Surface Properties and Conformation of Nephila clavipes Spider Recombinant Silk Proteins at the Air−Water Interface

ARTICLE *in* LANGMUIR · MAY 2009

Impact Factor: 4.46 · DOI: 10.1021/la900475q · Source: PubMed

8 AUTHORS, INCLUDING:

Thierry Lefèvre
Laval University
54 PUBLICATIONS **1,184** CITATIONS
SEE PROFILE

Sylvie Beaufils
Université de Rennes 1
64 PUBLICATIONS **355** CITATIONS
SEE PROFILE

Véronique Vié
Université de Rennes 1
122 PUBLICATIONS **1,140** CITATIONS
SEE PROFILE

Langmuir
Article
pubs.acs.org/Langmuir

Surface Properties and Conformation of *Nephila clavipes* Spider Recombinant Silk Proteins at the Air−Water Interface

Anne Renault,[‡] Jean-François Rioux-Dubé,[†] Thierry Lefèvre,[†] Stéphane Pezennec,[§] Sylvie Beaufils,[‡] Véronique Vié,[‡] Mélanie Tremblay,[†] and Michel Pézolet*,[†]

[†]*Centre de recherche sur les matériaux avancés, Département de chimie, Université Laval, Québec, Canada G1 V 0A6,* [‡]*Institut de Physique de Rennes, UMR CNRS 6251, Université de Rennes, Bat 11A Campus de Beaulieu, 35042 Rennes Cedex, France, and* [§]*Science et Technologie du Lait et de L'Oeuf, INRA, 35042 Rennes cedex, France*

Received February 8, 2009. Revised Manuscript Received March 19, 2009

The dragline fiber of spiders is composed of two proteins, the major ampullate spidroins I and II (MaSpI and MaSpII). To better understand the assembly mechanism and the properties of these proteins, the adsorption behavior of the recombinant proteins of the spider *Nephila clavipes* produced by Nexia Biotechnologies Inc. has been studied at the air−water interface using ellipsometry, surface pressure, rheological, and infrared measurements. The results show that the adsorption is more rapid and more molecules are present at the interface for MaSpII than for MaSpI. MaSpII has thus a higher affinity for the interface than MaSpI, which is consistent with its higher aggregation propensity in water. The films formed at the interface consist of networks containing a high content of intermolecular β-sheets as revealed by the *in situ* polarization modulation infrared absorption reflection spectra. The infrared results further demonstrate that, for MaSpI, the β-sheets are formed as soon as the proteins adsorb to the interface while for MaSpII the β-sheet formation occurs more slowly. The amount of β-sheets is lower for MaSpII than for MaSpI, most likely due to the presence of proline residues in its sequence. Both proteins form elastic films, but they are heterogeneous for MaSpI and homogeneous for MaSpII most probably as a result of a more ordered and slower aggregation process for MaSpII. This difference in their mechanism of assembly and interfacial behaviors does not seem to arise from their overall hydrophobicity or from a specific pattern of hydrophobicity, but rather from the longer polyalanine motifs, lower glycine content, and higher proline content of MaSpII. The propensity of both spidroins to form β-sheets, especially the polyalanine blocks, suggests the participation of both proteins in the silk's β-sheet crystallites.

Introduction

Spider silks are proteinaceous filaments that display a diversity of amino acid compositions, functions, and mechanical properties.[1−4] Although the protein composition has not yet been determined for all types of spider silk, several threads spun by *Ananeoidea* species seem to be constituted of different components. The minor ampullate and, above all, the major ampullate (MA) thread are well-characterized examples that have been shown to be composed of two proteins.[5−8] The MA silk, that forms the dragline and the radii and the frame of the web, is composed of a mixture of two spidroins named MaSpI and MaSpII (MaSp stands for major ampullate spidroin). For *Nephila clavipes* spiders, MaSpI accounts for 80% of the fiber.[8−10] Both proteins are rich in alanine and glycine residues and share the same repeat pattern, with a hard sequence composed of polyalanine runs and a soft sequence rich in glycine residues. It is well accepted that, due to this block copolymer structure, the MA silk is a semicrystalline biopolymer with amorphous flexible chains reinforced by stiff crystallites.[11,12] The crystalline regions, made of the alanine segments arranged in antiparallel β-sheets, are responsible for the silk tensile strength. Regions rich in glycine residues form the amorphous domains that give its extensibility to silk. The structure of the polypeptide backbone in the amorphous domains is mainly disordered but seems to contain 3_1-helices (also known as polyproline II helices)[13] and β-turns.[14]

MaSpI and MaSpII display some differences in their amino acid composition. First, the polyalanine sequences of MaSpI are shorter in average than those of MaSpII. Second, whereas the soft sequence of MaSpI contains the tripeptide GGX (X = Q, Y, L, or R), it rather consists of the pentapeptides GPGQQ and GPGGY for MaSpII. Third, MaSpII is particularly rich in proline residues (15% of the total amino acids), whereas MaSpI is almost devoid of proline. Finally, MaSpI contains more glycine residues than MaSpII (42% for MaSpI, 31% for MaSpII). These structural specificities are likely to confer a specific role to each protein in the structure and mechanical properties of MA silk. Indeed, the spatial distribution of MaSpI and MaSpII in the fiber is different, with MaSpI being uniformly distributed throughout the thread while MaSpII is concentrated in an inner area of the fiber's core.[15]

(1) Gosline, J. M.; Guerette, P. A.; Ortlepp, C. S.; Savage, K. N. *J. Exp. Biol.* **1999**, *202*, 3295–3303.
(2) Hu, X.; Vasanthavada, K.; Kohler, K.; McNary, S.; Moore, A. M. F.; Vierra, C. A. *Cell. Mol. Life Sci.* **2006**, *63*, 1986–1999.
(3) Lewis, R. V. *Chem. Rev.* **2006**, *106*, 3762–3774.
(4) Vollrath, F. *Rev. Mol. Biotechnol.* **2000**, *74*, 67–83.
(5) Colgin, M. A.; Lewis, R. V. *Protein Sci.* **1998**, *7*, 667–672.
(6) Gatesy, J.; Hayashi, C.; Motriuk, D.; Woods, J.; Lewis, R. *Science* **2001**, *291* (5513), 2603–2605.
(7) Guerette, P. A.; Ginzinger, D. G.; Weber, B. H. F.; Gosline, J. M. *Science* **1996**, *272*, 112–115.
(8) Hinman, M. B.; Lewis, R. V. *J. Biol. Chem.* **1992**, *267*(27), 19320–19324.
(9) Brooks, A. E.; Steinkraus, H. B.; Nelson, S. R. *Biomacromolecules* **2005**, *6*, 3095–3099.
(10) Sponner, A.; Schlott, B.; Vollrath, F.; Unger, E.; Grosse, F.; Weisshart, K. *Biochemistry* **2005**, *44*(12), 4727–4736.
(11) Simmons, A. H.; Michal, C. A.; Jelinski, L. W. *Science* **1996**, *271*(5245), 84–87.
(12) Termonia, Y. *Macromolecules* **1994**, *27*(25), 7378–7381.
(13) Kummerlen, J.; van Beek, J. D.; Vollrath, F.; Meier, B. H. *Macromolecules* **1996**, *29*(8), 2920–2928.
(14) Lefevre, T.; Rousseau, M.-E.; Pezolet, M. *Biophys. J.* **2007**, *92*, 2885–2895.

Preferential repartitions of MaSpI in the crystalline phase and MaSpII in the amorphous phase have thus been proposed[15] so that MaSpII would contribute mainly to the fiber's extensibility and not to its strength. As shown very recently, proline residues seem to play a key role in increasing the fiber's elasticity, extensibility, and capacity to shrink.[16−20] Moreover, Pomès and co-workers have revealed that the proline and glycine contents of proteins determine their amyloidogenic or elastomeric character.[21] In this context, MaSpII would be elastomeric whereas for MaSpI none of these characters seem to predominate.[21] Therefore, it can be predicted that the assembly of MaSpI and MaSpII either during the spinning process or *in vitro* should be different.

Huemmerich et al. have shown that the recombinant proteins ADF-3 and ADF-4, the MaSpI and MaSpII counterparts of *Araneus diadematus*, have opposite behaviors *in vitro*.[22] Whereas ADF-3 is water-soluble, ADF-4 is insoluble and self-assembles into filaments when expressed in insect cells.[22] This difference has been related to the protein sequences, as the authors underlined the fact that MaSpI/MaSpII couples encountered in various spider species display intrinsic physicochemical differences, in particular hydrophobicity and net charge.[22] Moreover, both proteins aggregate when involved in an elongation flow in a microfluidic system, but, whereas ADF-4 only forms spherical assemblies, ADF-3 forms fibers at high flow rates.[23]

Despite these results, a more complete comprehension of the self-assembly properties of these proteins is required to better understand the mechanisms governing the supramolecular assembly association as well as the resulting microstructure of spider silk. To this end, studies at the air−water interface are relevant to characterize the amphiphilicity of silk proteins and to highlight the nature of the interactions that dictate their self-assembly. We have thus used the monolayer technique to study the adsorption of MaSpI and MaSpII at the air−water interface and the physical and structural properties of the films formed. The experiments were performed using the recombinant *Nephila clavipes* spider BioSteel proteins produced by Nexia Biotechnologies Inc.[24,25] Several techniques were used to characterize *in situ* the time evolution of the physical properties and structure of the interfacial films. Ellipsometry and surface pressure measurements allowed us to estimate the amount of protein at the interface and to determine lateral interactions, respectively. On the other hand, surface rheology was used to detect the development of a rigid network at the interface while polarization modulation infrared absorption reflection (PM-IRRAS) combined with statistical analysis has allowed study of the conformation and orientation of the proteins.

