Cell, Vol. 113, 587–596, May 30, 2003, Copyright ©2003 by Cell Press

Structure and Biogenesis of the Capsular F1 Antigen from *Yersinia pestis*: Preserved Folding Energy Drives Fiber Formation

Anton V. Zavialov,[1] Jenny Berglund,[1] Alexander F. Pudney,[2] Laura J. Fooks,[2] Tara M. Ibrahim,[2] Sheila MacIntyre,[2,*] and Stefan D. Knight[1,*]

[1]Department of Molecular Biosciences/Structural Biology
Uppsala Biomedical Center
Swedish University of Agricultural Sciences
Box 590
SE-751 24 Uppsala
Sweden

[2]Microbiology Division
School of Animal and Microbial Sciences
University of Reading
Reading RG6 6AJ
United Kingdom

Summary

Most gram-negative pathogens express fibrous adhesive virulence organelles that mediate targeting to the sites of infection. The F1 capsular antigen from the plague pathogen *Yersinia pestis* consists of linear fibers of a single subunit (Caf1) and serves as a prototype for nonpilus organelles assembled via the chaperone/usher pathway. Genetic data together with high-resolution X-ray structures corresponding to snapshots of the assembly process reveal the structural basis of fiber formation. Comparison of chaperone bound Caf1 subunit with the subunit in the fiber reveals a novel type of conformational change involving the entire hydrophobic core of the protein. The observed conformational change suggests that the chaperone traps a high-energy folding intermediate of Caf1. A model is proposed in which release of the subunit allows folding to be completed, driving fiber formation.

Introduction

Most gram-negative pathogens express fibrous adhesive virulence organelles that mediate targeting to the sites of infection. A periplasmic chaperone/usher machinery is used for the controlled assembly of many such organelles (Thanassi et al., 1998). The periplasmic chaperones bind subunits as they emerge in the periplasm via the Sec pathway, assist in their folding, and partition them away from nonproductive aggregation pathways by capping their assembly surfaces. The chaperones deliver subunits to the outer membrane usher where they are released from the chaperone and assembled into linear fibers that are secreted through the oligomeric usher pore to the cell surface.

Two closely related but different classes of periplasmic chaperones with distinct sequence signatures are used for the assembly of rigid pili (FGS chaperones) and nonpilus organelles (FGL chaperones), respectively (Hung et al., 1996). Structural studies on FGS chaperones have shown that the chaperones consist of two immunoglobulin (Ig)-like domains joined at approximately right angles with a large cleft between the two domains (Knight et al., 2000). The F_1 and G_1 β strands in the first domain are connected by a long and flexible loop that protrudes like a handle from the body of the domain. The beginning of the G_1 β strand harbors a conserved motif of hydrophobic residues that is solvent exposed in the free chaperones and that is critical for subunit binding. FGS chaperones have a relatively short F_1–G_1 loop with three hydrophobic residues in the G_1 β-motif, FGL chaperones have a longer F_1–G_1 loop with five hydrophobic residues in the motif (MacIntyre et al., 2001). FGL chaperones are further characterized by longer N-terminal sequences and by a pair of invariant cysteine residues in the F_1 and G_1 β strands (Chapman et al., 1999; Hung et al., 1996; MacIntyre et al., 2001).

The crystal structures for the type 1 pilus FimC:FimH and the P pilus PapD:PapK complexes (Choudhury et al., 1999; Sauer et al., 1999) revealed that pilus subunits (pilins), like the chaperones, have Ig-like folds. However, the final (G) β strand of the fold is missing, creating a deep hydrophobic cleft on the surface of the subunit. The chaperones bind pilins by inserting their G_1 β strand into this cleft in a process called donor strand complementation (DSC). The hydrophobic side chains in the conserved G_1 motif are inserted into the hydrophobic acceptor cleft and become an integral part of the subunit hydrophobic core. The N terminus of pilus subunits, which is disordered in chaperone:subunit complexes, harbors a conserved β strand motif similar to the G_1 motif. Assembly of subunits is thought to proceed by a donor strand exchange (DSE) mechanism in which the chaperone G_1 donor strand is replaced by the N terminus of a second subunit, thereby joining subunits into a fiber. In contrast to the many ATP-dependent cellular chaperones (Saibil, 2000) no input of external energy is required for subunit release, and organelle assembly is independent of cellular energy (Jacob-Dubuisson et al., 1994). No structures showing how subunits interact in a fiber are available, and the molecular details of DSE as well as the energetics of the process remain to be determined.

The final architecture and morphology of the fibrous organelle will depend on subsequent subunit:subunit interactions that determine the coiling of secreted fibers into different structures such as rigid pili, thin fibrillae, or capsules. Rigid pili (e.g., type 1 and P pili) are rod-like appendages constructed from several different subunits. A specialized receptor binding protein (adhesin) is often incorporated into a more flexible structure at the tip of pili. The nonpilus organelles (e.g., F1 capsular antigen and Dr adhesins) are composed of only one or sometimes two types of subunits and do not contain a specialized adhesin. Nonetheless, most of them display adhesive properties important for cell binding and/or invasion. No structural information is available for any nonpilus chaperone or subunit, and the structural and

*Correspondence: stefan.knight@molbio.slu.se (S.D.K.); s.macintyre@reading.ac.uk (S.M.)

functional significance of the differences between FGS and FGL chaperones, as well as the structural organization of nonpilus adhesive organelles, is unclear.

The F1 capsular antigen from the plague pathogen *Yersinia pestis* consists of linear fibers of a single subunit (Caf1) and serves as a prototype for nonpilus organelles assembled via the FGL chaperone/usher pathway (Chapman et al., 1999; MacIntyre et al., 2001; Zavialov et al., 2002). Assembly of F1 is mediated by the Caf1M chaperone and the Caf1A usher. The capsule facilitates escape of the bacterium from phagocytosis (Du et al., 2002). It is nevertheless highly immunoprotective and together with V antigen is the basis for development of newer generation plague vaccines (Titball and Williamson, 2001). Plague, also known as the Black Death, is one of the most dangerous diseases known to mankind. Infection via the respiratory tract results in pneumonic plague and certain death within a few days unless rapidly treated with antibiotics. The increase in number of annual cases (Keeling and Gilligan, 2000), the appearance of multiple drug resistant strains (Galimand et al., 1997), and the potential use of *Y. pestis* as a bioweapon indicate that efforts to develop effective plague vaccines, diagnostics for rapid detection, and novel antibiotics are vital. Knowledge about the structure and assembly of F1 antigen will aid such efforts.

Here, we present two X-ray structures corresponding to snapshots of chaperone-assisted F1 fiber assembly. The structure for the Caf1M:Caf1 preassembly complex reveals the structural basis for the conserved differences between FGS and FGL chaperones. A Caf1M:Caf1:Caf1 complex corresponding to the smallest possible F1 fiber together with genetic data reveals how subunits interact in a fiber assembled via the chaperone/usher pathway and provides direct evidence for DSE assembly. Comparison of chaperone bound Caf1 subunit with the subunit in the fiber suggests that the chaperone traps a folding intermediate of Caf1, where the formation of a condensed hydrophobic core is not yet fully completed. A model is proposed in which release of the subunit allows folding to be completed, driving assembly.

Results and Discussion

Structure of Caf1M:Caf1 Preassembly Complex

To gain insight into the structure and assembly of nonpilus organelles, we decided to determine the X-ray structure for the Caf1M:Caf1 preassembly complex. Overexpression of Caf1M and Caf1 in the absence of Caf1A usher results in periplasmic accumulation of preassembly complex and also in formation of short periplasmic F1 fibers capped at one end by a single Caf1M molecule (Zavialov et al., 2002). Deletion mutations in the putative N-terminal donor strand of Caf1 blocked assembly. To simultaneously block assembly and facilitate purification, the Caf1 N-terminal donor sequence was replaced by a six-His tag. The resulting binary complex was purified and crystallized (Zavialov et al., 2003), and solved to 1.7 Å resolution using a combination of MAD and MIR techniques (Table 1).

