JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Internal Nanosecond Dynamics in the Intrinsically Disordered Myelin Basic Protein

Andreas M. Stadler,*[,†] Laura Stingaciu,[‡] Aurel Radulescu,[§] Olaf Holderer,[§] Michael Monkenbusch,[†] Ralf Biehl,[†] and Dieter Richter[†]

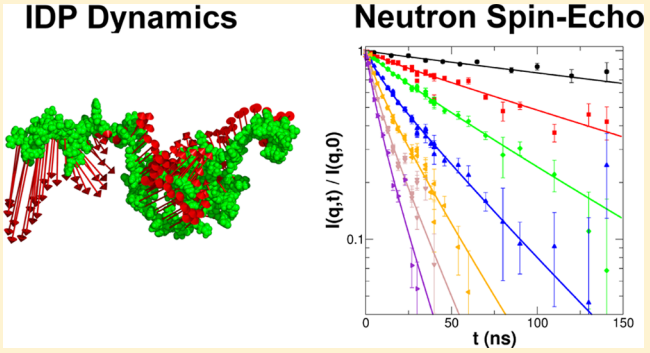
[†]Jülich Centre for Neutron Science JCNS and Institute for Complex Systems ICS, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

[‡]Jülich Centre for Neutron Science JCNS, Forschungszentrum Jülich GmbH, Outstation at SNS, POB 2008, 1 Bethel Valley Road, Oak Ridge, Tennessee 37831-6473, United States

[§]Jülich Centre for Neutron Science JCNS, Forschungszentrum Jülich GmbH, Outstation at MLZ, Lichtenbergstraße 1, 85747 Garching, Germany

Supporting Information

ABSTRACT: Intrinsically disordered proteins lack a well-defined folded structure and contain a high degree of structural freedom and conformational flexibility, which is expected to enhance binding to their physiological targets. In solution and in the lipid-free state, myelin basic protein belongs to that class of proteins. Using small-angle scattering, the protein was found to be structurally disordered similar to Gaussian chains. The combination of structural and hydrodynamic information revealed an intermediary compactness of the protein between globular proteins and random coil polymers. Modeling by a coarse-grained structural ensemble gave indications for a compact core with flexible ends. Neutron spin−echo spectroscopy measurements revealed a large contribution of internal dynamics to the overall diffusion. The experimental results showed a high flexibility of the structural ensemble. Displacement patterns along the first two normal modes demonstrated that collective stretching and bending motions dominate the internal modes. The observed dynamics represent nanosecond conformational fluctuations within the reconstructed coarse-grained structural ensemble, allowing the exploration of a large configurational space. In an alternative approach, we investigated if models from polymer theory, recently used for the interpretation of fluorescence spectroscopy experiments on disordered proteins, are suitable for the interpretation of the observed motions. Within the framework of the Zimm model with internal friction (ZIF), a large offset of 81.6 ns is needed as an addition to all relaxation times due to intrachain friction sources. The ZIF model, however, shows small but systematic deviations from the measured data. The large value of the internal friction leads to the breakdown of the Zimm model.



1. INTRODUCTION

The expected structural and dynamic properties of intrinsically disordered proteins (IDPs) reach from very soft structures, over folded elements connected by extended and flexible loops, to fully disordered polypeptide chains.[1,2] The biological role of IDPs is believed to be highly conformationally adaptive, which would be important for association with binding partners, or to respond rapidly to different environmental conditions.[3−5] IDPs that are involved in molecular recognition processes can even fold into different structures when bound to their different physiological target proteins, reviewed for example by Wright and Dyson.[6] The dynamics of unfolded peptide chains are essential for protein folding. Specifically, the sampling of the energy landscape and the exploration of a large conformational space prior to the collapse into a folded structure are strongly driven by conformational motions of the unfolded peptide chain. Concerning IDP involved in molecular binding, two extreme scenarios of folding and binding are discussed: on one hand, the "induced folding" mechanism, where folding of the IDP occurs after binding of the disordered protein to its physiological target;[7] on the other hand, the case of "conformational selection", where specific prefolded conformations are sampled in the accessible conformational space of the IDP, which are then selected by binding of the target protein.[8] In both scenarios, the exploration of a large conformational space of the IDP is the crucial point of the folding process. A characteristic physicochemical property of IDP, as pointed out by Uversky et al.,[9] is the combination of a low hydrophobicity

Received: January 13, 2014
Revised: March 7, 2014
Published: April 23, 2014

   dx.doi.org/10.1021/ja502343b | J. Am. Chem. Soc. 2014, 136, 6987−6994

with a large number of charged residues, which results in the destabilization of secondary structure. Crystallographic structures of IDP do not exist due to the existence of a large number of different conformational states. However, at low resolution, the protein structure in solution can be well-characterized by small-angle scattering of X-rays (SAXS) or neutrons (SANS), while neutron spin−echo spectroscopy (NSE) is a well-suited method to study polymer dynamics in solution[10] and functional relevant motions of protein domains[11−13] up to several hundred nanoseconds in the small-angle scattering range. Conformational dynamics in the nanosecond time-scale of unfolded proteins and IDP were investigated by single-molecule Förster resonance energy transfer (FRET) and nanosecond fluorescence correlation spectroscopy (nFCS).[14,15] Neutron and X-ray scattering measurements are highly complementary methods compared to fluorescence spectroscopy techniques and offer the advantage that native proteins without fluorescence dye labels can be investigated. Myelin basic protein (MBP) is a major component of the myelin sheath in the central nervous system. Several isoforms of MBP exist, but the major human and bovine isoform is the 18.5 kDa form.[16] *In vivo* MBP undergoes a large range of post-translational modifications, resulting in various charge isomers named C1 to C8, where C1 is the most cationic and least modified version.[16] MBP is not an integral membrane protein but, due to its positive charge, interacts with the lipid head groups and stabilizes the membranes.[16] From a medical aspect, MBP has significant importance as there are many neurological disorders, such as multiple sclerosis, which are related to MBP misfolding and malfunction.[16] In aqueous solution and in the lipid-free form, MBP is primarily unstructured and is classified as intrinsically disordered,[16] as judged from the net charge versus hydrophobicity scheme by Uversky et al.[9,17] Lipid-free MBP is not completely unfolded but retains some elements of α-helix and β-sheet.[16,18] Binding of lipids to MBP induces folding of the protein.[18] From a biophysical point of view, MBP is an interesting system to study the basic physical properties of disordered or partially unfolded proteins in solution. In this article, we present a study on the nature and extent of large conformational motions in MBP as an example for the role of these in IDP.

