# Novel Assembly Properties of Recombinant Spider Dragline Silk Proteins

Daniel Huemmerich,<sup>1,4</sup> Thomas Scheibel,<sup>1,\*</sup> Fritz Vollrath,<sup>2</sup> Shulamit Cohen,<sup>3</sup> Uri Gat,<sup>3,\*</sup> and Shmulik Ittah<sup>3,4</sup> <sup>1</sup>Department of Organic Chemistry and Biochemistry Technische Universität München 85747 Garching Germany <sup>2</sup>Department of Zoology University of Oxford South Parks Road Oxford OX1 3PS United Kingdom <sup>3</sup>Department of Cell and Animal Biology Silberman Life Sciences Institute Edmond Safra Campus at Givat Ram The Hebrew University Jerusalem 91904 Israel

## Summary

Spider dragline silk, which exhibits extraordinary strength and toughness, is primarily composed of two related proteins that largely consist of repetitive sequences. In most spiders, the repetitive region of one of these proteins is rich in prolines, which are not present in the repetitive region of the other [1]. The absence of prolines in one component was previously speculated to be essential for the thread structure [2]. Here, we analyzed dragline proteins of the garden spider Araneus diadematus, ADF-3 and ADF-4, which are both proline rich, by employing the baculovirus expression system. Whereas ADF-3 represented an intrinsically soluble protein, ADF-4 was insoluble in vitro and selfassembled into filaments in the cytosol of the host insect cells. These ADF-4 filaments displayed the exceptional chemical stability of authentic silk threads. We provide evidence that the observed properties of ADF-3 and ADF-4 strongly depend on intrinsic characteristics such as hydropathicity, which differs dramatically between the two proteins, as in most other pairs of dragline silk proteins from other Araneoidea species, but not on their proline content. Our findings shed new light on the structural components of spider dragline silk, allowing further elucidation of their assembly properties, which may open the door for commercial applications.

## Results

Spider dragline silk has extraordinary properties [3] originating in its composition as a semicrystalline polymer [4] that contains crystalline regions embedded in a less organized "amorphous" matrix. X-ray diffraction and NMR show the crystalline regions to consist of pleated  $\beta$ -sheets of polyalanine stretches that give strength to the thread [5, 6], and the predominant secondary structure of the amorphous matrix is the glycine-rich 3<sub>1</sub>-helix, providing elasticity [7]. Freshly secreted silk proteins are stored at high concentrations [8] as a liquid crystalline spinning dope [9, 10] that is altered by changes in ionic composition, pH (from pH 6.9 to 6.3) [11, 12], and water extraction [12, 13] during its passage through the spinning duct to be finally converted into a solid thread induced by extensional flow [14].

All dragline silks studied so far consist of at least two different proteins with molecular masses of up to several hundred kDa [15]. On the basis of sequence similarities, dragline silk proteins have been grouped into spidroin1like (MaSp1) and spidroin2-like (MaSp2) proteins [1]. Whereas the repetitive regions of the MaSp1 class are essentially proline free, the repetitive regions of MaSp2class proteins usually contain ~15% of proline residues. In most spider species, one protein of each class is involved in forming the dragline silk. It has been suggested that proline-free MaSp1 proteins are responsible for the formation of the crystalline regions, whereas proline-containing MaSp2 proteins are supposed to form the amorphous matrix around the crystals [2]. According to this model, the presence of a proline-free silk protein is essential for the integrity of the silk's semicrystalline structure. However, the dragline silk of Araneus diadematus is composed of the two spidroin2-like proteins ADF-3 and ADF-4. Because this silk displays similar mechanical characteristics when compared to dragline silks of other spider species [3], it can be concluded that two proline-rich proteins can also form structural features that result in the extraordinary physical properties of draglines. The question then arises whether there is another significant difference between ADF-3 and ADF-4, and whether this difference plays an important role during silk assembly or within the final silk structure.

In order to produce and investigate these two proteins, we chose the insect cell line Sf9 (derived from the fall armyworm *Spodoptera frugiperda*) as an expression host because insects belong to the same phylum as spiders and thus seem more suitable than previously used expression systems for producing spider silks. For gene transfer into the cells, baculoviruses containing previously established partial cDNAs of *adf-3* and *adf-4* were generated to produce single dragline silk proteins within the cytosol. Because so far no dragline silk gene has been cloned in its entirety, all previous studies used partial cDNA constructs to produce recombinant silk proteins in bacteria [16] and in mammalian cells [17].