The Caf1M chaperone has the typical fold of periplasmic chaperones with two domains oriented at ~90° angle (Figure 1A). Both domains are 7-stranded β sandwiches with Ig-like topology. The longer N-terminal and F_1–G_1 FGL sequences correspond to A_1 and G_1 edge strands that are significantly longer than in FGS chaperones. Only limited electron density is observed for the F_1–G_1 loop (residues 104–123), suggesting that it is flexible and disordered in the crystal. A disulfide bond between Cys98 and Cys137 that is important for folding of Caf1M (Zav'yalov et al., 1997) links the F_1 strand to the G_1 donor strand (Figure 1A). The two cysteines are conserved in FGL chaperones and are likely to form a disulfide bond in all of them.

Despite the lack of significant sequence similarity, Caf1 has the same incomplete Ig-like fold as pilus subunits (Figure 1A). The absence of a 7th (G) strand results in a 6-stranded β sandwich where the hydrophobic core of the subunit is partially exposed in a long and deep hydrophobic cleft (Figure 1B). Caf1 interacts mainly with domain 1 in Caf1M. The two proteins bind via edge strands in Caf1 and in the 1st domain of Caf1M to form a closed barrel with a common core (Figures 1A and 1C). A similar but less pronounced "super-barrel" is also formed in subunit complexes of the FGS chaperones FimC and PapD (data not shown) and is likely to be a general feature of chaperone:subunit complexes. Strand G_1 in Caf1M is hydrogen bonded to strand F in Caf1, the chaperone A_1 strand is hydrogen bonded to the subunit A strand. As in FGS chaperone:subunit complexes, hydrophobic residues from the chaperone G_1 strand are donated to the subunit to compensate for the missing G strand (Figures 1A and 1B). The longer G_1 donor strand in Caf1M contributes five rather than three hydrophobic side chains as in the FGS complexes, and the acceptor cleft of Caf1 is correspondingly much longer than in the pilus subunits. The longer A_1 strand in Caf1M also interacts more extensively with the subunit than the A_1 strand in FGS chaperone:subunit complexes. As in pilus chaperone:subunit complexes, the subunit C terminus is anchored in the cleft of the chaperone by hydrogen bonds to a pair of invariant positively charged residues (Arg20 and Lys139).

The G_1 donor strand starts at Val126. This residue is positioned between the short A′ strand and the F strand of Caf1 with main chain hydrogen bonds to Ile29′ and Gly136′ (Caf1 residues are primed to distinguish them from Caf1M residues). The G_1 donor strand runs parallel with and is hydrogen bonded to the subunit F strand almost all the way to the subunit C terminus. Val126, Val128, Val130, and Phe132 in the G_1 donor strand point into the acceptor cleft and are inserted deep into the hydrophobic core of Caf1 (Figure 1B). The importance of Val126 to subunit binding has been highlighted by mutagenesis (MacIntyre et al., 2001). Due to the twist of β strands, Ile134 is oriented more perpendicular to the acceptor cleft and is part of an additional layer of hydrophobic residues above the cleft. Tyr12, Val14, and Ile16 in the chaperone A_1 strand also contribute to this layer which is continuous with the hydrophobic core of the chaperone N-terminal domain (Figures 1B and 1C).

The Ig-like fold of pilins consists of two sheets packed against each other in a β sandwich. In Caf1, sheet 1 comprises strands ABED and sheet 2 comprises strands FCC′ (Figures 1A and 1D). The longer acceptor cleft in Caf1 compared to FGS subunits is the result mainly of differences in the path of strand A. In all the available

Table 1. Data Collection and Refinement

Data Collection Ternary Complex	
Wavelength (Å)	0.933
No. unique reflections	34,536
Multiplicity	5.4 (3.4)
Completeness	88.6 (82.0)
R_{merge}	0.086 (0.24)
I/σ	5.4 (3.0)

Refinement	Binary complex	Ternary complex
Data set	Native data	Native data
Resolution (Å)	20–1.8	20–2.0
Number of reflections		
total	34,097	34,445
work set	32,386	32,692
test set	1,711	1,753
R_{work}	0.194	0.169
R_{free}	0.241	0.225
Number of atoms		
protein	2,733	3,686
solvent	218	316
Wilson B-factor (Å^2)	20.3	23.9
Average B-factor (Å^2)		
main chain	19.5	29.7
side chain	22.2	31.5
solvent	33.8	41.0
Rmsd stereochemistry		
bonds (Å)	0.027	0.025
angles (°)	2.2	2.2
Rmsd B-factors (Å^2)		
main chain	1.6	1.7
side chain	5.7	3.8

Values for high resolution shells in parenthesis. $R_{merge} = \Sigma_h\Sigma_i|I(h)_i - (I)|/\Sigma_h\Sigma_i I(h)_i$, where $I(h)$ is the intensity of a reflection h, Σ_h is the sum over all reflections and Σ_i is the sum over i measurements of reflection h. $R_{work} = \Sigma|F_O - F_C|/\Sigma F_O$ where F_O and F_C are the observed and calculated structure factors respectively. R_{free} is calculated for a test set of reflections randomly excluded from refinement. B-factors are given with contribution from TLS tensors included. Rmsd stereochemistry is the deviation from ideal values. Rmsd B-factors is deviation between bonded atoms.

structures, this strand switches from the B strand toward the chaperone F_1–G_1 loop and hydrogen bonds to the G_1 strand. The switch in strand A defines one end of the acceptor cleft; the other end is defined by the subunit C terminus and the beginning of strand A. In Caf1, the switch occurs later so that more of the hydrophobic core of the subunit is exposed for interaction with the chaperone. The longer acceptor cleft is accommodated by the longer chaperone A_1 and G_1 strands. The corresponding increase in interaction area should result in increased specificity and perhaps affinity of FGL chaperones for their subunits compared to FGS chaperones. The principle difference between organelles assembled by FGS (pilus) and FGL (nonpilus) chaperones is that whereas rigid pili are complex structures composed of up to six different subunits, nonpilus organelles are polymers of only one, or sometimes two, different subunits. In either type of system a single chaperone is responsible for assembly. The smaller donor-acceptor recognition area of FGS chaperone:subunit complexes might reflect the need to recognize and bind several different subunits for assembly into complex pilus structures.

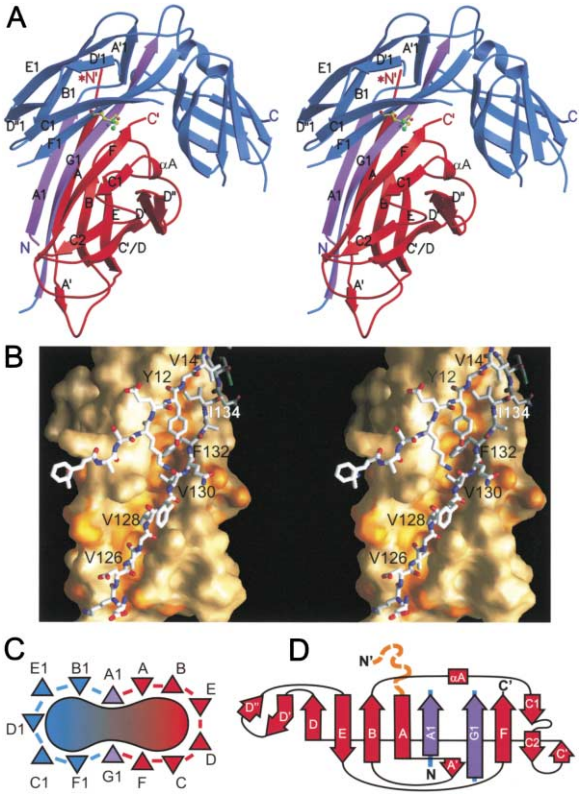

Figure 1. Caf1M:Caf1 Crystal Structure

(A) Stereo diagram of the Caf1M:Caf1 complex. Caf1M is blue with G_1 and A_1 edge strands in violet, Caf1 is red. The two conserved cysteine residues are shown as ball-and-stick. N and C termini are labeled in blue for Caf1M and in red for Caf1. The asterisk in the *N′ label indicates that the N-terminal donor sequence of Caf1 has been replaced by a His-tag.

