Refolding of a High Molecular Weight Protein: Salt Effect on Collapse

D. Lairez,* E. Pauthe,† and J. Pelta†
*Laboratoire Léon Brillouin, Gif-sur-Yvette, France; and †ERRMECe, Université de Cergy Pontoise, Cergy, France

ABSTRACT Small-angle neutron scattering experiments were performed on dilute solutions of a high molecular weight protein (fibronectin, $M = 580$ kg/mol) in four cases: native conditions; unfolded state obtained by a denaturing agent (urea); and two badly refolded (or collapsed) states obtained by progressive elimination of the denaturing agent in salt-containing or salt-free solutions. Our main result is concerned by the conformation of the protein as the attempt for refolding is driven with or without salt. In salt-containing solution, we observe unambiguously that the protein chain collapses at large length scales but still obeys to a Gaussian statistics at short length scales. In other words, the globule embodies a large quantity of solvent compared to the compact situation. In salt-free solutions, the badly refolded protein is not globular but displays both a coil-like and an open conformation at large length scales and a local high density area. This behavior is discussed with respect to the scaling theories for polymers and polyampholytes.

INTRODUCTION

The understanding of protein folding remains a challenge. Proteins control most of the events of life and their function are mainly determined by their three-dimensional structure; thus, many diseases are linked to protein misfolding (Dobson, 1999; Radford and Dobson, 1999). Over the past three decades extensive experimental and theoretical studies have been reported on this problem (Honig, 1999; Dobson and Karplus, 1999; Radford, 2000; Pande et al., 2000). New views describe protein folding in terms of folding funnels or landscapes, in which the protein conformation moves on a minimum free energy pathway from unfolded conformations to a unique refolded and globular state (Dill and Chan, 1997; Dinner et al., 2000). For instance, small globular proteins, or individual modules, generally refold spontaneously to their correct functional conformation after removal of the denaturing agent (Privalov, 1979; Baldwin and Rose, 1999; Galzitskaya et al., 2001). Theoretical models have been proposed to describe this process, and computer simulations have been performed. However, the present calculation power limits this approach to low molecular weight proteins (Fersht and Daggett, 2002). Experimentally, protein refolding has been studied by small-angle x-ray or neutron scattering that are favored techniques giving directly the overall size, shape, and conformation of proteins in solution (Kataoka and Goto, 1996). Probably because experimental studies greatly benefit from theoretical results (and vice versa), in our knowledge none of them are concerned with high molecular weight ($>10^5$ g/mol) and multidomain proteins. The experimental study and the theoretical understanding of such protein refolding is, therefore, a next major challenge.

From the only point of view of the exact calculation of the free energy associated to each conformation, the study of high molecular weight proteins is increasingly complex. The number of possible conformations in a protein globule increases roughly exponentially with the number of amino acids and is rapidly beyond the scope of such approaches. Even so, this apparent complexity would allow benefit from progresses in statistical physics, and scaling theories applied to polymers, polyelectrolytes, and polyampholytes (Higgs and Joanny, 1991; Dobrynin and Rubinstein, 1995; Kantor and Kardar, 1995). To a certain extent, the behavior of a high molecular weight protein should be simpler than expected. Polyampholytes are chains made of neutral, positively, and negatively charged monomers. That theories predict the chain conformation depends on the solvent quality for the neutral backbone and on the electrostatic interactions between charges, i.e., fraction of charged monomers, pH, and added salt. All these parameters introduce different length scales in the description of the chain conformation, rapidly making the scaling approach irrelevant for small proteins.

This article deals with fibronectin that is a high molecular weight, multidomain glycoprotein of the extracellular matrix ($M = 530$ kg/mol calculated from the primary structure; see Ruoslahti, 1988). Fibronectin is composed of two identical subunits covalently linked by one of their extremities; fibronectin itself consists of the repetition of 56 globular modules of three different types having a well-defined structure and a high degree of internal homology (Yamada, 1989). X-ray and NMR studies of individual modules of each type (I, II, and III) have shown that their secondary structure is mainly β-sheets (Dickinson et al., 1994; Williams et al., 1994; Sticht et al., 1998). However, the crystallization of the whole protein has never been reported, and thus the three-dimensional structure of whole fibronectin is undetermined. We have shown in a previous article that it is actually more convenient to speak about statistical conformation of fibronectin (Pelta et al., 2000).

Made of a little less than 5000 amino acids, fibronectin is a good candidate to test the relevance of scaling theories for

Submitted September 25, 2002, and accepted for publication January 15, 2003.

Address reprint requests to J. Pelta. Tel.: 33-13-425-6664; Fax: 33-13-425-6552; E-mail: Juan.Pelta@bio.u-cergy.fr.

0006-3495/03/06/3904/13 $2.00

understanding high molecular weight protein folding. Here, we report a small-angle neutron scattering (SANS) study performed on dilute solutions. In native condition, this protein adopts the conformation of a flexible string of 56 globules of 25-Å radius each (Pelta et al., 2000). In 8 M urea solution (denaturing condition) the protein is unfolded and swells as a linear polymer in good solvent. Our key observation is the following: as urea is slowly removed from the solution, fibronectin does not recover its native conformation; and two different collapsed conformations have been clearly identified, depending on the added salt concentration. At physiological ionic strength (150 mM NaCl), the protein collapses at large length scale but remains Gaussian at small length scales, and the globule still contains a large amount of solvent. In salt-free solution, the badly refolded protein is not globular but displays both a coil-like and an open conformation at large length scales and a local high density area. We will see that polyampholyte theories reasonably account for this behavior.

SAMPLES PREPARATION AND THE SANS EXPERIMENTS

Fibronectin was purified from human cryoprecipitated plasma as described by Poulouin et al. (1999), and Pelta et al. (2000). The purity of the preparation was determined equal to 96.6 ± 1.2% by densitometry analysis of silver nitrate-stained SDS-polyacrylamide gel electrophoresis. In this article we describe SANS experiments performed on seven protein samples that experienced different physico-chemical conditions (see Fig. 1). For each sample, 3 cm^3 of fibronectin solution at a concentration of 8 mg/cm^3 in H_2O with 10 mM Tris-HCl (pH = 7.4), were first dialyzed during 24 h, at room temperature using a microdialysis cassette (Slide-A-Lyzer, 10,000 Mw cutoff; Pierce, Rockford, IL) against 100 cm^3 of solution containing either D_2O with 10 mM Tris-DCl (pH = 7.4) and 150 mM NaCl (samples 1–3), or pure D_2O at a measured pH of 7 (samples 4–7). A second dialysis with the same volume ratio was performed against

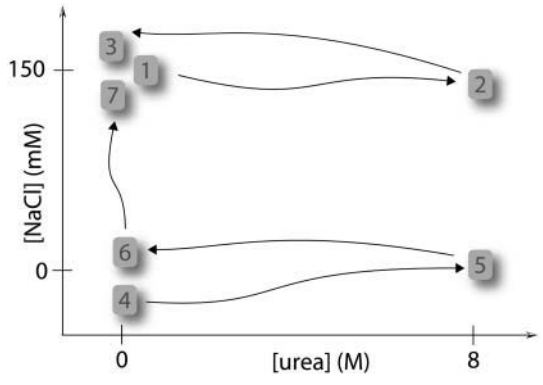

FIGURE 1 The seven samples here studied and their history.

D_2O with 10 mM Tris-DCl (pH = 7.4) and 150 mM NaCl (sample 1); or 8 M deuterated urea in D_2O with 10 mM Tris-DCl (pH = 7.4) and 150 mM NaCl (samples 2 and 3); or pure D_2O (sample 4); or 8 M deuterated urea in pure D_2O (samples 5–7). For samples 3, 6 and 7, urea was slowly eliminated by successive dialyzes against 100 cm^3 of solutions at the same ionic strength but decreasing deuterated urea concentration (6, 4, 2, 1, and 0 M) for a total dialysis duration of three days. Finally, for sample 7, 150 mM NaCl and Tris-DCl 10 mM were added in the solution. For each sample, fibronectin concentrations were measured after dialysis by optical absorbance measurements of $A^{1\%}_{280nm} = 12.8$, $A^{1\%}_{280nm}(\text{urea } 8\,\text{M}) = 14$. Here, it has to be noted that samples 4–6, in order to reach the lowest ionic strength possible at this protein concentration (8 mg/cm^3), do not contain any added salt or Tris buffer. For these samples, it was checked that protein buffering would be enough to ensure a measured pH always between 7 and 7.5. Although these pH variations are important, they do not allow modification of the charges distribution on the protein. SANS experiments in dilute solutions require, in most cases, the use of D_2O as solvent, and D_2O itself is known to affect folded-unfolded transitions. But only slight effects are expected (Russo et al., 2001). As far as this article is concerned, by extreme conformations in the phase diagram (0 M urea or 8 M urea) and not by intermediate states, such D_2O effects can be reasonably neglected.

