Salt-Induced Universal Slowing Down of the Short-Time Self-Diffusion of a Globular Protein in Aqueous Solution

Marco Grimaldo,†‡ Felix Roosen-Runge,† Marcus Hennig,†‡ Fabio Zanini,‡§ Fajun Zhang,‡ Michaela Zamponi,§‖ Niina Jalario,§‖ Frank Schreiber,‡ and Tilo Seydel‡

†Institut Max von Laue - Paul Langevin (ILL), CS 20156, 71 avenue des Martyrs, F-38042 Grenoble, France
‡Institut für Angewandte Physik, Universität Tübingen, Auf der Morgenstelle 10, D-72076 Tübingen, Germany
§Jülich Centre for Neutron Science (JCNS), Forschungszentrum Jülich GmbH, D-52425 Jülich, Germany
‖JCNS Outstation at the MLZ, Lichtenbergstraße 1, D-85747 Garching, Germany

Supporting Information

ABSTRACT: The short-time self-diffusion $D$ of the globular model protein bovine serum albumin in aqueous (D$_2$O) solutions has been measured comprehensively as a function of the protein and trivalent salt (YCl$_3$) concentration, noted $c_p$ and $c_s$ respectively. We observe that $D$ follows a universal master curve $D(c_p,c_s)/D(c_p=0,c_s)$ $g(c_s/c_p)$, where $D(c_p=0,c_s)$ is the diffusion coefficient in the absence of salt and $g(c_s/c_p)$ is a scalar function solely depending on the ratio of the salt and protein concentration. This observation is consistent with a universal scaling of the bonding probability in a picture of cluster formation of patchy particles. The finding corroborates the predictive power of the description of proteins as colloids with distinct attractive ion-activated surface patches.

Diffusion is a ubiquitous process in a broad variety of condensed matter systems. In the particular case of the biological cell, protein self-diffusion (i.e., the diffusion of single “tagged” protein macromolecules) governs, for example, transport processes, reaction kinetics, and information exchange in the biological cell. Importantly, self-diffusion is affected by single-particle properties as well as by environmental factors. One well-known environmental control parameter for self-diffusion is macromolecular crowding, that is, the dense packing fraction at which macromolecules such as proteins are usually found in the aqueous intracellular environment (refs 4–6 and references therein). The motion of proteins is affected by two kinds of interactions with surrounding macromolecules: direct interactions such as excluded volume and Coulomb interaction influence the dynamics on time scales of roughly microseconds and above. On shorter time scales, for example, on the order of several nanoseconds, proteins interact only through quasi-instantaneous hydrodynamic interactions.5

The presence of salts is another important control parameter that influences both structure and dynamics of protein solutions as well as, for example, DNA solutions, and may lead to the formation of protein aggregates. The interplay of macromolecular concentration and salts, physiological ones, and others to control condensation phenomena in protein solutions represents an important question with applications in medicine and biotechnology. Protein condensation is observed in several diseases, such as some forms of cataracts and neurodegenerative diseases.5 In Parkinson’s disease, raised levels of multivalent iron ions have been observed in sick neurons, pointing to a yet-to-be-explored role of salt. A deep understanding of the aggregation process of proteins is also of high interest for the tunable production of nanostructures through self-assembly. The formation of clusters in dependence on protein concentration has been previously studied using small-angle scattering (SAS), nuclear magnetic resonance (NMR), dynamic light scattering (DLS), and neutron spin-echo spectroscopy (NSE).

