LETTER

doi:10.1038/nature12007

# Structural and energetic basis of folded-protein transport by the FimD usher

Sebastian Geibel[1]*, Erik Procko[2]*, Scott J. Hultgren[3], David Baker[2] & Gabriel Waksman[1]

**Type 1 pili, produced by uropathogenic *Escherichia coli*, are multisubunit fibres crucial in recognition of and adhesion to host tissues[1]. During pilus biogenesis, subunits are recruited to an outer membrane assembly platform, the FimD usher, which catalyses their polymerization and mediates pilus secretion[2]. The recent determination of the crystal structure of an initiation complex provided insight into the initiation step of pilus biogenesis resulting in pore activation, but very little is known about the elongation steps that follow[3]. Here, to address this question, we determine the structure of an elongation complex in which the tip complex assembly composed of FimC, FimF, FimG and FimH passes through FimD. This structure demonstrates the conformational changes required to prevent backsliding of the nascent pilus through the FimD pore and also reveals unexpected properties of the usher pore. We show that the circular binding interface between the pore lumen and the folded substrate participates in transport by defining a low-energy pathway along which the nascent pilus polymer is guided during secretion.**

All pilus subunits exhibit an incomplete immunoglobulin-like fold lacking β-strand G, leaving a large hydrophobic groove across the subunit surface[4–7]. Folding of pilus subunits is catalysed by a periplasmic chaperone[8] (FimC for type 1 pili) (Supplementary Fig. 1a). In a mechanism known as donor-strand complementation, FimC donates the missing β-strand *in trans*, thereby complementing and stabilizing the pilus subunit's fold[4,6] (Supplementary Fig. 1b). The resulting binary chaperone–subunit complexes are next recruited to an outer membrane assembly 'nanomachine', the usher (FimD for type 1 pili), for pilus assembly and secretion. Subunit polymerization occurs through a mechanism known as donor-strand exchange (DSE), whereby the amino-terminal extension (Nte) of the subunit next in assembly replaces the chaperone strand in the groove of the previously assembled (or receiving) subunit[5] (Supplementary Fig. 1b).

Ushers consist of five functional domains: an N-terminal domain (NTD); a pore domain that, in the resting state, is occluded by a plug domain; and two carboxy-terminal domains (CTD1 and CTD2). Recently, the crystal structure of the pilus biogenesis initiation complex FimD–FimC–FimH (Supplementary Fig. 1c) inspired a model for subunit recruitment, polymerization and secretion by the usher[3] (Supplementary Fig. 1d). FimH consists of two domains: a pilin domain ($FimH_P$) similar to all other pilus subunits, and a receptor-binding domain referred to as the lectin domain[6] ($FimH_L$). In the FimD–FimC–FimH structure, $FimH_L$ is located inside the FimD pore. $FimH_P$ is still in donor-strand complementation interaction with the chaperone FimC bound to CTD1 and CTD2. The usher NTD, stabilized by the plug domain, is free and serves as a recruitment site for the next chaperone–subunit complex to be assembled (FimC–FimG; Supplementary Fig. 1d). Binding to the NTD places the FimG Nte peptide in an ideal position for DSE to occur[3], leading to the dissociation of FimC bound to $FimH_P$. FimC has no binding affinity for the usher CTDs on its own, and so dissociates from the complex. At this point, the entire nascent pilus translocates up inside the pore and the chaperone–subunit complex bound to the NTD transfers to the CTDs in a handover step (Supplementary Fig. 1d). The usher NTD is now free to enter another cycle of subunit incorporation by recruiting the next chaperone–subunit complex for assembly (FimC–FimF). This model for the subunit-incorporation cycle remains, however, to be structurally documented. Here we describe the structure of the elongation complex FimD–FimC–FimF–FimG–FimH, a post-initiation state during pilus biogenesis.

The complex FimD–FimC–FimF–FimG–FimH was purified to homogeneity and crystallized, and its structure was determined to a resolution of 3.8 Å (Supplementary Tables 1 and 2 and Supplementary Figs 2–4). The crystal structure of FimD–FimC–FimF–FimG–FimH captures the secretion of the type 1 pilus tip consisting of pilus subunits FimF, FimG and FimH through the usher pore across the outer bacterial membrane (Fig. 1a). The adhesin FimH is incorporated at the distal end of the pilus tip and found on the extracellular side of the transporter. FimG occupies the lumen of the usher β-barrel and contacts the usher CTD1, plug and β-barrel domains. As the last-assembled pilus subunit, FimF is found still in donor-strand complementation interaction with its chaperone FimC, which is bound by the usher C-terminal domains CTD1 and CTD2 in the periplasm. Thus, the structure provides the remarkable view of a transporter entirely traversed by its protein polymer substrate, with one subunit having entirely emerged from the transporter and another still remaining on the cellular side of the transporter. The type 1 pilus tip subunits FimH, FimG and FimF are linked non-covalently by DSE, where the N-terminal extension of FimF is inserted into FimG and that of FimG is inserted into $FimH_P$. The usher NTD shares only an interface with the plug domain and is otherwise exposed. Modelling shows that the NTD is indeed available to bind the next incoming chaperone–subunit complex, FimC–FimA, bringing it into close proximity to FimF and in the appropriate orientation to undergo DSE (Fig. 1b). This is consistent with the DSE model for pilus subunit incorporation introduced in ref. 3.

