CHAPTER FOUR

Translational diffusion of unfolded and intrinsically disordered proteins

Irina V. Nesmelova[a,b,*], Daria L. Melnikova[c], Venkatesh Ranjan[d], Vladimir D. Skirda[c]

[a]Department of Physics and Optical Sciences, University of North Carolina, Charlotte, NC, United States
[b]Center for Biomedical Engineering and Science, University of North Carolina, Charlotte, NC, United States
[c]Department of Physics, Kazan Federal University, Kazan, Russia
[d]Department of Chemistry, University of North Carolina, Charlotte, NC, United States
*Corresponding author: e-mail address: irina.nesmelova@uncc.edu

Contents

Abstract

Translational (or self-diffusion) coefficient in dilute solution is inversely proportional to the size of a diffusing molecule, and hence self-diffusion coefficient measurements have been applied to determine the effective hydrodynamic radii for a range of native and nonnative protein conformations. In particular, translational diffusion coefficient measurements are useful to estimate the hydrodynamic radius of natively (or intrinsically) disordered proteins in solution, and, thereby, probe the compactness of a protein as well as its change when environmental parameters such as temperature, solution pH, or protein concentration are varied. The situation becomes more complicated in concentrated solutions. In this review, we discuss the translational diffusion of disordered proteins in dilute and crowded solutions, focusing primarily on the information provided by pulsed-field gradient NMR technique, and draw analogies to well-structured globular proteins and synthetic polymers.

Progress in Molecular Biology and Translational Science, Volume 166
ISSN 1877-1173
https://doi.org/10.1016/bs.pmbts.2019.05.004

1. Introduction

"The three-dimensional structure of a native protein in its normal physiological milieu (solvent, pH, ionic strength, presence of other components such as metal ions or prosthetic groups, temperature, and other) is the one in which the Gibbs free energy of the whole system is lowest; that is, that the native conformation is determined by the totality of interatomic interactions and hence by the amino acid sequence, in a given environment."[1] Water, the primary solvent for proteins in native environments, is a poor solvent for protein backbone, and, until not that long ago, the prevailing view has been that a protein adopts a stable, well-structured conformation stabilized by hydrogen, ionic, and disulfide bonds and by hydrophobic interactions.[1–3] This view was strongly supported by theoretical predictions of the two main secondary structure elements in proteins, the α-helix and β-sheet, made by Pauling et al.[3] and by first crystallographic and NMR structures of proteins.[4–6]

However, the Anfinsen's "thermodynamic hypothesis,"[1] quoted above, clearly suggests the conformational plasticity of a protein in response to its environment. Indeed, the protein is formed as a linear polymer of monomer building blocks—amino acids, each of which has distinct physico-chemical properties due to the unique structure of a side-chain that can be neutral, polar, or charged. Accordingly, the protein can be considered as a charged heteropolymer, or a polyampholyte,[7] with keeping in mind two important distinctions of synthetic polymer polyampholytes, typically a much longer chain and the disorder of the charge distribution along the sequence. The conformation of flexible polyampholytes is controlled by electrostatic interactions between charged groups, and the composition and the placement of various amino acids in the protein sequence determine whether the intra-protein interactions are more favorable than the side-chain interactions with solvent (see reviews[7,8] and references therein). For example, in water, protein sequences with low hydrophobicity and high net charge have a preference to adopt disordered, extended conformations, rather than to form a compact well-ordered structure.[8–10] The abundance and functional significance of intrinsically disordered proteins (IDPs) is now broadly recognized.[11–22] Within the human proteome, approximately 50% of all proteins are predicted to contain long disordered segments ($\geq$30 residues).[23,24] The addition of small molecule solutes (osmolytes) modulates solvent quality and, consequently, protein conformation.[25] Denaturing osmolytes such as

urea and guanidine hydrochloride (GuHCl) promote the formation of extended protein conformations.[26,27] In contrast, trimethylamine N-oxide (TMAO) bias the protein structure toward the folded conformation.[28,29]

In attempt to classify the extent of conformations that a protein can adopt, four different thermodynamically stable states (ordered, molten globule, pre-molten globule, and unfolded) are proposed.[30–33] Transitions between molten globule and pre-molten globule might represent different phase states of the protein, as they are separated by the first-order phase transition.[34,35] Despite such classification is useful, it is important to realize that, in principle, any protein can adopt a continuum of conformational states that may gradually change depending on the solution conditions. Therefore, it becomes important to know the overall size of the protein in a given environment.

Unlike well-structured folded proteins that can be successfully characterized by X-ray crystallography or NMR spectroscopy,[5] the characterization of intrinsically disordered or partially unfolded proteins remains challenging as they sample an ensemble of rapidly interconverting alternative conformations ranging from random coils to more structured conformations with secondary structure and residual tertiary structure elements. Typically, such conformational exchange occurs on a much faster time scale than most experimental biophysical methods can access. Therefore, a quantitative measure of an average dimension of the conformational ensemble, e.g., an ensemble-averaged hydrodynamic radius (R_H), becomes an important characteristic as it provides the information about the nature of the structures adopted by a protein.[36–38]

Despite the long-sustained interest to protein folding and exponentially growing number of publications on IDPs, experimental values of R_H for partially/fully unfolded or intrinsically disordered proteins are rather sparse in literature. In this review, we focus primarily on the information provided by pulsed-field gradient NMR technique, and draw analogies to well-structured globular proteins and synthetic polymers.

2. Dilute solutions

Our view on the shape and size of a protein is largely based on experimental data obtained in dilute solutions primarily because of experimental requirements of utilized biophysical methods and the relative simplicity of theoretical treatment that avoids the need to take into account intermolecular interactions. In dilute solution, the overall shape and size of a

protein can be experimentally assessed by the variety of experimental methods, including small-angle X-ray scattering (SAXS),[39–41] size-exclusion chromatography (SEC),[42,43] fluorescence correlation spectroscopy and dynamic light scattering (DLS),[44] or pulsed-field gradient NMR diffusion measurements (PFG NMR).[45] While SAXS experiments provide the estimate of the radius of gyration (R_g), SEC, fluorescence correlation spectroscopy, and PFG NMR probe the hydrodynamic radius R_H of the protein chain. The non-invasive NMR approach has particular appeal as the measurements of the hydrodynamic radius can be carried out under the variety of experimental conditions and over the broad range of protein concentrations that are not accessible to other methods, from dilute to extremely concentrated solutions. In this review, we focus only on PFG NMR studies.

PFG NMR measures the self-diffusion diffusion coefficient of a protein D (further referred simply as the diffusion coefficient), which depends on protein size. To establish this dependence, a model of protein molecule is required. Two extreme cases, approximating the protein by a solid sphere (gives the upper limit for the diffusion coefficient) or by a random coil (gives the lower limit for the diffusion coefficient), can be considered.

