Supporting Information

Tracking Internal and Global Diffusive Dynamics During Protein Aggregation by High-Resolution Neutron Spectroscopy

Kevin Pounot[1,2], Hussein Chaaban[3], Vito Foderà[3]*, Giorgio Schirò[1], Martin Weik[1]*, Tilo Seydel[2]*

[1]Univ. Grenoble Alpes, CEA, CNRS, Institut de Biologie Structurale, F-38000 Grenoble, France
[2]Institut Max von Laue - Paul Langevin, 71 avenue des Martyrs, CS 20156, F-38042 Grenoble cedex 9, France
[3]Department of Pharmacy, Universitetsparken 2, University of Copenhagen, DK-2100 Copenhagen, Denmark
*Corresponding author Email addresses: vito.fodera@sund.ku.dk; weik@ibs.fr; seydel@ill.eu

1. Materials and methods

1.1. Lysozyme particulate formation followed by thioflavin T fluorescence kinetics

Lysozyme from chicken egg white (batch: 62970-5G-F, from by Sigma-Aldrich (Germany)) powder was kept at 4°C for storage. The lysozyme powder was dissolved in a D_2O (99.9%) solution at pD 10.5 and 0.1 M NaCl to a final concentration of 50 mg/mL. Subsequently, Thioflavine T (ThT) was added to a final concentration of 20 µM, the solution was sealed in a double-walled aluminum cylindrical sample container (outer diameter 22 mm, gap Δr = 0.15mm) with a 1 mm indium wire and incubated in a water bath at 90°C for either 0.5, 1, 1.5, 2, 3, 4 or 6 hours. After incubation, 300 µL of the solution was transferred in three wells (100 µL in each well) of a 96-well plate and the ThT fluorescence was measured on a BioTek Synergy H4 Hybrid plate-reader. ThT fluorescence was measured using 450 nm light for excitation and 490 nm filter for detection employing bottom optics and a manual gain of 65. The measurements in the three wells were averaged for each incubation time. The result is reported in figure 1b. In addition, the morphology of the particulates formed either in a plastic container or in the aluminum sample holder during the neutron scattering experiment (section 1.9) was observed to be similar as shown in figure S1.

1.2. Scanning Electron Microscopy

The morphology of the lysozyme particulates was studied by Scanning Electron Microscopy (SEM) using a Quanta™ 3D FEG scanning electron microscope (Thermo Fisher Scientific, Hillsboro, OR, US). The lysozyme powder was dissolved in a deuterated solution (D_2O, 99.9%) with 0.1 M NaCl at pD 10.5 to a final concentration of 50 mg/mL. The solution was incubated at 90°C for 90 minutes. Afterwards, the obtained lysozyme particulates were centrifuged at 10.000 rpm, 10 °C for 10 min. The samples were then washed in 1 ml deionized water (H_2O) and re-suspended. The centrifuge step was repeated. Subsequently, the samples were mounted on carbon tapes and sputter-coated with 0.2 nm gold using a Leica EM ACE200 (Leica Microsystems GmbH, Wetzlar, Germany) prior to imaging. The size distribution of the particulates was investigated by quantitative measurement of the size of 500 particulates from different representative SEM images (Main text Figure 1b).

1.3. Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy

Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy (ATR-FTIR) was used on native lysozyme and lysozyme particulates in transmission mode to obtain information on the secondary structure. The solid samples were run on a Bomem IR spectrometer (Bomem, Quebec, Canada). The spectra were collected during 256 scans at 2 cm^{-1} resolution in the 4000-600 cm^{-1} region. A double subtraction method was used tailed by the second derivative of 12 points, and a 2-3 point baseline correction in the amide I region. Finally the obtained spectra were area-normalized (Figure S2).

1.4. X-ray powder diffraction

The lysozyme powder was dissolved in deuterated solution (D_2O, 99.9%) and 0.1 M NaCl at pD 10.5 (100 µL) to a final concentration of 50 mg/mL in an Eppendorf tube. The solution was incubated at 90°C for 6 hours. Afterwards, the obtained lysozyme particulates were lyophilized by freezing them at -80°C and then using a Christ Martin lyophilizer operating at 0.01 mbar. A fraction of the powder (0.1 mg) was sealed in a Mylar capillary and mounted on the ID29 beamline of the ESRF synchrotron in Grenoble, France. Diffraction data were collected with a PILATUS 6M detector using the following scheme: 20 images with a total oscillation range of 10° (0.5° per image) and 1 s exposure time per image (100% transmission for 5.42×10^{12} ph/s at λ = 0.98 Å). The obtained diffraction images were averaged and processed using the pyFAI library[1] to perform the azimuthal integration (Figure S3).