Synthesis of both proteins could be monitored by providing a His<sub>6</sub>-Tag and using anti-His<sub>6</sub> antibodies. Versions without His<sub>6</sub>-Tag were also employed to exclude artificial influences caused by the tag. The recombinant viruses were used to infect Sf9 cells for production of the spider silk proteins. After 3 days of incubation, infected cells were lyzed by sonification, and insoluble

<sup>\*</sup>Correspondence: thomas.scheibel@fiberlab.de (T.S.); gatu@vms. huji.ac.il (U.G.)

<sup>&</sup>lt;sup>4</sup>These authors contributed equally to this work.





(A) Solubility of ADF-3 and ADF-4 after synthesis. Soluble (S) and insoluble components (P) of cell lysates were separated by sedimentation. Proteins were detected by immunoblotting with an anti-His<sub>6</sub> antibody. The sizes correlate well with the calculated MW of the constructs containing the His<sub>6</sub>-Tags. A minor lower band in the ADF-4 lane could be a product of premature termination.

(B) Filament in *adf-4*-expressing cell, as seen with light microscopy (upper panel), with regular fluorescence microscopy, or with confocal microscopy after immunocytochemistry (middle and lower panel, respectively). The scale bar represents 10  $\mu$ m.

(C) Filament in a cell synthesizing ADF-4 without the  $\text{His}_{\text{e}}\text{-}\text{Tag}.$  The scale bar represents 10  $\mu\text{m}.$ 

(D) Filaments of *adf-4*-expressing cells were purified, dissolved, and analyzed by SDS-PAGE and then by silver staining (SS). ADF-4 was detected by immunoblotting with an anti-His<sub>6</sub> antibody (IB).

(E) Solubility of cosynthesized ADF-3 and ADF-4. Soluble (S) were separated from insoluble (P) cell components by sedimentation. ADF-3 was detected with S-protein peroxidase conjugates after Western blotting, and ADF-4 was detected with anti-T7-tag antibodies.

cell contents were separated from soluble material by sedimentation. The sediment was dissolved in guanidinium thiocyanate (GdmSCN) before analysis by immunoblotting. Whereas a large fraction of ADF-3 was found to be soluble, ADF-4 was almost entirely insoluble 3 days after infection under the conditions employed (Figure 1A). For both proteins, few degradation products or smaller protein fragments could be detected, indicating the suitability of our expression system, which does not lead to translational pauses typical of some other expression systems [18]. Surprisingly, investigating the aggregates in adf-4-expressing cells revealed filaments that coiled throughout the cytoplasm, whereby most of the cells contained only one or few filaments (Figures 1B and 1C). In contrast, cells infected with control viruses or the adf-3 encoding virus never produced such filaments. Immunofluorescence performed on the infected cells with anti-His<sub>6</sub> antibodies showed specific staining of the filaments, thus confirming that the filaments were composed of ADF-4 (Figure 1B). ADF-4 not displaying a His<sub>6</sub>-Tag also formed filaments, indicating that fibrillization was not influenced by the His<sub>6</sub>-Tag (Figure 1C). Purifying and analyzing the filaments by SDS-PAGE revealed that the filaments consisted of one major protein



Figure 2. Morphology of Filaments and Aggregates

Purified filaments of *adf*-4-expressing cells were investigated by (A) scanning electron microscopy (SEM) (the scale bar represents 5  $\mu$ m); (B) transmission electron microscopy (TEM) (the scale bar represents 500 nm); (C) immunoelectron microscopy with mouse anti-His<sub>6</sub> antibodies and gold-conjugated anti-mouse antibodies (the scale bar represents 500 nm); (D) atomic force microscopy (AFM) (height image; the height of the filament is 0.7  $\mu$ m); and (E) AFM (deflection image; the scale bar represents 5  $\mu$ m).

(F) Purified filaments of *adf*-3- and *adf*-4-expressing cells; SEM; the scale bar represents 5  $\mu$ m.

(G) ADF-4 aggregate formed in vitro; SEM; the scale bar represents  $5\mu m.$ 

component, which could be identified to be ADF-4 by immunoblotting (Figure 1D).