Although studied since the pioneering work of Hofmeister more than a century ago, the effect of salt ions on the phase behavior of protein solutions presents still numerous questions. For the particular case of acidic proteins, trivalent salts such as FeCl$_3$ and YCl$_3$ induce complex reentrant phase diagrams depending on the protein concentration, $c_p$, and salt concentration, $c_s$, including macroscopic aggregation, liquid–liquid phase separation (LLPS), and protein crystallization and offering promising ways to manipulate the behavior of proteins. On the basis of the complex phase
behavior, it is tempting to speculate that dynamic precursor processes of protein aggregation and crystallization occur already at lower trivalent salt concentrations. Such precursor processes may be the formation of small static or transient protein clusters. Indeed, our group has also explored the formation of clusters due to the presence of multivalent salts by SAS \textsuperscript{19} and DLS.\textsuperscript{11}

Three factors render short-time self-diffusion a promising quantity to probe clustering of proteins as controlled by salt. First, the short-time self-diffusion of proteins can be described remarkably well by the theory of diffusion for effective hard spheres,\textsuperscript{6,30} which links experimental results to a physical picture only based on the diffusional tensor of the single particle and an effective viscosity of the environment. In particular, the shift of structural correlation peaks due to a varying cluster size does not influence the interpretation. Second, the addition of salt efficiently alters direct interactions such as Coulomb repulsion, while hydrodynamic interactions remain comparable, which allows us to study the nontrivial interplay of protein and salt concentration. Third, the short observation time scale much smaller than the collision time $t_c \approx 200 \text{ ns}$ ensures that the effects of particle collisions and specific interactions with obstacles do not affect the observed diffusion. This short-time scale is routinely accessible by neutron spectroscopy as opposed to NMR, DLS, and fluorescence correlation (FCS) techniques.

In this context, incoherent quasi-elastic neutron backscattering (QENS) is an essential tool because it unambiguously accesses the self-diffusion of globular proteins on a nanometer length-scale and a nanosecond time scale, even for crowded and possibly opaque samples.\textsuperscript{7} Employing QENS to systematically investigate the protein short-time self-diffusion can hence provide crucial information on the presence or absence of clusters with a lifetime down to nanoseconds.

Spectra of protein solutions as a function of the energy transfer, $h\omega$, and the scattering vector, $q$, were recorded at the backscattering spectrometer BASIS\textsuperscript{31} (SNS, ORNL) and fitted with the model described in detail elsewhere$^{6,32,33}$

$$S(q, \omega) = \mathcal{R}(q, \omega) \otimes \{ \beta(q) \cdot [A_0(q) \mathcal{L}(\gamma(q), \omega)] + (1 - A_0(q)) \mathcal{L}(\Gamma(q), \omega) \} + \beta_{D,O}(q) \mathcal{L}(\gamma_{D,O}(q), \omega)$$

(1)

Therein, $\mathcal{R}(q, \omega)$ represents the energy resolution function, and $\mathcal{L}(\bullet, \omega)$ stands for Lorentzian functions with a line width related to dynamical processes in the sample. $\gamma(q)$ describes the global diffusion of the proteins, consisting of both translational and rotational contributions. $\Gamma(q)$ represents the internal relaxations accessed by the dynamic window of the spectrometer and convoluted with the global diffusion. $\beta(q)$ and $A_0(q)$ are scalars. The contribution of the solvent is modeled by the fixed term $\beta_{D,O} \mathcal{L}(\gamma_{D,O}(q), \omega)$, as explained elsewhere.\textsuperscript{33} To increase the contrast between the hydrogenated proteins and the surrounding solvent, we use D$_2$O rather than H$_2$O, as commonly done in neutron scattering by virtue of the large difference of the incoherent neutron scattering cross sections of the two hydrogen isotopes.

An example spectrum is shown in the inset of Figure 1. The red circles represent the recorded spectrum, the blue dotted-dashed line depicts $\mathcal{L}(\gamma, \omega)$, and the green dashed line represents $\mathcal{L}(\Gamma, \omega)$. Finally, the turquoise solid line superimposed on the data is the result of the fit using eq 1. The widths $\gamma(q)$ obtained from the fit function in Figure 1 are plotted in the main part of Figure 1. The solid red line denotes the fit with the Fickian law $\gamma(q) = D q^2$, where $D$ is the global apparent self-diffusion coefficient of the proteins.\textsuperscript{6,30}