When the FimH proteins in the initiation complex FimD–FimC–FimH and the elongation complex FimD–FimC–FimF–FimG–FimH are superimposed, it becomes apparent that FimH undergoes a large conformational change when exiting the usher pore, with the angle between its two domains ($FimH_P$ and $FimH_L$) decreasing[9] (Fig. 2a and Supplementary Fig. 5). We speculate that such a conformational change may provide the necessary energy to favour translocation of FimH outside the pore and possibly prevent motion back into the periplasm. Another remarkable structural rearrangement affecting FimH during transport is the compression and decompression of the $FimH_L$ domain structure before and after transport, respectively (Fig. 2b).

Using the physically realistic ROSETTA energy function[10,11], we next investigated the interactions between the usher pore and its inserted substrate, $FimH_L$ in the initiation complex FimD–FimC–FimH, or FimG in the elongation complex FimD–FimC–FimF–FimG–FimH.

[1]Institute of Structural and Molecular Biology, University College London and Birkbeck College, Malet Street, London WC1E 7HX, UK. [2]The Howard Hughes Medical Institute and Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA. [3]Center for Women's Infectious Disease Research and Department of Molecular Microbiology, Washington University School of Medicine, St Louis, Missouri 63011, USA.
*These authors contributed equally to this work.

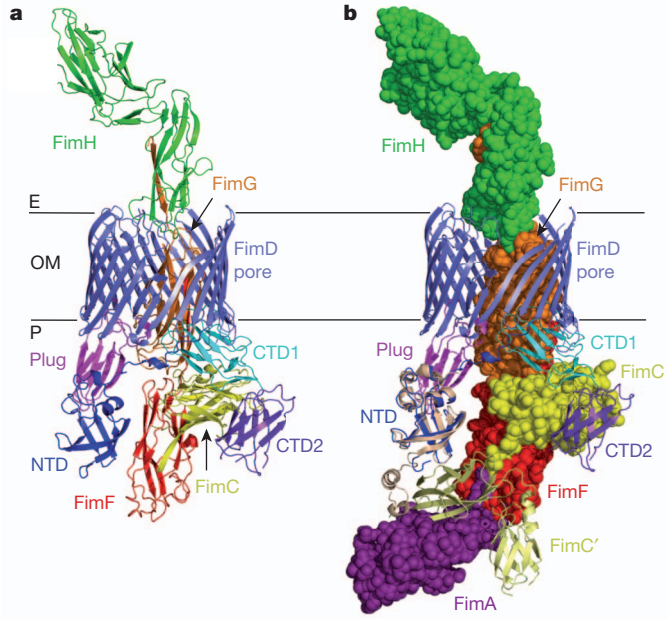


**Figure 1 | Structure of FimD–FimC–FimF–FimG–FimH.** **a**, Crystal structure of FimD–FimC–FimF–FimG–FimH. All proteins are in ribbon representation: FimD NTD, blue; β-barrel, slate; plug domain, magenta; CTD1, cyan; CTD2, purple; FimC, yellow; FimF, red; FimG, orange; FimH, green. E, extracellular; OM, outer membrane; P, periplasmic. **b**, Recruitment of the next chaperone–subunit complex FimC–FimA in assembly by FimD–FimC–FimF–FimG–FimH. FimD is in ribbon representation, coloured as in **a**. Chaperone FimC (yellow), pilus subunits FimF (red) and FimG (orange), and adhesin FimH (green) are in sphere representation. The recruited chaperone–subunit complex FimC′–FimA (PDB ID, 4DWH; FimC′, pale yellow ribbon; FimA, purple spheres) is modelled at the NTD on the basis of the crystal structure of the isolated FimD NTD domain bound to FimC–FimF (PDB ID, 3BWU; see Supplementary Methods for details).

