



Structure, Folding and Stability of FimA, the Main Structural Subunit of Type 1 Pili from Uropathogenic *Escherichia coli* Strains

Chasper Puorger, Michael Vetsch, Gerhard Wider*
and Rudi Glockshuber*

ETH Zürich, Institute of Molecular Biology and Biophysics, 8093 Zurich, Switzerland

Received 4 May 2011;
received in revised form
16 July 2011;
accepted 20 July 2011
Available online
27 July 2011

Edited by F. Schmid

Keywords:

type 1 pilus assembly;
protein folding;
FimA;
donor strand
complementation;
alternative folding
possibilities

Filamentous type 1 pili are responsible for attachment of uropathogenic *Escherichia coli* strains to host cells. They consist of a linear tip fibrillum and a helical rod formed by up to 3000 copies of the main structural pilus subunit FimA. The subunits in the pilus interact via donor strand complementation, where the incomplete, immunoglobulin-like fold of each subunit is complemented by an N-terminal donor strand of the subsequent subunit. Here, we show that folding of FimA occurs at an extremely slow rate (half-life: 1.6 h) and is catalyzed more than 400-fold by the pilus chaperone FimC. Moreover, FimA is capable of intramolecular self-complementation via its own donor strand, as evidenced by the loss of folding competence upon donor strand deletion. Folded FimA is an assembly-incompetent monomer of low thermodynamic stability ($-10.1 \text{ kJ mol}^{-1}$) that can be rescued for pilus assembly at 37 °C because FimC selectively pulls the fraction of unfolded FimA molecules from the FimA folding equilibrium and allows FimA refolding on its surface. Elongation of FimA at the C-terminus by its own donor strand generated a self-complemented variant (FimAa) with alternative folding possibilities that spontaneously adopts the more stable conformation ($-85.0 \text{ kJ mol}^{-1}$) in which the C-terminal donor strand is inserted in the opposite orientation relative to that in FimA. The solved NMR structure of FimAa revealed extensive β -sheet hydrogen bonding between the FimA pilin domain and the C-terminal donor strand and provides the basis for reconstruction of an atomic model of the pilus rod.

© 2011 Published by Elsevier Ltd.

Introduction

Attachment of Gram-negative pathogens to the host is the initial step in many bacterial infections and is mediated by filamentous protein complexes, termed pili, displayed on the surface of the pathogens. Among the large variety of adhesive pili that are assembled via the chaperone–usher pathway,^{1,2} the type 1 pilus system from uropathogenic *Escherichia coli* is one of the best characterized examples. Type 1 pili are long protein filaments anchored to the assembly platform FimD in the outer *E. coli* membrane and contain four different structural subunits (Fig. 1a). The main structural

*Corresponding authors. E-mail addresses:

gsw@mol.biol.ethz.ch; rudi@mol.biol.ethz.ch.

Present address: M. Vetsch, Novartis Pharma AG, Biotechnology Development, 4002 Basel, Switzerland.

Abbreviations used: DSE, donor strand exchange; GdmCl, guanidinium chloride; NOE, nuclear Overhauser enhancement; 2D, two-dimensional; HSQC, heteronuclear single quantum coherence; 3D, three-dimensional; TOCSY, total correlated spectroscopy; NOESY, nuclear Overhauser enhancement spectroscopy.

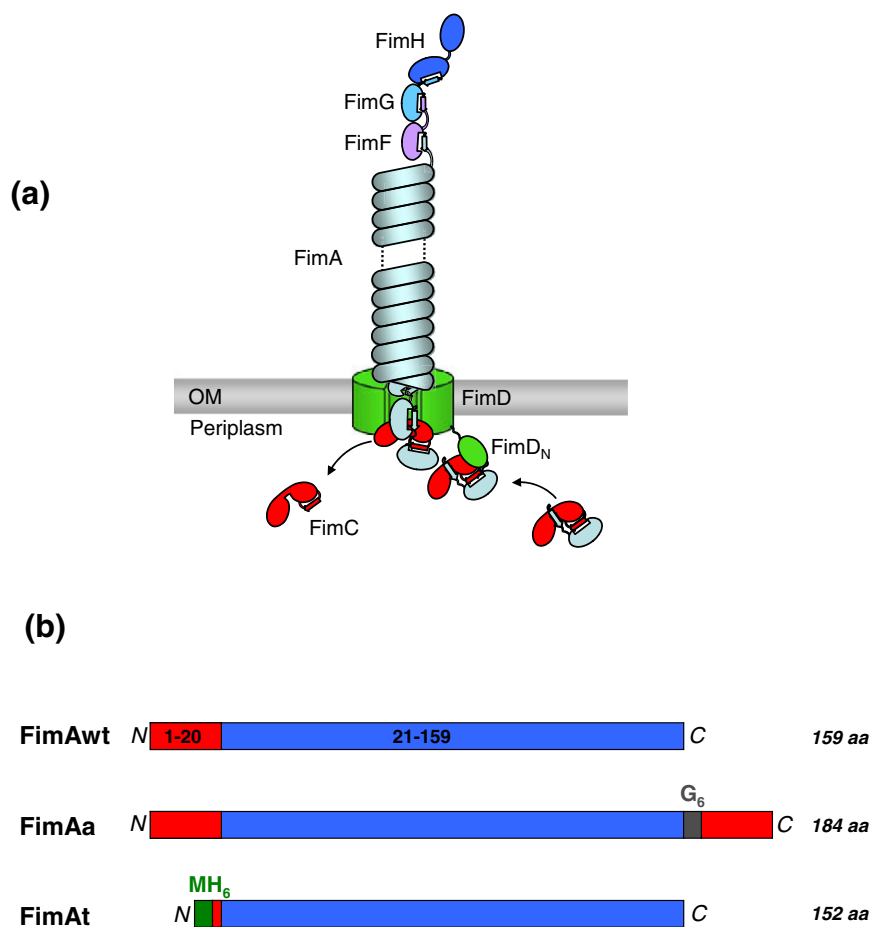


Fig. 1. Model of FimC- and FimD-catalyzed type 1 pilus assembly (a) and FimA constructs used in this study (b). (a) Model of type 1 pilus biogenesis according to the chaperone–usher pathway. The pilus is composed of a linear tip fibrillum formed by the subunits FimF, FimG and the adhesin FimH and the helical pilus rod formed by 500–3000 copies of the main structural pilus subunit FimA. The individual pilus subunits are secreted to the periplasm where the chaperone FimC catalyzes subunit folding and forms stoichiometric complexes with all subunits. The usher FimD specifically recognizes incoming FimC–subunit complexes via its periplasmic, amino-terminal domain, termed FimD_N. Subunit assembly takes place through a DSE mechanism, in which the N-terminal donor strand of the incoming FimC-bound subunit replaces the single copy of FimC capping the last subunit at the growing end of the pilus (see the text for details and references). (b) FimA constructs used in this study. (Top) Wild-type FimA (FimAwt). The N-terminal donor strand (residues 1–20) is indicated in red, and the FimA pilin domain, in blue. (Middle) FimAa: wild-type FimA extended at the C-terminus by a hexaglycine linker (gray) and its own donor strand segment (residues 1–20). (Bottom) FimAt, N-terminally truncated FimA. Residues 1–16 of the N-terminal donor strand were replaced by a hexahistidine affinity tag for purification via metal chelate affinity chromatography.

subunit FimA forms the helical pilus rod. The rod contains up to 3000 copies of FimA, arranged in a right-handed, helical quaternary structure with 3.4 subunits per turn, a diameter of 7 nm and varying length of up to 2 μ m. A single copy of the adhesin FimH at the tip of the pilus and one or several copies of the minor subunits FimF and FimG form the linear tip fibrillum, which is connected to the distal end of the rod.³ The structural pilus subunits are homologous, one-domain proteins, except for the adhesin FimH, which has an additional, N-terminal lectin domain⁴ that recognizes mannose units on the target receptors uroplakin Ia and Ib on epithelial

cells of the urinary tract.⁵ The pilin domain of FimH and all the other structural subunits share a common immunoglobulin (Ig)-like fold, which, in contrast to the original Ig-fold, lacks the seventh (C-terminal) β -strand. The incomplete fold of the monomeric subunits renders them very unstable and prone to aggregation.^{4,6} In the context of the quaternary structure of the pilus, however, the subunits are extremely stable and cannot be dissociated and unfolded by boiling or treatment with high denaturant concentrations.^{7,8} The helical quaternary structure of the rod can be stretched and unwound under mechanical stress without dissociation, so that

stripping force is mitigated and does not lead to the detachment of bacteria from the target cell.^{9–11} This extraordinary stability of type 1 pili is achieved through strong mutual stabilization of neighboring subunits in the pilus that is reminiscent of a domain-swapping mechanism¹² and termed “donor strand complementation.”⁴ Except for FimH at the tip, all pilus subunits possess an N-terminal extension of 15–20 residues that serves as a donor strand completing the fold of the preceding subunit. Each subunit thus donates its N-terminal extension to the previous subunit and accepts an N-terminal donor strand from the next subunit.

Prior to pilus assembly, the individual pilus subunits are translocated to the periplasm, fold in the periplasm and form 1:1 complexes with the soluble, periplasmic assembly chaperone FimC (Fig. 1a). In FimC–subunit complexes, the fold of the subunits is stabilized by a β -strand segment of FimC.⁴ The chaperone, which is not a component of the assembled pilus, thus transiently becomes part of the subunits' tertiary structure. However, it inserts its donor strand segment in a parallel orientation relative to the C-terminal F-strand of the subunits, while donor strand insertion in the opposite, “antiparallel” orientation is observed in subunit–subunit interactions.^{4,13–20} In addition, FimC catalyzes pilus subunit folding^{21,22} and prevents the premature assembly of the subunits in the periplasm. FimC-bound subunits are then transferred to the assembly platform FimD, which recognizes chaperone–subunit complexes and catalyzes subunit assembly and translocation through the outer membrane and release of free FimC to the

periplasm (Fig. 1).^{23–28} The actual assembly reaction, catalyzed by FimD, is a donor strand exchange (DSE) reaction in which the donor strand of FimC capping the last subunit of the growing pilus is replaced by the donor strand of the next incoming subunit.^{17,27,29,30} The assembly of the pilus starts with FimH and formation of the tip fibrillum, followed by assembly of the pilus rod formed by FimA.