In samples 1 and 4, fibronectin is at physiological pH, and has never undergone the denaturing condition; the corresponding protein will be called *native*. In samples 2 and 5, fibronectin was denatured by 8 M urea; it will be called *unfolded*. In samples 3 and 6, urea was eliminated, and these samples will be called *refolded*. Sample 7 has undergone the same history as sample 6, and NaCl was added only after refolding. We will refer to this sample as *refolded in salt-free solution with added salt*.

SANS experiments were performed on the PACE spectrometer at the LLB Institute, Saclay, France. To cover the widest range of scattering vector q, three different configurations were used for the sample-to-multidetector distance and for the wavelength: A), small q-range 4.7 m/15 Å; B), intermediate q-range 2.3 m/6 Å; and C), high q-range 0.6 m/4 Å. Data treatment was carried out following Cotton (Cotton, 1991). Measurements performed using the high q configuration were used for incoherent scattering subtraction. For absolute measurements, the contrast factor, K^2, was calculated, neglecting protein glycosylation, and following Jacrot for the values of the different amino-acid-specific volumes and effective H-D exchanges (Jacrot, 1976). For native fibronectin in D_2O, $K^2 = 1.10\,10^{-3}\,cm^2\,g^{-2}\,mol$, assuming that 80% of labile hydrogens is actually exchanged. For fibronectin in 8 M urea, $K^2 = 1.22\ 10^{-3}\ cm^2\ g^{-2}$ mol, assuming that all labile hydrogens are exchanged. For measurements in D_2O after denaturation (the *refolded* protein) $K^2 = 0.93\ 10^{-3}\ cm^2\ g^{-2}$ mol, assuming that all

labile hydrogens remain exchanged. For these measurements, protein solutions were contained in Boron-free quartz cuvettes of 5-mm thickness. At 6 Å, typical sample transmissions lie between 0.60 and 0.65 for this protein concentration. These transmission values, compared to the value 0.66 measured for the same thickness of pure D_2O, allow us to neglect coherent multiple scattering (incoherent multiple scattering being easily subtracted by the usual data treatment; see above).

THEORETICAL BACKGROUND OF THE SANS EXPERIMENTS

An incident plane wave of wavelength λ is scattered by the elementary scatterers (monomers of a macromolecule, small solvent molecules, etc.) of the sample. The interference pattern is measured as a function of the scattering vector $q = (4\pi/\lambda)\sin(\theta/2)$, with θ as the scattering angle. The measurement of the scattered intensity normalized to the incident intensity and expressed by the solid angle unit leads to the differential scattering cross section of the sample

$$\Sigma(q) = \sum_{\mathrm{i}}\sum_{\mathrm{j}} b_{\mathrm{i}} b_{\mathrm{j}} e^{\mathrm{iq}(\mathrm{r_i}-\mathrm{r_j})},$$

where i and j refer to different elementary scatterers, of scattering length b and position r. For macromolecules in solution, $\Sigma(q)$ has two contributions. The first one is the coherent differential scattering cross section, $\Sigma_{\mathrm{coh}}(q)$. It is the macromolecules contribution to the scattered intensity due to concentration fluctuations. The second contribution, Σ_{inc}, is the q independent incoherent scattering. $\Sigma_{\mathrm{coh}}(q)$ which is the quantity of interest, is globally a decreasing function and Σ_{inc} leads to a baseline which can be measured at sufficiently high q and thus properly subtracted. For a solution of n identical macromolecules, each one being made of N identical elementary scatterers, $\Sigma_{\mathrm{coh}}(q)$ can be expressed by the sum of interferences between scatterers belonging to the same macromolecules and interferences between scatterers belonging to different macromolecules. The normalized scattering function $F(q)$ is

$$F(q) = \frac{\Sigma_{\mathrm{coh}}(q)}{(\Delta b)^2 n N^2} = P(q) + nQ(q), \tag{1}$$

where Δb is the difference between the scattering length of one elementary scatterer of the macromolecule and the same volume of solvent.

$$P(q) = \frac{1}{N^2}\sum_{\mathrm{i}h}^{\mathrm{N}}\sum_{\mathrm{j}h}^{\mathrm{N}} e^{\mathrm{iq}(\mathrm{r}_{ih}-\mathrm{r}_{jh})},$$

and

$$Q(q) = \frac{1}{N^2}\sum_{\mathrm{i}h}^{\mathrm{N}}\sum_{\mathrm{j}k\neq h}^{\mathrm{N}} e^{\mathrm{iq}(\mathrm{r}_{ih}-\mathrm{r}_{jk})},$$

are normalized quantities which tend to unity at $q = 0$ and h and k refer to different macromolecules. In practice, the intermolecular term, $Q(q)$, can be neglected in dilute solutions ($n \to 0$) and at sufficiently high q-values compared to the reciprocal of the characteristic size, R^{-1}, of macromolecules ($qR > 1$). Then one accesses directly to the form factor or intramolecular term, $P(q)$. In the case of macromolecules with a spherical symmetry, Eq. 1 can be factorized introducing the center of mass, h and k, of macromolecules.

$$F(q) = S(q) \times P(q), \tag{2}$$

where

$$S(q) = \frac{1}{n}\sum_{\mathrm{h}}^{\mathrm{n}}\sum_{\mathrm{k}}^{\mathrm{n}} e^{\mathrm{ig}\vec{\mathrm{hk}}}$$

is the structure factor reflecting correlations between the centers of mass of the macromolecules. At $q \to \infty$, $S(q)$ tends to 1, and in practice this limit is reached at high q value compared to R^{-1}:

$$S(q)_{\mathrm{qR}\gg 1} = 1. \tag{3}$$

For $q \to 0$, molecules are pointlike, and $S(q)$ is proportional to the osmotic compressibility,

$$\kappa_{\mathrm{T}} = \left(C\frac{d\pi}{dC}\right)^{-1},$$

where C is the macromolecule concentration

$$S(q \to 0) = \frac{n}{V} \times kT\kappa_{\mathrm{T}}, \tag{4}$$

with kT the thermal energy.

Experimentally, the scattering function is deduced from the measured coherent scattered intensity per unit volume, I_{coh} expressed in cm^{-1}, following the equality

$$F(q) = \frac{\Sigma_{\mathrm{coh}}(q)}{(\Delta b)^2 n N^2} = \frac{I_{\mathrm{coh}}(q)}{K^2 C M}, \tag{5}$$

where $K^2 = (\Delta b)^2/m^2$ is the contrast factor, m the molecular mass of the elementary scatterers of scattering length b, C the concentration of macromolecules (mass per volume unit), and M their molecular mass.

$P(q)$ gives information about the internal structure and conformation of a single macromolecule. For molecules that can adopt different conformations, $P(q)$ is averaged over all the possible conformations. For the sake of clarity, the corresponding notation for this averaging is here omitted. In solution, macromolecules are freely rotating and thus $P(q)$ has to be averaged over all the possible orientations with respect to the scattering vector. This averaging leads to:

$$P(q) = \frac{1}{N^2}\sum_{\mathrm{i}}^{\mathrm{N}}\sum_{\mathrm{j}}^{\mathrm{N}} \frac{\sin qr_{\mathrm{ij}}}{qr_{\mathrm{ij}}}, \tag{6}$$

with r_{ij} the distance between two scatterers i and j.

At small q such as $qr_{\mathrm{ij}} < 1$, whatever the shape of the macromolecules, a Taylor's expansion of $\sin qr_{\mathrm{ij}}/qr_{\mathrm{ij}}$ leads to the Guinier's formula:

$$P_{\text{Guinier}}(qR_g) = 1 - \frac{(qR_g)^2}{3} + \cdots, \quad (7)$$

where

$$R_g^2 = \frac{1}{2N^2}\sum_i^N\sum_j^N r_{ij}^2,$$

is the square radius of gyration of macromolecules. This expression allows us to determine the radius of gyration. However, different representations are used allowing a better linearization of the data (Higgins and Benoît, 1994). For polymer chains, $1/P(q)$ versus q^2 is preferred (Zimm representation), whereas for dense particles $\ln(P(q))$ versus q^2 gives better results (Guinier representation).