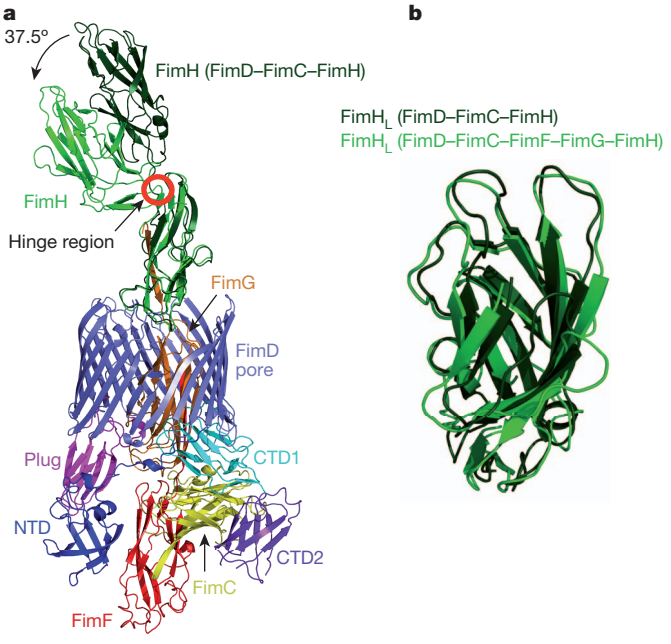


**Figure 2 | Comparison of the structures of FimH before and after translocation.** **a**, Superposition of the structure of FimH from the initiation complex FimD–FimC–FimH (in dark green) with that of FimH from the elongation complex FimD–FimC–FimF–FimG–FimH (same colour coding as in Fig. 1a). FimH$_P$ was used for the superposition. **b**, Superposition of the structure of FimH$_L$ from the initiation complex (dark green) with that of the same domain from the elongation complex (light green).

In these calculations, the protein assemblies were represented as essentially rigid components. Subunits were moved along a central pore axis (Supplementary Fig. 6a–c) with only minor deviations that can be sampled by rigid-body minimization. This is a reasonable assumption owing to the close steric fit of subunits within the pore that would otherwise clash.

Translating FimG and FimH$_L$ subunits within the pore laterally along a fine grid and plotting the calculated energies reveals a steep energy well, with forces on all sides returning the perturbed subunits to their central, ground-state positions (Methods, Fig. 3a, b and Supplementary Fig. 6d). However, when translated parallel to the pore axis to mimic subunit entry or exit, the calculated energies with FimG in the barrel increase slowly (even for perturbations with a root mean squared deviation of 6 Å), whereas with FimH$_L$ inside the pore there was a steep energy increase (Fig. 3c and Supplementary Fig. 6e). Thus, the subunit or domain inside the pore is more tightly held in the initiation complex than it is in the elongation complex.

In apo-FimD, the plug domain is found inside the pore[12,13] and must be expelled and replaced by FimH$_L$ in an activation process that might be triggered by engagement of FimH with FimD. When the plug domain (beginning with the structure of apo-FimD (Protein Data Bank (PDB) ID, 3OHN)) was translated towards the periplasm to mimic its extrusion from the pore, the energy increased slowly as it did for FimG (Fig. 3c). Hence, whereas the plug domain and FimG may readily exit the usher pore along the pore axis, FimH$_L$ forms tight interactions. We propose that this may be essential for activation, in which FimH$_L$ displaces the plug domain from inside the pore.

By sweeping a 60°-sector window emanating from the pore axis around the usher barrel and computing the binding energy within this

