Formation of Highly Stable Chimeric Trimers by Fusion of an Adenovirus Fiber Shaft Fragment with the Foldon Domain of Bacteriophage T4 Fibrin

Katerina Papanikolopoulou‡, Vincent Forge§, Pierrette Goeltz‡, and Anna Mitraki‡

From the ¶Institut de Biologie Structurale, 41 rue Jules Horowitz, Grenoble 38027 and the §Laboratoire de Biophysique Moléculaire et Cellulaire, Unité Mixte de Recherche 5090, Département Réponse et Dynamique Cellulaires, CEA-Grenoble, 17 rue des Martyrs, Grenoble 38054, France

The folding of β-structured, fibrous proteins is a largely unexplored area. A class of such proteins is used by viruses as adhesins, and recent studies revealed novel β-structured motifs for them. We have been studying the folding and assembly of adenovirus fibers that consist of a globular C-terminal domain, a central fibrous shaft, and an N-terminal part that attaches to the viral capsid. The globular C-terminal, or “head” domain, has been postulated to be necessary for the trimerization of the fiber and might act as a registration signal that directs its correct folding and assembly. In this work, we replaced the head of the fiber by the trimerization domain of the bacteriophage T4 fibrin, termed “foldon.” Two chimeric proteins, comprising the foldon domain connected at the C-terminal end of four fiber shaft repeats with or without the use of a natural linker sequence, fold into highly stable, SDS-resistant trimers. The structural signatures of the chimeric proteins as seen by CD and infrared spectroscopy are reported. The results suggest that the foldon domain can successfully replace the fiber head domain in ensuring correct trimerization of the shaft sequences. Biological implications and implications for engineering highly stable, β-structured nanorods are discussed.

Natural fibrous folds in proteins include collagen triple helices, α-helical coiled coils, or β-structured motifs (1, 2). Collagen and coiled coils remain well studied structural paradigms in the field of folding studies of fibrous proteins. However, many fewer studies have been done on the folding of elongated, β-structured proteins that very often are used by viruses as their attachment organelles. Recent studies revealed novel structural motifs for these proteins (3, 4). We use such a viral adhesin, the fiber protein of adenoviruses, as a model system to study folding and assembly of that class of proteins. The fiber is a homotrimeric protein (each monomer contains 582 amino acids) embedded into the viral capsid with a short, N-terminal part (about 45 residues). Its C-terminal, globular head domain is responsible for cell attachment. A central shaft domain contains a repeating sequence motif with an invariant glycine or proline and a conserved pattern of hydrophobic residues (5, 6).

The shaft from serotype 2 is 30 nm long, featuring 22 such repeats. The trimeric native fiber is resistant to proteases, heat, and to SDS at 4 °C (7, 8). This SDS resistance serves as the most reliable tool to assess correct folding and trimerization (8). At 4 °C, trimers migrate slowly in SDS-polyacrylamide gels and can be distinguished easily from partially folded or misfolded forms that migrate in the monomer position. The trimer band chases to the monomer band only upon boiling. In the presence of SDS at ambient temperatures, the full-length trimer starts to unfold from its N terminus. Under these conditions, a fragment of the protein comprising the head and four shaft repeats has been identified and proved to be SDS- and protease-resistant (8). The crystal structure of this stable fragment revealed a novel folding motif for the shaft sequences, termed the “triple β-spiral” (9). A β-propeller conformation for the globular head domain has already been reported in earlier studies (10, 11).

One natural question in the folding of rod-like molecules that contain repetitive sequences is how to avoid mismatched structures during folding and assembly. It has been established that collagens use noncollagenous domains at their C-terminal ends which act as nucleation points and ensure correct chain registration, preventing misfolded structures (12, 13). Registration motifs are also used by other fibrous proteins. In the case of phage T4 fibrin, a 27-amino acid C-terminal domain termed “foldon” serves as a registration motif for the segmented, triple coiled coil motif of the fibrin (14). This small domain in a β-propeller conformation can fold and trimerize autonomously (15). It has been reported that fibrinins with a deleted or mutated foldon domain fail to fold correctly, whereas N-terminal deletion mutants with an intact foldon domain trimerize successfully (16).