In the present study, we have characterized the structure, folding and stability of the main structural type 1 pilus subunit FimA toward a better understanding of the extraordinary stability of the rod and its assembly mechanism. For this purpose, we made use of three different FimA constructs (Fig. 1b): (i) wild-type FimA with its natural, 20-residue donor strand (FimAwt); (ii) a self-complemented FimA variant, termed FimAa, in which a second FimA donor strand was fused to the FimA C-terminus via a (Gly)₆-linker, allowing intramolecular self-complementation in FimAa and donor strand insertion in the antiparallel orientation as observed for subunit–subunit contacts; and (iii) the N-terminally truncated, assembly-incompetent variant FimAt, in which the major part (residues 1–14) of the natural FimA donor strand was replaced by a (His)₆-affinity tag.

We demonstrate that FimAa is the most stable, self-complemented pilus subunit analyzed so far and present the NMR structure of FimAa. The analysis of FimAt revealed that it is not capable of forming a defined tertiary structure, which is in contrast to all other donor-strand-depleted pilus subunits characterized so far, which, albeit being only marginally stable, retain the ability to form

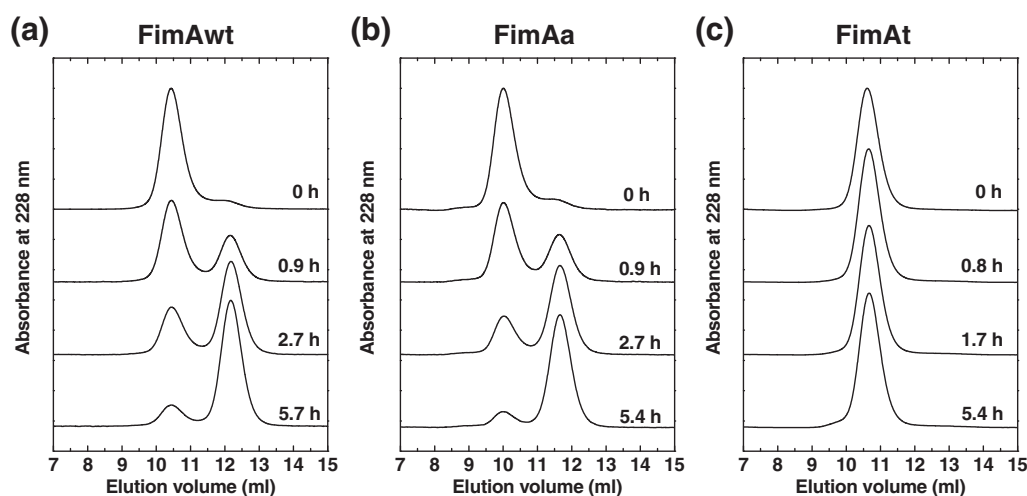


Fig. 2. Analysis of *in vitro* refolding of oxidized FimA and its variants at pH 7.0 and 25 °C by gel filtration. Disulfide-intact, unfolded FimAwt (a), FimAa (b) and FimAt (c) were refolded from 6 M GdmCl by rapid dilution (1:75) (final protein concentration: 2.5 μ M; final GdmCl concentration: 80 mM). After the reaction times indicated, aliquots were removed and immediately applied to a Superdex 75 gel-filtration column equilibrated with refolding buffer. The first peaks in (a) and (b) and the single species in (c) correspond to unfolded protein; the second peaks in (a) and (b) correspond to native protein.

tertiary structure.^{6,15,22} Surprisingly, wild-type FimA, in contrast to FimAt, proved to be folding competent and shows a previously unknown mechanism of self-complementation, in which its N-terminal donor strand stabilizes the FimA structure and keeps the protein monomeric in the absence of FimC. Finally, we show that spontaneous folding of FimAwt and FimAa is extremely slow (half-life ≈ 1.6 h), which explains the need for the catalyst FimC for pilus assembly *in vivo*. We show that FimC catalyzes FimA folding at least 400-fold and propose that the complex β -sheet topology of the native FimA structure is the main reason for its extremely slow, spontaneous folding rate.

Results

Spontaneous folding of FimA is extremely slow

Mature, wild-type FimA (FimAwt) and the variants FimAa and FimAt were produced in *E. coli* as cytoplasmic inclusion bodies. After solubilization of the inclusion bodies with concentrated guanidinium chloride (GdmCl) solution, the single disulfide bond in the constructs was formed by Cu^{2+} -catalyzed air oxidation in the presence of GdmCl. After removal of the denaturant by dialysis, a final anion-exchange chromatography step was applied, which yielded more than 20 mg of pure protein per liter of bacterial culture for each construct.

In a first folding experiment, disulfide-intact FimAwt, FimAa and FimAt were denatured with 6 M GdmCl, and refolding was initiated at pH 7.0 and 25 °C by rapid dilution to a final GdmCl concentration of 80 mM. After different reaction times, aliquots were removed and analyzed with rapid analytical gel filtration (15 min per run). The analysis shows that both FimAwt (Figs. 2a and 3a) and FimAa (Figs. 2b and 3b) fold extremely slowly to the more compact native structure. Within experimental error, both reactions exhibit comparable half-lives of 1.54 and 1.59 h, respectively. In contrast, no compaction was observed for FimAt, and the protein was always eluted during gel filtration at the retention volume corresponding to

the unfolded protein (Fig. 2c). Identical results were obtained from the folding reactions monitored with far-UV circular dichroism (CD) spectroscopy, which yielded the same rate constants of folding of FimAwt and FimAa within experimental error (Fig. 3c and d and Supplementary Fig. 1), and independently confirmed that FimAt is not capable of tertiary structure formation (Fig. 4c and d).

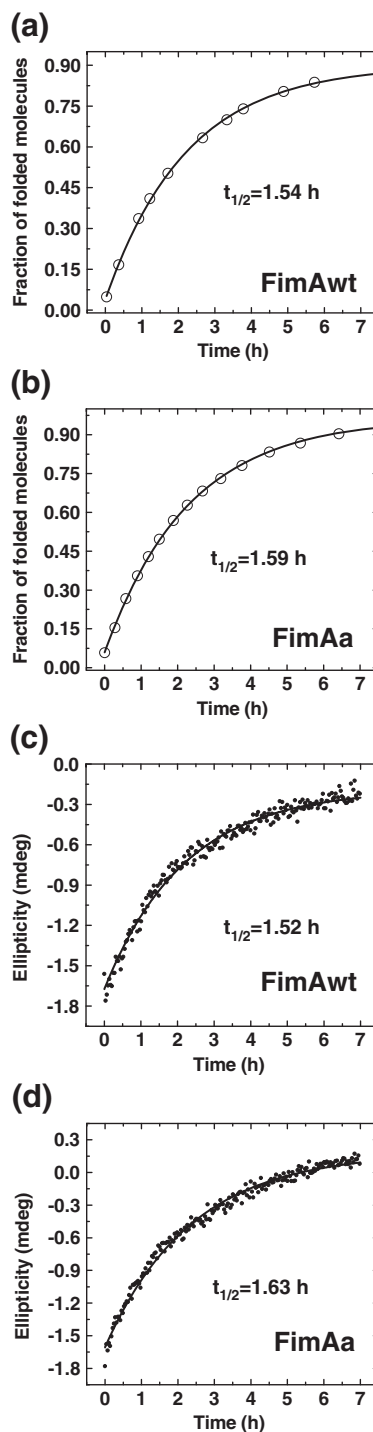
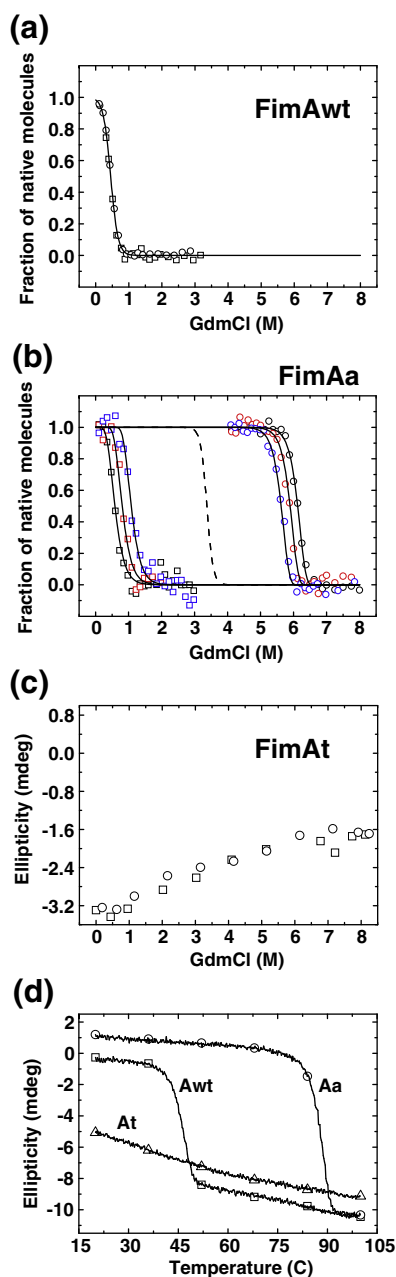


Fig. 3. Correlation between compaction (a and b) and secondary structure formation (c and d) during refolding of FimAwt and FimAa. Kinetics of formation of native molecules at pH 7.0 and 25°C of FimAwt (a and c) and FimAa (b and d) quantified with gel filtration (a and b) (cf. Fig. 2a and b) and secondary structure formation (c and d) analyzed by far-UV CD at 230 nm. The data in (a) to (d) could all be fitted with a single exponential function (continuous lines), and the deduced half-life times are indicated. The final GdmCl concentration during refolding was 80 mM in all cases.

