viii ix x xi
210
140
100
70
55
47
37
kDa
e
i ii iii iv
g
vi vii viii ix x xi
h
Anti-FimA
Anti-HBP35
WT
FimA-
WT
FimA-
FimA(Δ383)
FimA(Δ383)
i
10.2 Å
18.3 Å
K167
K40
K81
j
K40C, K81C
K40C, K167C
47
37
kDa
0 0.1 0.2 0.4 0 0.1 0.2 0.4 DTME (mM)
- - - - + + + + βMe

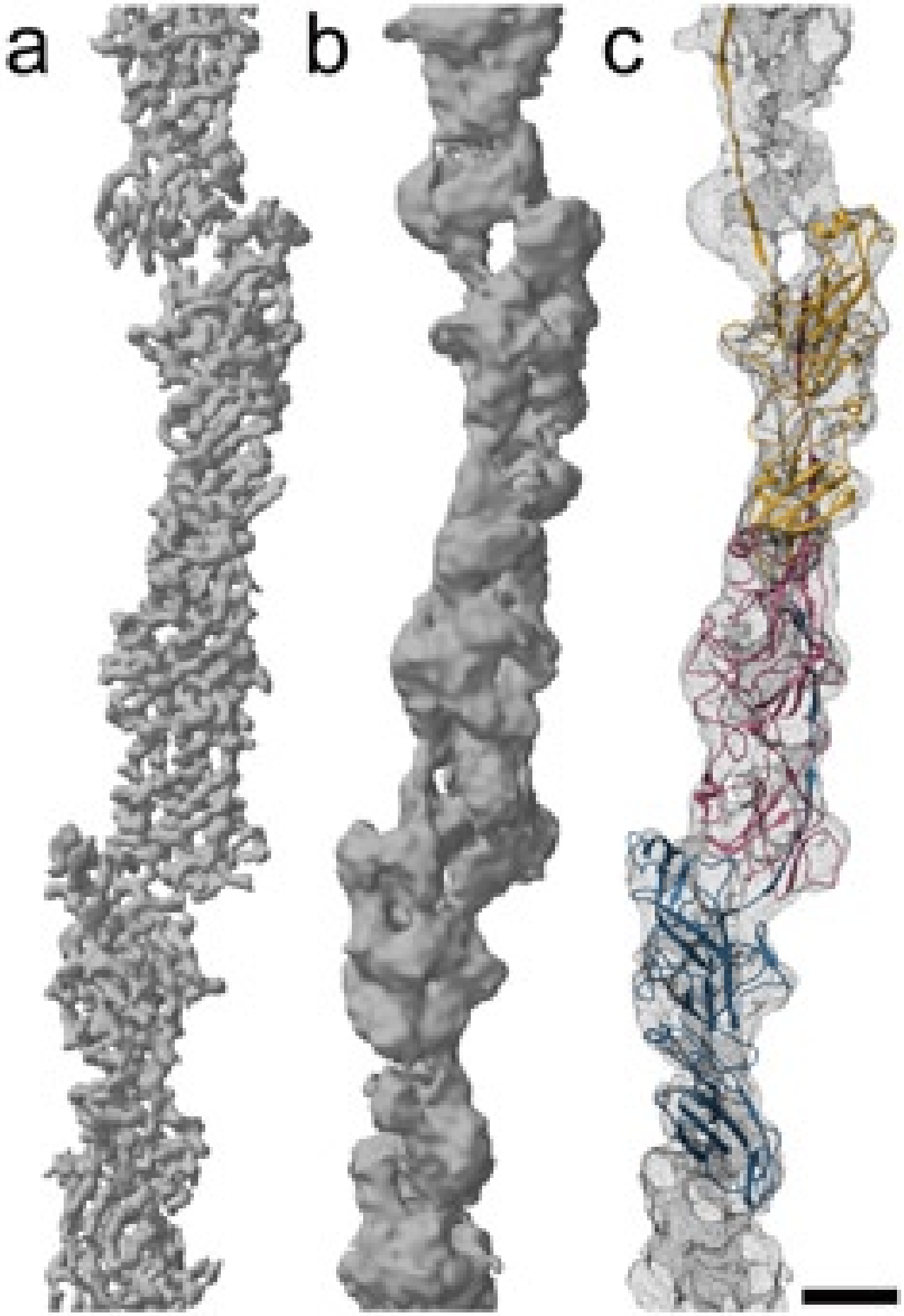

(I)
Stalk pilin
Tip pilin
Anchor pilin
NTD CTD
OM
(II)
RgpB
(III)
(IV)
(V)
(VI)
(VII)
(VIII)

