Effects of Macromolecular Crowding on an Intrinsically Disordered Protein Characterized by Small-Angle Neutron Scattering with Contrast Matching

Daniel Johansen,[†] Cy M. J. Jeffries,[‡] Boualem Hammouda,[§] Jill Trewhella,[‡¶] and David P. Goldenberg[†*]
[†]Department of Biology and [¶]Department of Chemistry, University of Utah, Salt Lake City, Utah; [‡]School of Molecular Biosciences, The University of Sydney, New South Wales, Australia; and [§]Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, Maryland

ABSTRACT   Small-angle neutron scattering was used to examine the effects of molecular crowding on an intrinsically disordered protein, the N protein of bacteriophage λ, in the presence of high concentrations of a small globular protein, bovine pancreatic trypsin inhibitor (BPTI). The N protein was labeled with deuterium, and the D_2O concentration of the solvent was adjusted to eliminate the scattering contrast between the solvent and unlabeled BPTI, leaving only the scattering signal from the unfolded protein. The scattering profile observed in the absence of BPTI closely matched that predicted for an ensemble of random conformations. With BPTI added to a concentration of 65 mg/mL, there was a clear change in the scattering profile representing an increase in the mass fractal dimension of the unfolded protein, from 1.7 to 1.9, as expected if crowding favors more compact conformations. The crowding protein also inhibited aggregation of the unfolded protein. At 130 mg/mL BPTI, however, the fractal dimension was not significantly different from that measured at the lower concentration, contrary to the predictions of models that treat the unfolded conformations as convex particles. These results are reminiscent of the behavior of polymers in concentrated melts, suggesting that these synthetic mixtures may provide useful insights into the properties of unfolded proteins under crowding conditions.

INTRODUCTION

It is widely recognized that the in vivo environments in which most proteins function contain much higher concentrations of macromolecules than are typically used for in vitro studies, and it has been suggested that macromolecular crowding may substantially alter the stabilities, conformations, and functional properties of protein molecules (1,2). Unfolded proteins may be particularly sensitive to crowding effects, as they represent broad ensembles of conformations that occupy a wide range of volumes. Because the volume available to molecules is decreased by increasing concentrations, the thermodynamic stabilities of more compact conformations are expected to increase relative to those of molecules that take up larger volumes (2). Native proteins are thus predicted to be stabilized relative to nearly all unfolded conformations, and the distribution of conformations making up an unfolded state is expected to shift, favoring the more compact conformations. Although the general features of these phenomena have been appreciated for some time, different theoretical and experimental attempts to measure the magnitudes of crowding effects on stability, both in vitro and in vivo, have yielded a wide range of estimates (2–11).

A major challenge in experimental studies of crowding is to devise methods for studying the properties of a protein of interest in the presence of very high concentrations of other macromolecules, especially proteins. Fluorescence techniques are well suited for this purpose (5,10,12), but generally provide information only about the sites of the fluorescent probes. Small angle x-ray scattering (SAXS) offers the advantage of providing information about the distribution of interatomic distances throughout the molecules of interest and has become one of the major techniques used to characterize unfolded proteins and other dynamic macromolecular structures (13–17). SAXS, however, is generally not suitable for studying the properties of an unfolded protein in the presence of high concentrations of other molecules, because the crowding molecules will also contribute to the scattering profile.

Neutrons can also be used for scattering studies and offer particular advantages for more complex mixtures (17,18). Unlike x-rays, which are scattered from the electrons of a molecule, neutrons are scattered by the atomic nuclei, and different isotopes of the same element can have clearly distinguishable scattering properties. In particular, scattering from the major isotope of hydrogen, ^1H, causes a change of $180°$ in the phase of the associated wave, whereas the scattering from deuterium, ^2H, does not. By careful adjustment of the ^2H content of different components in a mixture, the scattering contrast due to some components can be eliminated, making these species effectively invisible, whereas the scattering from other components remains detectable. Neutron scattering in conjunction with deuterium labeling has long been used to study the structures of macromolecular complexes (19–21), as well as the behavior of synthetic polymers (22,23). A recent publication (24) describes the use of this approach to examine the effects of crowding on a random

coil polymer, polyethylene glycol, but so far as we are aware, neutron scattering has not previously been used to study the effects of crowding on proteins.

In this study, neutron scattering was used to study the properties of an intrinsically disordered protein, the N protein of bacteriophage λ, in the presence of high concentrations of a small globular protein, bovine pancreatic trypsin inhibitor (BPTI). The N protein is a transcriptional anti-termination factor that plays a key role in the regulation of lysogeny when the bacteriophage infects its host, *Escherichia coli* (25). In isolation, the protein is highly disordered, as measured by NMR and circular dichroism spectroscopy (26), but the N-terminal segment takes on an α-helical conformation when it binds to a specific site on the RNA transcript (27). Other regions of the protein interact with RNA polymerase and other components of the transcription complex (28), and its intrinsically disordered nature may be an important factor in its ability to interact with multiple components of the highly dynamic transcription complex. BPTI was chosen as the crowding agent because of its small size (58 amino-acid residues), which is predicted to cause more pronounced effects than a larger molecule (9), and because both it and the N protein carry positive net charges, minimizing the likelihood of specific interactions between the two proteins.

The experimental results demonstrated a clear crowding effect at a relatively low crowding density, but this effect did not appear to increase at a higher density. This latter observation is contrary to the predictions of theoretical treatments based on the results of scaled particle theory, including a computational model introduced here, and suggests that the effects of crowding on unfolded proteins may be more complex than generally appreciated.

MATERIALS AND METHODS

Protein samples

Bacteriophage λ N protein was purified from *Escherichia coli* BL21(DE3) bacteria containing the expression plasmid pET-N1, described by Rees et al. (29). Cultures for preparing perdeuterated samples were grown in the PG medium of Studier (30) containing 90% D_2O (Cambridge Isotopes, Andover, MA) and 100 μg/mL ampicillin. (All D_2O concentrations are expressed as volume/volume percentages.) After lysis by sonication, the N protein was found in the insoluble fraction of the lysate and was resuspended and isolated as described by Van Gilst et al. (26) The protein was further purified by ion exchange chromatography using Whatman (Maidstone, Kent, UK) Express-Ion S cation exchange resin at pH 8. From sodium dodecyl sulfate gel electrophoresis, nondenaturing gel electrophoresis, and reversed-phase high-performance liquid chromatography, the protein was estimated to be at least 95% pure. The identity of the protein was confirmed by electrospray-ionization mass spectrometry.