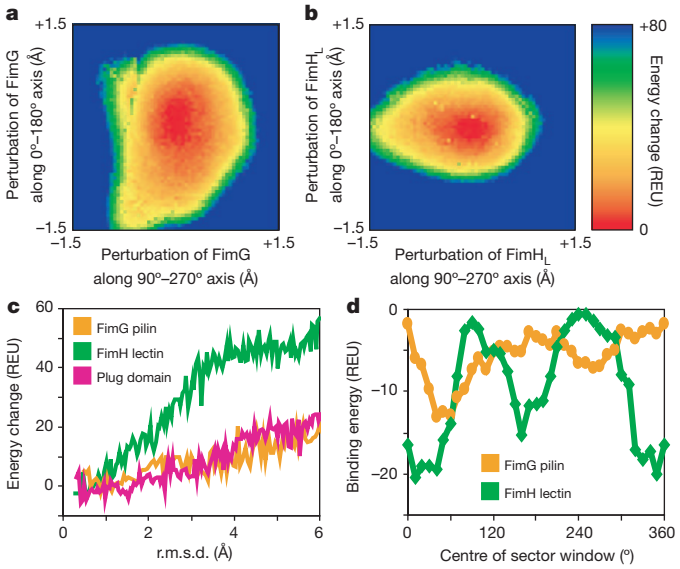


**Figure 3 | Steep energy funnels and opposing binding surfaces position the translocating substrate at the centre of the pore.** **a**, The heat map corresponding to the view down the pore axis, with the energy increase plotted as FimG is laterally translated (that is, perpendicular to the pore axis) along a finely spaced grid within the usher lumen. FimG located at any position on the grid will experience a force dependent on the slope of the potential energy well, returning FimG to its central ground state. REU, ROSETTA energy units. **b**, As in **a**, with FimH$_L$ now inside the usher lumen. **c**, Subunits or domains occupying the pore were randomly rotated by up to 8° about their geometric centres and translated parallel to the pore axis by up to 6 Å (for the plug domain, displacement was only in the direction of the periplasm to mimic plug extrusion during activation). Calculated energies are plotted against the root mean squared deviation (r.m.s.d.) from the respective minimized crystal structures. Eight thousand perturbations were made for each subunit, with the outline of the lowest-energy conformations shown. **d**, Plot of the respective binding energies as a 60°-sector window emanating from the pore axis is rotated around the FimG pilin structure or the FimH lectin structure.

window (Fig. 3d and Supplementary Figs 7 and 9), we observe two binding-energy troughs and two peaks that are roughly 180° apart for both $FimH_L$ inside the FimD pore of the FimD–FimC–FimH structure and FimG inside the FimD pore of the FimD–FimC–FimF–FimG–FimH structure. Regions of high binding energy correlate with closer atomic contacts as measured by solvent accessibility (Supplementary Fig. 7b), and regions that tightly bind FimG are distinct from those that tightly bind $FimH_L$, indicating specialized binding sites. Usher pores seem to have evolved binding sites that face each other on opposite sides of the pore, resulting in the placement of the translocating substrate at the very centre of the pore. Because pilin subunits share the same immunoglobulin-like fold and have high sequence similarity (compared with FimG, the pilin domains of FimH, FimF and FimA share 19–25% sequence identity and 62–67% similarity), it is likely that all pilin subunits occupy the same central position in the usher pore, and our results for FimG are generalizable to the rest of these proteins.

It has been previously proposed that incoming chaperone–pilin subunits are recruited to the usher NTD and then transfer after DSE to the usher CTDs during assembly (see above). This requires a rotation of the NTD-bound chaperone–subunit complex by about 100°–120° after DSE has occurred (Supplementary Fig. 8a). The torque produced by the transfer of the incoming subunit from the NTD to the CTDs might be accommodated through rotations around linker residues between pilus building blocks. Alternatively, rigid-body rotation of the entire tip assembly might necessitate a rotation of the translocating subunit within the pore.

To investigate such a possibility, we calculated an energy landscape as the FimG pilin was rotated in 2° steps and translated in steps of 1 Å up or down the usher pore. The calculated landscape reveals a continuous low-energy path for FimG as it enters and exits the FimD pore (visualized by following the 'low-energy' colours in Fig. 4a and Supplementary Fig. 8) that requires an anticlockwise twist. To predict a possible trajectory through the energy landscape, we introduced a torsional spring term to limit the extent to which FimG can rotate at each translation step (this ensures connectivity in the low-energy path); the entry–exit trajectory for FimG pilin is then derived from the lowest-energy states (Fig. 4b). By repeating these calculations using input structures minimized with and without constraints from the crystallographic electron density, different spring constants and structures refined to a resolution of either 3.8 or 4.1 Å, a twist of 1.6°–3.0° per ångström translation was observed in the trajectories (Supplementary Fig. 8). The presence of a low-energy exit path requiring pilin rotation was further validated with an alternatively defined pore axis based on the membrane-like packing of FimD within the crystal lattice (Supplementary Fig. 9). Visual inspection of the FimD–FimC–FimF–FimG–FimH crystal structure revealed that the three assembled pilin subunits respectively representing entrance into (FimF), residence within (FimG) and exit from ($FimH_P$) the usher pore are each separated by a translation by 53 Å and an anticlockwise rotation by 110° or 120° (Supplementary Fig. 8a). Hence, the twist undergone by FimG pilin as it enters or exits the pore and that is captured in the computed energy landscape matches the magnitude and direction of the rotations observed and anticipated from static crystallographic evidence. These results are probably generalizable to other pilin subunits owing to high sequence and structural similarity.

Similar analysis was applied to $FimH_L$ within the FimD lumen, using the FimD–FimC–FimH crystal structure, but no single, clear low-energy entry–exit path was observed (Supplementary Fig. 10). This is not surprising. Indeed, $FimH_L$ differs markedly from all other pilus subunits or domains: whereas $FimH_P$, FimG, FimF and FimA have a classical pilin fold that consists of six β-strands complemented by the donor strand of the next subunit in assembly, $FimH_L$ has a β-barrel 'jelly roll' fold. Furthermore, as shown here, the FimD pore forms tighter interactions with $FimH_L$ than with a representative pilin subunit. Thus, whereas all other pilus subunits and domains are easily extracted from and guided through the pore, unbinding and translocation of $FimH_L$ will require more energy, probably through conformational changes yet to be characterized.