2. RESULTS AND DISCUSSION

2.1. MBP Structure Measured by SAXS and SANS. Small-angle scattering was measured to gain information about the structure of MBP as a prerequisite for the NSE experiments. Measured SAXS and SANS form factors of MBP are given in Figure 1. Both curves show power law scattering at $q > 0.1$ Å^{-1} with power law coefficients of −2.12 and −2.13 ± 0.01 for SAXS and SANS, respectively. A power law coefficient of −2 is the characteristic sign of Gaussian chain polymers in Θ solvent and in the melt. The slightly steeper slope of the measured data might indicate a more compact conformation as compared to a Gaussian chain. The measured SAXS and SANS curves can be described using the Debye equation for Gaussian chains yielding Guinier radii R_G of 37.8 Å for SAXS and 33.6 Å for SANS. Using the Guinier approximation, R_G values of 33.0 Å for the SAXS and 33.8 Å for the SANS data were determined. The R_G from the SANS data are consistent for both the Debye model and the Gaussian approximation. The SAXS data deviate slightly from the Debye model at the smallest q values, which explains the larger R_G of the Debye model as compared to the Guinier approximation. Reverse Monte Carlo simulations were

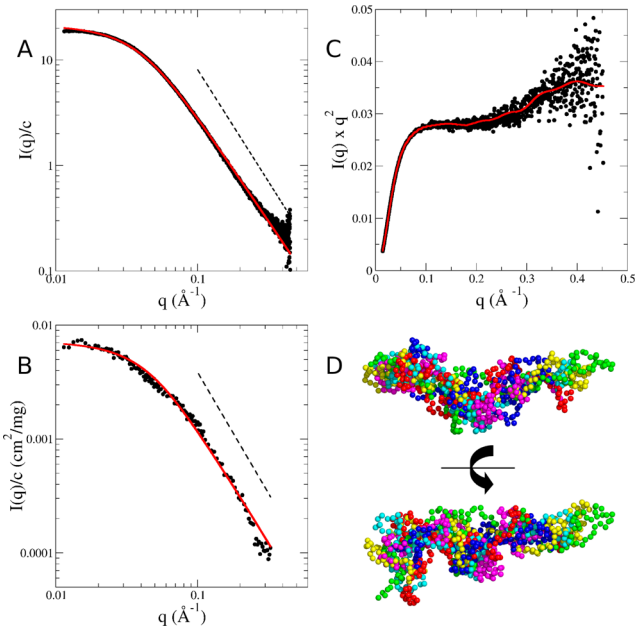


Figure 1. Measured small-angle scattering of MBP with structural models. (A) SAXS data of MBP at 4.5 mg/mL, (B) SANS intensity of MBP extrapolated to infinite dilution. The solid lines in (A) and (B) are fits with the Debye equation for Gaussian chains. Power law scattering behavior above $q > 0.1$ Å^{-1} is indicated by the straight lines. The power law coefficients for SAXS and SANS are −2.12 and −2.13, respectively. (C) Kratky plot of the measured SAXS data. The solid line is the calculated scattering curve of the conformational ensemble. (D) Representative coarse-grained conformations of MBP as determined by using the EOM software. The structures are rotated by 90° in the lower part of the figure. The blue structure is referred to as model A in the article.

applied to generate a coarse-grained ensemble representing the structural characteristics of MBP. The EOM software program[19] was used to generate 10 000 random-like Cα chains followed by the selection of a representative ensemble describing the measured SAXS data.

The distributions of R_G and of the maximal dimension D_{max} of the conformational ensemble are shown in Figure S1 in the Supporting Information. The average $\langle R_G \rangle$ and $\langle D_{max} \rangle$ from the selected ensemble are 35 and 115 Å, respectively. The six conformations, which occurred most often in the selection process (models A−F), are shown in Figure 1. Structure factors at 10 and 54 mg/mL were obtained by division of the measured SANS and SAXS data at high concentration by the protein form factor. At 10 mg/mL, the structure factor was found to be close to unity, and at 54 mg/mL, the structure factor showed modulations, which could be described by an effective hard sphere structure factor[20] with a fitted volume fraction of ϕ = 3.3% and an effective hard sphere radius of R_{HS} = 41 Å (see Figure S2). The MBP solution is well below the critical overlap concentration $c^* = M/(4/3\pi N_A R_G^3)$ of 202 (188) mg/mL for the R_G determined by SAXS (SANS).

2.2. Shape of MBP. The general features of the selected coarse-grained conformations indicate a central core region with flexible termini, while the overall shape appears to be slightly bent. The shape, the R_G, and the D_{max} of MBP are in agreement with SAXS measurements of recombinant murine MBP C1 and C8 charge isomers.[21] In that study, the GASBOR algorithm was used for ab initio modeling. That algorithm is

dx.doi.org/10.1021/ja502343b | J. Am. Chem. Soc. 2014, 136, 6987−6994

optimized for folded compact proteins, and information about disordered regions is lost during the averaging process.

Circular dichroism (CD) was measured to determine the secondary structure content of MBP in heavy water buffer and in deuterated 30% trifluoroethanol (TFE-d3) buffer (see Figure 2A). The secondary structure content was determined using the CDPro software package.[22] In D_2O, the protein was found to be 44% folded (α-helical, β-sheet, turns) and 56% unfolded, which is within the accuracy of the method identical to previous results of MBP in natural abundance buffer (47% folded, 53% unfolded).[18] The compact core of the coarse-grained structural ensemble of MBP might actually represent the folded structure. Folding of MBP can be induced by 30% v/v TFE,[23] and it was hypothesized that under those solvent conditions the in vivo environment of the protein is mimicked.[16] The folded structure content of MBP in 30% TFE is increased to 59% and the unfolded structure reduced to 41%. Cryo-electron microscopy of lipid-bound MBP revealed a pronounced bent shape of the protein,[24,25] which was later confirmed by SAXS in solution.[26] The lipid-bound model is shown in Figure 2B. Binding of lipids was interpreted to occur in the core of the bent protein inducing protein folding (72% folded and 28% unfolded structural content).[18,26] This indicates that lipid binding induces the transition from the slightly bent and largely unfolded lipid-free conformation to the strongly bent and folded structure in the lipid-bound state. The selected conformations of the disordered MBP were used as representative models for the interpretation of the NSE data as described below.

