Dynamic cluster formation determines viscosity and diffusion in dense protein solutions

Sören von Bülowa, Marc Siggela, Max Linkea, and Gerhard Hummera,b,1

aDepartment of Theoretical Biophysics, Max Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany; and bDepartment of Physics, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

Edited by Huan-Xiang Zhou, University of Illinois at Chicago, Chicago, IL, and accepted by Editorial Board Member J. A. McCammon April 1, 2019 (received for review October 11, 2018)

We develop a detailed description of protein translational and rotational diffusion in concentrated solution on the basis of all-atom molecular dynamics simulations in explicit solvent. Our systems contain up to 540 fully flexible proteins with 3.6 million atoms. In concentrated protein solutions (100 mg/mL and higher), the proteins ubiquitin and lysozyme, as well as the protein domains third IgG-binding domain of protein G and villin headpiece, diffuse not as isolated particles, but as members of transient clusters between which they constantly exchange. A dynamic cluster model nearly quantitatively explains the increase in viscosity and the decrease in protein diffusivity with protein volume fraction, which both exceed the predictions from widely used colloid models. The Stokes–Einstein relations for translational and rotational diffusion remain valid, but the effective hydrodynamic radius grows linearly with protein volume fraction. This increase follows the observed increase in cluster size and explains the more dramatic slowdown of protein rotation compared with translation. Baxter’s sticky-sphere model of colloidal suspensions captures the concentration dependence of cluster size, viscosity, and rotational and translational diffusion. The consistency between simulations and experiments for a diverse set of soluble globular proteins indicates that the cluster model applies broadly to concentrated protein solutions, with equilibrium dissociation constants for nonspecific protein–protein binding in the $K_g \approx 10^{-11}$ m$^{-1}$ regime.

The interior of cells is a densely crowded medium, in which macromolecular concentrations range from 90 mg/mL in red blood cells to 300 mg/mL in the mitochondrial matrix (1, 2). Macromolecular crowding influences the stability of proteins, reaction rates, the catalytic activity of enzymes, protein–protein association, and diffusion (3–13). Excluded volume, steric repulsion (14) and attractive protein–protein interactions as well as hydrodynamic interactions affect protein diffusion (6, 15–19). To address the influence of specific protein–protein interactions on protein diffusion (20), crowded solutions with proteins serving as both agents and readout have been studied (5, 14, 16, 21–30).

Experimental techniques to study the effects of macromolecular crowding on diffusion (15) include tracer boundary spreading (14), light scattering spectroscopy (31), fluorescence recovery after photobleaching (FRAP) (32–34), electron spin resonance (35), single-particle tracking (36), fluorescence correlation spectroscopy (FCS) (37–39), quasielastic neutron backscattering (27, 40), and NMR spectroscopy (24, 41, 42). Particle-based simulations complement these experiments (15), treating the proteins as spheres or ellipsoids (20, 43, 44), with residue-level coarse graining (45–47), or as rigid all-atom models (16, 48). Hydrodynamic interactions contribute significantly to the slowdown of protein diffusion in crowded environments (19). In implicit solvent, they are ignored or approximated via the diffusion tensor (16, 19, 44, 49).

Rapid advances in computing hardware and simulation algorithms have opened up the opportunity to study macromolecular crowding using atomistic molecular dynamics (MD) simulations. Explicit solvent accounts directly for excluded volume effects and hydrodynamic interactions and mediates short-range attractive and long-range electrostatic protein–protein interactions (5, 28–30, 50–52). Here, we use atomistic MD simulations of dense protein solutions to calculate the viscosity and protein diffusion coefficients as a function of protein concentration (Fig. 1). Ubiquitin (UBQ), the third IgG-binding domain of protein G (GB3), hen egg white lysozyme (LYZ), and villin headpiece (VIL) are used as model proteins.

Soluble proteins self-associate in concentrated solution to form transient and dynamic clusters (19, 24, 53–58). Clustering has also been reported for membrane proteins (59). The influence of cluster formation on the protein translational and rotational diffusivity has recently been addressed by atomistic simulation studies (29, 30). Here, we build on these findings and put cluster formation in the framework of the Stokes–Einstein relations connecting viscosity, cluster size, and diffusion. Central questions are (i) whether the Stokes–Einstein relations remain valid in concentrated protein solutions (60), (ii) how transient protein interactions affect the diffusivity and apparent hydrodynamic radii of proteins in concentrated solutions (42), (iii) how viscosity depends on protein concentration (39, 61, 62), and (iv) whether colloid models apply to concentrated protein solutions.

To address these questions, we perform extensive MD simulations, develop a cluster model of concentration-dependent

Author contributions: S.v.B. and G.H. designed research; S.v.B. and M.S. performed research; S.v.B., M.S., M.L., and G.H. analyzed data; and S.v.B., M.S., M.L., and G.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. H.-X.Z. is a guest editor invited by the Editorial Board.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

1To whom correspondence should be addressed. Email: gerhard.hummer@biophys.mpg.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1817564116/-/DCSupplemental.

Published online April 29, 2019.

www.pnas.org/cgi/doi/10.1073/pnas.1817564116

PNAS | May 14, 2019 | vol. 116 | no. 20 | 9843–9852
Translational Diffusion Slows Down at High Protein Density. As shown in Movie S1 for GB3 at 200 mg/mL with $N = 540$ proteins, translational and rotational diffusion in concentrated solution is strongly impacted by protein interactions. For each protein in the simulation box, mean-squared squared displacement (MSD) curves were calculated and fitted to the Einstein relation in SI Appendix, Eq. S6 (SI Appendix, Fig. S6). The MSD curves of the protein solutions averaged over starting times and proteins are linear at times exceeding 10 ns. The translational diffusion coefficients $D_\text{T}^{\text{PBC}}$ obtained by fitting the Einstein relation to the MSD from 10 ns to 30 ns are therefore long-time diffusion coefficients. The MSD curves of the dilute solutions (one protein in the simulation box) are linear at small delays and were fitted to the Einstein relation from 0 ns to 5 ns. We corrected $D_\text{T}^{\text{PBC}}$ for large finite-size effects using Eq. 7, where we used $\eta(\phi)$ from the quadratic fit, Eq. 3. The values before finite-size correction are listed in SI Appendix, Table S3.