At higher q-values, measurements become sensitive to the shape of the molecules. A dense sphere of radius, R, gives

$$P_{\text{sphere}}(qR) = \left[3\frac{\sin x - x\cos x}{x^3}\right]^2 \quad \text{with } x = qR, \quad (8)$$

with characteristic oscillations at $qR > 1$ which decrease in amplitude as $(qR)^{-4}$. An ideal chain obeys to Gaussian statistics—it corresponds to a random walk of N steps of length b. At length scale q^{-1} large enough compared to b, the form factor obeys to the Debye function:

$$P_{\text{Debye}}(qR_g) = \frac{2}{x^2}(e^{-x} - 1 + x) \quad \text{with } x = (qR_g)^2, \quad (9)$$

which displays a q^{-2} asymptotic behavior at $qR_g > 1$ reflecting a fractal dimension equal to 2. In the case of excluded volume interactions between monomers, one would rather expect a swollen conformation that decreases the fractal dimension down to $D = 5/3$. In this case, it can be useful to adopt an ad hoc expression that has the two correct asymptotic behaviors, in the Guinier range and at $qR_g > 1$:

$$P(qR_g) = \frac{1}{(1 + x^2)^{D/2}} \quad \text{with } x = qR_g. \quad (10)$$

In some cases, the object under study is made of n subsets of N scatterers each. The general expression of the form factor

$$P(q) = \frac{1}{(nN)^2}\sum_{i,j}^{nN} e^{iq(r_i - r_j)},$$

can be expressed introducing the centers of mass (h, k) of each subset:

$$P(q) = \frac{1}{(nN)^2}\sum_h^n\sum_k^n\sum_{ih}^N\sum_{jk}^N e^{iq\times(\vec{ih} + \vec{kj})} e^{iq\times\vec{hk}}.$$

In the particular case of subsets having a spherical symmetry, this latter expression can be factorized into two terms. The first, P_{subset}, is the form factor of one subset and the other reflects the correlation between the n centers of mass:

$$P(q) = P_{\text{subset}}(q) \times S(q), \quad (11)$$

with

$$P_{\text{subset}}(q) = \frac{1}{N^2}\sum_{ih}^N\sum_{jh}^N e^{iq(r_{ih} - r_{jh})},$$

and

$$S(q) = \frac{1}{n}\sum_{ih}^n\sum_{jk}^n e^{iq\times\vec{hk}}.$$

We have shown previously that the form factor of native fibronectin is adequately accounted for using the conformation of a Gaussian string of 56 spherical beads. This model is a particular case of Eq. 11 with $P_{\text{subset}}(q)$ taken as equal to the form factor of a sphere of radius $b/2$ (see Eq. 8) and $S(q)$ the form factor of a random walk:

$$P(q) = P_{\text{sphere}}\left(q\frac{b}{2}\right) \times S(q). \quad (12)$$

In this case, the finite length, b, of a step cannot be neglected and $S(q)$ cannot be approximated by $P_{\text{Debye}}(q)$ but must be obtained by computer simulations. Eq. 11 can also be used to account for the high q cutoff of q^{-D} power laws (Eqs. 9 and 10). This is valid for unfolded proteins because their persistence length (the length above which correlations between orientations of chain segments vanish) and chain thickness are of the same order of magnitude (Russo, 2000).

RESULTS

Salt-containing solutions

In Fig. 2, the measured coherent scattered intensity per concentration unit and the contrast unit ($I_{coh}/(CK^2)$, are plotted as functions of the scattering vector q for the samples in salt-containing solution (samples 1–3 and 7, see Fig. 1). These four solutions have an ionic strength allowing us to

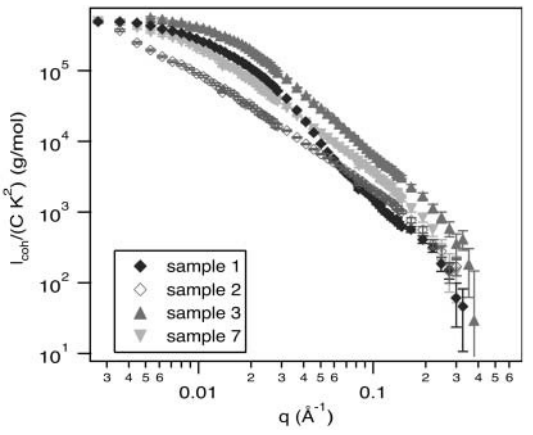

FIGURE 2 Scattered intensity per concentration unit and contrast unit as a function of scattering vector q for samples in salt-containing solution (see Fig. 1). (*Sample 1*) native fibronectin; (*sample 2*) unfolded fibronectin; (*sample 3*) refolded in salt-containing solution; and (*sample 7*) refolded in salt-free solution and added salt after refolding.

fully screen electrostatic interactions between proteins (see Discussion). In addition, these four samples were prepared at the same fibronectin concentration of 8 mg/cm^3. By light-scattering measurements, we have previously shown (Pelta et al., 2000) that fibronectin at this concentration is in dilute solution: in native condition, the two-body interaction term is un-measurable, whereas in an 8 M urea solution it leads to an overlap concentration equal to 17 ± 1 mg/cm^3. This result and the fact that in this article we deal with a high q-range ($qR_g > 1$), allow to consider the measured scattering function as mainly due to the form factor of a single molecule. This point is confirmed by the extrapolation at $q = 0$ of the scattered intensity which leads, within 10%, to a good estimation of the fibronectin molecular mass already measured by light scattering (Pelta et al., 2000): 580 ± 30 kg/mol. The measured value is in rather good agreement with the value, 530 kg/mol, calculated from the fibronectin primary structure. The calculated value is actually increased by glycosylation that contribute to ~4% of the molecular mass. Excepted for sample 2 (unfolded fibronectin), the form factor, $P(q)$, of fibronectin can be deduced from the ratio of these spectra to the extrapolated value at $q = 0$. As for the unfolded protein, due to a too-high radius of gyration, extrapolation at $q = 0$ cannot be performed from SANS measurement, the normalization of the spectra was done using the apparent molecular mass measured by light scattering at the same concentration (Pelta et al., 2000). A first examination of Fig. 2 shows that the spectra do not superpose at small q ($q < 0.02$ Å^{-1}). This indicates an important variation of the size of the protein in the four cases. This point is emphasized in Fig. 3, where the reciprocal of $P(q)$ is plotted as a function of q^2. Important variations of the slopes of the curves indicate clearly a variation of the characteristic size of the protein. For the 8 M urea solution of fibronectin (sample 2), the radius of gyration is too high to be determined by SANS experiments. The value reported in this article was previously determined by light-scattering measurements (Pelta et al., 2000). In the case of the native fibronectin sample (sample 1), the radius of gyration was determined using both light- and neutron-scattering techniques. For the SANS measurements, we have shown (Pelta et al., 2000) that an approximation of the Debye equation (Eq. 9) allows us to fit the spectrum in an extended Guinier q-range up to $qR_g = 4$. The R_g value so-determined is in excellent agreement with the light-scattering result. As for refolded fibronectin in salt-containing solution (sample 3) and in salt-free solutions with added salt (sample 7), the particular shape of the spectra in both cases allows us to fit the curves up to $qR_g = 2$ and 4, respectively. This point will be examined in more detail in the following, as the shape of the form factors are discussed. In Table 1 the measured values for the radius of gyration are reported for the four samples.

To emphasize the discrepancies between the different spectra, $P(q) \times (qR_g)^2$ as a function of qR_g is plotted in Fig. 4. Note that in this representation, uncertainty of R_g would only lead to a shift of the curves along the straight line which corresponds to $(qR_g)^2$. In the following, we will discuss in detail the shape of these curves.

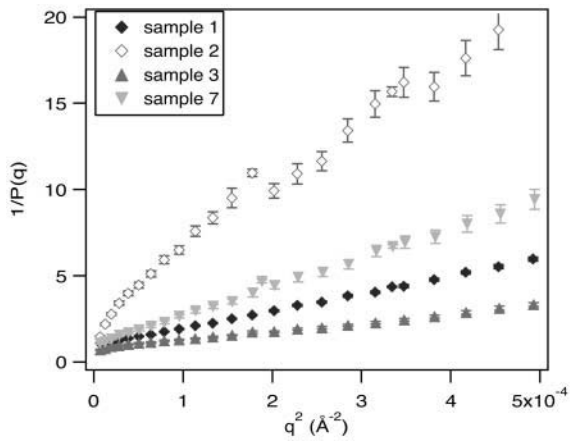

FIGURE 3 Reciprocal of the form factor, $1/P(q)$, as a function of the square scattering vector q^2 for samples in salt-containing solution. Important variations of the slope are characteristic of changes in the radius of gyration. The symbols' meaning is the same as in Fig. 2.