Figure 1. Inset: Example backscattering spectrum $I(q, \omega)$ (red circles) recorded at BASIS for BSA and YCl$_3$ in D$_2$O ($c_p = 150 \text{ mg/mL}$, $c_s = 8 \text{ mM}$, $T = 295 \text{ K}$, individual detector at $q = 0.65 \AA^{-1}$). The blue dotted-dashed line depicts $\mathcal{L}(\gamma, \omega)$ and the green dashed line pictures $\mathcal{L}(\Gamma, \omega)$ from eq 1. The turquoise solid line superimposed on the data is the result of the fit using eq 1. The widths $\gamma(q)$ obtained from the fit function in Figure 1 are plotted in the main part of Figure 1. The solid red line denotes the fit with the Fickian law $\gamma(q) = D q^2$, where $D$ is the global apparent self-diffusion coefficient of the proteins.\textsuperscript{6,30}

In this study, we explore the global diffusion coefficient $D$ in solutions of the globular protein bovine serum albumin (BSA) depending on its concentration $c_p$ and on the concentration $c_s$ of the salt YCl$_3$ (cf. dots in Figure 2b). For several fixed protein volume fractions of approximately 7–20%, we increase $c_s$ while remaining on an area of the phase diagram where no macroscopic aggregation or phase transition occurs.\textsuperscript{24}

Strikingly, when plotting the reduced self-diffusion coefficient $D(c_p c_s)/D(c_p = 0, c_s)$ as a function of the number of Y$^{3+}$ ions per protein $N_c = c_s/c_p$, in Figure 2a, we obtain a universal scaling for all measured protein concentrations. The master curve decreases continuously to 60% of the value with no added salt, when in the solution there are approximately six Y$^{3+}$ ions per protein. Essentially, this curve characterizes the diffusion of BSA in the presence of YCl$_3$ in the entire explored range of $c_p$ and $c_s$. The dashed line in Figure 2a represents an empirical function for the master curve

$$g(\epsilon_c / \epsilon_p) = 1 + a_3 \left( \frac{\epsilon_c}{\epsilon_p} \right)^2 + a_4 \left( \frac{\epsilon_c}{\epsilon_p} \right)^4$$

(2)

with the parameters $a_3 = (-17.32 \pm 2.76) \times 10^{-3}$ and $a_4 = (18.86 \pm 8.93) \times 10^{-3}$.

Having established this rather remarkable universal behavior, we now turn to its possible origin. There may be several rather complex ways to model the salt dependence. To keep the model simple and limit the number of parameters, we consider a salt-induced cluster formation, inspired by the theory of ion-activated patchy particles,\textsuperscript{34,35} which was successfully used to semi-quantitatively describe the phase diagram of the system of BSA and YCl$_3$. In this physical picture, Y$^{3+}$ ions bind to patches on the particle surface and thereby activate them for subsequent bridging between protein molecules. Thus, increasing the ions imposed on the data is the result of the fit using eq 1. The widths $\gamma(q)$ obtained from the fit function in Figure 1 are plotted in the main part of Figure 1. The solid red line denotes the fit with the Fickian law $\gamma(q) = D q^2$, where $D$ is the global apparent self-diffusion coefficient of the proteins.\textsuperscript{6,30}
in solution increases the probability $p_b$ of an ion-bridge between particles and so the formation of clusters.

The Flory–Stockmeyer theory provides a first-order estimation of the number density $\rho_b^s$ of n-clusters:\footnote{36,37}

$$\rho_b^s = (1 - p_b)^m [p_b (1 - p_b)^{m-2} - 1] \frac{m(m-n)!}{(m-n+2)!}!$$

where $m$ is the maximum number of ion-bridges per particle and $n$ is the number of particles in the n-cluster. We chose $m = 4$ consistent with ref 35, inspired by the crystal structure displaying four ion bridges per monomer.26