The crystal structure presented here provides unprecedented structural evidence supporting the model first proposed in ref. 3 for the usher-mediated subunit-incorporation cycle during pilus biogenesis. It also demonstrates that usher pores form finely tuned, circular protein–protein interfaces specifically geared to maximize transport, by forming diametrically opposed binding sites that position substrates at the very centre of the pore, and to facilitate specific steps during transport, by imposing rotational and translational constraints through defined energy paths. We anticipate that all transporters will demonstrate similar levels of sophistication by having evolved function-specific features facilitating defined steps during substrate translocation.

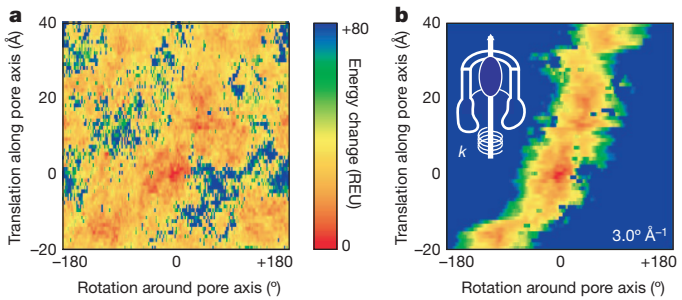


**Figure 4 | A low-energy pathway through the pore lumen facilitates translocation of subunits and their transfer from NTD to CTDs. a**, Starting from the native crystal structure with FimG and its complementing strand from FimF inside the FimD pore (subunits FimH, FimF and FimC were not considered), FimG was translated, in steps of 1 Å, up to 40 Å out of the pore and up to 20 Å back towards the periplasm, along the pore axis. At each translational step, FimG was rotated around the pore axis in 2° steps. Each sampled FimG rotation–translation conformation was minimized after rotamer repacking to resolve small clashes, and the calculated energy for each perturbation is plotted. Lower, more favourable, energies are hotter colours. **b**, The energy landscape calculated in **a** after modification by addition of a torsional spring potential ($E = k(\Delta\theta)^2/2$; represented by the coil with spring constant $k$ twisted around the pore axis in the sketch at top left) to derive a connected trajectory, with $\Delta\theta$ the angle from the lowest-energy FimG conformation before each step of 1 Å, and starting with FimG positioned at 0 Å, 0°.

## METHODS SUMMARY

**Expression and purification of FimD–FimC–FimF–FimG–FimH.** *Escherichia coli* Tuner cells (Novagen) were transformed with plasmids pNH237 encoding $fimC_{His}FGH$ under arabinose control and pAN2 encoding $fimD_{Strep}$ under isopropylthiogalactoside control. The complex FimD–FimC–FimF–FimG–FimH was then expressed and purified as described in ref. 3.

**Crystallization of FimD–FimC–FimF–FimG–FimH.** Crystals of the complex were grown in hanging drops at 20 °C using the vapour diffusion method, where 1 μl of the complex (8 mg ml$^{-1}$) was mixed with an equal volume of 1.2–2.0 M sodium formate.

**Structure determination of FimD–FimC–FimF–FimG–FimH.** Data were collected at the Diamond Light Source (beamline IO2) at 100 K and were indexed, integrated and scaled to a resolution of 3.8 Å using the XDS software package[14]. A high-resolution cut-off for the data was chosen following ref. 15. The space group was determined using POINTLESS[16]. The space group, cell dimensions and data collection statistics are reported in Supplementary Table 1. The phase problem was solved by molecular replacement using PHASER and the structures of FimD (PDB ID, 3RFZ), FimC–FimF (3BWU) and FimG–FimH (3JWN) as search models[17]. The molecular replacement model for FimD–FimC–FimF–FimG–FimH obtained by PHASER was refined by alternating rounds of model building with COOT[18] using B-factor-sharpened maps generated with PHENIX[19] and refinement cycles using CCP4 REFMAC[20], CNS[21] and PHENIX[19]. An initial rigid-body refinement of the model was carried out using REFMAC. For the next round of refinement, we followed the CNS protocol for low-resolution refinement, which includes torsion angle annealing, grouped B factor and deformable elastic network refinement[21,22]. Deformable elastic network parameters $\gamma$ and $\omega$ were set to 0.2 and 100, respectively.[23]. Final refinement was carried out with PHENIX. The atomic displacement parameters and stereochemistry weights were optimized during the refinement and resulted in tight restraints. Finally, $R_{free}$ converged to 29.8% (Supplementary

Tables 1 and 2 and Supplementary Figs 3 and 4). The same $R_{free}$ flag was maintained for cross-validation throughout the refinement process.