Salt-free solutions

In Fig. 5, the spectra measured for salt-containing solutions are compared to their counterpart measured in salt-free solution, each spectra being normalized by the scattered intensity extrapolated at $q = 0$ measured in salt-free solution. For salt-containing solutions, this ratio is mainly due to the form factor of the protein, whereas in salt-free solution an interaction peak appears due to unscreened repulsive electrostatic interactions responsible for a decrease of the osmotic compressibility and a corresponding decrease of the scattered intensity at $q \rightarrow 0$ (see Eq. 4). Actually, the isoelectric point of fibronectin is ~pH = 5 and the net charge of the protein is thus negative at pH = 7.4. In solution of hard spheres, repulsive interactions are responsible for a most probable distance, d, between first neighbors, leading to a peak in the structure factor at $q_m = 2\pi/d$. At the concentration used for these experiments, one calculates $d = (n/V)^{-1/3} = 475 \pm 25$ Å, leading to $q_m = (1.30 \pm 0.07)10^{-2}$ Å^{-1}, a value in agreement with our data. Due to these strong interactions, the structure factor cannot be neglected and the form factor cannot be approximated by the

TABLE 1 Measured radius of gyration for the four samples in salt-containing solution

		R_g (Å)
Sample 1	Native condition	153 ± 2
Sample 2	Unfolded in 8 M urea solution	300 ± 10
Sample 3	Refolded in salt-containing solution	88.5 ± 3.5
Sample 7	Refolded in salt-free solution with added salt	220 ± 5

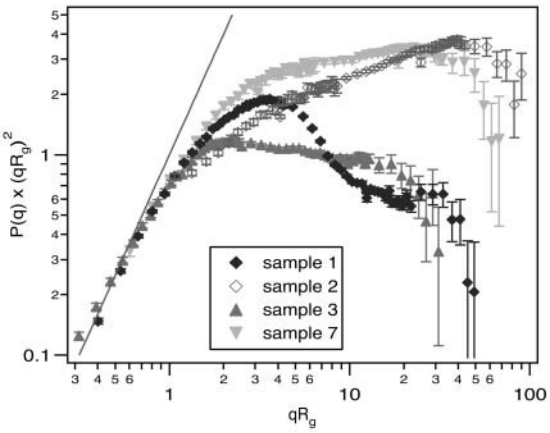

FIGURE 4 Form factors of samples in salt-containing solution in a Kratky representation $P(q) \times (qR_g)^2$ versus qR_g. The symbols meaning is the same as in Fig. 2. The straight line corresponds to $(qR_g)^2$.

scattered intensity profile as for salt-containing solutions. If the shape of the molecules remains unaffected by salt addition, the ratio of measurement performed for salt-free solution to that performed for salt-containing solution would lead to the intermolecular structure factor $S(q)$ (see Eq. 2). However, in the case of macromolecules having a soft conformation, intramolecular interactions distort their shape. This is typically the case of unfolded protein (samples 2 and 5) but also the case of native fibronectin (samples 1 and 4) because of its particular flexible string-of-beads conformation. Consequently, the corresponding scattered intensity spectra are more complex to interpret. This point will be discussed in the following section.

DISCUSSION

Two characteristic lengths have to be kept in mind. The Bjerrum length, l_B, is the distance at which the Coulomb energy between two charges is kT. In water at room temperature, $l_B = e^2/(4\pi\varepsilon_r\varepsilon_0 \times kT) = 7.2$ Å. The ionic strength, I, screens electrostatic interactions beyond the length $\kappa_0^{-1} = (4\pi l_B I)^{-1/2}$. Taking into account only added salt, $\kappa_0^{-1} = 10.7$ Å in our salt-containing solutions.

Native and unfolded states in salt-containing solutions

In the salt-containing solutions, one can assume that electrostatic intermolecular interactions are screened. In dilute solution and in the q-range of our measurements, the scattering intensity spectra give the form factor of the molecule. We have already reported (Pelta et al., 2000) the statistical conformation of fibronectin in native condition and 8 M urea solution in salt-containing solution (samples 1 and 2). Let us briefly recall these results. In Fig. 6, the two spectra are plotted in a Kratky representation and compared to the theoretical expectations. For native fibronectin, one can reasonably

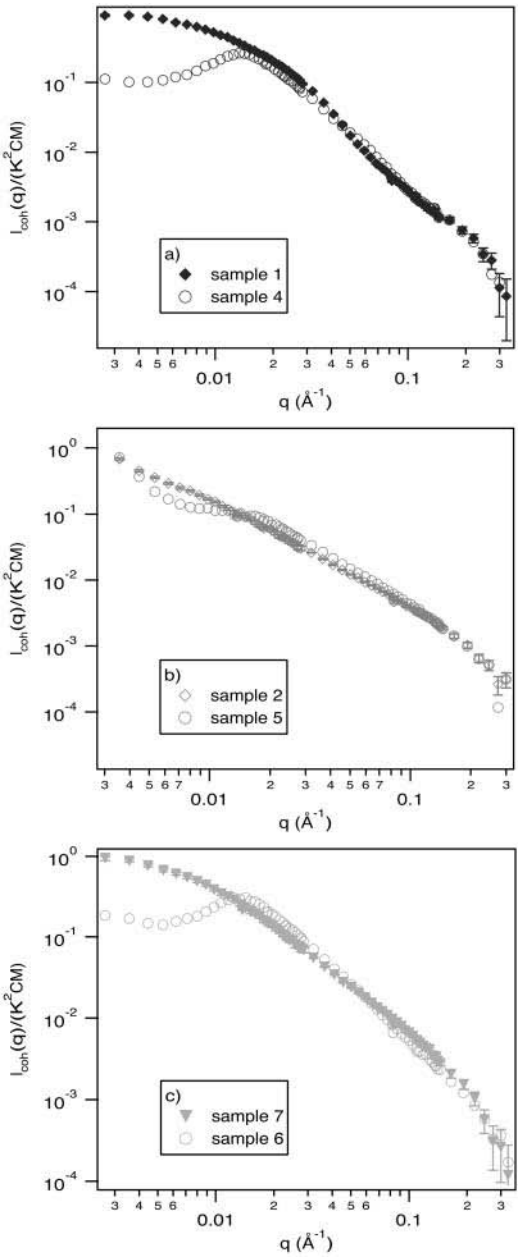

FIGURE 5 Scattering function versus q. (*a*) Native fibronectin in salt-containing and salt-free solution; (*b*) unfolded fibronectin in salt-containing and salt-free solution; and (*c*) refolded fibronectin in salt-free solution with and without added salt. For the sake of clarity, error bars are only displayed for salt-containing solutions; however, in both cases they are of the same order of magnitude.

account for the spectrum using a flexible string-of-spherical-beads (or dense globules) model that is obtained following Eq. 11. One has to note that in Fig. 6, the corresponding full line is not a fit of the data, but results from a model without any adjustable parameter. The number of 56 beads is the number of modules fibronectin is made of (modules mainly defined as being partly resistant to proteolysis). These modules are of three different types and sizes. They are separated by small polypeptide segments (highly sensitive to proteolysis) and

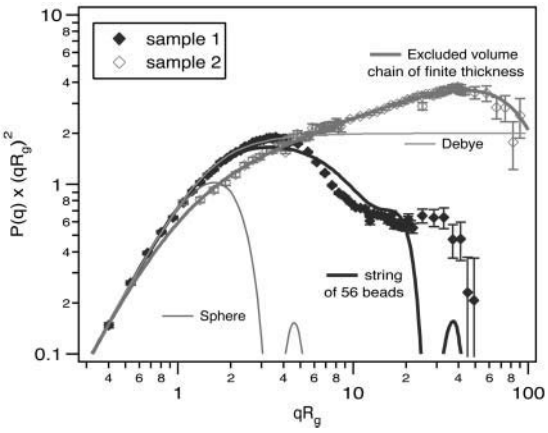

FIGURE 6 Form factor of the native and unfolded fibronectin plotted in a Kratky representation: $P(q) \times (qR_g)^2$ versus qR_g. The full lines correspond to the theoretical expectation for a sphere (Eq. 8), a string of 56 beads (Eq. 12), a Gaussian chain with infinitely small monomer (Debye function, Eq. 9), and a swollen chain with a finite thickness (Eq. 13).

should be more adequately described using an oblong shape. Our crude string-of-beads model does not take these refinements into consideration. From this model and the measured radius of gyration, one deduces an average radius of the order of 25 Å for each globule of the string.