1.5.Theoretical apparent center-of-mass diffusion coefficient of monomeric lysozyme

The theoretical diffusion coefficient $D_s^{(theo)}$ of lysozyme (50mg/mL) in D_2O was computed at 7°C and 90°C (Figure S5) for comparison with the experimental coefficient $D_s^{(exp)}$ (Figure S4) obtained from neutron scattering data.

This calculation involves (1) the theoretical diffusion coefficient in the limit of infinite dilution D_0, and (2) concepts from colloid physics to calculate the diffusion coefficient at the given finite protein concentration.

To obtain D_0, HYDROPRO[2] is employed, using the protein structure in the pdb format as input. The pdb structure for lysozyme is published only for ambient temperature (PDB entry code 3IJV used here). For the HYDROPRO calculation at 7°C*, i.e.* at the temperature at which full QENS spectra (Figure S4) were recorded prior to triggering aggregation, this native lysozyme structure (PDB entry code 3IJV) was used directly.

To compute the theoretical apparent center-of-mass diffusion coefficient for the unfolded[3] lysozyme monomer at 90°C, *i.e*. the temperature at which particulate-formation was followed in the neutron spectroscopy experiment, molecular dynamics simulations were employed to unfold the native structure. A starting model was constructed using the same published structure of lysozyme (PDB entry code 3IJV). This structure was solvated employing the VMD solvate plugin with a padding of 15 Å and charges were neutralized using the VMD autoionize plugin with Na and Cl ions[4]. The equilibration was performed using NAMD 2.13[5] with the TIP3 model for water[6] and the CHARMM36 force field[7]. The Nose-Hoover-Langevin piston algorithm maintained a constant pressure[8,9] and the stochastic velocity rescaling algorithm controlled the temperature[10]. Bonds with hydrogen atoms were constrained using the SHAKE algorithm[11], and the Verlet-I/r-RESPA multiple-time step scheme[12,13] was used to integrate the equations of motion, with time steps of 2 fs for long-range nonbonded forces, and 1 fs for short-range bonded forces. Electrostatic interactions were computed

using the smooth Particle Mesh-Ewald (PME) sum[14], with a cutoff set to 12 Å , a switching function starting at 10 Å, and a pair list distance of 16 Å was used.

Subsequent to the above steps, the native solvated lysozyme was unfolded using a 2 ns simulation at 327°C in the isothermal-isobaric (NPT) ensemble. Subsequently, a 100 ns equilibration at 90°C was performed, again in the NPT ensemble. After the simulation, the coordinates from six different frames along the simulation trajectory were used to produce pdb files. These files were used as input in the HYDROPRO software[2] to compute rotational and translational diffusion coefficients in the dilute limit.

In the procedure described above, the protonation state of the protein corresponds to a pH of 7 as obtained from CHARMM topology files. A calculation with PROPKA[15,16] gives mean pKa values of 10.1 for tyrosines and 11 for lysines for the unfolded lysozyme, thereby suggesting that less than half of tyrosines and a tenth of lysines are oxidized at pH 10, and that the structure at pH 7 is a reasonable representation of the monomeric lysozyme. The conformational ensemble probed during the 100 ns equilibration is sketched in Figure S5a where 25 structures obtained along the trajectory of 2500 frames are superposed. The radius of gyration of the unfolded lysozyme oscillates between 15 and 20 Å but remains larger than the monomer (Figure S5b). The radius of gyration is in agreement with a previous study on thermal unfolding of lysozyme[17], even though the authors used a pH lower than 7. Six conformers along the trajectory are used to compute the theoretical apparent diffusion coefficient, which yields 30 ± 1 Å²/ns.

1.6. Calculation of the diffusion coefficient at finite volume fraction ($\Phi > 0$)

To obtain the apparent diffusion coefficient observable in QENS (cf. section 1.10), we used an effective hydrodynamic protein volume fraction computed using:

$$\Phi_{\text{eff}} = \Phi \left(\frac{R_h}{R} \right)^3 \tag{1}$$

where $\Phi$is the dry volume fraction computed from the lysozyme specific volume ν = 0.726 mL/g[18] R_h is the hydrodynamic radius obtained from HYDROPRO by inserting the translational diffusion coefficient into the Stokes-Einstein relation, and R the dry protein radius obtained with

$$R = \sqrt[3]{\frac{3}{4\pi} \frac{\nu M_w}{\mathcal{N}_A}} \tag{2}$$

with M_w being the molar weight (14.3kDa) and $\mathcal{N}_A$ the Avogadro number.