(B) The G_1 donor strand and the A_1 strand of Caf1M (stick model) fill up and cover the acceptor cleft of Caf1 (molecular surface). Surfaces of hydrophobic side chains are orange. Caf1M side chains pointing toward the cleft are shown and labeled.

(C) Schematic illustration of the closed barrel formed by Caf1 and the N-terminal domain of Caf1M. β strands are viewed end on; triangles pointing up indicate β strands running out of the plane of the paper; triangles pointing down indicate β strands running into the plane. Lines indicate hydrogen bonding between adjacent strands. The colored area inside the triangles represents the fused hydrophobic core of the barrel. Color-coding as in (A).

(D) Caf1 topology diagram. Arrows indicate β strands, the single helix (αA) is shown as a rectangle. Caf1 is colored red (dashed orange line at the N terminus indicates that the N-terminal donor sequence is replaced by a His-tag in our construct); Caf1M G_1 and A_1 β strands are violet.

Double-Cysteine Mutagenesis Identifies Donor Strand Alignment in F1

We previously demonstrated that the N terminus of Caf1 mediates subunit DSC polymerization with residues 8–13 being the most critical (Zavialov et al., 2002). Modeling based on the structure for the Caf1M:Caf1 complex suggested that the N terminus would bind next to strand F as shown in Figure 2A. To test this hypothesis, a series of double-cysteine mutations were constructed within the subunit donor and F strands (Figure 2A). Pairs of cysteines positioned in close enough proximity should form disulfide bonds. These bonds would crosslink Caf1

subunits in the high molecular weight (HMW) Caf1M:$(Caf1)_n$ complexes that accumulate in the periplasm in the absence of the Caf1A usher (Zavialov et al., 2002). The formation of disulfide cross links can be easily assayed by SDS-PAGE of the mutants in the presence or absence of dithiothreitol (DTT). Seven mutants were prepared initially in plasmid pFM1, which encodes chaperone and subunit. All these mutants showed periplasmic accumulation of chaperone:subunit complexes. Three pairs of mutants (T8CA141C, T10CT139C, and T12CK137C) resulted in formation of DTT-sensitive HMW polymer. Caf1T10CT139C provided the optimal orientation of side chains for disulfide bond formation and formed the highest level of periplasmic HMW polymer (Figure 2B). Pairs of mutations corresponding to alignment of the donor strand two residues up (T8CT139C, T10CK137C) or down (T10CA141C, T12CT139C) with respect to the F strand, led to only low MW oligomers (Figure 2B), despite the fact that subunit accumulated in all mutants (Figure 2C). The results identify the alignment of donor strand and recipient F strand in subunit:subunit polymers as that shown in Figure 2A.

To confirm that the identified donor strand alignment was also formed on assembly of surface F1 polymer, the mutation T10CT139C was introduced into the plasmid pFMA1 (expresses chaperone, subunit, and outer membrane usher). Caf1T10CT139C was secreted to the cell surface, although somewhat less efficiently than the wild-type (Figure 2D). In the presence of DTT, after boiling, both native and mutant F1 polymers denatured to show only monomers of Caf1. In the absence of DTT, the surface-extracted Caf1T10CT139C was primarily visible as HMW polymer on both Coomassie blue staining (Figure 2D) and immunoblotting (Figure 2E) following SDS-PAGE. In contrast, native F1 polymer displayed only monomers after boiling. The experiment shows that subunits of the surface-extracted Caf1T10CT139C were linked by disulfide bonds. Thus, F1 capsular antigen consists of Caf1 subunits polymerized into fibers via DSC as in Figure 2A.

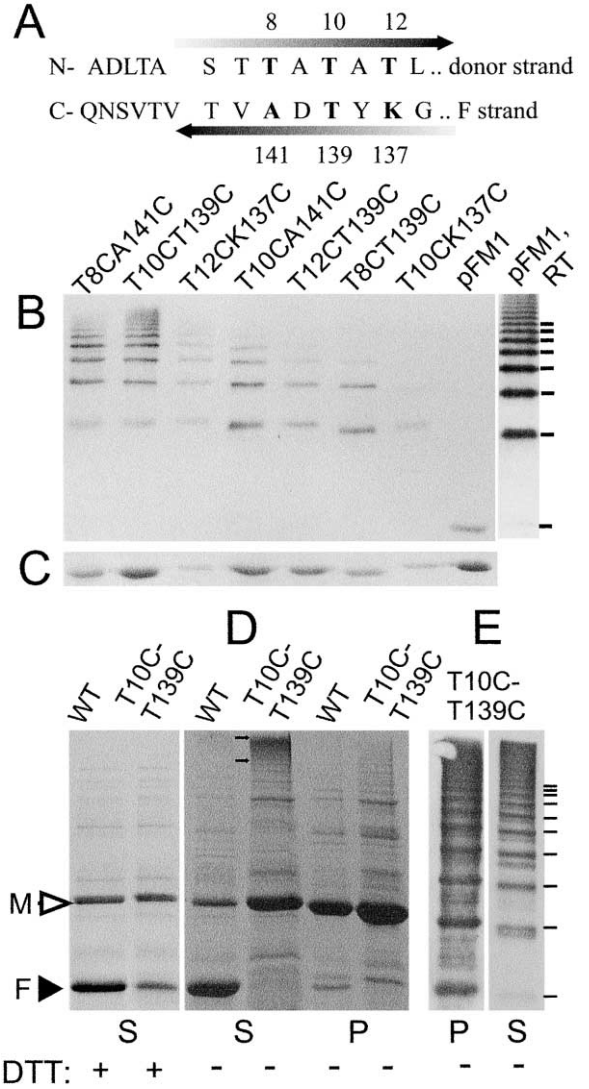

Figure 2. Disulfide Bond Formation Reveals Donor Strand Register in F1 Polymer

(A) Alignment of Caf1 donor and F strands. Residues in bold were mutated to Cys.

(B) Immunoblot of periplasmic subunit from *E. coli* expressing pFM1 mutants as indicated. 12.5% SDS-polyacrylamide gel, samples heated 100°C for 5 min in sample buffer in the absence of DTT. Periplasmic extracts from *E. coli* expressing pFM1 wt subunits before (pFM1, RT) and after boiling (pFM1) were loaded as controls. Bars from bottom to top indicate n = 1–8 subunits, larger oligomers are not clearly resolved.

(C) Denatured periplasmic subunit recovered from each mutant. Coomassie blue stained SDS-polyacrylamide gel, samples heated 100°C for 5 min in sample buffer in the presence of 0.1 M DTT.

(D) Disulfide-linked surface polymer of Caf1T10CT139C. Coomassie blue stained 14% SDS-polyacrylamide gel of F1 extracted from the surface (S) and periplasm (P) of *E. coli* DH5α/pFMA1-F1T10CT139C. Samples were heated at 100°C in the presence or absence of 0.1 M DTT to identify total denatured Caf1 and disulfide-linked polymer, respectively. Arrows encompass HMW disulfide-linked polymer of surface extracted Caf1T10CT139C. Open arrowhead, denatured Caf1M (M); filled arrowhead, denatured Caf1 monomer (F).

(E) Immunoblot of surface (S) extracted F1 and periplasmic (P) oligomers of Caf1 expressed in *E. coli* DH5α/pFMA1-F1T10CT139C. 12.5% SDS-polyacrylamide gel, samples heated at 100°C in the absence of DTT. Bars from bottom to top indicate n = 2–11 subunits.

Structure of the Minimal F1 Fiber

To better understand the fiber assembly process and the structural consequences of DSE, we wished to obtain detailed 3D structural information for the F1 fiber. Since high resolution 3D structural data cannot be obtained for fibers, we decided to crystallize the Caf1M:Caf1:Caf1 ternary complex, representing the minimal F1 fiber. Replacing Ala9 in the N-terminal donor strand of Caf1 with arginine was shown to suppress assembly of higher order complexes allowing ternary complex to be isolated (Zavialov et al., 2002). The complex was purified by ion exchange chromatography, crystallized, and solved by molecular replacement using the structure for the binary complex as a search model, and refined to 2.0 Å resolution (Table 1). To facilitate the following discussion, we have denoted the chaperone bound subunit Caf1′ and the second "fiber-subunit" Caf1″. To distinguish residues in different subunits in the complex, residue numbers are primed for Caf1′ and double-primed for Caf1″.