To compare our data with this theory we assume that at the short times probed with BASIS the clusters are in good approximation rigid. Thus, movements of proteins within the clusters are neglected. Given the assumed strong bonds, these motions might occur on similar time scales as, for example, protein interdomain motions\footnote{38,39} beyond the observation time scale of our experiment. Moreover, we assume that each n-cluster diffuses according to the theory of colloidal hard-sphere suspensions as if they were in monodisperse suspensions at the actual volume fraction of the protein solution, that is, for every cluster size $n$

$$D_i(\phi) = D_i^{(1)}(\phi = 0) n^{-\alpha} f(j)$$

$$D_i(\phi) = D_i^{(1)}(\phi = 0) n^{-\alpha} f(j)$$

where $D_i^{(n)}$ and $D_i^{(n)}$ are the translational and rotational diffusion coefficients of the n-cluster, $D_i^{(1)}(\phi = 0) = k_B T/(6 \pi \eta R_p)$, $D_i^{(1)}(\phi = 0) = k_B T/(8 \pi \eta R_p^3)$, and $R_p$ denotes, respectively, the dilute limit diffusion coefficients and the effective hydrodynamic radius of the monomers. $\eta$ represents the viscosity of the solvent and $f_j(\phi)$ is the theoretical reduced diffusion coefficient for hard-sphere suspensions.\footnote{39,40} Lacking a detailed description of the microscopic structure of the clusters, we choose $\nu = 1/3$ and $\alpha = 1$, valid in the limit of low fractality. Thus, the hydrodynamic radius of an n-cluster scales as that of a hard sphere of volume $nV_p$, where $V_p$ is the volume of a monomer.

The dynamic structure factor of the global diffusion is modeled for the cluster solution by

$$S_f(q, \omega) = s \sum_n \rho_b^s(p_b) n S_b(q, \omega)$$

Here $s$ is a scalar, and the well-known scattering function for a single n-cluster $S_b(q, \omega)$ is provided in eq 7. Fitting of eq 6 to the experimental $L(g, \omega)$ obtained from eq 1 returns the two free parameters $s$ and $p_b$. Interestingly, $p_b$ for different protein concentrations, as derived from the master curve of the dynamics, follows itself a master curve (green line in the inset of Figure 2b). Thus, the model reflects the damping of the diffusion, provided that the ion-bridge probability $p_b$ depends only on $c_b/c_p$ and not on $\phi$. In other words, the observed universality of the dynamics would suggest that the contour lines of $p_b$, in the explored area on the $c_b$-$c_p$ phase-diagram are straight lines with intercept 0 and increasing slope (cf. Figure 2b). A possible reason for the universality could be the nearly quantitative binding of cations to the protein due to a dominating ion binding energy. Such a result is important for a further refined picture along the model of ion-activated attractive patches in ref 35, also in connection with the generalized law of correspondent states for patchy interactions.\footnote{41}

Importantly, this cluster picture could also explain previous results from light scattering in BSA solutions with lower volume fractions.\footnote{11} In these experiments, both reduced long-time collective diffusion and reduced isothermal compressibility showed a universal scaling with respect to $c_b/c_p$. With increasing $c_b$, the diffusion of both a monomer and a cluster fraction showed a substantial slowing down and the compressibility diverged. Using the ion-bridge probability derived from QENS, the DLS observations could be explained by the occurrence of increasingly large clusters with a size distribution universal in protein concentration because the dynamics slows down and the total scattering power for light diverges once the clusters are large enough (cf. figure 3 in ref 35).

We emphasize that our experiment does not allow us to draw a conclusion of whether the observed clusters are dynamic or transient (i.e., have a finite lifetime) or static.