After finite-size correction, the translational diffusion coefficient $D_\text{T}$ of dilute UBO is consistent with results of NMR spectroscopy (71, 72) in dilute solution. Our calculated $D_\text{T}$ values of LYZ are bracketed by measurements in dilute and dense solutions (21, 24, 25, 71–76). The spread in the measured diffusion coefficients of LYZ is possibly due to differences in pH value, ionic strength, and temperature in the different experiments. All simulation values of $D_\text{T}$ in the dilute solution are very close to Hydropro (70) predictions (Fig. 3d). The translational diffusion coefficients calculated for the large systems (with $N \geq 120$ proteins) at 200 mg/mL concentration are similar to the values for the small systems ($N \leq 20$) at the same concentration, decreasing slightly with increasing box sizes (number of $N$). For all proteins, our dynamic cluster model Eq. 8, developed below, accounts nearly quantitatively for the slowdown of translational diffusion with increasing concentration.

Crowding Strongly Affects Rotational Diffusion. Rotational diffusion coefficients $D_\text{r}$ of the dense protein solutions were obtained by fitting quaternion correlations (80) and by integrating the orientational correlation function $\langle \langle P_3(\cos \theta(t)) \rangle\rangle$, respectively, over the time range 0–100 ns (SI Appendix). Fits to elements of the quaternion covariance matrix are shown in SI Appendix, Fig. S7. The resulting rotational diffusion coefficients $D_\text{T} - D_\text{r}$ are shown in SI Appendix, Fig. S8. Fits to the orientational correlation function are shown in SI Appendix, Fig. S9.

Orientationally averaged diffusion coefficients $\bar{D}_\text{T}$ and $\bar{D}_\text{r}$ after finite-size correction (81) decrease strongly with increasing $\phi$ (Fig. 3b). At infinite dilution, the UBO and GB3 results are bracketed by the rotational diffusion coefficients obtained from Hydropro (70) calculations and from NMR spectroscopy

---

**Fig. 1.** Representative simulation snapshots of dense UBO solution (200 mg/mL, $N = 405$) and GB3 solution (100 mg/mL, $N = 20$). Blue lines indicate the periodically replicated simulation boxes. Proteins are depicted as green (UBO) or green (GB3), Na$^+$ and Cl$^-$ ions as blue and cyan balls, respectively, and TIP4P-D water as red sticks.

**Fig. 2.** Viscosity of concentrated protein solutions as a function of protein volume fraction from all-atom MD simulations (symbols), experiments (dotted lines and gray shading), and HS colloid theory (dashed lines). Blue, UBO; orange, GB3; green, LYZ; red, VIL. The viscosity $\eta$ was calculated from MD using fluctuations in the pressure tensor (SI Appendix, Eq. S5, squares; fit of Eq. 3, solid lines) and the Stokes–Einstein (SE) relations (Eq. 6; small systems $N \leq 20$, circles; large systems $N \geq 120$, diamonds). Green and purple dotted lines: fits of Eq. 3 to experimental data on concentrated LYZ (61) and BSA (62) solutions. Magenta and olive dotted lines: exponential fit (equation 4 in Zorriola et al. (39)) to the experimental data on HSA and RNaseA solutions (39). Black dashed line: Einstein expression for HS colloidal suspensions (Eq. 4). Brown dashed-dotted line: Ladd expression (63, 64) for HSSs (Eq. 5).
The experimental rotational diffusion coefficient in dilute LYZ solution reported in ref. 73 is slightly lower than the calculated values, whereas the rotational diffusion coefficients reported in ref. 25 agree well with our data at all protein concentrations. The calculated rotational diffusion coefficients of dilute LYZ and VIL are in fair agreement with Hydropro (70) predictions. As for the translational diffusion, the rotational diffusion coefficients calculated for the large systems (N ≥ 120) at 200 mg/mL concentration are close to the values of the small systems (N ≤ 20) at the same concentration, being slightly lower for UBQ and slightly higher for GB3. The dynamic cluster model Eq. 9 predicts the rotational diffusion coefficients of UBQ, GB3, VIL, and LYZ accurately over the entire concentration range, except for the LYZ solution at 100 mg/mL concentration, where the effect of the weak clustering (Fig. 3C) is somewhat overestimated.

Diffusion in Dense Protein Solutions Follows the Stokes–Einstein Relation. Given translational and rotational diffusion coefficients, the viscosity can be estimated from the Stokes–Einstein relations (Eq. 6). For all small systems (N ≤ 20), we observe excellent agreement between the viscosity calculated from the autocorrelation of the pressure tensor fluctuations and from the diffusion coefficients (Fig. 2), indicating that the dense protein solutions show normal (Stokes–Einstein) diffusion for all protein volume fractions considered here. For the large systems (N ≥ 120), slight deviations in Dτ and Dφ (Fig. 3) lead to strongly overestimated viscosity (diamonds in Fig. 2). Therefore, in practical calculations we advise against calculating the viscosity via the Stokes–Einstein relation, Eq. 6, because the results are quite sensitive to the uncertainties in Dτ and Dφ.

Hydrodynamic Radius, Cluster Size, and Diffusion Are Related. We obtained very similar hydrodynamic radii from the Stokes–Einstein relations for translation and rotation, Eqs. 10 and 11, respectively (Fig. 3D). Therefore, after correcting for finite-size effects with actual shear viscosities η(φ), both translational diffusion and rotational diffusion follow Stokes–Einstein theory even at high protein volume fractions. Only values for the small systems (N ≤ 20) are shown, because only for these systems was the computationally expensive calculation of the viscosity from the pressure fluctuation autocorrelation function performed.

If the increase in viscosity were to capture all factors that contribute to the concentration-dependent slowdown of protein diffusivity, then the hydrodynamic radius, calculated from the Stokes–Einstein relations, Eqs. 10 and 11, should remain constant at all concentrations. Instead, we observe that the effective hydrodynamic radius cubed, R̃h, increases with protein volume fraction φ. This dependence is consistent with protein cluster formation (Fig. 3C). Indeed, when calculating the cluster size distribution based on an α-carbon distance cutoff criterion, the cluster size distribution shifts to larger clusters at increasing protein volume fraction (SI Appendix, Fig. S10). For protein concentrations up to 100 mg/mL, the mean number of proteins in a cluster grows linearly as m(φ) = 1 + ζφ, with clustering propensity ζ listed in Table 1 (Fig. 3C). The highest concentration (200 mg/mL) was not included in the fit, because the close proximity of proteins causes a significant dependence of the calculated mean cluster size on the cutoff criterion (SI Appendix, Fig. S11). Given the linear increase of cluster size with protein volume fraction φ, the effective hydrodynamic radius cubed should likewise increase linearly with φ: R̃h,eff(φ) = R̃h,φ=0(1 + ζφ), where we assumed that the hydrodynamic radius cubed is proportional to the cluster volume. Remarkably, the dynamic cluster model accounts nearly quantitatively for the relative increase of R̃h,eff(φ) (Fig. 3D).