HYDROPRO returns the translational and rotational diffusion coefficients in the dilute limit. Subsequently, molecular crowding can be considered using established analytical expressions of scaling factors derived from colloid models for translational motions[19] (eq. 11,12) and rotational motions[20] (eq. 21). Finally, the apparent diffusion coefficient D=D(D_r, D_t) is computed from the thus obtained rotational D_r and translational D_t components using the root finding algorithm scipy.optimize.root to solve the implicit relation[21,22]:

$$\sum_{l=0}^{n} B_l(q) \frac{D_r l(l+1) + (D_t - D)q^2}{[D_r l(l+1) + (D_t + D)q^2]^2} = 0 \tag{3}$$

with n being set to 100 to obtain good convergence and $B_l(q)$ being obtained using the radial hydrogen number density function (relative to the protein center of mass) $\rho(r,n)$:

$$B_l(q) = (2l+1)\int_0^\infty dr\rho(r,n)j_l^2(qr) \quad (4)$$

The Python/C scripts that were used to perform the analysis are available on github (github.com/kpounot/NAMDAnalyzer – ScatDiffusion module).

The resulting theoretical apparent diffusion coefficients as a function of volume fraction for native and unfolded lysozyme, respectively, are reported in figure S5c.

1.7. Dynamic Light Scattering

Lysozyme was prepared at final concentration of 20 mg/mL in 100 µL of a deuterated solution (D_2O, 99.9%) and 0.1 M NaCl, pD 10.5. The solution was filtered using a 20 µm filter and centrifuged at 17000 rpm for 15 minutes prior to measurements. All measurements were carried out on a Wyatt DynaPro NanoStar instrument, using standard cuvettes containing 10 µL of sample. The data were acquired and analyzed using the Dynamics software from Wyatt technology. Each sample showed an integrated intensity of about 700 000 counts and a sum of squared errors lower than 400 for cumulant fit. Instrument limitations, due to the lack of a Peltier element, restricted the accessible temperature range from 25 to 60°C. Dynamic Light Scattering (DLS) experiments were carried out at these two boundary temperature values (Figure S6).

1.8. Comparison of the long-time translational diffusion coefficient D_t obtained from DLS with the apparent short-time diffusion coefficient $D_s^{(exp)}$ obtained from neutron scattering, and the calculated diffusion coefficient $D_s^{(theo)}$

The diffusion coefficient D_t that is obtained from DLS (Figure S6a, by fitting the signal corresponding to the first peak in Figure S6b) reflects the long-time collective translational diffusion, where direct particle-particle interactions (collisions) influence the diffusion. Conversely, the coefficient $D_s^{(exp)}$ extracted from the analysis of the neutron scattering data reflects the short-time self-diffusion, where the time and length scales considered are too small for direct particle-particle interactions to affect the result. In the dilute limit, *i.e*. for a single protein in solution, long-time collective and short-time self diffusion coefficients are equal. The experimental values we obtained for both coefficients at finite volume fractions can be extrapolated to the dilute limit using eq. 20 and eq. 24 in the work by Banchio and Nägele[20] for $D_s^{(exp)}$ and D_t, respectively. In addition, assuming that the lysozyme - monomer or oligomer - is spherical, the result from the DLS measurement at 25°C can be extrapolated to 7°C (temperature of neutron scattering experiment prior to temperature raise) using the Stokes-Einstein relation. We thus extrapolate the DLS result, D_t at 7°C to the dilute limit to obtain $D_{0,t}$ = 3.52 Å²/ns, and we extrapolate the neutron scattering result $D_s^{(exp)}$ to the dilute limit at 7°C to obtain $D_{0,s}^{(exp)}$ = 3.2 Å²/ns. The theoretical diffusion coefficient in the dilute limit can be obtained by taking the value where the blue curve intersects the ordinate in Figure S5a (cf. section 1.6), which corresponds to a zero protein volume fraction. We obtain $D_{0,s}^{(theo)}$ = 6.3 Å²/ns, which indicates that both $D_{0,s}^{(exp)}$ and $D_{0,t}$ are approximately by a factor of two lower than the theoretical expectation for monomers. From this observation, we conclude that our experimental QENS and DLS results indicate the presence of oligomers prior to triggering the aggregation. We note again that the extrapolated experimental coefficient $D_{0,s}^{(exp)}$ from QENS contains contributions from both translational and

rotational motions[19] (cf. section 1.6). We also note that, since the aggregates are presumably transient, the different observation time scales of DLS (µs) and QENS (ns), respectively, do not necessarily give fully consistent results on the cluster size (cf. discussion and references in the main part).