FimAwt and FimAa adopt different conformations

We next measured GdmCl-induced unfolding and refolding transitions (Fig. 4a–c) and temperature-induced unfolding transitions (Fig. 4d) of FimAwt, FimAa and FimAt. The folding equilibrium at 25 °C of FimAwt was not attained after 1 day of incubation, and the unfolding midpoint was at higher (0.3 M) GdmCl concentrations than the refolding midpoint (data not shown). The equilibrium was, however, reached after 7 days of incubation, where unfolding and refolding were identical



within experimental error (Fig. 4a). Evaluation of these data according to the two-state model³¹ yielded a free energy of folding ($\Delta G_{\text{H}_2\text{O}}^0$) of $-10.1 \pm 0.6 \text{ kJ mol}^{-1}$ and a cooperativity of folding (m_{eq}) of $22.5 \pm 1.2 \text{ kJ mol}^{-1} \text{ M}^{-1}$. The experimental m_{eq} value is in good agreement with the expected value for a 15.8-kDa protein.³²

In contrast to FimAwt, the folding equilibrium of FimAa at 25 °C could never be attained. Figure 4b shows the unfolding and refolding transitions after 1, 7 and 17 days of incubation, all characterized by unfolding at high GdmCl concentration and refolding at low GdmCl concentration and the tendency of slowly moving toward equilibrium with increasing incubation times. The data were fully consistent with an unattained two-state equilibrium³³ [see Eqs. (1) and (2)] and yielded the following values for the folding and unfolding rate constants in the absence of denaturant (k_{F}^0 and k_{U}^0) and the kinetic m -values of folding and unfolding (m_{F} and m_{U}): $k_{\text{F}}^0 = 1.4 \times 10^{-4} \pm 2.2 \times 10^{-5} \text{ s}^{-1}$ ($t_{1/2} = 1.36 \text{ h}$), $k_{\text{U}}^0 = 1.8 \times 10^{-19} \text{ s}^{-1}$ ($t_{1/2} = 1.22 \times 10^{11} \text{ years}$; range of 95% confidence interval for k_{U}^0 is $1.3 \times 10^{-18} \text{ s}^{-1}$ to $1.6 \times 10^{-20} \text{ s}^{-1}$), $m_{\text{F}} = -5.3 \pm 0.2 \text{ M}^{-1}$ and $m_{\text{U}} = 5.4 \pm 0.3 \text{ M}^{-1}$. The extrapolated half-life of FimAa folding at zero denaturant is thus very close to that measured by CD spectroscopy and analytical gel filtration in 80 mM GdmCl (Figs. 2b and 3b). The extrapolated values for k_{F}^0 and k_{U}^0 yielded a free energy of folding of $-85 \pm 4 \text{ kJ mol}^{-1}$ [$\Delta G^0 = -RT \ln(k_{\text{F}}/k_{\text{U}})$], showing that FimAa is a thermodynamically extremely stable protein with very high activation energy barriers for unfolding and refolding, analogous to previously characterized, self-complemented pilus subunits with C-terminal donor strand.^{15,20,33–35}

The thermal unfolding transition of FimAa shows a very high melting temperature (T_{m}) of 87.8 °C, while the moderately stable FimAwt exhibits a T_{m} of 46.1 °C (Fig. 4d). The N-terminally truncated

Fig. 4. Dependence on GdmCl concentration and temperature of unfolding and refolding of FimAwt, FimAa and FimAt at pH 7.0. (a) GdmCl-dependent equilibrium unfolding (squares) and refolding (circles) of FimAwt at 25 °C after 7 days of incubation. The continuous line corresponds to a two-state state analysis. (b) Nonequilibrium unfolding (squares) and refolding (circles) of FimAa at 25 °C, recorded with far-UV CD at 230 nm after 1 day (black symbols), 4 days (red symbols) and 17 days (blue symbols). Continuous lines correspond to a global analysis according to an unattained two-state equilibrium [Eqs. (1) and (2)]. The broken line indicates the calculated equilibrium transition. (c) FimAt is essentially unfolded at pH 7.0 and 25 °C and does not show a cooperative, GdmCl-dependent unfolding/refolding transition. (d) Temperature-induced unfolding transitions of FimAwt (squares) and FimAa (circles). FimAt (triangles) is permanently unfolded under these conditions.

Table 1. Input for structure calculation and characterization of the energy-minimized NMR structures of FimAa

Quantity	Value ^a
NOE upper distance limits	4175
Residual target function (Å ²)	1.94±0.71
Residual NOE violations	
Number ≥0.1 Å	30±1
Maximum (Å)	0.15±0.04
Residual angle violations	
Number ≥2.5°	0±1
Maximum (°)	2.12±0.9
AMBER energies (kcal/mol)	
Total	-7171.69±72.31
van der Waals	-501.03±19.08
Electrostatic	-8051.60±62.86
r.m.s.d. to the mean coordinates	
for residues 20–160 and 167–184 (Å)	
N, C ^α , C ^γ (backbone)	0.55±0.06
All heavy atoms	0.78±0.06
Ramachandran plot statistics ^b (%)	
Most favored regions	72.3±2.5
Additional allowed regions	22.6±2.9
Generously allowed regions	3.6±1.1
Disallowed regions	1.5±0.9

^a Except for the first top entry, the average value for the 20 energy-minimized conformers with the lowest residual ATNOS/CANDID target function values and the standard deviation among them are given.

^b As determined by PROCHECK.³⁶

construct FimAt showed no transitions detectable with CD spectroscopy after incubation for 1 day (Fig. 4c), confirming that FimAt does not adopt tertiary structure.

The kinetics and thermodynamics of folding of FimAwt and FimAt differ from those of previously characterized pilus domains in several aspects. While donor-strand-depleted variants of the type 1 pilus subunits FimF and FimG retained the ability to adopt a defined tertiary structure (albeit with low thermodynamic stability in the range of -10 kJ mol⁻¹),^{6,15,22} the folding competence of FimAt is completely lost. Consequently, the ability of FimAwt to adopt a defined tertiary structure in our *in vitro* experiments in the absence of FimC is most likely a result of intramolecular insertion of residues of the natural, N-terminal FimA donor strand. This behavior has so far never been observed for an isolated pilus subunit and explains the high stability of folded FimA monomers against self-polymerization *in vitro*.⁷ Sterically, intramolecular self-complementation via the natural, N-terminal donor strand of FimAwt is only possible when the N-terminal extension of FimAwt inserts parallel relative to the C-terminal F-strand of FimA, that is, the same orientation of the donor strand that is observed in chaperone-subunit complexes (see Discussion for details). The self-complemented construct FimAa, in which antiparallel donor strand insertion is enforced (see below), shows the same nonequilibrium behavior in protein folding (unat-

tained two-state equilibrium) that has so far been observed for all other pilus subunit constructs with artificial C-terminal donor strand.^{15,20,33–35} However, FimAa shows by far the highest stability and slowest folding and unfolding kinetics relative to the other self-complemented pilus domains investigated so far.^{15,20,33–35}

NMR structure of FimAa

Based on the stability data above on FimAwt and FimAt, the construct FimAa has alternative folding possibilities in that it can incorporate either its N- or C-terminal donor strand segment into its tertiary structure. To confirm the anticipated incorporation of the C-terminal extension into the donor strand binding groove and to study the structural details of donor strand complementation in the type 1 pilus rod, we solved the solution structure of FimAa with NMR spectroscopy. Residues 1–19 of the natural, N-terminal FimA donor strand in FimAa were found to be flexibly disordered, while the structure of the self-complemented pilin domain of FimAa (residues 20–160 and 167–184) could be determined with high precision (r.m.s.d. values for 20 conformers of the final energy-refined ensemble are 0.55 and 0.78 Å for the backbone and heavy atoms, respectively; cf. Table 1 and Fig. 5a; residue numbering according to FimAwt). FimAa adopts the expected Ig-like fold, which has been observed for all other structures of pilus subunits from chaperone-usher systems determined so far.^{4,13–20,35} The artificial C-terminal extension (G-strand in the solved NMR structure) complements the Ig-fold in a canonical manner (i.e., runs antiparallel with the F-strand). The fold is mainly determined by two β-sheets, the first being composed of strand A (29–38), the donor strand G (167–182), strand F (146–159) and strand C (67–75), and the second being composed of strands B' (45–49), B'' (52–62), E (127–141), D' (99–105) and D'' (122–124), that are packed against each other in a wound fashion (Fig. 5b and c). In addition, a one-turn α-helix (residues 25–28) and two other one-turn 3₁₀ helices (39–41 and 80–82) could be identified in the structure. The disordered N-terminal segment of FimAa (1–19), which is supposed to act as a donor strand for a preceding subunit in the assembled pilus, shows no observable interaction with the rest of the protein (Fig. 5a) as was also observed before for the N-terminal segment of the subunit FimF.¹⁴ The hexaglycine linker (161–166), inserted between the natural FimA C-terminus and the C-terminal FimA donor strand segment, appears to be unstructured and dynamic, indicating that the chosen (Gly)₆ linker is long enough and does not impose any strain on the insertion of the additional, self-complementing donor strand G (167–182). All unstructured regions in FimAa (see high local r.m.s.d. values in Figs. 5a and 6a) almost perfectly