As shown in SI Appendix, the clustering propensity ζ is related to an effective dissociation constant Kd via the protein volume fraction φ: ζ = 1/(1 + Kdφ). Given protein volumes vP (UBQ) = 10.4 nm³, vP (GB3) = 7.2 nm³, vP (LYZ) = 17.2 nm³, and vP (VIL) = 5.1 nm³ (SI Appendix, Fig. S12), we obtain dissociation constants of Kd,eff(UBQ) = 1/(8.2 × 10.4 nm³ × NAv) ≈ 19 mM, Kd,eff(GB3) ≈ 26 mM, Kd,eff(LYZ) ≈ 26 mM, and Kd,eff(VIL) ≈ 28 mM with NAv Avogadro’s constant (Table 1).

Effective Viscosity Accounts for Hydrodynamic Interactions. We investigated whether the effective viscosity η(φ) captures the indirect (hydrodynamic) effects of increased protein concentration and cluster formation, i.e., all of the effects that are not accounted for by an increased effective hydrodynamic radius. For each protein, we recorded the times at which it is free, i.e., accounted for by an increased effective hydrodynamic radius. For all four protein species, the product η(φ)D̃free(φ) is approximately constant as a function of concentration (SI Appendix, Fig. S15). We conclude that the effective viscosity indeed accounts for the hydrodynamic contributions to the diffusivity slowdown.

Displacement Pair Correlation Shows Contribution from Direct and Hydrodynamic Interactions. We calculated the displacement pair correlation introduced by Ando and Skolnick (19) (SI Appendix). We analyzed the protein pair correlation for pairs at distances 0.6–3 nm. At distances corresponding to cluster formation, we observed highly correlated motion for all protein pairs at all concentrations and time delays (SI Appendix, Fig. S16). At increasing pair distance (~2–3 nm), the pair correlation decreased gradually.

Protein Binding Interfaces. The interactions between the proteins in clusters were loose but not entirely random in their orientation (SI Appendix, Fig. S17). For UBQ, the preferred binding interface coincides remarkably well with the noncovalent dimer interface reported from NMR measurements (82), more or less independent of protein concentration (Fig. 4). It includes residues 8–11, 20, 24, 25, 28, 31–42, 46–49, 54–60, and...
We can also relate the dimensionless Baxter parameter $\tau$ as $\tau = n_{HS} K_d$ (derived in SI Appendix, where $n_{HS}$ is the HS volume. Low $\tau$ indicates high protein stickiness. We conducted Monte Carlo (MC) simulations of Baxter sticky HSs with $\tau = 0.15$ and finite-range attractive interactions up to 1.05 $\sigma$. The dotted curve shows the fit $n_{HS} \phi^2 + n_{HS} \phi + 743.3 \phi^2$ to the MC data up to $\phi = 0.15$. Green crosses show experimental data on LYZ cluster formation (79). (C, Inset) Representative simulation snapshot of dense UBQ solution (100 mg/mL). Blue lines indicate the periodically replicated simulation boxes. Colors and colored labels indicate transient UBQ protein clusters and cluster size $m$, respectively (solvent not shown). (D) Dependence of reduced hydrodynamic radii cubed, $R_h^3/R_{h,\phi=0}^3 \phi$, on protein volume fraction $\phi$. The effective $R_h$ is calculated from the Stokes–Einstein relations for translation, Eq. 10 (circles), and for rotation, Eq. 11 (squares). The dashed lines show the prediction $1 + \zeta \phi$ from C with $\zeta$ values from Table 1.

Colloidal Suspension Model. Baxter’s attractive (sticky) HSs (68, 69, 84–86) are widely used as a model for suspensions of interacting colloidal particles. Their association constant is related to the dimensionless Baxter parameter $\tau$ as $\tau = n_{HS} K_d$ (derived in SI Appendix, where $n_{HS}$ is the HS volume. Low $\tau$ indicates high protein stickiness. We conducted Monte Carlo (MC) simulations of Baxter sticky HSs with $\tau = 0.15$ and finite-range attractive interactions up to 1.05 $\sigma$ of $N = 16$ and $N = 120$ particles in a box, respectively. Using $\tau = 0.15$ and finite-range attractive interactions up to 1.05 $\sigma$. The dotted curve shows the fit $n_{HS} \phi^2 + n_{HS} \phi + 743.3 \phi^2$ to the MC data up to $\phi = 0.15$. Green crosses show experimental data on LYZ cluster formation (79). (C, Inset) Representative simulation snapshot of dense UBQ solution (100 mg/mL). Blue lines indicate the periodically replicated simulation boxes. Colors and colored labels indicate transient UBQ protein clusters and cluster size $m$, respectively (solvent not shown). (D) Dependence of reduced hydrodynamic radii cubed, $R_h^3/R_{h,\phi=0}^3 \phi$, on protein volume fraction $\phi$. The effective $R_h$ is calculated from the Stokes–Einstein relations for translation, Eq. 10 (circles), and for rotation, Eq. 11 (squares). The dashed lines show the prediction $1 + \zeta \phi$ from C with $\zeta$ values from Table 1.

71–76. The C-terminal tail (residues 71–76) and an adjacent relatively hydrophobic surface patch show strong involvement, in line with experimental evidence (82). In dense LYZ solutions, we observed that residues Asp48 and Arg73 contribute most to LYZ–LYZ interaction (SI Appendix, Fig. S17). In a Brownian dynamics study (83), these residues were found to play crucial roles in the formation of a LYZ–LYZ encounter complex.
Concentration-dependent protein diffusion. (A) Dependence of the normalized translational diffusion coefficient $D_{t,vol}^{red}$ on protein volume fraction $\phi$. Solid circles show MD data from this study. Open circles show data from simulation studies (16, 26, 29, 87). Open triangles denote data from experimental studies (21–27, 75, 76, 88–93). Hb, hemoglobin; Mb, myoglobin; “Mixed,” C12 in different dense protein solutions; Ova, ovalbumin. The solid curves show the prediction of the dynamic cluster model $D_{t,vol}^{red}(\phi)$ from our protein simulation data for UBQ, GB3, LYZ, and VIL solutions, indicated by corresponding colors. The dashed curve shows the slowdown of the translational diffusion coefficient predicted from colloid theory on noninteracting HSs by van Blaaderen et al. (94) and Tokuyama et al. (97) ($\eta_\phi$) = 7 $\phi$; $D_B$ = 2 $\eta_\phi$, and substituting the Stokes–Einstein relation for noninteracting HSs by van Blaaderen et al. (94) and Tokuyama et al. (97) ($\eta_\phi$) = 7 $\phi$; $D_B$ = 2 $\eta_\phi$, and substituting the Stokes–Einstein relation for noninteracting HSs by van Blaaderen et al. (94) and Tokuyama et al. (97) ($\eta_\phi$) = 7 $\phi$.