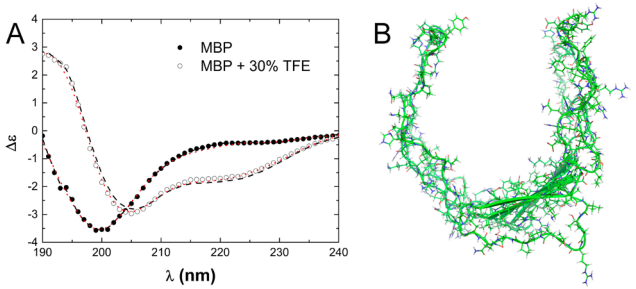


Figure 2. (A) Experimental CD data of the disordered state of MBP and the folded conformation induced by 30% TFE. The dashed and the dotted lines are reconstructions using the CDsstr and Contin algorithms, respectively. (B) Model of the folded conformation of lipid-bound MBP. The lipid molecules are not shown.

2.3. Combining Hydrodynamic and Structural Information. Diffusion coefficients of MBP in D_2O buffer were measured using dynamic light scattering (DLS) (see Figure S3). At infinite dilution, a hydrodynamic radius of R_H = 35.7 Å was obtained. The relation of hydrodynamic and gyration radii of MBP allows a comparison with ideal polymers and globular proteins. Gaussian polymer chains in good and Θ solvent are characterized by R_H/R_G ratios of 0.66 and 0.64,[27] respectively, while on the other extreme, the globular protein myoglobin, which has a similar molecular mass as MBP, has a much larger R_H/R_G ratio of 1.26,[28] being close to the theoretical limit of a sphere of $(5/3)^{0.5}$ = 1.29. The R_H/R_G ratios of MBP are 1.08 and 1.06 as obtained from the SAXS and SANS experiments, respectively, when the R_G from the Guinier approximation is taken. The R_H/R_G ratio serves as an indicator of compactness. MBP is more compact than an ideal flexible polymer chain but less compact than the folded protein. Compared to ideal polymer chains, it might also be speculated that a larger amount of interfacial water is bound to the surface of the protein, which would increase the R_H value.

2.4. Observation of Nanosecond Dynamics in MBP by NSE. Collective motions of MBP in 54 and 10 mg/mL solutions were explored using NSE. The NSE spectra of the 54 mg/mL solution are shown in Figure 3. Fourier times up to 140 ns were accessible for the 54 mg/mL sample, whereas for the 10 mg/mL solution, Fourier times up to 40 ns were measured.

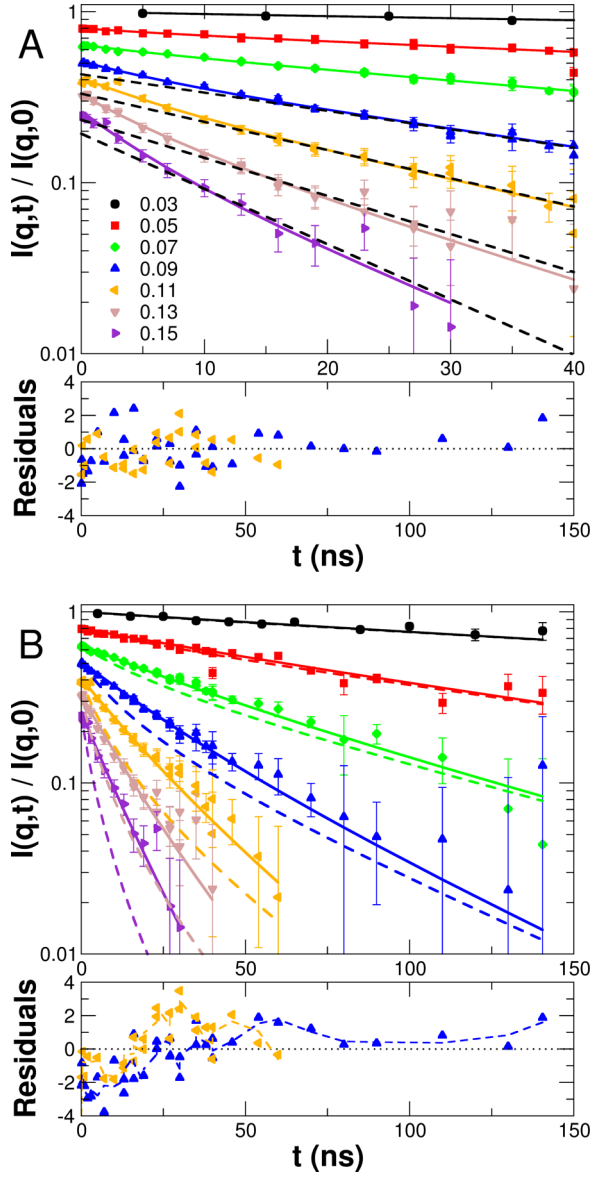


Figure 3. Measured NSE data of MBP. All spectra start at unity but are shifted consecutively by a factor of 0.8 for clarity. The q values in the legend are given in Å^{-1}. The residuals are given for q = 0.09 and 0.11 Å^{-1} below the graphs. (A) Solid lines are fits to the NSE data using eq 1 with structural model A. The data and fits are shown only up to 40 ns for better visibility of the short-time behavior. The full spectra and fits are given in the Supporting Information. The dashed lines are exponential fits for t > 20 ns at q = 0.09 Å^{-1} and for t > 10 ns at q = 0.11−0.15 Å^{-1} to extrapolate the long-time diffusion limit. A clear separation between internal dynamics and global protein diffusion is observed at short times with approximately t < 10 ns. (B) Dashed lines are calculated curves using the Zimm model; the solid lines are fits with the ZIF model.

dx.doi.org/10.1021/ja502343b | J. Am. Chem. Soc. 2014, 136, 6987−6994

In general, protein dynamics measured by NSE in solution consist of global diffusion and internal conformational motions. Typically, the observed relaxation due to internal protein dynamics decays significantly faster than global protein diffusion, which is the remaining contribution to the NSE spectra in the long-time limit. The effective diffusion coefficients D_{eff} from the initial slope of the spectra of the 54 mg/mL solution were determined by a cumulant analysis $I(q,t)/I(q,0) = \exp(K_1 t + 1/2K_2 t^2)$ with $D_{eff} = -K_1/q^2$. Within the errors, K_2 was found to be close to 0 below $q \leq 0.07$ Å^{-1} and a single-exponential fit was performed instead. A stretched exponential function $\exp(-(\Gamma t)^{\beta})$ could also describe the spectra. The stretching exponent β was found to be 0.86 for the range where internal motions are visible, being close to the Zimm prediction of 0.85 for Gaussian polymers in solution,[10] and approaching unity for low q as expected for translational diffusion. A single-exponential function was used for the determination of D_{eff} for the 10 mg/mL sample. To extrapolate the long-time diffusion limit—predominately related to global protein diffusion—the NSE spectra were fitted by single-exponential functions for $t > 20$ ns in the q range between 0.05 and 0.09 Å^{-1} and for $t > 10$ ns at $q = 0.11$ to 0.15 Å^{-1} (see Figure 3A). The D_{eff} obtained by NSE and DLS are plotted in Figure 4A.

