THE JOURNAL OF PHYSICAL CHEMISTRY Letters

Letter

pubs.acs.org/JPCL

Crowding-Controlled Cluster Size in Concentrated Aqueous Protein Solutions: Structure, Self- and Collective Diffusion

Michal K. Braun,[†] Marco Grimaldo,[†,‡] Felix Roosen-Runge,[*,‡,§] Ingo Hoffmann,[‡] Orsolya Czakkel,[‡] Michael Sztucki,[¶] Fajun Zhang,[*,†] Frank Schreiber,[†] and Tilo Seydel[*,‡]

[†]Institut für Angewandte Physik, Universität Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany

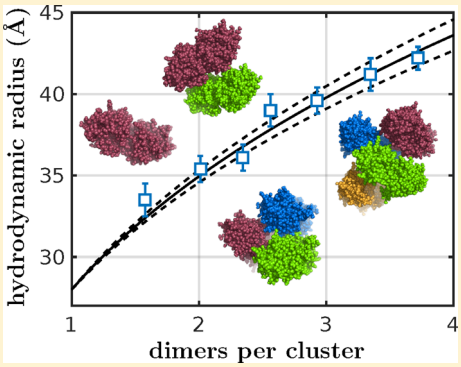
[‡]Institut Laue-Langevin, 71 Avenue des Martyrs, 38000 Grenoble, France

[¶]ESRF - The European Synchrotron, 71 Avenue des Martyrs, 38000 Grenoble, France

[§]Division of Physical Chemistry, Department of Chemistry, Lund University, Naturvetarvägen 14, 221 00 Lund, Sweden

Supporting Information

ABSTRACT: We investigate the concentration-controlled formation of clusters in β-lactoglobulin (BLG) protein solutions combining structural and dynamical scattering techniques. The static structure factor from small-angle X-ray scattering as well as de-Gennes narrowing in the nanosecond diffusion function $D(q)$ from neutron spin echo spectroscopy support a picture of cluster formation. Using neutron backscattering spectroscopy, a monotonous increase of the average hydrodynamic cluster radius is monitored over a broad protein concentration range, corresponding to oligomeric structures of BLG ranging from the native dimers up to roughly four dimers. The results suggest that BLG forms compact clusters that are static on the observation time scale of several nanoseconds. The presented analysis provides a general framework to access the structure and dynamics of macromolecular assemblies in solution.



The formation of protein clusters in aqueous solutions is of great fundamental and practical interest.[1−4] The understanding of cluster formation may help to rationally choose parameters for controlled self-assembly, and to circumvent undesired clustering in the case of the delivery of highly concentrated protein-based drugs.[4,5] Moreover, a systematic understanding of protein cluster formation is important in the context of exploring possible dynamic precursor processes of dynamical arrest, protein aggregation, and crystallization.[6−9] Various parameters influence the formation and size of protein clusters, including the protein concentration, pH, and salt-induced charge effects in the solutions.[4,10] So far, only very few protein solutions undergoing cluster formation have been investigated systematically with respect to both their nanometer structure and nanosecond dynamics.[1,5,7,10,11]

Lysozyme represents the most comprehensively studied system in this context.[1,2,11−14] The existence of lysozyme protein clusters in aqueous solutions has been shown by Stradner et al.[12] using small-angle X-ray and neutron scattering (SAXS/SANS). The small-angle scattering studies on lysozyme solutions suggest that isotropic charge-mediated interactions consisting of a short-range attraction and a long-range repulsion contribute to the formation of clusters. Depending on the protein concentration, three to ten proteins merge into a cluster.[11−13] From the static studies alone, the question of finite lifetime of the clusters remained controversial and inspired dynamics studies using neutron spin−echo spectroscopy (NSE),[1] setting the lower limit of the cluster lifetime to about 25 ns. Further studies observing diffusion on the much longer millisecond time scale found evidence for a finite lysozyme cluster lifetime on this scale.[15,16]

So far, as summarized above, systematic studies of protein cluster formation addressing both static and dynamic aspects have focused on the lysozyme model system, employing SAXS/SANS, NSE, and NMR methods, complemented by theory and simulations.[11,17−21] A recent study, by Grimaldo et al.,[10] systematically addressed the cation-induced formation of protein clusters of bovine serum albumin via neutron backscattering spectroscopy (NBS). This technique is promising for the study of clustering for two reasons: First, at short nanosecond time scales, the self-diffusion only depends on the individual cluster mobility in a hydrodynamic medium, and is not affected by encounters with other clusters. Second, by measuring incoherent scattering at large scattering vectors q, NBS unambiguously accesses the cluster self-dynamics independent of the cluster size. By contrast, NSE accesses collective diffusion, which approximates the self-diffusion only in the limit of large q, and the convergence of this limit depends on the protein concentration, aggregation, and charge state via the structure factor.[22,23]

Here, we report on solutions of bovine β-lactoglobulin (BLG), where clustering is induced solely by a high protein

Received: March 20, 2017
Accepted: May 19, 2017
Published: May 19, 2017

DOI: 10.1021/acs.jpclett.7b00658
J. Phys. Chem. Lett. 2017, 8, 2590−2596

volume fraction, i.e., by crowding, without the necessity for the presence of salt or a non-neutral pH. Under physiological conditions, BLG is found predominantly as a dimer.[24,25] We combine structural investigations using SAXS with dynamic investigations using both NSE and NBS to arrive at a consistent picture on cluster formation and structure.