The diffusion coefficient of a protein modeled by a solid sphere is given by the Stokes-Einstein formula[46] that relates it to the hydrodynamic radius of a sphere R_H:

$$D = \frac{kT}{6\pi\eta R_H}, \tag{1}$$

where η is the viscosity of pure solvent, k is the Boltzmann constant, and T is the temperature. The hydrodynamic radius of the protein is understood as the hydrodynamic radius of a sphere with the same diffusion coefficient. It is worth mentioning that the accurate estimate of R_H using Eq. (1) is not expected as the Stokes-Einstein formula was derived under several simplifying assumptions. First, the Stokes-Einstein formula only holds for infinite dilution, because it was derived for a single particle. Second, it was derived for a large solid sphere, whose size is much greater than the size of solvent molecules. Third, it was derived for a sphere with smooth surface, moving in a continuous fluid under no-slip conditions, even though this fluid consists of molecules at the microscopic scale. In general, the protein molecule has a non-spherical shape, rough surface (with amino acid side-chains sticking out), high degree of flexibility (where the degree of flexibility varies

dependent on the conformational state of the protein with disordered state possessing the highest flexibility), and the interactions with water (solvent) molecules cannot be ignored. Given the simplifying assumptions of Eq. (1), it is remarkable that simple Stokes-Einstein model satisfactory describes the large number of systems, including unfolded proteins,[47] and the agreement with experimental data can be further improved by taking into account the hydration layer around the protein[47] and by approximating the protein by the ellipsoid[48] or by describing its shape more accurately, using bead models.[49,50]

The major discrepancy of using Eq. (1) for describing the translational diffusion of proteins (and polymers in general) is related to how the viscosity is determined. Unlike in the case of a solid sphere with smooth surface, for proteins the viscosity must depend on the length and time scales of the experiment, which are determined by the size of diffusing molecule and by the diffusion time, at the very least over the distance of the protein's linear dimensions. Furthermore, in the case of flexible molecules, it is necessary to account for the local molecular dynamics, which can be found for homopolymers consisting of identical monomers.[51] However, the protein molecule consists of amino acids with different physicochemical properties. Thus, even if the local dynamics for each amino acid is known, the question on how to determine its mean value for the whole molecule in order to introduce a correcting coefficient for the viscosity in the Stokes-Einstein equation still remains. Consequently, despite some recent work in this direction,[52] there is no better alternative to Eq. (1) for proteins at present.

The application of the Stokes-Einstein approximation is particularly powerful for the comparative analysis of protein sizes, for example, for elucidating the magnitude of size changes due to unfolding in response to the change of environmental conditions. Calculated directly by using Eq. (1), the 37% increase of the hydrodynamic radius was found for the β-subunit of salt-mediated killer toxin from *Pichia farinose* due to the loss of compact structure upon dissociation of the heterodimer formed by α- and β-subunits.[53] Alternatively, Jones et al.[54] proposed to use a reference molecule to estimate the size change of a protein. If a reference molecule is used as an internal radius standard, then the following relationship for protein and reference molecules in the same solution follows from Eq. (1):

$$R_H = \frac{D_{ref}}{D} \times R_{ref} \tag{2}$$

where D_{ref} and R_{ref} are the diffusion coefficient and the hydrodynamic radius of the reference molecule, respectively. In order to obtain the accurate estimate of R_H using Eq. (2), the hydrodynamic radius of the reference molecule has to be known. Wilkins et al.[45] have calibrated their measurements using the radius of gyration of lysozyme in the folded state measured by SAXS and the hydrodynamic radius measured by PFG NMR. Any calibration though cannot be precise as the same simplifying assumptions as those that are used in the derivation of Stokes-Einstein equation (Eq. 1) would be applicable. However, the relative change in the hydrodynamic radius can be determined with high accuracy. Jones et al.[54] determined that the hydrodynamic radius of lysozyme increases by $38 \pm 1\%$ due to unfolding in urea, using dioxane as a reference molecule, and observed a good agreement with SAXS data[55]. They chose dioxane because the radius of dioxane did not change by the addition of urea. Using similar approach, the change of hydrodynamic radius of proteins, ranging from about 11% to 57%, due to unfolding under denaturing conditions (urea, GuHCl, extreme pH values, or disulfide bond reduction)[45,56–62] or associated with Ca^{2+} binding,[63] or, alternatively, upon folding in 40% trifluoroethanol[64] was determined.

In general, when the relative change of protein size is in question, using a reference molecule is not necessary. For example, in the case of temperature-induced unfolding, the estimate of protein size changes according to Eq. (2) can be done without using a reference molecule. Indeed, as it follows from Eq. (1), in the absence of processes that could lead to the change of protein size with temperature, i.e., protein aggregation or unfolding, the temperature dependence of D is determined only by the temperature dependence of solvent viscosity and has the slope of T/η. Thus, if the temperature dependence of protein diffusion coefficient is extrapolated to temperatures above the unfolding transition according to T/η, it represents the diffusion coefficient D_{ref} of a protein with unperturbed structure and unchanged hydrodynamic radius. Note that this analysis is done under the assumption of the independence of protein diffusion coefficient on the local dynamics. For flexible polymers, the diffusion coefficient was shown to change with changing local segmental dynamics.[51] It is likely though that the influence of local dynamics on the diffusion coefficient will be minimal in dilute solutions.

The relative size change caused by the temperature-induced protein expansion can be estimated from the ratio of experimentally measured diffusion coefficient and extrapolated D_{ref} taken at the same temperature. Using this approach, approximately the 22% change in R_H for lysozyme and ribonuclease A was detected,[65] whereas the average size of conformers

within the unfolded state ensemble for the HIV-1 nucleocapsid NCp7 protein was found to be 30–40% larger than the folded state structure.[66]

The Stokes-Einstein approximation was also applied to estimate the size of intrinsically disordered proteins.[61,67–72] For a 130-residue fragment (D1–D4) taken from a fibronectin-binding protein of *Staphylococcus aureus*, containing four fibronectin-binding repeats and unfolded but biologically active at neutral pH, the comparison to well-structured proteins showed on average a 75% larger hydrodynamic radius than would be expected for a compact fold,[68] while an 80% and a twofold increase of the hydrodynamic radius has been observed for the primary DNA-recognition subdomain of the Sleeping Beauty transposase[73] and for N-terminal domain of p53 protein,[70] respectively. A smaller increase of R_H as compared to an equivalent well-structured, folded protein was measured for the intrinsically disordered malaria surface protein MSP2 (about 48–70% depending on experimental conditions)[74] and for α-synuclein (about 50%).[58]

The important conclusion that follows from diffusion measurements in dilute solutions mentioned above is that the protein molecule remains compact[75] when it is unfolded under denaturing conditions or intrinsically disordered, with the largest increase of the hydrodynamic radius to the best of our knowledge being twofold, reported for N-terminal domain of p53 protein.[70] When comparison was done, all reported R_H values were found to be smaller than estimated for random coils.[60,61,66,69–71,74] In relation to this conclusion, the important question arises about the extent of validity of approximating the unfolded or IDP by Stokes-Einstein model assuming that the protein behaves as a solid sphere. Indeed, the disordered protein is highly flexible, exists as an ensemble of rapidly changing conformations, and, thus, can be expected to be easily penetrable for solvent (water) molecules. A few studies provided initial clues that the Stokes-Einstein model is a good approximation for describing the translational diffusion of unfolded and intrinsically disordered proteins in dilute solution.

For example, the evidence comes from the analysis of temperature dependences of protein diffusion coefficients acquired above the temperature-induced unfolding transition or for disordered proteins.[65,66,72,73,76,77] According to Eq. (1), the temperature dependence T/η of D is expected to follow the Arrhenius relation with the slope reflecting the activation energy of the self-diffusion of water (5.0 kcal/mol).[78,79] Fig. 1A exemplifies the temperature dependence of the diffusion coefficient of ribonuclease A acquired in the 3.2% D_2O solution at pH 2.5 (wt%).[65] Fig. 1B shows the temperature dependence of the diffusion coefficient of the

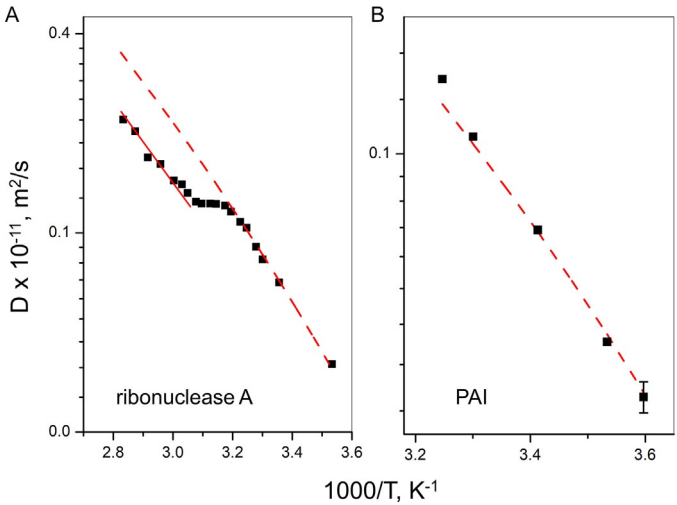

Fig. 1 Temperature dependencies of diffusion coefficients of ribonuclease A[63] and the primary DNA-recognition subdomain of the Sleeping Beauty transposase PAI.[71] Dashed lines in panels A and B and solid line in panel A show the slope of the temperature dependence T/η, calculated from water viscosity and shifted vertically for easier comparison.