For storage and transport, the purified N protein was dialyzed against 0.1 mol/L acetic acid and lyophilized. Before small-angle neutron scattering (SANS) measurements, the lyophilized protein was dissolved in 6 mol/L GuHCl, 50 mmol/L pH 7 Na-phosphate buffer. The protein was then dialyzed against 50 mmol/L pH 7 Na-phosphate buffer containing 0.5 mmol/L benzamidine.

Bovine pancreatic trypsin inhibitor (Aprotinin) was purchased from Roche Applied Science (Mannheim, Germany) and used without further purification.

Samples for SANS measurements were prepared by first dialyzing the N protein and BPTI, separately, against 50 mmol/L pH 7 Na-phosphate buffer in 46% D_2O. The concentrations of the proteins were then determined by ultraviolet absorbance, mixed in the appropriate ratios, and again dialyzed against the same buffer solution. The samples were then concentrated using Amicon Ultra-15 centrifugal filter units with Ultracel-3 cellulose membranes (3000 Da cutoff) (Millipore, Billerica, MA), and the total protein concentration was determined by ultraviolet absorbance. The final concentration of N protein in the samples was 8 mg/mL. BPTI samples for SAXS measurements were similarly prepared by dialysis and centrifugal concentration.

Small-angle x-ray scattering

SAXS data were recorded using an Anton Paar (Graz, Austria) SAXSess instrument with line-collimation and a two-dimensional charge-coupled device detector, as described by Jeffries et al. (31). Image data from 180 exposures of 10 s each were averaged and integrated using the Anton Paar SAXSquant software. For Guinier analysis, the slit-smeared data were corrected using the iterative method of Lake (32).

Simulated SAXS profiles for BPTI were calculated from the atomic coordinates of monomeric and decameric crystal structures (PDB entries 4PTI and 1BTZ, respectively) using the program CRYSOL (33). Models for oligomeric structures were as described by Hamiaux et al. (34). For direct comparison to the experimental SAXS data, the calculated profiles were smeared by convolution with the beam-length profile and the detector slit width.

Small-angle neutron scattering

SANS data were recorded using the NG3 30m instrument at the National Institute for Standards and Technology Center for Neutron Research (Gaithersburg, MD). Samples were placed in Hellma quartz cylindrical cells (Plainview, NY) with a 1-mm path-length and maintained at 20°C. The neutron wavelength was 6.0 $\pm$ 0.125 Å, and data were recorded using two sample-to-detector distances (1.33 and 4 m) to provide a combined Q range of 0.01–0.45 Å^{-1},

$$Q = (4\pi\sin\theta)/\lambda,$$

where θ is one-half of the scattering angle and λ is the neutron wavelength. The D_2O concentration for optimum contrast matching was determined by recording scattering profiles for BPTI ($\approx$ 40 mg/mL) dissolved in buffer containing 0, 20, 70, 85, and 100% D_2O. The scattering intensity (with a negative sign assumed for $[D_2O] > 60\%$) was plotted as a function of D_2O concentration and fit to a line, from which the interpolated concentration for zero scattering (46% D_2O) was estimated. For each of the samples containing N protein, data were recorded for 4–5 h at the shorter sample-to-detector distance and 1–2 h at the longer distance. Reference samples contained the same buffer and the same concentrations of D_2O and BPTI. The data were reduced using the IGOR Pro program and the SANS macros described by Kline (35).

As discussed in the Results, Guinier plots of the SANS data for some of the N-protein samples were biphasic, suggesting the presence of protein aggregates. To analyze these results, the scattering data were fit to a double-exponential function,

$$I(Q) = K_1 e^{-Q^2 R_{g,1}^2/3} + K_2 e^{-Q^2 R_{g,2}^2/3}, \tag{1}$$

where I is the scattering intensity. $R_{g,1}$ and $R_{g,2}$ are the radii of gyration of the sample components attributed to the two phases of the Guinier plot. The intensities of the two phases extrapolated to $Q = 0$, K_1 and K_2, represent the

squares of the scattering density difference between the particles and the solvent.

To estimate the relative concentrations of the particles giving rise to the two phases in the Guinier plots, K_1 and K_2 were each assumed to be proportional to the mass concentration (C) and the molecular mass (M) of a single population of particles. The particles were further assumed to be made up of well-solvated polymers, such that the chain length (and molecular mass) is proportional to $R_g^{1.7}$. Thus,

$$\frac{K_1}{K_2} = \frac{C_1 M_1}{C_2 M_2} = \frac{C_1}{C_2}\left(\frac{R_{g,1}}{R_{g,2}}\right)^{1.7} \qquad (2)$$

and

$$\frac{C_1}{C_2} = \frac{K_1}{K_2}\left(\frac{R_{g,2}}{R_{g,1}}\right)^{1.7}. \qquad (3)$$

Computational modeling

The procedure used to generate random ensembles of calculated structures using the program DYANA (36) (recently superseded by CYANA) has been described previously (16,37). Although the original procedure resulted in quite good correlations with experimental data for denaturant-unfolded proteins, it generates distributions of backbone dihedrals that deviate significantly from the standard Ramachandran allowed regions. To address this problem, the sampling procedure was modified so that the initial conformations were generated by setting the backbone dihedral angles to values selected at random from the allowed regions (rather than all possible values, as before), whereas side-chain dihedrals were set to random values between 0 and 360°. All of the dihedral angles were then adjusted to minimize the CYANA target function, which considers only steric repulsion terms. This procedure resulted in distributions very similar to the classical Ramachandran maps (see Fig. S1 in the Supporting Material). The major effect of the changed sampling procedure was to increase slightly the average dimensions of the chain. For the N protein, the new procedure resulted in a root-mean square (RMS) radius of gyration of 36.9 Å, versus 33.8 Å for the original procedure. A more detailed description of the new sampling algorithm will be presented elsewhere.