The SAXS scattering profiles shown in Figure 1A display a clear concentration-dependent correlation peak at intermediate

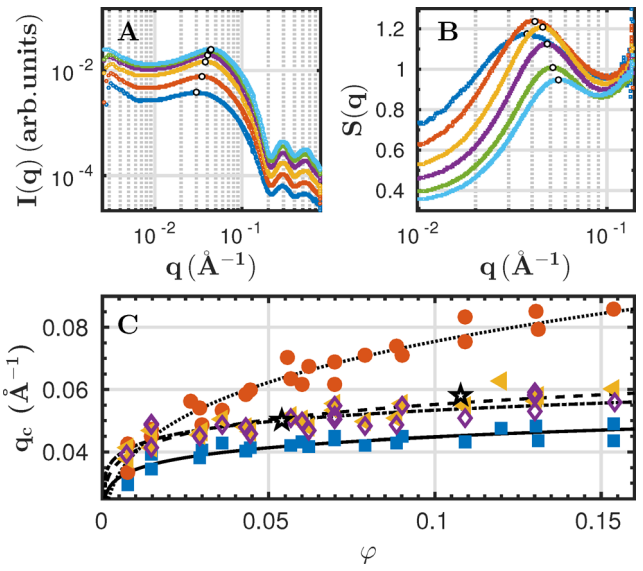


Figure 1. (A) Representative SAXS profiles of BLG solutions with protein concentrations of 10, 20, 40, 60, 80, and 100 mg/mL (from bottom to top). (B) Experimental structure factors obtained from the SAXS profiles by dividing by the form factor of the BLG dimer (cf. Methods). The open circles mark the positions of the peak maxima. (C) Position q_c of correlation peaks as a function of the protein volume fraction φ from different approaches: maxima of the SAXS intensity according to part A (blue squares); maxima of the experimental structure factor according to part B (yellow triangles); based on a fit using a SC structure factor (orange circles); based on a fit using a Two-Yukawa structure factor (purple diamonds). The open star symbols denote q_c obtained from NSE (Figure 2).

q. At high q, the scattering profiles are concentration-independent, and are clearly consistent with the form factor of a BLG dimer. This observation indicates the dominance of dimeric building blocks at all concentrations, because no longer-range order is required in the form factor model. We observe a similar behavior for BLG in both H_2O and D_2O (Figure S1). The SAXS measurements have been repeated for different protein concentrations and BLG proteins from different batches (Figure S2).

The pronounced correlation peak in the SAXS profiles (Figure 1A) indicates an overall repulsive interaction, and the peak position q_c is related to the averaged center-to-center distance $d_{cc} = 2\pi/q_c$ between proteins. Assuming a solution consisting solely of BLG dimers, the volume per dimer $V \propto d_{cc}^3$ shrinks linearly in dimer volume fraction φ. Thus, the peak position q_c should scale with $\varphi^{1/3}$ for a monodisperse system.[26−28]

In contrast to this expectation, the peak position scales with a significantly smaller exponent. We employed four different methods to extract q_c from the data (Figure 1C; see Methods for details), and fitted the function $q_c = a\cdot\varphi^b$ (lines in Figure 1C, Table S1). First, we used the model-free peak positions of the SAXS intensity curves (blue squares; for method see Figure 1A), resulting in an exponent of $b = 0.11 \pm 0.03$. Second, the peak positions of the experimental structure factors (yellow triangles; for method see Figure 1B) yield $b = 0.13 \pm 0.03$. Third, peak positions of fitted structure factors based on a Two-Yukawa potential with short-range attraction and long-range repulsion[2] result in $b = 0.10 \pm 0.03$ (purple diamonds). Fourth, peak positions of fitted structure factors based on a screened-Coulomb potential yield an exponent of $b = 0.28 \pm 0.03$ (orange circles), deviating from the other methods.

To further quantify the deviation from a monodisperse system, we estimate the number density of clusters

$$n_{\text{cluster}} = \varphi_{\text{cp}} \Big/ \left(\frac{1}{6}\pi d_{\text{cc}}^3\right) = \frac{q_c^3}{\pi^3 4\sqrt{2}} \tag{1}$$

Here, we assume that imaginary spheres around the clusters with the diameter of the intercluster distance effectively have a packing fraction close to $\varphi_{\text{cp}} = \frac{\pi}{3\sqrt{2}} \approx 0.74$.[29] The number density for dimers as the basic building blocks of the clusters is given from the sample preparation as

$$n_{\text{dimer}} = \frac{m}{M_w}\cdot\frac{1}{(V_s + \theta_p\cdot m)} \tag{2}$$

where m is the protein mass, $M_w = 36.8$ kDa is the molecular weight of the BLG dimer, V_s is the volume of added water and $\theta_p = 0.75$ mL/g[30] denotes the specific volume of the protein.

From these two number densities, the number of dimers per cluster $N_{\text{dimers}} = n_{\text{dimer}}/n_{\text{cluster}}$ is calculated, yielding an increase from $N_{\text{dimers}} \approx 1.6$ at $\varphi = 0.054$ to 3.7 at $\varphi = 0.18$.

Complementing this structural information on the dimers per cluster, we use neutron spectroscopy, in particular, neutron spin echo (NSE) and neutron backscattering (NBS), to obtain detailed information on cluster formation and structure. While NSE accesses collective motions on length scales of several tenths of nanometers, NBS allows to address the ensemble-averaged self-diffusion coefficients of the proteins on nanometer length scales. The intermediate scattering functions from NSE are well described by a single exponential with relaxation rate $\gamma(q)$ (Figure 2A). The NBS spectra are fitted following the previously published protocol[31,32] for each recorded q-value individually and independently, thus not imposing any q-dependence for the Lorentzian line width $\gamma(q)$ during the fitting (Figure 2B).