primary DNA-recognition domain of the Sleeping Beauty transposase, PAI, acquired in 25 mM sodium phosphate buffer prepared using 100% D_2O at pH 5.0,[73] where the protein is unfolded at all temperatures. Lines represent the slope of the temperature dependence T/η, calculated from water viscosity. The slope of the temperature dependence of the diffusion coefficient is close (5.8 kcal/mol), albeit not exactly the same, to that of the temperature dependence T/η, suggesting that hydrodynamically the behavior of these proteins is close to a solid particle and justifying the use of the Stokes-Einstein model to describe unfolded and disordered proteins in dilute solutions. Similar data (not shown) were obtained for lysozyme[65] and for the HIV-1 nucleocapsid NCp7 protein.[66] Furthermore, the comparison of experimental diffusion coefficients of the antifreeze glycoprotein AFGP8 with the hydrodynamic calculations using the viscosity of water showed good agreement, also suggesting that it behaves like a well-structured globular protein despite that it does not assume a particular secondary structure and is segmentally flexible.[72]

Additionally, in dilute solution, the diffusion coefficient of α-casein shows the same trend as the theoretically-derived concentration dependence of the diffusion coefficient of rigid Brownian spheres,[80] suggesting that the

molecule of α-casein remains in compact state, which is hydrodynamically comparable to that of the rigid Brownian particle, even though the molecule of α-casein does not form a stable structure.[81] Although fibrinogen is not a fully intrinsically disordered molecule, it has highly flexible and largely unstructured ~390 amino acid residue long extensions.[76] The comparison of experimental data to the theoretical curve for rigid Brownian particles shows that in fibrinogen solution the regime of dilute solution is observed only up to ~0.2% (wt%).[76]

While the Stokes-Einstein approximation seems to describe the translational diffusion of unfolded or IDPs in dilute solutions quite satisfactory, the protein can be viewed as a charged heteropolymer. The unfolded or disordered protein represent an extreme case, where the "polymer" (e.g., flexible) nature of a protein is likely to manifest itself the most. In fact, Tanford et al.[82] measured the intrinsic viscosities and sedimentation coefficients of various protein polypeptide chains in concentrated GuHCl solutions and demonstrated that both properties depend on molecular weight exactly as predicted for randomly coiled linear polymer chains. The dimensions of randomly coiled polymers are strongly affected by polymer-solvent interactions, and random coils are typically attained when polymer molecules are dissolved in a good or indifferent solvent, in which the attractive forces between the polymer segments and the solvent are stronger than, or at least as strong as, the attractive forces between one polymer segment and another. It is inherently difficult to find such a solvent for protein molecules consisting of amino acids with a diverse variety of physicochemical properties.[83] However, the hydrodynamic radius estimate for a protein represented by a random coil would provide an upper limit boundary for protein size.

In the theory developed by Kirkwood and Riseman,[84] the flexible polymer is approximated as a linear chain of N segments. The center of gravity of the polymer molecule moves relative to the solvent in which it is immersed. The mean force exerted by the fluid on the molecular center of gravity is the sum of the mean forces exerted by individual monomers, calculated by using the Oseen tensor averaged over all internal configurations of the molecule. The friction coefficient of the polymer molecule is given by:

$$f = \zeta N \left[1 + \frac{N\zeta}{6\pi\eta R_H} \right]^{-1} \tag{3}$$

where the hydrodynamic radius of the polymer molecule calculated according to $R_H^{-1} = \frac{1}{N^2} \sum_{i \neq j} \left\langle r_{ij}^{-1} \right\rangle$ with the mean separation $\langle r_{ij} \rangle$ between

monomers i and j, and ζ is the friction coefficient of monomer. The diffusion coefficient can be calculated in accordance to the following equation:

$$D = \frac{kT}{N\zeta}\left[1 + \frac{N\zeta}{6\pi\eta R_H}\right] \quad (4)$$

It follows from the Eq. (4) that for $N \gg 1$ the diffusion of a polymer in dilute solutions is determined by hydrodynamic interactions, because $R_H \sim N^{\alpha}$, where α changes from 0.5 to 0.6 when the quality of solvent changes from theta (indifferent) to good.

For random coiled ideal polymer chains, the end-to-end distance obeys a Gaussian distribution to a good approximation. The size of an ideal random coil can be characterized by the Flory[85] radius $R_F = bN_k^{1/2}$. Here, the idealized polymer chain of N_k freely-joined Kuhn segments is considered, where each Kuhn segment consists of several (typically 6–8) monomers and the length of the Kuhn segment b typically ranges from 0.5 to 2. For the real polymer chain, when the excluded volume effect is taken into account, the end-to-end distance scales with the number of monomers as N^{ν}, where ν is the Flory exponent that depends on the solvent quality.[85] In good, theta (indifferent), and poor solvents the values of ν for long ($N \gg 1$) homopolymers are 3/5, 1/2, and 1/3, respectively. In human proteins, the median number of amino acid residues is 375,[86] so that the deviation from long chain behavior may be expected.[87] Several empirical expressions relating the hydrodynamic radius of a protein to the number of amino acid residues comprising it were proposed based on the analysis of experimental data.[45,88–91] They are summarized in Table 1. Several trends can be noticed. First, the difference between well-folded, globular and unfolded or intrinsically disordered proteins is revealed in the scaling factor ν. For well-structured globular proteins, it ranges from 0.285 (Ref. 89) to 0.382 (Ref. 91), and is close to the value 0.333 expected for a homopolymer in a poor solvent, whereas for unfolded proteins or IDPs the value of ν is greater. Second, there is a difference between protein unfolded by adding denaturing agents (urea or GuHCl) and IDPs.

The value of ν reported for chemically denatured proteins is the largest and ranges from 0.543 (Ref. 90) to 0.57 (Ref. 45), close to the predicted value of 0.6 for homopolymers in good solvents. In the case of IDPs, reported ν values range from 0.492 (Ref. 91) to 0.509 (Ref. 89), and correspond to the value of ν for homopolymers in theta (indifferent) solvents. Although the experimental dataset is limited, empirically established relations suggest that polymer theory can be applied to predict the hydrodynamic radius of chemically unfolded proteins and IDPs.

Table 1 Empirical expressions relating the hydrodynamic radius of a protein R_H to the number of amino acid residues N for different states.

Empirical relation[a]	Method	References
$R_H^F = 4.92N^{0.285}$ $R_H^D = 2.33N^{0.549}$ $R_N^{ID} = 2.49N^{0.509}$	PFG NMR, SEC	Marsh et al.[89]
$R_H^F = (3.35 \pm 1.04)N^{0.358\pm0.005}$ $R_H^D = (2.43 \pm 1.03)N^{0.543\pm0.007}$ $R_H^{ID} = (2.84 \pm 1.04)N^{0.493\pm0.008}$ $R_H^{MG} = (4.26 \pm 1.12)N^{0.334\pm0.021}$ $R_H^{PMG} = (3.86 \pm 1.06)N^{0.402\pm0.012}$	Viscometry, SEC, DLS	Uversky et al.[90]
$R_H^F = 4.75N^{0.29}$ $R_H^D = 2.21N^{0.57}$	PFG NMR	Wilkins et al.[45]
$R_H^F = 3.405N^{0.382}$ $R_H^{ID} = 3.128N^{0.492}$	PFG NMR	Dudás et al.[91]
$R_H = 3.57N^{0.415}$	PFG NMR (short unstructured peptides)	Danielsson et al.[88]

[a] *F*—folded, *D*—chemically denatured, *ID*—intrinsically disordered, *MG*—molten globule, and *PMG*—pre-molten globule.

