A Model for the Structure of the C-Terminal Domain of Dragline Spider Silk and the Role of Its Conserved Cysteine

Shmulik Ittah,[†] Amit Michaeli,[‡] Amiram Goldblum,[‡] and Uri Gat*[,†]

Department of Cell and Animal Biology, Silberman Life Sciences Institute, Edmond Safra Campus at Givat-Ram, and Department of Medicinal Chemistry, School of Pharmacy, Faculty of Medicine, The Hebrew University, Jerusalem 91904, Israel

Received April 25, 2007; Revised Manuscript Received July 6, 2007

Dragline spider silk fibers have extraordinary attributes as biomaterials of superior strength and toughness. Previously we have shown that the conserved C-terminal domain of a dragline spider silk protein is necessary for directing oriented microfiber formation. Here we present for the first time a state-of-the-art model of the three-dimensional structure of this domain, and, by comparing several dragline proteins, identify its key evolutionarily conserved features. Further, using the baculovirus expression system, we produced recombinant proteins that are mutated in the unique cysteine residue present in the domain. While a conservative mutation to serine allows fiber formation, thus demonstrating that there is no need for disulfide bond formation in this system, a mutation to arginine significantly alters the local surface properties, preventing fiber formation. These experimental results are in agreement with our model, wherein the cysteine is localized in a highly conserved hydrophobic loop that we predict to be important for the protein−protein interactions of this domain and hence also for fiber formation.

Introduction

Among the different types of spider silks, the "dragline silk" secreted by the major ampullate gland is studied most intensely. Dragline silk is used by orb (circular) web-weaving spiders to build the frame and radii of their nets as well as to serve as a lifeline when a spider drops to escape a threat.[1−3] Additionally, dragline silk displays a large degree of elasticity, which, together with its strength, results in a toughness that is higher than that of most other known materials whether natural or man-made.[2−6]

These extraordinary properties are the result of the unique architecture of the silk threads, which are composed of spider silk proteins, termed spidroins as they belong to a larger family of structural proteins called fibroins. Spider silk genes were found to encode for large proteins with a major central repetitive part flanked by much shorter nonrepetitive N- and C-terminal domains.[6,7] The repetitive regions are composed of short reiterated sequence motifs of largely nonessential amino acids such as alanine and glycine. The spidroins have their unique peptide repeats, which vary between the different proteins but always harbor a peptide containing multiple alanines and one or more glycine-rich motifs such as GPGXX or GGX. X-ray diffraction measurements, NMR, and other physical assays have shown that the A_n (alanines) block modules adopt a crystalline β-sheet structure, which confers toughness to the fiber. The glycine-rich sequences adopt a yet uncharacterized conformation, generally termed "amorphous", which is thought to confer the elasticity property of the fiber.[3,8−10] Dragline spider silk fibers were shown to be composed of two proteins, spidroin 1 (MaSp1) and spidroin 2 (MaSp2) (MaSp = major ampullate gland spidroin), which are similar in their overall structure but differ in some of the sequence elements within their glycine-rich repetitive domains.[3,5,11−13] The molecular structure of spider silks is of utmost interest in proteins and material sciences, but is still far from being resolved and constitutes a classical riddle of nature left to be deciphered.

Many studies have been performed on the main repetitive part of spidroins, but what is the role of the nonrepetitive N- and C-terminal domains? It was difficult to obtain N-termini sequence data because of the length of the spidroins' messenger RNAs (mRNAs) and because of their repetitive nature. The first sequence information became available for the flagelliform (spiral capture silk) spidroins, analysis of which indicated that, like in other fibroins, this domain is hydrophilic and may assume a globular structure.[7] Sequences of genomic clones obtained from several species and a complementary DNA (cDNA) sequence allowed the derivation of several dragline N-terminal coding regions and demonstrated that this region contains a signal peptide for secretion, is largely α-helical, and is the most highly conserved domain in the entire coding region.[14,15]

When multiple spidroin C-terminal sequences were compared, it was found that this particular domain is highly conserved between major ampullate proteins from different species (about 75% identity), is closely related to the C-termini of minor ampullate spidroins,[7,16,17] and even resembles (30% identity) those of the more "primitive" cribellate spider silks.[18] Studies using C-terminal domain-specific antibodies demonstrated that, indeed, the C-terminal domains of *Nephila clavipes* dragline spidroins are part of the spinning dope of the gland and are also present in the final fiber.[19] Experiments using recombinant C-terminal domains expressed in *Escherichia coli*, and analysis of extracts made of ampullate gland secretions in sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) under denaturing versus nondenaturing conditions have suggested that the conserved cysteine in the dragline C-terminal domains may elicit dimer formation.[19,20] Thus, the conserved cysteines are thought to form intermolecular dimers of possible importance for polymerization, which may affect the fiber's mechanical qualities.[17,19]

* Corresponding author. Tel/Fax: 972 2 6585920. E-mail: gatu@vms.huji.ac.il.