The q-dependent diffusion coefficients $D(q) = \gamma/q^2$ for two samples obtained from NSE and NBS are shown in Figure 2C, and the finite, nonvanishing $D(q)$ as well as the fact that $S(q = 0) < 1$ supports the picture of overall repulsive protein clusters. The collective diffusion function $D(q)$ observed in NSE is given by

$$D(q) = D_0\frac{H(q)}{S(q)} \tag{3}$$

where D_0 denotes the self-diffusion coefficient in the limit of infinite dilution, and $S(q)$ is the static structure factor. The hydrodynamic function $H(q)$ scales in a comparable way to $S(q)$, with generally weaker features.[33,34] The minimum in $D(q)$ thus corresponds to the peak position q_c of the structure factor, and the obtained values for q_c using a dynamic technique agree well with the values from SAXS (Figure 1C, black stars). As theoretically expected, the collective diffusion coefficient $D(q)$ from NSE converges to the self-diffusion coefficient $D(q)$ from

DOI: 10.1021/acs.jpclett.7b00658
J. Phys. Chem. Lett. 2017, 8, 2590−2596

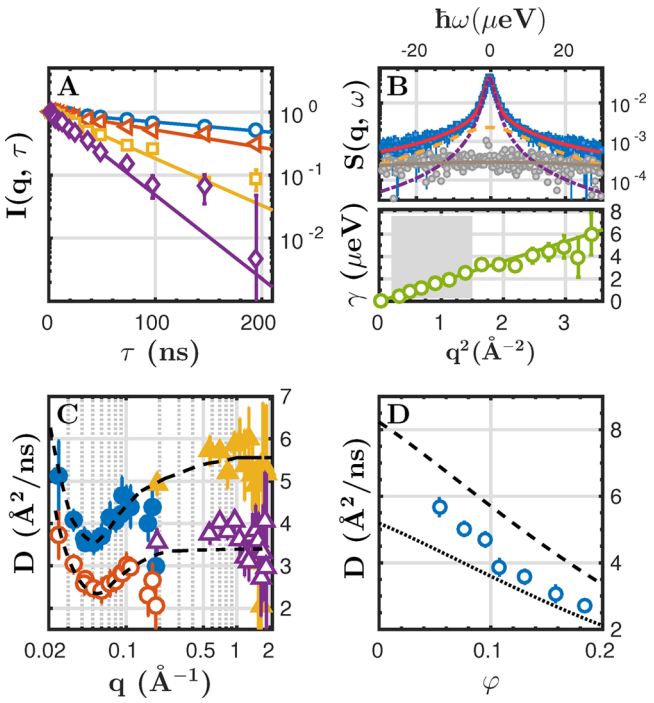


Figure 2. (A) Example intermediate scattering functions $I(q, \tau)$ from neutron spin−echo spectroscopy (BLG concentration 161 mg/mL) versus relaxation time τ for q = 0.034, 0.051, 0.081, 0.10 Å^{-1} (from top to bottom). (B) Example neutron backscattering results for a BLG concentration of 300 mg/mL. Top: Spectrum at q = 0.70 Å^{-1} (blue) along with model fits (lines, see Methods), and corresponding D_2O background (gray). Bottom: Line widths $\gamma = Dq^2$ from the narrow contribution (dash-dotted) represent simple diffusion of the entire protein molecule (fit region in gray). (C) $D(q)$ from neutron spin echo (circles) and backscattering (triangles) for 76 (open symbols) and 161 mg/mL (filled symbols) BLG solutions. The dashed lines are guides to the eye. (D) Apparent diffusion coefficients D from neutron backscattering versus the dry protein volume fraction φ, eq 4. The dashed and dotted lines are the theoretical D for a BLG dimer and a cluster consisting of 4 dimers, respectively (cf. text).

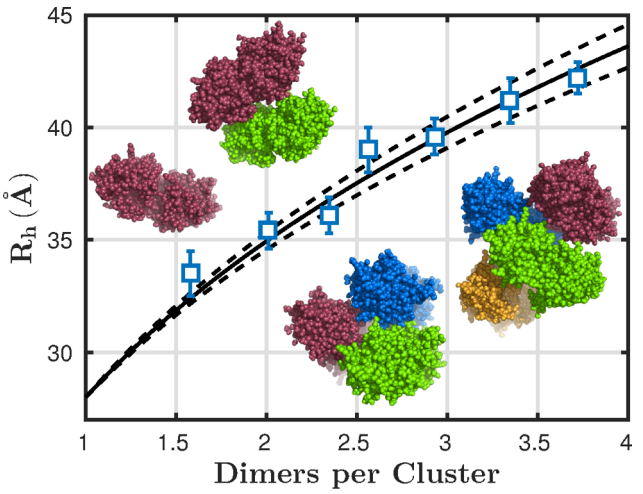


Figure 3. Effective hydrodynamic radii R_h of the clusters obtained from backscattering data versus number of dimers per cluster $N_{dimers} = n_{dimer}/n_{cluster}$ obtained from the SAXS data using eqs 1,2. The solid line is a fit of the data yielding $R_h = 28.0 \cdot N_{dimers}^{0.32 \pm 0.02}$. The dashed lines indicate the 95% confidence limits for the fitted exponent. Note that the result for R_h is sensitive to the position of the correlation peak extracted from the SAXS and NSE data. The schematic illustrations of the clusters were created by assembling one to four dimers (1beb.pdb) using PyMOL.[37]

NBS at high q, since collective effects become weak at small length scales clearly below the interparticle distance and short times below the interparticle interaction time.

An even more detailed picture of cluster formation is obtained from the self-diffusion coefficients measured by NBS for several protein volume fractions. Consistent with the information from SAXS, the resulting apparent diffusion coefficients D for all samples (symbols in Figure 2D, Table S2) are not in agreement with the theoretical diffusion of BLG dimers (dashed line in Figure 2D) at high φ. The theoretical D was calculated following ref 35, assuming R_h = 2.80 nm for the dimer (HYDROPRO,[36] dashed line) and R_h = 4.44 nm for a compact cluster of 4 dimers (dotted line). Given the small length scales, the apparent diffusion coefficients D depend on the translational D_t, and rotational D_r diffusion coefficients via an implicit function.[32,35] By modeling the clusters as effective hard spheres with volume $\frac{4}{3}\pi R_h^3$, D_t and D_r are only determined by the effective hydrodynamic radius R_h and the known dependencies on the volume fraction (see Methods for details). Fitting the apparent diffusion coefficients D for each protein concentration with the effective hydrodynamic radii $R_h(\varphi)$ as only free parameter, we obtain detailed information on the average hydrodynamic size of the clusters (Figure 3, Table S2).