**Computational methods.** Before binding-energy calculations were performed, crystal structure coordinates were relaxed in the ROSETTA force field using FASTRELAX with electron density constraints[11,24] (Methods and Supplementary Figs 11 and 12). FimD sectors were generated on the basis of the angle a residue's Cα atom makes with FimD–D208 (set to 0°) when projected on a plane perpendicular to the pore axis. Only FimD and the subunit within the pore were considered (FimG pilin (amino acids 12–144)/FimF Nte (1–12) or $FimH_L$ (1–157)), and calculations of binding energy and buried solvent-accessible surface area were made using ROSETTASCRIPTS[25] (Supplementary Fig. 13). Translational and rotational perturbations of the translocating subunit were made with the MMTSB suite[26] followed by rotamer repacking and side-chain and backbone gradient minimization in ROSETTA (Supplementary Figs 14–16).

**Full Methods** and any associated references are available in the online version of the paper.

**Received 2 October 2012; accepted 14 February 2013.**


1. Waksman, G. & Hultgren, S. J. Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nature Rev. Microbiol.* **7,** 765–774 (2009).
2. Nishiyama, M., Ishikawa, T., Rechsteiner, H. & Glockshuber, R. Reconstitution of pilus assembly reveals a bacterial outer membrane catalyst. *Science* **320,** 376–379 (2008).
3. Phan, G. *et al.* Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate. *Nature* **474,** 49–53 (2011).
4. Sauer, F. G. *et al.* Structural basis of chaperone function and pilus biogenesis. *Science* **285,** 1058–1061 (1999).
5. Sauer, F. G., Pinkner, J. S., Waksman, G. & Hultgren, S. J. Chaperone priming of pilus subunits facilitates a topological transition that drives fiber formation. *Cell* **111,** 543–551 (2002).
6. Choudhury, D. *et al.* X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic *Escherichia coli*. *Science* **285,** 1061–1066 (1999).
7. Zavialov, A. V. *et al.* Structure and biogenesis of the capsular F1 antigen from *Yersinia pestis*: preserved folding energy drives fiber formation. *Cell* **113,** 587–596 (2003).
8. Vetsch, M. *et al.* Pilus chaperones represent a new type of protein-folding catalyst. *Nature* **431,** 329–333 (2004).
9. Le Trong, I. *et al.* Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like β-sheet twisting. *Cell* **141,** 645–655 (2010).
10. Das, R. & Baker, D. Macromolecular modeling with Rosetta. *Annu. Rev. Biochem.* **77,** 363–382 (2008).
11. Leaver-Fay, A. *et al.* Rosetta3: an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol.* **487,** 545–574 (2011).
12. Huang, Y., Smith, B. S., Chen, L. X., Baxter, R. H. & Deisenhofer, J. Insights into pilus assembly and secretion from the structure and functional characterization of usher PapC. *Proc. Natl Acad. Sci. USA* **106,** 7403–7407 (2009).
13. Remaut, H. *et al.* Fiber formation across the bacterial outer membrane by the chaperone/usher pathway. *Cell* **133,** 640–652 (2008).
14. Kabsch, W. XDS. *Acta Crystallogr. D* **66,** 125–132 (2010).
15. Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336,** 1030–1033 (2012).
16. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D* **62,** 72–82 (2006).
17. McCoy, A. J. *et al.* PHASER crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).
18. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of COOT. *Acta Crystallogr. D* **66,** 486–501 (2010).
19. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66,** 213–221 (2010).
20. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **67,** 355–367 (2011).
21. Schröder, G. F., Levitt, M. & Brunger, A. T. Super-resolution biomolecular crystallography with low-resolution data. *Nature* **464,** 1218–1222 (2010).
22. Brunger, A. T. Version 1.2 of the Crystallography and NMR system. *Nature Protocols* **2,** 2728–2733 (2007).
23. O'Donovan, D. J. *et al.* A grid-enabled web service for low-resolution crystal structure refinement. *Acta Crystallogr. D* **68,** 261–267 (2012).
24. DiMaio, F., Tyka, M. D., Baker, M. L., Chiu, W. & Baker, D. Refinement of protein structures into low-resolution density maps using Rosetta. *J. Mol. Biol.* **392,** 181–190 (2009).
25. Fleishman, S. J. *et al.* RosettaScripts: a scripting language interface to the Rosetta macromolecular modeling suite. *PLoS ONE* **6,** e20161 (2011).
26. Feig, M., Karanicolas, J. & Brooks, C. L. III. MMTSB tool set: enhanced sampling and multiscale modeling methods for applications in structural biology. *J. Mol. Graph. Model.* **22,** 377–395 (2004).