At high protein concentration, interparticle effects and hydrodynamic interactions influence the observed dynamics. Therefore, structure factors and hydrodynamic functions were used to correct the $D_{eff}(q)$ and to determine $D_0(q)$ at infinite dilution. Hydrodynamic functions $H(q,c)$ were approximated by constant factors H_c, which were determined at 10 and 54 mg/mL by $H_c = D_c S(q \rightarrow 0)/D_0$, where D_0 is the extrapolated value at infinite dilution, D_c is the diffusion coefficient at the concentration c measured by DLS, and $S(q)$ is the structure factor. Hydrodynamic functions with $H_{10mg/mL} = 0.98$ and $H_{54mg/mL} = 0.71$ were obtained. After the correction, the $D_0(q)$ values from the initial slope of both concentrations overlap within the errors, demonstrating the validity of our approach (see Figure 4B). At small q values, the $D_0(q)$ determined by NSE converge toward the center of mass diffusion coefficient measured by DLS. At larger q values, the contribution of rotational diffusion and internal dynamics leads to a q-dependent increase of $D_0(q)$.

Protein aggregation could occur at the high concentrations used for the NSE measurements. The intensity- and volume-weighted size distributions were determined using DLS (see Figure S4 in the Supporting Information). The highly concentrated solutions were found to be bimodal in the intensity-weighted distributions consisting of the monomer and an aggregated fraction. The 10 mg/mL solution contained an intensity-weighted fraction of 77% with a R_H of 22 nm, while the 54 mg/mL sample had an intensity-weighted fraction of 55% with an R_H of 13 nm. At 2 and 6 mg/mL, the samples were found to be monodisperse. It is important to emphasize that in Rayleigh scattering of light the scattered intensities depend on the protein radius to the sixth power, and a small fraction of large aggregates even dominates the measured data. The volume-weighted size distributions revealed 100% monodispersity of the MBP monomer at all investigated concentrations with negligible contribution from aggregates. Due to the large size of the aggregates compared to the monomer, scattering of the aggregates is peaked at small q values but is strongly reduced in the q range covered by the NSE measurement (less than 0.2% of the scattered signal for the 54 mg/mL solution

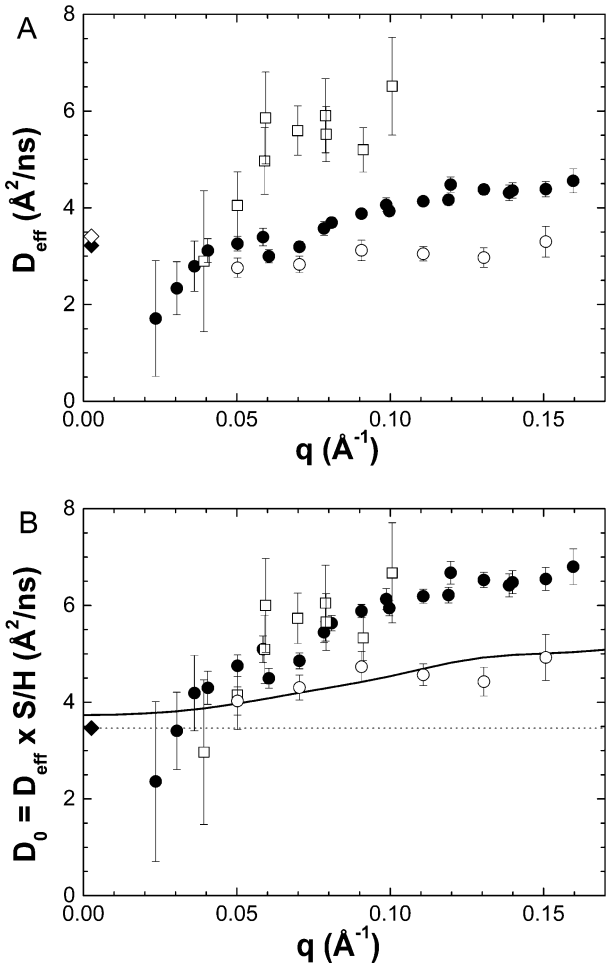


Figure 4. (A) Effective diffusion coefficient D_{eff} of MBP measured with NSE of the 54 (●) and 10 mg/mL (□) solutions. Diffusion coefficient of MBP in the long-time limit (○) of the 54 mg/mL sample. Diffusion coefficients of MBP measured using DLS at 54 (♦) and 10 mg/mL (◊). (B) D_{eff} corrected for structure factor and hydrodynamic effects to yield the diffusion coefficients D_0 at infinite dilution. DLS data extrapolated to infinite dilution (♦). The dotted line indicates the q-independent contribution of center of mass diffusion. The solid line is the calculated rigid body diffusion at infinite dilution of the structural model A.

and zero intensity for the 10 mg/mL sample at $q = 0.05$ Å^{-1}, where internal dynamics start to be seen by NSE). A significant amount of aggregated particles would also be visible in the experimentally determined structure factor as deviations from the theoretical hard sphere model, which is not the case. Furthermore, the presence of protein aggregates would result in deviations of the D_0 between the 10 and 54 mg/mL solutions, which is not observed. Therefore, it can be excluded that the small amount of aggregated protein modifies the NSE measurement.