Experimental Section

MaSpI and MaSpII Solutions. The recombinant spidroins MaSpI and MaSpII were obtained from Nexia Biotechnologies Inc. (Montréal, Québec, Canada).[24,25] Their sequences are composed of the repetitve subunits and the C-terminal part, with a molecular weight of about 60−65 kDa. The protein powders were dissolved in 6 M guanidine hydrochloride (GdnHCl) at a concentration of 3.1 mg/mL and heated at 60 °C for 60 min. To remove GdnHCl, the solutions were eluted through PD-10 Sephadex G-25 desalting columns (Amersham Biosciences, Baie d'Urfé, Québec, Canada) first equilibrated with 50 mL of the desired buffer. MaSpI was dissolved in H_2O (pH 7), whereas MaSpII was dissolved in a 20.0 mM phosphate buffer pH 11, so that the gap between the pH and the isoelectric point, pI, is comparable ($|pH - pI| \sim 3$ and 2.5 for MaSpI and MaSpII, respectively, which corresponds to net charges of about +17 and −28, respectively). The protein concentration was verified by UV absorption with a UV−vis-NIR spectrophotometer Cary 500 (Varian, Palo Alto, CA) and the solution diluted with the same buffer to obtain the desired final concentration. For all the techniques used, the protein solutions were poured directly into the trough immediately after the preparation. The same solutions were used for PM-IRRAS and ellipsometric/surface tension experiments, with the measurements being recorded simultaneously. Transmission Fourier transform infrared spectroscopy and circular dichroism analyses performed on the recombinant proteins at 10 and 1 mg/mL, respectively, did not reveal the presence of any intermolecular β-sheets that could be due to self-aggregated proteins in the solutions used for the experiments at the air−water interface.

Ellipsometry, Surface Tension, And Shear Elastic Constant Measurements. The ellipsometric angles (Δ and ψ) and surface pressure (Π) were recorded simultaneously in the same trough. The ellipsometric measurements were carried out with a conventional null ellipsometer using a He−Ne laser operating at 632.8 nm as described elsewhere.[26] The angle of incidence was 52.12° (Brewster angle of the air−water interface for pure water of 53.12°). The surface pressure was measured with a Nima PS4 film balance (Coventry, England) using a Wilhelmy plate as a filter paper (0.5 × 1.0 cm Whatman paper, no. 1). All experiments were performed at room temperature between 19 and 21 °C. Protein concentrations in the subphase, C_b, between 1 and 300 μg/mL were investigated.

The rheometer used is based on the action of a very light float applying a small rotational strain (from 10^{-3} down to 10^{-6}) to the water surface through a magnetic couple (with a pair of Helmoltz coils and a small magnet pin deposited in the float). A detailed description of this instrument can be found elsewhere,[27,28] A sinusoidal torque excitation was applied to the float in the 0.01−100 Hz frequency range, and the resistance that the monolayer opposes to the rotation of the float directly measured. The amplitude and phase of the mechanical response of the pure subphase was first analyzed as a function of strain frequency to control the absence of rigidity. The protein solution was then poured into the trough, and the mechanical response of the interfacial layer was recorded at a fixed frequency of 5 Hz as a function of time. When the shear elastic constant, μ (expressed in mN/m), reached a nearly constant value (after ∼8 h), a new measurement was recorded as a function of frequency to determine whether the system behaves as an elastic layer. Measurements were carried out at 18 °C in parallel to ellipsometry.

***In Situ* Infrared Spectroscopy.** The polarization modulation infrared reflection absorption spectrometer used has been described in detail elsewhere.[29] The measurements were performed

(15) Sponner, A.; Unger, E.; Grosse, F.; Weisshart, K. *Nat. Mater.* **2005**, *4*, 772–775.
(16) Brooks, A. E.; Stricker, S. M.; Joshi, S. B.; Kamerzell, T. J.; Middaugh, C. R.; Lewis, R. V. *Biomacromolecules* **2008**, *9*, 1506–1510.
(17) Liu, Y.; Shao, Z.; Vollrath, F. *Biomacromolecules* **2008**, *9*, 1782–1786.
(18) Liu, Y.; Sponner, A.; Porter, D.; Vollrath, F. *Biomacromolecules* **2008**, *9*, 116–121.
(19) Savage, K. N.; Gosline, J. M. *J. Exp. Biol.* **2008**, *211*, 1937–1947.
(20) Savage, K. N.; Gosline, J. M. *J. Exp. Biol.* **2008**, *211*, 1948–1957.
(21) Rauscher, S.; Baud, S.; Miao, M.; Keeley, F. W.; Pomès, R. *Structure* **2006**, *14*, 1667–1676.
(22) Hummerich, D.; Scheibel, T.; Vollrath, F.; Cohen, S.; Gat, U.; Ittah, S. *Curr. Biol.* **2004**, *14*, 2070–2074.
(23) Rammensee, S.; Slotta, U.; Scheibel, T.; Baush, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6590–6595.
(24) Lazaris, A.; Arcidianacono, S.; Huang, Y.; Zhou, J.-F.; Duguay, F.; Chretien, N.; Welsh, E. A.; Soares, J. W.; Karatzas, C. N. *Science* **2002**, *18*, 472–476.
(25) Karatzas, C. N.; Chretien, N.; Duguay, F.; Bellemare, A.; Zhou, J.-F.; Rodenhiser, A.; Islam, S. A.; Turcotte, C.; Huang, Y.; Lazaris, A. High-toughness spider silk fibers spun from soluble recombinant silk produced in mammalian cells. In *Biotechnology of Biopolymers*; Doi, S., Ed.; Wiley-VCH: Weinheim, 2005; Vol. *2*, pp 945−967.
(26) Berge, B.; Renault, A. *Europhys. Lett.* **1993**, *21*, 773–777.
(27) Vénien-Bryan, C.; Lenne, P.-F.; Zakri, C.; Renault, A.; Brisson, A.; Legrand, J.-F.; Berge, B. *Biophys. J.* **1998**, *74*, 2649–2657.
(28) Zakri, C.; Renault, A.; Berge, B. *Phys. B* **1998**, *248*, 208–210.
(29) Bourque, H.; Laurin, I.; Pézolet, M. *Langmuir* **2001**, *17*, 5842–5849.

using a 5 mm deep Teflon trough. Spectra were recorded at 8 cm^{-1} resolution using a Nicolet Magna 850 Fourier transform IR spectrometer (Thermo Scientific, Madison, WI) equipped with a photovoltaic MCT detector (Kolmar Technologies, Newburyport, MA). The polarization of the beam was modulated with a photoelastic modulator PEM-90 (Hinds Instruments, Hillsboro, OR) that was set for optimum efficiency at 1400 cm^{-1}. The spectra were recorded at a scanning mirror velocity of 0.47 cm/s by coadding 1024 interferograms, resulting in a total acquisition time of ~16 min. Normalized PM-IRRAS spectra were obtained using eq 1:

$$\Delta S/S = [S(d) - S(0)]/S(0) \quad (1)$$

where $S(d)$ and $S(0)$ are the PM-IRRAS signals of the covered and uncovered subphase, respectively.

The spectra were recorded as a function of time for at least 4 h after the beginning of the adsorption. They were corrected for baseline deviations using a cubic polynomial function, and corrected for the water vapor contribution by subtracting a reference water vapor spectrum ponderated such that the curvimetric length of the corrected spectra is minimized (Goormaghtigh, E., personal communication).

Statistical Spectral Analyses. The PM-IRRAS spectra were analyzed by principal component analysis (PCA).[30] The purpose of PCA is to summarize data with a controlled, optimized loss of information. It allows one to identify spectral patterns which are the most representative of the variability in the whole data set and to locate individual spectra on similarity maps based on these spectral patterns. It is widely used as an exploratory technique to optimally characterize large data sets.[31,32] If the n spectra consist of the intensities measured at p wavenumbers, any spectrum is a point in a p-dimensional space. Then, PCA can be viewed as the projection of the data on a subspace which maximizes the dispersion of the projected data (variance maximization). The axes that define this subspace are orthogonal (thus uncorrelated), ordered by decreasing order of projected variance, and are called principal components. Since they are linear combinations of the p initial variables (intensities measured at p wavenumbers), each principal component can be associated to a spectral pattern ("load" on the principal component as a function of the wavenumber). Each spectrum is characterized by its scores on the different principal components. These scores express the differences between the spectrum and the overall average spectrum, and represent the coordinates of the spectrum on a factorial map whose axes origin corresponds to the average spectrum.