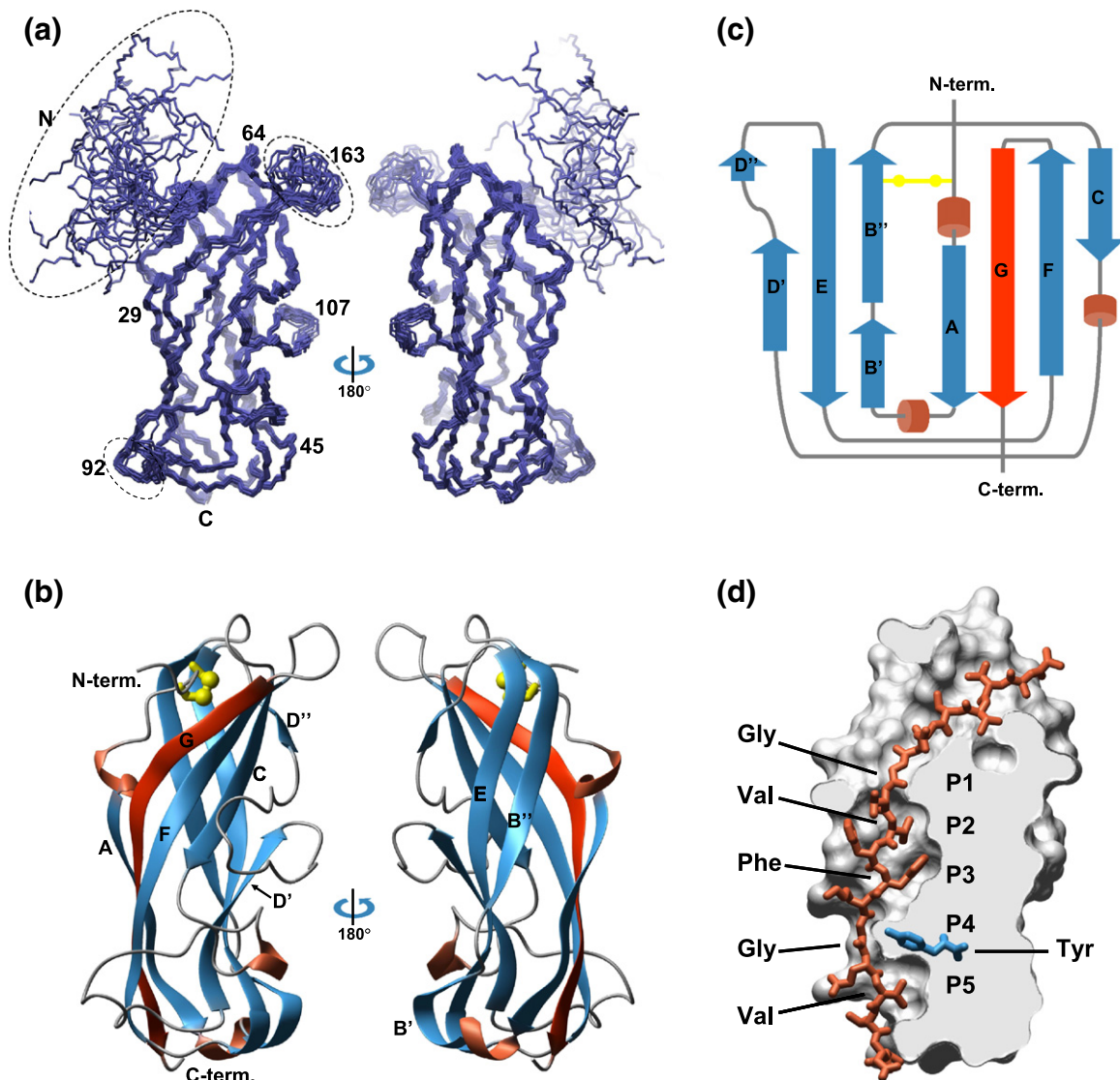


Fig. 5. NMR structure of FimAa. (a) Two-side view of the polypeptide backbone for a bundle of 20 energy-refined DYANA conformers. The superposition was performed by minimizing the r.m.s.d. value of the backbone atoms N, C $^{\alpha}$ and C' of residues 21–160 and 167–184 [best-defined regions (Table 1)]. The N- and C-termini and sequence numbers of selected amino acids are shown, and the three regions with the highest local r.m.s.d. (see Fig. 6a) are highlighted with a broken line. (b) Two-side view (ribbon drawing) of one (with the best target function) of the 20 energy-refined DYANA conformers of FimAa from (a), which was selected to represent the NMR structure. The engineered, C-terminal β -strand is shown in red. The protein orientations are the same as in (a). (c) Topology model of FimAa. Helices are represented by cylinders, and β -strands, by arrows. The disulfide bond is depicted in yellow. The individual β -strands are sequentially labeled by letters. The color coding is the same as in (b). (d) Slice through the surface representation of FimAa, with the engineered G-strand represented in red by a sticks model. Binding pockets P1–P5 (general nomenclature according to Sauer *et al.*¹⁷) interact with the donor strand side chains oriented toward the protein core. Pockets P2, P3 and P5 are deep and occupied by Val175, Phe177 and Val181, respectively, while the P1 and P4 pockets are shallow and occupied by glycine residues. (a), (b) and (d) were generated with the program MOLMOL.³⁷

coincide with the regions showing low relative intensities of the heteronuclear $^{15}\text{N}\{^1\text{H}\}$ nuclear Overhauser enhancement (NOE) values (I_{rel}) (Fig. 6b), indicating that the flexible regions are not a result of missing constraints in the structure determination.

The introduced C-terminal donor strand in FimAa occupies the groove between the F- and the A-strands as observed in donor-strand-complemented structures of other pilin proteins.^{4,13–20,35} Amino acid side chains facing the interior of the protein are located in five commonly found binding pockets P1–P5

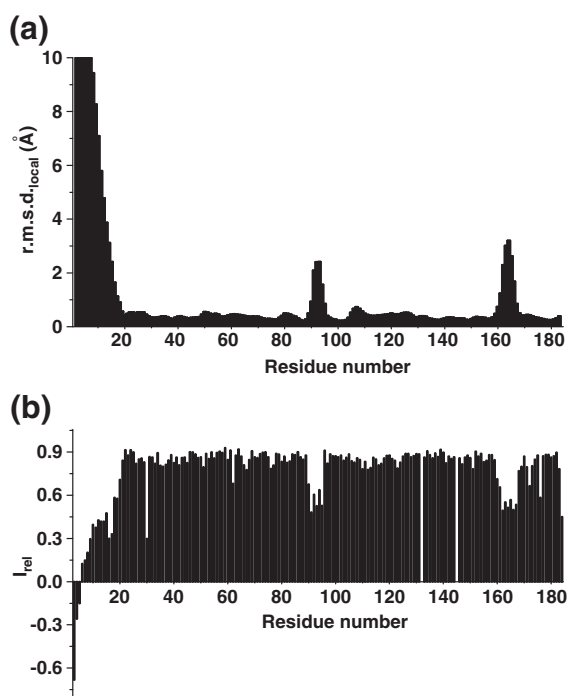


Fig. 6. Heteronuclear $^{15}\text{N}\{^1\text{H}\}$ -NOE measurements of FimAa and local r.m.s.d. values. High local r.m.s.d. values (a) correlate with the relative intensities of the heteronuclear $^{15}\text{N}\{^1\text{H}\}$ -NOE values (I_{rel}) smaller than 0.6 (b). The local r.m.s.d. value of a given residue was calculated by superimposing three residues centered at this residue for minimal r.m.s.d. There are three distinct regions in the protein showing low values of the heteronuclear NOEs: the N-terminal extension (residues 1–19), the engineered glycine linker region preceding the C-terminal G donor strand (residues 161–167) and a short loop in the middle of the sequence (residues 90–95). All three polypeptide segments appeared to possess variable conformations (high local r.m.s.d. values) in the bundle of 20 representative FimAa NMR structures (Fig. 5a). The rest of the protein shows homogeneously high heteronuclear NOE values around 0.85.

(Fig. 5d).¹⁷ P2, P3 and P5 are deep and accommodate the bulky hydrophobic residues V175, F177 and V181, respectively (corresponding to V11, F13 and V17 of the natural FimA donor strand), while positions P1 and P4 are shallow and are occupied by glycines (natural donor strand residues G9 and G15). The shallow form of the P4 pocket is largely caused by residue Y137 (depicted in blue in Fig. 5d). Pocket P4 must be occupied by a Gly residue to preserve the contiguous β -sheet hydrogen-bonding network between the inserted C-terminal donor strand and strands A and F of the FimA pilin domain and to define the register of donor strand insertion. This result is fully consistent with the high conservation of corresponding Gly residue opposite to the P4 site in donor strands of pilin subunits.^{4,15}

FimC selectively binds unfolded FimAwt and accelerates folding of FimAwt at least 400-fold

Intramolecular self-complementation in FimAwt via its natural N-terminal donor strand predicted that folded FimAwt can no longer be recognized by FimC. To test the ability of FimC to bind folded and unfolded FimAwt monomers, we refolded GdmCl-denatured FimAwt by dilution at pH 7.0. Aliquots of the refolding reaction were withdrawn after different incubation times and rapidly mixed with a 1.6-fold molar excess of FimC, and formation of native FimC–FimAwt complexes was quantified immediately by rapid cation-exchange chromatography (Fig. 7a). When FimC was added immediately after the onset of refolding, all FimAwt molecules formed a complex with FimC. With increasing refolding times prior to FimC addition, the ability of FimC to bind FimAwt was continuously reduced. The FimC–FimAwt peak areas were plotted against FimAwt refolding time and could be fitted to a mono-exponential decay with a half-life of 1.8 h (Fig. 7b), a value very similar to the half-life of FimAwt folding (1.54 h, cf. Fig. 3). In contrast, the same experiment with FimAt instead of FimAwt demonstrated that FimC was capable of forming native FimC–FimAt complexes at any time of FimAt refolding (Fig. 7c). As isolated FimAt is unfolded under physiological conditions (Figs. 2c and 4c and d), the results show that FimAt can only fold in the presence of FimC, that FimC only recognizes the unfolded state of FimA (followed by catalysis of FimA folding; see below) and that binding of unfolded FimA occurred rapidly within the dead time of the experiment (less than 1 min), that is, the time required for applying the FimC–FimA mixture onto the analytical cation-exchange column. Assuming that this upper time limit corresponds to more than three half-lives of the binding reaction and taking into account the initial concentrations of FimC and FimAt, we calculated a lower limit for the association rate constant k_{on} of $10^4 \text{ s}^{-1} \text{ M}^{-1}$ for binding of unfolded FimAt by FimC. Consequently, a lower limit for the rate of FimAt folding (k_{F}) on the surface of FimC of about 0.05 s^{-1} could be estimated. FimC thus accelerates folding of FimA relative to the uncatalyzed reaction ($k_{\text{F}} = 1.25 \times 10^{-4} \text{ s}^{-1}$) by at least 400-fold.

Folded, monomeric FimAwt can be recovered for the assembly process via unfolding and rebinding to FimC

Native FimAwt that had been refolded in the absence of FimC was not capable of forming FimC–FimAwt complexes upon addition of a 1.6-fold excess of FimC on the timescale of several minutes at concentrations of $5 \mu\text{M}$ FimA and $8 \mu\text{M}$ FimC (see above) and behaved like a stable monomer. As FimA can, however, only be incorporated efficiently

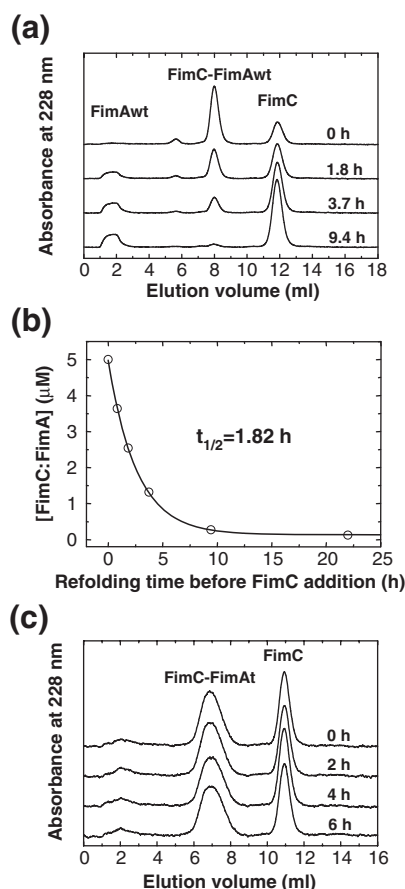


Fig. 7. FimAwt, in contrast to FimAt, loses the ability of rapid binding to FimC during spontaneous refolding. (a) FimAwt in 6 M GdmCl was diluted at (1:75) 25 °C with refolding buffer (pH 7.0; final FimAwt concentration: 5 μ M). After the refolding times indicated, excess FimC (8 μ M) was added, and the reaction products were immediately separated by fast cation-exchange chromatography at 4 °C and pH 6.7 on a Resource S column with a linear NaCl gradient. Representative elution profiles are shown. (b) Concentration of formed FimC–FimA complexes as a function of FimAwt refolding time. The decrease in the FimC–FimA peak area from (a) was fitted according to a single exponential decay (continuous line). (c) Addition of FimC after different times of dilution of FimAt (in 6 M GdmCl) with refolding buffer [pH 7.0, 25 °C, same conditions as in (a)]. The experiment shows that all FimAt molecules are quantitatively bound by FimC at any “refolding” time and that FimAt can adopt native structure only when bound to FimC.