It has been noticed long ago by the virology community that N-terminal deletions of the adenovirus fiber as well as the globular head alone can fold and trimerize correctly both inside eukaryotic and bacterial cells (17–20). On the contrary, C-terminal deletions or mutations in the head domain prevented correct folding and assembly of the full-length fiber in vivo. Furthermore, a 41-amino acid synthetic peptide corresponding to shaft sequences 354–396 that are immediately adjacent to the head failed to fold into the triple β-spiral conformation; instead, it polymerized into amyloid fibrils (21). Thus, isolated shaft segments of any length fail to trimerize when produced in either recombinant or synthetic form (17, 18, 21). Taking this evidence into account, it has been postulated that the globular head is required for correct folding and assembly of the fiber and might play a registration role similar to other trimerization domains.

To understand the folding and trimerization mechanisms of...
Adenovirus Fiber Shaft-Fibritin Foldon Chimeras

**Experimental Procedures**

E. coli Strains and Plasmids—The strain DH5α (Invitrogen) was used for the selection of recombinant clones and plasmid DNA production. Protein expression was performed in the strain JM109(DE3) (Promega) which contains the λ-lysozyme DE3 encoding the T7 RNA polymerase under control of the isopropyl-β-D-thiogalactopyranoside-inducible lacUV5 promoter. DNA fragments were cloned in the pT7.7 vector that allows transcription from the T7 RNA polymerase promoter and contains the ribosome-binding site for effective translation in the cell (25).

Conformation of the genetic clones

<table>
<thead>
<tr>
<th>Primer</th>
<th>5'-Sequence-3'</th>
<th>Restriction site</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>CAATCTAAACAAAAAAACATAAGTTGAGCATAAAAA</td>
<td>Ndel</td>
</tr>
<tr>
<td>B</td>
<td>CGGCTGACGATCCTTTGTTACCTGTTGCGCC</td>
<td>BamHI</td>
</tr>
<tr>
<td>C</td>
<td>GAGGGGAGATTAGTTAAAAAAGGTTGAGCATAAAAA</td>
<td>Ndel</td>
</tr>
<tr>
<td>D</td>
<td>TTGCTAGACGGGCTTGTGCTTTGCTATCCTGTTG</td>
<td>BamHI</td>
</tr>
<tr>
<td>E</td>
<td>GGTAGTTGTTGCTATGTTGCTTGTATTGTTAAT</td>
<td>BamHI</td>
</tr>
<tr>
<td>F</td>
<td>CAAGGCTCCTCAAAGGATCCTGGTTATTT</td>
<td>BamHI</td>
</tr>
<tr>
<td>G</td>
<td>CTTGCGGGCCCTCACAATTCGTTGATTGCTTTAAGGGTAA</td>
<td>Clal</td>
</tr>
<tr>
<td>H</td>
<td>GAAGGCTCCTCAAAGGATCCTGGTTATTT</td>
<td>BamHI</td>
</tr>
<tr>
<td>I</td>
<td>CTTGCGGGCCCTCACAATTCGTTGATTGCTTTAAGGGTAA</td>
<td>Clal</td>
</tr>
<tr>
<td>J</td>
<td>CAACCAACTAAGTTGATGTTGAGCTGATGAGCATAAAAA</td>
<td>Ndel</td>
</tr>
<tr>
<td>K</td>
<td>CAAGGCTCCTACAAGGATCCTGGTTATTT</td>
<td>BamHI</td>
</tr>
</tbody>
</table>

The fiber, one approach we took was to replace the head domain with a smaller trimerization motif, namely the foldon, prompting by recent reports on its successful use as a foreign trimerization motif. It has been indeed shown that the foldon can successfully play a trimerizing and stabilizing role when fused with other fibrous motifs such as collagen sequences, phase T4 short tail fibers, and human immunodeficiency virus glycoproteins (15, 22, 23). Fusion proteins of a fibrin fragment comprising the foldon domain and part of the coiled coil domain have also been reported to be correctly folded and trimeric (24).

In this work, we used the adenovirus fiber stable fragment (residues 319–582), whose structure has been reported previously (9), as a minimal working model. This stable fragment comprises the shaft residues 319–392, a flexible linker sequence (residues 393–398) and the globular head (residues 399–582). Three chimeric proteins were constructed, two comprising the shaft segments (residues 319–392) with the foldon domain in its N-terminal end (in replacement of the natural head domain) and one with the foldon domain in the N-terminal end of the shaft segment. In one of the chimeric proteins with the foldon at the C-terminal end, the natural linker sequence (Asn-Lys-Asn-Asp-Asp-Lys, residues 393–398) was introduced because of the cloning strategy.