PCA was applied to a set of 98 spectra obtained from 10 experiments carried out with MaSpI and MaSpII at three concentrations as a function of time. The spectra were first baseline-corrected, truncated between 2000 and 1200 cm^{-1}, and centered (subtraction of the average spectrum) without variance scaling. PCA was computed using the prcomp function of the GNU software package "R".[33] We have used the analysis of the variance (ANOVA) to quantify the statistical significance of the relation between experimental parameters (nature of the protein, bulk protein concentration, adsorption time) and features of the spectra. ANOVA was performed using the coordinates of individual spectra on the similarity map provided by PCA as the responses, adsorption time, and protein concentration as quantitative predictive variables and the nature of the protein as a two-level qualitative predictive variable. For all factors, the main and, for quantitative ones, quadratic effects and all the two-way interactions have been taken into account. In the following, only effects significant at the $p \leq 0.01$ level will be discussed. The significance level p is the probability of making the so-called type I error consisting of rejecting the null hypothesis of absence of effect when this hypothesis is actually true. ANOVA was also performed using "R".[33]

To identify the spectral features discriminating MaSpI and MaSpII, we have used discriminant function analysis (DFA). This provides a set of orthogonal axes which are also linear combinations of the initial variables, and which maximize the "separation" of projected spectra of MaSpI from those of MaSpII (minimization of within-group variance, maximization of between-group variance). As the input of DFA, we used the coordinates of the spectra on the eight first principal components provided by PCA, representing more than 96% of the total variance of the PM-IRRAS data set. DFA was performed using the function lda of the "MASS" R package.[34]

Calculation of the Hydrophilicity of the Proteins. The hydrophilicity of MaSpI and MaSpII has been calculated using the normalized scale of Roseman[35] and Kyte and Doolittle[36] using ProtScale on the ExPASy server (URL: http://us.expasy.org/tools/protscale.html).[37] A window size of five amino acids and a relative weight of the window edges compared to the window center of 100% were used.

Results

MaSpI and MaSpII Adsorption at the Air−Water Interface. This section reports the adsorption kinetics of MaSpI and MaSpII at the air−water interface as probed by ellipsometry and surface pressure measurements. Figures 1 and 2 show the ellipsometric angle, Δ, and the surface pressure, Π, as a function of time for a bulk (subphase) protein concentration, C_b, ranging between 1 and 300 μg/mL for MaSpI and MaSpII. For the benefit of the reader, we recall that Δ essentially depends on the protein interfacial concentration, whereas Π mainly reflects protein−protein lateral interactions. The kinetics of adsorption at the interface has been studied for a minimum of 10 h, a time period after which water evaporation begins to be detectable. The slight lag at the start of the adsorption kinetics observed at high bulk concentrations is due to the fact that the time before the start of the measurement is longer than the time the protein takes to start to adsorb at the interface.

Figures 1 and 2 show that Π and Δ increase with time, revealing that both MaSpI and MaSpII are amphiphilic enough to adsorb at the interface and to form films. The evolution of Π and Δ with time is comparable to globular proteins and can be basically divided into two regions: for short times, Π and Δ increase rapidly, while for long times most Π curves reach a plateau or start to decrease slightly while Δ still increases slightly and regularly. Basically, the first step is generally acknowledged to correspond to the anchoring of the proteins onto the water surface, whereas the latter usually involves a conformational reoganization.[38−40] The conformation of a protein at the air−water interface depends on the affinity of the hydrophobic and hydrophilic residues with the air and water phases, respectively, as well as on the protein flexibility. Table 1 gives the surface pressure

(30) Jolliffe, I. T. *Principal component analysis*; Springer: New York, 1986.
(31) Bertrand, D.; Scotter, C. N. G. *Appl. Spectrosc.* **1993**, *46*, 1420–1425.
(32) Dufour, E.; Dalgalarrondo, M.; Adam, L. *J. Colloid Interface Sci.* **1998**, *207*, 264–272.
(33) R-Development-Core-Team R: a language and environment for statistical computing, http://www.R-project.org.
(34) Venables, W. N.; Ripley, B. D. *Modern applied statistics with S*; Springer: New York, 2002.
(35) Roseman, M. A. *J. Mol. Biol.* **1988**, *200*, 513–522.
(36) Kyte, J.; Doolittle, T. F. *J. Mol. Biol.* **1982**, *157*, 105–132.
(37) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. Protein identification and analysis tools on the ExPASy server. In *The proteomics protocols handbook*; Walker, J. M., Ed.; Humana Press: Totowa, NJ, 2005; pp 571−607.
(38) Damodaran, S. Interfaces, protein films, and foams. In *Advances in Food and Nutrition Research*; Kinsella, J. E., Ed.; Academic Press: New York, 1990; Vol. *34*.
(39) Magdassi, S.; Kamyshny, A. *Surface activity of proteins*; Marcel Dekker: New York, 1996.
(40) Graham, D. E.; Philips, M. C. *J. Colloid Interface Sci.* **1979**, *70*, 403–414.

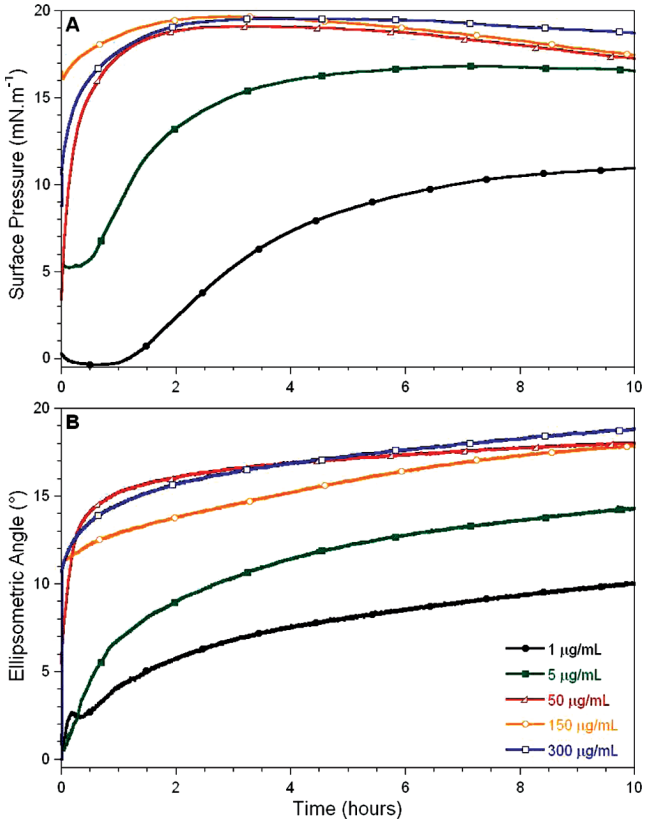

Figure 1. (A) Surface pressure, Π, and (B) ellipsometric angle, Δ, as a function of time for a MaSpI subphase concentration between 1 and 300 μg/mL. For the clarity of the figure, only 10% of the data points are presented.

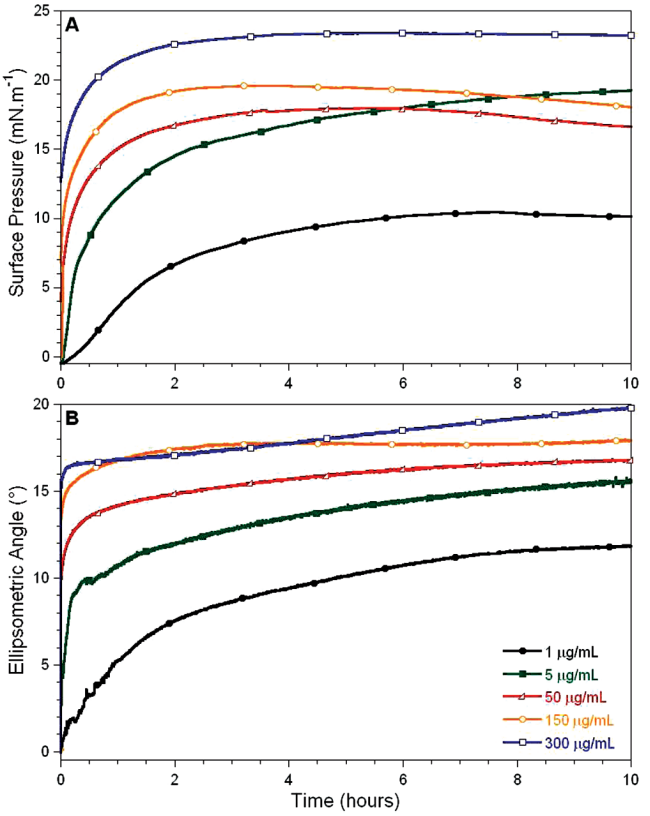

Figure 2. (A) Surface pressure, Π, and (B) ellipsometric angle, Δ, as a function of time for a MaSpII subphase concentration between 1 and 300 μg/mL. For the clarity of the figure, only 10% of the data points are presented.

and ellipsometric values obtained for the different bulk concentrations in the plateau region after 8 h, that is, in a region corresponding to an almost stable film.

Figure 1 shows that the adsorption of MaSpI is more rapid and that the interfacial concentration increases as the bulk concentration is increased, although the difference is reduced at and above 50 μg/mL. For these concentrations, the Π curves converge toward the same value of ~18 mN/m. Figure 1B shows that, for $C_b \geq 5$ μg/mL, the Δ curves tend at long times toward the same value of 18°, suggesting a saturation of the interface by the MaSpI molecules. Interestingly, for these concentrations, a decrease of Π is observed after a few hours, whereas Δ still slightly increases. This decrease of Π is not due to water evaporation but is representative of a decrease in the lateral interactions, which can also be interpreted as an insertion of water molecules within the film. These opposite tendencies of Π and Δ suggest a reorganization of the interfacial layer characterized by some sort of segregation of the proteins from water.