[†] Department of Cell and Animal Biology.

[‡] Department of Medicinal Chemistry.

10.1021/bm7004559 CCC: $37.00
Published on Web 08/14/2007

Previously, we have established a baculovirus-based system for the high-level expression and exploration of the structure of fibers produced by the self-assembly of dragline silk proteins in insect cells.[12] Although this system of polymerization is different from the natural mechanical drawing performed by the spider, many of its attributes are still in common, as the resulting structures share similarities in shape and chemical resistance.[12] As this system is the closest heterologous expression system to spiders used to date and allows facile genetic engineering of the dragline spider silks, we have directly examined the role of the C-terminus of the recombinant *Araneus diadematus* fibroin (rADF)-4 protein in their polymerization. Our results clearly demonstrated that the C-terminal domain of ADF-4 is necessary for the proper structure of the basic nanometer-scale fibrils as well as for the oriented assembly of the final micrometer-scale rADF-4 fibers.[21]

Here we further study the structure of the conserved C-terminal domain by presenting a three-dimensional energy-minimized model for its structure. In addition, we provide direct experimental evidence showing that the conserved cysteine residue is structurally important for a critical loop within this domain while providing strong evidence that fiber formation can take place in the absence of covalent disulfide bonds at least in cultured insect cells.

Materials and Methods

Donor Plasmid Construction. The construction of the basic donor plasmid pFastBacHTa, containing partial cDNA of ADF-4 (gene bank entry U47856), has been previously described. Two variants of this donor plasmid were constructed coding for the replacement of the C-terminal cysteine residue by serine or arginine. This replacement was mediated by point mutations using a QuikChange II site-directed mutagenesis kit (Stratagene) according to the manufacturer's instruction manual.

For each mutation, two primers were used: one sense and the other antisense (i.e., reverse complement of the sense). Only the sense primer is shown, with the codon for the amino acid underlined and the mutated nucleotide in bold:

Serine: 5′-CCAGGTTTATCGGGAT**C**TGATGCTCTTGTGCAGGC-3′

Arginine: 5′-CCAGGTTTATCGGGA**C**GTGATGCTCTTGT-GCAGGC-3′

Cell Culture, **Production of Recombinant Baculovirus**, **and Expression of rADF-4-Based Mutated Proteins.** These were performed as previously described.[21]

Purification of Fibers and Aggregates. Infected cells were harvested 3−5 days post-infection and centrifuged for 10 min at 16 000*g*. The cell pellets were resuspended in a 10% SDS solution and boiled for 1 h, and the protein assemblies were sedimented as above. Purified fibers/aggregates were resuspended at the desired solution and volume.

Immunocytochemisty. Cells grown on cover slips at 50% confluency were infected with recombinant viruses at a multiplicity of infection (MOI) of 10. Three days post-infection, cells were fixed with methanol at −20 °C. Cover slips were incubated with mouse anti-His_6 monoclonal antibody (Roche) at a 1:300 dilution followed by Texas Red-conjugated anti-mouse secondary IgG at 1:500 dilution. Cells were observed with an Olympus BX51 fluorescence microscope, and images were taken with a Magnafire SP camera or analyzed by confocal microscopy.

Protein Homology Modeling. Multiple structure alignments were performed using the FUGUE module of the TRIPOS Sybyl7.3 program.[22] The model was then generated using the ORCHESTRAR module in SYBYL 7.3 by the Tripos Company on the basis of the highest scoring (FUGUE z-score) structure. Model mutations were done using the Biopolymer module of the Insight II program (Accelrys, Inc.). Following modeling and mutation, energy minimizations were performed. Hydrogen atoms at pH 6.3 were added to the protein structures using the Insight II program. In order to simulate an uncharged N-terminus of the C-terminal domain, which is more similar to the natural, full length state, atoms were added to the model up to the Cα of the preceding residue. Minimization was performed using Discover with the AMBER force field by the steepest descent algorithm (500 steps) and conjugate gradient (10 000 iteration steps until reaching convergence of less than 0.001 kcal/Å). Hydrogen atoms were minimized first with a constraint on all heavy atoms, then side chains were minimized with constraints on the backbone, and finally the structure was minimized without constraints. A distance-dependent dielectric $\epsilon = 4r$ and cut-off distance of 18 Å (with a 1.5 Å switch) were used for coulomb electrostatic energies. Backbone root-mean-square deviation (rmsd) calculations were performed following superimposition using the Insight II program. Modeling of atomic clashes and Ramachandran plots were performed using MolProbity.[23]