3. Concentrated solutions

Even though studies in dilute solutions contribute to our general understanding of the translational diffusion of proteins, many proteins function inside cells and tissues, e.g., in extremely crowded environments, where they have to move to the site of action (the primary mechanism is the translational diffusion). In particular, IDPs are of interest because they are commonly found in cellular compartments and regions with very high local concentrations of proteins, DNA, and RNA.[92–95] A few studies highlighted the difference of translational diffusion of disordered and globular proteins.[54,60,68,70,73] However, a qualitative and quantitative description of the translational diffusion of unfolded proteins and IDPs in crowded environments is far from complete.[96,97] In the discussion below, we distinguish self-crowding conditions, e.g., a concentrated solution of a single type of protein (self-crowding), and crowding conditions created by the addition of different type of molecules (crowders). In the latter case, the concentration of the protein of interest may be low, however, the approximation of the dilute solution and laminar flow around the protein molecule is not valid because the size of the protein may be comparable to the size

of a crowder molecule and the interactions with the crowder are not negligible. We further note that below we discuss only IDPs, because to the best of our knowledge, there is no data in the literature on the diffusion of proteins unfolded under denaturing conditions at high protein concentrations.

3.1 Self-crowding

While in dilute solutions, the translational diffusion of an unfolded or intrinsically disordered protein and the pervaded solvent can be modeled by approximating the protein by a compact solid particle, it is not evident that it should retain this similarity in the crowded milieu, where, in contrast to globular proteins, the shape and compactness of unfolded or intrinsically disordered proteins may change more readily. Furthermore, because the amino acid composition of IDPs is deficient in hydrophobic residues and rich in charged and polar residues, they may have similarities to polyampholytes,[10] for which, due to the long-range of the electrostatic potential, the electrostatic interactions between macromolecules may significantly affect their translational diffusion.[98] In particular, polyampholytes demonstrate a strong tendency for self-association driven by the reduction of electrostatic energy of individual chains.[99]

The conditions of self-crowding are the simplest conditions for studying the diffusion of proteins in concentrated solutions. At such conditions, the experiments were carried out for a few intrinsically disordered proteins or proteins with extended disordered regions. These experiments showed that the diffusion coefficient significantly decreases with increasing protein concentration.[76,81,100–102] Interestingly, the comparison of the concentration dependence of the diffusion coefficient of an intrinsically disordered protein α-casein to master curves describing the concentration dependence of the diffusion coefficient of globular proteins[103] (black symbols) and flexible polymers[51] (cyan line), and to the theoretically-derived concentration dependence for the diffusion coefficient of rigid Brownian spheres[80] (blue line) (Fig. 2) revealed that it differs from all of them.[81] To enable the comparison, the diffusion coefficients of α-casein shown in Fig. 2 were normalized at each concentration by the diffusion coefficient D_0 extrapolated to zero protein concentration to exclude the dependence on temperature and by the critical volume fraction $\hat{\varphi}$, determined from the intersection of the asymptote $(\varphi)^0$ (infinite dilution, no interactions) and the tangent $(\varphi)^{-3}$ drawn to experimental dependence.[51,81] Note that the division of

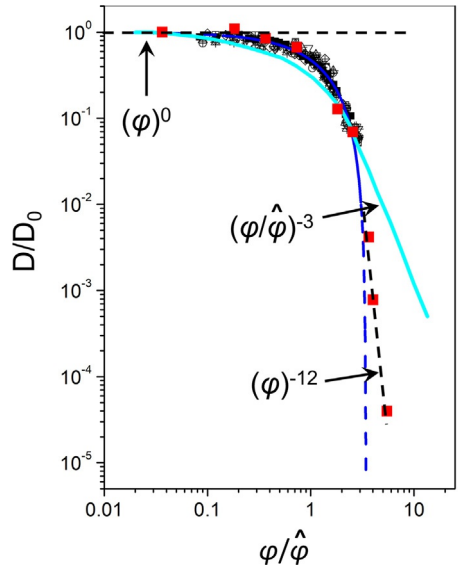

Fig. 2 Concentration dependence of the diffusion coefficient of α-casein (red squares),[79] the master curve representing the concentration dependence of the diffusion coefficient for globular proteins (black symbols)[101] and for flexible polymers (light-blue line),[49] and the theoretical concentration dependence for Brownian rigid spheres (blue line).[78] The asymptote with zero slope is shown by black dashed line. All curves are normalized by the procedure that merely leads to the shift of the concentration dependence along the horizontal and vertical axes.[49,79]

each value of D by D_0 and of the volume fraction φ by the factor $\hat{\varphi}$ merely shifts the whole curve, without altering its shape, along vertical or horizontal axes.

In stark contrast to globular proteins, for which the high concentration region of the concentration dependence of the diffusion coefficient is described by the tangent with the slope of $(\varphi)^{-3}$, the concentration dependence of the diffusion coefficient of α-casein can be approximated by the straight line with the slope of $(\varphi)^{-12}$ at the protein concentrations >5% (wt%). In the case of globular proteins, the master curve[103] describing the concentration dependence of their diffusion coefficients coincides in the range of volume fractions from 0.015 to 0.4 (the measurements at higher concentrations were limited by protein solubility) with the theoretically-derived concentration dependence of the diffusion coefficient of rigid Brownian spheres[80,104] and shows the same sharp crossover between the concentration-independent regime of dilute solution, where the interactions between protein molecules are negligible, to the regime of the

concentrated solution where protein molecules interact (Fig. 2). Thus, common dynamic characteristics of globular proteins remain close to rigid Brownian particles in concentrated solution (at least, up to the volume fraction of 0.4).[103,104] The diffusion coefficient of α-casein shows the same trend as globular proteins or rigid Brownian spheres up to the volume fraction of about 0.05, but deviates from both dependences at higher concentrations, demonstrating a stronger concentration dependence than globular proteins, a weaker dependence than rigid Brownian spheres, and, thus, a qualitatively different behavior than globular proteins or rigid Brownian spheres. Furthermore, the concentration dependence of α-casein diffusion coefficient differs from the master curve established for flexible polymers, except for the asymptotic behavior $(\varphi)^0$ in dilute solution. The universal concentration dependence for polymers shows a gradual transition from dilute to concentrated solution, where the asymptote $(\varphi/\hat{\varphi})^{-3}$ corresponds to a qualitatively different diffusion regime of entangled polymer chains, and the movement of the polymer chain as a whole results from the movement of its individual segments.[105–108] In contrast, the diffusion coefficient of α-casein shows a weaker dependence on concentration in the transition region of the curve and a significantly stronger dependence on concentration beyond that, indicating that the diffusion of α-casein is also qualitatively different from that of a flexible polymer. Note that as with flexible polymers, averaged diffusion coefficients of α-casein are discussed after normalizing them for the local dynamics determined from spin-spin and spin-lattice relaxation.

The strong dependence of α-casein or fibrinogen diffusion coefficient on concentration was attributed to their self-association.[76,81] Based on the analysis of the dependence of the shape of the diffusion attenuation on the diffusion time t_d it was established that at the concentration of α-casein 15% (wt%) only 10% of α-casein molecules diffused freely in solution with t_d-independent diffusion coefficient.[81] Remaining 90% of α-casein molecules showed anomalous diffusion with $D \sim t_d^{-1}$, typically observed for physical or chemical gels.[109–111] Furthermore, from the t_d-dependence of the shape of the diffusion attenuation the life time of α-casein molecules in gel was estimated. Thus, at the concentration of α-casein 15%, ~90% of α-casein molecules formed a labile gel-like network stabilized by noncovalent intermolecular interactions with the lifetime of α-casein molecules within the network of ~3.5 s.[81] Similarly, fibrinogen molecules formed noncovalent linkages through highly extendable unstructured αC regions and coupled into a dynamic network.[76] Given the limited information on the translational diffusion of IDPs in concentrated solutions and the

amino acid sequence variability between them,[8,31,33,112,113] it is premature to draw unifying conclusions for all IDPs. However, based on three cases discussed here, we propose that self-association of IDPs into some kind of supramolecular structures may be a common dominant feature affecting their translational diffusion in the concentrated solution at self-crowding conditions. First, it follows from the similarity of IDPs to polyampholytes that tend to self-associate to neutralize the charge and to decrease the electrostatic energy of individual molecules and demonstrate various phase-separation behaviors.[7,99,114] Second, it would be in line with a well-documented strong tendency of IDPs to self-associate or to engage in polyvalent binding events (for review, see Refs. 115, 116 and references therein).