The structure of Caf1M and of the chaperone bound

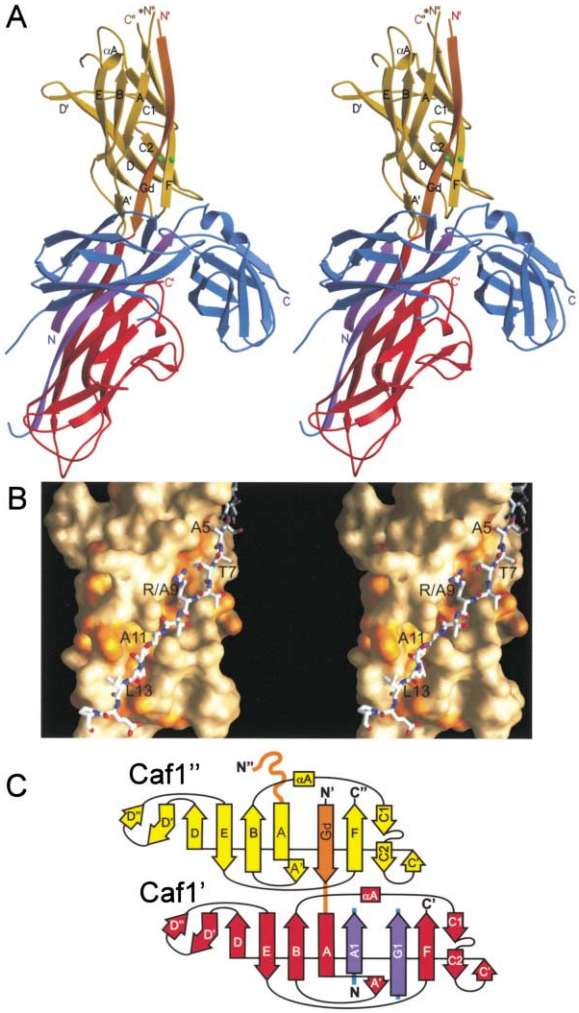

Figure 3. Caf1M:Caf1:Caf1 Crystal Structure

(A) Stereo diagram of the Caf1M:Caf1:Caf1 complex. Caf1M is in blue except for G_1 and A_1 β strands (violet). The chaperone bound Caf1′ subunit is red (N-terminal donor strand (G_d) in orange); the Caf1″ subunit corresponding to the tip of a growing fiber is in yellow. N and C termini are labeled in the same colors as the ribbons. The asterisk in the *N″ label indicates that the N-terminal sequence of Caf1″ up to this point is disordered. The two residue positions affording maximum crosslinking in our double-cysteine assay (Thr10 and Thr139; Figure 2) are indicated by green spheres.

(B) The N-terminal G_d donor strand (stick model) of Caf1′ is inserted into the acceptor cleft of Caf1″ (molecular surface). Surfaces of hydrophobic side chains are shown in orange. Residues pointing into the cleft are indicated. R/A9 indicates that wild-type Caf1 has Ala instead of Arg in position 9. The suppression of HMW polymer formation observed for the Caf1A9R mutant might be caused by local nonoptimal docking of the mutated G_d strand in the acceptor cleft resulting in somewhat weaker subunit:subunit binding. A small interface (560 Å^2) between Caf1″ and Caf1M presumably stabilizes the ternary complex relative to higher-order polymers, explaining the observed effect.

(C) Caf1′ (red)-Caf1″ (yellow) topology diagram. The disordered G_d donor strand of Caf1″ is shown as a wavy orange line, the ordered G_d donor strand of Caf1′ is labeled and colored orange; Caf1M G_1 and A_1 β strands are violet.

subunit is virtually the same as in the binary complex. Superpositioning of Caf1M:Caf1 from the two structures aligns 324 Cα positions with a rmsd of 0.49 Å. The N-terminal region of Caf1″ is disordered. In contrast, the N-terminal region of Caf1′ is ordered and forms an antiparallel donor strand interaction with the last (F) strand of Caf1″ (Figure 3). The donor strand binds to the subunit F strand as predicted by our double-cysteine assay (Figures 2A and 3A). DSE occurs with a change in direction of the donated strand to produce a bona fide Ig-like topology in the fiber-subunit, as predicted (Choudhury et al., 1999; Sauer et al., 1999). A similar "topological transition" has also recently been observed for the P pilus subunit PapE bound to a peptide (Sauer et al., 2002). The N terminus of the Caf1′ donor strand is next to the Caf1″ C terminus, and there is a one-by-one exchange of subunit for chaperone donor residues in the acceptor cleft (compare Figures 1B and 3B). In contrast, the N terminus of the donor peptide bound to PapE was far from the subunit C terminus, and the peptide did not bind in the same register as the chaperone G_1 donor strand in the PapE acceptor cleft (Sauer et al., 2002). We have denoted the N-terminal donor strand "G_d" (d for donor) because in the fiber it plays the same structural role as the (C-terminal) G stand of the canonical Ig fold.

The two Caf1 subunits in the complex are significantly different with a rmsd of 1.20 Å for 130 equivalent Cα atoms. This difference is due to a conformational change resulting in a more shallow and narrow hydrophobic groove on the surface of Caf1″ compared to Caf1′ (Figures 1B and 3B). The side chain of Leu13′ occupies the space filled by Val126 in the chaperone:subunit complex and is inserted deeply into the subunit core. In contrast, the remaining G_d donor residues do not penetrate the core of the subunit to the same extent as the corresponding residues in the chaperone:subunit complex. Instead the side chains pack into rather shallow subpockets on the surface of Caf1″. This is probably a general feature of the nonpilus systems since their G_d donor residues tend to be relatively small compared to the corresponding chaperone G_1 donor residues (Figure 4; Hung et al., 1996).

Apart from G_d DSC there is only limited contact between the two Caf1 subunits in the complex. The two subunits are separated by a short linker (residues 15′–18′) between the G_d and A strands in Caf1′. The relative orientation of the two subunits is stabilized by hydrogen bonds and electrostatic interactions involving two of the linker residues (Figure 5A). The side chain of Glu15′ bridges the two subunits by a salt link to Lys52′ and a hydrogen bond to the peptide amino group of Ala135″. The side chain of Arg18′ is hydrogen bonded to the main chain carbonyl oxygen of Ile31″, and to a water molecule bound between the main chain carbonyl of Gly50′ and the side chain of Asp33″.

In order to visualize an F1 fiber, we constructed a model fiber based on our crystal structure for the ternary complex. The tilt and rotation between Caf1′ and Caf1″ in the complex is 30° and 132°, respectively. If a similar end-to-end orientation between subunits is maintained throughout an assembled chain of Caf1 subunits, a thin (~20 Å diameter) extended linear fiber would result (Figure 5B). The model fiber is wound into an open right-handed helical structure with a rise per subunit of ~47 Å (corresponding roughly to the length of the subunit) and

gi	Subunit	Organelle	N-terminal sequence
115437	Caf1	F1 capsular antigen	ADLTASTTATATLV
462674	MyfA	Myf fibrillae	EPTVINSKDISATKTVK
2494479	PsaA	pH6 antigen	STVINSKDVSGEVTVK
1168383	AggA	AAF-I	ALERPPIKATETIR
2576413	AafA	AAF-III	ATAVAKTATSTIT
2506404	AfaE-1	AFA-I	NFTSSGTNGKVDLT
2196861	AfaE-2	AFA-II	AVDKHATGTYTLN
2494477	AfaE-3	AFA-III	GFTPSGTTGTTKLT
120409	DraE	Dr hemagglutinin	GFTPSGTTGTTKLT
2196863	AfaE-5	AFA-V	AFTGSGSTGTTKLT
1843574	DrbE-122	Drb122	AFTASGNTGTTKLT
1843572	DrbE-121	Drb121	TFQPSGTTGGVELT
120414	DaaE	F1845 fimbriae	TFQASGTTGITTLT
4836791	CseA	CS22 fimbriae	ATVVGDVATVQAPVVFS
120403	CS-3	CS-3 fimbriae	AAGPTLTKELALNVLS
1842443	NfaE-111	NFA	NDNVLNGVGGADGIRLGTATASGTIT
2290538	Dra2E	Dr-II	DANGLNTVNAGDGKNLGTATATATIT
1171697	NfaA	NFA-I	DANGLNTVNAGDGKNLGTATATIT
462119	SefA	SEF-14 fimbriae	AGFVGNKAVVQAAVTIA
	Caf1M G_1 donor sequence		138ICNNIAFQVFVGVD125

Figure 4. Alignment of N-Terminal Amino Acid Sequences of FGL Subunits with Caf1 G_d Donor Strand

Residues corresponding to G_d donor residues in Caf1 are shaded. Most of the positions are quite variable although there is a slight predominance of small residues, in particular threonine and serine. Two positions (dark shaded) appear more restricted. The last donor residue most frequently has a large hydrophobic side chain (leucine or isoleucine), and the residue two donor positions upstream of this is always a small residue (alanine or glycine). For comparison, the Caf1M G_1 donor sequence is shown under the G_d sequences. gi, GenBank sequence accession number.