Results and Discussion

Several studies using different experimental strategies, including ours, have implicated the conserved C-termini of dragline spider silks as being important for fiber assembly.[7,16,19−21,24] Secondary structure predictions have been made on *N. clavipes* spidroin 1 and 2 C-terminal domains, and the expected structures deduced included three largely hydrophobic domains that were expected to adopt amphipathic α-helices with the central helix containing the conserved cysteine residue.[20] However, no three-dimensional model attempting to predict the complete conformation of the C-terminal domain and the shape it may hold has been presented, and thus we set out to establish such a model in order to be able to start testing it experimentally.

We aligned the C-terminal domain sequences of all available MaSp's to date using ClustalW (Figure 1). As can be seen, these sequences share a high level of similarity, with about 75% identity as reported before[7] and close to 90% similarity taking into consideration the chemical characteristics of amino acids, which are color coded in the alignment. Among the most conserved amino acids is the cysteine (Figure 1, shaded), which is found at the same position in all sequences except that of *Araneus ventricosus* MaSp1, which is a divergent major ampullate spidroin showing many characteristics of flagelliform silk in its repetitive sequence.[7] Interestingly, all of the sequences show a conserved aspartic acid following the cysteine (Figure 1, shaded) except for *A. ventricosus* MaSp2, which harbors a functionally similar glutamic acid at this position.

We then chose the C-terminal sequence of ADF-4, which we have used in our previous studies (Figure 1, bottom),[12,21] and made use of the advanced ORCHESTRAR homology modeling program suite in order to predict a three-dimensional structure for this domain. This program suite is based on FUGUE technology, which finds distant homologues based on sequence−structure comparison.[22] The inherent advantage of this modeling strategy is its ability to use homology modeling on a protein domain for which there is no close homologue that has been solved.[25] We hypothesize that the C-terminal domain is mostly embedded in a protein environment, which poses a problem for alternative "ab initio" methods that are based on the assumption of an aqueous environment. We thus based our model on the best output protein sequence of the program, which is, in this case, the structure of PAS factor (PDB code 2B8I).[26] This factor belongs to a family of proteins that interact closely with lipid membranes via a hydrophobic patch, and thus we find this similarity important and relevant, as other