Combining the information on the φ-dependent hydrodynamic and static size of the clusters (Figure 3), we obtain the scaling $R_h = R_{h,dimer} \cdot (N_{dimers})^f$ with f = 0.32 ± 0.02. $R_{h,dimer}$ = 28.0 Å^{-1} is given by the hydrodynamic radius of the BLG dimer calculated using HYDROPRO[36] and the dimer pdb file 1beb.pdb. The fitted exponent f = 0.32 ± 0.02 supports a picture of compact assemblies of BLG dimers.

In this context, it is important to stress that the result of compact clusters is robust against changes in the central assumption in eq 1. Using a very different packing fraction $\varphi_{cp} = \pi/6 \approx 0.52$ corresponding to virtual spheres on a cubic grid, we obtain a scaling exponent of f = 0.29 ± 0.01, while R_h and the numbers of dimers per cluster increase to compensate for the decreased $n_{cluster}$.

The combination of results from the SAXS, NBS, and NSE experiments permits to infer on several properties of the protein solutions.

First, the comparison of SAXS and NSE data allows to robustly extract information on the structure factor maximum q_c (Figure 1C). While the dynamic NSE data directly access the structure factor (see eq 3), different models for the SAXS profiles return different q_c values due to the influence of changing form factor contributions. In particular, both the peak position from the experimental structure factor and the fit using the Two-Yukawa potential agree with the NSE result, whereas the model of a screened Coulomb potential and the intensity maximum of the SAXS curves fail. From this observation we tentatively conclude that our BLG solutions are governed by a superposition of a short-range attraction and a long-range repulsion, favoring the formation of clusters.[38] The corresponding scaling $q_c \propto \varphi^b$ with $b \approx 0.12 \ll 1/3$ indicates the presence of protein clusters in our samples.

Second, the high-q limit of the collective diffusion function $D(q)$ observed in the NSE results is in good agreement with the self-diffusion coefficients observed using NBS on the same samples, as expected for length scales much shorter than the

DOI: 10.1021/acs.jpclett.7b00658
J. Phys. Chem. Lett. 2017, 8, 2590−2596

interparticle distance. This observation corroborates the general consistency of the NSE and NBS experiments. This agreement also supports the picture that the assumed protein clusters in our samples have a lifetime that is larger than both the NSE and NBS observation time scales. In the opposing picture of a cluster lifetime that would cross the NSE or NBS time scales, we would expect inconsistent results from the two techniques.

Third, NBS allows one to systematically explore the dependence of the self-diffusion coefficient on the protein concentration in the samples. We observe that the NBS data are not consistent with the presence of dimers over the explored concentration range (Figure 2D). This observation gives rise to the assumption that clusters are present in our samples. Using the presented analysis, we access the concentration-dependence of the cluster hydrodynamic radius R_h in our samples. The increase of R_h with rising protein concentration (Figure 3) can be associated with the presence of clusters with increasing size. We can conclude that these clusters are static on the observation time scale of our NBS experiment: if the clusters were transient, they would diffuse with a smaller effective hydrodynamic radius.

Fourth, by combining the information from NBS with the SAXS results, we can associate R_h with the number of dimers per cluster (Figure 3). The obtained scaling $R_h \propto N_{dimers}^{1/3}$ suggests the formation of compact clusters with a size increasing with protein concentration.

Finally, our results are consistent with the picture that BLG dimers constitute the elementary building blocks of the observed clusters. Notably, these dimers consistently model the SAXS form factor in the Two-Yukawa fits. On a higher level, the assembly of the dimers within the clusters may be structurally disordered, because a specific arrangement of dimers into clusters would induce additional features such as shoulders in the SAXS profiles beyond the observed dimeric characteristics.

We note that studies of the dynamic properties of protein clusters and the details of the static structure factor for a system with competing interaction potentials have shown that a reduced power-law behavior $q_c \propto d_{cc}^{-1} \propto \varphi^b$ with $b < 1/3$ alone is not a sufficient condition for the existence of equilibrium, i.e., static protein clusters. Rather, an intermediate-range order may exist in a system with long-range repulsion and short-range attraction, which could obscure an unambiguous discrimination of static or dynamic clusters.[2] However, our additional dynamic experiment confirms the existence of static clusters for our system at least on the NSE time scale on the order of 50 ns.

For BLG, previous studies by Piazza and Iacopini[39] on BLGA suggest that the proteins form oligomer-type clusters. Their results using static and dynamic light scattering indicate that these clusters have a limited lifetime on the microsecond observation time scale of DLS and should be considered as "transient". These findings are consistent with PFG-NMR results on millisecond time scales by Le Bon et al.,[40] which concluded that BLG self-diffusion in concentrated solutions is in agreement with dimers. We remark that these results on these long observation time scale do not contradict our results for BLGA/BLGB mixtures on a much shorter observation scale.