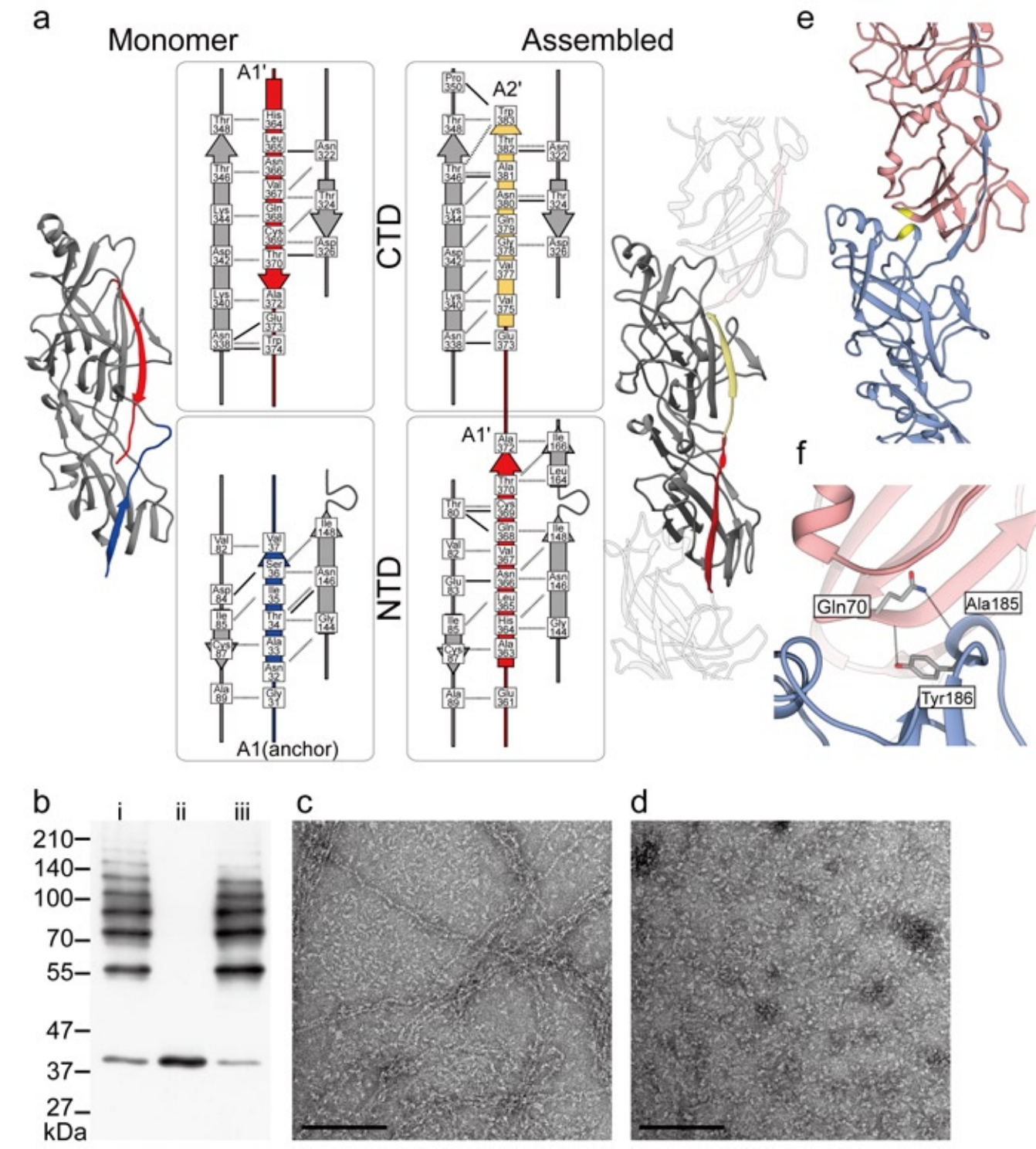

a
47
FimA1 MKKTKFFLLGLAALAMTACNKDNEAEPVTEGNATISVVLKTSNSNRAFG-VGD--DESKV
FimA2 MKKTKFFLLGLAALAMTACNKDNEAEPVTEGNATISVVLKTSNSNRAFG-VGD--DESKV
FimA3 MKKTKFFLLGLAALAMTACNKDNEAEPVTEGNATISVVLKTSNPNRAFGNAG---DEAKV
FimA4 MKKTKFFLLGLAALAMTACNKDNEAEPVVETNATVSFIIKSGES-RAVG-DDL--TDAKI
FimA5 MKKTKFFLLGLAALAMTACNKDNEAEPIVETDATVSFIIKAGSPQRETEPNSLLDSDAKI
FimA1 AKLTVMVYNGEQQEAIKSAENAT---KVEDIKCSAG-QRTLVVMANT----G--AMELVG
FimA2 AKLTVMVYNGEQQEAIKSAENAT---KVEDIKCSAG-QRTLVVMANT----G--AMELVG
FimA3 AKLTVMVYKGEQQEAIKSAENAT---KVENIKCSAG-QRTLVVMANT----G--GMELAG
FimA4 TKLTAMVYAGQVQEGIKTVEEDGGVLKVEGIPCKSGANRVLVVVANH--N-----YELTG
FimA5 TKLTAMVYAGQVQEGIKTVEDADNVLKVEGIKCKSGANKVLVVVANYDKNAGGDAIDFTG
CD14 (115-136)
i (151-162)
FimA1 KTLAEVKALTTELTAENQEAAGLIMTAEPKTIVLKAGKNYIGYS---GTGEGNHIEND-P
FimA2 KTLAEVKALTTELTAENQEATGLIMTAEPVDVTLVAGNNYYGYD---GSQGGNQISQDTP
FimA3 KTLAEVKALTTELTEGNQEAAGLIMTAEPVEVTLVAGNNYYGYD---GSQGGNQISQDTP
FimA4 KSLNEVEALTTSLTAENQNAKNLIMTGKSAAFTIKPGSNHYGYP-G-GTASDNLVSAGTP
FimA5 KTLDQVKAMTIQLTQDNQSAKFLIMTGESNAFTIKPGTNYYGYPAGTGTTQDNLIETGMA
CD11b/CD18 (212-231)
FimA1 LKIKRVHARMAFTEIKVQMSAAY-DNIYTFVPEK--IYGLIAKKQSNLFGATLVNADANY
FimA2 LEIKRVHARMAFTEIKVQMSQSY-VNKYNFAPEN--IYALVAKKESNLFGASLANSDDAY
FimA3 LEIKRVHARIAFTKIEVTMSQSY-ANKYNFAPEN--IYALVAKKKSNLFGASLANNDDAY
FimA4 LAVTRVHAGISFAGVEVNMATQY-QNYYSFKPADAKIAALVAKKDSKIFGNSLVSNTNAY
