The Role of Alanine Sequences in Forming β-Sheets of Spider Dragline Silk

Erik J. Spek, Huai-Chuan Wu, and Neville R. Kallenbach*

Department of Chemistry, New York University
New York, New York 10003

Received September 13, 1996

Spider dragline silk is a protein biomaterial endowed with an exceptional combination of strength and elasticity.[1] The tensile strength and its high elongation-to-break ratio are comparable to the strongest synthetic macromolecules. The dragline fiber forms both the frame of the web of orb-spinning spiders and the dragline itself.[2] X-ray analysis reveals that dragline silk fibers consist of a mixture of crystalline (β-sheet) and amorphous regions. Xu and Lewis isolated cDNA's encoding dragline silk sequences, which allowed them to identify two fibroin-like proteins, MaSp 1 and MaSp 2.[3] The amino acid sequences contain extensive repeats of a motif in which short runs of Ala_n ($n = 4-7$) are interspersed between glycine-rich segments. Several models have been proposed to identify the side chains that are involved in the crystalline and amorphous regions.[4] Since alanine has the highest tendency among the natural amino acids to form a helix, one proposal assigned the glycine regions to the β-sheet.[5] However, the 5.3 Å spacing seen in fiber diffraction patterns is consistent with an alanine layer, seen in synthetic sheets formed from oligomers of alanine.[4] Solid state ^{13}C NMR spectroscopy on the fibers indicates that essentially all of the alanines have a β-sheet conformation.[6] Relaxation studies suggest that the alanines in dragline silk comprise two dynamic subpopulations. While both subpopulations are β-sheet, 40% are thought to be highly oriented, while 60% are less ordered.[7] These data have led to a revised picture in which stretches of polyalanine comprise the core of the crystalline region, while the glycine-rich sequences make up the amorphous region.

Peptide model studies on consensus sequence fragments of spider dragline silk have attempted to address structural issues more directly.[4] Three synthetic peptides studied so far (P15, P30, and P47) correspond respectively to a glycine-rich repeat from MaSp 1, an entire glycine- and alanine-rich repeat from MaSp 1, and a longer repeat unit from MaSp 2. None of these peptides exhibits β-sheet structure at low temperature and concentrations (<1 mg/mL). The shorter peptides P15 and P30 can be induced to form β-sheets at temperatures higher than 40 °C. At concentrations above 10 mg/mL, P30 also forms a precipitate with the FT-IR spectrum of a β-sheet. In this paper, we have adopted a minimalist approach and synthesized two short peptide fragments with consensus sequences taken from MaSp 1 of *Nephila clavipes* (Figure 1). The sequences consist of stretches of four (A4) and seven (A7) alanines flanked by GAG and GGAG termini, truncated versions of the glycine-rich repeats in this protein. At the C-terminus, the sequence GGY was added to facilitate concentration determination.

A4: GAG-AAAA-GGAG-GGY-NH_2

A7: GAG-AAAAAAA-GGAG-GGY-NH_2

Figure 1. Consensus sequence of dragline silk of *N. clavipes*.

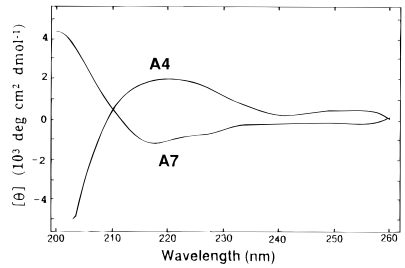

Figure 2. Circular dichroism spectra of A4 and A7 in H_2O at 4 °C. The concentrations are 100 μM as determined by absorbance.

Peptide Synthesis. The peptides were prepared by using solid phase synthesis on a Milligen/Biosearch 9600 synthesizer. Fmoc protection was employed for the α-function of amino acids. PAL resin was used as the solid support. The coupling was performed by using the BOP/HOBT combination. Cleavage and deprotection were effected by treatment with TFA reagent (90% trifluoroacetic acid, 5% thioanisole, 3% ethanediol, and 2% anisole) at room temperature for 2 h and precipitated with ether. Crude peptides were purified by reverse phase preparative HPLC on a Delta-Pak 15 μM-C18 column (300 mm × 7.8 mm) monitored at 270 nm with a gradient of 0−40% acetonitrile in 0.1% trifluoroacetic acid.