Equating $b = 5.931 + 1.899/\tau$ with the coefficient $b$ from the $\phi^2$ term in the viscosity expansion fitted to MD data (Eq. 3 and Table 1), we obtain values of $\tau$ and $K_d^{red}$ (Table 1), which are again close to those obtained from the cluster model, $g(\tau)$, and $D_t(\phi)$ (Table 1).

**Dissociation Constant from Off Rate of Nonspecific Complexes.**

The cumulative distribution functions of the lifetimes of protein pairs (SI Appendix, Fig. S19) show that most pairs stay together for 1–50 ns, indicating dynamic clustering according to Liu et al.’s (24) terminology. The lifetimes of protein pairs are independent of the protein concentration, supporting the presence of dynamic protein clusters rather than protein aggregation. From the cumulative distribution function, we obtained the same median protein pair lifetime of $\tau_{off}$ = 5 ns for all proteins and concentrations and defined an off rate for non-specific complexes as $k_{off} = 1/\tau_{off} = ln(2)/\tau_{off} \approx 1.4 \times 10^9 M^{-1}s^{-1}$, assuming exponential kinetics. Assuming in addition a Smoluchowski on rate, $k_{on} = 4\pi D_r R_0$, and substituting the Stokes–Einstein relation for translational diffusion, we expect an on rate for nonspecific complexes of $k_{on} = 2k_B TN_A/3\eta_\phi \approx 1.8 \times 10^9 M^{-1}s^{-1}$. The resulting “kinetic” dissociation constant $K_d^{kinetic} = k_{off}/k_{on} \approx 80 \text{ mM}$ agrees well with those obtained from the other methods (Table 1).

**Fig. 5.** Concentration-dependent protein diffusion. (A) Dependence of the normalized translational diffusion coefficient $D_{t,vol}^{red}$ on protein volume fraction $\phi$. Solid circles show MD data from this study. Open circles show data from simulation studies (16, 26, 29, 87). Open triangles denote data from experimental studies (21–27, 75, 76, 88–93). Hb, hemoglobin; Mb, myoglobin; “Mixed,” C12 in different dense protein solutions; Ova, ovalbumin. The solid curves show the prediction of the dynamic cluster model $D_{t,vol}^{red}(\phi)$ (Eq. 8, no adjustable parameters) for UBQ, GB3, LYZ, and VIL solutions, indicated by corresponding colors. The dashed and dotted curves show the slowdown of $D_{t,vol}^{red}$ predicted from colloid theory on noninteracting HSs by van Blaaderen et al. (94) and Tokuyama and Oppenheim (95), respectively (SI Appendix). The dashed-dotted line shows a linear fit to $D_{t,vol}^{red}$ of UBQ and GB3 at $\phi \leq 0.04$. A plot with experimental studies resolved by symbols is shown in SI Appendix, Fig. S20A. (B) Dependence of the normalized rotational diffusion coefficient $D_t^{red}$ on protein volume fraction $\phi$. Solid circles show MD data from this study. Open circles show data from simulation studies (16, 29). Open triangles denote data from experimental studies (22, 25, 96). “Cell”: Hb or Mb in different cell types. The solid curves show the prediction of the dynamic cluster model $D_t^{red}(\phi)$ (Eq. 9, no adjustable parameters) for UBQ, GB3, LYZ, and VIL solutions, indicated by corresponding colors. The dashed curve shows the slowdown of $D_t^{red}$ predicted from colloid theory for noninteracting HSs (97) (SI Appendix). A plot with experimental studies resolved by symbols is shown in SI Appendix, Fig. S20B.
Separating the Effects of Clustering and Hydrodynamics. Using an elegant MD simulation setup, Nawrocki et al. (30) found that direct protein interactions are the dominant contributors to the slowdown of rotational diffusivity, whereas hydrodynamics play only a minor role. Here, we could show that the effective viscosity accounts for the indirect, hydrodynamic effects of dense solutions on the diffusivity slowdown (SI Appendix, Fig. S15). In dense UBQ, GB3, and VIL solutions at 200 mg/mL, the viscosity is ~2.5-fold increased compared to the solvent viscosity (Fig. 2). The translational diffusion decreases by a factor of ~4, whereas the rotational diffusion decreases by a factor of ~6 (Fig. 3 A and B). The direct effect of protein clustering (i.e., the increase in the effective hydrodynamic radius) accounts for an additional factor ~4/2.5 = 1.6 on translational diffusion and a factor ~6/2.5 = 2.4 on rotational diffusion. We therefore find that direct clustering interactions contribute to ~40% and 50% of the total slowdown in translational and rotational diffusivity, respectively, of dense UBQ, GB3, and VIL solutions.

The displacement pair correlation function assesses concerted protein motion and is used to distinguish short-range and long-range interactions in simulations of crowded systems (19). At short distances, the proteins form clusters, and their motion is highly correlated (SI Appendix, Fig. S16). The pair correlation decreases at larger protein pair distances, because the short-range interactions fade out and only contributions from long-range effects (electrostatics, hydrodynamics) remain. Interestingly, at low protein concentrations (30–50 mg/mL), the protein motions are correlated up to high distances, whereas the correlation decreases quickly with increasing distance at high protein concentration. This suggests that the motion of proteins at low concentration is more effectively coupled by hydrodynamic and electrostatic forces. At high concentrations, the coupled motion of a protein pair at short distances appears to be effectively quenched by interference of competing proteins. This trend is surprising, because proteins may serve as bridging intermediates at high concentrations, which should effectively increase the pair correlation for the distances considered.

Protein Solutions as Colloidal Suspensions with Attractions. The effective dissociation constants $K_d$ and Baxter parameters $\tau$ were derived according to the low-concentration behavior of the sticky-HS model for attractive colloidal particles. The $K_d$ values for nonspecific protein–protein binding derived from the structure, dynamics, and viscosity of the protein solutions are fairly consistent with each other (Table 1). As listed in Table 1, $K_d$ estimates from $\pi(\phi, g(\tau), D_\ell(\phi), k_\text{eq},$ and $\eta(\phi)$ vary within about one order of magnitude for a given protein. The corresponding values of $\tau$ also agree well with each other. Interestingly, the decreased clustering propensity of LYZ solutions compared with UBQ, GB3, and VIL solutions is reflected only by a higher $\tau$, but not by an increase in the effective dissociation constant $K_d = \tau/v_\text{p}$, because the increased $\tau$ value of LYZ is compensated by its larger volume $v_\text{p}$. In turn, this results in a relatively small $\xi = 1/(K_d v_\text{p})$ for LYZ. For our limited set of four proteins, protein stickiness decreases with size, thereby ameliorating the aggregation problem for large proteins pointed out by Ando and Solnick (19).