As for fibronectin in 8 M urea, $P(q)$ shows a near-$q^{-5/3}$ behavior at $qR_g > 1$ that is characteristic of a chain with an excluded volume conformation. $P(q)$ can be described in the whole q-range by the product of Eq. 10 and Eq. 8 (or 7) to account for the finite thickness and size of monomers. In Fig. 6, the full line corresponds to

$$P(q) = \frac{1}{(1 + (qR_g)^2)^{D/2}} \times P_{\text{sphere}}(a \times qR_g) \quad \text{with} \quad \begin{cases} D = 1.65 \pm 0.05 \\ a = (22 \pm 1) \times 10^{-3} \end{cases} . \tag{13}$$

Here, the high q behavior is fitted using a sphere form factor rather than the Guinier approximation, but this is only to fit the data up to $qa = 2$ and we do not claim that the monomer is spherical. The parameter a is no more than the average monomer size in the longitudinal and transverse directions with respect to the chain sequence, i.e., persistence length and chain half-thickness, respectively. Actually, using a more general expression for the form factor (Pedersen and Schurtenberger, 1996) of swollen chain, it has been shown for unfolded protein that these two lengths are of the same order of magnitude (Russo, 2000). This justifies the use of Eq. 13. In this equation, a is expressed in R_g unit. The actual value of R_g (see Table 1) leads to

$$a = (6.6 \pm 0.5)\,\text{Å}. \tag{14}$$

This result is in good agreement with the value reported in the literature for unfolded protein thickness and/or persistence length (Russo, 2000). Eq. 14 leads to a Kuhn length (i.e., the length of the statistical segment that is twice the persistence length) $b = 2a = 13.2$ Å for the unfolded fibronectin, which has to be compared to $\kappa_0^{-1} = 10.7$ Å and is consistent with the assumption of screened electrostatic interactions. Assuming a length of 3.45 Å for each amino acid (Russo et al., 2000), from the value for b the unfolded protein is a flexible chain made of 1250 statistical segments.

Native and unfolded states: salt effect

A preliminary question is whether the conformations of protein in native and unfolded states are affected by salt removing. As mentioned above, the form factors of molecules in salt-free solution cannot be deduced from measurements. Only qualitative comparisons of experiments with theoretical expectations can give a general idea of structure changes with salt removing. In Fig. 7 the salt effect on the scattering spectra is shown in a Kratky representation.

First, in the case of native and unfolded protein, notice that the spectra match, by pairs, at high q; i.e., the measurements for native protein with and without salt (samples 1 and 4) on one hand and the measurements on unfolded protein (samples 2 and 5) on the other hand. This indicates that the local conformation of the protein is not affected by the ionic strength but depends on urea concentration. In particular, fibronectin is presumably just as unfolded without salt (sample 2) as with salt (sample 5). Another argument supporting this idea concerns the structure factor. A major difference between interaction peaks observed for dense structure (hard spheres) and those observed for unfolded chains (polyelectrolytes) is that the former increases in intensity with concentration whereas the latter decreases (Stevens and Kremer, 1995; Shew and Yethiraj, 1999, 2000). Thus, at a given concentration one expects a less marked peak for an unfolded conformation. This is what we observe in Fig. 5, *a* and *b*.

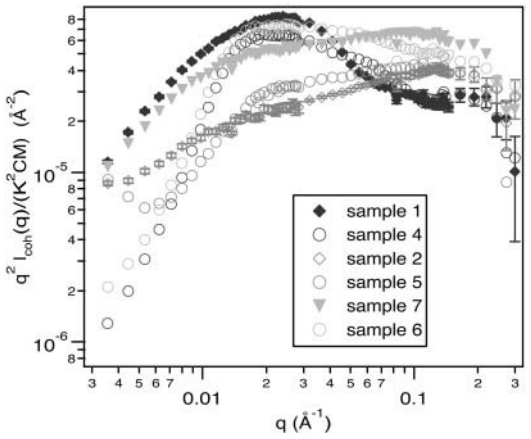

FIGURE 7 Same data as in Fig. 5, *a* and *b*, but in a Kratky representation ($q^2 I_{coh}$ versus q).

Refolded states

Starting from an unfolded conformation a progressive elimination of urea does not allow fibronectin to recover its native conformation. This point appears clearly in Fig. 4 where the spectra, obtained after two different attempts for protein refolding (with and without salt), are quite different from the spectrum obtained in native condition. In salt-containing and urea solution (sample 2), fibronectin is unfolded and adopts a swollen conformation similar to that of a neutral polymer chain in good solvent. As urea is removed from the solution, fibronectin partly collapses as proved by the decreasing radius of gyration (see Table 1). This collapse has to be considered with respect to the theoretical expectations (Higgs and Joanny, 1991; Dobrynin and Rubinstein, 1995; Kantor and Kardar, 1995) for polyampholytes chains. From theory, different terms are involved in the free energy governing the conformation of a chain:

Neutral backbone contribution. 1), Conformational entropy of the chain which acts as a return spring constant to the Gaussian chain conformation; and 2), volume interactions of the neutral backbone, i.e., two-body (either attractive or repulsive depending on the solvent quality), and three-body (always repulsive) interactions.

Charges contribution. 3), Repulsive electrostatic interactions due to the net charge of the chain (asymmetry between charges of opposite sign); and 4), attractive electrostatic interactions due to fluctuations in the local charge distribution leading to sequences of opposite sign along the chain.

Expressions for the two later contributions (terms 3 and 4) are controversial. However, at least in the simple case of neutral polyampholytes, the different theories agree. In that case, electrostatic interactions reduce to the attractive term 4 that can be included within an effective two-body volume interaction term (Higgs and Joanny, 1991). The chain is expected to collapse in a globule with an internal concentration resulting from a balance between attractive two-body and repulsive three-body interactions (the conformational entropy, term 1, being negligible). The former need a minimum volume (or distance) to be significant, meaning that small chain segments remain Gaussian, whereas the overall chain collapses at large length scale. This is also expected for neutral polymer chains in poor solvent (de Gennes, 1996). In both cases, the resulting chain conformation is the one of a compact set of Gaussian chain segments of size ξ. In other words, the chain statistics remain Gaussian up to the length scale ξ but correspond to an homogenous concentration distribution beyond ξ. Such chain segments of size ξ are named ''blobs'' by polymer physicists (de Gennes, 1996). In the case of polyampholyte, ξ is mainly governed by electrostatic attractions and their screening. Let us denote $\kappa_0 = (4\pi l_B I)^{1/2}$ the reciprocal of the screening length due to added salt, and κ_p the one inside the globule due to polymer charge concentration, $\kappa_p = (4\pi(2\varphi N l_B)/R^3)^{1/2}$, where N is the number of statistical segments of the chain, R the radius of the globule, and 2φ the total fraction of charged segments (φ positive and φ negative). Depending on κ_0/κ_p, two regimes are expected (Higgs and Joanny, 1991). At low salt concentration

$$\xi_{\kappa_0 < \kappa_p} \propto b \times \frac{b}{\varphi l_B}, \tag{15}$$

whereas at higher salt concentration, the added salt governs the screening and reduces the attractive forces. ξ increases with the added salt concentration as

$$\xi_{\kappa_p < \kappa_0} \propto b \times \kappa_0 \frac{b^3}{\varphi^2 l_B^2}. \tag{16}$$

In both cases, denoting g the number of statistical segments per blob, the Gaussian conformation at small length scales leads to $g = 6(\xi/a)^2$ (the factor 6 comes from the ratio of the end-to-end distance to the radius of gyration of a Gaussian chain). The large scale dense structure writes

$$(R/a)^3 = \frac{N}{g}(\xi/a)^3.$$

Thus, the volume fraction of the chain inside the globule scales as

$$\phi^* = N/(R/a)^3 = \frac{3}{2}(\xi/a)^{-1}. \tag{17}$$

In the case of non-neutral or asymmetric polyampholytes in salt-free solution, polyelectrolyte-like repulsions are expected to highly distort the shape of the collapsed globule at large length scales. A cigar-shape, rather than a spherical conformation, minimizes the free energy (Dobrynin and Rubinstein, 1995). However, the cigar-shape is unstable and splits into smaller spherical globules linked by narrow strings (Kantor and Kardar, 1995) leading to a necklace-conformation. Depending on the net charge of the chain (long-range repulsions), on the solvent quality and on charge asymmetry (short-range attractions), cascade of transitions is expected between necklace conformations of various numbers of beads (Dobrynin et al., 1996; Chodanowski and Stoll, 1999).

Our SANS data interpretations are guided by these expected different shapes for the collapsed protein.

Protein refolded in salt-containing solution

At sufficiently high salt concentration, one expects a spherical and globular overall conformation with a remaining Gaussian conformation up to the length scale ξ. However, even if scaling concepts would apply, due to the finite size of the unfolded chain and with regard to the three different length scales involved (segment length b, blob size ξ, overall chain size R) significant power law behaviors characteristic of the different regimes for the chain conformation cannot be

observed on the form factor. Thus, the compact set-of-Gaussian-blobs picture was tested by Monte Carlo computer simulations based on the observation that a chain confined in a sphere adopts the same conformation but for a different physical reason. The corresponding $P(q)$ was computed as follows.