FimA5 LKVTRVHAAMSIQNVTVTFDPQYSSNYY-FKPQN--VAGLICKKQSKIFGASLDFGTD-Y
ii (224-228)
iii (247-250)
CD11b/CD18 (252-271)
FimA1 LTGSLTTFNGAYTP----ANYA-NVPWLSRNYVA---PAADAPQGFYVLENDYSANGGTI
FimA2 LTGSLTNFSGAYTP----ANYT-HVDWLGRDYTEIGAATVNTPKGFYVLESTYAQNAG-L
FimA3 LTGSLTTFNGAYTP----ANYT-HVDWLGRDFTE---PSNNAPQGFYVLESTYAQNAG-L
FimA4 LYGVQTP-AGLYTPDAAGETYELEAS-LNTNYAV-----T--GA--GFYVLESHYDASNE-L
FimA5 LGGVATT-AAAYTP----TSYDNNVSWLTKPYAA-----T--KA--GFYIMESVYQVGNN-L
iv (265-273)
v (285-292)
vi (305-309)
Salivary protein (312-231)
FimA1 HPTILCVYGKLQK-NGADLAGADLAAAQAANWMDAEG--KTYYPVLVNFNSNNYTYDSNY
FimA2 RPTILCVKGKLTKHDGAPLSPEEMTAAFNAGWIVADNDPTTYYPVLVNFNSNNYTYDNGY
FimA3 RPTILCVKGKLTKHDGTPLSSEEMTAAFNAGWIVANNDPTTYYPVLVNFESNNYTYTGEA
FimA4 RPTILCIYGKLLDKDGNPLTEPALTDAINAGFGDGDG--TTYYPVLVNYDGNGYIYSGAI
FimA5 RPTILCVYGKLKKTETQDFSQEELDAAVAAGYQDGNA--DTYYPVLVNYNGYGYTYTGEN
statherin (339-372)
vii (359-364)
PRP1 (364-383)
FimA1 T-PKNKIERNHKYDIKLTITGPGTNNPENRITESAHLNVQCTVAEWVLVGQNATW-
FimA2 T-PKNKIERNHKYDIKLTITGPGTNNPENPITESAHLNVKCTVAEWVLVGQNATW-
FimA3 V-EKGKIVRNHKFDINLTITGPGTNNPENPITESANLNVNCVVAAWKGVVQNVIW-
FimA4 TQGQNKIVRNNHYKISLNITGPGTNTPENRQPVQANLNVTCQVTPWVVVNQAATW-
FimA5 T-GLNKILRNNHYKISLTVKGPGTNTPEGRLPEEANLNVNCEVVSWNVVNQSAIWN
b
vii
iii
iv
vi
v
i
c
vii
iii
iv
v
ii
i
d
vii
iii
iv
vi
ii
90°
e
f
PRP1
PRP1
Statherin
Salivary protein
Statherin
CD11b
CD18
CD14