**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** This work was funded by Medical Research Council grant 85602 to G.W. D.B. and E.P. are supported by grant P41 GM103533 from the National Institute of General Medical Studies at the US National Institutes of Health (NIH). S.J.H. was supported by grant AI029549 from the National Institute of Allergy and Infectious Disease at the NIH. We thank the staff of beamline ID23-1 at the European Synchrotron Radiation Facility, the staff of beamline IO2 at the Diamond Light source and A. Cole for technical assistance during data collection.

**Author Contributions** S.G. and E.P. carried out the crystallographic and computational work, respectively. D.B. and G.W. supervised the work. S.G., E.P., D.B. and G.W. analysed the data. S.G., E.P., S.J.H., D.B. and G.W. wrote the paper.

**Author Information** The atomic coordinates and structure factors of FimD–FimC–FimF–FimG–FimH have been deposited in the Protein Data Bank under accession ID 4J3O. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.W. (g.waksman@ucl.ac.uk) or D.B. (dabaker@u.washington.edu).

## METHODS

**Expression and purification of FimD–FimC–FimF–FimG–FimH.** *Escherichia coli* Tuner (Novagen) was transformed with plasmids pNH237 encoding $fimC_{His}FGH$ under arabinose control and pAN2 encoding $fimD_{Strep}$ under isopropylthiogalactoside (IPTG) control[3,13]. Bacteria were grown in TB media containing 35 mg ml$^{-1}$ kanamycin and 25 mg ml$^{-1}$ chloramphenicol at 37 °C. Protein overexpression was induced by addition of 100 μM IPTG and 0.1% (w/v) L-arabinose supplemented with 0.1% (v/v) glycerol at $D_{600\,nm}$ 1.0 for 48 h at 16 °C. By not providing the FimA subunit to the assembly system, pilus production stalls after incorporation of the last tip subunit, FimF, resulting in production of homogeneous FimD–tip complexes.

**Protein purification of FimD–FimC–FimF–FimG–FimH.** FimD–FimC–FimF–FimG–FimH was purified as described in ref. 3 for FimD–FimC–FimH.

**Crystallization of FimD–FimC–FimF–FimG–FimH.** Plate-like crystals of FimD–FimC–FimF–FimG–FimH were grown in hanging drops at 20 °C using the vapour diffusion method, where 1 μl of the complex (8 mg ml$^{-1}$) was mixed with an equal volume of 1.2–2.0 M sodium formate and equilibrated against 600 μl reservoir solution. Crystals were transferred into mother liquor with 25% (v/v) glycerol as cryoprotectant and then flash-frozen in liquid nitrogen.

**Structure determination of FimD–FimC–FimF–FimG–FimH.** Data were collected at the Diamond Light Source (beamline IO2) at 100 K and were indexed, integrated and scaled to a resolution of 3.8 Å using the XDS software package[14]. A high-resolution cut-off for the data was chosen following ref. 15 (Supplementary Table 1). The space group was determined using POINTLESS[16]. The space group, cell dimensions and data collection statistics are reported in Supplementary Table 1.

The phase problem was solved by molecular replacement using PHASER and the structures of FimD (PDB ID, 3RFZ), FimC–FimF (3BWU) and FimG–FimH (3JWN) as search models[17]. The molecular replacement model for FimD–FimC–FimF–FimG–FimH obtained by PHASER was refined by alternating rounds of model building with COOT[18] using B-factor-sharpened maps generated with PHENIX[19] and refinement cycles using CCP4 REFMAC[20], CNS[21] and PHENIX[19]. Initial rigid-body refinement (rigid bodies were assigned for subunits FimC, FimF, FimG and FimH and for the β-barrel, the plug and the N-terminal and two C-terminal domains of FimD) was carried out using REFMAC and resulted in an $R_{free}$ value of 45%. For the next round of refinement, we followed the CNS protocol for low-resolution refinement, which includes torsion angle annealing, grouped B factor and deformable elastic network (DEN) refinement[21,22]. Two B factors were assigned to each amino acid, one to the main-chain atoms and one to the side-chain atoms. FimC (PDB ID, 3BWU), FimD (3RFZ), FimF (3BWU), FimG (3JWN) and FimH (3JWN) were used as DEN reference files. After a two-dimensional grid search for the DEN parameters using the SBGrid Science Portal web service, DEN parameters $\gamma$ and $\omega$ were set to 0.2 and 100, respectively[23]. DEN parameter $\gamma$ balances the influences of the diffraction data and the reference model, whereas $\omega$ is the weighting factor for the DEN potentials. The starting annealing temperature was also optimized using the SBGrid Science Portal DEN web service, and was set to 1,500 K. In the final rounds of DEN refinement, the starting and reference models were the same. DEN refinement changed the positioning of the N- and C-terminal domains, which seem to be quite flexible, as well as the shape of the usher β-barrel, and therefore improved $R_{free}$ considerably. DEN refinement converged to $R_{free} = 31\%$. The final refinement rounds were carried out with PHENIX. The refinement protocol included three rounds of Cartesian and individual B-factor refinement. The atomic displacement parameters and stereochemistry weights were optimized during the refinement and resulted in tight restraints. Finally, $R_{free}$ converged to 29.8%. The same $R_{free}$ flag was maintained for cross-validation throughout the refinement process. Note that high-resolution data to a resolution of 3.8 Å were included because the resulting final model had a better $R_{free}$ value, improved stereochemistry and improved electron density (Supplementary Tables 1 and 2 and Supplementary Figs 3 and 4).