Rigid body translational D_{0t} and rotational D_{0r} diffusion coefficients at infinite dilution of the structural ensemble determined by SAXS were calculated using the HYDROPRO program after side chains were positioned by the REMO algorithm.[29,30] Averaged over all models, the calculated translational and rotational diffusion coefficients are $D_{0t} = 3.7 \pm 0.2$ Å^2/ns and $D_{0r} = 2.35 \pm 0.36 \times 10^6$ s^{-1}, respectively. The corresponding average rotational correlation time is $\tau_{0r} = 1/(6D_{0r}) = 70.9$ ns. The average center of mass diffusion agrees

dx.doi.org/10.1021/ja502343b | J. Am. Chem. Soc. 2014, 136, 6987−6994

well with the measured DLS result at infinite dilution of 3.46 ± 0.05 Å^2/ns. See Table S1 for the calculated diffusion coefficients of all models. The calculated rigid body diffusion of model A is compared to the measured diffusion coefficients in Figure 4B. Model A is just one configuration of the selected structural ensemble. The calculated rigid body diffusion converges toward the measured center of mass coefficient at small q. Rotational diffusion of the protein results in an increase of the diffusion coefficient with q. The calculated rigid body diffusion of the models is in good agreement with the long-time limit determined by NSE.

Both translational and rotational diffusion of the largely disordered MBP can be described by rigid body diffusion of the structural ensemble. Internal protein dynamics are visible as an additional contribution sitting on top of the long-time limit resulting from rigid body diffusion. The difference between the measured $D_0(q)$ and the experimentally determined long-time diffusion coefficients, therefore, is attributed to internal protein dynamics, which enter the focus of the NSE instrument. A full interpretation including internal dynamics and protein diffusion is presented in the following sections.

In a first approach, we use the structural ensemble to interpret the full NSE spectra. The full NSE spectra of the 54 mg/mL solution were described by rigid body diffusion and internal protein dynamics according to[12]

$$I(q,t)/I(q,0) = [(1 - A(q)) + A(q)\exp(-\Gamma t)] \times \exp\left(-q^2 D_t \frac{H_t}{S(q)} t\right) \times \left(\sum_{l=0}^{15} S_l(q)\exp(-l(l+1)D_r H_r t)\right) / \sum_{l=0}^{15} S_l(q)$$

$$S_l(q) = \sum_m \left| \sum_i b_i j_l(qr_i) Y_{l,m}(\Omega_i) \right|^2 \quad (1)$$

where D_t and D_r are the calculated scalar translational and rotational diffusion coefficients of the rigid protein at infinite dilution from the structural models. Rotational diffusion of the rigid protein is expressed in spherical harmonics with spherical Bessel functions $j_l(qr)$, spherical harmonics $Y_{l,m}$, and scattering length densities b_l. Hydrodynamic interactions are considered by the parameters H_t and H_r. Interparticle effects are taken into account by the structure factor $S(q)$. Internal protein dynamics were described by an exponential decay with a q-independent rate Γ and a q-dependent contribution $A(q)$ of internal dynamics to the NSE spectra. Equation 1 was simultaneously fitted to all NSE spectra.

The fit with structural model A to the NSE spectra is shown in Figure 3A up to 40 ns. The full spectra with the fit are given in the Supporting Information Figure S5. At shorter times, the contribution of internal protein dynamics to the spectra becomes directly visible as compared to the long-time limit. In general, the structure-based model gives a good fit to the NSE data. Averaged over all models, a χ^2 value of 2.7 is obtained for a simultaneous fit to the whole data set. The residuals, illustratively shown for the spectra at q = 0.09 and 0.11 Å^{-1}, fluctuate randomly. The contribution of internal dynamics $A(q)$ contains the relevant information on the geometry of motion, and its interpretation is presented below. The q-independent hydrodynamic functions and the internal protein relaxation rates of all models are summarized in Table S1. On average, a single relaxation time $1/\Gamma$ = 8.4 ± 2.0 ns of the whole structural ensemble could be determined from the NSE measurement. The average translational hydrodynamic function is H_t = 0.59 ± 0.07, while the rotational hydrodynamic function was found to be H_r = 0.8. The average value of the translational hydrodynamic function is within the errors slightly below the experimentally determined result by DLS. Intrinsic viscosities $[\eta]$ of the structural models were calculated using the programs HYDROPRO[29] and HYDRO++,[31] which are related to H_t by $H_t = 1 - c[\eta]$. HYDRO++ in contrast to HYDROPRO employs a coarse-grained bead-modeling approach, where each bead represents an amino acid. HYDRO++ was run in a configuration where each bead has a uniform radius of half the distance between the residues. The calculated $[\eta]$ of all models are summarized in Table S1. The average values of $[\eta]$ using HYDROPRO and HYDRO++ are 15.6 ± 3.3 and 8.6 ± 2.2 mL/g, corresponding to H_t = 0.16 ± 0.18 and 0.54 ± 0.12, respectively, at c = 54 mg/mL. The experimentally determined hydrodynamic factor by NSE agrees well with the HYDRO++ calculation of the structural ensemble, validating both the experimental results and the structural models. The shell-modeling approach of HYDROPRO, however, fails for the disordered structures as it yields a very small H_t value, which is out of the physically meaningful range. As expected, the experimentally determined translational and rotational hydrodynamic functions of MBP are also smaller than theoretical calculations of hard sphere suspensions, which give values of H_t = 0.94, H_r = 0.98, with a volume fraction ϕ = 0.033 as obtained from the structure factor fit.[32] The lower values of H_t and H_r of MBP compared to hard spheres are related to the Perrin friction factor of the strong asymmetric shape of the protein.[33]

In an alternative approach, we test whether the dynamics of MBP can be described using simplified models from polymer theory. The Zimm model is a coarse-grained description of the dynamics of polymers in solution including hydrodynamic interactions.[27] The Zimm model of a finite chain consists of N beads, which are connected by entropic springs with a uniform bead distance l. The internal motions of the chain including rotational diffusion are represented by relaxation modes with mode number p and characteristic times τ_p given by

$$\tau_p = \frac{\eta R_e^3}{\sqrt{3\pi} k_B T} p^{-3\nu} \quad (2)$$

with the D_2O viscosity η = 1.679 cP at 10 °C. As the SAXS and SANS measurements could be described well with the Debye form factor, we are using the values for Gaussian chain statistics with ν = 0.5 and an end-to-end distance of $R_e = \sqrt{6} R_G$ = 80.8 Å with the measured R_G of 33 Å from the Guinier approximation of the SAXS data. Using 40 beads then yields l = 12.8 Å with $R_e = (l^2 N^{2\nu})^{1/2}$. Using more beads does not change the calculated curves. The first Zimm mode, corresponding to the rotational relaxation time,[27] equals τ_1 = 73.9 ns being close to the value of the structural models with τ_{0r} = 70.9 ns.