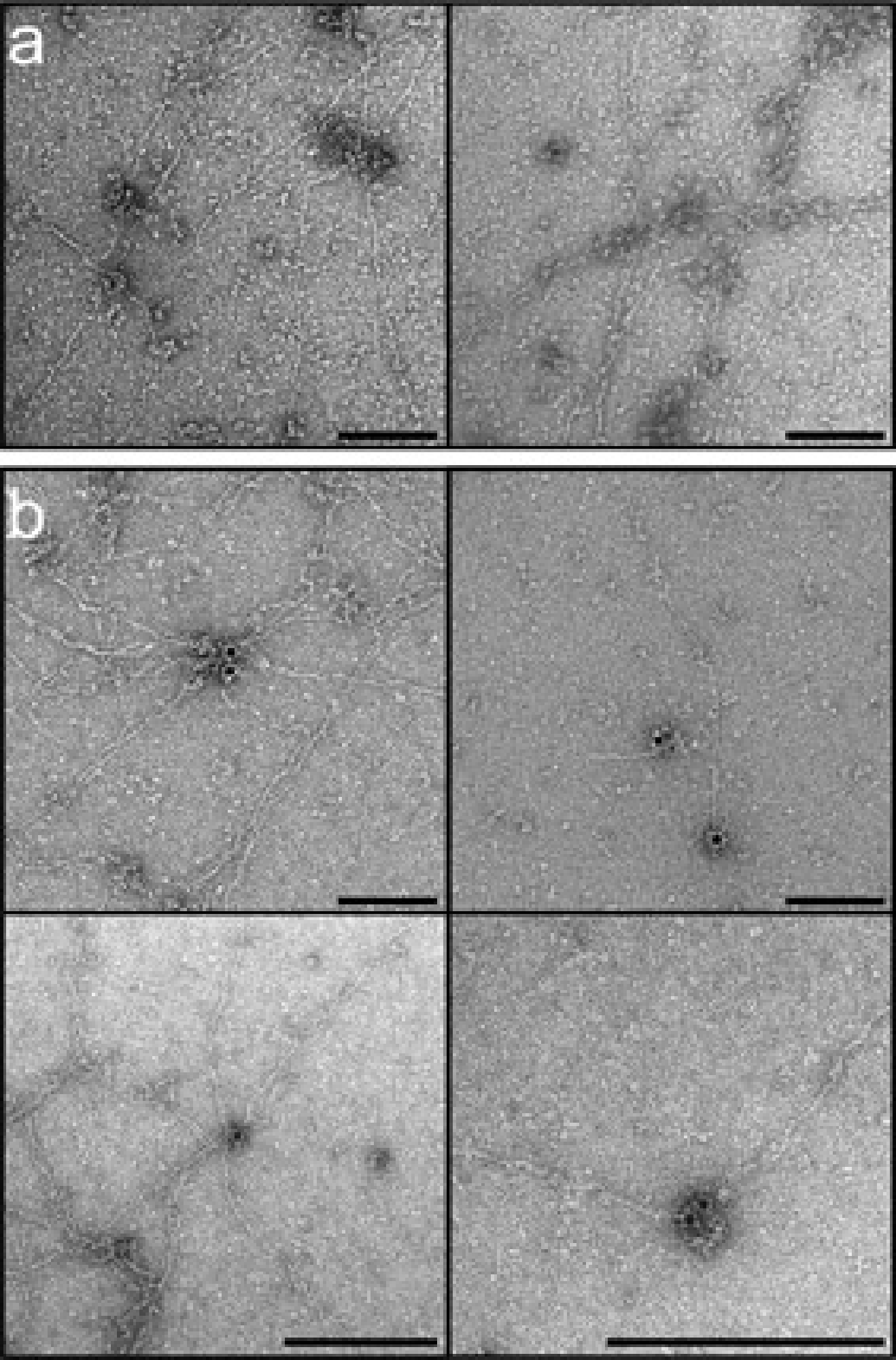

a
i
ii
iii
iv
250
150
100
75
50
37
25
20
kDa
*
**
b
c
d
40
20
Frequency (%)
100
500
1000
1500
Filament length (nm)
e
i
ii
iii
iv
v
vi
vii
250
150
100
75
50
37
25
kDa
*


a


b
1
0.8
0.6
0.4
0.2
0
Fourier Shell Correlation
12
11
10
9
8
7
6
5
4
3
Resolution [Å]
c
φ=90°
φ=135°
φ=45°
θ=60°
θ=30°
φ=180°
φ=0°
φ=225°
φ=315°
φ=270°
11+
0


a
b
c


a
Monomer
Assembled
A1'
A2'
CTD
NTD
A1(anchor)
e
f
Gln70
Ala185
Tyr186
b
i
ii
iii
210–
140–
100–
70–
55–
47–
37–
27–
kDa
c
d


a
b

Extended Data Figure 7

a. X-ray data collection and refinement statistics

	FimA1 (PDB 6JZJ)	FimA2 (PDB 6JZK)
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	35.8, 85.5, 242.6	58.0, 86.0, 102.1
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	48.5-2.1 (2.16-2.10)*	65.8-1.6 (1.63-1.60)
R_{merge}	0.050 (0.397)	0.051 (0.514)
$I / \sigma I$	21.3 (4.4)	13.7 (2.7)
Completeness (%)	100.0 (100.0)	99.9 (99.9)
Redundancy	6.6 (6.9)	4.7 (4.7)
Refinement		
Resolution (Å)	2.1 (2.15)	1.6 (1.6)
No. reflections	44,790 (2,982)	67,987 (2,564)
R_{work} / R_{free}	18.9/24.0 (24.0/27.9)	17.8/19.5 (25.0/25.4)
No. atoms		
Protein	5,231	2,643
Ligand/ion	21	10