~2.7 subunits/turn, giving it a wavy appearance in projection.

Assembly of P and of type 1 pili is initiated by binding of the binary chaperone:adhesin complex to an empty usher in the outer membrane (Thanassi et al., 1998). The adhesins are specialized carbohydrate binding subunits that consist of an N-terminal lectin domain fused to a pilin domain (Knight et al., 2000; Thanassi et al., 1998). Given the conservation of periplasmic chaperones and outer membrane ushers, an intriguing question is then how assembly is initiated in FGL systems where there is no specialized two-domain adhesin. Comparison of the structures for Caf1M:Caf1:Caf1 and FimC:FimH chaperone:adhesin complex shows a striking similarity in quaternary structure (Figure 5C). Since Caf1M:Caf1:Caf1 complex forms readily in the periplasm, it is tempting to speculate that this similarity reflects a case of molecular mimicry and that assembly of FGL organelles is initiated by targeting and binding of a ternary chaperone:subunit:subunit complex to the usher.

Assembly Is Driven by Folding Energy Preserved by the Chaperone

Two main events occur upon DSE: the chaperone G_1 strand is replaced by the G_d strand of a neighboring subunit, and the chaperone A_1 strand is removed without any compensating replacement (compare Figures 1D and 3C). Owing to this switch from a two-strand interaction between chaperone and subunit to a one-strand interaction between subunits, chaperone:subunit contacts involve a much larger surface area than subunit:subunit contacts. For example, in F1, ~3600 Å^2 total and ~2250 Å^2 hydrophobic surface area is buried at the Caf1M:Caf1 interface. Only ~2250 Å^2 total and ~1400 Å^2 hydrophobic surface area is buried between the two Caf1 subunits in the ternary complex. This suggests that subunits bind tighter to the chaperone than to each other, unless the fit between chaperone and subunit is much poorer than between two subunits.

To examine the geometrical fit between protein surfaces in the complexes, we calculated the shape correlation statistic (S_c) (Lawrence and Colman, 1993) for the Caf1M:Caf1 and Caf1′:Caf1″ interfaces. This statistic provides a measure of the packing of two protein surfaces. A value of $S_c = 0$ indicates no geometrical fit, whereas a value of $S_c = 1$ corresponds to two perfectly packed surfaces. Calculation of the shape correlation statistic for the Caf1M:Caf1 interface gave a value of $S_c = 0.74$ for the whole interface and of $S_c = 0.76$ for the G_1 donor strand interface alone. The S_c value for the G_d donor strand interface between two Caf1 subunits was 0.78. For both interfaces, the values are at the upper limit of S_c values observed for protein:protein interfaces (Lawrence and Colman, 1993). Both the chaperone G_1 and the subunit G_d donor strand have good fits to their partner acceptor cleft. The presence of internal cavities representing a destabilizing energetic cost (Xu et al., 1998) might not necessarily be detected by the S_c statistic. Application of a cavity-detection algorithm showed that there are no such cavities in the Caf1M:Caf1 superbarrel. The chaperone:subunit complex would thus appear to be tighter than the subunit:subunit complex, which begs the intriguing question as to how assembly can occur to produce stable fibers without input of energy?

An answer to this question is suggested by comparing the structures of Caf1 before and after DSE (Figure 6A; also compare Figures 1B and 3B). DSE is accompanied by an overall rearrangement of the subunit from an expanded conformation in the chaperone bound form to a more collapsed conformation in the fiber. The two sheets of the subunit β sandwich are rotated ~6 degrees with respect to each other bringing the two sheets closer in Caf1″ (Figure 6A). Sheet 2 of the β sandwich remains essentially unchanged but there is some distortion of sheet 1, mainly in the A strand and at the ends of the sheet. One example that illustrates how Caf1 repacks as a result of DSE is given in Figure 6B. On removal of the bulky side chain of Phe132 in the chaperone G_1 donor strand, the side chains of Leu21″, Leu46″, and Val142″ move in to fill the space. The small remaining pocket at the surface of Caf1″ is filled by the methyl group of Thr7′ in the G_d donor strand.

To further investigate the β sheet packing in the two forms of Caf1, we calculated the shape correlation statistic for the sheet 1-sheet 2 interface. The difference in packing is reflected in a much higher S_c-value for the two sheets in Caf1″ ($S_c = 0.71$) than in Caf1′ ($S_c = 0.58$). The value of S_c for Caf1″ is similar to that found for packing of β sheets in stable β-barrels (data not shown), whereas the low S_c value for chaperone bound Caf1′ reflects a poorly packed hydrophobic core. The low S_c

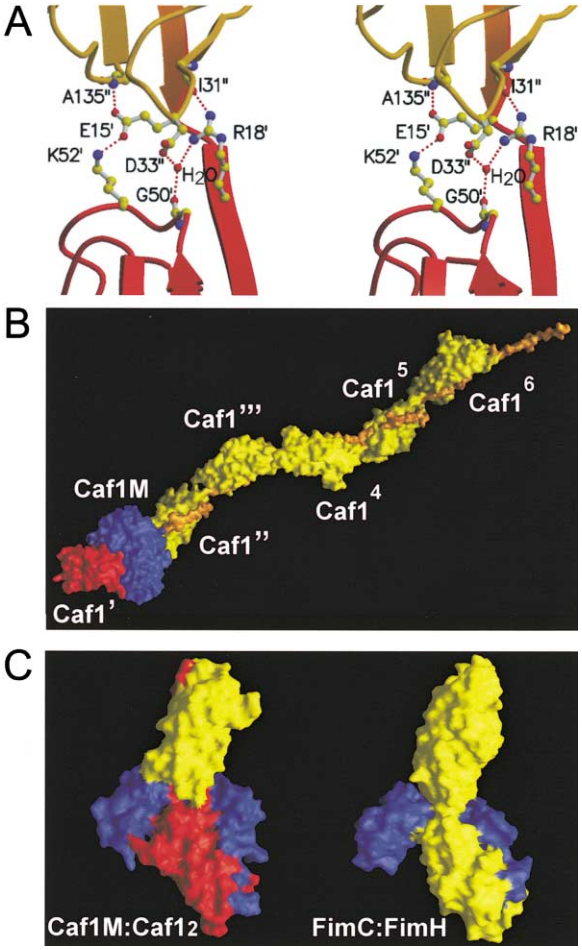

Figure 5. Assembly and Architecture of F1 Fibers

(A) A network of hydrogen and ionic bonds stabilizes the relative orientation between Caf1′ (red; G_d strand in orange) and Caf1″ (yellow) subunits in the complex. The orientation is further stabilized by additional interactions between Caf1″ and Caf1M (data not shown).

(B) Molecular surface rendering of a model for a F1 fiber generated by assuming the same orientation between successive subunits as observed for Caf1′ and Caf1″. Caf1M is blue, chaperone bound Caf1′ is red, remaining Caf1 subunits are yellow; the G_d β strand is orange.