An approximation of the diffusion coefficient for the lysozyme solution at 90°C can be obtained also from the experimental value at 7°C. Using the Stokes-Einstein relation with tabled values for the viscosity of D_2O[23], we obtain 363 η_{280K}/280 η_{363K}=6.6 where η_{280K} is the viscosity of D_2O at 280K, which gives $D_s^{(\text{exp approx})}$ = 19.8 Å²/ns for the folded lysozyme at 90°C. This value is slightly higher than the one obtained experimentally, which can be explained by the fact that lysozyme is unfolded at 90°C, while this value is approximated from a folded structure. Hence, we can assume a similar distribution of oligomers at 7°C and 90°C.

1.9. Neutron scattering experiments

We have used the cold neutron backscattering spectrometer IN16B[24,25] with Si(111) analyzer crystals, corresponding to an elastic wavelength of 6.27 Å. A mechanical Doppler drive was used to set the incident energy, and a phase space transformation (PST) chopper was employed to optimize the neutron flux at the sample position[26]. The lysozyme was dissolved at 50 mg/mL in 1 mL of D_2O (99.9%) and 0.1 M NaCl at pD 10.5 and was sealed in a double-walled cylindrical aluminum sample container (outer diameter 22 mm, gap Δr = 0.15 mm) and introduced at 7°C into a standard Orange cryofurnace that allowed changing the sample temperature during the experiment. The choice of the instrument was motivated by the high flux available and its unique capability to perform elastic and inelastic fixed-window scans (E/IFWS) that allowed us to acquire data with a reasonable time resolution - around 20 minutes for a full cycle of elastic and three inelastic data points (Figure 2)[26]. A schematic illustration of the concept of E/IFWS is depicted in figure S11.

Quasi-elastic neutron scattering (QENS) spectra were first recorded at 7°C (Figure S4), with the momentum transfer q range $0.19 < q < 1.90$ Å^{-1} explored. Subsequently, the temperature was increased to 90°C (heating time approximately 30 min; temperature vs. time displayed in Figure S8) to trigger the aggregation process, and E/IFWS were acquired at 90°C using the following scheme: 1 min counting time at 0 µeV, 2 min at 0.6 µeV, 6 min at 1.5 µeV, and 6 min at 3 µeV. This sequence was repeated continuously for 6.5 hours. Finally, QENS spectra were recorded at 90°C (Figure S7) to obtain a more accurate measurement of the final state as a control.
No visible corrosion of the employed Al sample container for the lysozyme solution was detected at the end of the neutron experiment. This absence of corrosion is in agreement with published work on the behavior of Al surfaces[27].
The neutron data are permanently curated by the ILL and accessible at http://dx.doi.org/10.5291/ILL-DATA.8-04-811

1.10. Neutron scattering: Basic concepts and details of modeling and data analysis

Basic concepts: Crucially, the analysis of the relatively scarce fixed-window neutron scattering data (E/IFWS. section 1.9 and schematic figure S11) is achieved by a so-called global fit for the data at all momentum transfer values q and energy transfer values ΔE simultaneously. The model contains two Lorentzian contributions (one accounting for the center-of-mass diffusion and one for the

protein-internal dynamics), the *q*-dependence of which is fixed in the model. Moreover, the aqueous solvent contribution in the fits is fixed by results from separate fits and tabled values of corresponding measurements on pure water (D_2O). In this way, the number of free parameters is drastically reduced at the expense of imposing the *q*-dependence, which involves assumptions on the type of diffusion. However, the validity of these assumptions was shown previously in stable-state QENS where full spectra could be recorded[28,29].