Next, we investigated whether ADF-3 and ADF-4 can coassemble into filaments. We generated a recombinant baculovirus containing both *adf-3* and *adf-4* under the control of the independent p10 and polyhedrin promoters with the pFastbacDUAL donor plasmid. Infection of Sf9 cells with this virus resulted in synthesis of both proteins. Interestingly, ADF-3 again was entirely soluble, whereas ADF-4 was only found in the insoluble fraction, indicating that at this time of infection there was no stable interaction between these two proteins within the cytosol (Figure 1E).

Morphology of purified ADF-4 filaments was characterized with scanning electron microscopy (SEM), atomic force microscopy (AFM), transmission electron microscopy (TEM), and immunogold labeling. The diameters of filaments ranged from 200 nm to 1  $\mu$ m; however, the diameter was found to be constant for each single filament. Furthermore, the filaments showed lengths up to 100  $\mu$ m and often terminated in knots or branches or formed closed circles (Figures 2A–2F). Filaments displayed a smooth surface and were often associated with nanofibers (diameter  $\sim$ 5 nm) and protein aggregates (Figures 2A and 2B). Immunoelectron microscopy illus-

trated the presence of ADF-4 within the filaments (Figure 2C). From cells cosynthesizing ADF-3 and ADF-4 in the cytosol, filaments could be isolated that revealed a similar morphology when compared to filaments formed by synthesis of ADF-4 alone (Figure 2F). The low number of filaments per cell and the recruitment of almost the entire cellular ADF-4 into the aggregates indicated that self-assembly of ADF-4 in Sf9 cells is likely to be a nucleated process, which previously has been also suggested for the silk spinning process of *Bombyx mori* [19].

To investigate whether filament formation is an intrinsic property of ADF-4, we analyzed its self-assembly properties in vitro. Soluble ADF-4 was readily obtained by dissolving filaments in 6 M GdmSCN. Dissolved ADF-4 rapidly aggregated upon removal of GdmSCN by dialysis or dilution in 10 mM Tris (pH 8.0). However, the ADF-4 aggregates formed in vitro did not show fibrillar structures (Figure 2G), indicating that folding and assembly of ADF-4 requires factors and conditions that are present in the cytosol of the insect cells. Because we have been able to self-assemble other recombinant spider silk proteins into fibrillar structures after chemical denaturation in vitro (data not shown), determining the cellular factors involved in ADF-4 folding and assembly in vivo might lead to successful ADF-4 filament formation in vitro.

The length of the filaments formed in the Sf9 cells seemed to be constrained by the volume of the cells, making them too short for mechanical force measurements typically performed with silk threads [17]. However, we were able to analyze the chemical stability of ADF-4 filaments and aggregates formed in vitro in comparison to natural dragline silk threads of A. diadematus. Dragline threads have been reported to be insoluble in many denaturing agents [20]. Application of 2% sodium dodecylsulfate (SDS) and 8 M urea readily dissolved ADF-4 aggregates formed in vitro (Figure 3A) but apparently had no effect on the structure of ADF-4 filaments and dragline threads after 30 s of exposure (Figures 3B and 3C). Immersion of the filaments in 6 M guanidinium chloride (GdmCl) did not lead to solubilization of either ADF-4 filaments or dragline threads, although it did lead to swelling of dragline silk. Such swelling is likely caused by fiber supercontraction [17], which has previously been described for spider silks immersed in aqueous solutions and which results from reformation of hydrogen bonds in the amorphous matrix [21]. In contrast to the denaturants mentioned above, a small drop of 6 M GdmSCN completely dissolved ADF-4 filaments as well as dragline threads within seconds (Figures 3B and 3C). In consequence, we conclude that both structures share molecular interactions that are responsible for chemical resistance to specific denaturants.

## Discussion

Thus far, little is known about the structure, function, and possible interplay between protein components of spider dragline silk threads. We observed that despite their similar proline content, ADF-3 and ADF-4, which represent the repetitive parts and carboxyl-termini of



Figure 3. Chemical Stability of ADF-4 Aggregates, Filaments, and Dragline Silk Threads

(A) ADF-4 aggregates formed in vitro and (B) ADF-4 filaments were treated with denaturants as indicated. Dissolved ADF-4 was detected by immunoblotting with an anti-His<sub>6</sub> antibody.