The interfacial behavior of MaSpII (Figure 2) shares obvious similarities with that of MaSpI, but some differences can also be noted. First, MaSpII adsorbs more rapidly than MaSpI. Second, a regular increase in the values of Δ or Π is observed at 10 h as a function of C_b without reaching saturation. Therefore, MaSpII is able to adsorb at a higher rate and in larger amounts than MaSpI, suggesting a higher affinity of MaSpII for the interface. For a bulk concentration of 50 and 150 μg/mL, a decrease of Π is observed as a function of time with a concomitant increase of Δ, like for MaSpI, which suggests that for $C_b > 5$ μg/mL a similar protein segregation occurs. The values of Π and Δ after 8 h (Table 1) are in general slightly higher for MaSpII, reflecting the larger amounts of proteins. Since values of Δ between 10° and 16° generally correspond to a saturated monolayer in the case of globular proteins,[41] the results of Table 1 suggest that both MaSpI and MaSpII form monolayers for $C_b < 50$ μg/mL, whereas they assemble into multilayers and form a protein network for $C_b \geq 50$ μg/mL.

From the measured ellipsometric angle and an increment of the refractive index of the protein of 0.200 mL/g,[42] the surface concentration of the adsorbed proteins, Γ (mg/m^2), can be calculated using the relationship between Γ and Δ derived by de Feijter et al.[42] ($\Gamma = \Delta/5$). Typically, $\Gamma \sim 3$ mg/m^2 for a globular protein monolayer.[41,43,44] Figure 3 shows the value of Γ after 8 h of adsorption as a function of C_b. For MaSpI, the amount of protein adsorbed at the interface increases with C_b and is constant above 50 μg/mL. This plot confirms that MaSpI forms monolayers below 50 μg/mL, whereas above this value the films cannot accommodate more molecules (saturation) and probably form multilayers. In the case of MaSpII, a progressive increase of Γ with C_b is observed even at high concentrations, with the value obtained at 300 μg/mL being significantly higher than that of MaSpI.

At bulk concentrations of 1 μg/mL or less, it is possible from the value of Γ to characterize the very first step of adsorption of the proteins onto a free interface by assuming a diffusion-driven

(41) De Feijter, J. A.; Benjamins, J. Adsorption kinetics of protein at the air-water interface. In *Food emulsions and foams*; Dickinson, E., Stainsby, G., Eds.; Royal Society of Chemistry: London, 1987.
(42) De Feijter, J. A.; Benjamins, J.; Veer, F. A. *Biopolymers* **1978**, *17*, 1759–1772.
(43) Benjamins, J.; Lucassen-Reynders, E. H. Surface dilational rheology of proteins adsorbed at air/water and oil/water interfaces. In *Proteins at liquid interfaces*; Möbius, R. M. D., Ed.; Elsevier: Amsterdam, 1998; pp 341−383.
(44) Pezennec, S.; Gauthier, F.; Alonso, C.; Graner, F.; Croguennec, T.; Brulé, G.; Renault, A. *Food Hydrocolloids* **2000**, *14*, 463–472.

Table 1. Values of the Surface Pressure, Ellipsometric Angle, and Shear Elastic Constant for MaSpI and MaSpII as a Function of Concentration[a]

	MaSpI			MaSpII		
C_b (μg/mL)	Π (mN/m) ($\pm$1.0)	Δ (°) ($\pm$1.0)	μ (mN/m) ($\pm$1.0)	Π (mN/m) ($\pm$0.5)	Δ (°) ($\pm$0.5)	μ (mN/m) ($\pm$1)
1	10	9	9	11	12	0
5	17	13	7	18	15	5
50	18	18	10	18	17	17
150	18	17	nm[b]	19	18	nm[b]
300	19	18	nm[b]	23	19	nm[b]

[a] Values of Π and Δ are taken after 8 h, and those of μ are taken at the end of the kinetics. [b] nm: not measured.

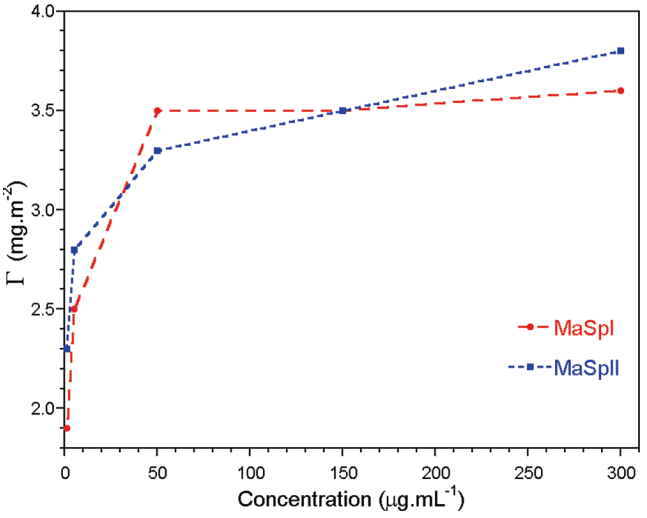

Figure 3. Plot of the surface concentration, Γ, measured after 8 h of adsorption as a function of the bulk concentration of MaSpI and MaSpII.

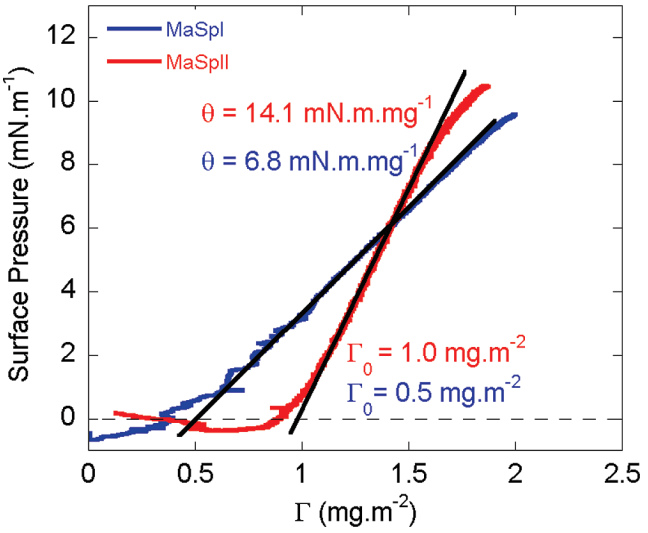

Figure 4. Plot of surface pressure, Π, as a function of the surface concentration, Γ, for MaSpI and MaSpII at a bulk concentration C_{bulk} = 1 μg/mL. The values of the parameters Γ_0 (intersection with the abscissa axis) and θ (slope of the linear part of the curves) are given.

adsorption process. Under these conditions, Ward and Tordai[45] have shown that the surface concentration varies as the square root of the time (eq 2):

$$\Gamma = 2C_b(Dt/\pi)^{1/2} \quad (2)$$

where D is the diffusion coefficient from the bulk to the interface and t is time. Therefore, the initial part of the $\Gamma - t^{1/2}$ plot should be linear and the diffusion coefficient can be extracted experimentally. However, for many proteins, energy barriers are present for adsorption so that the diffusion coefficient extracted from the linear regime is an apparent one.[46] Nevertheless, this parameter allows a comparison between proteins and is usually found in the 10^{-10} m^2/s range for small objects in solution. The linear domain corresponding to the apparent diffusive regime is actually observed for the two proteins (data not shown), and the D values extracted from the $\Gamma - t^{1/2}$ plots are 0.8×10^{-10} m^2/s for MaSpI and 2.0×10^{-10} m^2/s for MaSpII. They are thus of the same order of magnitude and are comparable to the values found in the literature for globular proteins.[44] The higher value obtained for MaSpII is consistent with a more rapid adsorption.

For proteins, a lag time may exist between the onset of the surface coverage and the rise of the surface pressure.[47] To determine the critical surface concentration at which the proteins at the interface start to reduce the surface tension of water, data were analyzed by plotting Γ against Π at C_b = 1 μg/mL (Figure 4). From these curves, it is possible to determine Γ_0 (mg/m^2), the surface concentration at which lateral interactions begin to be detected (i.e., start of the rise of the surface pressure), and θ_0, the slope of the curve (mN·m·mg^{-1}), which characterizes the rate at which the surface pressure increases after it starts to increase. The values of Γ_0 for MaSpI and MaSpII are 1.0 and 0.5 mg/m^2, respectively, indicating that a higher amount of adsorbed proteins is needed to initiate lateral organization for MaSpI than for MaSpII, which in turn suggests that, at the beginning of the adsorption kinetics, MaSpII occupies a larger molecular area. In addition, the higher value of θ_0 for MaSpI (14 mN·m·mg^{-1}) compared to MaSpII (7 mN·m·mg^{-1}) indicates that lateral interactions develop more efficiently for MaSpI than for MaSpII, with the former being more active once at the interface than the latter.

Mechanical Properties at the Air−Water Interface. To further characterize the interfacial films, surface rheology was performed at different bulk concentrations. However, due to our experimental setup and to the high rigidity of the layer, it was not possible to perform measurements above 50 μg/mL. Figure 5 shows the shear elastic constant, μ, as a function of time at C_b = 50 μg/mL as a typical example. The values measured after 8 h of adsorption are presented in Table 1 for the different subphase concentrations. For MaSpI, μ increases with time (i.e., as the surface concentration increases), but the evolution is rather irregular with sudden drops of μ. This behavior was observed for all experiments and suggests that the interfacial films of MaSpI are very heterogeneous and composed of coarse particles and patches. The experiments are thus poorly reproducible, and μ displays rather dispersed values in the 7−20 mN m^{-1} range as a function of C_b (Table 1), without clear tendency. The mechanical response of the layer after 8 h of adsorption has nevertheless been measured as a function of the excitation frequency as shown in

(45) Ward, A. F. H.; Tordai, L. *J. Chem. Phys.* **1946**, *14*, 453–461.
(46) Xu, S.; Damodaran, S. *Langmuir* **1994**, *10*, 472–480.
(47) Razumovsky, L.; Damodaran, S. *Langmuir* **1999**, *15*, 1392–1399.