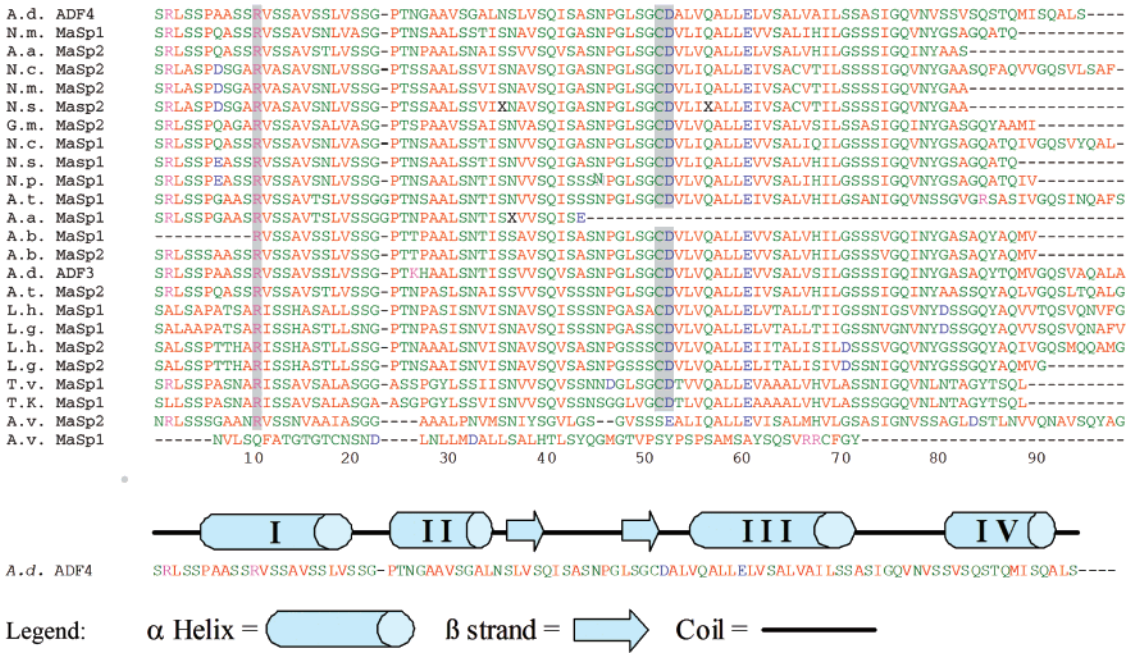

Figure 1. Multiple sequence alignment of the C-terminal domains of different MaSp proteins. The colors of amino acids represent common chemical properties (e.g., blue for acidic). The conserved cysteine, aspartic acid, and arginine amino acids that are mentioned in this work are shaded. Bottom: The secondary structural elements that are predicted by our model are depicted above the sequence of ADF-4, which we have used in this study. The two letters preceding the sequences represent the initials of the genus and species: *N.c*, *Nephila clavipes*; *N.m*, *Nephila madagascariensis*; *N.s*, *Nephila senegalensis*; *N.p*, *Nephila pilipes*; *A.t*, *Argiope trifasciata*; *A.a*, *Argiope aurantia*; *A.b*, *Araneus bicentenarius*; *A.d*, *Araneus diadematus*; *L.h*, *Latrodectus hesperus; L.g*, *Latrodectus geometricus*; *A.v*, *Araneus ventricosus*; *G.m*, *Gasteracantha mammosa; T.k, Tetragnatha kauaiensis*; *T.v, Tetragnatha versicolor.*

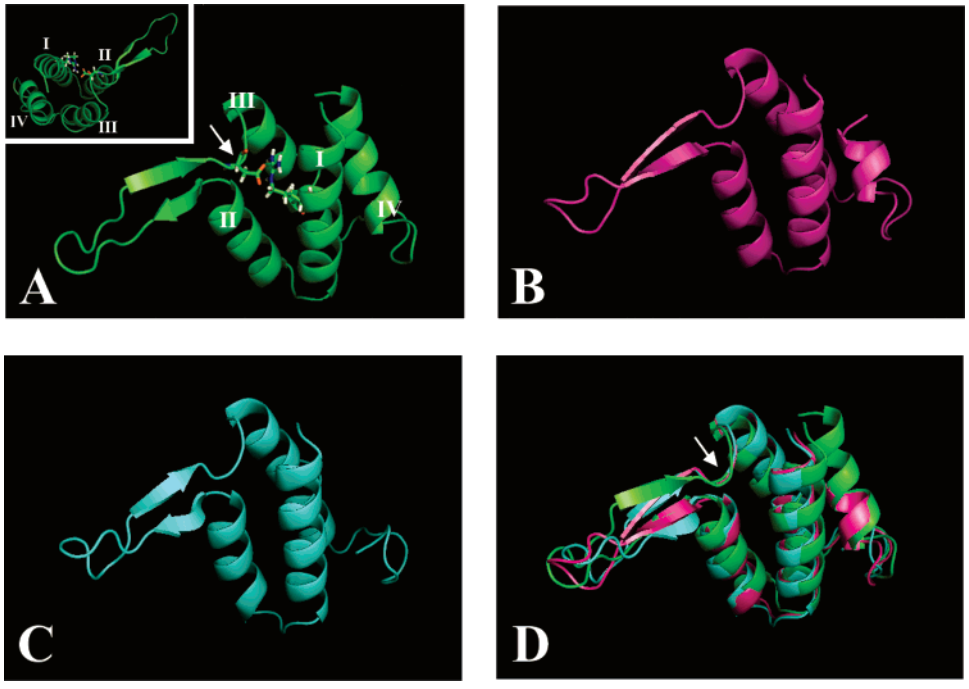

Figure 2. A proposed structure for the C-terminal domain of MaSp silk proteins. (A) Predicted tertiary structure for C-terminal domain of *A. diadematus* ADF-4. Inset: A different angle of the structure showing the circular arrangement of the helices in the imaginary barrel structure. The cysteine-containing loop that is N-terminal to helix III is marked by an arrow, and the D and R residues that are involved in a salt bridge between this loop and helix I are shown in a ball and stick format. (B) Model of *N. madagascariensis* MaSp1. (C) Model of *A. aurantia* MaSp2. Notice that this protein has no helix IV. (D) Superimposition of all three predicted structures. The conserved loop mentioned in panel A is marked by an arrow. All tertiary structures were generated using the ORCHESTRAR homology modeling program.

parts of the spider silk, such as the polyalanine repeats, resemble the alkane chains of a phospholipid more closely than an aqueous environment. The results predicted an overall barrel-like structure containing four parallel-oriented α-helices (I−IV), one of which is broken into two short segments (helix 3), and two antiparallel β-strands, which form the base of a loop, oriented orthogonal relative to the axis of the α-helices (Figure 2A and Figure 1, bottom), which we termed "the thumb".