1. A random walk of 1250 steps is generated (this number comes from the unfolded conformation results; see Native and Unfolded States in Salt-Containing Solutions) with the constraint that the walker cannot reach a distance larger than R_c (radius of confinement) from the center of mass of the walk. This is obtained by reflection of the walk as R_c is reached.
2. $P(q)$ and R_g of the random walk are computed by averaging over all orientations (see Eq. 6).
3. $P(q)$ and R_g of the random walk are averaged over several iterations depending on R_c. For small R_c values, the number of possibilities for the chain conformation is small and the calculation of $P(q)$ converges rapidly in less than 100 iterations, whereas high R_c values require up to 10^4 iterations.
4. To account for the finite width and radius a of statistical segments of the chain, the averaged $P(q)$ of the random walk is multiplied by the form factor of a sphere of radius a (Eq. 11 and 13).

The resulting form factors are plotted in Fig. 8. The two border full lines correspond to the theoretical form factors of Gaussian chain and compact sphere, respectively, whereas the intermediate full lines correspond to a progressive confinement of the chain. The measured spectra for fibronectin refolded in salt-containing solution nicely fits within the bundle of confined chain curves. To free from the discrete number of simulations, i.e., discrete number of R_c,

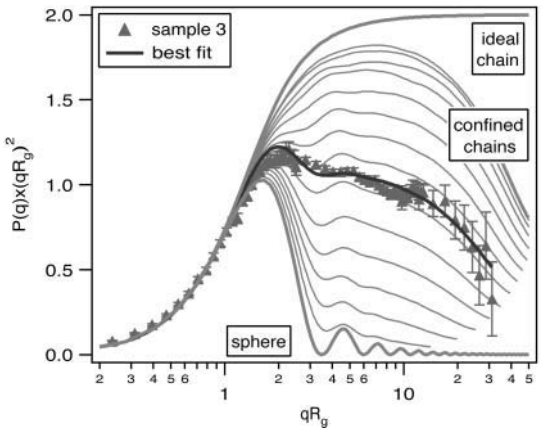

FIGURE 8 $P(q) \times (qR_g)^2$ versus qR_g for fibronectin refolded in salt-containing solution (*data points*) compared to confined chains (*lines*). (*Top* to *bottom*), lines correspond to: Debye function (unconfined Gaussian chain with infinitely small monomer); confined chains with radius of confinement, R_c from 30- to 4-step-length unit; and compact sphere. The line fitting the data corresponds to $\overline{R_c} = 16 \pm 2$ step-length unit.

the experimental data were fitted using a linear combination of the spectra obtained by computer simulation weighted by a Gaussian distribution in R_c:

$$p(R_c) = \left(\sqrt{2\pi\sigma}\right)^{-1/2} e^{-(R_c-\overline{R_c})^2/(2\sigma^2)},$$

where one gets $\overline{R_c}/a = (16 \pm 2)$, and $\overline{R_c}/R_g = 1.57$. The actual radius of gyration measured by SANS (see Table 1) leads to $\overline{R_c} = 139 \pm 15$ Å.

Note that although the curve fitting is obtained with only two adjustable parameters ($\overline{R_c}$ and σ), it describes well our measurements in the whole q-range. In particular, the second-order oscillation due to the finite size of the spherical confinement volume appears on the fitted curve and on the data also. As for the first-order oscillation, we observe, experimentally, a smoothing that can be ascribed to the spectrometer resolution (Lairez, 1999). Actually, this oscillation lies in a q-range corresponding to the smallest q-values of the intermediate q configuration of our experiments (see Samples Preparation and the SANS Experiments, configuration *B*), i.e., data points with the lowest resolution.

Our result proves unambiguously that the chain in salt-containing solution collapses at large length scales but remains Gaussian at small length scales. In other words, the globule embodies a large quantity of solvent compared to the compact situation. The volume fraction of protein inside the globule can be estimated to be

$$\phi^* = N\left(\frac{a}{\overline{R_c}}\right)^3 = 0.3 \pm 0.1. \qquad (18)$$

From Eq. 17 and the reported value for a (Eq. 14) the blob size, ξ, can thus be estimated as being $\sim 33 \pm 12$ Å.

Let us come back to the problem of R_g measurement. As $P(q)$ of the refolded protein in salt-containing solution has been identified, one can check the validity of R_g determined in an extended Guinier q-range, i.e., up to $qR_g = 2$. Fig. 9 is a Guinier representation ($\ln(P(q))$ versus q^2) of the small q part of our measurements and simulation. Because of the globular shape of the protein, this representation is expected to be the most appropriate for the radius determination (see Eq. 7 and remark below). The data points obtained from simulation display a reasonable linear behavior up to $qR_g = 2$ and differ very little from the expected exponential decay $S(q) = e^{-(qR_q)^{2/3}}$ (*straight line* in Fig. 9). From this difference at $qR_g = 2$, one deduces that fitting our data in this q-range by a linear approximation leads us to underestimate the actual radius of gyration by 4%. This value corresponds to the error bar reported on Table 1. As for measurements (*triangles* in Fig. 9), apart from the first four data points, which clearly indicate the presence of a few aggregates in the solution, the same linear behavior is found. Note that the fit leads to the correct value of the scattered intensity extrapolated at $q = 0$ in view of the molecular mass of the protein. This makes us confident on the reported R_g value.

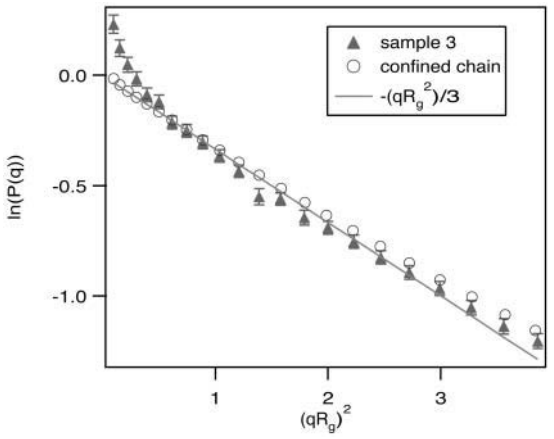

FIGURE 9 Logarithm of the form factor $P(q)$ versus $(qR_g)^2$ (Guinier representation) for refolded protein in salt-containing solution (*triangles*) and for the form factor of a confined chain obtained by simulation (*circles*). The line corresponds to $-(qR_g)^2/3$.

Protein refolded in salt-free solution

In salt-containing solution, fibronectin is not fully collapsed, but still contains a large amount of solvent. This suggests that the screening of attractive interaction is governed by the added salt contribution (Eq. 16). The ratio of Eq. 16 to 15 and Eq. 17 give $\phi^*_{\kappa_0<\kappa_p}/\phi^*_{\kappa_p<\kappa_0} = \kappa_0 b(b/\varphi l_B) > 5$ (in 150 mM NaCl solution, $\kappa_0^{-1} = 10.7$ Å, $l_B = 7.2$ Å, $b = 13.2$ Å and by definition $\varphi < 0.5$). Thus, one expects in salt-free solution a globule density 5× higher than in salt-containing solution (Eq. 18). A fully collapsed and locally dense conformation is thus expected in salt-free solution. At pH = 7.4 fibronectin is globally negatively charged and this charge asymmetry may split the globule into a necklace of smaller ones. The main experimental problem is that these expectations would concern sample 6 for which no form factor can be directly measured. Two questions arise:

1. Does salt addition (from sample 6 to sample 7) affect protein conformation?
2. What is the shape of the protein when collapsed in salt-free solution and the salt added only after the collapse (sample 7)?