The formation of 3D clusters explains the strong increase of the mean cluster size $\pi(\phi)$ in sticky-HS solutions at large $\phi$ (Fig. 3C). Sticky spheres have no orientational preference, which allows the formation of compact clusters. By contrast, protein–protein interactions are directional and only some protein orientations result in favorable interactions (SI Appendix, Fig. S17). The orientational preference disfavors the formation of compact clusters, as can be seen in the formation of only $m – 1$ protein connections for a cluster of size $m$ up to intermediate (100 mg/mL) protein concentration (SI Appendix, Fig. S21).
In consideration of these differences, \( \tau \) should be considered an effective parameter whose value depends on the property that is fitted and on the fit range. Therefore, we emphasize that the sticky HS model is only a rough approximation of a complex protein solution, which nonetheless explains the behavior of the complex protein systems surprisingly well up to intermediate protein concentrations.

From light-scattering experiments, Scherer et al. (57) estimated \( \tau = 0.6 \) for concentrated monoclonal antibody mAB2 solution (150 mM NaCl concentration), which is of the same order of magnitude as our simulation results (Table 1). They observed a strong dependence of \( \tau \) on the ionic strength, ranging from \( \sim 1.000 \) to \( \sim 0.5 \) when increasing the ion concentration from 40 mM to 600 mM. It will be interesting to study the effect of varying ionic strengths on \( \tau \) by MD simulations. In contrast to mAB2 solutions, the scattering data for antibody mAB1 solutions could not be represented by Baxter's sticky HS model due to stronger (and possibly long-range) interactions, whereas an alternative model based on two oligomeric species (57, 58) fitted the data well. This suggests that only weakly interacting proteins are well described by the attractive HS model (57).

**Limitations.** The generality of our findings on the connection of cluster size, viscosity, and diffusion is subject to several limitations.

All four proteins studied here are small and globular. This justifies the use of averaged rotational diffusion coefficients \( D_r \) and facilitates the comparison with colloid theory, but does not consider the effects of different shapes on the concentration dependence of diffusion. Dense solutions of larger proteins and proteins with varying degree of anisotropy will have to be simulated atomistically to address these issues.

Slight finite-size effects on protein clustering and diffusivity were observed. In the smallest simulation of dense protein solutions (UBQ at 200 mg/mL with \( N = 15 \) proteins), occasional formation of a box-spanning cluster was observed, as seen in the contribution of cluster sizes \( m = 12 - 15 \) to the cluster size distribution for this system (SI Appendix, Fig. S10). This artifact was not seen in any of the other simulations and does not appear to strongly affect the diffusivity and viscosity of the UBQ solution (Figs. 2 and 3). The translational and rotational diffusion coefficients from high concentration (200 mg/mL) vary slightly for higher protein numbers in the simulation box. The absolute deviations in \( D_t \) and \( D_r \) between small and large systems are small. However, the relative deviations in \( D_t \) and \( D_r \) are substantial, overestimating the viscosity calculated from the Stokes–Einstein relations (Fig. 2, circles). On the one hand, we attribute these deviations to incomplete sampling of the extremely large simulation systems. We expect that the simulations with large system size (\( N \geq 120 \)) take more time to equilibrate, thus introducing a bias on the diffusion coefficients, which is not accounted for by the current finite-size corrections (Eq. 7 and SI Appendix, Eq. S10). On the other hand, we observe a significant tail to high cluster numbers in the cluster distribution in these systems (SI Appendix, Fig. S10). Whereas the mean cluster sizes in the small and large systems are similar (Fig. 3C), occasional large clusters in the large systems appear to slightly suppress diffusion in a way not captured by the smaller simulation systems. Assessing and quantifying both effects in detail would merit a separate study.

Based on the clustering data for UBQ and GB3 solutions, we assume that the cluster size increases linearly with protein concentration up to intermediate (100 mg/mL) protein concentration. Although this simple model works surprisingly well in explaining our calculated diffusion data even at higher concentrations (Fig. 3A and B), the actual functional dependence of cluster size on protein concentration may be nonlinear and depend on the specifics of the system (protein type, pH, ionic strength, temperature). Much longer simulations would be needed to precisely determine the cluster distribution and mean cluster size at high protein concentration.

The representation of dense protein solutions by Baxter’s sticky HS suspensions is limited to weak, short-range protein–protein interactions. As scattering data show (57), the sticky HS model does not represent experimental data well for strongly interacting particles or particles with significant attractive long-range (electrostatic) interactions. In our simulations and in cellular conditions, these interactions are effectively shielded by the ions in the solution, making this limitation less relevant in vivo. Nevertheless, it would be interesting to test the applicability of the cluster model for these cases.

**Conclusions**

By performing all-atom molecular dynamics simulations of dense protein solutions, we found an increase in the viscosity of the solutions at higher protein volume fractions, consistent with experimental results (39, 61, 62). This increase is considerably higher than predicted by colloidal models of noninteracting HSs, stressing the importance of measuring or calculating rather than approximating the viscosity at protein volume fractions approaching cellular crowding conditions. We calculated translational and rotational diffusion coefficients and corrected them for finite-size effects using the respective viscosity of the solution. Translational diffusion and rotational diffusion are strongly affected by protein crowding. For LYZ solutions, experimentally measured diffusion coefficients are available also at high concentration and are in excellent agreement with our simulation data (Fig. 3A and B). We calculated effective hydrodynamic radii using the Stokes–Einstein relations and found that a similar increase in the effective hydrodynamic radius can be inferred from the slowdown of translational and rotational diffusion caused by the formation of dynamic protein clusters. Indeed, establishing consistency with the Stokes–Einstein relations requires accounting for protein cluster formation (42) as a result of attractive interactions (60). Overall, we conclude that the concentration dependence of protein cluster size, the translational and rotational diffusion coefficient, and viscosity are consistent with each other (exception: LYZ at 100 mg/mL) and—for the proteins studied here—are explained well by Baxter’s sticky HS model of colloidal suspensions.