(C) Molecular surfaces of the Caf1M:Caf1:Caf1 complex and type 1 pilus FimC:FimH chaperone:adhesin complex. Chaperones are colored blue, Caf1′ is red, Caf1″ and FimH adhesin are yellow.

value for Caf1′ arises because chaperone G_1 donor residues penetrate the core of the subunit and keep the two β sheets apart. To investigate if a similar repacking of the hydrophobic core also occurs in pilus subunits, we analyzed the structures for PapE bound to its PapD chaperone and to a donor peptide (Sauer et al., 2002). The conformational difference observed for PapE before ($S_c = 0.57$) and after ($S_c = 0.72$) release from the chaperone is due to a similar collapse of the hydrophobic core as in Caf1. In PapE the collapse is localized to the tip-most half of the subunit and is the result of entirely removing a large G_1 donor residue, rather than a one-by-one replacement of large G_1 for smaller G_d donor residues as in Caf1.

The chaperone bound conformation of pilins in some ways resembles a molten globule where secondary structure elements are formed but the hydrophobic core is not yet fully developed. We do not imply that the bound structure is a molten globule, merely that it has several of the characteristics of this state. This conformation is expected to be highly unstable in solution. The fiber form of Caf1 (and the peptide bound form of PapE) has a much more native-like structure that might be at least marginally stable in solution. This suggests that the chaperones trap subunits in a molten globule-like conformation and that folding energy might contribute to driving assembly (Figure 6C).

Vetsch et al. (2002) recently showed that the pilin domain of FimH can refold by itself in the absence of chaperone. However, folding was inefficient, with an apparent rate constant of 0.0066 s^{-1}. The observed rate appears very slow given that pilus assembly can be completed in less than five minutes in vivo (Eshdat et al., 1981). Periplasmic chaperones might catalyze subunit folding by providing a folding platform consisting of large hydrophobic side chains that nucleate hydrophobic collapse as previously suggested (Soto et al., 1998). Subunit folding onto this platform would result in chaperone donor residues being incorporated into the core of the subunit and formation of a fused β-barrel. In this way the chaperone would effectively jam the folding process and trap the subunit in an expanded high-energy conformation as observed in the structures of chaperone:subunit complexes.

Release of the chaperone bound "frozen globule" would remove the jam provided by the chaperone and allow folding to proceed to a collapsed final conformation similar to that observed for the Caf1″ subunit. The drop in free energy resulting from subunit collapse would offset the larger binding area of the chaperone:subunit complexes (Figure 6C). Vetsch et al. (2002) reported that self-folded protein rebound inefficiently to the chaperone whereas binding of unfolded protein was much more efficient. This result is expected if rebinding of the folded solution structure requires a reopening of the β sandwich. In vitro, following release and collapse, subunits might then assemble spontaneously by binding of the G_d donor strand in the shallow hydrophobic groove on the surface of the collapsed subunit.

Although in vitro assembly could in principle proceed as outlined above, in vivo assembly must involve the control and facilitation of several of the steps. The most likely site for controlling and facilitating assembly is the outer membrane usher where chaperone:subunit complexes dock prior to assembly. Since isolated subunits can be at least marginally stable (Vetsch et al., 2002), DSE might proceed stepwise with dissociation of chaperone from the subunit at the base of a growing fiber followed by binding of the donor strand from the attacking complex (Figure 6D, left image). If this is the case, the usher pore could provide a shielding environment that would limit nonspecific aggregation and promote DSE. The usher for type 1 pili, FimD, has been reported to have two binding sites for chaperone:subunit complexes (Thanassi et al., 2002). This suggests that a proximity effect may play a role in catalyzing DSE by proper alignment of the attacking donor strand with the base of a growing fiber. A further possibility for catalyzing DSE is concerted release of the chaperone

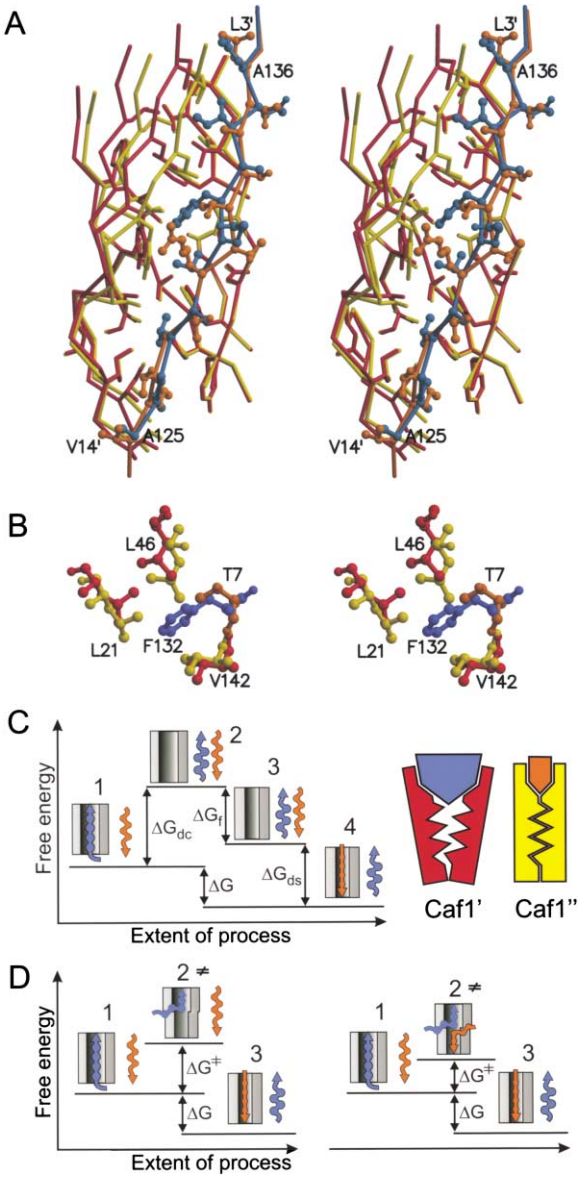

Figure 6. Assembly Is Driven by Folding Energy

(A) Superposition of Caf1′ (red) bound to the chaperone G_1 donor strand (blue) and Caf1″ (yellow) bound to the Caf1′ G_d donor strand (orange) (stereo view). To avoid clutter, only β strands in the two sheets of each subunit are shown. Core residues are shown as sticks; donor strand residues as ball-and-stick. The ends of the donor strands are labeled.

(B) Example illustrating repacking of the Caf1 hydrophobic core following DSE (stereo view). Caf1′ and Caf1″ are superpositioned as in (A). Residues shown are the two corresponding donor residues Phe132 in G_1 (blue) and Thr7′ (orange) in G_d, and Caf1 core residues Leu21, Leu46, and Val142 (red for Caf1′, yellow for Caf1″).

(C) Conceptual DSE thermodynamic energy diagram. (1) and (4) correspond to the crystallographically observed chaperone:subunit and subunit:subunit structures, respectively. The inset on the right is a schematic representation of these structures with the Caf1M:Caf1 complex on the left and the Caf1′:Caf1″ complex on the right. (2) and (3) are the (imaginary) structures that would result from rigid-body dissociation of the observed complexes. ΔG_{dc} represents the free energy difference between the chaperone:subunit complex and the free proteins in the same conformation. ΔG_f represents the free energy difference between the imaginary isolated chaperone bound G_1 donor strand and insertion of the subunit G_d donor strand in a zip-out-zip-in mechanism (Figure 6D, right image). In this mechanism, the process would avoid both of the high-energy steps of the uncatalyzed reaction (Figure 6C, left image, steps 2 and 3) by going through a series of transition states separated by low activation energy barriers.

A more complete understanding of chaperone/usher-mediated assembly awaits further experimentation to determine the kinetics and thermodynamics of the process. The data and the models presented here provide a solid framework for such work.

Concluding Remarks

Conformational changes are of fundamental importance in many if not all biological processes. Changes ranging from local side chain rearrangements to large scale loop and domain movements are well documented. Here, a different type of conformational change involving the entire hydrophobic core of a protein is described. This conformational change provides an explanation for the mechanism by which periplasmic chaperones mediate subunit folding and polymerization in the absence of cellular energy. By arresting subunit folding and trapping the subunit in a molten globule-like high-energy conformation, the chaperone preserves folding energy that can drive assembly even when chaperone:subunit interactions are more extensive than subunit:subunit interactions in the fiber. In this model, assembly is driven by the sum of the free energies for subunit collapse (ΔG_f) and subunit-subunit binding (ΔG_{ds}) (Figure 6C). The contribution of the two ΔG terms might vary between systems, but the analysis of our structures shows that the ΔG_f term cannot be ignored.