It is important to note that the neutron backscattering experiment with its energy resolution of approximately 0.8 μeV FWHM corresponds to an observation time scale of the diffusion in the sample of a few nanoseconds (cf. sections 1.6 and 1.8). On this short observation time scale, the proteins diffuse by less than their radius, and protein-protein collisions are negligible[30]. Therefore, the observed center-of-mass diffusion coefficient of the proteins $D_s^{(exp)}$ can be interpreted in terms of a colloidal short-time coefficient. Since the scattering from the proteins with its prevailing hydrogen atoms is mostly incoherent, this coefficient $D_s^{(exp)}$ can be identified with the self-diffusion or, synonymously, tracer diffusion coefficient. Due to the large momentum transfers involved in the backscattering experiment, $D_s^{(exp)} = D_s^{(exp)}(D_r, D_t)$ is an implicit function that contains contributions from both rotational ("tumbling") D_r and translational D_t diffusion[30]. Therefore, $D_s^{(exp)}$ is also denoted *apparent* or *observable* diffusion coefficient (cf. sections 1.5 and 1.6). The internal diffusive dynamics arising mainly from protein backbone is convoluted with the center-of-mass diffusion in the measured signal. On the observation scale of our experiment, this internal diffusion can be well modeled by the inclusion of an additional Lorentzian function in the model (equation 5 below), as shown in previous work[30]. This additional Lorentzian accounts mainly for the protein backbone fluctuations, while the protein side chains diffuse too fast to be detected within the range of our E/IFWS experiment[28]. The resulting combined model function for our neutron spectra will be explained in the following.

Details of modeling and data analysis: To analyze the neutron scattering data, we assumed that the center-of-mass diffusion follows Fick's law and is, thus, reflected in the incoherent QENS signal by a Lorentzian contribution with the width[21,30] $\Gamma(q) = D_s^{(exp)} q^2$. The validity of this relation was shown in previous work for various protein solutions both at physiological and elevated temperatures[28,29]. The superimposed contribution to the scattering function from the backbone and side-chain motions - that is, internal dynamics - within the protein was assumed to follow a more complicated functional dependence, the so-called jump diffusion model[30]. In brief, it is approximated by an additional Lorentzian contribution, the *q*-dependence of which reads $\gamma(q) = \frac{D_i q^2}{1 + D_i q^2 \tau_0}$ where *q* is the momentum transfer,

D_i is the diffusion coefficient associated with internal dynamics and τ_0 is the residence time for atoms between diffusive jumps. The jump-diffusion model has been shown to capture the backbone contribution of the protein internal dynamics reasonably well[30]. In contrast, it does not describe the fast side-chain motions that are beyond the energy range probed by the present experiment.

The additional aqueous solvent contribution to the signal was analyzed by measuring the pure solvent separately. The solvent was then fitted separately using a single Lorentzian lineshape and the result was added to the model used for the protein solution[22]. This approach of fixing the solvent contribution further reduced the number of free parameters in the fit of the protein solution data.

Data reduction, *i.e*. normalization to the incident beam intensity recorded by a so-called monitor device, integration over the regions of interest of the vertically position-sensitive detector tubes,

calculation of the energy axis, and centering of the elastic line positions – using a separate vanadium measurement - was performed using Mantid's data reduction module for IN16B[31].

Empty cell subtraction, as well as the fitting of the resolution function, D_2O signal, normalization of the detector efficiency employing the Vanadium signal, and fitting of the fully reduced data were performed using an custom-made python API, with the source code available on github (github.com/kpounot/nPDyn). The resolution function was fitted using both a pseudo-Voigt profile and a double Gaussian model – presented results used pseudo-Voigt profile. The D_2O spectrum was fitted using published linewidths already available from measurements on the time-of-flight spectrometer IN6, following the procedure described previously[21,30].

Hence, the final model reads:

$$S(q,\Delta E) = R(q,\Delta E) \otimes \{\beta[a_0\mathcal{L}_\Gamma(q,\Delta E) + (1-a_0)\mathcal{L}_{\Gamma+\gamma}(q,\Delta E)] + \beta_{\mathrm{D_2O}}\mathcal{L}_{\mathrm{D_2O}}\} \quad (5)$$

where $R(q,\Delta E)$ is the resolution function, β accounts for the overall amplitude, a_0 can be identified with the Elastic Incoherent Structure Factor EISF[21,30] and represents the relative contribution of the center-of-mass diffusion Lorentzian, $\mathcal{L}_\Gamma$ is the center-of-mass diffusion Lorentzian, the Lorentzian with the width $\mathcal{L}_{\Gamma+\gamma}$ is the convolution between the Lorentzians for center-of-mass diffusion and internal dynamics, and $\beta_{\mathrm{D_2O}}\mathcal{L}_{\mathrm{D_2O}}$ is the D_2O signal, which is fixed in the fits as explained above. The convolution with the resolution function is performed analytically. The convolution of the Lorentzians representing the resolution and the model is straightforward and simply gives a Lorentzian of linewidth being the sum of the two convoluted Lorentzians. The convolution of the Gaussian representing the resolution function and the Lorentzians in the model is analytically given by the Voigt function[32].