3.2 Diffusion in the presence of crowders

As with self-crowding, the diffusion of an IDP significantly slows down in the presence of other type crowder molecules.[117,118] The diffusion measurements carried out on intracellular α-synuclein expressed in *E. coli* cells or purified and added separately to the extracellular medium showed the reduction of the diffusion coefficient by almost three orders of magnitude due to both the crowding effect and the effect of restricted diffusion inside the cell.[118]

The diffusion of intrinsically disordered proteins in crowded environment depends on the type of crowder and may demonstrate a non-trivial behavior. For example, Wang et al.[117] discovered that α-synuclein, an IDP with molecular weight of 14 kDa, diffuses slower than a globular protein of a smaller size, chymotrypsin inhibitor 2 (CI2, 7.4 kDa) in simple solvents or glycerol, as expected in accordance with the Stoke-Einstein formula (Eq. 1). However, when the crowding is created by the addition of large-size molecules such as polyvinylpyrrolidone (PVP, 40 kDa), Ficoll (70 kDa), lysozyme (15 kDa), and bovine serum albumin (BSA, 67 kDa), unexpectedly, α-synuclein diffuses faster than CI2 despite the fact that α-synuclein has a larger molecular weight. The result did not arise from a structural change under crowded conditions because CI2 remained compact and α-synuclein remained collapsed, hence, the authors proposed that α-synuclein adopted a different diffusion strategy related to the inherent internal motion of disordered proteins.[117] A crossover behavior of diffusion coefficient was also experimentally observed for IDPs MAP2c and $p21^{Cip1}$ relative to a globular protein RNase T1 by using fluorescence correlation

spectroscopy (FCS).[119] However, their relatively faster diffusion was associated with a gradual shift of the conformational ensemble toward more compact states without actual cooperative structural transition to a well-folded globule.

Wang et al.[120] proposed a disorder plus collapse model to analyze the diffusion coefficient of an IDP in semi-dilute polymer solution. In this model, the IDP is represented by an ensemble of interchanging conformations with a broad range of gyration radii. The mean gyration radius, describing the conformational ensemble, decreases due to compaction in the presence of crowding. The average diffusion coefficient is then calculated using scaling relation obtained for the diffusion coefficient of the probe molecule in polymer solution:[121]

$$D = D_0 exp\left[-\frac{\gamma}{RT}\left(\frac{\mathrm{R}_{eff}}{\xi}\right)^{\beta}\right]$$
$$R_{eff}^{-2} = r_h^{-2} + R_H^{-2}; \langle R_g \rangle = \left\langle R_g^0 \right\rangle \left(\frac{1}{1+\lambda\phi/\phi^*}\right)^{1/5} \quad (5)$$

Here D_0 is given by Eq. (1); β, γ, and λ are parameters depending on the system (polymer and protein), r_h^{-2} is the hydrodynamic radius of the polymer molecules, ξ is the correlation length of the polymer solution, $\langle R_g^0 \rangle$ is the mean gyration radius of the IDP in pure solvent in the absence of crowding, and ϕ^* is the overlap concentration for polymer molecules at the crossover from the dilute to semi-dilute regime. Applying the disorder plus collapse model, Wang et al.[120] were able to show the crossover behavior with IDP diffusing slower than the globular protein of smaller size in simple solutions and faster under crowded conditions, and to achieve a good agreement with experimental data.

Several other factors contribute to the change of the diffusion coefficient of a protein in the crowded environment, including the excluded volume, interactions between crowder and protein molecules, interactions (association) of protein molecules, and steric obstruction from crowders (e.g., cell or organelle membranes).[122–127] While the effect of excluded volume is the simplest to account for and, hence, is the most studied effect,[128,129] the role of other factors becomes increasingly recognized as reviewed by Kuznetsova et al.[130] This area remains largely unexplored by PFG NMR, however, the diffusion measurements are likely to provide a novel insight into the behavior of proteins in crowded conditions.

4. Conclusions and future directions

Given the abundance of intrinsically disordered proteins, the broad assortment of their physicochemical properties, and the limited literature on their translational diffusion, it is currently difficult to deduce a unifying theme. Therefore, translational diffusion data on IDPs in general are still in high demand. For the development of theoretical models, the diffusion data on IDPs in dilute solutions would be useful. For example, the investigation of the temperature dependence of protein diffusion coefficient would be necessary to verify whether the translational diffusion of an IDP can be satisfactory described by the Arrhenius dependence with the slope reflecting the activation energy of the self-diffusion of solvent (water). In this regard, the diffusion of a protein unfolded using denaturing agents (urea, GuHCl) can provide an additional insight.

For understanding biologically relevant conditions, the translational diffusion of IDPs (and proteins in general) requires thorough investigation both under the conditions of self-crowding and crowding by dissimilar molecules. The Stokes-Einstein equation is derived for a sphere diffusing in a continuum fluid of viscosity η. In crowded solutions, diffusing proteins experience different micro viscosities, dependent on the concentration of molecules[131] and the size ratio of protein and macromolecular co-solute.[132,133] Furthermore, it is likely that the mean square displacement does not follow the simple relation $\langle r^2 \rangle = 6Dt_d$.[134] Interestingly, the crowding may affect the translational and rotational diffusion of proteins differently. For a set of globular proteins, it was shown that translational diffusion scales with macroviscosity, whereas the rotational diffusion depends on microviscosity and reveals protein-specific behavior.[135]

The conditions of self-crowding provide a simplified model for studying the effect of protein-protein interactions on their translational diffusion. At self-crowding, it is interesting to establish whether a common law exists that can describe the translational diffusion of IDPs as with globular proteins[103] or flexible polymers.[51] If found, the practical significance of the result will be the possibility to predict the behavior of any IDP protein in dilute or concentrated solution when the direct measurement of the diffusion coefficient is not possible and to facilitate the development of theoretical models. It would also be interesting to establish the extent of similarities between synthetic polymer polyampholytes and intrinsically disordered proteins. When

established, some concepts developed for translational diffusion in solutions of synthetic polyampholytes might find applicability in the field of IDPs.

In the case of crowding created by dissimilar molecules, the investigation of translational diffusion of an IDP using different types of crowders (size, charge, etc.) as model systems could improve our understanding of protein behavior inside the cell[123] and the complex relation between reaction rates, diffusion coefficients, and protein concentrations.

Finally, PFG NMR is particularly useful for studying the translational diffusion of proteins, including IDPs, under crowding conditions. However, it is desirable to use large pulse-filed gradients to observe the whole spectrum of diffusion coefficients that are significantly slowed down in crowded milieu, to utilize specific labeling of the protein of interest to discriminate its signal from other molecules in solution; and to analyze the diffusion coefficient as a function of diffusion time. The latter will provide the information on the supramolecular organization in the solution, on whether the protein undergoes the restricted diffusion, and on possible exchange processes related to conformational changes of the protein or fluctuations in its surroundings.