Let us first deal with the second question. Although samples 3 and 7 are at the same pH, ionic strength and urea concentration, they show different conformations (see Fig. 4). Consequently, at least one of these two samples is not at the thermodynamical equilibrium. Comparing the locally low density of the globule obtained in salt-containing solution (sample 3) and the locally high density expected for sample 6, it seems reasonable to assume that sample 7 is out of equilibrium. A possible quenched dense structure, present in sample 6, may remain in sample 7. As the spectra cannot be fitted using a sphere form factor, we tried to use a necklace model made of a Gaussian string scattered with a number of small globules, i.e., one of the two theoretical expectations for sample 6. In this model, the center of mass of the amino acids belonging to the chain merges, in average, with the one of the set of the small globules. In addition, neglecting interactions between globules, the form factor of the necklace can be approximated by

$$P_{\text{necklace}}(q) = [xP_{\text{Debye}}(qR_1)^{1/2} + (1-x)P_{\text{sphere}}(qR_2)^{1/2}]^2, \tag{19}$$

with x the fraction of amino acids belonging to the chain, R_1 the radius of gyration of the chain and R_2 the radius of one globule. In Fig. 10, the best fit using this model is plotted for sample 7. We found: $x = 0.976 \pm 0.002$, $R_1 = (206 \pm 2)$ Å, and $R_2 = (21 \pm 1)$ Å. From the fraction x, the radius R_2 of one globule, and assuming a dense structure for these globules, one deduces the number, n, of globules along the chain: $n = 1250(1-x)/[R_2/(b/2)]^3 = 0.93 \pm 0.3$. This means the necklace contains only one globule (a result consistent with negligible interactions). Thus, from the approximation of Eq. 19, the protein conformation is the one of a small globule along a Gaussian chain. To strengthen this result, we tried to fit the data with the rigorous expression for the form factor of a Gaussian star polymer (made of Gaussian chains linked together by one of their extremities) with a dense core,

$$P_{\text{star with core}}(q) = [xP_{\text{star}}(q,R_1,f)^{1/2} + (1-x)P_{\text{sphere}}(qR_2)^{1/2}]^2, \tag{20}$$

where x is the fraction of amino acids belonging to the f Gaussian arms of the star, R_1 is the radius of gyration of one of these chains, and R_2 is the radius of the dense core. Following Higgins and Benoît (1994), P_{star} is

$$P_{\text{star}}(q,R_1,f) = 2\left[1-\frac{1}{f}\right]P_{\text{Debye}}(q,\sqrt{2}R_1) - \left[1-\frac{2}{f}\right]P_{\text{Debye}}(q,R_1). \tag{21}$$

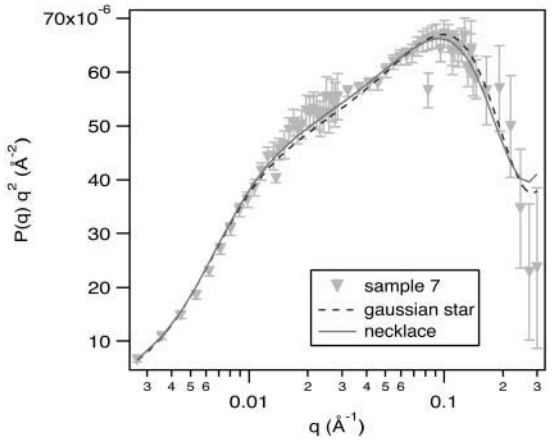

FIGURE 10 $P(q) \times q^2$ versus q for the refolded protein in salt-free solution and salt added after refolding (*sample 7*). The solid and dashed lines correspond to the necklace model (Eq. 19) and a Gaussian star with core (Eq. 20), respectively.

The best fit using this model is plotted on Fig. 10 with: $x = 0.976 \pm 0.002$, $f = 1.5 \pm 0.002$, $R_1 = (209 \pm 2)$ Å, and $R_2 = (21 \pm 1)$ Å. The f-value is consistent with the result obtained using the necklace model (one chain with a single globule). For sample 7, these two results are consistent with a still unfolded and Gaussian conformation at large length scale with a remaining locally dense area.

As the form factor of sample 7 is now identified, the same approach as for sample 3 (see Fig. 9) can be used for the R_g value reported in Table 1. Results are plotted in Fig. 11 in a Zimm representation that is the more appropriate in the case of a coil-like conformation. The data and the necklace model as well display a nice linear behavior up to $qR_g = 4$. This makes us confident on the reported radius of gyration despite its high value.

Let us come back to Fig. 7 and to the effect of salt addition from sample 6 to sample 7. In this figure, it appears that the local conformation of the collapsed chain (sample 6) is affected by salt addition, contrary to native and unfolded protein. At high q, the scattered intensity measured for sample 6 is lower than for sample 7, which indicates a higher average local density for sample 6 (same thing as comparing native and unfolded protein). A reasonable interpretation should be that the protein collapsed in salt-free solution (sample 6) adopts a necklace conformation with many small globules that are widely, but not all, dissolved by salt addition (sample 7). This interpretation is supported by two additional observations; first, at high q, the spectra measured for sample 6 greatly differs from the sphere form factor; and second, the interaction peak observed for sample 6 (see Figs. 5 and 7) is well-pronounced, indicating a dense structure closely related to the native rather than to coil-like conformations (samples 2, 5, or 7). However, the partial dissolution of this necklace structure by salt addition indicates that sample 6 is certainly not yet in a native conformation, even if it is closer to it.

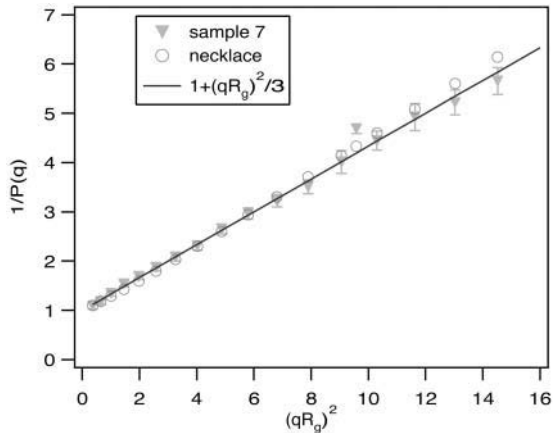

FIGURE 11 Reciprocal of the form factor $1/P(q)$ versus $(qR_g)^2$ (Zimm representation) for refolded protein in salt-free solution and salt added after refolding (*sample 7*, *triangles*) and for the necklace model (*circles*). The line corresponds to $1 + (qR_g)^2/3$.

CONCLUSION

In this article we have reported a SANS study of salt effect on the refolding of fibronectin, a high molecular weight protein. To our knowledge, this is the first study of such a problem. On the one hand, protein refolding studies have been reported on proteins 10× smaller; but on the other hand, salt effect has only been studied in terms of protein stability and denaturation, using local probes techniques (Pace et al., 2000; Hung and Chang, 2001).

Fibronectin looks to be a flexible string of 56 modules of 25-Å radius. In urea solution, the protein unfolds and adopts the conformation of a swollen coil. As urea is slowly removed from the solution, salt effects on the protein collapse have been studied. In salt-containing and dilute solutions, the form factor can be directly deduced from measurements and the conformation adopted by the molecule clearly identified. Results are summarized and schematized on Fig. 12.

Our main result concerns the two different conformations obtained when salt is added after or before the collapse stage. When the collapse occurs in an already salt-containing solution (sample 3), the protein adopts a globular shape at large length scale with a low density and a Gaussian local structure up to the length ξ, which has been found of the order of 30-Å radius. When the collapse occurs in salt-free solution and salt is added only after this stage (sample 7), the protein adopts a still open conformation at large length scale but displays a locally dense area of the order of 20-Å radius. Although the conformation of the collapsed protein in salt-free solution (sample 6) was not unambiguously determined, arguments have been given favoring a necklace conforma-

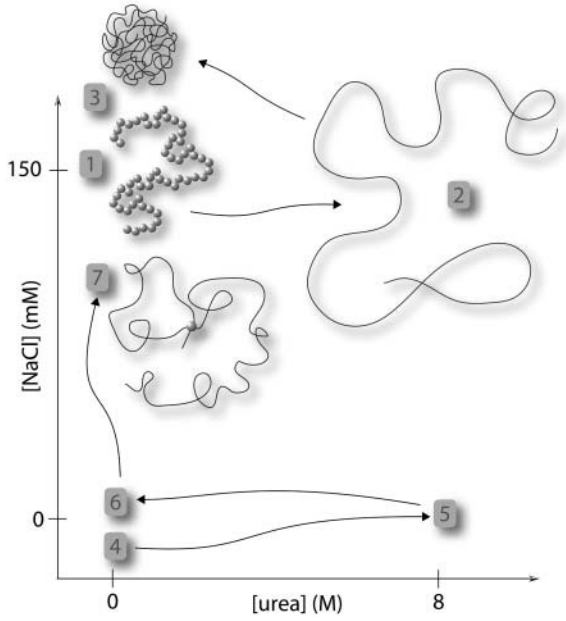

FIGURE 12 Conformation of fibronectin depending on the sample history.

tion (partly dissolved by salt addition), which may be reminiscent of the native structure. We have shown that the polyampholytes scaling theory qualitatively explains this behavior and should be considered with attention in the experimental investigation of refolding of high molecular weight proteins. Here, one has to emphasize that such a polyampholyte-type behavior cannot be observed on small proteins or fibronectin isolated modules. The blob size, ξ, in sample 3, and of the local dense area in sample 7, are of the order of the radius of gyration of protein usually studied with respect to the unfold-refold transition.