Mass Spectrometry. The primary ion molecular weights of the purified peptides were confirmed by MALDI mass spectrometry (MALDI TOF, Shimadzu). For A4: M + H obsd at *m/z* 1006, calcd 1009. For A7: M + H obsd at *m/z* 1223, calcd 1219.

Circular Dichroism Measurements. Circular dichroism spectra were run on a modified Cary 60 spectropolarimeter (Aviv DS 60) equipped with an HP Model 89100A temperature controller. The wavelength was calibrated by use of a (+)-10-camphorsulfonic acid standard. The concentration of stock peptide was determined by using the absorption of Tyr in 6 M guanidine hydrochloride at 275 nm as reference.[8] Measurements were carried out in water using 1 mm path length cells. The concentrations in the experiments were 10 μM to 2 mM.

FT-IR Spectroscopy. Spectra were recorded by the Michigan Molecular Institute using a Magan 550 instrument at 4 cm^{-1} resolution averaging over 100 scans. Samples were ca. 1 mM in D_2O as determined by weight.

Circular dichroism measurements on A7 show a spectrum with a minimum at 215 nm, consistent with a β-sheet−coil mix (Figure 2). This spectrum is observed at all concentrations measured, from 10 μM to 2 mM. Deconvolution of the spectrum indicates that it consists of 67% β-sheet with the remainder random coil.[9] The spectrum of A7 is independent of concentration in the range from 10 μM to 1 mM. Above this value, the intensity of the β-sheet signal increases with concentration, indicating association. Above 2 mM concentrations, a precipitate forms. This behavior precludes use of NMR to characterize the sheet structure in detail. Temperature measurements indicate that the β-sheet structure is maintained up to 70 °C. Above this temperature, a random coil spectrum is detected (Figure 3). The thermal denaturing process is irreversible, and the soluble β-sheet form could not be recovered at lower temperatures. Measurements on A4 show a spectrum resembling a random coil for all measured concentrations (up

(1) Termonia, Y. *Macromolecules* **1994**, *27*, 7378.
(2) Vollrath, F. *Sci. Am.* **1992**, *March*, 70.
(3) (a) Xu, M.; Lewis, R. V. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 7120. (b) Hinman, M. B.; Lewis, R. V. *J. Biol. Chem.* **1992**, *267*, 19320.
(4) Lewis, R. V. *Acc. Chem. Res.* **1992**, *25*, 392.
(5) (a) Lyu, P. C.; Liff, M. I.; Marky, L. A.; Kallenbach, N. R. *Science* **1990**, *250*, 669. (b) O'Neal, K. T.; DeGrado, W. F. *Science* **1990**, *250*, 646. (c) Chakrabartty, A.; Kortemme, T.; Baldwin, R. L. *Protein Sci.* **1994**, *3*, 843.
(6) Simmons, A. M.; Ray, E.; Jelinski, L. W. *Macromolecules* **1994**, *27*, 5225.
(7) Simmons, A. M.; Michal, C. A.; Jelinski, L. W. *Science* **1996**, *271*, 84.
(8) Brandts, J. F.; Kaplan, L. *Biochemistry* **1973**, *12*, 2011.
(9) (a) Greenfield, N.; Fasman, G. D. *Biochemistry* **1969**, *8*, 4108. (b) Yang, J. T.; Wu, C. S.; Martinez, H. M. *Methods Enzymol.* **1986**, *130*, 228.