References

1. Anfinsen CB. Principles that govern the folding of protein chains. *Science*. 1973;181:223–230.
2. Kauzmann W. Some factors in the interpretation of protein denaturation11the preparation of this article has been assisted by a grant from the National Science Foundation. In: Anfinsen CB, Anson ML, Bailey K, Edsall JT, eds. *Advances in Protein Chemistry*. Academic Press; 1959:1–63. vol. 14.
3. Pauling L, Corey RB, Branson HR. The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc Natl Acad Sci U S A*. 1951;37:205–211.
4. Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature*. 1958;181:662–666.
5. Berman HM, Westbrook J, Feng Z, et al. The protein data bank. *Nucleic Acids Res*. 2000;28:235–242.
6. Williamson MP, Havel TF, Wuthrich K. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by 1H nuclear magnetic resonance and distance geometry. *J Mol Biol*. 1985;182:295–315.
7. Dobrynin AV, Colby RH, Rubinstein M. Polyampholytes. *J Polym Sci Polym Phys*. 2004;42:3513–3538.
8. Uversky VN, Gillespie JR, Fink AL. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins*. 2000;41:415–427.
9. Weathers EA, Paulaitis ME, Woolf TB, Hoh JH. Reduced amino acid alphabet is sufficient to accurately recognize intrinsically disordered protein. *FEBS Lett*. 2004;576:348–352.

10. Mao AH, Crick SL, Vitalis A, Chicoine CL, Pappu RV. Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc Natl Acad Sci U S A*. 2010;107:8183–8188.
11. Theillet FX, Binolfi A, Frembgen-Kesner T, et al. Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs). *Chem Rev*. 2014;114:6661–6714.
12. Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol*. 2015;16:18–29.
13. Oldfield CJ, Dunker AK. Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu Rev Biochem*. 2014;83:553–584.
14. Tompa P. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem Sci*. 2012;37:509–516.
15. Uversky VN, Oldfield CJ, Dunker AK. Intrinsically disordered proteins in human diseases: introducing the D(2) concept. *Annu Rev Biophys*. 2008;37:215–246.
16. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z. Intrinsic disorder and protein function. *Biochemistry*. 2002;41:6573–6582.
17. Liu ZR, Huang YQ. Advantages of proteins being disordered. *Protein Sci*. 2014;23:539–550.
18. Tantos A, Han KH, Tompa P. Intrinsic disorder in cell signaling and gene transcription. *Mol Cell Endocrinol*. 2012;348:457–465.
19. Uversky VN. Functional roles of transiently and intrinsically disordered regions within proteins. *FEBS J*. 2015;282:1182–1189.
20. Babu MM, van der Lee R, de Groot NS, Gsponer J. Intrinsically disordered proteins: regulation and disease. *Curr Opin Struct Biol*. 2011;21:432–440.
21. Uversky VN, Oldfield CJ, Midic U, et al. Unfoldomics of human diseases: linking protein intrinsic disorder with diseases. *BMC Genomics*. 2009;10(Suppl 1):S7.
22. Vacic V, Markwick PR, Oldfield CJ, et al. Disease-associated mutations disrupt functionally important regions of intrinsic protein disorder. *PLoS Comput Biol*. 2012;8:e1002709.
23. Oates ME, Romero P, Ishida T, et al. D(2)P(2): Database of disordered protein predictions. *Nucleic Acids Res*. 2013;41:D508–D516.
24. van der Lee R, Buljan M, Lang B, et al. Classification of intrinsically disordered regions and proteins. *Chem Rev*. 2014;114:6589–6631.
25. Timasheff SN. Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components. *Proc Natl Acad Sci U S A*. 2002;99:9721–9726.
26. Whitney PL, Tanford C. Solubility of amino acids in aqueous urea solutions and its implications for the denaturation of proteins by urea. *J Biol Chem*. 1962;237:PC1735–PC1737.
27. Nozaki Y, Tanford C. The solubility of amino acids, diglycine, and triglycine in aqueous guanidine hydrochloride solutions. *J Biol Chem*. 1970;245:1648–1652.
28. Arakawa T, Timasheff SN. The stabilization of proteins by osmolytes. *Biophys J*. 1985;47:411–414.
29. Yancey PH, Somero GN. Counteraction of urea destabilization of protein structure by methylamine osmoregulatory compounds of elasmobranch fishes. *Biochem J*. 1979;183:317–323.
30. Ptitsyn OB, Uversky VN. The molten globule is a third thermodynamical state of protein molecules. *FEBS Lett*. 1994;341:15–18.
31. Uversky VN. Diversity of equilibrium compact forms of denatured globular proteins. *Protein Pept Lett*. 1997;4:355–367.
32. Uversky VN. Unusual biophysics of intrinsically disordered proteins. *Biochim Biophys Acta*. 2013;1834:932–951.

33. Holehouse AS, Pappu RV. Collapse transitions of proteins and the interplay among backbone, sidechain, and solvent interactions. *Annu Rev Biophys*. 2018; https://doi.org/10.1146/annurev-biophys-070317-032838. [Epub ahead of print].
34. Finkelstein AV, Shakhnovich EI. Theory of cooperative transitions in protein molecules. II. Phase diagram for a protein molecule in solution. *Biopolymers*. 1989;28:1681–1694.
35. Shakhnovich EI, Finkelstein AV. Theory of cooperative transitions in protein molecules. I. Why denaturation of globular protein is a first-order phase transition. *Biopolymers*. 1989;28:1667–1680.
36. Bernadó P, Blackledge M. A self-consistent description of the conformational behavior of chemically denatured proteins from NMR and small angle scattering. *Biophys J*. 2009;97:2839–2845.
37. Garcia de la Torre JG. Hydrodynamics of segmentally flexible macromolecules. *Eur Biophys J*. 1994;23:307–322.
38. Mao AH, Lyle N, Pappu RV. Describing sequence-ensemble relationships for intrinsically disordered proteins. *Biochem J*. 2013;449:307–318.
39. Bernado P, Mylonas E, Petoukhov MV, Blackledge M, Svergun DI. Structural characterization of flexible proteins using small-angle X-ray scattering. *J Am Chem Soc*. 2007;129:5656–5664.
40. Fitzkee NC, Rose GD. Reassessing random-coil statistics in unfolded proteins. *Proc Natl Acad Sci U S A*. 2004;101:12497–12502.
41. Kohn JE, Millett IS, Jacob J, et al. Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc Natl Acad Sci U S A*. 2004;101:12491–12496.
42. Wang YW, Teraoka I, Hansen FY, Peters GH, Hassager O. A theoretical study of the separation principle in size exclusion chromatography. *Macromolecules*. 2010;43:1651–1659.
43. Uversky VN. Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule. *Biochemistry*. 1993;32:13288–13298.
44. Nettels D, Muller-Spath S, Kuster F, et al. Single-molecule spectroscopy of the temperature-induced collapse of unfolded proteins. *Proc Natl Acad Sci U S A*. 2009;106:20740–20745.
45. Wilkins DK, Grimshaw SB, Receveur V, Dobson CM, Jones JA, Smith LJ. Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry*. 1999;38:16424–16431.
46. Einstein A, Fürth R. *Investigations on the Theory of Brownian Movement*. New York, N.Y.: Dover Publications; 1956.
47. Cantor CR, Schimmel PR. *Biophysical chemistry*. San Francisco: W. H. Freeman; 1980. p v. <1>.
48. Perrin F. Brownian movement of an ellipsoid (II).—Free rotation and depolarisation of fluourescences.—Translation and diffusion of ellipsoidal molecules. *J Phys Paris*. 1936;7:1–11.
49. Bloomfield V, Van Holde KE, Dalton WO. Frictional coefficients of multisubunit structures. II. Application to proteins and viruses. *Biopolymers*. 1967;5:149–159.
50. Garcia de la Torre JG, Bloomfield VA. Hydrodynamic properties of complex, rigid, biological macromolecules: theory and applications. *Q Rev Biophys*. 1981;14:81–139.
51. Skirda VD, Sundukov VI, Maklakov AI, Zgadzai OE, Gafurov IR, Vasiljev GI. On the generalized concentration and molecular mass dependencies of macromolecular self-diffusion in polymer-solutions. *Polymer*. 1988;29:1294–1300.
52. Colmenero J. Polymer chain diffusion in polymer blends. A theoretical interpretation based on a memory function formalism. *J Polym Sci B*. 2019. https://doi.org/10.1002/polb.24811.