The N-protein sequence was represented with all non-hydrogen atoms, and a total of 259,633 random conformations were generated. Structures for which the CYANA target function was greater than 0.01 Å^2 were discarded, leaving a total of 228,259. For each of these structures, the excluded covolume of the N protein and a sphere with radius 15 Å (see below) was calculated using a grid-search algorithm. The volume occupied by protein hydrogen atoms was accounted for implicitly by using the group atomic radii of Chothia (38).

To determine an appropriate radius for the sphere used to represent the crowding molecule (r_{bpti}), the covolumes of native BPTI and spheres of radius 2.5–30 Å were calculated using the same method as applied to the individual N-protein conformations. The calculated covolumes were then plotted as a function of sphere radius (r_s) and fit to the expression

$$V = \frac{4\pi}{3}\left(r_{bpti} + r_s\right)^3. \qquad (4)$$

An excellent fit ($R^2 = 0.998$) was obtained with a value of $r_{bpti} = 15$ Å.

The calculated excluded covolumes of the N-protein conformations were used to calculate the change in chemical potential, $\Delta\mu$, for transferring a molecule from a dilute to crowded solution, using the results of scaled particle theory for spheres (39), following Eq. 33 of Minton (6). A Boltzmann weighting factor for each conformation i at a given crowding density was calculated according to

$$w_i = \frac{e^{-\Delta\mu_i(\phi)/RT}}{\sum_{i=1}^{N} e^{-\Delta\mu_i(\phi)/RT}}, \qquad (5)$$

where N is the total number of conformations in the ensemble and ϕ is the fraction of total volume occupied by the crowding molecules.

The predicted neutron scattering profile for each structure was calculated using the program CRYSON (40), assuming a solvent deuteration fraction of 46% and 85% fractional perdeuteration of the N protein. The default values were used for all other CRYSON parameters. The Boltzmann factors described above were then used to calculate weighted average SANS profiles for different crowding densities.

Data fitting

The experimental SAXS and SANS data were fit to calculated curves by using the method of least squares to adjust a single scaling parameter introduced to account for differences in protein concentration. The experimental data were weighted according to their uncertainties (standard errors), and the individual fits to a given experimental data set were compared using the reduced χ^2 statistic, as described previously (16).

RESULTS

Characterization of BPTI at high concentrations

Because the effects of crowding are predicted to be sensitive to the dimensions of the crowding molecule (9), it was important to establish the oligomeric state of BPTI under the conditions of the experiments, especially because previous studies have shown that BPTI can form decamers under some solution conditions (34,41). Small-angle x-ray scattering was used to characterize BPTI at concentrations ranging from 32 mg/mL to 150 mg/mL at pH 7.0 and 46% D$_2$O, as shown in Fig. 1. In panel A, the data from very small scattering angles ($Q < 0.045$ Å^{-1}) are presented in the form of a Guinier plot, in which $ln(I)$ is plotted as a function of Q^2. In the limit of small scattering angles ($Q \lesssim 1/R_g$, where R_g is the radius of gyration), the plot is predicted to be linear for homogeneous particles, reflecting the Guinier approximation (42)

$$I(Q) = I(0)e^{-Q^2 R_g^2/3}, \qquad (6)$$

where $I(0)$ is the extrapolated scattering intensity at zero angle and is proportional to both the mass concentration and the molecular mass of the particles. The radii of gyration, derived from the slope of the Guinier plot, are plotted as a function of protein concentration in panel B, showing that the apparent R_g increased with protein concentration, from 12.3 Å to 15.7 Å, as compared to 12 Å calculated from the crystal structure of native BPTI. Although this trend could be interpreted as evidence of an oligomeric form in equilibrium with the monomer, the linear dependence of $I(0)$ on concentration (Fig. 1 B) indicates that only a single species contributes significantly to the scattering profiles: An equilibrium among oligomeric forms is predicted to cause a higher-order increase in $I(0)$.

The BPTI SAXS profiles were analyzed further by comparison with profiles calculated from atomic coordinates of different oligomeric forms, as described in Materials and Methods. As shown in Fig. 1 C, only the
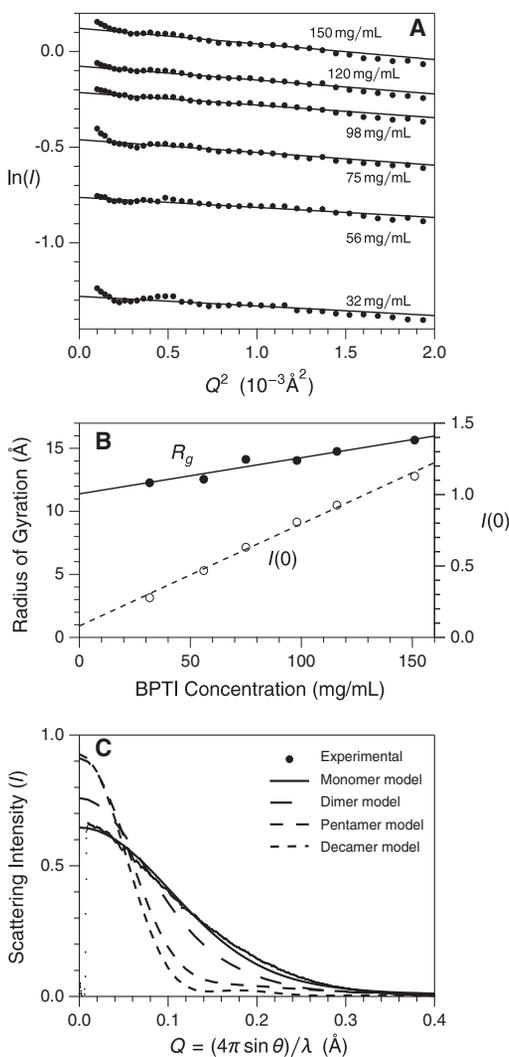
FIGURE 1 Small-angle x-ray scattering from BPTI in 46% D$_2$O at pH 7. Error bars representing the standard deviations of the scattering intensities are of approximately the same size as the solid circles. (A) Guinier plots. The apparent radii of gyration and extrapolated scattering intensity at zero angle, $I(0)$, were determined from a least-squares fit to the data, represented by the straight lines. (B) Apparent radius of gyration and $I(0)$ as a function of BPTI concentration. The lines represent least-squares fits to the data. (C) Fits of the scattering profile from 150 mg/mL BPTI to calculated profiles for monomeric, dimeric, pentameric, and decameric models.

monomer model yielded a reasonable fit to the experimental data, even at the highest BPTI concentration. The SAXS data thus indicate that the monomeric form predominates under the conditions of these experiments, consistent with previous studies demonstrating that the monomer and decamer are the only significant forms, with the decamer detectable only in the presence of high salt concentrations or pH values <5 or >9 (34,41).