In all cases the amplification was performed for 30 cycles of 30 s denaturing at 94 °C, 30 s annealing at 60 °C, and 45 s extension at 72 °C. The amplified DNA fragments were isolated by preparative gel electrophoresis in 4% Nusieve GTG agarose gel and purified using the QIAquick Gel Extraction Kit (Qiagen). The purified bands and pT7.7 vector were digested for 1 h 30 min at 37 °C with restriction endonucleases purchased from Roche Applied Science. The digested DNA was purified using the QIAquick Gel Extraction Kit. Inserts and vectors were ligated enzymatically using the Roche Rapid DNA Ligation Kit in a 1:3 ratio. The ligation mixture was transformed into DH5α-competent cells by incubation for 30 min on ice followed by a temperature heat shock at 42 °C for 2 min. The transformed bacteria were plated on LB plates supplemented with 100 μg/ml ampicillin. When colonies were visible (onset of blue color), they were cultured for 18 h at 37 °C. The next day, the culture was used to inoculate 1 liter of LB medium containing 330 mM sorbitol, 2.5 mM betaine hydrochloride, and 100 μg/ml ampicillin. Cells were harvested and the plasmid DNA was isolated using the QIAprep® Spin Miniprep Kit (Qiagen). After digestion of the purified plasmid DNA, positive clones were identified by restriction enzyme digestion and were sequenced by the company Genome Express (Grenoble, France) using the T7 universal primer.

Protein Expression—Cells of the JM109(DE3) bacterial strain were transformed with 200 ng of the respective vector and plated on LB plates supplemented with 100 μg/ml ampicillin and 1% glucose. Single colonies were grown overnight in 10 mL of LB medium supplemented with 5.5 mM isopropyl-β-D-thiogalactopyranoside to a 0.5 final concentration. After 18 h at 37 °C, they were cultured for 18 h at 37 °C. The next day, the culture was used to inoculate 1 liter of LB medium containing 330 mM sorbitol, 2.5 mM betaine hydrochloride, and 100 μg/ml ampicillin. The cultures were grown at 37 °C until the Abs 600nm reached 0.4, then cooled to 22 °C. Protein expression was induced by the addition of isopropyl-β-D-thiogalactopyranoside to a 0.5 mM final concentration, and the incubation was continued further for 14 h at 22 °C. The cells were collected by centrifugation at 14,000 × g for 10 min and then resuspended in 20 mL of Tris buffer (50 mM Tris-HCl, pH 8.0, 2 mM EDTA, 20 mM NaCl) containing a tablet of Roche Complete™ protease inhibitors. The cells were lysed using a French pressure cell; pellet-supernatant fractionations were performed by centrifugation at 43,000 × g for 4°C for 20 min.

Protein Purification—The chimeric protein with the foldon domain at its C-terminal end and the linker sequence was purified as follows. After French pressure cell passage and removal of the insoluble material, streptomycin sulfate (Sigma) was added to the supernatant of the lysate to a final concentration of 1% (w/w) to precipitate nucleic acids. The pellet on solution was kept on ice for 15 min before centrifugation. After centrifugation, ammonium sulfate was added to the supernatant (with vigorous mixing) to a final concentration of 60% saturation. The solution was incubated for 15 min on ice and centrifuged at 43,000 × g for 20 min. The supernatant was recovered, and precipitation by ammonium sulfate was repeated going up to 80% saturation. The protein precipitate was dissolved in 10 mM Tris-HCl buffer, pH 8.5, 1 mM EDTA and dialyzed against the same buffer in the cold room. The protein solution was loaded onto a Resource Q column (Amersham Biosciences) equilibrated with 10 mM Tris-HCl buffer, pH 8.5, 1 mM EDTA at a flow rate of 3 mL/min. The protein was
eluted with a gradient of 0–300 mM NaCl. Fraction containing the protein were pooled and brought to 1.7 M ammonium sulfate. The protein solution was dialyzed against a phosphate buffer (25 mM Na₂HPO₄, 25 mM NaH₂PO₄, 1 mM EDTA, 1.7 M ammonium sulfate, pH 6.5) and applied to a phenyl-Superose 5/5 column (Amersham Biosciences). For elution, a linear gradient of 1.7–0 M ammonium sulfate was applied; the protein eluted at about 1.5 M. The purified protein was precipitated by adding ammonium sulfate to 80% saturation and stored at 4 °C. The original adenovirus stable fragment as well as the globular head domain were purified as described previously (27). The identities of all proteins produced in this study were confirmed with N-terminal amino acid sequencing, using standard procedures as described previously (8). The specific absorption coefficient of the purified chimeric protein with the foldon domain at the C-terminal end and the linker sequence was determined by UV spectroscopy combined with in-house quantitative amino acid composition analysis. The protein concentration was subsequently determined by UV spectroscopy using a specific absorption coefficient A₂₈₀ nm, 1 cm, of 0.89.