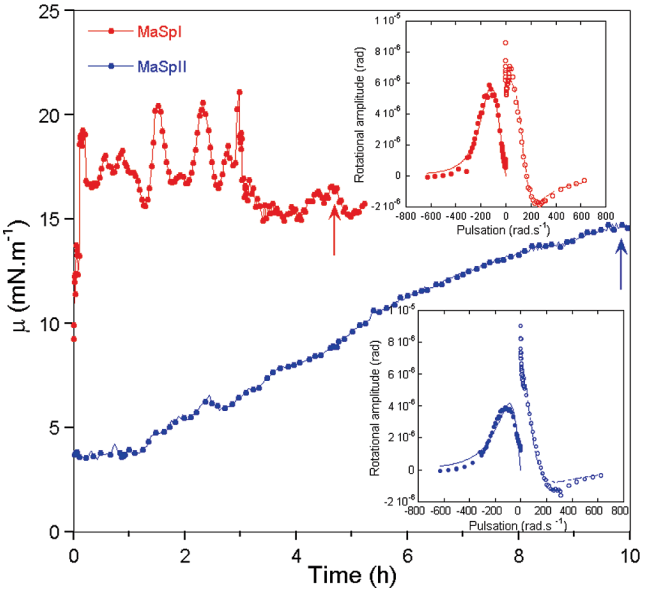

Figure 5. Shear elastic constant, μ, as a function of time during the adsorption kinetics of MaSpI (red curve) and MaSpII (blue curve) for a bulk concentration of 50 μg/mL. Measurements were made at a fixed frequency of 5 Hz. The insets show the real (empty symbols) and imaginary (full symbols) parts of the rotational amplitude of the float measured (mechanical responses of the layer) for MaSpI (red) and MaSpII (blue) as a function of the excitation frequency at the end of the adsorption kinetics (indicated by arrows). The fit of the data using an elastic model (full lines) is also shown (see text for details).

Figure 5. It is noteworthy that the fit of the data (real and imaginary parts of the rotational amplitude of the float) using an elastic model[27,28] (insets in Figure 5) is satisfactory, suggesting that the film is elastic. As opposed to MaSpI, μ increases very regularly with time for MaSpII, thus reflecting a much more homogeneous film. The shear elastic constant increases with C_b (Table 1), which may be expected since μ should increase with the surface density as it is often found for globular proteins that form elastic films. The fit of the data using an elastic model is also satisfactory, which ascertains that the film behaves like an elastic layer. However, as opposed to MaSpI, an increase of the real part of the mechanical response of MaSpII at very low frequencies indicates that slow, nonoscillating motion takes place additionnaly to the elastic response.

Protein Conformation at the Interface. Figure 6 shows the PM-IRRAS spectra in the amide I (1700−1600 cm^{-1}) and amide II (1580−1480 cm^{-1}) regions as a function of time for MaSpI (Figure 6A) and MaSpII (Figure 6B) solutions at C_b = 50 μg/mL. The overall intensity increases with time, showing the progressive adsorption of the proteins. In the case of MaSpI, an amide I component at 1621 cm^{-1} assigned to β-sheets[48] is observed on the very first spectrum recorded. This indicates that the first protein molecules to adsorb at the interface at very low surface pressure already contain β-sheets, while transmission infrared spectrum demonstrates that MaSpI does not contain β-sheets in solution (data not shown). The unresolved shoulder at ~1650 cm^{-1} is due to other secondary structures such as α- and 3_1-helices and/or unordered structures[48] that can account for residual native or unfolded chain segments. As the adsorption progresses, the

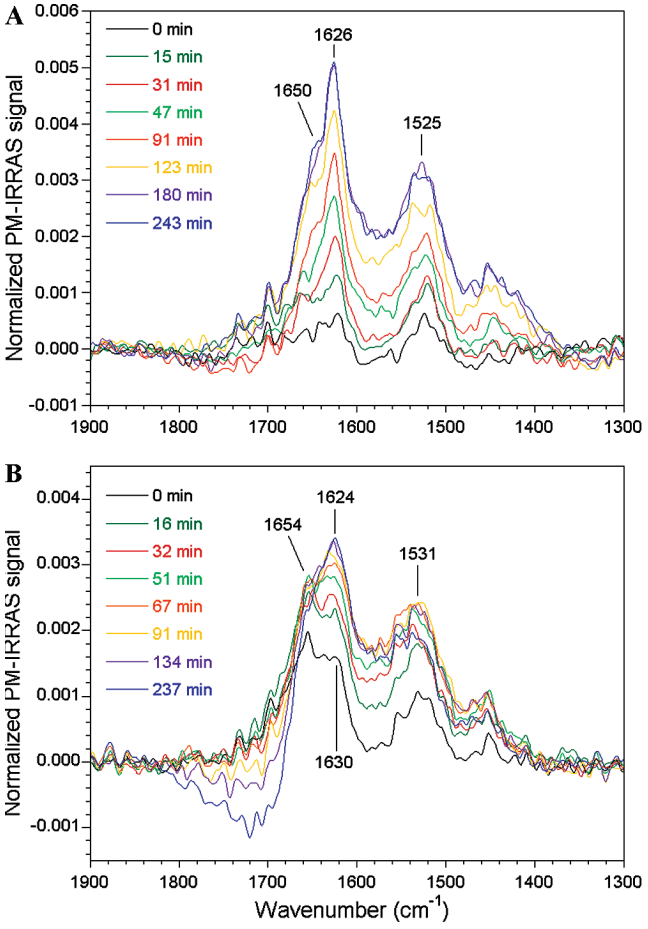

Figure 6. PM-IRRAS spectra of (A) MaSpI and (B) MaSpII as a function of time for a bulk concentration of 50 μg/mL.

component at 1621 cm^{-1} shifts to 1624 cm^{-1}. In addition, the 1624/1650 cm^{-1} intensity ratio slightly increases, indicating that the proportion of β-sheet rises during the adsorption kinetics. The last spectra obtained are typical of a protein in an aggregated state containing a large amount of intermolecular β-sheets,[49,50] which is consistent with the formation of a protein network. The fact that the sign of these bands is positive reveals that the β-sheets are predominantly oriented in the plane of the film.[51] Moreover, the amide I/amide II intensity ratio increases with time, suggesting a reorientation of the β-sheets.

As seen in Figure 6B, the first spectrum recorded for MaSpII is more intense than that of MaSpI, indicating that MaSpII adsorbs more readily than MaSpI, in agreement with the ellipsometry data. The shape of the spectrum is also different, indicating a different conformation at the interface. The amide I band is broad, suggesting that the protein is composed of various secondary structures. A sharp and intense band located at 1654 cm^{-1} dominates the spectrum. Its position and width suggest that it is due to the α-helical conformation. The presence of the component at 1630 cm^{-1} is associated with β-sheets. The shape of the amide I band changes significantly with time as the component at 1630 cm^{-1} shifts to 1624 cm^{-1} and as it increases in intensity at the expense of the 1654 cm^{-1} component. After approximately 2 h, the band at 1624 cm^{-1} is the dominant feature

(48) Goormaghtigh, E.; Cabiaux, V.; Ruysschaert, J.-M. Determination of soluble and membrane protein structure by Fourier transform spectroscopy. II. Experimental aspects, side-chain structure and H/D exchange. In *Physicochemical methods in the study of biomembranes*; Hilderson, H. J., Ralston, G. B., Eds.; Plenum Press: New York, 1994; Vol. *23*, pp 363−403.

(49) Casal, H. L.; Köhler, U.; Mantsch, H. H. *Biochim. Biophys. Acta* **1988**, *957*, 11–20.

(50) Surewicz, W. K.; Mantsch, H. H.; Stahl, G. L.; Epand, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7028–7030.

(51) Blaudez, D.; Turlet, J.-M.; Dufourcq, J.; Bard, D.; Buffeteau, T.; Desbat B. *J. Chem. Soc., Faraday Trans.* **1996**, *92*, 525–530.

in the amide I region, emphasizing the predominance of intermolecular β-sheets due to a protein interfacial self-aggregation. However, as judged from the relative intensity of the band at 1624 cm^{-1}, MaSpII forms less β-sheets than MaSpI while the component at 1654 cm^{-1} is still clearly visible as a shoulder. One can also note the progressive appearance of a broad negative dip above 1700 cm^{-1}. Such a feature has also been observed in the PM-IRRAS spectra of phospholipids, proteins, and peptides and has been associated with optical effects when the water surface is covered by thin films.[52] In the case of protein films, this broad negative dip is also sensitive to the orientation of the amide groups at the interface. Figure 6B shows clearly that this negative feature is more important in the case of MaSpII, suggesting that the peptide bonds of this protein in contact with the interface are more oriented. As for MaSpI, the amide I/amide II intensity ratio progressively increases as a function of time, indicating that, once the film is formed, it undergoes a reoganization with time.