53. Price WS, Tsuchiya F, Suzuki C, Arata Y. Characterization of the solution properties of Pichia farinosa killer toxin using PGSE NMR diffusion measurements. *J Biomol NMR*. 1999;13:113–117.
54. Jones JA, Wilkins DK, Smith LJ, Dobson CM. Characterisation of protein unfolding by NMR diffusion measurements. *J Biomol NMR*. 1997;10:199–203.
55. Chen L, Hodgson KO, Doniach S. A lysozyme folding intermediate revealed by solution X-ray scattering. *J Mol Biol*. 1996;261:658–671.
56. Morgan CJ, Wilkins DK, Smith LJ, Kawata Y, Dobson CM. A compact monomeric intermediate identified by NMR in the denaturation of dimeric triose phosphate isomerase. *J Mol Biol*. 2000;300:11–16.
57. Bhattacharjya S, Xu P, Xiang H, Chretien M, Seidah NG, Ni F. pH-induced conformational transitions of a molten-globule-like state of the inhibitory prodomain of furin: implications for zymogen activation. *Protein Sci*. 2001;10: 934–942.
58. Morar AS, Olteanu A, Young GB, Pielak GJ. Solvent-induced collapse of alpha-synuclein and acid-denatured cytochrome c. *Protein Sci*. 2001;10:2195–2199.
59. Li H, Frieden C. Fluorine-19 NMR studies on the acid state of the intestinal fatty acid binding protein. *Biochemistry*. 2006;45:6272–6278.
60. Pan H, Barany G, Woodward G. Reduced BPTI is collapsed. A pulsed field gradient NMR study of unfolded and partially folded bovine pancreatic trypsin inhibitor. *Protein Sci*. 1997;6:1985–1992.
61. Li Y, Shan B, Raleigh DP. The cold denatured state is compact but expands at low temperatures: hydrodynamic properties of the cold denatured state of the C-terminal domain of L9. *J Mol Biol*. 2007;368:256–262.
62. Balbach J. Compaction during protein folding studied by real-time NMR diffusion experiments. *J Am Chem Soc*. 2000;122:5887–5888.
63. Weljie AM, Yamniuk AP, Yoshino H, Izumi Y, Vogel HJ. Protein conformational changes studied by diffusion NMR spectroscopy: application to helix-loop-helix calcium binding proteins. *Protein Sci*. 2003;12:228–236.
64. Chatterjee C, Martinez D, Gerig JT. Interactions of trifluoroethanol with [val5]angiotensin II. *J Phys Chem B*. 2007;111:9355–9362.
65. Molchanov S, Faizullin DA, Nesmelova IV. Theoretical and experimental investigation of the translational diffusion of proteins in the vicinity of temperature-induced unfolding transition. *J Phys Chem B*. 2016;120:10192–10198.
66. Ramboarina S, Srividya N, Atkinson RA, et al. Effects of temperature on the dynamic behaviour of the HIV-1 nucleocapsid NCp7 and its DNA complex. *J Mol Biol*. 2002;316:611–627.
67. Lane AN, Hays LM, Tsvetkova N, Feeney RE, Crowe LM, Crowe JH. Comparison of the solution conformation and dynamics of antifreeze glycoproteins from Antarctic fish. *Biophys J*. 2000;78:3195–3207.
68. Penkett CJ, Redfield C, Jones JA, et al. Structural and dynamical characterization of a biologically active unfolded fibronectin-binding protein from Staphylococcus aureus. *Biochemistry*. 1998;37:17054–17067.
69. Anil B, Li Y, Cho JH, Raleigh DP. The unfolded state of NTL9 is compact in the absence of denaturant. *Biochemistry*. 2006;45:10110–10116.
70. Dehner A, Kessler H. Diffusion NMR spectroscopy: folding and aggregation of domains in p53. *Chembiochem*. 2005;6:1550–1565.
71. Krishnan VV, Lau EY, Yamada J, et al. Intramolecular cohesion of coils mediated by phenylalanine—glycine motifs in the natively unfolded domain of a nucleoporin. *PLoS Comput Biol*. 2008;4:e1000145.
72. Krishnan VV, Fink WH, Feeney RE, Yeh Y. Translational dynamics of antifreeze glycoprotein in supercooled water. *Biophys Chem*. 2004;110:223–230.

73. Leighton GO, Konnova TA, Idiyatullin B, Hurr SH, Zuev YF, Nesmelova IV. The folding of the specific DNA recognition subdomain of the sleeping beauty transposase is temperature-dependent and is required for its binding to the transposon DNA. *PLoS One*. 2014;9:e112114.
74. Zhang X, Perugini MA, Yao S, et al. Solution conformation, backbone dynamics and lipid interactions of the intrinsically unstructured malaria surface protein MSP2. *J Mol Biol*. 2008;379:105–121.
75. Berry GC. The hydrodynamic and conformational properties of denatured proteins in dilute solutions. *Protein Sci*. 2010;19:94–98.
76. Zuev YF, Litvinov RI, Sitnitsky AE, et al. Conformational flexibility and self-association of fibrinogen in concentrated solutions. *J Phys Chem B*. 2017;121: 7833–7843.
77. Yuwen T, Sekhar A, Baldwin AJ, Vallurupalli P, Kay LE. Measuring diffusion constants of invisible protein conformers by triple-quantum (1) H CPMG relaxation dispersion. *Angew Chem Int Ed Engl*. 2018;57:16777–16780.
78. Mccall DW, Douglass DC, Anderson EW. Diffusion in Liquids. *J Chem Phys*. 1959;31:1555–1557.
79. Mills R. Self-diffusion in normal and heavy-water in range 1-45 degrees. *J Phys Chem*. 1973;77:685–688.
80. Tokuyama M, Oppenheim I. Dynamics of hard-sphere suspensions. *Phys Rev E*. 1994;50:R16–R19.
81. Melnikova DL, Skirda VD, Nesmelova IV. Effect of intrinsic disorder and self-association on the translational diffusion of proteins: the case of alpha-casein. *J Phys Chem B*. 2017;121:2980–2988.
82. Tanford C, Kawahara K, Lapanje S. Proteins as random coils. I. Intrinsic viscosities and sedimentation coefficients in concentrated guanidine hydrochloride. *J Am Chem Soc*. 1967;89:729–736.
83. Smith LJ, Fiebig KM, Schwalbe H, Dobson CM. The concept of a random coil. Residual structure in peptides and denatured proteins. *Fold Des*. 1996;1:R95–106.
84. Kirkwood JG, Riseman J. The intrinsic viscosities and diffusion constants of flexible macromolecules in solution. *J Chem Phys*. 1948;16:565–573.
85. Flory PJ. *Statistical Mechanics of Chain Molecules*. Interscience Publishers; 1969.
86. Milo R, Phillips R. *Cell Biology by the Numbers*. Garland Science; 2016.
87. Fuertes G, Banterle N, Ruff KM, et al. Decoupling of size and shape fluctuations in heteropolymeric sequences reconciles discrepancies in SAXS vs. FRET measurements. *Proc Natl Acad Sci U S A*. 2017;114:E6342–E6351.
88. Danielsson J, Jarvet J, Damberg P, Gräslund A. Translational diffusion measured by PFG-NMR on full length and fragments of the Alzheimer Aβ(1–40) peptide. Determination of hydrodynamic radii of random coil peptides of varying length. *Magn Reson Chem*. 2002;40:S89–S97.
89. Marsh JA, Forman-Kay JD. Sequence determinants of compaction in intrinsically disordered proteins. *Biophys J*. 2010;98:2383–2390.
90. Uversky VN, Santambrogio C, Brocca S, Grandori R. Length-dependent compaction of intrinsically disordered proteins. *FEBS Lett*. 2012;586:70–73.
91. Dudas EF, Bodor A. Quantitative, diffusion NMR based analytical tool to distinguish folded, disordered, and denatured biomolecules. *Anal Chem*. 2019;91:4929–4933.
92. Barbar E, Nyarko A. Polybivalency and disordered proteins in ordering macromolecular assemblies. *Semin Cell Dev Biol*. 2015;37:20–25.
93. Rout MP, Aitchison JD, Suprapto A, Hjertaas K, Zhao Y, Chait BT. The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J Cell Biol*. 2000;148:635–651.