In order to check the generality of our three-dimensional model, we constructed further models for the C-terminal domains of MaSp's of two other species: the MaSp1 of *Nephila madagascariensis* (*N.m*) and the MaSp2 of *Argiope aurantia*

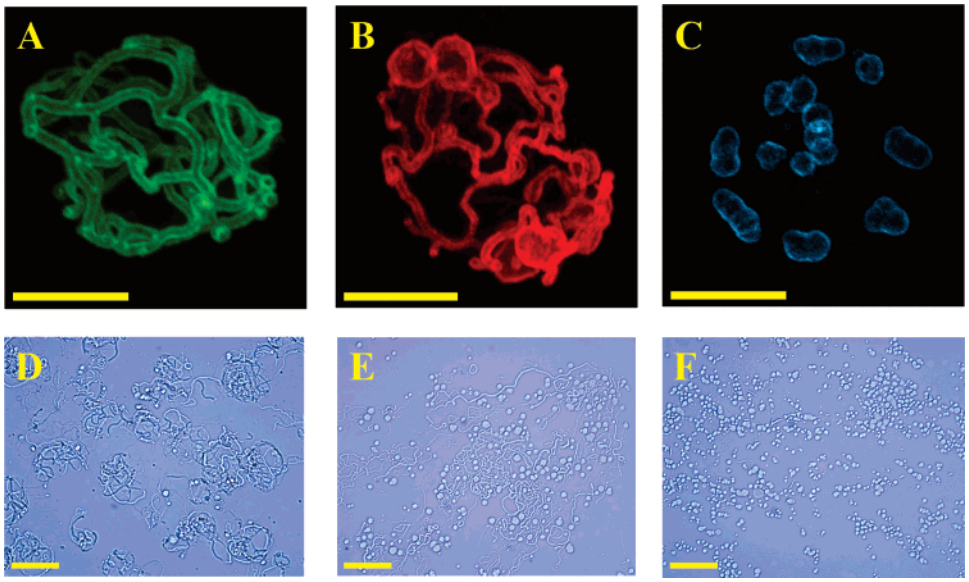

Figure 3. Expression and isolation of rADF-4 and its mutant versions. (A) An rADF-4 fiber in the cytoplasm of an infected *Sf9* insect cell viewed by confocal microscopy using anti-His_6-tag antibodies. (B) Same as panel A, but a C52S fiber. (C) Same as panel A, but C52R aggregates. (D−F) Fibers or aggregates purified from infected insect cells corresponding to panels A−C. The scale bars in panels A−C represent 10 μm, and those in panels D−F represent 25 μm.

(*A.a*) (Figure 2B,C, respectively). The models were constructed by changing the relevant residues of ADF-4 (18 conversions to form *A. aurantia* MaSp2 and 24 for *N. madagascariensis* MaSp1) and by the deletion of extra C-terminal residues. The models were all energy minimized and analyzed for clashes and Ramachandran plot deviations. All models were clash free, and Ramachandran outliers (<0.2%) were observed in all models at leucine 66 (α-helix), as well as in position 81 (coil) of *A. diadematus* (alanine) and of *A. aurantia* (serine). The models maintained structural similarity to ADF-4, as can be observed by their superimposition (Figure 2D), with backbone rmsd's of 1.89 (MaSp1) and 1.96 (MaSp2). This overall similarity is expected since the FUGUE program uses multiple sequence alignments;[22] however, the substructures that were still found to vary between the three MaSp domains may be biologically relevant. The only major difference between the three structures is that the C-terminal most helix IV of *N. madagascariensis* is shorter, and that *A. aurantia* does not have this helix at all. Importantly, the cysteine residue initiates a short loop between the second β-strand and helix III (C52−V56 of *A. diadematus*), a loop whose conformation is of the most highly conserved elements in the three models. Interestingly, the aspartic acid, which follows the cysteine in the loop (D53), creates a salt bridge in our model with arginine 11 of helix 1 (Figure 2A), which may serve as a stabilizing factor determining the right angle between the helices and the "thumb". This bond is also present in all three of our models, and since these two residues are highly conserved between species, it may well be found in all MaSp structures (Figure 1) and beyond − in other spider silks like the minor ampullate and flagelliform silks and even in fibroins of some cribellate weavers, which contain these conserved residues.[18]