Small angle neutron scattering

To detect neutron scattering from the N protein in the presence of high crowder concentrations, the protein was
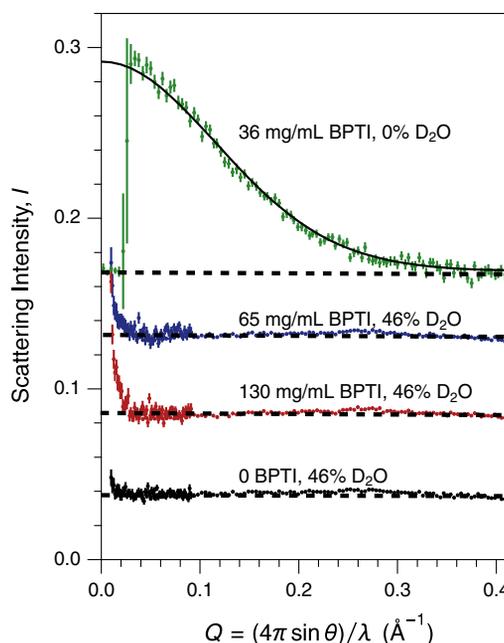

FIGURE 2 Neutron scattering from BPTI with and without solvent contrast matching. The error bars in this and the other SANS profiles represent standard deviations calculated from counting statistics. The scattering curves are displaced vertically for clarity, and the dashed lines are drawn to aid in detecting deviations from the baseline. The smooth curve superimposed on the data for BPTI in 0% D$_2$O was calculated from the atomic coordinates of BPTI (PDB entry 4PTI) using the program CRYSON.

biosynthetically perdeuterated at a level of ~85%, whereas the crowding protein, BPTI, was unlabeled. Fig. 2 shows the neutron scattering from the unlabeled BPTI dissolved in either H$_2$O or 46% D$_2$O. The scattering curve from 36 mg/mL BPTI in H$_2$O was readily detectable and closely matched the curve predicted for monomeric BPTI (*black curve*). The scattering from BPTI was almost completely masked, however, when the protein was dissolved in 46% D$_2$O .

The scattering profiles from N protein in the presence of 0, 65, and 130 mg/mL BPTI are shown in Fig. 3 *A*. For comparison, a simulated scattering curve for a globular protein of similar chain length, ribonuclease A, is shown as a dashed curve. All three of the N-protein scattering curves have the general shape expected for a highly disordered protein, as discussed further below, and are clearly distinguishable from that of a folded protein.

In Fig. 3 *B*, the data from very small scattering angles are presented in a Guinier plot. As shown in the figure, the plots for the N protein are distinctly biphasic, suggesting the presence of protein aggregates. SAXS experiments indicate that deuteration does not affect the conformation of the N protein or its tendency to aggregate, but aggregation is enhanced in the presence of D$_2$O (D. Johansen, J. Trewhella, and D. P. Goldenberg, unpublished). To account for the presence of aggregates, the data were fit to a double-exponential function of Q^2 (Eq. 1); the parameters are listed in Table 1. For the samples containing 0 or 65 mg/mL BPTI, the fits
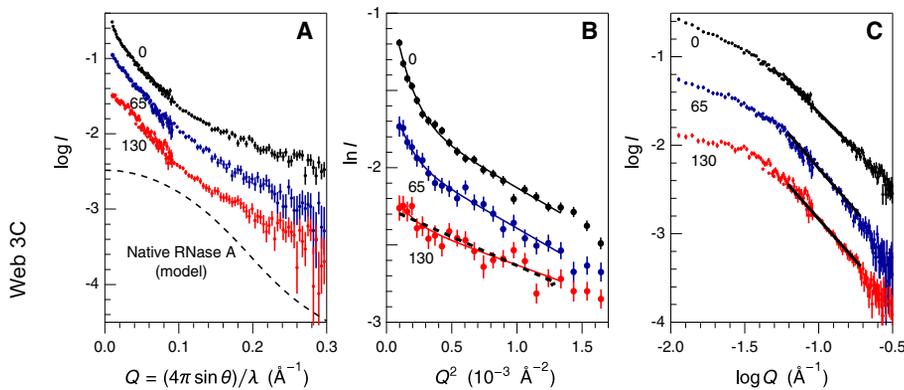
FIGURE 3 Neutron scattering from unfolded N protein (8 mg/mL) in the presence of 0 (*black*), 65 (*blue*), or 130 mg/mL (*red*) of crowding protein (BPTI), and 46% D_2O. In panel *A*, the experimental scattering profiles are compared with the calculated profile for a globular protein of similar size (ribonuclease A, 124 amino-acid residues, PDB entry 7RSA). (*B*) Guinier plots of the very-small angle region ($Q < 0.04$ Å^{-1}. The curves represent least-squares fit of the data to a double-exponential function (Eq. 1), with the parameters listed in Table 1. The data from the sample containing 130 mg/mL BPTI were also fit to a single-exponential function (*black dashes*). (*C*) Log-log plot showing fits to the power-law regions of the scattering curves. The slopes of the fit lines are -1.70 ± 0.02 (0 BPTI), -1.92 ± 0.03 (65 mg/mL BPTI), and -1.89 ± 0.03 (130 mg/mL BPTI). In each of the panels, the scattering curves are displaced arbitrarily along the vertical axis for clarity.

were consistent with each sample containing a mixture of two classes of particles, one with $R_g \approx 160$ Å. The radius of gyration for the second component was estimated to be 38 Å, consistent with a monomeric unfolded protein composed of 107 residues (14). The data for the sample containing 130 mg/mL BPTI were consistent with a mixture containing components with similar sizes, but the data were fit equally well by a single exponential function.