into the pilus when folded and bound to FimC,^{25,27,29,30} this raised the question of whether folded FimAwt monomers can form complexes with FimC under certain conditions and thus potentially be recovered for pilus assembly. We therefore performed long-term *in vitro* incubation experiments in which folded FimAwt was incubated at different concentrations (2.5 μ M, 10 μ M and 40 μ M)

in the presence of equimolar amounts of FimC, and the reaction products were analyzed at different time points by size-exclusion chromatography (Fig. 8). At 25 °C, we could not detect formation of FimC–FimAwt complexes or formation of oligomers of FimAwt after incubation for several hours at any protein concentration (data not shown). We then investigated the binding of FimAwt to FimC at 37 °C. Figure 8a shows that FimAwt alone was only marginally stable at 37 °C, with about 15% of the molecules unfolded at equilibrium. This is consistent with its low thermal unfolding midpoint of 46.1 °C and low free energy of folding (-10.1 kJ mol⁻¹) at 25 °C (Fig. 4d). In the absence of FimC, only a very small fraction of FimAwt oligomers were formed after 6.7 h at 37 °C (Fig. 8a). In the presence of equimolar amounts of FimC and low FimAwt concentrations (2.5 μ M), formation of the FimC–FimAwt complex occurred at 37 °C, and an apparent equilibrium was established after about 2 h of incubation, with a significant amount of monomeric FimC and folded FimAwt still present (Fig. 8b). At medium (10 μ M) and high (40 μ M) FimAwt concentrations, not only was the equilibrium shifted toward the FimC–FimAwt complex, a significant fraction of FimC-bound oligomers of FimAwt with higher molecular mass also appeared after several hours of incubation (Fig. 8c and d). Based on previous results on FimA self-assembly in the presence of FimC³⁰ and data on PapA oligomerization from PapC–PapA complexes,³⁸ we interpret the appearance of the latter species as a slow, spontaneous oligomerization reaction through DSE between FimC–FimAwt complexes after unfolded FimAwt is pulled from the FimAwt folding equilibrium through binding to FimC. The reactions can be described according to Scheme 1, where C corresponds to FimC; A^{N*} and A^U correspond to the native and unfolded FimAwt monomers, respectively; CA is the complex between FimC and folded, assembly competent FimA; and CAA and CAAA are oligomers of FimAwt with the last FimA bound to FimC. The kinetics of formation of FimC-bound FimAwt polymers are in agreement with the previously measured rate constant (k_{DSE}) of 4 s⁻¹ M⁻¹ for uncatalyzed DSE between a FimC-bound donor and a FimC-bound acceptor subunit.³⁰ As the kinetics of formation of the FimC–FimAwt complex showed essentially no dependence on protein concentration (Fig. 8b–d), we conclude that unfolding of FimAwt, which is a prerequisite for binding to FimC, is rate limiting for the initial formation of FimC–FimAwt complexes. In contrast, DSE between FimC–FimAwt complexes was strongly favored with increasing protein concentrations (Fig. 8b–d). Together, the results indicate that folded FimAwt monomers, albeit assembly incompetent, may not necessarily be lost for pilus rod assembly at 37 °C. Folded

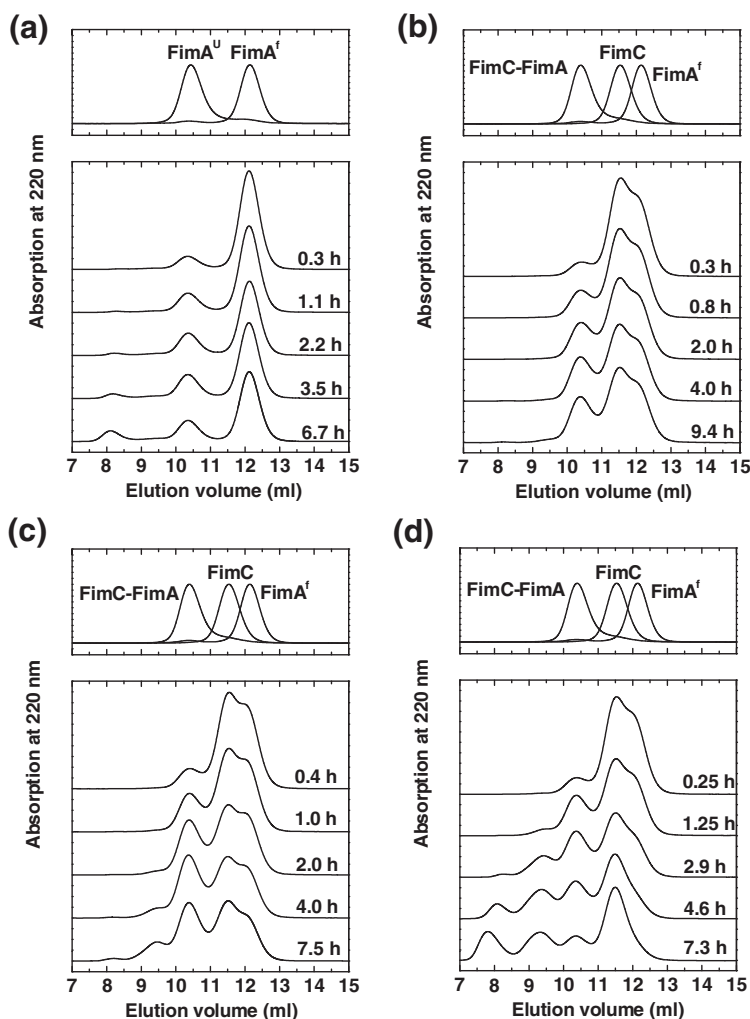
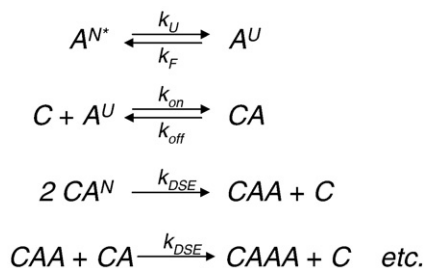


Fig. 8. Folded FimAwt monomers can be recovered for pilus rod assembly at 37 °C through rebinding to FimC. (a) Folded FimAwt (40 μM) was incubated alone at 37 °C and pH 7.0 and analyzed after different times by analytical gel filtration. The results show that about 15% of the FimAwt molecules are unfolded (FimA^U) at 37 °C and that FimAwt has a low tendency of slowly forming oligomers under these conditions. (b–d) FimAwt monomers at different concentrations [2.5 μM (b), 10 μM (c) and 40 μM (d)] were incubated at pH 7.0 and 37 °C in the presence of equimolar amounts of FimC, and reaction products were analyzed by analytical gel filtration after the incubation times indicated. Panels at the top show the elution profiles of the individual compounds.

FimAwt monomers may, for example, accumulate *in vivo* when the periplasmic concentrations of FimC do not suffice to complex every newly synthesized FimAwt molecule entering the periplasm. As soon as free FimC molecules become available, the fraction of unfolded FimAwt molecules will be pulled from the FimAwt folding equilibrium by binding to FimC and become assembly competent again.



Scheme 1. Reaction mechanism for FimA recovery by FimC followed by FimA oligomerization.

Discussion

The role of donor strand complementation for FimA stability and folding

FimA is the main structural subunit of type 1 pili forming the pilus rod, which is connected to the tip fibrillum formed by FimF, FimG and the adhesin FimH. The self-complemented variant FimAa was constructed in order to mimic the state of FimAwt in the context of the quaternary structure of the pilus rod. Denaturant-induced unfolding/refolding transitions and thermal unfolding transitions (Fig. 4) showed that FimAa is an extraordinarily stable protein with extremely slow rates of unfolding and refolding. The extremely slow unfolding of FimA in the assembled pilus is essential for pilus stability and function, as stochastic unfolding of a single FimA subunit in an assembled pilus rod would cause fragmentation of the pilus and loss of the functional tip fibrillum to the extracellular medium.

Given that one pilus rod consists of 500–3000 FimA subunits, one can calculate that only 50% of the pili would survive a 10-h incubation at 25 °C if a single FimA subunit unfolded only once within several years (corresponding to $k_U^0 \approx 10^{-8} \text{ s}^{-1}$). The extrapolated half-life of FimAa unfolding (1.2×10^{11} years) thus guarantees the infinite stability of pilus rods against dissociation in the extracellular environment. Very high activation barriers for unfolding and refolding appear to be a general feature of donor-strand-complemented pilus subunits, as similar observations were made for other subunits of the type 1 pilus tip fibrillum^{15,33} and subunits from the F1 antigen fiber,²⁰ Dr fimbriae³⁴ and F4 fimbriae.³⁵ Despite the similarities above between FimA and previously characterized pilus subunits, the present study also revealed two unique features of FimA that have not been observed so far for pilus subunits. First, FimAwt has the ability of intramolecular self-complementation via its own N-terminal extension. Second, the N-terminally truncated construct FimAt is intrinsically unstable and not capable of adopting a defined tertiary structure, while other donor-strand-depleted pilus subunits such as FimGt and the FimH pilin domain retained the ability to fold spontaneously, albeit with relatively low free energies of folding in the range of -10 kJ mol^{-1} .^{6,15,22} A strict dependence on the chaperone or a donor strand from another subunit had indeed been postulated originally when the structures of the first chaperone–subunit complexes were solved^{4,16} but never been demonstrated experimentally. Assuming that the far-UV CD signal change in FimAt upon folding would be the same as for FimAwt and the accuracy of the measured CD signal is around 5% of this signal change, the free energy of FimAt folding must be above $+6 \text{ kJ mol}^{-1}$. The unusual, intramolecular self-complementation via the N-terminal donor strand in FimA thus increases the stability of FimAwt relative to FimAt by at least 16 kJ mol^{-1} . The stabilization most likely comes from formation of β -sheet hydrogen bonds between the N-terminal extension and the first and last strands of the FimA pilin domain (A- and F-strands, respectively) and further interactions of donor strand side chains with side-chain binding pockets in the pilin domain (Fig. 5). The fact that FimC does not bind folded FimAwt (Figs. 7 and 8) supports the idea that the groove between the A- and the F-strands in the FimAwt structure is occupied by its own N-terminal extension.