In reduced form as a function of protein volume fraction showed that the relative slowdown in translational diffusion is consistent with results from previous studies. The relative slowdown in rotational diffusion shows a larger spread, consistent with the notion that rotational diffusivity depends more sensibly on clustering propensity and thus on the specifics of the protein interactions. Dynamic cluster formation has recently been observed also for membrane proteins (59) and shown to slow down rotational diffusion. In light of our analysis in terms of the cluster and colloidal models, we would expect similar affinities \( K_f \approx 10 \) mM for other abundant, soluble proteins with similar size and shape. It will be interesting to explore the limits of our cluster model, i.e., if the cluster model remains valid for large, anisotropic, and more sticky proteins. In this context, the effect of liquid–liquid phase separation (99) on protein diffusivity is an exciting question to address.

We find that the proteins favor certain orientations for interactions and our findings on UBQ contact interfaces are consistent with experiments (82). The protein interactions lead to highly correlated motion at short distances and the correlation is sustained up to larger distances at low concentration. At high concentration, despite increased protein cluster formation, the pair correlation (19) at similar distances is decreased.

In the cellular environment, the situation is complicated by molecular heterogeneity, reactions, partitioning in microenvironments by phase separation, interactions with membranes.
and structural proteins, and other factors (99–102). Nevertheless, the findings here and in earlier work (103), as well as the observation that diffusion in cell lysates is similar to diffusion in crowded protein solutions (22), suggest that both in concentrated solution and in cells, proteins appear to diffuse not as isolated particles, but as members of dynamic clusters between which they constantly exchange. From the consistency of our diffusivity results with experiments in solution and on the basis of our cluster model, we conclude that—in crowded conditions corresponding to the cellular concentration—the strength of nonspecific protein–protein interactions for abundant proteins such as UBQ should correspond to low-millimolar binding.

We can now carry out atomistic MD simulations of crowded simulations at an unprecedented scale (29), here with up to 540 proteins and 3.6 million atoms in the box simulated over microseconds. Atomistic simulations of solutions of protein mixtures, possibly reflecting the distribution of proteins in the cell, no longer seem out of reach (52). Developments in nucleic acid force fields (104–109) will make it attractive to test the above findings on dense nucleic acid solutions and dense protein–nucleic acid mixtures. Ultimately, the macromolecular diversity in the cell will have to be considered (101) to predict passive diffusion in vivo.

SI Appendix. SI Appendix contains supplementary text: SI Appendix, Figs. S1–S21; SI Appendix, Tables S1–S3; Movie S1; and SI references.

Movie S1. Shown is an atomistic MD simulation of 540 GB3 proteins in concentrated solution (200 mg/mL) at simulation time 0–500 ns. The fully flexible proteins are shown in surface representation and differentiated by color. For clarity, water and ions are omitted. Proteins that seem to appear and disappear traverse the periodic boundaries.

Materials and Methods

MD Simulations of Dense Protein Solutions. We performed all-atom MD simulations of solutions of human UBQ (PDB code 1UBQ (110)), GB3 (PDB code 1P7F (111)), Lyz (PDB code 1E8L (112)), and VIL (PDB code 1VJL (113)) at up to five different densities with N = 15 (UBQ) and N = 20 (GB3, Lyz, VIL) proteins in the simulation box. In addition, the most concentrated UBQ and GB3 solutions (200 mg/mL) were simulated with large simulation boxes containing from N = 120 to N = 540 proteins (SI Appendix, Table S1). To mimic an infinitely dilute system, MD simulations with a single protein copy were carried out. The simulation procedures are detailed in SI Appendix.

Viscosity Calculation and Approximations. The low-frequency, low-shear viscosity η(φ) of dense protein solutions differs from the viscosity η0 of the pure solvent consisting of only water and ions. We determined η(φ) and η0 from MD simulations by integration of the autocorrelation functions of the pressure tensor fluctuations (114), as detailed in SI Appendix. The dependence of η on the protein volume fraction φ is well captured by a quadratic function,

\[ \eta(\phi) = (1 + 2.5\phi + 0.5\phi^2)\eta_0. \]

with parameter b fitted to the calculated viscosities, η0 the solvent viscosity averaged over the NaCl concentrations used in this study, and the coefficient 2.5 adopted from Einstein's colloid theory (66). The experiments of Woutersen and De Kruif (69) have shown that the coefficient b increases with the strength of attractions between colloidal particles.

We compared the calculated viscosities to predictions from colloid theory. At low solute volume fractions φ, Einstein (66) predicted a linear dependence of the viscosity of HS suspensions on φ,

\[ \eta(\phi) = (1 + 2.5\phi)\eta_0. \]

For higher solute volume fractions, this expression was modified by Ladd (63, 64) to

\[ \eta(\phi) = 1 + 5\phi(1 + S(\phi))^{-1} - \phi(1 + S(\phi))^{-2}. \]

with S(\phi) = \phi^2 - 2.3\phi^3.

In addition, we also estimated the viscosity by assuming that the Stokes–Einstein relations for rotational diffusion, \( D_r = k_BT/(8\pi\eta R_p^2) \), and for translational diffusion, \( D_t = k_BT/(6\pi\eta R_p^3) \), are exactly satisfied, such that

\[ \eta(\phi) = k_BT \frac{D_t(\phi)}{27D_r^2(\phi)}. \]

where the bar indicates that we average over orientational asymmetries.

Translational Diffusion. Translational diffusion coefficients \( D_t^{\text{obs}} \) were obtained for each protein density by fitting the Einstein relation to the MSD, as detailed in SI Appendix. The diffusion coefficients were corrected for finite-size effects using (115)

\[ D_t = D_t^{\text{obs}} + \frac{k_BT}{6\pi\eta(\phi)L^2}, \]

with \( \xi = 2.837297 \) and \( L \) the edge length of the cubic simulation box. In the simulations of dilute proteins, where the protein radius \( R_p \) is comparable to the box dimension \( L \), we used the expanded correction, \( \xi = 2.837297 - 4R_p^2/L^2 \) (115).

We compared the calculated translational diffusion coefficients to the predictions of a dynamic cluster model without any free parameters,

\[ D_t(\phi) = D_t(\phi = 0) = \frac{\eta_0}{\eta(\phi)(1 + \phi / \phi_c)^1/2}. \]

where 1 + \( \phi_c \) is the mean cluster size, as defined in SI Appendix and calculated directly from the MD structures. This model is based on the assumption that the slowdown in translational diffusion is linked to the increase in the effective hydrodynamic radius and the viscosity via the Stokes–Einstein relation for translational diffusion. We also compared the reduced translational diffusion coefficients \( D_t(\phi)/D_t(\phi = 0) \) to the approximate form for monodisperse noninteracting HS colloidal suspensions (94, 95) (SI Appendix).