To find the global minimum of the fit, *i.e.* to remove the dependence on the starting parameters ("initial guess") during the fitting procedure, we used the *scipy basinhopping* algorithm, which performs multiple iterations of the following scheme: (1) random shift of the starting parameter values, (2) use of a minimization algorithm to find the local minimum, (3) compare the local minimum with the previously found minimum by the Metropolis criterion to accept or reject the parameters fitted in the iteration[33]. The basinhopping procedure thus allows to better explore the solution landscape and maximizes the chance to find the true global minimum. In the case of the global fit, *i.e.* fitting all momentum transfer values at once, the aforementioned formula for the Lorentzian widths Γ and γ were used, with an explicit dependence on momentum transfer *q*. For momentum transfer-wise fits, where each momentum transfer value is fitted individually, the two Lorentzian widths become scalar parameters.

In addition, the robustness of the fitting procedure for E/IFWS was assessed by performing momentum transfer-wise fitting of data (Figure S9), which shows that the information about center-of-mass diffusion and protein-internal dynamics is already present at each momentum-transfer. Moreover, we observed that the subtraction of the empty cell signal suppresses most of the elastic peak in the data, and might lead to a bias of the result when the elastic point is lower than the others. We thus performed a fit of a subset of the data which does not contain the EFWS (Figure S10), yielding again the same result. The errors at the energy transfer ΔE=0.6 μeV are larger than at higher energy transfers, such that the fitting procedure favors the latter due to the error weighting.

2. Supplementary figures

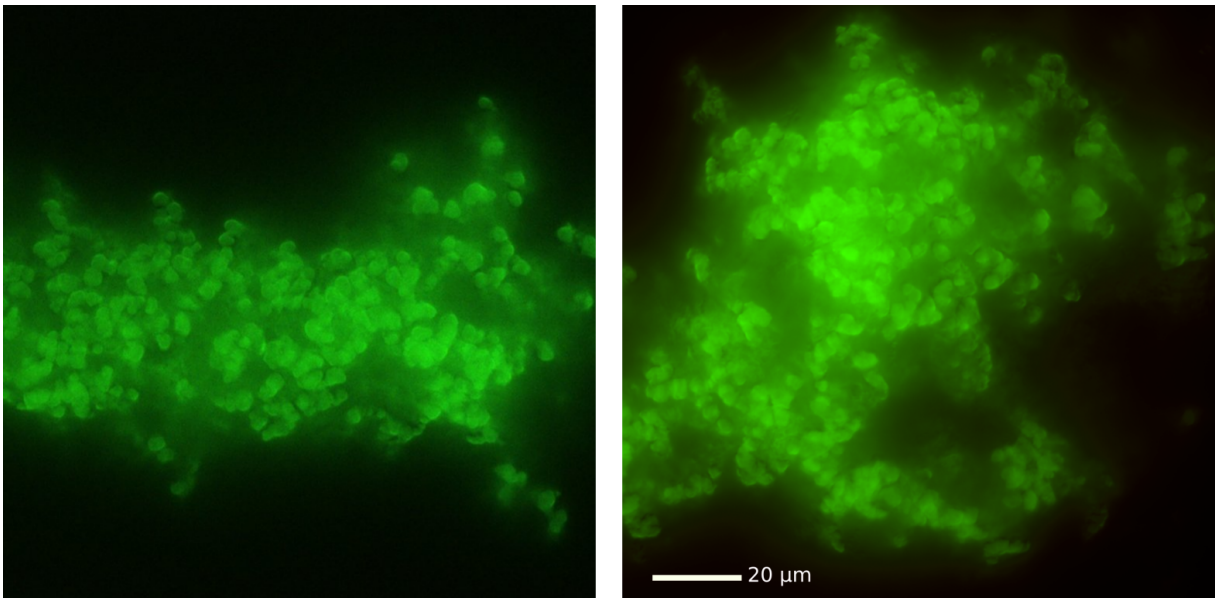

Figure S1. Lysozyme particulates formed in plastic and aluminum containers are similar. Observation of particulates by ThT fluorescence microscopy. Particulates were formed either in the aluminum sample holder during the neutron scattering experiment **(right),** or, as a control, in a plastic container **(left),** showing the same morphology. The scale bar holds for both images. We note that the samples for the experiments in the figures S1, S2 and S3 were not identical, but were prepared independently according to the identical protocol.