Biological polymers are formed by head (proteins and fatty acids) or tail (DNA, RNA, and polysaccharides) polymerization reactions. Head polymerization means that each monomer carries a high-energy bond that is used for addition of the next monomer. In tail polymerization, each monomer carries a high-energy bond for its own addition. Our study illustrates formation of fibers by head polymerization. The folding energy preserved in the chaperone bound subunit at the base of a growing fiber is used for incorporation of the next subunit. Is the use of preserved folding energy essential for fiber formation? Could the process be driven simply by very strong binding between subunits in the fiber? A large interactive surface between subunits would be needed to drive such a pathway. In the case of DSE assembly this would require large hydrophobic residues and/or a significantly longer N-terminal donor sequence. The

(2) and fiber bound (3) forms of the subunit. ΔG_{ds} represents the free energy difference between the unbound and bound fiber form of the subunit. ΔG corresponds to the free energy for fiber formation from chaperone:subunit assembly complex and must necessarily be <0. Values for the different free energy terms are not known, and the figure is not meant to indicate even relative sizes of these terms.

(D) Models for usher-catalyzed assembly in vivo. Left image: Stepwise DSE in which the entire G_1 donor strand is removed before the G_d strand is bound. Right image: Sequential concerted DSE in which G_1 is gradually replaced by G_d in a zip-in-zip-out mechanism.

released subunit would have to be reactivated to allow insertion of these larger side chains into the acceptor cleft. A more hydrophobic N-terminal sequence may conceivably increase unproductive aggregation. Reactivation of the subunit would need input of additional energy. Therefore, despite its apparent simplicity, such a tail mode of polymerization would not be favorable.

Adhesive organelles offer attractive targets for antibacterial vaccines. Our study shows that the structure of organelle subunits assembled via the chaperone/usher pathway changes following assembly and that each Ig-like fiber module is made from two protein subunits linked by DSC. Hence, vaccines based on isolated or chaperone bound subunits may not be appropriate, in particular when blocking antibodies are desired. The structures for Caf1 provide an invaluable tool for development of new plague vaccines.

Experimental Procedures

Plasmids and Mutagenesis

The Trc99a (Pharmacia) based plasmid, pFM1, expressing Caf1M plus Caf1 and pFMA1, expressing Caf1M, Caf1, and Caf1A, are described (Chapman et al., 1999). *E. coli* DH5α [*supE44 recA1 gyr*A *relA1* Δ*lacU169 deoR* (ϕ*80 lacZ* ΔM15)] was used as host for DNA manipulations and analysis of protein expression.

Cys mutations were created by two rounds of mutagenesis within the *caf1* gene by inverse polymerase chain reaction (PCR) using *Pfu* polymerase (Stratagene), pFM1 as template, and 16 cycles of amplification, followed by DpnI digestion as described (MacIntyre et al., 2001). Coding primers (mutations in lower case) were as follows: F1T8C-1, GATTTAACTGCAAGCACCtgtGCAACGGCAACTCTTG; F1T10C-1, CTGCAAGCACCACTGCAtgcGCAACTCTTGTTGAACC; F1T12C-1, GCACCACTGCAACGGCAtgtCTTGTTGAACCAGCC; F1K137C-1, CGGTAAACTTGCAGCAGGTtgcTACACTGATGCTG TAACCG; F1T139C-1, CTTGCAGCAGGTAAATACtgcGATGCTGTA ACCGTAACCGTATC; and F1A141C-1, GCAGCAGGTAAATACACT GATtgcGTAACCGTAACCGTATCTAACC. For each construct, two transformants were selected, *caf1* and *caf1M* DNA sequences confirmed, and intersubunit disulfide bond formation assessed. pFMA1 derivative of T10CT139C was prepared by gel purifying the 713 bp fragment from pFM1-F1T10CT139C and ligating this with PstI digested pFMA (Chapman et al., 1999). Transformants were screened for a single insert using BglI and SpeI and for the correct orientation using AatII and SpeI.

Isolation and Analyses of Periplasmic and Cell Surface F1

Exponentially growing cells were induced with 0.75 mM IPTG for 2 hr (pFM1 constructs) or 0.5 mM IPTG for 1.5 hr (pFMA1 constructs). Cells (harvested from an equivalent of 10 OD_{600} units following induction) were subjected to osmotic shock to yield 200 μl periplasmic fraction (MacIntyre et al., 2001) or were incubated in 200 μl PBS for 1 hr at 55°C with shaking to extract surface F1. Cells were removed from surface extracted F1 by centrifugation for 5 min at 12,000 $\times$ g. Periplasmic Caf1M was a contaminant in these preparations. SDS-PAGE (samples heated at 98–100°C or incubated at RT) and immunoblotting of Caf1 (1:20,000 dilution of rabbit polyclonal anti-Caf1 polymer antibody) was performed as described in MacIntyre et al. (2001).

Expression, Purification, and Crystallization of Caf1M:Caf1 and Caf1M:Caf1:Caf1 Complexes

Binary Caf1M:Caf1 and ternary Caf1M:Caf1:Caf1 complexes were purified from periplasmic extracts of *E. coli* B834 (DE3) cells transformed with pFM1-F1-6H (Zavialov et al., 2003) and pFM1-F1A9R (Zavialov et al., 2002) plasmids, respectively. Purification and crystallization of the binary complex is described in Zavialov et al. (2003). Ternary complex was purified by anion and cation exchange chromatography on Mono Q HR 10/10 and Mono S HR 5/5 columns (Pharmacia, Sweden). Anion exchange chromatography was performed in 20 mM Tris-HCl buffer, [pH 7.7]. Protein was eluted with a 0–200 mM NaCl gradient. Fractions containing ternary complex were diluted with 30 mM Na acetate buffer, [pH 4.5], and loaded onto the cation exchange column. Pure protein was obtained by elution with a 0–200 mM NaCl gradient. Before crystallization, the ternary complex was concentrated using Vivaspin 6 ml (Vivaspin, UK) to 55 mg/ml. Crystallization was performed by the hanging-drop vapor-diffusion method at 293 K. Large plate-like crystals were obtained in drops with 16%–17% PEG 8000 in 0.1 M Na cacodylate and 0.2 M Ca acetate, [pH 5.9]. The crystals belong to space group $P2_12_12$ with unit cell dimensions a = 178.8 Å, b = 69.6 Å, c = 45.3 Å, and one copy of the complex in the asymmetric unit.

Structure Determination

Data collection and initial phasing for the binary complex are reported in Zavialov et al. (2003). Diffraction data (λ = 0.933 Å) for the ternary complex were collected under liquid-nitrogen cryoconditions at 100 K on an ADSC Q4 CCD-based detector at beamline ID14-2, ESRF, France. Prior to freezing, crystals were soaked for 10–30 s in a cryoprotection solution consisting of 1 part of precipitant solution and 1 part of 25% PEG 400. Data to 2.0 Å Bragg spacing were collected, processed and reduced using MOSFILM, and scaled using SCALA (CCP4, 1994) (Table 1). The structure was solved by molecular replacement using AMORE (Navaza, 1994) with the refined binary complex as a search model.

For both structures, an initial model was constructed using ARP/wARP 5.0 (Lamzin and Wilson, 1997) and O (Jones et al., 1991). Positional, bulk solvent, and isotropic B factor refinement was performed using CNS (Brünger et al., 1998) and REFMAC5 (Murshudov et al., 1997) with TLS refinement. Two TLS groups were used for the binary complex; one consisting of the Caf1 subunit and the N-terminal domain of Caf1M, and one consisting of the C-terminal domain of Caf1M. For the ternary complex a third TLS group consisting of the second Caf1 subunit was also used. Progress of refinement and selection of refinement schemes were monitored by the R_{free} for a test set comprising 5% of the data. Refinement statistics are given in Table 1.