In light of the important function suggested for the aforementioned cysteine residue in the C-terminal domain of MaSp proteins and its conservation also in our model,[19,20] we prepared two different replacement mutations: In the first, we replaced it with serine, which is the most similar amino acid in terms of atomic structure (exchange of S for O) and chemical properties (both are nucleophilic with similar hydrogen bond geometries), and is quite close in size, even though the sulfur atom is bigger. In the second, we exchanged cysteine for arginine, which has a totally different atomic composition and shape, is much larger and is positively charged.

To study the function of these mutated dragline proteins, we employed our baculovirus-based expression system. Expression of the original rADF-4 monomers resulted in their self-assembly to create fibers within the cytosol of the expressing insect cell as assayed by confocal microscopy of virus-infected *Spodoptera frugiperda* 9 (*Sf9*) cells, using antibodies directed against the N-terminal His_6-tag (Figure 3A). In addition, we obtained purified fibers, which are relatively easy to isolate after lysis of the cells (Figure 3D) and possess some properties of native dragline fibers.[12,21] One of the advantages of this system is that it is easy to create recombinant baculovirus coding for a protein that contains a desired mutation. In order to examine our hypothesis regarding the conserved C-terminal cysteine residue, we created two new baculoviruses, coding for above-described mutated versions of the rADF-4. The two recombinant baculoviruses were used to infect insect cells, the expression of which was detected by immunoblotting infected cell lysates using the aforementioned anti-His_6-tag antibodies (results not shown). We then used light microscopy as well as immunofluorescence as described above to observe whether any differences exist between the three proteins: the original and the two mutated ones. The C52S (number relative to the start of the C-terminus) mutant maintained the property of self-assembly into fiber (Figure 3B), in that infected cells displayed long threads in their cytoplasm, but did show a difference from the original product by containing a significant proportion of large and round insoluble aggregates that can also be purified along with the fibers (Figure 3E). Similar to the rADF-4 fibers, the diameter of the fibers varied from about 100−800 nm, but, within each cell, the diameter was remarkably constant, probably attesting to a nucleation event leading to rapid elongation of the fibers. In contrast to the nearly 80% fiber formation ratio observed in the original rADF-4 virus-infected cells, in the C52S virus-infected cells only about 50% displayed long fibers, which is probably due to lower initiation efficiency.

The C52R mutant displayed almost no ability to self-assemble into fiber, but rather created abundant sphere-like aggregates

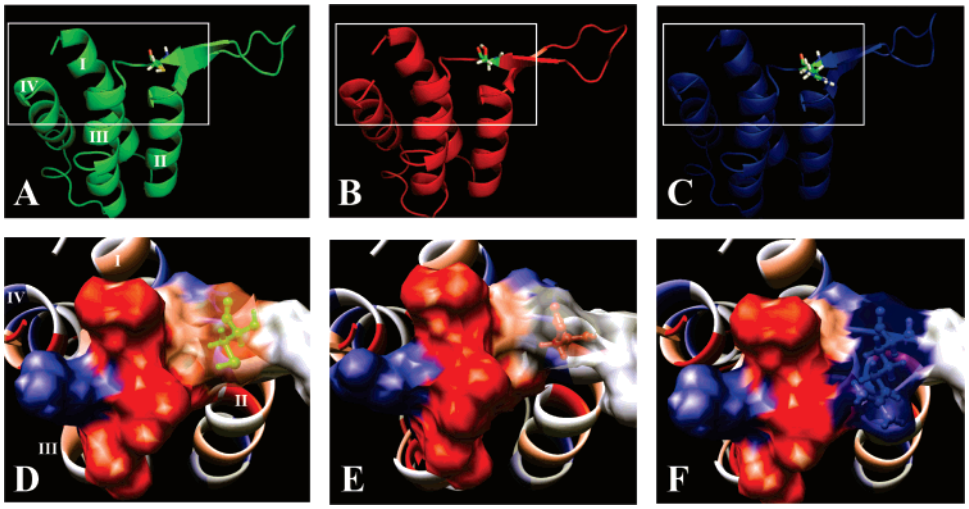

Figure 4. Modeling the structural effect of cysteine mutations. (A) Predicted tertiary structures for the C-terminal domains of the native rADF-4 sequence with the cysteine in ball and stick format. (B) The C52S (the number relates to the C-terminus) mutant. (C) The C52R mutant. (D−F) Surface illustrations of the region containing the cysteine residue, SG**C**DALVQALL (D) and the mutated forms (E,F) corresponding to panels B and C. The colors indicate the relative surface hydrophobicity of the amino acid residues, according to Kyte and Doolittle: Red, hydrophobic; White, neutral; Blue, hydrophilic. The cysteine, serine, and arginine are depicted as semi-transparent to show the backbone in a ball and stick format. All tertiary structures were generated using the ORCHESTRAR homology modeling program.