Raw modeling based on the FimAa structure showed that the P2 and P3 pockets can be reached without steric clashes by the equivalent Val and Phe residues (V10 and F12) in the N-terminal extension (occupied by V175 and F177 in FimAa), but in a reversed order. As the N-terminal donor strand of FimA must have the opposite orientation to the C-terminal donor strand of FimAa (parallel *versus*

antiparallel with F-strand), main-chain hydrogen bonding and hydrophobic core packing are presumably less optimal, consistent with the lower stability of FimAwt compared to FimAa.

Remarkably, the construct FimAa has alternative folding possibilities in that it can incorporate either the N- or the C-terminal extension into its tertiary structure. However, the significantly higher thermodynamic stability of FimAa relative to FimAwt and the NMR structure of FimAa show that FimAa spontaneously folds to the thermodynamically most stable conformation, in which the C-terminal extension is incorporated into the tertiary structure.

The dynamics of the folding equilibria of FimAwt and FimAa differ essentially only in the extrapolated unfolding rate, which is about 13 orders of magnitude lower in FimAa than in FimAwt, while the folding rates are identical (Fig. 2). This finding is in line with data on the third fibronectin type III domain from human tenascin (TNfn3) also possessing an Ig-like fold.³⁹ Shortening the C-terminal G-strand by two amino acids accelerated unfolding by almost 2 orders of magnitude while leaving the folding rate unchanged. High-resolution Φ -value analysis of TNfn3 showed that the N- and C-terminal strands A and G are the latest to be formed in the folding process and are not present in the structure of the transition state; hence, they do not contribute to the free-energy difference between unfolded and transition states and therefore do not influence the folding rate.⁴⁰ However, it does contribute to the energy of the native state; thus, it influences both the thermodynamic stability of the protein and its unfolding rate. A similar mechanism for FimA (i.e., that incorporation of the G-strand happens at the later stages of the folding process and is not yet present in the structure of the transition state) would explain why FimAa and FimAwt fold with the same rates.

For FimD-catalyzed incorporation of FimA into a growing pilus, FimA must be bound to FimC.^{27,30} Our results show that FimA binds to FimC only in the unfolded state and then folds on the surface of the chaperone. Regarding the fact that a single type 1 pilus contains up to 3000 copies of FimA, we speculate that the folded, assembly-incompetent FimAwt monomer could represent a periplasmic storage form of FimA under conditions where FimA is present at excess over FimC. Folding through self-complementation via its own N-terminal donor strand could possibly protect FimA from proteolytic degradation, as the pilin domain alone would be permanently unfolded and prone to degradation. In addition, our data show that folded FimAwt monomers can, in principle, be recovered for pilus assembly at 37 °C (Fig. 8). As free periplasmic FimC is regenerated after incorporation of FimA into the pilus, this mechanism could become effective *in vivo* as soon as FimA is no longer present at excess over FimC in the periplasm.

The slow spontaneous folding of FimA

The apparent nonequilibrium unfolding/refolding behavior of FimAa and the folding data of FimAwt proved to be fully consistent with two-state folding. Both proteins show identical and extremely low spontaneous folding rates in the range of 10^{-4} s^{-1} (1.5-h half-life). This is about 50 times slower than the previously reported folding rates of the subunits FimH⁶ and FimG.²² The extremely slow spontaneous folding rate of FimA cannot be explained by proline cis–trans isomerization, since FimA has only two trans prolines, and cis-to-trans isomerization of prolyl peptide bonds is 2–3 orders of magnitude faster than FimA folding.⁴¹ In fact, spontaneous FimA folding is 1–2 orders of magnitude slower than the slowest known folding rates of one-domain proteins that show two-state folding^{42–48} (see Supplementary Material).

As the rate of FimA folding is an extreme outlier in the statistics of known folding rates of small, one-domain two-state folders, we calculated the contact order (CO) and the average contact order (ACO) of the FimAa NMR structure, which, together with other parameters of topological tertiary structure complexity, correlates well with experimentally determined folding rates.^{49,50} Another parameter of topological complexity that predicts folding rates with reasonable accuracy is the number of sequence-distant native pairs (Q_d).^{47,51} The calculated values of these measures of topological complexity for FimAa are extraordinarily high. Specifically, an ACO value of 30.4 is calculated for contacts present in all conformers (Supplementary Material). The entropic contribution of topological complexity of the FimA structure to its free-energy barrier of folding may thus be the most important reason for its extremely slow folding rate of $1.3 \times 10^{-4} \text{ s}^{-1}$.

The process of *in vivo* pilus assembly was reported to be very rapid and presumably completed in less than 5 min.⁷ The slow spontaneous folding of FimA cannot satisfy the demand of bacteria for rapid growth of its pili and underlines the crucial importance of FimC as a folding catalyst. An acceleration factor of 100-fold for folding of the subunit FimG subunit has been reported previously.²² In the case of FimA, folding is also catalyzed by at least 2 orders of magnitude (see above). As FimC interacts with both the N- and the C-terminal β -strands of all subunits in solved structures of chaperone–subunit complexes, an intriguing possible mechanism of FimC catalysis is the simultaneous recognition of the N- and C-terminal segments of the subunit while the rest of the subunit is still unfolded, which in turn would lead to a reduction in contact order. Finally, we are confident that the NMR structure of FimAa, which represents the conformational state of FimAwt in the context of the assembled pilus and provided the

information on the register of donor strand insertion, will provide the basis for reconstruction of an atomic model of the entire type 1 pilus rod together with the known electron density maps of the rod obtained from electron microscopy data.⁵²

Materials and Methods

Expression and purification

FimAwt (159 amino acid residues, 15.8 kDa), FimAt with N-terminal (His)₆-tag (153 amino acid residues, 15.5 kDa) and FimAa (184 amino acid residues, 18.0 kDa) (see Supplementary Material for the primary structures of all constructs) were expressed at 37 °C in 2×YT medium (16 g/l tryptone, 10 g/l yeast extract, 5 g/l NaCl) as cytoplasmic inclusion bodies in the *E. coli* strain BL21 (DE3) using the T7-promoter-based expression plasmid pET-11a (Novagen). Protein expression was induced at an OD_{600 nm} of 0.5 with 1 mM IPTG; bacteria were grown for additional 4 h. After cell disruption, inclusion bodies of FimAwt, FimAa and FimAt were isolated as described⁵³ and dissolved in 6 M GdmCl, 50 mM Tris–HCl, pH 8.0, 1 mM ethylenediaminetetraacetic acid and 50 mM DTT. The proteins were applied to a gel-filtration column equilibrated with 20 mM Tris–HCl, pH 8.0, 6 M GdmCl and 1 mM DTT. Eluted proteins were diluted to final concentration of 30 μ M, and 0.1 mM CuCl₂ was added to catalyze formation of the single disulfide bond in the constructs through air oxidation. Residual DTT also became oxidized under these conditions. The absence of free thiols after air oxidation was verified with Ellman's assay.⁵⁴ The oxidized proteins were refolded through dialysis against 10 mM Tris–HCl, pH 8.0, for 12 h. The refolded proteins were then subjected to anion-exchange chromatography as a final purification step: proteins were applied to a Resource Q column (Amersham Biosciences) in 10 mM Tris–HCl, pH 8.0, and eluted with a gradient from 0 to 500 mM NaCl. Purified proteins were concentrated by ultrafiltration under buffer exchange to 10 mM sodium phosphate, pH 7.0, and 200 mM NaCl. The yields of purified protein were above 20 mg per liter of bacterial culture for all constructs. FimC was expressed and purified as described previously.⁵⁵

Protein concentrations

Protein concentrations were determined via their extinction coefficients at 280 nm (FimC: 22,900 M⁻¹ cm⁻¹; FimAwt, FimAa and FimAt: 2680 M⁻¹ cm⁻¹) and by NMR spectroscopy using the method PULCON.⁵⁶

Analytical gel filtration

Analytical gel-filtration experiments were performed on an ÄKTA™ purifier system (Amersham Biosciences) using a Superdex™ 75 HR 10/30 column (Amersham Biosciences) equilibrated with 10 mM sodium phosphate, pH 7.0 and 200 mM NaCl at a flow rate of 1 ml min⁻¹ at room temperature (duration of a single run: 15 min). Eluted proteins were detected by their absorbance at either 228 nm or 220 nm.

Analytical ion-exchange chromatography

For following the kinetics of FimA–FimC complex formation, the reactions were stopped after different times by cooling on ice, and the reaction products were separated by analytical cation-exchange chromatography at 4 °C and pH 6.7 on a Resource S column using a linear NaCl gradient. Eluted proteins were detected via the absorbance at 228 nm.

Refolding kinetics

Chemically denatured proteins (188 μM in 6 M GdmCl, 10 mM sodium phosphate, pH 7.0 and 200 mM NaCl) were refolded at 25 °C through rapid dilution (1:75) with 10 mM sodium phosphate, pH 7.0, and 200 mM NaCl (final GdmCl concentration: 80 mM; final protein concentration: 2.5 μM). After different incubation times, aliquots were removed and analyzed by gel filtration as described above. The peak areas corresponding to the native proteins were quantified and plotted against refolding time. Alternatively, refolding of FimAwt and FimAa was followed via the change in the far-UV CD signal at 230 nm using a JASCO J-710 spectropolarimeter. All refolding reactions could be fitted with a single exponential function.

Temperature-induced unfolding transitions

Thermal unfolding transitions were followed by the change in the far-UV CD signal at 230 nm. Transitions were recorded in 10 mM sodium phosphate pH 7.0 and 200 mM NaCl at a protein concentration of 7.5 μM in a 0.1-cm quartz cuvette, using a Peltier element for temperature control and a heating rate of 0.5 °C per minute.

GdmCl-dependent unfolding and refolding transitions

Stock solutions of native protein (0.8 mM in 10 mM sodium phosphate, pH 7.0 and 200 mM NaCl) and unfolded protein (0.35 mM in 6 M GdmCl, 10 mM sodium phosphate, pH 7.0 and 200 mM NaCl) were diluted with 10 mM sodium phosphate, pH 7.0 and 200 mM NaCl containing different GdmCl concentrations and incubated at 25 °C. The final denaturant concentrations were determined via the refractive index of the solutions.⁵⁷ After different incubation times, the fractions of folded (f_N) and unfolded ($f_U = 1 - f_N$) molecules were assessed via the far-UV CD signal at 230 nm.