Within the Zimm model the dynamic structure factor is then given by

dx.doi.org/10.1021/ja502343b | J. Am. Chem. Soc. 2014, 136, 6987−6994

$$B(n, m, t) = (n - m)^{2\nu} l^2 + \frac{4R_e^2}{\pi^2} \sum_{p=1}^{p_{max}} \frac{1}{p^{2\nu+1}} \cos\left(\frac{\pi p n}{N}\right) \cos\left(\frac{\pi p m}{N}\right) \left[1 - \exp\left(-\frac{t}{\tau_p}\right)\right]$$

$$I(q, t) = \frac{\exp(-q^2 D_{cm} t)}{N} \sum_{n,m}^{N} \exp\left(\frac{-q^2 B(n, m, t)}{6}\right) \quad (3)$$

The maximal number of modes was set to 20 as above that value no changes in the spectra were observed. To consider modifications of the translational diffusion due to interprotein interactions, D_{cm} is replaced by $D_t/S(q)$ with the measured structure factor $S(q)$ and the translational diffusion coefficient D_t = 2.46 Å^2/ns of the 54 mg/mL solution measured by DLS after structure factor correction ($D_{DLS,5\%} \times 0.78$ = 2.46 Å^2/ns with $D_{DLS,5\%}$ = 3.15 Å^2/ns). In this way, the Zimm model describes only the internal dynamics including rotational diffusion.

The Zimm model, however, neglects motions that occur at length scales shorter than the typical distances between the coarse-grained beads. To account for additional microscopic interactions between neighboring beads—such as, for example, internal barriers, hindered dihedral rotations, side chain interactions, or hydrogen bonding—the classical Rouse and Zimm models were extended to the Rouse and Zimm models with internal friction (RIF and ZIF).[34,35] Internal friction within the ZIF model is represented by a relaxation time τ_i, which is added to each Zimm mode in eq 2 according to $\tau_p^{ZIF} = \tau_p + \tau_i$. The Zimm and ZIF models are compared to the NSE spectra in Figure 3B. For the internal friction as the single free parameter of the ZIF model, we obtain τ_i = 81.6 ± 3.2 ns and a χ^2= 3.6 for a simultaneous fit to the whole data set.

2.5. Internal Nanosecond Dynamics in MBP. We now come back to the interpretation of the observed internal dynamics in MBP. The Zimm model does not reproduce the shape of the spectra at larger q values where internal modes are visible and predicts a decay of the dynamic structure factor $I(q,t)/I(q,0)$, which is too fast. As the dynamics of other IDP and unfolded proteins have recently been characterized by nFCS measurements and interpreted using the RIF model,[15] we consider here the ZIF approach. At a first glance, the ZIF model appears to describe the dynamic structure factor in a reasonable approximation. However, a close inspection of the residuals shown for q = 0.09 and 0.11 Å^{-1} appears to reveal small but systematic deviations for the ZIF model: the model lies above the measured data at shorter times, while it is below the data at larger times. The residuals of all q values are given in Figure S6 in the Supporting Information. Furthermore, in order to make the ZIF model fit to the data, a large value of the internal friction τ_i = 81.6 ns needs to be imposed, which is even larger than the first mode of the pure Zimm model. The internal friction dominates all higher relaxation times to such an extent that the limiting value τ_i is quickly approached, and the characteristic $\tau_p \propto p^{-3/2}$ signature of the Zimm model completely breaks down. In addition, the first Zimm mode including internal friction equals 155.5 ns, which is far from the rotational relaxation time of the structural models and does not give the expected rotational relaxation time of the protein.[27] Adding the contribution of internal friction starting at the second Zimm mode and leaving the first mode unchanged does not reproduce the NSE spectra. Thus, the predicted line shape of the ZIF model does not describe the data, and the parameter τ_i leads to an unphysical result of the rotational diffusion. The treatment of rotational diffusion would need to be included separately into the ZIF model, while the framework of the ZIF model should only be applied for the description of internal dynamics.

As expected, the motions of MBP are significantly slower than compared to the ideal behavior of Gaussian polymer chains in solution represented by the pure Zimm model. This result is related to the fact that MBP is not fully unfolded but retains a compact core and a folded secondary structure content of 44%. The dynamics of folded protein structure elements do not possess the same flexibility as loosely connected polymers. Thus, the detailed q (space) and time dependence of the dynamic structure factor cannot be reconciled with a polymer-like (Zimm) relaxation. An open question here would be if the dynamics of fully unfolded proteins approach the properties of polymers.

For the case of MBP, the structural ensemble gives a significantly better description of the NSE spectra than the ZIF model as judged from the χ^2 values. To interpret the observed internal protein dynamics in more detail, the lowest lying collective excitations of the structural ensemble were calculated using normal mode (NM) analysis. The calculated displacement patterns of the first and second nontrivial NMs 7 and 8 of a representative conformation (model A) are shown in Figure 5A.

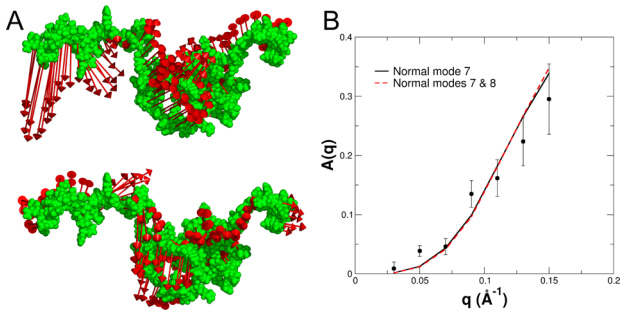


Figure 5. (A) Calculated NMs of the representative MBP conformation model A. NM 7 in the upper part, NM 8 in the lower part of the figure. The lengths of the vectors are increased for better visibility. (B) Contribution of internal protein dynamics to the NSE spectra $A(q)$. The solid and dashed lines are fits to the data using eq 4 with NM 7 and the combination of NMs 7 and 8.