Time-dependent spectra were also recorded at bulk protein concentrations of 5 and 150 μg/mL, most often in duplicate. The resulting data set consists of 98 spectra whose spectral variability has been examined by PCA. Figure 7 illustrates the spectral patterns associated with the first two principal components, which account for 76% and 9% of the total variance. Any other principal component accounts for less than 5% of the total variance and is not considered thereafter. The overall average spectrum is shown for comparison. The spectral pattern associated with the first principal component, PC1, shows a broad, asymmetric shape ranging from 1700 to 1400 cm^{-1} with a maximum at 1627 cm^{-1} which is reminiscent of the contribution of β-sheets. The broad and ill-defined feature that encompasses the 1600−1400-cm^{-1} region represents the shape of the baseline, especially between the amide I and amide II bands (i.e., between 1600 and 1580 cm^{-1}). PC2 shows a negative broad band centered at ~1683 cm^{-1} that seems to represent mainly the broad negative dip at 1700 cm^{-1} and possibly another small band near 1660 cm^{-1} due to other secondary structures. The shoulder at 1630 cm^{-1} is representative of the β-sheets, the band at 1517 cm^{-1} seems to be due to the amide II band, whereas the region centered at ~1593 cm^{-1} is assigned to baseline or amino acid variations.

Using the scores of the spectra on PC1 and PC2 as coordinates, the factorial map has been plotted (Figure 8). Considering the proportions of variance accounted for by PC1 and PC2, this map expresses more than 84% of the total variance of the data. Some coarse tendencies are apparent, for example, the fact that scores on PC1 tend to increase with time and bulk protein concentration. Some of these trends are confirmed by ANOVA. Indeed, the scores on PC1 were affected, in order of importance, by protein concentration, time, and nature of the protein at the $p < 0.01$ level. Scores on PC1 significantly increase with time and concentration. This means that PC1 mainly reflects the overall intensity of the spectra. As a matter of fact, the area of the spectra is hightly correlated with the PC1 score with a correlation coefficient of 0.98 (data not shown). These results are in good agreement with ellipsometry and surface pressure measurements, showing that the amount of protein at the interface increases with time and concentration.

Scores on PC1 are slightly higher for MaSpII, suggesting that it adsorbs in higher amounts, in agreement with ellipsometry. Scores on PC2 are also slightly higher for MaSpII than for MaSpI, which can be interpreted as more negative contributions at ~1700 and 1630 cm^{-1} for MaSpII. Thus, the broad negative dip is stronger and the β-sheet content is lower for MaSpII, as illustrated by the spectra of Figure 6. Scores on PC1 and PC2 evolve similarly with time for MaSpI and MaSpII, with the effects of other factors on the scores on PC2 being not significant. Consequently, the time-dependent conformational changes at the interface seem similar for both spidroins. Overall, the PCA results are in good agreement with Figure 6, but the strength of such a statistical analysis is that the conclusions regarding the behaviors of MaSpI and MaSpII are obtained with the whole data set and not only with a single series of spectra.

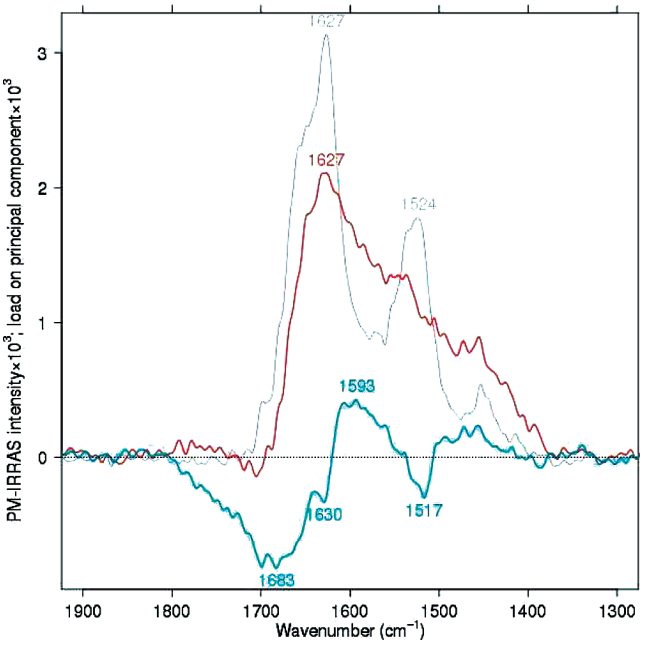

Figure 7. Overall average spectrum (gray spectrum) and spectral patterns associated with the first two principal components, PC1 (red spectrum) and PC2 (blue spectrum). For comparison, the spectral patterns have been ponderated by 2× their standard deviation.

In the factorial map of Figure 8, the spectra of MaSpI and MaSpII are not well separated. This result is noteworthy, since it shows that the spectral variability of MaSpI and MaSpII is dominated by bulk concentration and time, and not much by the type of protein. To discriminate between MaSpI and MaSpII, we took advantage of DFA which, in this particular case, has been found to be very efficient. Figure 9 shows the spectral pattern associated with the discriminant factor computed from the scores of the first eight principal components, as well as the distribution (histograms) of the scores of the spectra of MaSpI (scores negative) and MaSpII (scores positive). This discriminant factor represents the spectral pattern that discriminates the two proteins at best. It displays three well-defined negative bands at 1702, 1627, and 1516 cm^{-1}. The first two bands are due to the PM-IRRAS negative dips and β-sheets, respectively. Considering the positive score of MaSpII, DFA confirms the higher contribution of the negative dip and lower contribution of β-sheets for this protein. The component at 1516 cm^{-1}, although located in the amide II region, is assigned to tyrosine residues,[53] because it is narrow and often encountered in proteins. Its negative contribution reveals that it is more intense in MaSpI than in MaSpII. The content of tyrosine in the protein cannot explain this feature, since the former contains less tyrosine residues than the latter (~3 and ~5%, respectively). In fact, it is due to the pH used. MaSpII being

(52) Blaudez, D.; Buffeteau, T.; Cornut, J. C.; Desbat, B.; Escafre, N.; Pézolet, M.; Turlet, J. M. *Appl. Spectrosc.* **1993**, *47*, 869–874.

(53) Venyaminov, S. Y.; Kalnin, N. N. *Bioploymers* **1990**, *30*, 2143–1257.

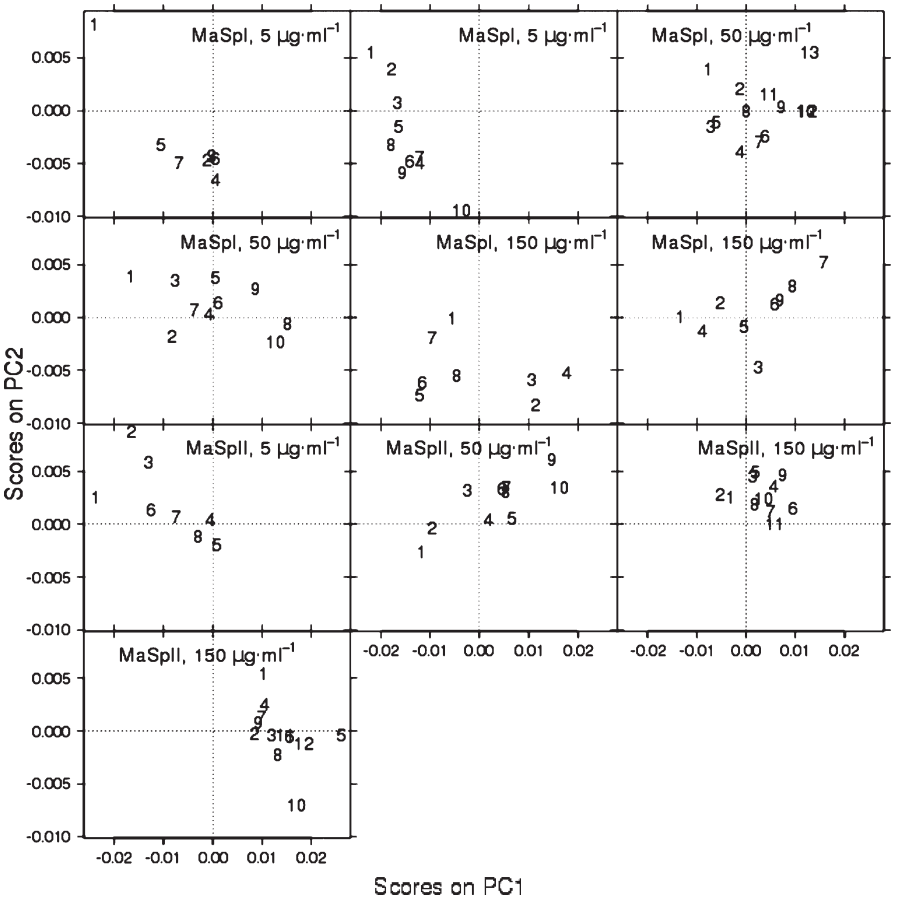

Figure 8. PC1−PC2 factorial map calculated for the different PM-IRRAS experiments. The axis origin corresponds to the position of the overall average spectrum. Each point represents a particular spectrum of the data set.