(C) ADF-4 filaments and dragline silk threads were treated with denaturants as indicated and examined by light microscopy afterwards; the scale bar represents 25  $\mu$ m.

the proteins involved in forming the dragline thread of *A. diadematus*, display surprisingly different properties. ADF-3 was soluble within the cytosol of the insect cells and has been shown to be highly soluble even at high protein concentrations in vitro [22]. Whereas ADF-3 thus represents an intrinsically soluble protein, ADF-4 is virtually insoluble under our experimental conditions.

Because thread formation has to be fast at natural reeling speeds of 1–10 cm/s [23], an easily assembling compound, such as ADF-4, is mandatory for silk formation. However, the tendency of ADF-4 to aggregate implies that other factors within the dope are likely required to keep it from premature polymerization in the gland. Although ADF-3 did not influence solubility of ADF-4 within the cytosol of Sf9 cells, we presume that during or after secretion, the two proteins interact in a way that could not be assessed by the methods shown in this work, thus influencing each other's solubility and assembly.

The different solubilities of ADF-3 and ADF-4 can be explained by the overall hydrophobicities of the two proteins (Table 1). The more hydrophilic ADF-3 interacts favorably with the aqueous solvent and thus remains soluble under most conditions. In contrast, the more hydrophobic ADF-4 favors interactions with other protein molecules and thus tends to aggregate. Experiments with synthetic spider silk proteins have shown that proteins with similar hydrophobicity compared to that of ADF-3 and ADF-4 display a similar solubility [24].

Interestingly, all pairs of dragline silk proteins from different spider species display a common distinct distribution of hydrophobicity and charge. MaSp1/ADF-4

Table 1. Biophysical Characteristics of Dragline Silk Proteins							
	A.d.	<i>A.a</i> .	<i>A.t</i> .	<i>L.g.</i>	<i>N.c.</i>	<i>N.m</i> .	<i>N.s.</i>
	ADF-4/ADF-3	MaSp1/MaSp2	MaSp1/MaSp2	MaSp1/MaSp2	MaSp1/MaSp2	MaSp1/Masp2	MaSp1/MaSp2
Repeat Units	8/12	8/5	14/7	8/11	16/11	3/9	5/5
Hydropathicity	-0.3/-0.9	-0.0/-0.5	-0.1/-0.7	-0.1/-0.2	-0.0/-0.6	-0.1/-0.7	0.1/-0.2
Relative Charge (%)	1.0/0.0	2.0/0.9	1.9/0.9	1.7/0.9	2.5/0.7	1.9/0.3	2.4/0.9

The maximal number (given in the table) of complete repeat units extending from one polyalanine sequence to the next were analyzed with published sequences. Hydropathicity was calculated according to Kyte and Doolittle [25]. The higher the values of hydropathicity, the more hydrophobic the protein. The relative charge is given as the percentage of all charged amino acids. The following abbreviations were used: *A.d., Araneus diadematus; A.a., Argiope aurantia; A.t., Argiope trifasciata; L.g., Latrodectus geometricus; N.c., Nephila clavipes; N.m., Nephila madagascariensis; and N.s., Nephila senegalensis.* 

proteins generally display relatively high hydrophobicity and at least 1% charged residues, whereas the corresponding MaSp2/ADF-3 partner protein is more hydrophilic, with less than 1% charge (Table 1). Although there are variations in absolute values of these parameters, and the differences between two protein components vary between spider species, there is a clear general tendency, which indicates that spider dragline silks display an underlying universality common to *Araneoidea* (orb weaving) species.

## Conclusions

In this study we have employed the baculovirus expression system to efficiently produce dragline silk components. We have compared similarly sized parts of the two A. diadematus dragline silk proteins, which showed surprisingly different assembly properties, in that ADF-3 is much more soluble than its counterpart, ADF-4, which formed insoluble fibers in the cytoplasm of the host insect cells. Strikingly, these fibers displayed the chemical resilience typical of native dragline silk. On the basis of analysis of available dragline spider silk protein sequences from different Araneoidea species, we conclude that the repetitive regions of the two components of dragline silks differ in their overall hydrophobic nature and not necessarily in their proline content, which corresponds well to our experimental findings on the behavior of recombinant ADF-3 and ADF-4. Because one component can form stable fibers on its own, the question of the mechanistic and structural role of the second protein arises and has to be investigated in the future. Solving this question will provide a further step toward commercially using recombinantly produced spider silks as new materials. Such knowledge is required for the spinning of silk threads from recombinant proteins and for the manufacturing of a new generation of fibrous biomaterials, which may be based on the natural sequences or be engineered for selected purposes.