The relative concentrations of the aggregates in the samples were estimated from the relative values of the amplitudes of the two exponential terms (K_1 and K_2 in Eq. 1) for the two components, with the assumption that both forms behaved as well-solvated random polymers, as described in Materials and Methods. On this basis, the aggregates were estimated to represent 12%, 5%, and 2% of the total N protein contributing to scattering in the presence of 0, 65, and 130 mg/mL BPTI, respectively. If, as is likely, the aggregates were more densely packed than the unfolded monomers, then the relative concentrations of the aggregated forms would be significantly less than estimated here. Because the relative scattering intensity from larger particles decreases very rapidly with scattering angle, the aggregates are expected to make minimal contributions to the observed scattering intensities outside of the Guinier region, and the remainder of our analysis is based on the data from larger *Q*-values.

The scattering from intermediate angles (0.05 Å$^{-1} \lesssim Q \lesssim 0.15$ Å^{-1}) contains information about the distribution of interatomic distances within a molecule. For many different particle types, the intensity within this range decreases with *Q* according to a power-law dependence (43) that is revealed as a linear region of a plot of log (*I*) versus log (*Q*), as shown in Fig. 3 *C*. The slope of the linear region in the log-log plot reflects the distribution of interatomic distances on different length scales, and can be interpreted in terms of a mass fractal dimension, D_m, which is equal to the negative of the slope (16,44–48). This parameter represents the scaling relationship between the mass enclosed by a volume and the linear dimensions of the volume. For an idealized random-flight polymer (which is allowed to cross itself), $D_m = 2$. For real polymers, however, the chains are more expanded because of excluded volume effects, and D_m for a well-solvated chain is predicted to approach 1.7. The fractal dimension is also the reciprocal of the Flory exponent, ν, which relates the linear dimensions of a polymer to the chain length.

From the log-log plots shown in Fig. 3 *D*, the fractal dimension of the N protein was estimated to be 1.70 ± 0.02, 1.92 ± 0.03, and 1.89 ± 0.03 in the presence of 0, 65, and 130 mg/mL BPTI. In the absence of the crowding molecule, the N protein thus appeared to display exactly the scaling behavior expected of a solvated excluded-volume

TABLE 1 Guinier plot fitting parameters

(BPTI) (mg/mL)	K_1	$R_{g,1}$ (Å)	K_2	$R_{g,2}$ (Å)	χ^2_ν	C_{agg}/C_{mon}*
0[†]	0.25 ± 0.03	156 ± 11	0.19 ± 0.01	38 ± 2	1.0	0.12
65[†]	0.10 ± 0.05	165 ± 42	0.15 ± 0.01	38 ± 3	0.46	0.05
130[†]	0.04 ± 0.05	163 ± 103	0.12 ± 0.01	30 ± 4	0.97	0.02
130[‡]	—	—	0.13 ± 0.003	33 ± 2	0.97	—

*Relative mass concentrations of aggregates and monomers were estimated from Eq. 3 in the text.
[†]Parameters from least-squares fits of neutron scattering data to the double-exponential function given in Eq. 1 of the text, as illustrated by the curves in Fig. 3 *B*.
[‡]Parameters from least-squares fit of the data for the 130 mg/mL BPTI sample to a single-exponential function, as illustrated by the dashed line in Fig. 3 *B*.

chain. In the presence of the crowder, the fractal dimension increased, as expected if compact conformations are favored. It is striking, however, that there was no significant increase in D_m as the concentration of BPTI was increased from 65 to 130 mg/mL.

Comparison with a computational model

To gain further insights into the crowding effects detected by SANS, the scattering profiles were compared to those predicted by a computational model. Using a previously described approach, ~230,000 random conformations of the N protein were generated, with bond lengths and angles fixed to standard values and the dihedral angles set to random values constrained only by steric repulsion, both short- and long-range (16,37). As shown previously, this procedure generates ensembles of conformations that closely match experimental measurements of the scaling relationships and overall dimensions of unfolded proteins in denaturants. For each of these conformations, the excluded covolume of the N protein and a sphere representing the crowding protein was calculated. The energetic effects of crowding were then estimated using the results of scaled particle theory to calculate the change in free energy for transferring each N-protein conformation from a dilute solution to a solution containing crowding spheres, as described in Materials and Methods. The crowding free energies were used to calculate Boltzmann factors and weighted distributions of the N-protein conformations. This procedure is conceptually similar to the postprocessing approach introduced by Dong et al. (9).

In Fig. 4 A, the weighted distributions are shown as histograms of the radii of gyration. The histogram for the original unweighted distribution ($\phi = 0$) is characterized by conformations with radii of gyration ranging from 17 to 68 Å and an RMS average of 37 Å. The presence of the crowding agent shifts the distribution toward more compact conformations, with the RMS radius of gyration decreasing to 24 Å at $\phi = 0.2$, as shown in the inset.

For each of the calculated conformations, the SANS profile was predicted, and average profiles for the ensemble were calculated using the Boltzmann weighting factors for $\phi = 0$–0.2 (Fig. 4 B). As the crowding density increases, the calculated scattering curves at low and intermediate Q-values are shifted to the right, reflecting the shift toward shorter interatomic distances. Within the intermediate Q-range, the curves also become steeper, reflecting the increased fractal dimension of the more compact conformations. The fractal dimension for each weighted ensemble was calculated from the slope of the linear region of a log-log plot of the averaged scattering curve. For the unweighted ensemble ($\phi = 0$), the fractal dimension was 1.63, in good agreement with polymer theory, and increased to 2.67 for $\phi = 0.2$, as shown in the inset to Fig. 4 B.

To compare the computational and experimental results, each of the calculated average profiles was fit to the experimental SANS data, as shown in Fig. 5. The scattering profile from the sample containing no crowding protein was best fit by the calculated profile for the unweighted distribution, corresponding to $\phi = 0$. The only significant deviations between the observed and calculated profiles were in the very small-angle region ($Q < 0.025$), due to the aggregation discussed above.