**Cell-free Expression of Chimeric Proteins and Immunoblotting—**The Rapid Translation System (RTS 500) E. coli Circular Template Kit, supplied by Roche Applied Science, was also used for the cell-free expression of the chimeric proteins with the foldon at the C-terminal end (with and without linker). The lyophilized components of the kit were reconstituted according to the product instructions using the ready-to-use buffer solution supplemented with 10 mM EDTA and one tablet of Complete™ mini-EDTA-free. 10 μg of plasmid DNA was used for a 1-ml reaction. For the coexpression of the two proteins, 7.5 μg of each plasmid was added to the 1-ml reaction mixture. The RTS Instrument was set to 22 °C with a stirring speed of 120 rpm. After 96 h of incubation a 5-μl aliquot was removed, electrophoresed on a 12.5% SDS-polyacrylamide gel, and transferred to a polyvinylidene difluoride membrane (Bio-Rad) at 12 V overnight under non-denaturing conditions. Both SDS-PAGE and electrophoretic transfer were performed at 4 °C. After transfer the blot was soaked in phosphate-buffered saline containing 5% powdered milk and 1% Tween 20, followed by incubation with an anti-fiber polyclonal antibody for 2 h. To probe the primary antibody, alkaline phosphatase conjugated to goat anti-rabbit Ig (Sigma) was used. The color reaction for alkaline phosphatase was developed with nitro blue tetrazolium and 5-bromo-4-chloro-3-indolyl phosphate. The membrane was rinsed with distilled water and air dried. The immunoblotting procedure was used for total E. coli lysates. Dual color precision plus protein standards (Bio-Rad) were used.

**Protease Digestion—**Chymotrypsin (sequencing grade from bovine pancreas, Roche Applied Science) was added to a purified protein solution (0.3 mg/ml in 50 mM phosphate buffer, 1 mM EDTA) at a 1:10 ratio w/w, in the presence or absence of 0.1% SDS. One sample of the protein, containing 0.1% SDS, was also boiled for 4 min prior to protease digestion. Digestion mixtures were carried out at 22 °C. At defined times, 10-μl aliquots were removed and mixed with 5 μl of 3X concentrated loading buffer (28) (to give a final SDS concentration of 2%) and frozen at −20 °C to stop the reaction. The resulting samples were analyzed by 15% SDS-polyacrylamide gel, without prior boiling, and visualized with Coomassie Blue staining. The same protocol was followed with endoprotease Glu-C, also obtained from Roche Applied Science.

**Circular Dichroism Spectroscopy—**CD measurements were performed on a JASCO J-810 spectropolarimeter equipped with a thermostated cell holder, using a quartz cell of 1.0-mm path length. Spectra were recorded at 22 °C over the wavelength range of 250–195 nm at 0.5-nm intervals and corrected for the contribution of the buffer. Each spectrum was the average of 30 scans. 50 mM phosphate buffer pH 6.5 was used for all proteins, and protein concentrations were as follows: 2.65 μM for the chimeric protein and 1.86 μM for the stable fiber fragment and the globular head domain.

**Fourier Transform Infrared Spectroscopy—**FTRIR spectra were collected on a JASCO 610 Fourier transform spectrometer. The protein was dissolved in D₂O at a final concentration of 0.86 mM and placed between a pair of CaF₂ windows (Spectra tech) separated by a 100-μm spacer. 1,000 interferograms were accumulated at 22 °C as a spectral resolution of 4 cm⁻¹. The water vapor spectrum was subtracted from the protein spectrum. The second derivative of the amide I band was used to identify the different spectral components. The second derivative was calculated using JASCO software with a window of seven points. The fit of the sum of the various components detected in the second derivative spectrum to the raw data was performed with the curve fitting software provided by JASCO. The area under each component was used to estimate the percentage of the various secondary structures. We made the assumption that the extinction coefficient was the same for each type of secondary structure.