Unfolding and refolding transitions of FimAwt reached equilibrium within 7 days of incubation were evaluated according to the two-state model and normalized using a six-parameter fit.⁵⁸ Nonequilibrium unfolding and refolding transitions of FimAa were recorded after different incubation times and analyzed according to the theory of an unattained two-state equilibrium³³ using Eqs.(1) and (2):

$$f_N(D, t) = \frac{k_F^0 \cdot e^{m_F \cdot D}}{k_F^0 \cdot e^{m_F \cdot D} + k_U^0 \cdot e^{m_U \cdot D}} + \left(f_N(0) - \frac{k_F^0 \cdot e^{m_F \cdot D}}{k_F^0 \cdot e^{m_F \cdot D} + k_U^0 \cdot e^{m_U \cdot D}} \right) \times \exp(-(k_F^0 \cdot e^{m_F \cdot D} + k_U^0 \cdot e^{m_U \cdot D}) \cdot t) \quad (1)$$

where k_F and k_U are the rates of folding and unfolding, respectively; m_F and m_U describe the dependence of $\ln(k_F)$

and $\ln(k_U)$ on denaturant concentration, respectively; D is the denaturant concentration; and t is the incubation time, with $f_N(0) = 1$ for the unfolding transition and $f_N(0) = 0$ for the refolding transition.

Original spectroscopic data were normalized according to Eq. (2)

$$S(D) = S_U^0 + n_U \cdot D + (S_N^0 + n_N \cdot D - S_U^0 - n_U \cdot D) \cdot f_N(D) \quad (2)$$

where S_N^0 and S_U^0 are the spectroscopic signals of the native (N) and unfolded (U) states at zero denaturant, respectively, and n_N and n_U are the dependencies of the signal of N and U on denaturant concentration, respectively. The original data of nonequilibrium unfolding and refolding transitions were fitted globally with ORIGIN (MicroCal Inc.).

NMR sample preparation, data collection and structure determination

For the production of uniformly $^{15}\text{N}/^{13}\text{C}$ -labeled FimAa, cells of *E. coli* strain BL21 (DE3) were grown in M9 minimal medium containing $^{15}\text{NH}_4\text{Cl}$ (1 g/l) and ^{13}C -glucose (2 g/l) as the sole nitrogen and carbon sources. Induction with IPTG and protein purification were performed as described above for growth in rich medium. All NMR measurements were performed at 25 °C in 20 mM phosphate, pH 7.0, either on a Bruker DRX 750-MHz spectrometer or on an Avance 900-MHz spectrometer. Data were processed using XWINNMR (Bruker) and analyzed with CARAM. Sequence-specific backbone assignment was achieved using two-dimensional (2D) [$^{15}\text{N}, ^1\text{H}$] heteronuclear single quantum coherence (HSQC), 2D [$^{13}\text{C}, ^1\text{H}$]-HSQC, three-dimensional (3D) ct-HNCA, 3D HN(CO)CA and 3D HNCACB experiments. ^1H and ^{13}C side-chain assignments were obtained based on 3D HC(C)H total correlated spectroscopy (TOCSY), 3D ^{15}N -[$^1\text{H}, ^1\text{H}$]-TOCSY, 3D nuclear Overhauser enhancement spectroscopy (NOESY)-[$^{15}\text{N}, ^1\text{H}$]-HSQC and 3D NOESY-[$^{13}\text{C}, ^1\text{H}$]-HSQC experiments.⁵⁹ The mixing times in HC(C)H-TOCSY and 3D ^{15}N -[$^1\text{H}, ^1\text{H}$]-TOCSY experiments were 20 ms and 60 ms. All NOESY spectra were recorded with a mixing time of 50 ms. The ^{15}N [^1H] heteronuclear NOEs were recorded in an interleaved fashion.⁶⁰ Assignment of aromatic residues was performed using a 2D CBHD experiment, a 2D ct-[$^{15}\text{C}, ^1\text{H}$]-HSQC and a 3D NOESY-[$^{13}\text{C}, ^1\text{H}$]-HSQC, all optimized to observe resonances of aromatic residues.⁵⁹ The assignment completeness of backbone, all protons and all nitrogen plus carbon atoms are 98%, 92% and 89%, respectively. The first two residues of the N-terminal donor strand were completely missing, apparently due to their high flexibility.

The structure calculation was performed with the software packages ATNOS/CANDID^{61,62} and DYANA,⁶³ using the amino acid sequence, the chemical shift list and the three aforementioned 3D NOESY spectra as input (the disulfide bond between C21 and C61 was assumed to be formed). The standard protocol with seven cycles of peak picking using ATNOS,⁶¹ NOE assignment using CANDID⁶² and structure calculation using DYANA⁶³ was

applied (Table 1). The 20 conformers with the lowest target function values, out of 100 conformers from the seventh cycle of ATNOS/CANDID/DYANA, were energy refined in a water shell using the program OPAL⁶⁴ with the AMBER force field⁶⁵ (Table 1). The program MOLMOL³⁷ was used to analyze the protein structure and to prepare the figures showing molecular models.

Calculation of CO, ACO and Q_d parameters

The absolute contact order ACO⁶⁶ is calculated via arithmetical mean of the distance in primary sequence between two amino acid residues *i* and *j* (where *i* > *j*) making contact in the native structure, multiplied by *n_{ij}*, the number of atomic contacts between these residues:

$$\text{ACO} = \frac{1}{N} \sum_{i=2}^L \sum_{j=1}^{i-1} (i-j)n_{ij}$$

where *N* is the total number of atomic contacts and *L* is the number of residues in the polypeptide chain. An atomic contact was defined as two non-hydrogen atoms of different residues within 6 Å in space in the native structure.

The relative contact order CO⁴⁹ is given by:

$$\text{CO} = \frac{\text{ACO}}{L}$$

The number of sequence-distant native pairs, Q_d,⁴⁷ is defined as:

$$Q_d = \sum_{i=2}^L \sum_{j=1}^{i-1} \Delta_{ij}$$

where *i* and *j* are the residue numbers of two residues for which the C^α–C^α distance in space in the native structure is within 6 Å. Δ_{ij} = 1 if *i* – *j* > 12; otherwise, Δ_{ij} = 0.

The MOLMOL³⁷ software was used to calculate the aforementioned parameters.

Accession numbers

The chemical shift list has been deposited with the BioMagResBank[‡] with the accession number 15423.

The coordinates of the ensemble of 20 conformers of FimAa have been deposited in the Protein Data Bank[§] with the entry code 2JTY.

Supplementary materials related to this article can be found online at [doi:10.1016/j.jmb.2011.07.044](https://doi.org/10.1016/j.jmb.2011.07.044)

Acknowledgements

This work was supported by the Swiss National Science Foundation [grants 3100A0-100787 (R.G.) and 3100A0-113730 (G.W.)] and the Swiss Federal

Institute of Technology Zurich within the framework of the National Center for Competence in Research Structural Biology Program.

References

- Zavialov, A., Zav'yalova, G., Korpela, T. & Zav'yalov, V. (2007). FGL chaperone-assembled fimbrial polyadhesins: anti-immune armament of Gram-negative bacterial pathogens. *FEMS Microbiol. Rev.* **31**, 478–514.
- Waksman, G. & Hultgren, S. J. (2009). Structural biology of the chaperone–usher pathway of pilus biogenesis. *Nat. Rev., Microbiol.* **7**, 765–774.
- Jones, C. H., Pinkner, J. S., Roth, R., Heuser, J., Nicholes, A. V., Abraham, S. N. & Hultgren, S. J. (1995). FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the *Enterobacteriaceae*. *Proc. Natl Acad. Sci. USA*, **92**, 2081–2085.
- Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S. J. & Knight, S. D. (1999). X-ray structure of the FimC–FimH chaperone–adhesin complex from uropathogenic *Escherichia coli*. *Science*, **285**, 1061–1066.
- Zhou, G., Mo, W. J., Sebbel, P., Min, G., Neubert, T. A., Glockshuber, R. *et al.* (2001). Uroplakin Ia is the urothelial receptor for uropathogenic *Escherichia coli*: evidence from *in vitro* FimH binding. *J. Cell Sci.* **114**, 4095–4103.
- Vetsch, M., Sebbel, P. & Glockshuber, R. (2002). Chaperone-independent folding of type 1 pilus domains. *J. Mol. Biol.* **322**, 827–840.
- Eshdat, Y., Silverblatt, F. J. & Sharon, N. (1981). Dissociation and reassembly of *Escherichia coli* type 1 pili. *J. Bacteriol.* **148**, 308–314.
- McMichael, J. C. & Ou, J. T. (1979). Structure of common pili from *Escherichia coli*. *J. Bacteriol.* **138**, 969–975.
- Andersson, M., Fallman, E., Uhlin, B. E. & Axner, O. (2006). A sticky chain model of the elongation and unfolding of *Escherichia coli* P pili under stress. *Biophys. J.* **90**, 1521–1534.
- Fallman, E., Schedin, S., Jass, J., Uhlin, B. E. & Axner, O. (2005). The unfolding of the P pili quaternary structure by stretching is reversible, not plastic. *EMBO Rep.* **6**, 52–56.
- Jass, J., Schedin, S., Fallman, E., Ohlsson, J., Nilsson, U. J., Uhlin, B. E. & Axner, O. (2004). Physical properties of *Escherichia coli* P pili measured by optical tweezers. *Biophys. J.* **87**, 4271–4283.
- Bennett, M. J., Choe, S. & Eisenberg, D. (1994). Domain swapping: entangling alliances between proteins. *Proc. Natl Acad. Sci. USA*, **91**, 3127–3131.
- Anderson, K. L., Billington, J., Pettigrew, D., Cota, E., Simpson, P., Roversi, P. *et al.* (2004). An atomic resolution model for assembly, architecture, and function of the Dr adhesins. *Mol. Cell*, **15**, 647–657.
- Gossert, A. D., Bettendorff, P., Puorger, C., Vetsch, M., Herrmann, T., Glockshuber, R. & Wuthrich, K. (2008). NMR structure of the *Escherichia coli* type 1 pilus subunit FimF and its interactions with other pilus subunits. *J. Mol. Biol.* **375**, 752–763.
- Puorger, C., Eidam, O., Capitani, G., Erilov, D., Grutter, M. G. & Glockshuber, R. (2008). Infinite