In contrast to the uncrowded sample, the scattering profile from the sample containing 65 mg/mL was poorly matched by the $\phi = 0$ calculated curve but was well fit with $\phi = 0.1$ (Fig. 5 B). The fractal dimension calculated from the simulated curve was 1.96, in good agreement with the value from the experimental curve, 1.92 ± 0.03. If the crowding molecule is approximated as a sphere with radius 15 Å, the concentration of 65 mg/mL corresponds to a crowding density of $\phi = 0.085$, consistent with the fitting result.
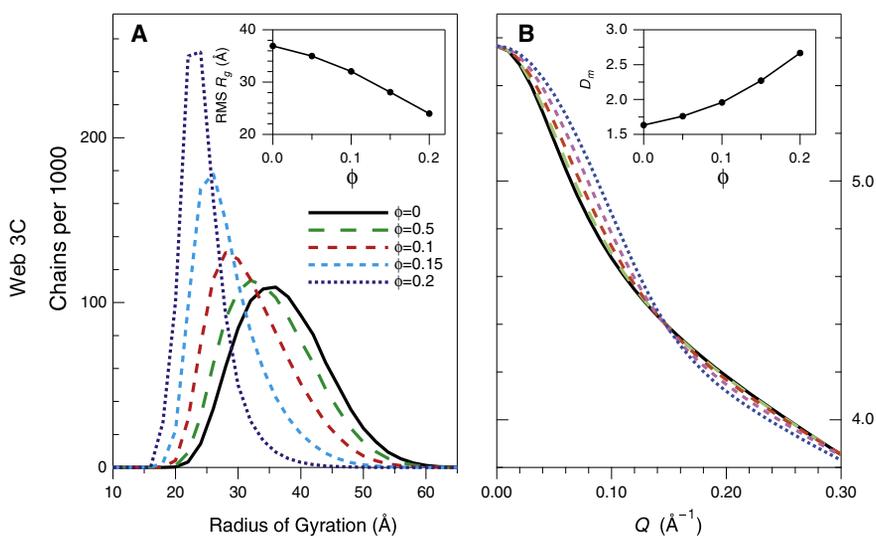

FIGURE 4 Computational simulation of molecular crowding effects. (A) Distribution of calculated conformations, as defined by the radius of gyration. The solid curve represents the unweighted distribution of ~230,000 conformations, expressed as the number of chains per 1000 with radii of gyration lying within an interval of 2 Å. The weighted distributions represented by the dashed curves for different crowding densities were calculated from the excluded covolumes of the individual conformations and a spherical crowding molecule with a radius of 15 Å, as described in the text. The inset shows the RMS radius gyration of the weighted ensembles as a function of crowding density. (B) Simulated SANS profiles calculated from the individual conformations and Boltzmann weighting factors for the indicated crowding densities. The inset shows the fractal dimension, D_m, calculated from the simulated scattering curves, as a function of ϕ.
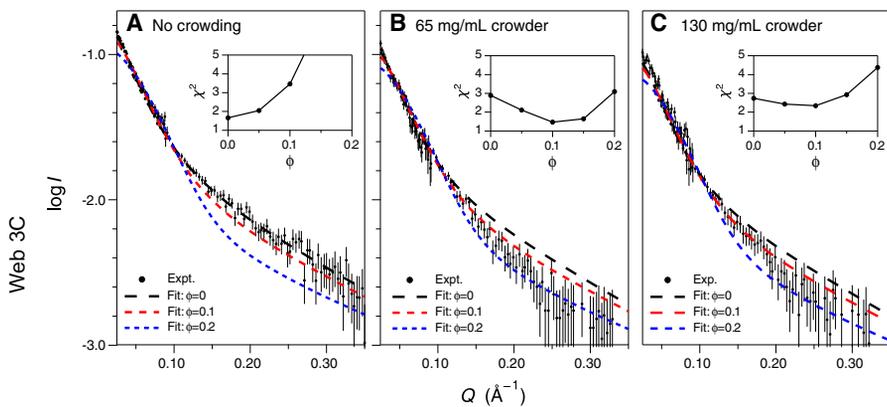
FIGURE 5 Fits of experimental SANS data to predicted scattering curves. For each experimental scattering curve, each of the computed curves shown in Fig. 4 *B* was fit to the data by introducing a single adjustable scaling factor representing the total scattering intensity. The best fit was identified by the minimum value of the reduced χ^2 statistic, as shown in the insets. In each panel, the experimental data with error bars are plotted along with the best fits for the simulated ensembles with $\phi = 0$ (*black*), 0.1 (*red*), and 0.2 (*blue*).

Contrary to expectations, but consistent with the fractal-dimension analysis, the $\phi = 0.1$ simulated scattering curve also yielded the best fit to the data for 130 mg/mL crowding protein. From the model, doubling the crowding density is expected to cause a pronounced shift in the distribution of conformations and an increase in the fractal dimension to ~2.4. The absence of such a shift in the experimental data suggests that there are features of the crowding effect that are not fully accounted for in the model, as discussed further below.

DISCUSSION

Theoretical treatments of molecular crowding predict that the average overall dimensions of molecules should decrease as the crowding density increases. This effect should, in principle, be measurable from the Guinier region of the scattering curves (Fig. 3 *C*). In practice, however, the Guinier plot is very sensitive to the presence of any aggregates, as well as effects from interparticle interference (49). In the data shown here, deviations from linearity in the Guinier plots can be accounted for by aggregates representing 12% or less of the total sample mass, but these deviations introduce considerable uncertainty in estimating the radius of gyration. None the less, the estimated radii of gyration (30–40 Å) are consistent with the results for other unfolded proteins of similar chain length, as well as our computational simulations.

The analysis of the Guinier plots suggests that the concentration of aggregates decreased significantly at higher crowder concentrations (Table 1). This result is somewhat surprising, because association reactions are generally expected to be favored by macromolecular crowding (2). If, however, the associated form is highly asymmetric, the monomer may be favored by crowding (2). In addition, the crowding molecules may screen interactions among the N-protein molecules or shift the distribution of unfolded molecules to favor conformations that are less prone to aggregate. It is also possible that the observed effect is kinetic, so that the rate of aggregation is reduced

by the presence of crowding molecules, even though the aggregates may be favored thermodynamically.

At greater scattering angles, the contribution from larger particles falls off very rapidly. For this reason, our analysis focused on the portion of the curve outside of the Guinier region, using log-log plots to estimate the fractal dimension of the molecules and using direct comparisons with computational results. For the case of relatively low crowding density (65 mg/mL BPTI, $\phi \approx 0.085$), there was a clear increase in the fractal dimension and a very good fit of the experimental data to the computational model for $\phi = 0.1$.