**RESULTS**

Design of the Chimeric Proteins—In designing the chimeric proteins the structural compatibility of the two parts (the shaft domain residues 319–392 and the fibritin foldon domain residues 457–483) must be considered and, if necessary, linkers used to avoid potential structural conflicts. Inspection of the crystal structures revealed that the three Gly residues of the fiber shaft lie on a triangle with sides of 13.5 Å, and the three Gly residues of the foldon domain lie on a triangle with sides of 12.5 Å. These values are close enough to ensure a priori a compatibility. The schematic representations of the various constructs that we decided to build are shown in Fig. 1A. Construct 1 depicts the original adenovirus stable fragment, comprising the shaft segment (residues 319–392), the natural linker sequence (residues 393–398, Asn-Lys-Asn-Asp-Asp-Lys) that connects the shaft to the fiber head, and the globular head (residues 399–582). There is no clearly defined electron density for the natural linker in the crystal structure of the stable fragment (9), and the crystallographic axis of the shaft is tilted by 2° relative to the axis of the head, suggesting that the linker ensures flexibility of the two domains relative to each other. In construct 2, the foldon domain was joined to the C-terminal end of the shaft with the natural linker sequence. Apart from this linker, this construct also contains an additional Gly-Ser sequence as a consequence of inserting the BamHI site. In construct 3, the foldon domain was joined to the C-terminal end of the shaft domain with just the Gly-Ser insert; this construct will be herein referred to as having no linker from here on. For the chimeric protein that has the foldon domain at the N terminus of the shaft domain (construct 4), no linker sequence was used; three additional amino acids (Gly-Ser-Gly) were introduced between the domains as a consequence of the cloning strategy. The amino acid sequences of the shaft and the foldon domains are shown in Fig. 1B.

Expression of the Chimeric Proteins in E. coli—Fig. 2 depicts supernatants of E. coli lysates for the two chimeric proteins with the foldon domain at the C terminus. For comparison, the supernatant of a lysate of the original fiber stable fragment (residues 319–582) is shown (lanes 1 and 2). This original stable fragment is expressed in soluble form and it is SDS-resistant: the trimer, migrating at an apparent molecular mass of about 70 kDa (lane 1), dissociates to its monomeric form (apparent mass of 28 kDa, lane 2) only after boiling. The supernatant of the chimeric protein with the linker, nonboiled and boiled, is shown in lanes 3 and 4, respectively. The protein is produced in soluble form that migrates at an apparent molecular mass compatible with a trimer, i.e. about 33 kDa. After 3 min of boiling in 2% SDS, a substantial part of the chains still migrates in the position of the trimer, indicating a remarkable stability of the protein. The chimeric protein without the linker sequence migrates, as expected, slightly lower than the linker-containing protein in the nonboiled sample (lane 5) and still appears stable to SDS, although less so compared with the linker-containing protein because a greater percentage of chains is dissociated upon boiling (lane 6). The boiled forms of both chimeric proteins migrate at positions expected from their monomer theoretical molecular masses of 11,514 and 10,800 Da for the linker and nonlinker protein, respectively. These results strongly suggest that the chimeric proteins not only can fold properly in E. coli, but are also remarkably resistant to the combined heat and SDS action. The enhanced production of the chimeric proteins relative to the original fragment is also noteworthy.

---

1 The abbreviation used is: FTIR, Fourier transform infrared.
Oligomerization State of the Chimeric Proteins—In their oligomeric and nondissociated forms, SDS-resistant proteins do not bind SDS efficiently and do not migrate as normal SDS-polyacrylamide gels: as a result, very often they migrate anomalously in SDS-polyacrylamide gels depending on their shape, overall charge, and eventual degree of unfolding (8, 29). Thus, the migration of the chimeric proteins with an apparent mass compatible with a trimer has to be corroborated by another criterion to assess their trimeric form safely. The linker and monomer positions for the chimeric proteins are marked with arrows. Lane M, molecular mass markers.

Expression of the Chimeric Protein Carrying the Foldon Domain in Its N terminus in E. coli—Total lysates of the chimeric protein carrying the foldon domain in its N terminus after SDS-PAGE and immunoblotting with a polyclonal anti-fiber antibody are shown in Fig. 4. A partition of chains migrating in the trimeric apparent molecular mass position and in the monomer position is observed (lane 2). The chains migrating in the trimer position chase to the monomer upon boiling (lane 3). A smear of bands migrating slower than the trimeric band can be seen as well. Formation of slower migrating bands has been documented previously for the full-length fiber (8); these bands can result from partial unfolding of the native trimers induced by SDS and/or temperature. Alternatively, the observed smear might result from partially folded chains that failed to reach a compact, SDS-resistant trimeric state; the bands seem also to chase to the monomer position after boiling. Except for an extra glycine residue (see Fig. 1), this protein has exactly the same chemical composition as the chimeric protein that has the foldon domain in its C terminus without a linker; the trimer and monomer positions for the two proteins are the same. Total lysates, nonboiled and boiled, of this latter protein are shown (lanes 4 and 5, respectively). These results suggest that either some of the chimeric chains with the foldon in their N-terminal end fail to fold into SDS-resistant trimers, or, alternatively, they can fold into trimers that are SDS-sensitive. Thus, we can conclude that the foldon domain fails to lead to efficient formation of SDS-resistant trimers when appended to the N-terminal end of the shaft segment.