in all the cells that were infected, most of which were nonsymmetric. These aggregates were different from those of C52S but were likewise insoluble and could be purified out of the expressing cells (Figure 3C,F). Thus, a rather conservative change of cysteine to serine, which nonetheless abolishes the possibility of disulfide bond formation, still enables fiber formation, while a drastic change of cysteine to arginine eliminated fiber formation altogether.

We propose that the conserved cysteine is important for the proper structure of a critical region within the C-terminal domain and thus possibly for the formation of the dragline fiber. Using modeling helps understand the molecular rational for our experimental observations; the natural (Figure 4A), C52S (Figure 4B), and C52R (Figure 4C) ADF-4 C-terminal domain models are shown with the respective residues illustrated. In the C-terminus model, both replacements caused a small global structural change, with backbone rmsd's of 1.07 (C52S) and 0.91 (C52R). The mutations, however, were found to affect the local structure and surface properties of the loop in which they reside.

The cysteine of the original C-terminus is present in the hydrophobic to neutral environment of the highly conserved loop, which we previously identified in the model (Figure 4D). The C52S mutation confers a more negative potential to the loop projecting outward, which is caused by the more negative partial charge of the oxygen atom (compare panel E to panel D in Figure 4), while the overall physical structure (volume and geometry) is retained. However, the C52R mutation places a positive charge within the conserved loop and causes a large steric interference in its vicinity (Figure 4F), which most likely disrupts the conformation of this region that is necessary for fiber organization.

Thus, both our experimental results and modeling support our hypothesis that the conserved cysteine is required for the correct structure of the C-terminal domain of dragline spider silks and that it does not form disulfide bonds in the insect cell system. Although we do not expect that disulfide bonds will form in the reducing environment of the cytoplasm of the insect cells, where self-assembly of the fibers takes place, it was important to show by mutation that indeed these bonds do not occur, since, in some rare circumstances, such bonds can form in the cytosol,[27] and since the cells are highly infected by the virus, which may affect their redox potential or cause disruption of the endoplasmic reticulum compartment. Another support to the claim that the cysteine may not be necessary for fiber assembly is that the *A. ventricosus* MaSp2 contains a serine in the position of the conserved cysteine (Figure 1). In addition, unless we consider the cysteine as a unique hallmark of major ampullate/ampullate silks, the C-terminal fibroin sequences of cribellate spiders do not contain cysteine at this location, which, interestingly, is usually replaced by serine or threonine. In addition, flagelliform silks do contain a cysteine in the domain, but its position is shifted seven residues toward the N-terminus in the alignment and may not play the same role as in ampullate-derived fibers.[18] This said, we cannot negate the possibility that disulfide bond formation may contribute to the spinning process as well as to the superior mechanical attributes of natural spider silk threads.

In this work we have proposed a three-dimensional model for the conserved C-terminal domain of a dragline spider silk protein, in which the number and position of the secondary structures that have been predicted before are different.[18,20] We have shown that the cysteine residue is localized in a structurally conserved loop, which we predict to be essential for the protein−protein interactions of the dragline monomers, a prediction that our mutational analyses support, as fiber formation is abolished when the cysteine of this loop is replaced by the structurally different arginine. Further mutational analyses will help to increase our knowledge of the structure elements of this domain and its interactions, such as the role of the "thumb" structure and, most challenging, defining the interface with the main repetitive region of these interesting and unique proteins.

Acknowledgment. We wish to thank Dr. Tsafi Danieli for help and advice with insect cells and baculovirus work. Many thanks also to Dr. Naomi Melamed-Book for excellent confocal analysis. We also express our gratitude to Yael Ittah and Eli Safrai for their devoted help in figure generation.

References and Notes

(1) Winkler, S.; Kaplan, D. L. *J. Biotechnol.* **2000**, *74*, 85−93.
(2) Vollrath, F.; Knight, D. P. *Nature* **2001**, *410*, 541−548.