At the higher crowding density examined (130 mg/mL BPTI, $\phi \approx 0.17$), there was no further increase in the fractal dimension of the unfolded protein, and the best fit to the simulated ensembles was obtained with $\phi = 0.1$. The discrepancy between the experimental and computational results for higher crowding densities suggests that alternative theoretical approaches are necessary. One of the important simplifications introduced in the model used here is the treatment of the unfolded conformations as spheres when calculating the crowding free energy. Although the atomic coordinates were used to compute the excluded covolume of each conformation, the structures were then approximated as spheres with equivalent volumes when applying the scaled particle theory. In reality, the individual conformations in an ensemble of random-coil chains are expected to display considerable asymmetry, and it is possible that the distribution of asymmetric conformations is able to accommodate high crowding densities with less perturbation than is expected for a population of spherical particles. In particular, random polymers are typically characterized by relatively compact chain segments, sometimes referred to as blobs, separated by more extended segments. Crowding may favor local compact structures, while still allowing extended segments to pass between the particles.

The scaled particle theory can be extended to other convex shapes, but even these shapes may not be an adequate representation of unfolded protein conformations, which are likely to present both convex and concave surfaces to crowding molecules. Similar treatments of

macromolecular crowding by other authors also assume convex shapes for the unfolded molecules (6,50,51), an issue requiring further consideration.

Additional insights may come from studies of synthetic polymers at high densities. In 1949, Flory (52) predicted that the molecular dimensions of real polymers in concentrated melts should scale to the 1/2 power of chain length, rather than $\approx 3/5$ as predicted for well-solvated excluded-volume chains in dilute solution (i.e., fully swollen). That is, the behavior of real chains at very high concentrations approaches that of an idealized random flight chain that can cross itself, with a fractal dimension of 2. This rather counter-intuitive result has since been expanded upon and confirmed using neutron scattering experiments analogous to those described here (reviewed in de Gennes (53)). The situation involving crowding of disordered chains by globular structures is different from that of a mixture of random polymers, because the globular structures are presumed to be insensitive to the crowding. However, it may be that the value 2 represents a limiting value for the fractal dimension of a polymer at high densities, and polymer melts may be a useful analog for unfolded proteins under crowding conditions.

The results presented here demonstrate the potential of neutron scattering as a method for studying unfolded proteins, or other structures, in the presence of high concentrations of other molecules. This approach offers a unique perspective into the behavior of these molecules under conditions that may resemble those found in vivo, at least with respect to the fraction of volume accessible to macromolecules. The very good agreement with a computational simulation for relatively low crowding densities is encouraging, whereas the results at a higher density indicate that there is more to be learned. Further understanding of the crowding effects on unfolded molecules will likely require both more sophisticated simulations and more complete experimental data, including measurements at higher concentrations and with other unfolded proteins and crowding agents. Experiments with proteins of different relative sizes may be particularly valuable in further developing and testing theories based on steric exclusion.

SUPPORTING MATERIAL

One figure is available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(11)00066-X.

REFERENCES

1. Ellis, R. J. 2001. Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci.* 26:597–604.

2. Zhou, H.-X., G. Rivas, and A. P. Minton. 2008. Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.* 37:375–397.

3. Ghaemmaghami, S., and T. G. Oas. 2001. Quantitative protein stability measurement in vivo. *Nat. Struct. Biol.* 8:879–882.

4. Qu, Y., and D. W. Bolen. 2002. Efficacy of macromolecular crowding in forcing proteins to fold. *Biophys. Chem.* 101-102:155–165.

5. Ignatova, Z., and L. M. Gierasch. 2004. Monitoring protein stability and aggregation in vivo by real-time fluorescent labeling. *Proc. Natl. Acad. Sci. USA.* 101:523–528.

6. Minton, A. P. 2005. Models for excluded volume interaction between an unfolded protein and rigid macromolecular cosolutes: macromolecular crowding and protein stability revisited. *Biophys. J.* 88:971–985.

7. Cheung, M. S., D. K. Klimov, and D. Thirumalai. 2005. Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc. Natl. Acad. Sci. USA.* 102:4753–4758.

8. Stagg, L., S.-Q. Zhang, …, P. Wittung-Stafshede. 2007. Molecular crowding enhances native structure and stability of α/β protein flavodoxin. *Proc. Natl. Acad. Sci. USA.* 104:18976–18981.

9. Dong, H., S. Qin, and H.-X. Zhou. 2010. Effects of macromolecular crowding on protein conformational changes. *PLOS Comput. Biol.* 6:e1000833.

10. Hong, J., and L. M. Gierasch. 2010. Macromolecular crowding remodels the energy landscape of a protein by favoring a more compact unfolded state. *J. Am. Chem. Soc.* 132:10445–10452.

11. Miklos, A. C., C. Li, …, G. J. Pielak. 2010. Volume exclusion and soft interaction effects on protein stability under crowded conditions. *Biochemistry.* 49:6984–6991.

12. Ittah, V., E. Kahana, …, E. Haas. 2004. Applications of time-resolved resonance energy transfer measurements in studies of the molecular crowding effect. *J. Mol. Recognit.* 17:448–455.

13. Millett, I. S., S. Doniach, and K. W. Plaxco. 2002. Toward a taxonomy of the denatured state: small angle scattering studies of unfolded proteins. *Adv. Protein Chem.* 62:241–262.

14. Kohn, J. E., I. S. Millett, …, K. W. Plaxco. 2004. Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci. USA.* 101:12491–12496.

15. Jacob, J., R. S. Dothager, …, T. R. Sosnick. 2007. Fully reduced ribonuclease A does not expand at high denaturant concentration or temperature. *J. Mol. Biol.* 367:609–615.

16. Wang, Y., J. Trewhella, and D. P. Goldenberg. 2008. Small-angle x-ray scattering of reduced ribonuclease A: effects of solution conditions and comparisons with a computational model of unfolded proteins. *J. Mol. Biol.* 377:1576–1592.

17. Jacques, D. A., and J. Trewhella. 2010. Small-angle scattering for structural biology—expanding the frontier while avoiding the pitfalls. *Protein Sci.* 19:642–657.