Structural Characterization of the Chimeric Protein Carrying the Linker Sequence—The ability of the chimeric proteins with the foldon at their C terminus to form stable, SDS-resistant trimers suggests that the shaft and the foldon domain...
sequences can *a priori* be accommodated in the same molecule to form a “foldable,” viable and stable protein. The question immediately raised about these proteins is whether the two domain sequences fold into their original structure within the chimeric protein. To carry the structural characterization of the chimeric proteins we opted to continue with the purification and study of the protein with the foldon at the C-terminal end and the linker sequence.

CD spectra of the purified protein (Fig. 5, upper panel) show a single minimum at 203 nm, a shoulder at 215 nm, and a positive peak at 230 nm. According to the previously reported spectra of the foldon domain alone (15), the shoulder at 215 nm and the positive signal around 230 nm can be attributed to the foldon. The minimum at 203 nm can be compared with the strong minimum at 208 nm reported previously for the spectrum of the full-length fiber (5); this unusual minimum position most probably reflects a high content of β-turns and loops in combination with β-structure. The minimum at 203 nm is also detected in the spectrum of the stable adenovirus fiber fragment (Fig. 5, lower panel). This minimum is not present in the spectrum of globular head alone which shows, in agreement with its β-propeller conformation (11), a minimum around 215 nm (Fig. 5, lower panel). Therefore, in the spectrum of the stable adenovirus fiber fragment as well as in the one of the chimeric protein, it seems reasonable to assign the minimum at 203 nm to the novel folding motif revealed in the crystal structure of the stable fragment for the shaft structure, the triple β-spiral (9).

We also sought to study the chimeric protein by infrared spectroscopy, known to give more reliable estimates for β-sheet content in proteins. The FTIR spectrum of the chimeric protein in the amide I region is shown in Fig. 6 (upper panel). The spectral components are revealed in the second derivative of the spectrum (lower panel). The three bands between 1638 cm⁻¹ and 1620 cm⁻¹ together with the one at 1680 cm⁻¹ indicate the presence of a high content of β-structures. The band at 1638 cm⁻¹ has been related to the presence of parallel β-strands (30); therefore the amide I band of the chimeric protein is in agreement with a conformation that contains parallel together with anti-parallel β-strands. The numerous bands around 1660–1690 cm⁻¹ show the presence of a large amount of turns in the structure. Finally, the peak at 1647 cm⁻¹ indicates the presence of a significant amount of the random coil conformation. The peaks detected in the second derivative spectrum were fitted to the raw amide I band (Fig. 6, upper panel), and the area under each peak was used to analyze the percentage of the various secondary structure components (30, 31). We found the following decomposition: 42% β-structure, 42% turns, and 16% random coil. When, as an approximation, we used the band at 1638 cm⁻¹ to estimate the amount of parallel β-structures, we found that about 25% of the β-structure should be parallel. These percentages are in fair agreement with those estimated from crystallographic data. If the structure of the chimeric protein was the same as the sum of its parts in their native context, the secondary structure composition deduced from the crystallographic data would be: 34% β-sheet, of which 10% parallel and 24% antiparallel, 36% turns, 7% 3₁₀ helix, and 23% random coil. It should be noted that some residues in the crystal structure could be classified as being in an extended conformation, i.e. having ϕ and ψ angles of β-structure but not being engaged in hydrogen bonding. The above percentages are calculated by classifying these residues as random coil. If these residues were classified as β-structure, the overall β-sheet percentage would be 47%, with random coil decreasing to 10%. The small amount of 3₁₀ helix expected according to the crystallographic data is not detected in the amide I band of the chimeric protein. Taken together, the CD and FTIR results suggest that the secondary structure...
Adenovirus Fiber Shaft-Fibritin Foldon Chimeras