**Computational methods.** For analysis of static structures, the refined crystal structures were relaxed in the ROSETTA force field using two rounds of FASTRELAX and Cartesian space minimization (F. DiMaio, manuscript in preparation) with electron density constraints[24]. B-factor-sharpened $2mF_o - DF_c$ density maps were generated in PHENIX[19] using deposited structure factors for FimD–FimC–FimH or structure factors to a resolution of 4.1 Å for FimD–FimC–FimF–FimG–FimH. ROSETTA command line and score term weights are detailed in Supplementary Fig. 11. The electron-density-constraining score term accounts for ~40% and ~55% of the total score for FimD–FimC–FimF–FimG–FimH and FimD–FimC–FimH, respectively. Relaxed decoys from 50 relaxation runs were well converged both in total score (scores ranged from −5,964 to −5,911 for FimD–FimC–FimF–FimG–FimH and from −6,376 to −6,326 for FimD–FimC–FimH) and structure (root mean squared deviation from an arbitrary reference decoy was ≤0.22 Å for FimD–FimC–FimF–FimG–FimH and ≤0.14 Å for FimD–FimC–FimH). The lowest-scoring decoys (Supplementary Fig. 12 shows fits of relaxed structures to electron density) were chosen for final analysis in Fig. 3 and Supplementary Figs 7, 8g, h and 9e–h. The ROSETTA correlation coefficient for measuring agreement between coordinates and a density map[24] decreased by 3% for FimD–FimC–FimF–FimG–FimH and by 2% for FimD–FimC–FimH following relaxation, indicating a small trade-off between minimizing structural features within the ROSETTA force field and maintaining a close fit to the experimental electron density map.

To generate FimD sectors, residues were included if the Cα atom's projection on a plane perpendicular to the pore axis fell within the sector's desired angle range, with the Cα projection of FimD–D208 set to 0°. This creates a PDB file containing FimG pilin (amino acids 12–144)/FimF Nte (1–12) or FimH lectin (1–157) bound to only a subset of FimD residues. Binding energy and buried solvent-accessible surface area were calculated using ROSETTASCRIPTS (Supplementary Fig. 13) without rotamer repacking in either the bound or unbound state[25] (rotamer repacking at the edges of a FimD sector would be meaningless). We confirmed that the combined binding energy from any set of sectors adding up to a full 360° sweep consistently approximates the binding energy calculated for the complete FimD protein bound to its transported substrate.

Translational and rotational perturbations were made using the convpdb.pl application in the MMTSB suite[26]. Perturbed structures were minimized in ROSETTA using full-atom score12 with rounds of rotamer repacking, and side-chain and backbone gradient minimization[11,25] (Supplementary Figs 14–16). In these cases, where large dynamic subunit motions were explored (Figs 3c and 4 and Supplementary Figs 6, 8b–f, 9i and 10), starting structures were also first pre-minimized in full-atom score12, ensuring that any measurements of root mean squared deviation from the initial ground state accurately reflected the applied perturbations, not simply differences from minimizing in different score functions.

For determining potential FimG trajectories through the FimD pore, noise in the energy landscape was dampened by averaging the energy at each sampled FimG rotation–translation with the energies calculated for FimG rotations within 4°. Starting with FimG at 0° and 0 Å, a torsional spring potential ($E = k(\Delta\theta)^2/2$) was added to the energy calculated using ROSETTA. The angle $\Delta\theta$ was measured relative to the lowest-energy conformation in the previous translational step.

Structure figures were generated using The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC (http://www.pymol.org/).