‡ www.bmr.b.wisc.edu

§ www.pdb.org

- kinetic stability against dissociation of supramolecular protein complexes through donor strand complementation. *Structure*, **16**, 631–642.
16. Sauer, F. G., Futterer, K., Pinkner, J. S., Dodson, K. W., Hultgren, S. J. & Waksman, G. (1999). Structural basis of chaperone function and pilus biogenesis. *Science*, **285**, 1058–1061.
 17. Sauer, F. G., Pinkner, J. S., Waksman, G. & Hultgren, S. J. (2002). Chaperone priming of pilus subunits facilitates a topological transition that drives fiber formation. *Cell*, **111**, 543–551.
 18. Verger, D., Bullitt, E., Hultgren, S. J. & Waksman, G. (2007). Crystal structure of the P pilus rod subunit PapA. *PLoS Pathog.* **3**, e73.
 19. Zavialov, A. V., Berglund, J., Pudney, A. F., Fooks, L. J., Ibrahim, T. M., MacIntyre, S. & Knight, S. D. (2003). Structure and biogenesis of the capsular F1 antigen from *Yersinia pestis*: preserved folding energy drives fiber formation. *Cell*, **113**, 587–596.
 20. Zavialov, A. V., Tischenko, V. M., Fooks, L. J., Brandsdal, B. O., Aqvist, J., Zav'yalov, V. P. *et al.* (2005). Resolving the energy paradox of chaperone/usher-mediated fibre assembly. *Biochem. J.* **389**, 685–694.
 21. Bann, J. G., Pinkner, J. S., Frieden, C. & Hultgren, S. J. (2004). Catalysis of protein folding by chaperones in pathogenic bacteria. *Proc. Natl Acad. Sci. USA*, **101**, 17389–17393.
 22. Vetsch, M., Puorger, C., Spirig, T., Grauschopf, U., Weber-Ban, E. U. & Glockshuber, R. (2004). Pilus chaperones represent a new type of protein-folding catalyst. *Nature*, **431**, 329–333.
 23. Nishiyama, M., Vetsch, M., Puorger, C., Jelesarov, I. & Glockshuber, R. (2003). Identification and characterization of the chaperone-subunit complex-binding domain from the type 1 pilus assembly platform FimD. *J. Mol. Biol.* **330**, 513–525.
 24. Thanassi, D. G., Stathopoulos, C., Dodson, K., Geiger, D. & Hultgren, S. J. (2002). Bacterial outer membrane ushers contain distinct targeting and assembly domains for pilus biogenesis. *J. Bacteriol.* **184**, 6260–6269.
 25. Nishiyama, M., Horst, R., Eidam, O., Herrmann, T., Ignatov, O., Vetsch, M. *et al.* (2005). Structural basis of chaperone-subunit complex recognition by the type 1 pilus assembly platform FimD. *EMBO J.* **24**, 2075–2086.
 26. Huang, Y., Smith, B. S., Chen, L. X., Baxter, R. H. & Deisenhofer, J. (2009). Insights into pilus assembly and secretion from the structure and functional characterization of usher PapC. *Proc. Natl Acad. Sci. USA*, **106**, 7403–7407.
 27. Nishiyama, M., Ishikawa, T., Rechsteiner, H. & Glockshuber, R. (2008). Reconstitution of pilus assembly reveals a bacterial outer membrane catalyst. *Science*, **320**, 376–379.
 28. Phan, G., Remaut, H., Wang, T., Allen, W. J., Pirker, K. F., Lebedev, A. *et al.* (2011). Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate. *Nature*, **474**, 49–53.
 29. Remaut, H., Rose, R. J., Hannan, T. J., Hultgren, S. J., Radford, S. E., Ashcroft, A. E. & Waksman, G. (2006). Donor-strand exchange in chaperone-assisted pilus assembly proceeds through a concerted beta strand displacement mechanism. *Mol. Cell*, **22**, 831–842.
 30. Vetsch, M., Erilov, D., Moliere, N., Nishiyama, M., Ignatov, O. & Glockshuber, R. (2006). Mechanism of fibre assembly through the chaperone-usher pathway. *EMBO Rep.* **7**, 734–738.
 31. Pace, C. N. G., Grimsley, G. R. & Scholtz, J. M. (2005). *Protein Folding Handbook*. WILEY-VCH, Weinheim, Germany.
 32. Myers, J. K., Pace, C. N. & Scholtz, J. M. (1995). Denaturant *m* values and heat capacity changes: relation to changes in accessible surface-areas of protein unfolding. *Protein Sci.* **4**, 2138–2148.
 33. Erilov, D., Puorger, C. & Glockshuber, R. (2007). Quantitative analysis of nonequilibrium, denaturant-dependent protein folding transitions. *J. Am. Chem. Soc.* **129**, 8938–8939.
 34. Piatek, R., Bruzdziak, P., Zalewska-Piatek, B., Kur, J. & Stangret, J. (2009). Preclusion of irreversible destruction of Dr adhesin structures by a high activation barrier for the unfolding stage of the fimbrial DraE subunit. *Biochemistry*, **48**, 11807–11816.
 35. Van Molle, I., Moonens, K., Garcia-Pino, A., Buts, L., De Kerpel, M., Wyns, L. *et al.* (2009). Structural and thermodynamic characterization of pre- and postpolymerization states in the F4 fimbrial subunit FaeG. *J. Mol. Biol.* **394**, 957–967.
 36. Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
 37. Koradi, R., Billeter, M. & Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics*, **14**, 51–55, 29–32.
 38. Bullitt, E., Jones, C. H., Striker, R., Soto, G., Jacob-Dubuisson, F., Pinkner, J. *et al.* (1996). Development of pilus organelle subassemblies *in vitro* depends on chaperone uncapping of a beta zipper. *Proc. Natl Acad. Sci. USA*, **93**, 12890–12895.
 39. Hamill, S. J., Meekhof, A. E. & Clarke, J. (1998). The effect of boundary selection on the stability and folding of the third fibronectin type III domain from human tenascin. *Biochemistry*, **37**, 8071–8079.
 40. Hamill, S. J., Steward, A. & Clarke, J. (2000). The folding of an immunoglobulin-like Greek key protein is defined by a common-core nucleus and regions constrained by topology. *J. Mol. Biol.* **297**, 165–178.
 41. Balbach, J. S. & Schmid, F. X. (2000). Proline isomerization as a rate limiting step. In *Protein Folding: Frontiers in Molecular Biology* (Pain, R., ed.), Oxford University Press, Oxford, UK.
 42. Galzitskaya, O. V., Garbuzynskiy, S. O., Ivankov, D. N. & Finkelstein, A. V. (2003). Chain length is the main determinant of the folding rate for proteins with three-state folding kinetics. *Proteins*, **51**, 162–166.
 43. Horng, J. C., Tracz, S. M., Lumb, K. J. & Raleigh, D. P. (2005). Slow folding of a three-helix protein via a compact intermediate. *Biochemistry*, **44**, 627–634.
 44. Huang, J. T., Cheng, J. P. & Chen, H. (2007). Secondary structure length as a determinant of folding rate of proteins with two- and three-state kinetics. *Proteins*, **67**, 12–17.
 45. Ivankov, D. N., Garbuzynskiy, S. O., Alm, E., Plaxco, K. W., Baker, D. & Finkelstein, A. V. (2003). Contact order revisited: influence of protein size on the folding rate. *Protein Sci.* **12**, 2057–2062.

46. Jackson, S. E. (1998). How do small single-domain proteins fold? *Folding Des.* **3**, R81–R91.
47. Kamagata, K., Arai, M. & Kuwajima, K. (2004). Unification of the folding mechanisms of non-two-state and two-state proteins. *J. Mol. Biol.* **339**, 951–965.
48. Weikl, T. R. & Dill, K. A. (2003). Folding rates and low-entropy-loss routes of two-state proteins. *J. Mol. Biol.* **329**, 585–598.
49. Plaxco, K. W., Simons, K. T. & Baker, D. (1998). Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* **277**, 985–994.
50. Plaxco, K. W., Simons, K. T., Ruczinski, I. & Baker, D. (2000). Topology, stability, sequence, and length: defining the determinants of two-state protein folding kinetics. *Biochemistry*, **39**, 11177–11183.
51. Makarov, D. E. & Plaxco, K. W. (2003). The topomer search model: a simple, quantitative theory of two-state protein folding kinetics. *Protein Sci.* **12**, 17–26.
52. Hahn, E., Wild, P., Hermanns, U., Sebbel, P., Glockshuber, R., Haner, M. *et al.* (2002). Exploring the 3D molecular architecture of *Escherichia coli* type 1 pili. *J. Mol. Biol.* **323**, 845–857.
53. Rudolph, R. & Lilie, H. (1996). *In vitro* folding of inclusion body proteins. *FASEB J.* **10**, 49–56.
54. Ellman, G. L. (1959). Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* **82**, 70–77.
55. Hermanns, U., Sebbel, P., Eggli, V. & Glockshuber, R. (2000). Characterization of FimC, a periplasmic assembly factor for biogenesis of type 1 pili in *Escherichia coli*. *Biochemistry*, **39**, 11564–11570.
56. Wider, G. & Dreier, L. (2006). Measuring protein concentrations by NMR spectroscopy. *J. Am. Chem. Soc.* **128**, 2571–2576.
57. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280.
58. Santoro, M. M. & Bolen, D. W. (1988). Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry*, **27**, 8063–8068.
59. Cavanagh, J., Fairbrother, W. J., Palmer, A. G. I., Rance, M. & Skelton, N. J. (2007). *Protein NMR Spectroscopy: Principles and Practice*, 2nd edit. Elsevier Academic Press, Amsterdam, Netherlands.
60. Renner, C., Schleicher, M., Moroder, L. & Holak, T. A. (2002). Practical aspects of the 2D ¹⁵N-[¹H]-NOE experiment. *J. Biomol. NMR*, **23**, 23–33.
61. Herrmann, T., Güntert, P. & Wüthrich, K. (2002). Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J. Biomol. NMR*, **24**, 171–189.
62. Herrmann, T., Güntert, P. & Wüthrich, K. (2002). Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* **319**, 209–227.
63. Güntert, P., Mumenthaler, C. & Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283–298.
64. Luginbühl, P., Güntert, P., Billeter, M. & Wüthrich, K. (1996). The new program OPAL for molecular dynamics simulations and energy refinements of biological macromolecules. *J. Biomol. NMR*, **8**, 136–146.
65. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M. *et al.* (1995). A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. *J. Am. Chem. Soc.* **117**, 5179–5197.
66. Grantcharova, V., Alm, E. J., Baker, D. & Horwich, A. L. (2001). Mechanisms of protein folding. *Curr. Opin. Struct. Biol.* **11**, 70–82.