(3) Gosline, J. M.; Guerette, P. A.; Ortlepp, C. S.; Savage, K. N. *J. Exp. Biol.* **1999**, *202* (Pt. 23), 3295−3303.
(4) Vollrath, F. *J. Biotechnol.* **2000**, *74*, 67−83.
(5) Hinman, M. B.; Jones, J. A.; Lewis, R. V. *Trends Biotechnol.* **2000**, *18*, 374−379.
(6) Rising, A.; Nimmervoll, H.; Grip, S.; Fernandez-Arias, A.; Storckenfeldt, E.; Knight, D. P.; Vollrath, F.; Engstrom, W. *Zool. Sci.* **2005**, *22*, 273−281.
(7) Bini, E.; Knight, D. P.; Kaplan, D. L. *J. Mol. Biol.* **2004**, *335*, 27−40.
(8) Oroudjev, E.; Soares, J.; Arcdiacono, S.; Thompson, J. B.; Fossey, S. A.; Hansma, H. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (Suppl. 2), 6460−6465.
(9) Simmons, A. H.; Michal, C. A.; Jelinski, L. W. *Science* **1996**, *271*, 84−87.
(10) van Beek, J. D.; Hess, S.; Vollrath, F.; Meier, B. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10266−10271.
(11) Hinman, M. B.; Lewis, R. V. *J. Biol. Chem.* **1992**, *267*, 19320−19324.
(12) Huemmerich, D.; Scheibel, T.; Vollrath, F.; Cohen, S.; Gat, U.; Ittah, S. *Curr. Biol.* **2004**, *14*, 2070−2074.
(13) Gatesy, J.; Hayashi, C.; Motriuk, D.; Woods, J.; Lewis, R. *Science* **2001**, *291*, 2603−2605.
(14) Motriuk-Smith, D.; Smith, A.; Hayashi, C. Y.; Lewis, R. V. *Biomacromolecules* **2005**, *6*, 3152−3159.
(15) Rising, A.; Hjalm, G.; Engstrom, W.; Johansson, J. *Biomacromolecules* **2006**, *7*, 3120−3124.
(16) Beckwitt, R.; Arcidiacono, S. *J. Biol. Chem.* **1994**, *269*, 6661−6663.
(17) Guerette, P. A.; Ginzinger, D. G.; Weber, B. H.; Gosline, J. M. *Science* **1996**, *272*, 112−115.
(18) Challis, R. J.; Goodacre, S. L.; Hewitt, G. M. *Insect Mol. Biol.* **2006**, *15*, 45−56.
(19) Sponner, A.; Unger, E.; Grosse, F.; Weisshart, K. *Biomacromolecules* **2004**, *5*, 840−845.
(20) Sponner, A.; Vater, W.; Rommerskirch, W.; Vollrath, F.; Unger, E.; Grosse, F.; Weisshart, K. *Biochem. Biophys. Res. Commun.* **2005**, *338*, 897−902.
(21) Ittah, S.; Cohen, S.; Garty, S.; Cohn, D.; Gat, U. *Biomacromolecules* **2006**, *7*, 1790−1795.
(22) Shi, J.; Blundell, T. L.; Mizuguchi, K. *J. Mol. Biol.* **2001**, *310*, 243−257.
(23) Davis, I. W.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Nucleic Acids Res.* **2004**, *32*, W615−619.
(24) Sponner, A.; Unger, E.; Grosse, F.; Weisshart, K. *Nat. Mater.* **2005**, *4*, 772−775.
(25) Williams, M. G.; Shirai, H.; Shi, J.; Nagendra, H. G.; Mueller, J.; Mizuguchi, K.; Miguel, R. N.; Lovell, S. C.; Innis, C. A.; Deane, C. M.; Chen, L.; Campillo, N.; Burke, D. F.; Blundell, T. L.; de Bakker, P. I. *Proteins* **2001**, *Suppl. 5*, 92−97.
(26) Lee, J. H.; Yang, S. T.; Rho, S. H.; Im, Y. J.; Kim, S. Y.; Kim, Y. R.; Kim, M. K.; Kang, G. B.; Kim, J. I.; Rhee, J. H.; Eom, S. H. *J. Mol. Biol.* **2006**, *355*, 491−500.
(27) Cumming, R. C.; Andon, N. L.; Haynes, P. A.; Park, M.; Fischer, W. H.; Schubert, D. *J. Biol. Chem.* **2004**, *279*, 21749−21758.

BM7004559