In conclusion, we have studied the formation of BLG clusters both using static (SAXS) and neutron spectroscopic (NSE and NBS) methods. By combining the results from these static and dynamic methods, we obtain a conclusive picture of the presence of clusters in BLG solutions that are static on the accessible observation time scale. The SAXS data are consistent with a disordered assembly of BLG dimers that form compact clusters. The experimental SAXS structure factor agrees with a model of short-ranged attraction, and long-ranged repulsion, consistent with the presence of clusters. The spectroscopic data from both the NBS and NSE experiments, respectively, yield equal diffusion coefficients at the overlapping q within the errors. These diffusion coefficients are in agreement with static BLG clusters on the experimental observation time scale (coherence time) of up to 50 ns in the case of NSE. Moreover, the NSE data are consistent with the SAXS data. The combination of the SAXS and NSE results points to the general possibility to refine the analysis of static SAXS data by employing complementary NSE experiments. Our analysis shows that the combination of the SAXS, NSE, and NBS results provides a robust quantitative picture of the cluster size and compactness depending on the protein concentration in the solution, a criterion for the choice of the SAXS model, and information on the cluster lifetime.

METHODS

Sample Preparation. β-Lactoglobulin (BLG) from bovine milk was purchased from Sigma-Aldrich (L3908, guaranteed purity of 90%) and used without further purification. This product is a mixture of the genetic variants A and B that differ at two positions in the primary sequence of 162 amino acids in total.[24] Under physiological conditions, BLG is found predominantly as a dimer.[24] With an isoelectric point (pI) of 5.2, BLG is acidic and carries a net charge of −10 e at neutral pH.[41]

All samples were prepared by mixing the required amount of protein stock solution with D_2O (for NSE and NBS) or deionized degassed Millipore H_2O (for SAXS). The pH of the protein solutions was monitored using a Seven Easy pH meter from Mettler Toledo. Freshly degassed Millipore water has a pH of 7.1, and the resulting BLG solutions display a pH of about 7.2 independent from protein concentration.

The protein volume fraction for a given protein concentration $c_p = m_p/V_s$ is given by[35]

$$\varphi = \frac{c_p}{1 + c_p \cdot \theta_p} \cdot \theta_p \tag{4}$$

Small-Angle X-ray Scattering. SAXS measurements were performed at the European Synchrotron Radiation Facility (ESRF), Grenoble, France, at beamline ID02. Protein solutions as well as H_2O were measured using a flow-through capillary cell with a wall thickness of about 10 μm at a sample-to-detector distance of 2 m. The beam energy was set to 16 keV, and the accessed q values ranged from 0.006 to 0.34 Å^{-1}. For each sample, 10 exposures of 0.1 s or 20 exposures of 0.05 s each were measured on fresh sample solution, respectively. The 2D intensity pattern was corrected and azimuthally averaged to obtain the intensity profiles. Finally, the solvent background was subtracted. More detailed information on data reduction and q-resolution calibration can be found in the literature.[42,43]

Additional SAXS data were collected on a laboratory source SAXS instrument—Xeuss 2.0 (Xenocs, Grenoble, France)—employing a GeniX 3D microfocus X-ray tube consisting of a copper anode, using an X-ray wavelength of 1.54 Å. With a sample-to-detector distance of 1.85 m, the employed Pilatus 300 K detector covered a q range of 0.0055 to 0.3 Å^{-1}. Protein solutions as well as the buffer (D_2O) were measured in a flow-through quartz capillary with a wall thickness of about 10 μm.

DOI: 10.1021/acs.jpclett.7b00658
J. Phys. Chem. Lett. 2017, 8, 2590−2596

Experimental structure factors (examples shown in Figure 1B) were calculated from the SAXS intensity by dividing with an ellipsoid form factor with rotation axis R_a = 38 Å and axis R_b = 19 Å.[44] For structure factor fits, we used the Two-Yukawa potential and the Hayter-Penfold mean spherical approximation (HPMSA) for the screened Coulomb potential as implemented in IGOR.[45] For the Two-Yukawa potential, an ellipsoid form factor with R_a = 38 Å and R_b = 19 Å was used. For the screened Coulomb potential, we employed an ellipsoid form factor with R_b = 19 Å, and R_a was allowed to take values between 39 and 55 Å.

In general, coarse-grained form factors have been shown to potentially induce incorrect features.[46] The ellipsoidal form factor is in very good agreement with experimental data and calculated profiles from the molecular structure up to 0.15 Å^{-1} (cf. Figure S3). Furthermore, similar results for q_c are obtained for the average structure factor and the decoupling approximation, and our conclusions are thus robust.

Neutron Spectroscopy. The neutron spin−echo measurements were performed at IN15 at Institut Laue-Langevin (ILL), Grenoble, using the wavelengths of 8 and 12 Å and detector angles between 3.5 and 12.5° covering a q-range from 0.025 to 0.19 Å^{-1} and Fourier times of up to 190 ns at 12 Å and 58 ns at 8 Å. Details of the experiment setup are explained elsewhere.[47,48] The samples on IN15 were kept in quartz cuvettes at T = 295 K. For calibration and background treatment, a graphite sample and the pure buffer (D_2O) signal were used.

The neutron backscattering experiment was carried out using the spectrometer IN16B[49] (ILL) employing a Phase Space Transformer[50] and using Si(111) single crystals for the monochromator and analyzers in exact backscattering geometry, corresponding to a wavelength of 6.27 Å. The energy was scanned by mechanically Doppler shifting the monochromator crystal on a sinus velocity profile, achieving a maximum energy transfer of approximately ±30 μeV. The nearly perfect Gaussian resolution line shape had a width of approximately 0.9 μeV FWHM. The experiment was carried out using the measurement, calibration, and data reduction procedures described earlier.[31] The fit function for the NBS spectra consists of two Lorentzian profiles convoluted with the instrumental resolution (dash-dotted and dashed line in Figure 2B), plus a fixed D_2O background.[31,32]

Apparent Diffusion Coefficient D. Using the NBS results, we obtain the apparent self-diffusion coefficient D from $\gamma(q) = Dq^2$ (cf. Figure 2B). The apparent D, translational D_t, and rotational D_r diffusion coefficients are related via an implicit function[32,35]

$$\sum_{l=0}^{\infty} B_l(q)\frac{D_r l(l+1) + (D_t - D)q^2}{[D_r l(l+1) + (D_t + D)q^2]^2} = 0 \quad (5)$$