The binary complex model includes residues 7–49, 59–103, 124–205, and 213–233 in Caf1M, and all residues following the 6-His tag in Caf1 (residues 14–149). 99.0% of the residues are in the core region of the Ramachandran plot. The model for the ternary complex includes residues 9–54, 59–107, and 122–234 in Caf1M; all residues in Caf1′; and residues 16–107 and 111–149 in Caf1″. 97.0% of the residues are in the core region of the Ramachandran plot.

Accessible surface area calculations were performed with programs based on the Yale algorithm (Miller et al., 1987) using a probe radius of 1.4 Å. Cavity calculations were done with Voidoo (Kleywegt and Jones, 1994). Figures were generated with MOLSCRIPT (Kraulis, 1991) and GRASP (Nicholls, 1993).

Acknowledgments

We thank the staff at the ID14-2 and ID14-4 beamlines at the ESRF (Grenoble, France) and at the Maxlab 7-11 beamline in Lund, Sweden, for their kind help during data collection. We thank H. Eklund, J. Hajdu, A. Jones, and V. Zav'yalov for stimulating discussions and critical reading of the manuscript. S.D.K. thanks Mia for technical assistance. A.V.Z. was supported by a stipend from the Wenner-Gren Foundations and T.M.I. by a Nuffield Foundation Bursary. This work was supported by grants from the Swedish Research Council (to S.D.K.) and the Biotechnology and Biological Sciences Research Council/Ministry of Defense (UK) (45/B16926 to S.M.).

Received: January 16, 2003
Revised: April 22, 2003
Accepted: April 29, 2003
Published: May 29, 2003

References

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new

software suite for macromolecular structure determination. Acta Crystallogr. D *54*, 905–921.

CCP4 (Collaborative Computational Project 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D. Biol. Crystallogr. *50*, 760–763.

Chapman, D.A., Zavialov, A.V., Chernovskaya, T.V., Karlyshev, A.V., Zav'yalova, G.A., Vasiliev, A.M., Dudich, I.V., Abramov, V.M., Zav'yalov, V.P., and MacIntyre, S. (1999). Structural and functional significance of the FGL sequence of the periplasmic chaperone Caf1M of *Yersinia pestis*. J. Bacteriol. *181*, 2422–2429.

Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S.L., and Knight, S.D. (1999). X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic *Escherichia coli*. Science *285*, 1061–1066.

Du, Y., Rosqvist, R., and Forsberg, A. (2002). Role of fraction 1 antigen of *Yersinia pestis* in inhibition of phagocytosis. Infect. Immun. *70*, 1453–1460.

Eshdat, Y., Silverblatt, F.J., and Sharon, N. (1981). Dissociation and reassembly of *Escherichia coli* type 1 pili. J. Bacteriol. *148*, 308–314.

Galimand, M., Guiyoule, A., Gerbaud, G., Rasoamanana, B., Chanteau, S., Carniel, E., and Courvalin, P. (1997). Multidrug resistance in *Yersinia pestis* mediated by a transferable plasmid. N. Engl. J. Med. *337*, 677–680.

Hung, D.L., Knight, S.D., Woods, R.M., Pinkner, J.S., and Hultgren, S.J. (1996). Molecular basis of two subfamilies of immunoglobulin-like chaperones. EMBO J. *15*, 3792–3805.

Jacob-Dubuisson, F., Striker, R., and Hultgren, S.J. (1994). Chaperone-assisted self-assembly of pili independent of cellular energy. J. Biol. Chem. *269*, 12447–12455.

Jones, T.A., Zou, J.-Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. *A47*, 110–119.

Keeling, M.J., and Gilligan, C.A. (2000). Metapopulation dynamics of bubonic plague. Nature *407*, 903–906.

Kleywegt, G.J., and Jones, T.A. (1994). Detection, delineation, measurement and display of cavities in macromolecular structures. Acta Crystallogr. D *50*, 178–185.

Knight, S.D., Berglund, J., and Choudhury, D. (2000). Bacterial adhesins: structural studies reveal chaperone function and pilus biogenesis. Curr. Opin. Chem. Biol. *4*, 653–660.

Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

Lamzin, V.S., and Wilson, K.S. (1997). Automated refinement for protein crystallography. Methods Enzymol. *277*, 269–305.

Lawrence, M.C., and Colman, P.M. (1993). Shape complementarity at protein/protein interfaces. J. Mol. Biol. *234*, 946–950.

MacIntyre, S., Zyrianova, I.M., Chernovskaya, T.V., Leonard, M., Rudenko, E.G., Zav'yalov, V.P., and Chapman, D.A. (2001). An extended hydrophobic interactive surface of *Yersinia pestis* Caf1M chaperone is essential for subunit binding and F1 capsule assembly. Mol. Microbiol. *39*, 12–25.

Miller, S., Janin, J., Lesk, A.M., and Chothia, C. (1987). Interior and surface of monomeric proteins. J. Mol. Biol. *196*, 641–656.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D *53*, 240–255.

Navaza, J. (1994). AMoRe: an automated package for molecular replacement. Acta Crystallogr. A *50*, 157–163.

Nicholls, A. (1993). GRASP: Graphical Representation and Analysis of Surface Properties (New York: Columbia University Press).

Saibil, H. (2000). Molecular chaperones: containers and surfaces for folding, stabilising or unfolding proteins. Curr. Opin. Struct. Biol. *10*, 251–258.

Sauer, F.G., Futterer, K., Pinkner, J.S., Dodson, K.W., Hultgren, S.J., and Waksman, G. (1999). Structural basis of chaperone function and pilus biogenesis. Science *285*, 1058–1061.

Sauer, F.G., Pinkner, J.S., Waksman, G., and Hultgren, S.J. (2002). Chaperone priming of pilus subunits facilitates a topological transition that drives fiber formation. Cell *111*, 543–551.

Soto, G.E., Dodson, K.W., Ogg, D., Liu, C., Heuser, J., Knight, S., Kihlberg, J., Jones, C.H., and Hultgren, S.J. (1998). Periplasmic chaperone recognition motif of subunits mediates quaternary interactions in the pilus. EMBO J. *17*, 6155–6167.

Thanassi, D.G., Saulino, E.T., and Hultgren, S.J. (1998). The chaperone/usher pathway: a major terminal branch of the general secretory pathway. Curr. Opin. Microbiol. *1*, 223–231.

Thanassi, D.G., Stathopoulos, C., Dodson, K., Geiger, D., and Hultgren, S.J. (2002). Bacterial outer membrane ushers contain distinct targeting and assembly domains for pilus biogenesis. J. Bacteriol. *184*, 6260–6269.

Titball, R.W., and Williamson, E.D. (2001). Vaccination against bubonic and pneumonic plague. Vaccine *19*, 4175–4184.

Vetsch, M., Sebbel, P., and Glockshuber, R. (2002). Chaperone-independent folding of type 1 pilus domains. J. Mol. Biol. *322*, 827-840.

Xu, J., Baase, W.A., Baldwin, E., and Matthews, B.W. (1998). The response of T4 lysozyme to large-to-small substitutions within the core and its relation to the hydrophobic effect. Protein Sci. *7*, 158–177.

Zav'yalov, V.P., Chernovskaya, T.V., Chapman, D.A., Karlyshev, A.V., MacIntyre, S., Zavialov, A.V., Vasiliev, A.M., Denesyuk, A.I., Zav'yalova, G.A., Dudich, I.V., et al. (1997). Influence of the conserved disulphide bond, exposed to the putative binding pocket, on the structure and function of the immunoglobulin-like molecular chaperone Caf1M of *Yersinia pestis*. Biochem. J. *324*, 571–578.

Zavialov, A., Berglund, J., and Knight, S.D. (2003). Overexpression, purification, crystallization and preliminary X-ray diffraction analysis of the F1 antigen Caf1M–Caf1 chaperone–subunit pre-assembly complex from *Yersinia pestis*. Acta Crystallogr. D Biol. Crystallogr. *59*, 359–362.

Zavialov, A.V., Kersley, J., Korpela, T., Zav'yalov, V.P., MacIntyre, S., and Knight, S.D. (2002). Donor strand complementation mechanism in the biogenesis of non-pilus systems. Mol. Microbiol. *45*, 983–995.

Accession Numbers

The PDB codes are 1P5V for the binary and 1P5U for the ternary complex.