In general, the NMs 7 and 8 correspond to bending and stretching motions of all structural models. The NMs 7 and 8 were used to describe the measured contribution of internal dynamics $A(q)$ (see Figure 5B). NM 7 is the dominating collective excitation and fully reproduces the measured $A(q)$. The root mean square displacements (rmsd) were calculated as average values of the whole structural ensemble and the errors as standard deviations. The rmsd of NM 7 was found to be 2.1 ± 0.3 Å, while the weighted combination of NMs 7 and 8 yielded a rmsd of 2.2 ± 0.3 Å. Including higher NMs does not modify the rmsd. The amplitudes of motions using NM analysis of all models are summarized in Table S2. The displacement vectors from NM analysis indicate that the center part of MBP remains rather rigid, while the termini are more flexible. In

dx.doi.org/10.1021/ja502343b | J. Am. Chem. Soc. 2014, 136, 6987−6994

general, the amplitudes of motion correspond to distances between the end termini of the coarse-grained structures. An intriguing idea would be that the observed dynamics represent movements with an average relaxation time of 8.4 ns between the different structures of the selected configurational ensemble.

3. CONCLUSION

In conclusion, we investigated the disordered state of MBP as model system for the properties of IDP. MBP was found to be structurally disordered although having a higher compactness than ideal Gaussian polymers in solution. Global protein diffusion and internal collective motions were studied by using NSE. A significant contribution of internal protein dynamics with a relaxation rate of 8.4 ns was found using an interpretation based on a representative structural ensemble. NM analysis of the structural ensemble indicates that the observed motions are essentially governed by low-frequency collective stretching and bending motions, where the termini were found to be especially flexible. The large-scale conformational motions increase the accessible protein surface, which facilitates the interaction with binding partners.[36] Due to the folded secondary structure and the compact core, the pure Zimm model cannot describe the observed dynamics. A further result of our work is that the ZIF approach that was applied in the interpretation of recent nFCS experiments of IDP and unfolded proteins approximately describes the NSE spectra but yields small systematic deviations of the dynamic structure factor of MBP. A value of the internal friction is obtained, which is larger than the first mode of the pure Zimm model.

4. MATERIALS AND METHODS

4.1. Sample Preparation. Bovine myelin basic protein was purchased from Sigma-Aldrich and dissolved in D_2O buffer (20 mM NaH_2PO_4/Na_2HPO_4, pH 4.8, 99.9 atom % D for D_2O) or in the D_2O buffer containing also 30% v/v deuterated TFE. Protein concentration was determined by optical absorption at 280 nm with an absorbance of $E_{1\%} = 5.89$. Sample temperature was 10 °C if not otherwise stated.

4.2. Circular Dichroism and Dynamic Light Scattering. Circular dichroism was measured on a J-810 spectropolarimeter (JASCO, Tokyo, Japan). The samples were measured in 1 mm thick quartz cuvettes under constant nitrogen flow at a concentration of 0.3 mg/mL at 20 °C. Dynamic light scattering was measured on a Zetasizer Nano ZS instrument (Malvern Instruments, Malvern, United Kingdom). Autocorrelation functions were analyzed by the CONTIN-like algorithm.[37] The hydrodynamic radius R_H was determined according to $R_H = k_BT/(6\pi\eta D)$ with the D_2O viscosity $\eta = 1.679$ cP at 10 °C.

4.3. Small-Angle Scattering. SAXS was measured on BM29 at the ESRF, Grenoble, France. The X-ray wavelength was 1 Å. Protein concentrations of 4.5 and 54 mg/mL were measured. SANS was measured on KWS-1 at the MLZ in Garching. The neutron wavelength was set to 4.5 Å. Protein concentrations were 2.9, 5.6, and 10 mg/mL. The measured background-corrected intensities were linearly extrapolated to infinite dilution to extract the form factor per unit mass. The scattering vector q is defined as $q = 4\pi/\lambda\cdot\sin(\theta/2)$ with the incident X-ray or neutron wavelength λ and the scattering angle θ.

The form factor of a Gaussian chain can be described by the Debye equation

$$I(q) = \frac{2I_0}{q^4 {R_G}^4}[\exp(-q^2 {R_G}^2) - 1 + q^2 {R_G}^2] \quad (4)$$

where R_G is the Guinier radius of the chain. The Guinier radius from the Guinier approximation was determined according to $I(q) = I_0\cdot\exp(-R_G\cdot q^2/3)$ in the range up to $qR_G < 1$ where the plot of ln $I(q)$ versus q^2 was found to be linear.

4.4. Neutron Spin–Echo Spectroscopy. NSE was measured on J-NSE at the MLZ. Three incident neutron wavelengths of 8, 10, and 12.8 Å were used. For better statistics in the determination of the long-time limit and for the interpretation with the models, the data were binned over the whole detector.

The rigid body diffusion of a structural model at infinite dilution was calculated according to[11]

$$D_0(q) = \frac{1}{q^2 F(q)} \sum_{j,k} \left\langle b_j \exp(-i\vec{q}\vec{r}_j) \begin{pmatrix} \vec{q} \\ \vec{q}\times\vec{r}_j \end{pmatrix} \hat{D} \begin{pmatrix} \vec{q} \\ \vec{q}\times\vec{r}_k \end{pmatrix} \times b_k \exp(i\vec{q}\vec{r}_k) \right\rangle \quad (5)$$

where $\hat{D}$ is the 6×6 diffusion tensor, which was calculated using the HYDROPRO program.[29]

Elastic normal modes were calculated using the MMTK software package.[38] In a first-order approximation for small displacements, internal protein dynamics can be expressed in terms of overdamped normal modes according to[12]

$$I(q, t) \propto I(q) + \sum_\alpha a_\alpha e^{-\lambda_\alpha t} P_\alpha(q)$$

$$P_\alpha = \left\langle \sum_{k,l}^{N} b_k b_l \exp(i\vec{q}(\vec{r}_k - \vec{r}_l))(q e_k^\alpha)(q e_l^\alpha) \right\rangle \quad (6)$$

where $I(q)$ is the elastic form factor, a_α is the squared mode amplitude, $\vec{e}^\alpha$ is the eigenvector of mode α, and λ_α is the relaxation rate of the overdamped mode. The normal modes are weighted using the equipartition principle that each mode has the same elastic energy. A general amplitude factor a is used according to $a_\alpha = a\omega_7^2/\omega_\alpha^2$ with the hypothetical eigenfrequencies ω of mode α and the lowest nontrivial mode 7. The resulting mean displacement of mode α is $\langle a_\alpha^{0.5} e^\alpha \rangle$ calculated as average over all atoms. For several modes, the mean displacements were summed. Using a common relaxation rate for all normal modes, the contribution of internal dynamics to the NSE spectra can be expressed as[12]

$$A(q) = \frac{\sum_\alpha a_\alpha P_\alpha(q)}{I(q) + \sum_\alpha a_\alpha P_\alpha(q)} \quad (7)$$

■ ASSOCIATED CONTENT

Supporting Information

R_G and D_{max} distribution, structure factor. Diffusion coefficients and size distributions measured with dynamic light scattering, NSE spectra. Translational and rotational diffusion coefficients and intrinsic viscosities of the structural models. Hydrodynamic functions, relaxation times of internal protein dynamics determined from the NSE measurement. Amplitudes of motion from NM analysis of the structural models. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

a.stadler@fz-juelich.de

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work is based on experiments performed at the ESRF, Grenoble, France, and at the Heinz Maier-Leibnitz Zentrum (MLZ), Garching, Germany.