Although fibronectin in samples 1, 3, and 7 is at the same pH and ionic strength, its conformation depends on the sample history. Clearly two of these samples are not at the thermodynamical equilibrium. Slow kinetics effects on protein conformation can thus be suspected and need further experimental investigations. As a high local density may favor quenched and metastable states, it can be assumed that samples 1 and 7 are out of equilibrium. This result concerning the native conformation (sample 1) is quite puzzling and opens biological questions.

Our attempts to refold fibronectin have led to collapsed but still badly refolded conformations. In comparison, most in vitro studies of small proteins or modules of multidomain proteins undergo reversible unfolding process with urea. Experiments have been reported on individuals modules of fibronectin, which spontaneously refold after removal of the denaturing agent. In vivo, as far as the whole fibronectin is concerned, one has to account for posttranslational modifications (glycosylation, disulfide bonds) that may occur after folding (Ruoslahti, 1988), i.e., the protein folded in vivo may be different from the one studied in vitro. In addition, one has to keep in mind a possible sequential folding of the newly translated and nascent fibronectin. Moreover, it has been shown that efficient folding of high molecular weight nascent proteins depends on assistance of molecular chaperones (Ellis and Hartl, 1999; Frydman, 2001; Hartl and Hayer-Hartl, 2002). The study of in vitro refolding assisted by such molecular chaperones should be of great interest, but has not been considered up to now.

We thank P. Calmettes for a careful reading of the manuscript and constructive remarks. We thank le Laboratorie Francais du Fractionnement et des Biotechnologies for the gift of human cryoprecipitated plasma.

REFERENCES

Baldwin, R. L., and G. D. Rose. 1999. Is protein folding hierarchic? II. Folding intermediates and transition states. *Trends Biochem. Sci.* 24: 77–83.

Chodanowski, P., and S. Stoll. 1999. Monte Carlo simulations of hydrophobic polyelectrolytes: evidence of complete configurational transitions. *J. Chem. Phys.* 111:6069–6081.

Cotton, J.-P. 1991. Neutron, X-ray and Light Scattering. North Holland Publishing, Amsterdam. pp. 3–31.

de Gennes, P.-G. 1996. Scaling Concepts in Polymer Physics. Cornell University Press, Ithaca, NY.

Dickinson, C., B. Veerapadian, X.-P. Dai, R. Hamlin, N.-H. Xuong, E. Ruoslahti, and K. Ely. 1994. Crystal-structure of the 10th type-III cell-adhesion module of human fibronectin. *J. Mol. Biol.* 236:1079–1092.

Dill, A. K., and H. S. Chan. 1997. From Leventhal to pathways to funnels. *Nat. Struct. Biol.* 4:10–19.

Dinner, A. R., A. Sali, L. J. Smith, C. M. Dobson, and M. Karplus. 2000. Understanding protein folding via free-energy surfaces from theory and experiment. *Trends Biochem. Sci.* 25:331–339.

Dobrynin, A., and M. Rubinstein. 1995. Flory theory of a polyampholyte chain. *J. Phys. II France.* 5:677–695.

Dobrynin, A., M. Rubinstein, and S. Obukhov. 1996. Cascade transitions of polyelectrolytes in poor solvents. *Macromolecules.* 29:2974–2979.

Dobson, C. M. 1999. Protein misfolding, evolution and disease. *Trends Biochem. Sci.* 24:329–332.

Dobson, C. M., and M. Karplus. 1999. The fundamentals of protein folding: bringing together theory and experiment. *Curr. Opin. Struct. Biol.* 9: 92–101.

Ellis, R. J., and F. U. Hartl. 1999. Principles of protein folding in the cellular environment. *Curr. Opin. Struct. Biol.* 9:102–110.

Fersht, A., and V. Daggett. 2002. Protein folding and unfolding at atomic resolution. *Cell.* 108:573–582.

Frydman, J. 2001. Folding of newly translated proteins in vivo: the role of molecular chaperones. *Annu. Rev. Biochem.* 70:603–647.

Galzitskaya, O. V., D. N. Ivankov, and A. V. Finkelstein. 2001, Folding nuclei in proteins. *FEBS Lett.* 489:113–118.

Hartl, F. U., and M. Hayer-Hartl. 2002. Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science.* 295:1852–1858.

Higgins, J. C., and H. C. Benoît. 1994, Polymers and Neutron Scattering. Clarendon Press, New York.

Higgs, P. G., and J.-F. Joanny. 1991. Theory of polyampholyte solutions. *J. Chem. Phys.* 94:1543–1554.

Honig, B. 1999. Protein folding: from the Leventhal paradox to structure prediction. *J. Mol. Biol.* 293:283–293.

Hung, H., and G. Chang. 2001. Multiple unfolding intermediates of human placental alkaline phosphatase in equilibrium urea denaturation. *Biophys. J.* 81:3456–3471.

Jacrot, B. 1976. The study of biological structures by neutron scattering from solution. *Rep. Prog. Phys.* 39:911–953.

Kantor, Y., and M. Kardar. 1995. Instabilities of charged polyampholytes. *Phys. Rev. E.* 51:1299–1312.

Kataoka, M., and Y. Goto. 1996. X-ray solution scattering studies of protein folding. *Fold. Des.* 1:107–114.

Lairez, D. 1999. Résolution d'un spectromètre de diffusion de neutrons aux petits angles. *J. Phys. IV.* 9:67–81.

Pace, C., R. Alston, and K. Shaw. 2000. Charge-charge interactions influence the denatured state ensemble and contribute to protein stability. *Prot. Sci.* 9:1395–1398.

Pande, V. S., A. Y. Grosberg, and T. Tanaka. 2000. Heteropolymer freezing and design: towards physical models of protein folding. *Rev. Mod. Phys.* 72:260–314.

Pedersen, J. S., and P. Schurtenberger. 1996. Scattering functions of semiflexible polymers with and without excluded volume effects. *Macromolecules.* 29:7602–7612.

Pelta, J., H. Berry, G. C. Fadda, E. Pauthe, and D. Lairez. 2000. Statistical conformation of human plasma fibronectin. *Biochemistry.* 39:5146–5154.

Poulouin, L., O. Gallet, M. Rouahi, and J.-M. Imhoff. 1999. Plasma fibronectin: three steps to purification and stability. *Prot. Expres. Purif.* 17:146–152.

Privalov, P. L. 1979. Stability of proteins: small globular proteins. *Adv. Protein Chem.* 33:167–236.

Radford, S., and C. M. Dobson. 1999. From computer simulations to human disease: emerging themes in protein folding. *Cell.* 97:291–298.

Radford, S. E. 2000. Protein folding: progress made and promises ahead. *Trends Biochem. Sci.* 25:611–618.

Ruoslahti, E. 1988. Fibronectin and its receptors. *Annu. Rev. Biochem.* 57:375–413.

Russo, D. 2000. Etude structurale et dynamique de l'état natif et des états dénaturés de la néocarzinostatine, par microcalorimétrie différentielle, spectroscopies optiques et diffusion de neutrons et rayons X. Ph.D. thesis. Université de Paris XI. Paris, France.

Russo, D., D. Durand, P. Calmettes, and M. Desmadril. 2001. Characterization of the denatured states distribution of neocarzinostatin by small-angle neutron scattering and differential scanning calorimetry. *Biochemistry.* 40:3958–3966.

Russo, D., D. Durand, M. Desmadril, and P. Calmettes. 2000. Study of thermally and chemically unfolded conformations of a small β-protein by means of small-angle neutron scattering. *Phys. B.* 276:520–521.

Shew, C.-Y., and A. Yethiraj. 1999. Monte Carlo simulations and self-consistent integral equation theory for polyelectrolyte solutions. *J. Chem. Phys.* 110:5437–5443.

Shew, C.-Y., and A. Yethiraj. 2000. Self-consistent integral equation theory for semiflexible chain polyelectrolyte solutions. *J. Chem. Phys.* 113: 8841–8847.

Stevens, M. J., and K. Kremer. 1995. The nature of flexible linear polyelectrolytes in salt free solution: a molecular dynamics study. *J. Chem. Phys.* 103:1669–1690.

Sticht, H., A. Pickford, J. Potts, and I. Campbell. 1998. Solution structure of the glycosylated second type-II module of fibronectin. *J. Mol. Biol.* 276:177–187.

Williams, M., I. Phan, T. Harvey, A. Rostagno, L. Gold, and I. Campbell. 1994. Solution structure of a pair of fibronectin type-I modules with fibrin binding-activity. *J. Mol. Biol.* 235:1302–1311.

Yamada, K. 1989. Fibronectin. Academic Press, New York. pp. 48–121.