## **Experimental Procedures**

#### Detection and Solubility of ADF-3 and ADF-4

Cells were resuspended at  $1.2 \times 10^7$  cells/ml in 100 mM NaCl and 20 mM *N*-(2-hydroxyethyl)piperazine-*N'*-(2-ethanesulfonic acid) (HEPES) (pH 7.5) and lyzed by sonification. Soluble and insoluble components were separated by centrifugation at 125,000  $\times$  g for 30 min. For further analysis, pellets were resuspended in 6 M GdmSCN and dialyzed against 8 M urea. Supernatant and pellet derived from  $1.5 \times 10^5$  cells were loaded on 10% Tris-glycine sodium dodecylsulfate polyacrylamide gels under reducing conditions and blotted onto polyvinylidene fluoride (PVDF) membranes (Millipore).

Spider silk proteins were detected with a mouse anti-His<sub>6</sub> monoclonal antibody (Sigma-Aldrich, 1:10,000) or a mouse anti-T7 monoclonal antibody (Novagen, 1:10,000) and anti-mouse IgG peroxidase conjugate (Sigma-Aldrich, 1:5,000) as secondary antibody. An S-Protein peroxidase conjugate (Novagen, 1:5,000) was used to directly detect S-tagged ADF-3.

#### **ADF-4 Thread Purification**

Cells were resuspended at  $1.2\times10^7$  cells/ml in 100 mM NaCl and 20 mM HEPES (pH 7.5) and lyzed by adding 2% w/v sodium dodecyl-sulfate and then incubating at 95°C for 5 min. Threads were sedimented at 5000  $\times$  g and then washed with 8 M urea and water\_bidest-

## Testing Chemical Stability of ADF-4 Aggregates, Filaments, and Dragline Silk

ADF-4 aggregates formed in vitro and purified filaments of  $5 \times 10^5$  cells were resuspended in 0.5 ml denaturant. Undissolved material was removed by centrifugation at 125,000  $\times$  g for 30 min. The supernatant was dialyzed against 8 M urea and analyzed by immunoblotting (see above). Alternatively purified filaments loaded on mica as described for AFM analysis (see Supplemental Data) and dragline silk threads mounted on polypropylene were incubated for 30 s with ~0.1 µl of denaturant. After being rinsed with water, samples were examined by light microscopy.

#### Supplemental Data

Supplemental Experimental Procedures are available at http://www.current-biology.com/cig/content/full/14/22/2070/DC1/.

#### Acknowledgments

We thank John Gosline and Paul Guerette for kindly providing clones of *adf-3* and *adf-4*, Bettina Richter for technical assistance, and Alexander Sponner and Michael Wise for critical comments on the manuscript. Special thanks to Naomi Melamed-Book for confocal microscopy/imaging and to Dr. Tsafi Danieli for help with baculovirus techniques. This work was supported by the Bundesministerium für Bildung und Forschung (D.H.), Fonds der Chemischen Industrie (D.H. and T.S.), and the Deutsche Forschungsgemeinschaft (T.S.).

Received: August 30, 2004 Revised: September 30, 2004 Accepted: September 30, 2004 Published: November 23, 2004

#### References

- Gatesy, J., Hayashi, C., Motriuk, D., Woods, J., and Lewis, R. (2001). Extreme diversity, conservation, and convergence of spider silk fibroin sequences. Science 291, 2603–2605.
- Thiel, B.L., Guess, K.B., and Viney, C. (1997). Non-periodic lattice crystals in the hierarchical microstructure of spider (major ampullate) silk. Biopolymers 41, 703–719.
- Gosline, J.M., Guerette, P.A., Ortlepp, C.S., and Savage, K.N. (1999). The mechanical design of spider silks: From fibroin sequence to mechanical function. J. Exp. Biol. 202, 3295–3303.