Using an effective sphere with hydrodynamic radius R_h, the translational and rotational self-diffusion are given by $D_t = f(\tilde{\varphi})k_BT/(6\pi\eta R_h)$ and $D_r = g(\tilde{\varphi})k_BT/(8\pi\eta R_h^3)$, respectively. The dependencies on the concentration, $f(\tilde{\varphi})$ and $g(\tilde{\varphi})$, are the known functions for the translational[51] and rotational[33] self-diffusion of colloidal hard spheres. The effective cluster volume fraction reads $\tilde{\varphi} = n_{cluster}\frac{4}{3}\pi R_h^3$, where we used q_c from the Two-Yukawa fits to calculate $n_{cluster}$ from eq 1. To calculate $B_l(q)$ analytically, we model the radial distribution function by $\rho_H(r) = \rho_0 \cdot 4\pi \cdot r^2 \cdot \Theta(R_h - r)$, where ρ_0 = 0.0415 Å^{-3} is obtained from the actual hydrogen density distribution in the BLG monomer (Figure S4). We obtain (with $x = qR_h$)

$$B_l(q) = (2l+1)\rho_0 2\pi x^3[j_l^2(x) - j_{l+1}(x)j_{l-1}(x)] \quad (6)$$

We solve eq 5 with an upper summation limit of l_{max} = 230 numerically using the above assumptions, yielding R_h as the sole fit parameter for each protein concentration.

The hydrodynamic radius R_h of dimers has been obtained from the dilute limit of the translational diffusion and calculated using HYDROPRO,[36] which is in good agreement with experiments for BLG.[39]

■ ASSOCIATED CONTENT

Supporting Information

The neutron data are permanently curated by the ILL and accessible under http://doi.ill.fr/10.5291/ILL-DATA.8-04-724 and http://doi.ill.fr/10.5291/ILL-DATA.8-04-754. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpclett.7b00658.

Further SAXS data and analysis, tables of the measured samples and fit parameters. (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: felix.roosen-runge@fkem1.lu.se.
*E-mail: fajun.zhang@uni-tuebingen.de.
*E-mail: seydel@ill.eu.

ORCID

Michal K. Braun: 0000-0002-9642-7492
Felix Roosen-Runge: 0000-0001-5106-4360
Fajun Zhang: 0000-0001-7639-8594
Tilo Seydel: 0000-0001-9630-1630

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG) and the Knut and Alice Wallenberg Foundation. We are grateful to B. Farago, P. Falus (ILL), C. Beck (ILL and Tübingen) and S. Da Vela (Tübingen) for stimulating discussions, and to R. Ammer (ILL) for technical support. We thank T. Stehle and C. Schall (Tübingen) for sharing lab resources, and assistance during characterization experiments. Furthermore, we thank the ILL and ESRF for allocation of beamtime.

■ REFERENCES

(1) Porcar, L.; Falus, P.; Chen, W.-R.; Faraone, A.; Fratini, E.; Hong, K.; Baglioni, P.; Liu, Y. Formation of the Dynamic Clusters in Concentrated Lysozyme Protein Solutions. *J. Phys. Chem. Lett.* **2010**, *1*, 126−129.

(2) Liu, Y.; Porcar, L.; Chen, J.; Chen, W.-R.; Falus, P.; Faraone, A.; Fratini, E.; Hong, K.; Baglioni, P. Lysozyme Protein Solution with an Intermediate Range Order Structure. *J. Phys. Chem. B* **2011**, *115*, 7238−7247.

(3) Dinsmore, A. D.; Dubin, P. L.; Grason, G. M. Clustering in Complex Fluids. *J. Phys. Chem. B* **2011**, *115*, 7173−7174.

(4) Johnston, K. P.; Maynard, J. A.; Truskett, T. M.; Borwankar, A. U.; Miller, M. A.; Wilson, B. K.; Dinin, A. K.; Khan, T. A.; Kaczorowski, K. J. Concentrated Dispersions of Equilibrium Protein Nanoclusters that Reversibly Dissociate into Active Monomers. *ACS Nano* **2012**, *6*, 1357−1369.

(5) Yearley, E.; Godfrin, P.; Perevozchikova, T.; Zhang, H.; Falus, P.; Porcar, L.; Nagao, M.; Curtis, J.; Gawande, P.; Taing, R.; et al. Observation of Small Cluster Formation in Concentrated Monoclonal