18. Engelman, D. M., and P. B. Moore. 1972. A new method for the determination of biological quaternary structure by neutron scattering. *Proc. Natl. Acad. Sci. USA.* 69:1997–1999.

19. Capel, M. S., D. M. Engelman, …, P. B. Moore. 1987. A complete mapping of the proteins in the small ribosomal subunit of *Escherichia coli*. *Science.* 238:1403–1406.

20. Pardon, J. F., D. L. Worcester, …, B. M. Richards. 1975. Low-angle neutron scattering from chromatin subunit particles. *Nucleic Acids Res.* 2:2163–2176.

21. Neylon, C. 2008. Small angle neutron and x-ray scattering in structural biology: recent examples from the literature. *Eur. Biophys. J.* 37:531–541.

22. Daoud, M., J. P. Cotton, …, P. G. de Gennes. 1975. Solutions of flexible polymers. neutron experiments and interpretation. *Macromolecules.* 8:804–818.

23. Hammouda, B. 1993. SANS from homogeneous polymer mixtures: a unified overview. *Adv. Polym. Sci.* 106:87–133.

24. Le Coeur, C., B. Demé, and S. Longeville. 2009. Compression of random coils due to macromolecular crowding. *Phys. Rev. E.* 79:031910.

25. Weisberg, R. A., and M. E. Gottesman. 1999. Processive antitermination. *J. Bacteriol.* 181:359–367.

26. Van Gilst, M. R., W. A. Rees, …, P. H. von Hippel. 1997. Complexes of N antitermination protein of phage λ with specific and nonspecific RNA target sites on the nascent transcript. *Biochemistry.* 36:1514–1524.

27. Legault, P., J. Li, …, J. Greenblatt. 1998. NMR structure of the bacteriophage λ N peptide/*boxB* RNA complex: recognition of a GNRA fold by an arginine-rich motif. *Cell.* 93:289–299.

28. Conant, C. R., J. P. Goodarzi, …, P. H. von Hippel. 2008. The antitermination activity of bacteriophage λ N protein is controlled by the kinetics of an RNA-looping-facilitated interaction with the transcription complex. *J. Mol. Biol.* 384:87–108.

29. Rees, W. A., S. E. Weitzel, …, P. H. von Hippel. 1996. Bacteriophage λ N protein alone can induce transcription antitermination in vitro. *Proc. Natl. Acad. Sci. USA.* 93:342–346.

30. Studier, F. W. 2005. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* 41:207–234.

31. Jeffries, C. M., A. E. Whitten, …, J. Trewhella. 2008. Small-angle x-ray scattering reveals the N-terminal domain organization of cardiac myosin binding protein C. *J. Mol. Biol.* 377:1186–1199.

32. Lake, J. A. 1967. An iterative method of slit-correcting small angle x-ray data. *Acta Crystallogr.* 23:191–194.

33. Svergun, D., C. Barberato, and M. H. J. Koch. 1995. CRYSOL—a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Cryst.* 28:768–773.

34. Hamiaux, C., J. Pérez, …, P. Vachette. 2000. The BPTI decamer observed in acidic pH crystal forms pre-exists as a stable species in solution. *J. Mol. Biol.* 297:697–712.

35. Kline, S. R. 2006. Reduction and analysis of SANS and USANS data using IGOR Pro. *J. Appl. Cryst.* 39:895–900.

36. Güntert, P., C. Mumenthaler, and K. Wüthrich. 1997. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* 273:283–298.

37. Goldenberg, D. P. 2003. Computational simulation of the statistical properties of unfolded proteins. *J. Mol. Biol.* 326:1615–1633.

38. Chothia, C. 1975. Structural invariants in protein folding. *Nature.* 254:304–308.

39. Lebowitz, J. L., E. Helfand, and E. Praestgaard. 1965. Scaled particle theory of fluid mixtures. *J. Chem. Phys.* 43:774–779.

40. Svergun, D. I., S. Richard, …, G. Zaccai. 1998. Protein hydration in solution: experimental observation by x-ray and neutron scattering. *Proc. Natl. Acad. Sci. USA.* 95:2267–2272.

41. Gottschalk, M., K. Venu, and B. Halle. 2003. Protein self-association in solution: the bovine pancreatic trypsin inhibitor decamer. *Biophys. J.* 84:3941–3958.

42. Guinier, A. 1939. X-ray diffraction at very small angles: application to the study of related ultramicroscopic phenomena [La diffraction des rayons x aux très petits angles, application à l'étude de phénomèmes ultramicroscopiques]. *Ann. Phys.* 12:161–237.

43. Porod, G. 1982. General theory. *In* Small-Angle X-Ray Scattering. O. Glatter and O. Kratky, editors. Academic Press, London, UK. 17–51.

44. Teixeira, J. 1988. Small-angle scattering by fractal systems. *J. Appl. Cryst.* 21:781–785.

45. Schmidt, P. W. 1989. Use of scattering to determine the fractal dimension. *In* The Fractal Approach to Heterogeneous Chemistry. D. Avnir, editor. John Wiley & Sons, Chichester, UK. 67–79.

46. Beaucage, G. 1996. Small-angle scattering from polymeric mass fractals of arbitrary mass-fractal dimension. *J. Appl. Cryst.* 29:134–136.

47. Dewey, T. G. 1997. Fractals in Molecular Biophysics. Oxford University Press, Oxford, UK.

48. Beaucage, G. 2008. Toward resolution of ambiguity for the unfolded state. *Biophys. J.* 95:503–509.

49. Chen, S.-H., and D. Bendedouch. 1986. Structure and interactions of proteins in solution studied by small-angle neutron scattering. *Methods Enzymol.* 130:79–116.

50. Qin, S., and H.-X. Zhou. 2009. Atomistic modeling of macromolecular crowding predicts modest increases in protein folding and binding stability. *Biophys. J.* 97:12–19.

51. Qin, S., and H.-X. Zhou. 2010. Generalized fundamental measure theory for atomistic modeling of macromolecular crowding. *Phys. Rev. E.* 81:031919.

52. Flory, P. J. 1949. The configuration of real polymer chains. *J. Chem. Phys.* 17:303–310.

53. de Gennes, P.-G. 1979. Chapter II. Polymer melts. *In* Scaling Concepts in Polymer Physics.. Cornell University Press, Ithaca, NY. 54–68.