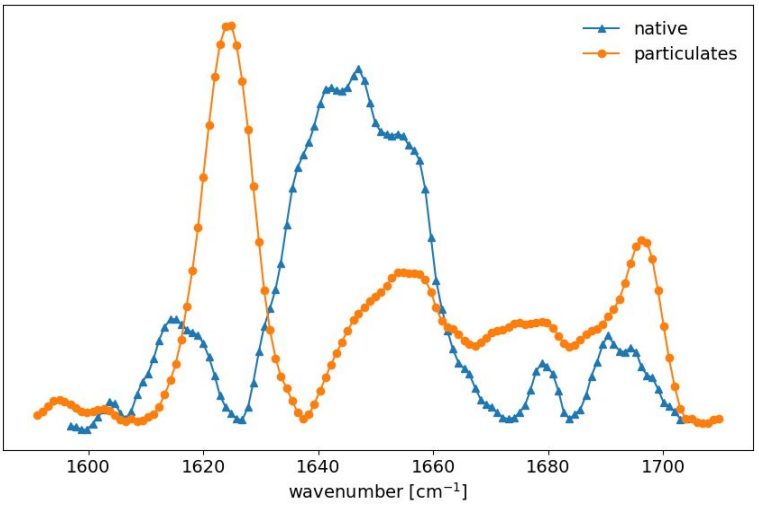

Figure S2. Lysozyme undergoes substantial secondary structure change upon particulates formation. ATR-FTIR spectra (section 1.3) for native lysozyme (blue) and lysozyme particulates (orange), respectively. Native lysozyme is mainly alpha-helical in accordance with other studies[34]. However a transition to β -sheet is observed in the particulate samples, which was also shown by Krebs et al.[35]. We note that the samples for the experiments in the figures S1, S2 and S3 were not identical, but were prepared independently according to the identical protocol.

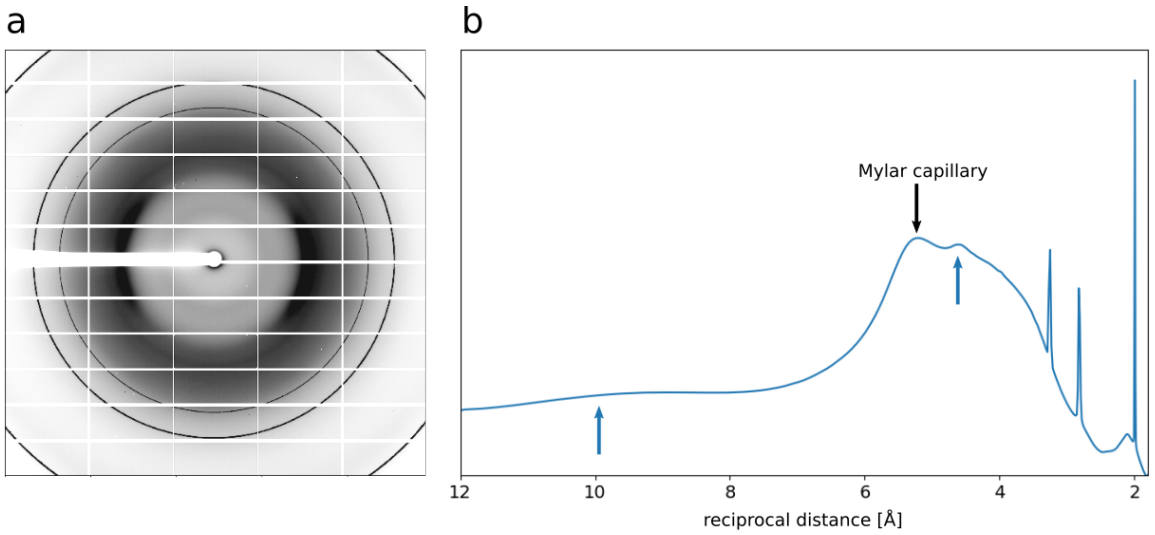

Figure S3. Lysozyme particulates exhibit an amyloid-like signal in X-ray powder diffraction patterns. X-ray powder diffraction pattern for lysozyme particulates (section 1.4); **a** Average over the 20 measurements of the diffraction patterns of the lysozyme particulates powder. **b** The azimuthal integration yields a curve that presents three sharp peaks at small reciprocal distance, corresponding to salt crystals. The blue arrows indicate amyloid-related rings[36] and the black arrow the position of the Mylar capillary peak. We note that the samples for the experiments in figures S1, S2 and S3 were not identical, but were prepared independently according to the identical protocol.