DOI: 10.1021/acs.jpclett.7b00658
J. Phys. Chem. Lett. 2017, 8, 2590−2596

Antibody Solutions and its Implications to Solution Viscosity. *Biophys. J.* **2014**, *106*, 1763−1770.
(6) Bucciarelli, S.; Myung, J. S.; Farago, B.; Das, S.; Vliegenthart, G. A.; Holderer, O.; Winkler, R. G.; Schurtenberger, P.; Gompper, G.; Stradner, A. Dramatic Influence of Patchy Attractions on Short-Time Protein Diffusion under Crowded Conditions. *Sci. Adv.* **2016**, *2*, e1601432.
(7) Bucciarelli, S.; Casal-Dujat, L.; De Michele, C.; Sciortino, F.; Dhont, J.; Bergenholtz, J.; Farago, B.; Schurtenberger, P.; Stradner, A. Unusual Dynamics of Concentration Fluctuations in Solutions of Weakly Attractive Globular Proteins. *J. Phys. Chem. Lett.* **2015**, *6*, 4470−4474.
(8) Vekilov, P. G. The Two-step Mechanism of Nucleation of Crystals in Solution. *Nanoscale* **2010**, *2*, 2346−2357.
(9) Sauter, A.; Roosen-Runge, F.; Zhang, F.; Lotze, G.; Jacobs, R. M. J.; Schreiber, F. Real-Time Observation of Nonclassical Protein Crystallization Kinetics. *J. Am. Chem. Soc.* **2015**, *137*, 1485−1491.
(10) Grimaldo, M.; Roosen-Runge, F.; Hennig, M.; Zanini, F.; Zhang, F.; Zamponi, M.; Jalarvo, N.; Schreiber, F.; Seydel, T. Salt-Induced Universal Slowing Down of the Short-Time Self-Diffusion of a Globular Protein in Aqueous Solution. *J. Phys. Chem. Lett.* **2015**, *6*, 2577−2582.
(11) Cardinaux, F.; Zaccarelli, E.; Stradner, A.; Bucciarelli, S.; Farago, B.; Egelhaaf, S. U.; Sciortino, F.; Schurtenberger, P. Cluster-Driven Dynamical Arrest in Concentrated Lysozyme Solutions. *J. Phys. Chem. B* **2011**, *115*, 7227−7237.
(12) Stradner, A.; Sedgwick, H.; Cardinaux, F.; Poon, W. C. K.; Egelhaaf, S. U.; Schurtenberger, P. Equilibrium Cluster Formation in Concentrated Protein Solutions and Colloids. *Nature* **2004**, *432*, 492−495.
(13) Stradner, A.; Cardinaux, F.; Schurtenberger, P. A Small-Angle Scattering Study on Equilibrium Clusters in Lysozyme Solutions. *J. Phys. Chem. B* **2006**, *110*, 21222−31.
(14) Shukla, A.; Mylonas, E.; Di Cola, E.; Finet, S.; Timmins, P.; Narayanan, T.; Svergun, D. I. Absence of Equilibrium Cluster Phase in Concentrated Lysozyme Solutions. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 5075−5080.
(15) Barhoum, S.; Yethiraj, A. NMR Detection of an Equilibrium Phase Consisting of Monomers and Clusters in Concentrated Lysozyme Solutions. *J. Phys. Chem. B* **2010**, *114*, 17062−17067.
(16) Liu, Y.; Wang, X.; Ching, C. B. Toward Further Understanding of Lysozyme Crystallization: Phase Diagram, Protein-Protein Interaction, Nucleation Kinetics, and Growth Kinetics. *Cryst. Growth Des.* **2010**, *10*, 548−558.
(17) Sciortino, F.; Mossa, S.; Zaccarelli, E.; Tartaglia, P. Equilibrium Cluster Phases and Low-Density Arrested Disordered States: The Role of Short-Range Attraction and Long-Range Repulsion. *Phys. Rev. Lett.* **2004**, *93*, 055701.
(18) Campbell, A. I.; Anderson, V. J.; van Duijneveldt, J. S.; Bartlett, P. Dynamical Arrest in Attractive Colloids: The Effect of Long-Range Repulsion. *Phys. Rev. Lett.* **2005**, *94*, 208301.
(19) Hutchens, S. B.; Wang, Z.-G. Metastable Cluster Intermediates in the Condensation of Charged Macromolecule Solutions. *J. Chem. Phys.* **2007**, *127*, 084912.
(20) Pan, W.; Galkin, O.; Filobelo, L.; Nagel, R. L.; Vekilov, P. G. Metastable Mesoscopic Clusters in Solutions of Sickle-Cell Hemoglobin. *Biophys. J.* **2007**, *92*, 267−277.
(21) Jadrich, R. B.; Bollinger, J. A.; Lindquist, B. A.; Truskett, T. M. Equilibrium Cluster Fluids: Pair Interactions via Inverse Design. *Soft Matter* **2015**, *11*, 9342−9354.
(22) Segre, P.; Pusey, P. Scaling of the Dynamic Scattering Function of Concentrated Colloidal Suspensions. *Phys. Rev. Lett.* **1996**, *77*, 771.
(23) Holmqvist, P.; Nägele, G. Long-time Dynamics of Concentrated Charge-Stabilized Colloids. *Phys. Rev. Lett.* **2010**, *104*, 058301.
(24) Verheul, M.; Pedersen, J. S.; Roefs, S. P. F. M.; De Kruif, K. G. Association Behavior of Native β-Lactoglobulin. *Biopolymers* **1999**, *49*, 11−20.
(25) Zhang, F.; Roosen-Runge, F.; Sauter, A.; Roth, R.; Skoda, M. W.; Jacobs, R. M.; Sztucki, M.; Schreiber, F. The Role of Cluster Formation and Metastable Liquid-Liquid Phase Separation in Protein Crystallization. *Faraday Discuss.* **2012**, *159*, 313−325.
(26) Häußler, W.; Wilk, A.; Gapinski, J.; Patkowski, A. Interparticle Correlations Due to Electrostatic Interactions: A Small Angle X-ray and Dynamic Light Scattering Study. I. Apoferritin. *J. Chem. Phys.* **2002**, *117*, 413.
(27) Häußler, W. Neutron Spin Echo Studies on Ferritin: Free-Particle Diffusion and Interacting Solutions. *Eur. Biophys. J.* **2008**, *37*, 563−571.
(28) Baglioni, P.; Fratini, E.; Lonetti, B.; Chen, S. H. Structural Arrest in Concentrated Cytochrome C Solutions: The Effect of pH and Salts. *J. Phys.: Condens. Matter* **2004**, *16*, S5003.
(29) Roosen-Runge, F.; Hennig, M.; Seydel, T.; Zhang, F.; Skoda, M. W. A.; Zorn, S.; Jacobs, R. M. J.; Maccarini, M.; Fouquet, P.; Schreiber, F. Protein Diffusion in Crowded Electrolyte Solutions. *Biochim. Biophys. Acta, Proteins Proteomics* **2010**, *1804*, 68−75.
(30) Lee, J. C.; Timasheff, S. N. Partial Specific Volumes and Interactions with Solvent Components of Proteins in Guanidine Hydrochloride. *Biochemistry* **1974**, *13*, 257−265.
(31) Grimaldo, M.; Roosen-Runge, F.; Zhang, F.; Seydel, T.; Schreiber, F. Diffusion and Dynamics of γ-Globulin in Crowded Aqueous Solutions. *J. Phys. Chem. B* **2014**, *118*, 7203−7209.
(32) Grimaldo, M.; Roosen-Runge, F.; Jalarvo, N.; Zamponi, M.; Zanini, F.; Hennig, M.; Zhang, F.; Schreiber, F.; Seydel, T. High-Resolution Neutron Spectroscopy on Protein Solution Samples. *EPJ Web Conf.* **2015**, *83*, 02005.
(33) Banchio, A. J.; Nägele, G. Short-time Transport Properties in Dense Suspensions: From Neutral to Charge-Stabilized Colloidal Spheres. *J. Chem. Phys.* **2008**, *128*, 104903.
(34) Gapinski, J.; Wilk, A.; Patkowski, A.; Häussler, W.; Banchio, A.; Pecora, R.; Nägele, G. Diffusion and Microstructural Properties of Solutions of Charged Nanosized Proteins: Experiment versus Theory. *J. Chem. Phys.* **2005**, *123*, 054708.
(35) Roosen-Runge, F.; Hennig, M.; Zhang, F.; Jacobs, R. M. J.; Sztucki, M.; Schober, H.; Seydel, T.; Schreiber, F. Protein Self-diffusion in Crowded Solutions. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 11815−11820.
(36) Ortega, A.; Amorós, D.; Garcia de la Torre, J. Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins From Atomic- and Residue-Level Models. *Biophys. J.* **2011**, *101*, 892−898.
(37) *The PyMOL Molecular Graphics System*, version 1.7; Schrödinger, LLC: New York.
(38) Groenewold, J.; Kegel, W. K. Anomalously Large Equilibrium Clusters of Colloids. *J. Phys. Chem. B* **2001**, *105*, 11702−11709.
(39) Piazza, R.; Iacopini, S. Transient Clustering in a Protein Solution. *Eur. Phys. J. E: Soft Matter Biol. Phys.* **2002**, *7*, 45−48.
(40) Le Bon, C.; Nicolai, T.; Kuil, M. E.; Hollander, J. G. Self-Diffusion and Cooperative Diffusion of Globular Proteins in Solution. *J. Phys. Chem. B* **1999**, *103*, 10294−10299.
(41) Elofsson, U. M.; Paulsson, M. A.; Arnebrant, T. Adsorption of β-Lactoglobulin A and B in Relation to Self-Association: Effect of Concentration and pH. *Langmuir* **1997**, *13*, 1695−1700.
(42) Narayanan, T. Synchroton Small-Angle X-ray Scattering. In *Soft Matter Characterization*; Borsali, R., Pecora, R., Eds.; Springer: New York, 2008; Chapter 17.
(43) Ianeselli, L.; Zhang, F.; Skoda, M. W. A.; Jacobs, R. M. J.; Martin, R. A.; Callow, S.; Prévost, S.; Schreiber, F. Protein-Protein Interactions in Ovalbumin Solutions Studied by Small-Angle Scattering: Effect of Ionic Strength and the Chemical Nature of Cations. *J. Phys. Chem. B* **2010**, *114*, 3776−3783.
(44) Sauter, A.; Oelker, M.; Zocher, G.; Zhang, F.; Stehle, T.; Schreiber, F. Nonclassical Pathways of Protein Crystallization in the Presence of Multivalent Metal Ions. *Cryst. Growth Des.* **2014**, *14*, 6357−6366.
(45) Kline, S. R. Reduction and Analysis of SANS and USANS Data Using IGOR Pro. *J. Appl. Crystallogr.* **2006**, *39*, 895−900.
(46) Castellanos, M. M.; Clark, N. J.; Watson, M. C.; Krueger, S.; McAuley, A.; Curtis, J. E. Role of Molecular Flexibility and Colloidal