On the nanosecond time scale explored by our experiment, our model explains the measured short-time self-diffusion by at least temporarily "rigid" clusters, similar to NSE results on the formation of lysozyme clusters induced by the increase in the protein concentration (i.e., not by additional salt), where the

Figure 2. (a) Reduced short-time apparent self-diffusion coefficient $D(c_s,c_p)/D(c_s = 0, c_p)$ as a function of $c_s/c_p$ (symbols). Independently from the protein concentration, all of the points lie on a master curve, which is parametrized by a polynomial function (dashed line, see text). (b) Contour plot of the bonding probability, $p_b$, as a function of $c_p$ and $c_s$. The dots denote the coordinates where experimental data were recorded. The area with the solid lines corresponds to the range of protein concentrations measured in this study. The dashed lines are extrapolations to lower and higher protein concentrations. Inset: Bonding probability, $p_b$, as a function of the number of $Y^{3+}$ ions per protein $c_s/c_p$ as obtained from the experiment under the assumptions explained in the text, using eqs 2 and 6. (See the Experimental Methods section.)
lifetime of the observed clusters is longer than a few tens of nanoseconds.\textsuperscript{22,42}

As an alternative scenario, the increase in attractive interactions between proteins due to bound Y\textsuperscript{3+} ions on the protein surface could enhance fluctuations of the local volume fraction $\varphi$ of monomers and thus decrease the averaged apparent self-diffusion coefficient without the formation of bonds between particles because most proteins would experience a denser packing. Under these assumptions, we would be able to describe our data assuming a well-separated bimodal distribution of $\varphi$; however, two points seem not to be consistent with a physical explanation along this scenario: (i) There is no obvious physical reason for a well-separated bimodal distribution $G(\varphi)$ away from a phase separation. (ii) There is no obvious physical reason why density fluctuations should produce a universal slowing down because this implies a very specific relation between the fluctuation amplitude and the overall protein and salt concentration. (See Figure S1 in the SI.)

In conclusion, we experimentally establish the existence of a remarkably universal slowing down of the short-time self-diffusion of the protein BSA as a function of the number of multivalent Yttrium ions per protein. The observation can be connected to previous results\textsuperscript{11} on a universal slowing down of the long-time collective diffusion of proteins in the same system (BSA + YCl\textsubscript{3} in solution) at lower protein concentration. In combination, the two results suggest that dynamics and thermodynamics in protein solutions can be tuned in a relatively universal manner for a broad range of protein concentration by the addition of multivalent cations.

Salt-induced cluster formation via ion bridges renders a clear physical explanation for the universality and strength of the slowing down, provided that the ion-bridge probability depends only on the number of ions per protein, as, for example, expected for nearly quantitative binding of cations to the protein surface, embedding well into a general qualitative understanding of dynamics and thermodynamics in protein solutions with multivalent cations by assuming ion-activated attractive patches.\textsuperscript{35}

Such a result is promising in several respects. First, similar results might also be observed in other systems with strong cross-linking. Second, if the observations were indeed to be explained with the formation of clusters distributed according to a remarkably universal cluster distribution, the finding would be extremely relevant for a rational choice of pathways for self-assembly.

## EXPERIMENTAL METHODS

### Sample Preparation

BSA and YCl\textsubscript{3} were purchased from Sigma-Aldrich with a purity of 98 (A3059) and 99.99% (451363), respectively. D\textsubscript{2}O was purchased from Acros Organics with an indicated purity of 99.8%. Samples were prepared as described in ref 6. After some hours on a roller mixer, all BSA powder was completely dissolved and the solutions appeared clear. For every protein concentration, a series of samples with increasing $c_\text{p}$ were prepared. The solutions were filled into double-walled aluminum cylinders (23 mm outer diameter, 0.15 mm gap, that is, difference between inner and outer radius), sealed against vacuum, and kept at $T = 295$ K for the measurements.