Water	381	351
B-factors		
Protein	39.5	20.3
Ligand/ion	51.3	31.2
Water	40.7	29.8
R.m.s. deviations		
Bond lengths (Å)	0.007	0.005
Bond angles (°)	0.818	0.757

*Values in parentheses are for highest-resolution shell.
One crystal was used for each structure.

b. Cryo-EM data collection, model refinement and validation statistics

	FimA1 (EMD-0724) (PDB 6KMF)
Data collection and processing	
Magnification	95,000x
Voltage (kV)	200
Electron exposure ($e^-/Å^2$)	46
Defocus range (μm)	-1.5 to -2.5
Pixel size (Å)	1.12
Symmetry imposed	None (C1)
Initial particle images (no.)	159,300
Final particle images (no.)	61,728
Map resolution (Å)	3.6
FSC threshold	0.143
Map resolution range (Å)	2.5 – 4.5
Model Refinement	
Initial model used (PDB code)	6JZJ
Model resolution (Å)	4.0
FSC threshold	0.5
Model resolution range (Å)	4.0 – 4.2
Map sharpening *B* factor ($Å^2$)	-150
Model composition	
Non-hydrogen atoms	7,704
Protein residues	1,011
B factors ($Å^2$)	
Protein	128.9
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	0.89
Validation	
MolProbity score	1.9
Clashscore	5.3
Poor rotamers (%)	0.62
Ramachandran plot	
Favored (%)	86.6
Allowed (%)	13.4
Disallowed (%)	0