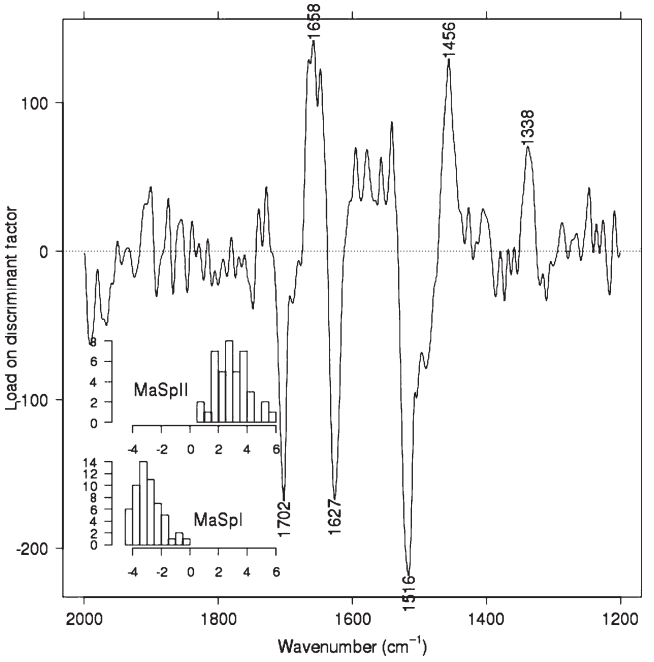

Figure 9. Discriminant factor computed from the scores of the first eight principal components and distribution (histograms) of the scores of the spectra of MaSpI (scores negative) and MaSpII (scores positive) on the discriminant factor.

dissolved at pH 11, the O−H groups of the tyrosine side chains are largely ionized, so that the band at 1516 cm^{-1} is shifted to ∼1500 cm^{-1}, resulting in a more intense band at 1516 cm^{-1} for MaSpI. The discriminant spectral pattern also exhibits positive narrow bands at 1658, 1456, and 1338 cm^{-1} that are thus more intense for MaSpII than for MaSpI. The band at 1658 cm^{-1} band being positive, it is more important in MaSpII and thus most probably due to α-helices. The presence of α-helices for MaSpII can be accounted for due to the fact that, although proline residues are known to prevent the formation of α-helices in aqueous solution, prolines nonetheless allow their formation in a hydrophobic environment.[54] The 1456 and 1338 cm^{-1} bands may be due to the CH_3 asymmetric/CH_2 bending modes and CH_3 symmetric bending mode, respectively, or to proline side chains.[55,56] The higher intensity of these two bands for MaSpII thus reflects a difference in the amino acid composition.

Discussion

This work shows that, due to their amphiphilic nature, MaSpI and MaSpII readily adsorb and form films at the air−water interface. The evolution of Π and Δ reveals that the adsorption is more rapid for MaSpII and that it can adsorb in larger amounts without saturation. Taken together, the results indicate that MaSpII has a higher affinity for the air−water interface than MaSpI. This is not due to an electrostatic effect, since, at the pHs used, the net charge of MaSpI in absolute value is smaller than that of MaSpII. It has been suggested that the two spidroins have a different solubility in water.[22] The absence of turbidity of the solutions indicates that MaSpI and MaSpII are soluble in water,

(54) Li, S.-C.; Goto, N. K.; Williams, K. A.; Deber, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6676–6681.
(55) Krimm, S.; Bandekar, J. *Adv. Protein Chem.* **1986**, *38*, 181–364.
(56) Barth, A. *Prog. Biophys. Mol. Biol.* **2000**, *74*, 141–173.

but transmission IR experiments (not shown) have revealed that MaSpI is quite stable in aqueous solution whereas MaSpII readily aggregates with time. This process is temperature- and concentration-dependent. Thus, the higher affinity of MaSpII for the interface with respect to MaSpI is consistent with the lower stability of MaSpII in water.

As seen by PM-IRRAS, the adsorbed proteins contain a predominant amount of β-sheets with other conformational elements. The spectra are typical of protein aggregates, characterized by a large content of intermolecular β-sheets, as found in denatured globular proteins[49,50] or in amyloids.[57] Thus, MaSpI and MaSpII spontaneously and readily aggregate to form β-sheets at the interface, to finally develop into a two-dimensional network as proteins reach the interface. It is noteworthy that the spectra of MaSpI, the most abundant spidroin of the MA thread, are very similar to the attenuated total reflection (ATR)-IR spectrum of a dragline silk fiber (not shown). Thus, it seems that the self-association of MaSpI at the air−water interface is not fundamentally different than that occurring in the spider's gland. Nevertheless, the process is much more rapid at the interface (in the order of minutes) than in solution (in the order of hours, data not shown) as the PM-IRRAS spectra recorded at low adsorption times are already typical of aggregated proteins.

MaSpII forms less β-sheets than MaSpI, which is probably due to the higher proline content of MaSpII. This amino acid has, among the 20 natural ones, the lowest β-sheet propensity[58−60] and is a β-sheet breaker,[54,61,62] so that it is likely that the glycine-rich (i.e., proline-rich region) region of MaSpII is not able to form β-sheets but is rather involved in α-helices. Consequently, like those found in natural fibers, the β-sheets generated at the air−water interface are mostly formed by the polyalanine blocks. This is probably also true for MaSpI, although the glycine-rich sequence may be involved in some β-sheets. It has been shown using ATR infrared spectroscopy that LB films formed by *Bombyx mori* fibroin at the air−water interface and transferred on a germanium substrate give a strong band at 1624 cm^{-1} characteristic of β-sheets (silk II polymorph).[63] It was proposed that the crystalline segments of the fibroin form β-sheets at the interface while the amorphous region would lie in the subphase.[63] In other studies, a new polymorph of *Bombyx mori* fibroin was found with a trigonal unit cell and a hexagonal packing of the chains, similar to polyglycine II.[64,65]

Ellipsometry measurements at a protein bulk concentration of 1 μg/mL show that a lower surface concentration of MaSpII is necessary to initiate lateral interactions. Thus, at low surface coverages, MaSpII occupies a molecular area sufficiently large to develop intermolecular interactions (larger Γ_0). This larger molecular area probably results from a conformational restriction imposed by the proline residues[21,66] and to a lower structural flexibility of the protein due to its lower Gly content.[21] MaSpI at low surface concentrations is already dominated by the β-sheet structure and is probably extensively aggregated, but the aggregates appear to be compact (small Γ_0) due to a higher chain flexibility. As the surface concentration increases, interactions develop rapidly (θ_0 positive). Since an equivalent increase in surface concentration leads to a smaller increase of the lateral interactions for MaSpII, the protein self-assembly appears slower.

Surface rheology gives clear evidence that MaSpI forms heterogeneous elastic films whereas MaSpII forms more homogeneous and elastic ones. This supports a rapid and disordered aggregation process for the former spidroin and a more progressive and more ordered one for the latter. The higher value of θ_0 for MaSpI is consistent with this result. In addition, its lower affinity for the interface should lead to a higher predominance of intermolecular interactions which contributes to the higher rate and the more disordered aggregation process of this protein. Nevertheless, protein−protein interactions are high enough for both MaSpI and MaSpII to result in a segregation between the proteins and water at high concentrations, although the phenomena are less important for MaSpII, since its affinity for the interface is higher.

These differences in the mechanisms of assembly of MaSpI and MaSpII can be fruitfully compared with the protein gel networks formed by globular proteins in solution. Globular protein gels, especially heat-induced ones, exhibit a wide range of macroscopic properties, due to different shapes and sizes of the aggregates that result from specific aggregation mechanisms. For example, elastic and transparent gels can be formed upon heating when electrostatic interactions are strong, whereas stiff and opaque gels occur for low elecrostatic repulsions.[67,68] The former gels are constituted by fine filamentous strands (fine-stranded gels), whereas the latters are made of coarse, roughly spherical particles and have a lower water-holding capacity (particulate gels). Mixed gels can also be found. Fine-stranded and particulate gel networks are made of aggregates in which the proteins contain a large amount of β-sheets, but, whereas the former gels result from an ordered aggregation process, the latter ones are formed through a disordered aggregation that occurs at a higher rate.[67,69] The protein aggregation results from a delicate balance between protein−water and protein−protein interactions, with the latter being favored for particulate gels. As extreme cases, amyloid fibrils are the perfect example of aggregates resulting from a slow and hierachical aggregation,[70] whereas protein coagulation is the archetype of rapid and random aggregation.[69] Consequently, the films formed at the air−water interface by MaSpI may be more of the "particulate" type, whereas those formed by MaSpII may be closer to "fine-stranded".