- Warwicker, J. (1960). Comparative studies of fibroins. II. The crystal structures of various fibroins. J. Mol. Biol. 2, 350–362.
- Simmons, A.H., Ray, E., and Jelinski, L.W. (1994). Solid-state <sup>13</sup>C NMR of *Nephila clavipes* dragline silk establishes structure and identity of crystalline regions. Macromolecules 27, 5235– 5237.
- Parkhe, A.D., Seeley, S.K., Gardner, K., Thompson, L., and Lewis, R.V. (1997). Structural studies of spider silk proteins in the fiber. J. Mol. Recognit. *10*, 1–6.
- van Beek, J.D., Hess, S., Vollrath, F., and Meier, B.H. (2002). The molecular structure of spider dragline silk: Folding and orientation of the protein backbone. Proc. Natl. Acad. Sci. USA 99, 10266–10271.
- Hijirida, D.H., Do, K.G., Michal, C., Wong, S., Zax, D., and Jelinski, L.W. (1996). 13C NMR of Nephila clavipes major ampullate silk gland. Biophys. J. 71, 3442–3447.
- Kerkam, K., Viney, C., Kaplan, D., and Lombardi, S. (1991). Liquid crystallinity of natural silk secretions. Nature 349, 596–598.
- Knight, D.P., and Vollrath, F. (1999). Liquid crystals and flow elongation in a spider's silk production line. Proc. R. Soc. Lond. B 266, 519–523.
- Knight, D.P., and Vollrath, F. (2001). Changes in element composition along the spinning duct in a Nephila spider. Naturwissenschaften 88, 179–182.
- Vollrath, F., Knight, D., and Hu, X.W. (1998). Silk production in a spider involves acid bath treatment. Proc. R. Soc. Lond. B. Biol. Sci. 265, 817–820.
- Tillinghast, E.K., Chase, S.F., and Townley, M.A. (1984). Water extraction by the major ampullate duct during silk formation in the spider, *Argiope aurantia* Lucas. J. Insect Physiol. 30, 591–596.
- Knight, D.P., Knight, M.M., and Vollrath, F. (2000). Beta transition and stress-induced phase separation in the spinning of spider dragline silk. Int. J. Biol. Macromol. 27, 205–210.
- Winkler, S., and Kaplan, D.L. (2000). Molecular biology of spider silk. J. Biotechnol. 74, 85–93.
- Arcidiacono, S., Mello, C., Kaplan, D., Cheley, S., and Bayley, H. (1998). Purification and characterization of recombinant spider silk expressed in Escherichia coli. Appl. Microbiol. Biotechnol. 49, 31–38.
- Shao, Z., Young, R.J., and Vollrath, F. (1999). The effect of solvents on spider silk studied by mechanical testing and single-fibre Raman spectroscopy. Int. J. Biol. Macromol. 24, 295–300.
- Fahnestock, S.R., and Irwin, S.L. (1997). Synthetic spider dragline silk proteins and their production in Escherichia coli. Appl. Microbiol. Biotechnol. 47, 23–32.
- Li, G., Zhou, P., Shao, Z., Xie, X., Chen, X., Wang, H., Chunyu, L., and Yu, T. (2001). The natural silk spinning process. Eur. J. Biochem. 268, 6600–6606.
- Lombardi, S., and Kaplan, D. (1990). The amino acid composition of major ampullate gland silk (dragline) of *Nephila clavipes* (araneae, tetragnathidae). J. Arachnol. *18*, 297–306.
- Shao, Z.Z., and Vollrath, F. (1999). The effect of solvents on the contraction and mechanical properties of spider silk. Polym. 40, 1799–1806.
- Lazaris, A., Arcidiacono, S., Huang, Y., Zhou, J.F., Duguay, F., Chretien, N., Welsh, E.A., Soares, J.W., and Karatzas, C.N. (2002). Spider silk fibers spun from soluble recombinant silk produced in mammalian cells. Science 295, 472–476.
- Carmichael, S., and Viney, C. (1999). Molecular order in spider major ampullate silk (dragline): Effects of spinning rate and postspin drawing. J. Appl. Polym. Sci. 72, 895–903.
- Huemmerich, D., Helsen, C.W., Oschmann, J., Rudolph, R., and Scheibel, T. (2004). Primary structure elements of dragline silks and their contribution to protein solubility and assembly. Biochemistry 43, 13604–13612.
- Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105–132.