The ensemble of the above results implies that the chimeric protein in the presence of chymotrypsin. Except for one minor band that appears after 30 min of digestion at a 10:1 w/w protein:protease ratio at 22 °C (Fig. 7A, lanes 2–5), the trimer band remains essentially intact after 5 h of proteolysis, indicating a fairly good resistance to the protease. If proteolysis is carried out in the presence of 0.1% SDS at 22 °C the protein becomes susceptible to digestion, as indicated by the decrease of the trimer band and the concomitant appearance of proteolytic bands (Fig. 7B, lanes 2–5). If the protein is boiled at 100 °C for 4 min in the presence of 0.1% SDS, cooled on ice for 1 min, then subjected to proteolysis, the following pattern is observed: a small amount of monomer band is produced (lane 6) which disappears from the first minutes of digestion (lanes 7–9). Again because of its exceptional stability, the majority of the protein still remains in the trimer band after boiling (lane 6). However, the protein becomes increasingly susceptible to proteolysis, as indicated by the substantial degradation of the trimer band, and appearance of a ladder of proteolytic products (lanes 7–9). 0.1% SDS is employed instead of 2% SDS, so that proteases can still act (both chymotrypsin and endoproteinase Glu-C are active in 0.1% SDS). A similar pattern of results was obtained with endoproteinase Glu-C under the same conditions (not shown). The ensemble of the above results implies that the chimeric trimer is fairly robust and essentially resistant to proteases, becoming susceptible to digestion after boiling in 0.1% SDS.

**FIG. 7.** Protolytic digestion of the purified protein. A, time course of proteolytic digestion of the purified protein with the linker in the absence of SDS. Chymotrypsin was added to a purified protein solution (0.3 mg/ml in 50 mM phosphate buffer, pH 7, 1 mM EDTA) at a 1:10 ratio w/w, and digestion was allowed to proceed for the following times: 30 min (lane 2), 1 h (lane 3), 3 h (lane 4), and 5 h (lane 5). Lane 1, protein control in the absence of protease; lane 6, protease control in the absence of protein. B, time course of proteolytic digestion in the presence of SDS. The digestion protocol is the same as in A, except that the protein was incubated in the presence of 0.1% SDS before adding the protease. Lane 1, protein control in the absence of protease. Lanes 2–5, same times of digestion as the corresponding lanes in A. Lane 6, the protein was boiled for 4 min in 0.1% SDS, cooled on ice for 1 min, and then the protease was added. Digestion was subsequently carried out at 22 °C for 30 min (lane 7), 1 h (lane 8), and 3 h (lane 9). Aliquots were taken at the corresponding times, and sample buffer was added to give a 2% final SDS concentration; the samples were frozen at −20 °C until electrophoresis. The samples have not been boiled prior to loading in the gels. All gels (15%) were stained with Coomassie Blue. Lane M, molecular mass markers.

**DISCUSSION**

In this work, we created chimeric proteins by fusing the phage T4 fibritin trimerization domain to either the C or N terminus of the shaft segment 319–392 of the adenovirus fiber. The ensemble of our results described above proves that the totality of the synthesized chains with the foldon domain in their C terminus forms trimeric, SDS-resistant proteins. The successful trimer formation by the chimeric proteins suggests that the foldon domain can act as a trimerization motif when appended to the C terminus of the shaft segment. In addition, the formed trimers are remarkably resistant to the combined heat and SDS action. However, when the foldon domain is fused at the N terminus of the shaft fragment, the chains partition among monomeric, trimeric, and slower migrating forms that chase to monomers after boiling. This partition suggests the following possibilities: either some of the chains fail to reach the trimeric form, or alternatively, they might be able to fold into unstable trimers that are SDS-sensitive. The ensemble of these results clearly suggests that the foldon domain has to be appended to the C-terminal end of the shaft sequences to lead to efficient formation of SDS-resistant trimers. Previously postulated mechanisms for the action of the foldon imply that, as an autonomous folding domain, it is the first to trimerize, providing a “template” that brings the sequences into close proximity and correct register (15, 16). In the case of the shaft fragment, the role of the foldon would not be merely to provide an efficient “clamp.” This purpose is also certainly served when the foldon is located at the N-terminal domain of the shaft frag-
ment; however, in this position the foldon domain fails to act efficiently. To act efficiently, the foldon domain must be located at the C-terminal end of the shaft fragment, implying that besides “tethering,” directionality in folding and assembly of the shaft sequences, i.e. going from C to N terminus, has to be assured. To this date there is no direct kinetic evidence for this directionality in the folding of the fiber; such a mechanism has to be proved explicitly by folding studies of the full-length fibers, where it should be a priori possible to distinguish the folding of the head from the shaft domain. However, obtaining high quantities of recombinant material for the full-length fibers still represents a considerable challenge. It is noteworthy that a number of fibrous superhelical proteins such as collagens (12), fibrin (16), myosin, and paramyosin (32) have C-terminal domains that seem to play a crucial role in their folding and assembly. Serving as recognition domains that initiate folding and provide the correct register for the assembling chains seems to be the underlying role for these otherwise structurally different domains. The C-terminal location of these domains is a common feature that is certainly thought provoking; however, a pertinent biological explanation remains to be found.