Such behaviors of the MA silk proteins are related to their particular primary structures that are different but share the same pattern, which explains why their behaviors are not so drastically different. Thus, more subtle structural particularities should describe their properties. Hummerich et al. reported that MaSpI is more hydrophobic than MaSpII for several spider species. They associate this characteristic and the relative net charges of the spidroins to the different water solubility of MaSpI and MaSpII.[22] Using the non-normalized scale of Kyte and Doolittle,[36]

(57) Fabian, H.; Choo, L.-P.; Szendrei, G. I.; Jackson, M.; Halliday, W. C.; Otvos, L.Jr.; Mantsch, H. H. *Appl. Spectrosc.* **1993**, *47*, 1513–1518.
(58) Pal, D.; Chakrabarti, P. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2000**, *56*, 589–594.
(59) Street, A. G.; Mayo, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9074–7076.
(60) Deléage, G.; Roux, B. *Protein Eng.* **1987**, *1*(4), 289–294.
(61) Forrod, B.; Pérez-Payá, E.; Houghten, R. A.; Blondelle, S. *Biochem. Biophys. Res. Commun.* **1995**, *211*, 7–13.
(62) Shamir, M. S.; Dalby, A. R. *Biophys. J.* **2007**, *92*, 2080–2089.
(63) Muller, S. W.; Samuelson, L. A.; Fossey, S. A.; Kaplan, D. L. *Langmuir* **1993**, *9*, 1857–1861.
(64) Valluzzi, R.; Gido, S. P.; Muller, S. W.; Kaplan, D. L. *Int. J. Biol. Macromol.* **1999**, *24*, 237–242.
(65) Valluzzi, R.; Gido, S. P.; Zhang, W.; Muller, S. W.; Kaplan, D. L. *Macromolecules* **1996**, *29*, 8606–8614.
(66) MacArthur, M. W.; Thornton, J. M. *J. Mol. Biol.* **1991**, *218*, 397–412.
(67) Clark, A. H.; Judge, F. J.; Richards, J. B.; Stubbs, J. M.; Suggett, A. *Int. J. Pept. Protein Res.* **1981**, *17*, 380–392.
(68) Stading, M.; Hermansson, A.-M. *Food Hydrocolloids* **1991**, *5*, 339–352.
(69) Schmidt, R. H. Gelation and coagulation. In *Protein functionality in food*; Cherry, J. P., Ed.; American Chemical Society: Washington, DC, 1981; Vol. *147*, pp 131−145.
(70) Jarrett, J. T.; Lansbury, P. T.Jr *Cell* **1991**, *73*, 1055–1058.

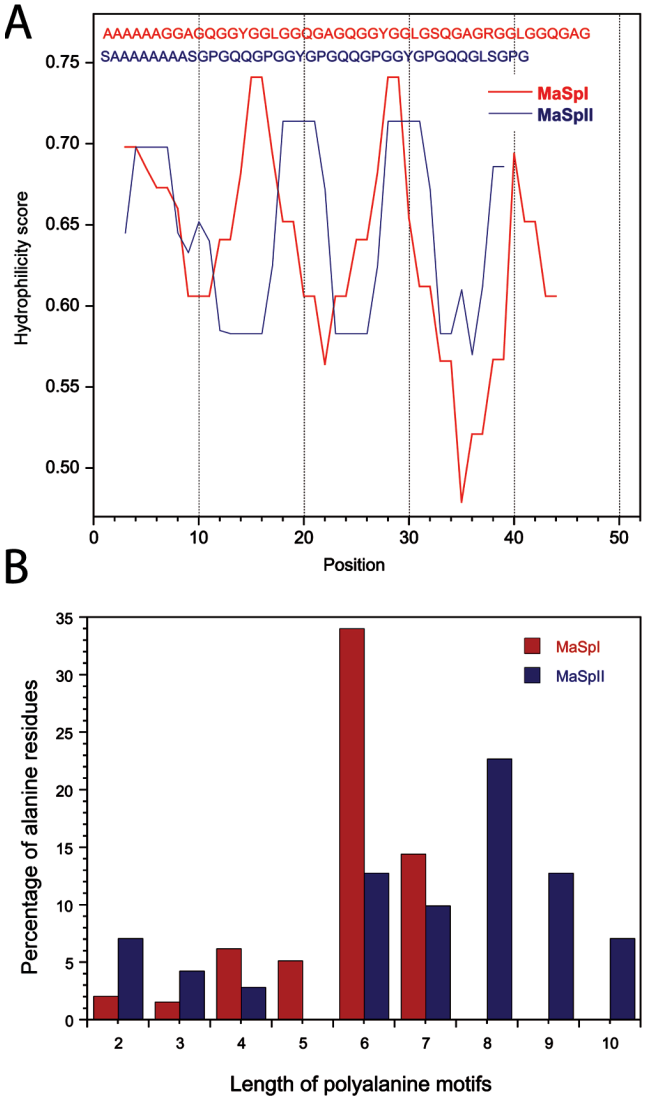

Figure 10. Characterization of the repetitive sequences of MaSpI and MaSpII. (A) Hydrophilicity calculated with the normalized scale of Roseman.[35] The values of 0 and 1 correspond to the most (phenylalanine) and least (arginine) hydrophobic residues, respectively, and the mean value for the 20 natural amino acids is 0.59. The typical repetitive sequences used for the calculation are shown on the top on the graph. (B) Length distribution of the polyalanine sequences of the recombinant proteins MaSpI and MaSpII calculated over the entire sequences.

these authors have found hydrophobicity values of −0.1 and 0.7 for MaSpI and MaSpII, respectively.[22] Using the scale of Kyte and Doolittle normalized between 0 and 1 (0 and 1 correspond to the least hydrophobic (arginine) and the most hydrophobic (isoleucine) residue, respectively), one obtains an overall hydrophobicity of 0.50 and 0.46 for MaSpI and MaSpII, respectively, which represents a difference of 8%. Using the more recent scale of Roseman (in this case, the limit values of 0 and 1 correspond to the arginine and phenylalanine, respectively),[35] one finds an even smaller disparity (0.63 for MaSpI and 0.64 for MaSpII). Hence, the difference in hydrophobicity between MaSpI and MaSpII is very small and probably not large enough to account for the specificity of their behavior at the interface and in the bulk.

Other structural characters have to be found to explain the present data. It appears that the mean hydrophobicities of the polyalanine- and glycine-rich segments are also very close for both proteins (not shown). Other particularities in their primary structures however do exist. First, whereas MaSpI contains essentially the GXG block, MaSpII is mainly constituted of the GPGXX motifs As a consequence, although the overall hydrophobicity is very close for MaSpI and MaSpII, the distribution of the local hydrophobicity along the polypeptide chain might be different. Figure 10A shows the hydrophobicity calculated along the typical repetitive sequences of MaSpI and MaSpII using the normalized hydrophilicity scale of Roseman.[35] As seen on these graphs, the two spidroins exhibit similar alternations of hydrophobic/hydrophilic blocks. Such a pattern may play a role in the aggregation process during the spinning process, since it has been shown that the alternating periodicity of polar and apolar amino acids favors protein oligomerization[71] and the formation of amyloid fibrils.[72] Moreover, the role of a pattern of hydrophobic and hydrophilic blocks has already been emphasized to describe the assembly of cocoon silk fibroins.[73] However, no significant difference seems to arise between the repetitive segments of MaSpI and MaSpII. Likewise, the C-terminal sequences of MaSpI and MaSpII are very similar (data not shown). The second structural disparity lies in the polyalanine blocks of the spidroin sequences. Figure 10B shows the distribution of the length of the polyalanine motifs for the two proteins. It can be seen that MaSpII contains longer polyalanine motifs than MaSpI. Indeed, 42.6% of the polyalanine segments of MaSpII are composed of 7−10 residues, whereas all the polyalanine segments of MaSpI contain 7 residues or less. From the present data, it may be suggested that longer polyalanine blocks prevent or slow down the formation β-sheets at the air−water interface. Interestingly, it is recognized that the inverse occurs in aqueous solution: the longer the number of alanine in a sequence, the higher the propensity to form β-sheets, with a value of 8−9 consecutive alanine residues constituting a threshold for the conformational conversion into β-sheets.[74,75] This behavior seems to be the counterpart of the present one occurring at the interface and is consistent with the fact that alanine is more hydrophobic than hydrophilic (0.7 on the normalized scales of Roseman and Kyte−Doolittle) and that the presence of lysine residues is required to allow the solubilization of polyalanine peptide in water.[61,74]

Conclusion

Our results show clearly that, due to their specific amino acid sequences, MaSpI and MaSpII adsorb with different kinetics and form films with different structures at the air−water interface. Films of MaSpII contain α-helices and less β-sheets than MaSpI due to the presence of proline residues. In conjunction with a lower glycine content, they also make MaSpII more conformationally constrained than MaSpI. Finally the length of the polyalanine motifs seems also to influence the assembly mechanism at the interface. Therefore, because of these particularities of the sequences of MaSpI and MaSpII, each protein probably has a specific behavior in solution. In particular, the fact that MaSpI and MaSpII readily aggregate at the interface suggests that they have a high propensity to form β-sheets, which may be important for the efficiency of the spinning process. It is therefore probable that the polyalanine sequences of both MaSpI and MaSpII are

(71) Xiong, H.; Buckwalter, B. L.; Shieh, H.-M.; Hecht, M. H. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6349–6353.
(72) Wang, W.; Hecht, M. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2760–2765.
(73) Jin, H.-J.; Kaplan, D. L. *Nature* **2003**, *424*, 1057–1061.
(74) Blondelle, S. E.; Forood, B.; Houghten, R. A.; Pérez-Paya, E. *Biochemistry* **1997**, *36*, 8393–8400.
(75) Shinchuk, L.; Sharma, D.; Blondelle, S. E.; Rixach, N.; Inouye, H.; Kirschner, D. A. *Proteins* **2005**, *61*, 579–589.

involved in the formation of β-sheets in MA silk and very unlikely that MaSpII is confined solely in the amorphous phase of the fiber, although it may play a particular role in the mechanical properties.

Acknowledgment. This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada, the Fonds Québécois de Recherche sur la Nature et les Technologies (FQRNT), and the Centre National de la Recherche Scientifique (CNRS) of France. The authors are also grateful to Costas Karatzas formerly with Nexia Biotechnologies Inc. for providing the recombinant proteins MaSpI and MaSpII and for helpful discussions, and to Christian Salesse for the use of the air−water interface ellipsometer.