Rotational Diffusion. Rotational diffusion coefficients were calculated following the procedure by Linke et al. (80, 116), as detailed in SI Appendix. Additionally, an effective rotational diffusion coefficient was obtained from fits to the orientational correlation function \( \langle P(t_1 \cos \theta(t)) \rangle \) = \( \langle \cos \theta(t) \rangle \) (117, 118), as detailed in SI Appendix.

We compared the calculated rotational diffusion coefficients to the predictions of the dynamic cluster model,

\[ D_r(\phi = 0) = D_r(\phi = 0) = \frac{\eta_0}{\eta(\phi)(1 + \phi / \phi_c)}. \]

where we again assumed the Stokes–Einstein relation to apply with a hydrodynamic radius cubed proportional to the mean cluster size 1 + \( \phi_c \). We also compared the reduced rotational diffusion coefficients \( D_r(\phi)/D_r(\phi = 0) \) to predictions from colloidal models of noninteracting HSs (97) (SI Appendix).

Hydrodynamic Radius. We solved the Stokes–Einstein relations for translational and rotational diffusion to define effective hydrodynamic radii for translation,

\[ R_{h,t}(\phi) = \frac{k_BT}{6\pi\eta(\phi)D_t(\phi)}. \]

and for rotation,

\[ R_{h,r}(\phi) = \left( \frac{k_BT}{8\pi\eta(\phi)D_r(\phi)} \right)^{1/3}. \]

Gaussian error propagation of \( D_t(\phi) \), \( D_r(\phi) \), and \( \eta(\phi) \) was employed to estimate the errors of \( R_{h,t} \) and \( R_{h,r} \).

ACKNOWLEDGMENTS. We thank Jürgen Köfinger and Martin Vögele for helpful discussions and technical assistance and Kara Grotz for help with the simulation setup. This research was supported by the Max Planck Society (S.v.B., M.S., M.L., and G.H.), the Human Frontier Science Program RGP0002/2017 (S.v.B. and G.H.), and the Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz (LOEWE) Dynamem program of the state of Hesse (M.S. and G.H.).
36. Mastro AM, Babich MA, Taylor WD, Keith AD (1984) Diffusion of a small molecule in aqueous solutions. J Chem Phys 81:3414–3418.
37. Jin S, Verkman AS (2007) Single particle tracking of complex diffusion in membranes: Simulation and detection of barrier, raft, and interaction phenomena. J Phys Chem B 111:3625–3632.
38. Gorti S, Ware BR (1985) Probe diffusion in an aqueous polyelectrolyte solution. J Phys Chem 89:4465–4469.
39. Prados-Cerdeña MM, Ferrer-Costa I, Sastre-Campos M (2003) Diffusion of proteins in crowded protein solutions. FEBS J 270:1195–1201.
40. Roos-Runge F, et al. (2011) Protein self-diffusion in crowded solutions. Proc Natl Acad Sci USA 108:11815–11820.
41. Li C, Liu M (2013) Protein dynamics in living cells studied by cell-IM NMR spectroscopy. FEBS Lett 587:1008–1011.
42. Rothe M, et al. (2016) Transient binding accounts for apparent violation of the generalized Stokes-Einstein relation in crowded protein solutions. Phys Chem Chem Phys 18:17920–17928.
43. Trovato F, Tozzini V (2014) Diffusion within the cytoplasm: A mesoscale model of crowded systems. J Chem Phys 141:035102.
44. Park S, Agmon N (2008) Theory and simulation of diffusion-controlled Michaelis-Menten kinetics for a static enzyme in solution. J Phys Chem B 112:5977–5987.
45. Gusev AJ, Perez-Gonzalves GM, Sweer SL, Young GB, Pielak GJ (2018) Protein shape modulates crowding effects. Proc Natl Acad Sci USA 115:10965–10970.
46. Neves MA, Perevezikhova T, Martorana V, Manno M, Roberts CJ (2018) Protein—protein interactions in dilute solutions of conditioned protein solutions with a mean field treatment of hydrodynamic interactions. J Phys Chem B 122:8523–8533.
47. Ansari A, Jones CM, Henry ER, Hofrichter J, Eaton WA (1992) The role of solvent viscosity in the dynamics of protein conformational changes. Science 256:1796–1798.
48. Elcock AH (2010) Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. Curr Opin Struct Biol 20:196–206.
49. Ando T, Skolnick J (2010) Crowding and hydrodynamic interactions likely dominate in vivo macromolecular motion. Proc Natl Acad Sci USA 107:18457–18462.
50. Treviño F, Tozzini V (2014) Diffusion within the cytoplasm: A mesoscale model of interacting macromolecules. Biophys J 107:2579–2591.
51. Nesmelova IV, Skirda VD, Fedotov DV (2002) Generalized concentration dependence of globular protein self-diffusion coefficients in aqueous solutions. Biopolymers 63:132–146.
52. Wang Y, Lu C, Pielak GJ (2010) Effects of proteins on protein diffusion. J Am Chem Soc 132:9392–9397.
53. Porcar L, et al. (2010) Formation of the dynamic clusters in concentrated lysozyme protein solutions. J Am Chem Soc 132:126–129.
54. Liu Y, et al. (2011) Lysozyme protein solution with an intermediate range order parameter. J Phys Chem B 115:7238–7247.
55. Roos M, et al. (2016) Coupling and decoupling of rotational and translational diffusion of proteins: crowding conditions and temperature. Soft Matter 12:10365–10372.
56. Balbo J, Mereghetti P, Herten DP, Wade RC (2013) The shape of protein crowders is a major determinant of protein diffusion. Biophys J 104:1576–1584.
57. Grimaudo M, Roosen-Runge F, Zhang F, Seydel T, Schreiber F (2014) Diffusion and dynamics of globular protein crowders in aqueous solutions. J Phys Chem B 118:7203–7209.
58. Feig M, Sugita Y (2013) Reaching new levels of realism in modeling biological macromolecules in cellular environments. J Mol Graph Model 35:144–156.
59. Navrocki G, Wang PH, Yu L, Sugita Y, Feig M (2017) Slow-down in diffusion of globular proteins in crowded environments. Phys Chem Chem Phys 19:11260–11267.
60. Phillips GD, Ullmann GS, Ullmann K, Lin TH (1985) Phenomenological scaling laws for “semidilute” macromolecule solutions from light scattering by optical probe particles. J Chem Phys 83:6449–6456.
61. Konopka MC, et al. (2009) Cytoplasmic protein mobility in osmotically stressed Esherichia coli. J Bacteriol 191:231–237.
62. Verkman AS (2003) Diffusion of proteins in cells measured by fluorescence recovery after photobleaching. Methods Enzymol 360:635–648.
63. Mastro AM, Babich MA, Taylor WD, Keith AD (1984) Diffusion of a small molecule in the cytoplasm of mammalian cells. Proc Natl Acad Sci USA 81:3414–3418.
64. Jin S, Verkman AS (2007) Single particle tracking of complex diffusion in membranes: Simulation and detection of barrier, raft, and interaction phenomena. J Phys Chem B 111:3625–3632.
65. Ortega A, et al. (2013) Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. Biophys J 101:892–898.
66. Altieri AS, Byrd RA, Hinton DP (1995) Association of biomolecular systems via pulsed field gradient NMR self-diffusion measurements. J Am Chem Soc 117:5756–5761.
67. Mayh KH, Iyini E, Park H (1996) A recipe for designing water-soluble, beta-sheet-forming peptides. Protein Sci 5:1301–1315.
68. Dubin SB, Clark NA, Benedek GB (1971) Measurement of the rotational diffusion coefficient of lysozyme by depolarized light scattering: Configuration of lysozyme in vivo. Biophys J 7:1002–1006.
69. Cossins RP, Jacobson MP, Guiller V (2011) A new view of the bacterial cytosol structure. Proc Natl Acad Sci USA 108:5010–5015.
70. Liu C, Pradz-Garcia D, Rar H (2013) Structure and dynamics of water in crowded environments slows down peptide conformational changes. J Chem Phys 141:45101.
71. Yu L, et al. (2016) Biomolecular interactions modulate macromolecular structure and dynamics in atomistic model of a bacterial cytoplasm. Elife 5:e19274.
72. Minton AP (1980) Excluded volume as a determinant of protein structure and stability. Biophys J 32:77–99.
73. Minton AP (1983) The effect of volume occupancy upon the thermodynamic activity of proteins: Some biochemical consequences. Mol Cell Biochem 55:119–140.
74. Muramatsu N, Minton AP (1989) Hidden self-association of proteins. J Mol Recognit 1:166–171.
75. Stradner A, et al. (2004) Equilibrium cluster formation in crowded protein solutions and colloids. Nature 432:492–495.
76. Scherer TM, Liu J, Shire SJ, Minton AP (2010) Intermacromolecular interactions of IgG1 monoclonal antibodies at high concentrations characterized by light scattering. J Phys Chem B 114:12948–12957.
77. Lillystrom WG, Yadav S, Shire SJ, Scherer TM (2013) Monoclonal antibody self-association, cluster formation, and rheology at high concentrations. J Phys Chem B 117:6373–6384.
Wang D, Kreutzer U, Chung Y, Jue T (1997) Myoglobin and hemoglobin rotational diffusion in the cell. Biophys J 73:2764–2770.