The conformational signatures of the chimeric proteins as seen by CD and FTIR spectroscopy are a priori compatible with a combination of the structural signatures of the foldon and shaft domains. However, this has to be corroborated by a crystal structure; crystallization trials are currently under way. The possibility of having a correctly folded shaft domain without the head can open interesting perspectives for a functional characterization of the shaft. It has long been considered that the shaft plays only a mechanical role in adenoviruses; however, recent studies suggest that it may play a functional role in infection (33–35). It has also been suggested recently that elongated, β-fibrillar structures in viruses may have evolved as carbohydrate-binding motifs (4, 36, 37). Obtaining the shaft motif in the absence of the globular head can help to address this point explicitly in vitro. The foldon domain has no known receptor function and therefore should not interfere with this kind of study.

Replacing the head domain of the adenovirus fiber with the foldon domain could also be of practical relevance. Adenoviruses are used for gene therapy, and new generations of gene therapy vectors seek to target certain tissues specifically. To avoid targeting to several tissues, one needs first to remove the natural receptor-binding head domain and introduce a new binding motif specific for the target tissue. Given the double role of the head domain (receptor binding and trimerization motif), removing this domain impairs proper folding and assembly of the fiber. To ensure correct folding the head domain must be replaced with another trimerization motif; a number of such approaches have been published in the field of gene therapy (for review, see Ref. 38). The trimerizing and stabilizing action of the foldon, together with its absence of binding to receptors, could make it ideal for this purpose. Fusion proteins of a fibrin fragment comprising the foldon domain and part of the coiled coil domain with the two most N-terminal repeats of the adenovirus shaft have been proposed previously for gene therapy vector use (24). To our knowledge, this is the first report that proves that the fibrin foldon domain alone (without a coiled coil part) can be fused directly to fiber shaft sequences to give trimeric and extremely stable proteins. Thus, the engineered shaft-foldon chimeric proteins described here, further modified with tissue-targeting motifs, could also prove useful in experimental gene therapy.

The exceptional resistance to the combined action of heat and SDS is another attractive feature of the chimeric proteins described here. Engineering fibrous proteins that can withstand extreme conditions is of considerable interest for biomaterials and nanotechnology applications (39–41). Adding more shaft repeats to the chimeric proteins described here to make longer fibrous constructs with controlled dimensions is one interesting perspective. If the shaft sequences indeed adopt the triple β-spiral conformation within the chimeric proteins, then it should be possible to functionalize these constructs using the surface-exposed loops. Furthermore, the results of the present study further corroborate previous reports that the foldon domain is a versatile trimerization motif and can be combined with a variety of fibrous motifs, provided the appropriate linkers are used. A future challenge would be to combine different motifs within the same fibrillar molecule, to create stable rods with multiple functions.

Acknowledgments—We thank Jean-Pierre Andrieu for N-terminal sequencing and amino acid analysis and Dr. Jean Gagnon for help with protein purification. We are grateful to Dr. Mark van Raaij (University of Santiago de Compostela, Spain) for crystallographic advice on the design of the chimeric proteins, calculating the secondary structure content from crystallographic data, and critical reading of the manuscript. We also thank Dr. van Raaij for the gift of the adenovirus plasmid pT7.Ad2fib388.

REFERENCES
Adenovirus Fiber Shaft-Fibritin Foldon Chimeras

Formation of Highly Stable Chimeric Trimers by Fusion of an Adenovirus Fiber Shaft Fragment with the Foldon Domain of Bacteriophage T4 Fibrin
Katerina Papanikolopoulou, Vincent Forge, Pierrette Goeltz and Anna Mitraki

doi: 10.1074/jbc.M311791200 originally published online December 29, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M311791200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 40 references, 11 of which can be accessed free at http://www.jbc.org/content/279/10/8991.full.html#ref-list-1