Quasielastic Neutron Backscattering. We used the backscattering spectrometer BASIS\textsuperscript{31} at the ORNL’s Spallation Neutron Source in Oak Ridge, Tennessee in the standard configuration with unpolished Si(111) analyzer crystals, which set the selected neutron wavelength to 6.27 Å. The energy resolution defined by the fwhm of the resolution function is $\sim 3.5$ μeV. The accessible energy transfer range is $-100 \leq \Delta E \leq +100$ μeV. The resulting accessible time range is thus

$$41 \text{ ps} \leq \tau = \frac{\hbar}{\Delta E} \leq 1.2 \text{ ns}$$

The $q$ range $0.25 \leq q \leq 1.95$ Å\textsuperscript{−1} sets the probed length scale $l = 2\pi/|q|$ to

$$25 \geq l \geq 3.22 \text{ Å}$$

The raw data were normalized to the incident neutron flux and relative detector efficiency. To remove the contribution of the sample holder from the spectra, we have used the Paalman–Pings coefficients, accounting for the $q$-dependent absorption of neutrons by the sample and cell walls. The details can be found in the Supporting Information.

### Determination of $p_\text{v}$

The scattering function of an $n$-cluster is given by

$$S_n(q, \omega) = \sum_{l=0}^{\infty} B_l(q) \frac{D_l^{(n)}(l + 1) + D_l^{(n)}(q^2)}{\omega^2 + [D_l^{(n)}(l + 1) + D_l^{(n)}(q^2)]^2}$$  (7)

Therein, $D_l^{(n)}$ and $D_l^{(n)}$ are defined by the eqs 4 and 5, respectively. $B_l(q)$ is defined as

$$B_l(q) = \int_0^{\infty} dr \rho(r)(2l + 1)j_l^2(qr)$$

where $\rho(r)$ describes the radial density distribution of the hydrogen atoms in the molecule and $j_l(x)$ denotes the $l$th-order spherical Bessel function of first kind. Under the assumption of homogeneously distributed H atoms in effective hard spheres, $\rho(r) = 4\pi r^2$. To obtain $p_\text{v}$, one can fit eqs 6 and 7 to a Lorentzian function $L(\gamma, \omega)$ with $\gamma$ defined by

$$\gamma = D(c_\text{p} = 0, c_\text{p})g(c_\text{p})q^2$$

de where $g(c_\text{p})$ denotes the master curve and $D(c_\text{p} = 0, c_\text{p})$ is the apparent diffusion coefficient of the solution at protein concentration $c_\text{p}$ and no added salt, which can be determined experimentally or calculated for an effective hard sphere by solving the equation

$$\sum_{l=0}^{\infty} B_l(q) \frac{D_l(l + 1) + (D_l - D_q)^2}{[D_l(l + 1) + (D_l + D_q)^2]^2} = 0$$

where $D_l \equiv D_l^{(1)}$ and $D_q \equiv D_l^{(3)}$, because at $c_\text{p} = 0$ the protein solution is a monodisperse suspension of monomers (i.e., $n = 1$).

## ASSOCIATED CONTENT

### Supporting Information

Discussion of the alternative picture of dynamic density fluctuations. Explanation of the subtraction of the scattering contribution from the sample holder using Paalman-Pings coefficients. Figure displaying the number density $\rho_n$ of $n$-clusters according to eq 3. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcllett.5b01073.

### AUTHOR INFORMATION

**Corresponding Author**

*E-mail: seydel@ill.eu.*
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The research at Oak Ridge National Laboratory’s Spallation Neutron Source was sponsored by the Scientific User Facilities Division, Office of Basic Energy Sciences, U.S. Department of Energy. The spectrometer BASIS is supported by the Jülich Center for Neutron Science (JCSNS), Germany, via the partner user programme and M.G. and F.Za. acknowledge travel funding by ILL to conduct the experiments as well as a student grant by the Institut Laue-Langevin (ILL). M.H., M.G. and T.S. have obtained travel funding by ILL to conduct the experiments. We acknowledge E. Mamontov, M. Oettel, R. Roth, and H. Schober for fruitful discussions and R. Moody and V. Glenisson for assistance. We further acknowledge financial support by the DFG.

REFERENCES

(3) Dhont, J. An Introduction to Dynamics of Colloids; Elsevier Science: New York, 1996.