97. Zuzovsky M, Adler PM, Brenner H (1983) Spatially periodic suspensions of convex particles in linear shear flows. Ill. Dilute arrays of spheres suspended in Newtonian fluids. Phys Fluids 26:1714.

98. Shukla A, et al. (2008) Absence of equilibrium cluster phase in concentrated lysozyme solutions. Proc Natl Acad Sci USA 105:5075–5080.

99. Hyman AA, Weber CA, Jülicher F (2014) Liquid-liquid phase separation in biology. Annu Rev Cell Dev Biol 30:39–58.

100. Luby-Phelps K (2013) The physical chemistry of cytoplasm and its influence on cell function: An update. Mol Biol Cell 24:2593–2596.

101. Feig M, Yu Y, Wang PH, Navrocki G, Sugita Y (2017) Crowding in cellular environments at an atomistic level from computer simulations. J Phys Chem B 121:8009–8025.

102. Rivas G, Minton AP (2018) Toward an understanding of biochemical equilibria within living cells. Biophys Rev 10:241–253.

103. Johnson ME, Hummer G (2011) Nonspecific binding limits the number of proteins in a cell and shapes their interaction networks. Proc Natl Acad Sci USA 108:603–608.

104. Ivani J, et al. (2015) Parmbsc1: A refined force field for DNA simulations. Nat Methods 13:55–58.

105. Zgarbová M, et al. (2015) Refinement of the sugar-phosphate backbone torsion beta for AMBER force fields improves the description of Z- and B-DNA. J Chem Theory Comput 11:5723–5736.

106. Pérez A, et al. (2007) Refinement of the AMBER force field for nucleic acids: Improving the description of α/β conformers. Biophys J 92:3817–3829.

107. Bana P, et al. (2010) Performance of molecular mechanics force fields for RNA simulations: Stability of UUCG and GNRa hairpins. J Chem Theory Comput 6:3836–3849.

108. Zgarbová M, et al. (2011) Refinement of the Cornell, et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. J Chem Theory Comput 7:2886–2902.

109. Tan D, Piana S, Dirks RM, Shaw DE (2018) RNA force field with accuracy comparable to state-of-the-art protein force fields. Proc Natl Acad Sci USA 115:20171027.

110. Vijay-Kumar S, Bugg CE, Cook WI (1987) Structure of ubiquitin refined at 1.8 Å resolution. J Mol Biol 194:531–544.

111. Ulmer TS, et al. (2003) Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal NMR spectroscopy. J Am Chem Soc 125:9179–9191.

112. Schwalbe H, et al. (2001) A refined solution structure of hen lysozyme determined using residual dipolar coupling data. Protein Sci 10:677–688.

113. McKnight CJ, Mutsudaira PT, Kim PS (1997) NMR structure of the 35-residue villin headpiece subdomain. Nat Struct Biol 4:180–184.

114. Hess B (2002) Determining the shear viscosity of model liquids from molecular dynamics simulations. J Chem Phys 116:209.

115. Yeh IC, Hummer G (2004) System-size dependence of diffusion coefficients and viscosities from molecular dynamics simulations with periodic boundary conditions. J Phys Chem B 108:15873–15879.

116. Linke M (2018) Pydiffusion–A python software implementation to calculate the quaternion covariances from simulations, fit the rotational diffusion tensor, and run rotational Brownian dynamics simulations. Available at https://github.com/bio-phys/. Accessed April 10, 2018.

117. Favro LD (1960) Theory of the rotational Brownian motion of a free rigid body. Phys Rev 119:53–62.

118. Woessner DE (1962) Nuclear spin relaxation in ellipsoids undergoing rotational Brownian motion. J Chem Phys 37:647–654.