REFERENCES

(1) Dyson, H. J. *Q. Rev. Biophys.* **2011**, *44*, 467.
(2) Fink, A. L. *Curr. Opin. Struct. Biol.* **2005**, *15*, 35.
(3) Tompa, P. *Trends Biochem. Sci.* **2002**, *27*, 527.
(4) Tompa, P. *Trends Biochem. Sci.* **2012**, *37*, 1.
(5) Dyson, H. J.; Wright, P. E. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197.
(6) Wright, P. E.; Dyson, H. J. *Curr. Opin. Struct. Biol.* **2009**, *19*, 31.
(7) Sugase, K.; Dyson, H. J.; Wright, P. E. *Nature* **2007**, *447*, 1021.
(8) Song, J.; Guo, L.-W.; Muradov, H.; Artemyev, N. O.; Ruoho, A. E.; Markley, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 1505.
(9) Uversky, V. N.; Gillespie, J. R.; Fink, A. L. *Proteins* **2000**, *41*, 415.
(10) Richter, D.; Monkenbusch, M.; Arbe, A.; Colmenero, J. *Neutron Spin Echo in Polymer Systems*; Springer: Berlin, 2005; Vol. *174*.
(11) Biehl, R.; Hoffmann, B.; Monkenbusch, M.; Falus, P.; Preost, S.; Merkel, R.; Richter, D. *Phys. Rev. Lett.* **2008**, *101*, 138102.
(12) Inoue, R.; Biehl, R.; Rosenkranz, T.; Fitter, J.; Monkenbusch, M.; Radulescu, A.; Farago, B.; Richter, D. *Biophys. J.* **2010**, *99*, 2309.
(13) Bu, Z.; Biehl, R.; Monkenbusch, M.; Richter, D.; Callaway, D. J. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17646.
(14) Mukhopadhyay, S.; Krishnan, R.; Lemke, E. A.; Lindquist, S.; Deniz, A. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2649.
(15) Soranno, A.; Buchli, B.; Nettels, D.; Cheng, R. R.; Muller-Spath, S.; Pfeil, S. H.; Hoffmann, A.; Lipman, E. A.; Makarov, D. E.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 17800.
(16) Harauz, G.; Ishiyama, N.; Hill, C. M.; Bates, I. R.; Libich, D. S.; Fares, C. *Micron* **2004**, *35*, 503.
(17) Uversky, V. N. *Protein Sci.* **2002**, *11*, 739.
(18) Polverini, E.; Fasano, A.; Zito, F.; Riccio, P.; Cavatorta, P. *Eur. Biophys. J.* **1999**, *28*, 351.
(19) Bernadó, P.; Mylonas, E.; Petoukhov, M. V.; Blackledge, M.; Svergun, D. I. *J. Am. Chem. Soc.* **2007**, *129*, 5656.
(20) Klebanov, I.; Gritsay, P.; Ginchitskii, N. *Phys. Rev. Lett.* **2006**, *10*, 3.
(21) Wang, C.; Neugebauer, U.; Bürck, J.; Myllykoski, M.; Baumgärtel, P.; Popp, J.; Kursula, P. *PLoS One* **2011**, *6*, e19915.
(22) Sreerama, N.; Woody, R. W. *Anal. Biochem.* **2000**, *287*, 252.
(23) Liebes, L. F.; Zand, R.; Phillips, W. D. *Biochim. Biophys. Acta, Protein Struct.* **1975**, *405*, 27.
(24) Beniac, D. R.; Luckevich, M. D.; Czarnota, G. J.; Tompkins, T. A.; Ridsdale, R. A.; Ottensmeyer, F. P.; Moscarello, M. A.; Harauz, G. *J. Biol. Chem.* **1997**, *272*, 4261.
(25) Ridsdale, R. A.; Beniac, D. R.; Tompkins, T. A.; Moscarello, M. A.; Harauz, G. *J. Biol. Chem.* **1997**, *272*, 4269.
(26) Haas, H.; Oliveira, C. L. P.; Torriani, I. L.; Polverini, E.; Fasano, A.; Carlone, G.; Cavatorta, P.; Riccio, P. *Biophys. J.* **2004**, *86*, 455.
(27) Doi, M.; Edwards, S. F. *The Theory of Polymer Dynamics*; Clarendon Press: Oxford, 1986.
(28) Stadler, A. M.; Pellegrini, E.; Johnson, M.; Fitter, J.; Zaccai, G. *Biophys. J.* **2012**, *102*, 351.
(29) de la Torre, J. G.; Huertas, M. L.; Carrasco, B. *Biophys. J.* **2000**, *78*, 719.
(30) Li, Y.; Zhang, Y. *Proteins* **2009**, *76*, 665.
(31) de la Torre, J. G.; Echenique, G. D. R.; Ortega, A. *J. Phys. Chem. B* **2007**, *111*, 955.
(32) Degiorgio, V.; Piazza, R.; Jones, R. B. *Phys. Rev. E* **1995**, *52*, 2707.
(33) Perrin, F. *J. Phys. Radium* **1934**, *5*, 497.
(34) Cheng, R. R.; Hawk, A. T.; Makarov, D. E. *J. Chem. Phys.* **2013**, *138*, 074112.
(35) Khatri, B. S.; McLeish, T. C. B. *Macromolecules* **2007**, *40*, 6770.
(36) Harauz, G.; Ladizhansky, V.; Boggs, J. M. *Biochemistry* **2009**, *48*, 8094.
(37) Provencher, S. W. *Comput. Phys. Commun.* **1982**, *27*, 229.
(38) Hinsen, K. *J. Comput. Chem.* **2000**, *21*, 79.