94. Han TW, Kato M, Xie S, et al. Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. *Cell*. 2012;149:768–779.
95. Denning DP, Patel SS, Uversky V, Fink AL, Rexach M. Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. *Proc Natl Acad Sci U S A*. 2003;100:2450–2455.
96. Elcock AH. Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. *Curr Opin Struct Biol*. 2010;20:196–206.
97. Uversky VN. A decade and a half of protein intrinsic disorder: biology still waits for physics. *Protein Sci*. 2013;22:693–724.
98. Ando T, Skolnick J. Crowding and hydrodynamic interactions likely dominate in vivo macromolecular motion. *Proc Natl Acad Sci U S A*. 2010;107:18457–18462.
99. Everaers R, Johner A, Joanny JF. Polyampholytes: from single chains to solutions. *Macromolecules*. 1997;30:8478–8498.
100. Melnikova DL, Skirda VD, Nesmelova IV. Effect of reducing agent TCEP on translational diffusion and supramolecular assembly in aqueous solutions of alpha-casein. *J Phys Chem B*. 2019;123:2305–2315.
101. Kusova AM, Sitnitsky AE, Idiyatullin BZ, Bakirova DR, Zuev YF. The effect of shape and concentration on translational diffusion of proteins measured by PFG NMR. *Appl Magn Reson*. 2018;49:35–51.
102. Kusova AM, Sitnitsky AE, Zuev YF. Effect of structural disorder on hydrodynamic behavior of alpha-casein according to PFG NMR spectroscopy. *Appl Magn Reson*. 2018;49:499–509.
103. Nesmelova IV, Skirda VD, Fedotov VD. Generalized concentration dependence of globular protein self-diffusion coefficients in aqueous solutions. *Biopolymers*. 2002;63:132–140.
104. Tokuyama M, Moriki T, Kimura Y. Self-diffusion of biomolecules in solution. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2011;83:051402.
105. de Gennes PG. *Scaling Concepts in Polymer Physics*. Ithaca, N.Y: Cornell University Press; 1979: 324.
106. Doi M, Edwards SF. Dynamics of concentrated polymer systems .2. Molecular-motion under flow. *J Chem Soc Farad Trans 2*. 1978;74:1802–1817.
107. Doi M, Edwards SF. Dynamics of concentrated polymer Systems.1. Brownian-motion in equilibrium state. *J Chem Soc Farad Trans 2*. 1978;74:1789–1801.
108. Doi M, Edwards SF. Dynamics of concentrated polymer Systems.3. Constitutive equation. *J Chem Soc Farad Trans 2*. 1978;74:1818–1832.
109. Gafurov IR, Skirda VD, Maklakov AI, Perevezentseva SP, Zimkin YA. NMR study of the structure of aqueous gelatine gels and the process of their formation. *Polymer Sci USSR* 1989;31:292–300.
110. Gafurov IR, Skirda VD, Maklakov AI, Ryskina II. Self-diffusion and gelation in benzyl alcohol solutions of cellulose triacetate. *Polymer Sci U.S.S.R*. 1988;30:1639–1644.
111. Skirda VD, Gafurov IR, Maklakov AI, Doroginitskii MM, Flyaisher G. Translational mobility of macromolecules in networks. *Vysokomol Soedin B*. 1988;30:313–314.
112. Das RK, Pappu RV. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc Natl Acad Sci U S A*. 2013;110:13392–13397.
113. Uversky VN. Natively unfolded proteins: a point where biology waits for physics. *Protein Sci*. 2002;11:739–756.
114. Dobrynin AV. 1.05—solutions of charged polymers. In: Matyjaszewski K, Möller M, eds. *Polymer Science: A Comprehensive Reference*. Amsterdam: Elsevier; 2012:81–132.
115. Tompa P, Schad E, Tantos A, Kalmar L. Intrinsically disordered proteins: emerging interaction specialists. *Curr Opin Struct Biol*. 2015;35:49–59.

116. Uversky VN. Dancing protein clouds: the strange biology and chaotic physics of intrinsically disordered proteins. *J Biol Chem*. 2016;291:6681–6688.
117. Wang Y, Benton LA, Singh V, Pielak GJ. Disordered protein diffusion under crowded conditions. *J Phys Chem Lett*. 2012;3:2703–2706.
118. Waudby CA, Mantle MD, Cabrita LD, Gladden LF, Dobson CM, Christodoulou J. Rapid distinction of intracellular and extracellular proteins using NMR diffusion measurements. *J Am Chem Soc*. 2012;134:11312–11315.
119. Szasz CS, Alexa A, Toth K, Rakacs M, Langowski J, Tompa P. Protein disorder prevails under crowded conditions. *Biochemistry*. 2011;50:5834–5844.
120. Wang J, Bian YK, Cao XL, Zhao NR. Understanding diffusion of intrinsically disordered proteins in polymer solutions: a disorder plus collapse model. *AIP Adv* 2017;7:115120.
121. Holyst R, Bielejewska A, Szymanski J, et al. Scaling form of viscosity at all length-scales in poly(ethylene glycol) solutions studied by fluorescence correlation spectroscopy and capillary electrophoresis. *Phys Chem Chem Phys*. 2009;11:9025–9032.
122. Nawrocki G, Wang PH, Yu I, Sugita Y, Feig M. Slow-down in diffusion in crowded protein solutions correlates with transient cluster formation. *J Phys Chem B*. 2017;121:11072–11084.
123. Schavemaker PE, Boersma AJ, Poolman B. How important is protein diffusion in prokaryotes? *Front Mol Biosci*. 2018;5:93.
124. Banks DS, Fradin C. Anomalous diffusion of proteins due to molecular crowding. *Biophys J*. 2005;89:2960–2971.
125. Roosen-Runge F, Hennig M, Zhang F, et al. Protein self-diffusion in crowded solutions. *Proc Natl Acad Sci U S A*. 2011;108:11815–11820.
126. Dey P, Bhattacherjee A. Disparity in anomalous diffusion of proteins searching for their target DNA sites in a crowded medium is controlled by the size, shape and mobility of macromolecular crowders. *Soft Matter*. 2019;15:1960–1969.
127. Stefferson MW, Norris SL, Vernerey FJ, Betterton MD, Hough LE. Effects of soft interactions and bound mobility on diffusion in crowded environments: a model of sticky and slippery obstacles. *Phys Biol*. 2017;14:045008.
128. Minton AP. Models for excluded volume interaction between an unfolded protein and rigid macromolecular cosolutes: macromolecular crowding and protein stability revisited. *Biophys J*. 2005;88:971–985.
129. Zimmerman SB, Minton AP. Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annu Rev Biophys Biomol Struct*. 1993;22:27–65.
130. Kuznetsova IM, Zaslavsky BY, Breydo L, Turoverov KK, Uversky VN. Beyond the excluded volume effects: mechanistic complexity of the crowded milieu. *Molecules*. 2015;20:1377–1409.
131. Dix JA, Verkman AS. Crowding effects on diffusion in solutions and cells. *Annu Rev Biophys*. 2008;37:247–263.
132. Lavalette D, Hink MA, Tourbez M, Tetreau C, Visser AJ. Proteins as micro viscosimeters: brownian motion revisited. *Eur Biophys J*. 2006;35:517–522.
133. Lavalette D, Tetreau C, Tourbez M, Blouquit Y. Microscopic viscosity and rotational diffusion of proteins in a macromolecular environment. *Biophys J*. 1999;76:2744–2751.
134. Zhou HX, Rivas G, Minton AP. Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu Rev Biophys* 2008;37:375–397.
135. Roos M, Ott M, Hofmann M, et al. Coupling and decoupling of rotational and translational diffusion of proteins under crowding conditions. *J Am Chem Soc*. 2016;138:10365–10372.