S0002-7863(96)03218-0 CCC: $14.00

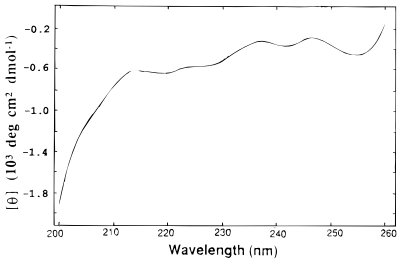

Figure 3. Circular dichroism spectrum of A7 in H_2O at 80 °C. The concentration is 100 μM as determined by absorbance.

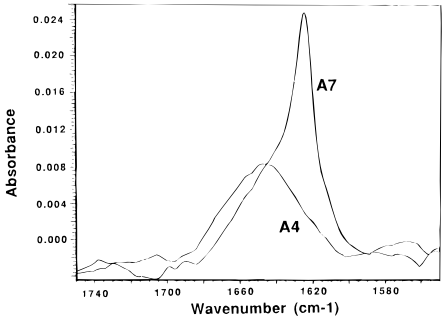

Figure 4. FT-IR spectra of A4 and A7 in D_2O at 4 °C. Sample concentrations were ca. 1 mM as determined by weight.

to 5 mM) and temperatures up to 70 °C. Assignment of A4 to coil and A7 to β-sheet is confirmed by FT-IR spectral analysis using multiple internal reflection to enhance path length (Figure 4). The spectrum shows a band for A7 at 1625 cm^{-1}, in the range that corresponds for the amide stretch vibration of a peptide in β-sheet formation.[10]

At this stage, we do not have high-resolution structural data on the β-sheets. It seems unlikely that these peptides do not aggregate at high concentrations. This would imply formation in A7 of an intramolecular β-sheet with a very unlikely structure in which the alanines would have to make up the β-turns. A more plausible model could be a β-sheet buildup from stretches of alanine, with the glycines accounting for the turns in the polypeptide. This model would predict concentration dependence, and we observe this at concentrations above 2 mM. At lower concentrations, the dependence upon concentration is weaker and the CD spectra appear invariant. This is consistent with initial formation of a tight nucleating complex which propagates at the higher concentrations. This model can also account for the fact that all alanines are in β-sheet conformation as observed by the solid state NMR experiments.[6,7] Smaller strands of alanines (A4 for example) that are incapable of forming β-sheets on their own could assemble on a β-sheet core formed by longer alanine sequences.

It is interesting in this connection that a second group of proteins containing stretches of seven adjacent alanines are the antifreeze polypeptides found in arctic fish.[11] In contrast to the A7 peptide studied here, these peptides are highly α-helical. This striking difference in structure between these molecules can be accounted for by the presence in the antifreeze proteins of charged and polar side chains at key positions that favor helical formation. There is a strong capping box at the N-terminus[12] as well as an i, $i + 4$ salt bridge in several of these sequences. By contrast, A7 contains only alanines and glycines, which adopt a β-sheet or β-turn formation as observed in our experiments. This confirms the idea that intrinsic helical propensity is not the major factor in secondary structure formation.

The study of A4 and A7 provides the first direct evidence that the longer stretches of seven alanines account in principle for the β-sheet in the MaSp 1, while the shorter runs of four alanines alone do not. A buildup of the polypeptide strand by accretion of stretches of alanines with different length, nucleated by A7 runs, could account for the crystalline region of the dragline silk.

Acknowledgment. This work was supported by grant GM 40746 from the NIH and a grant from the Human Frontiers of Science Program. We thank Dr. Wood at the Michigan Molecular Institute for measuring the FT-IR spectra

JA963218N

(10) Cantor, C. R.; Schimmel, P. R. In *Biophysical Chemistry*; W. H. Freeman: New York, 1980; Part II, p 468. The A4 spectrum shown is in the range of α-helical frequencies, but the signal is weak and not consistent with the CD, which is a more sensitive indicator of peptide conformation.

(11) Sicheri, F.; Yang, D. S. C. *Nature* **1995**, *375*, 427.
(12) Harper, E. T.; Rose, G. D. *Biochemistry* **1993**, *32*, 7605.