DOI: 10.1021/acs.jpclett.7b00658
J. Phys. Chem. Lett. 2017, 8, 2590−2596

Descriptions of Proteins in Crowded Environments From Small-Angle Scattering. *J. Phys. Chem. B* **2016**, *120*, 12511−12518.
(47) Farago, B.; Falus, P.; Hoffmann, I.; Gradzielski, M.; Thomas, F.; Gomez, C. The IN15 Upgrade. *Neutron News* **2015**, *26*, 15−17.
(48) Schleger, P.; Alefeld, B.; Barthelemy, J.; Ehlers, G.; Farago, B.; Giraud, P.; Hayes, C.; Kollmar, A.; Lartigue, C.; Mezei, F.; et al. The Long-Wavelength Neutron Spin-Echo Spectrometer IN15 at the Institut Laue-Langevin. *Phys. B* **1997**, *241-243*, 164−165.
(49) Frick, B.; Mamontov, E.; van Eijck, L.; Seydel, T. Recent Backscattering Instrument Developments at the ILL and SNS. *Z. Phys. Chem.* **2010**, *224*, 33−60.
(50) Hennig, M.; Frick, B.; Seydel, T. Optimum Velocity of a Phase-Space Transformer for Cold-Neutron Backscattering Spectroscopy. *J. Appl. Crystallogr.* **2011**, *44*, 467−472.
(51) Tokuyama, M.; Oppenheim, I. Dynamics of Hard-Sphere Suspensions. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1994**, *50*, R16−R19.

DOI: 10.1021/acs.jpclett.7b00658
J. Phys. Chem. Lett. 2017, 8, 2590−2596