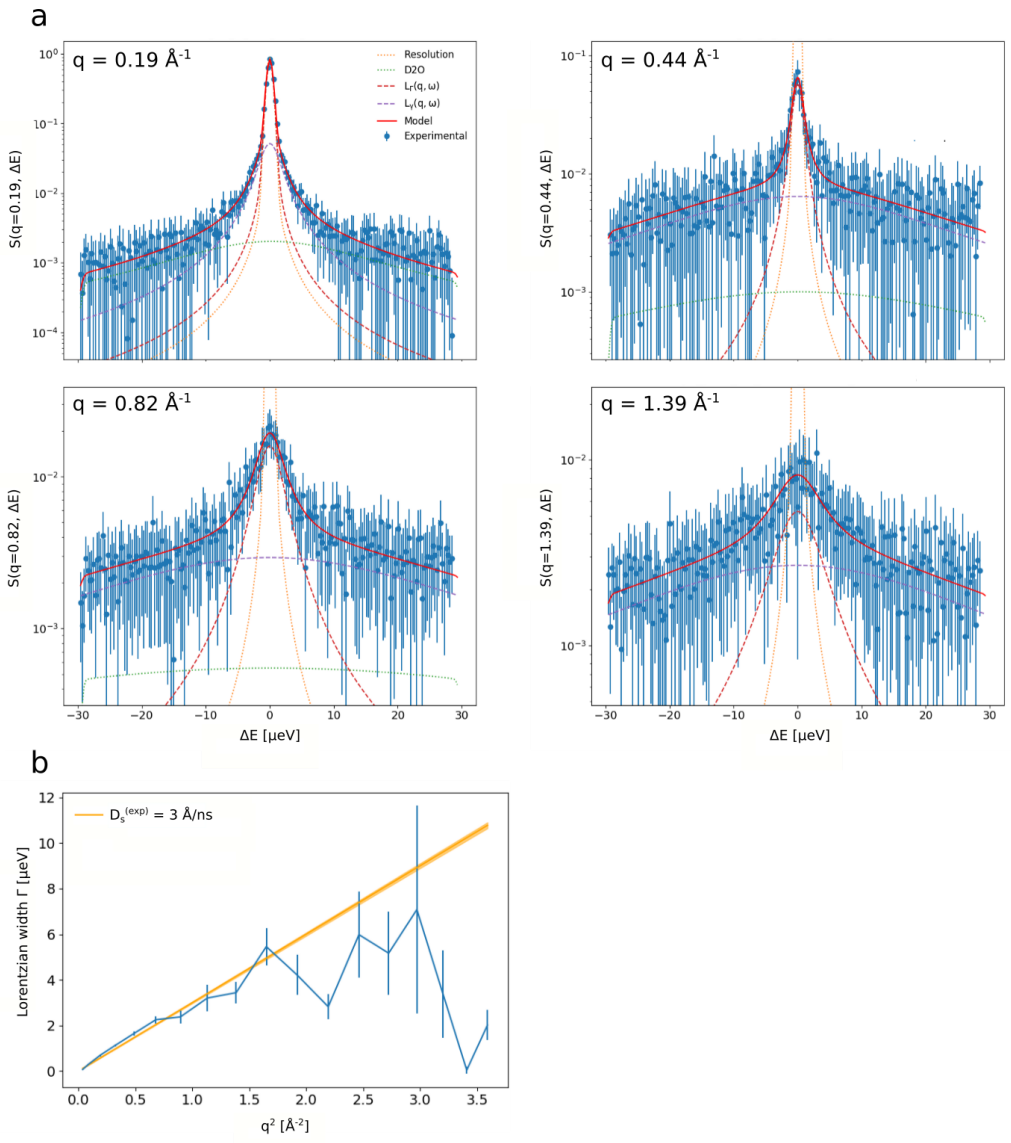

Figure S4. Determination of the center-of-mass diffusion coefficient $D_s^{(exp)}$ of lysozyme with QENS prior to aggregation. **a** QENS spectra at four different momentum transfer (*q*) values measured at 7°C prior to aggregation (blue circles, errors represented by blue vertical bars) and fit results using the momentum transfer-wise procedure (sections 1.9 and 1.10). The dotted orange line represents the resolution function, the green dotted line the D_2O signal, the dashed red line the Lorentzian describing center-of-mass diffusion, the dashed purple line the Lorentzian describing the protein-internal dynamics, and the red solid line the resulting model. **b** The width of the Lorentzian describing center-of-mass diffusion (blue line with error bars) is plotted against q² and fitted using the relation $\Gamma(q) = D_s^{(exp)} q^2$ (orange line with shaded area to represent the standard deviation) to obtain the indicated apparent diffusion coefficient $D_s^{(exp)}$.

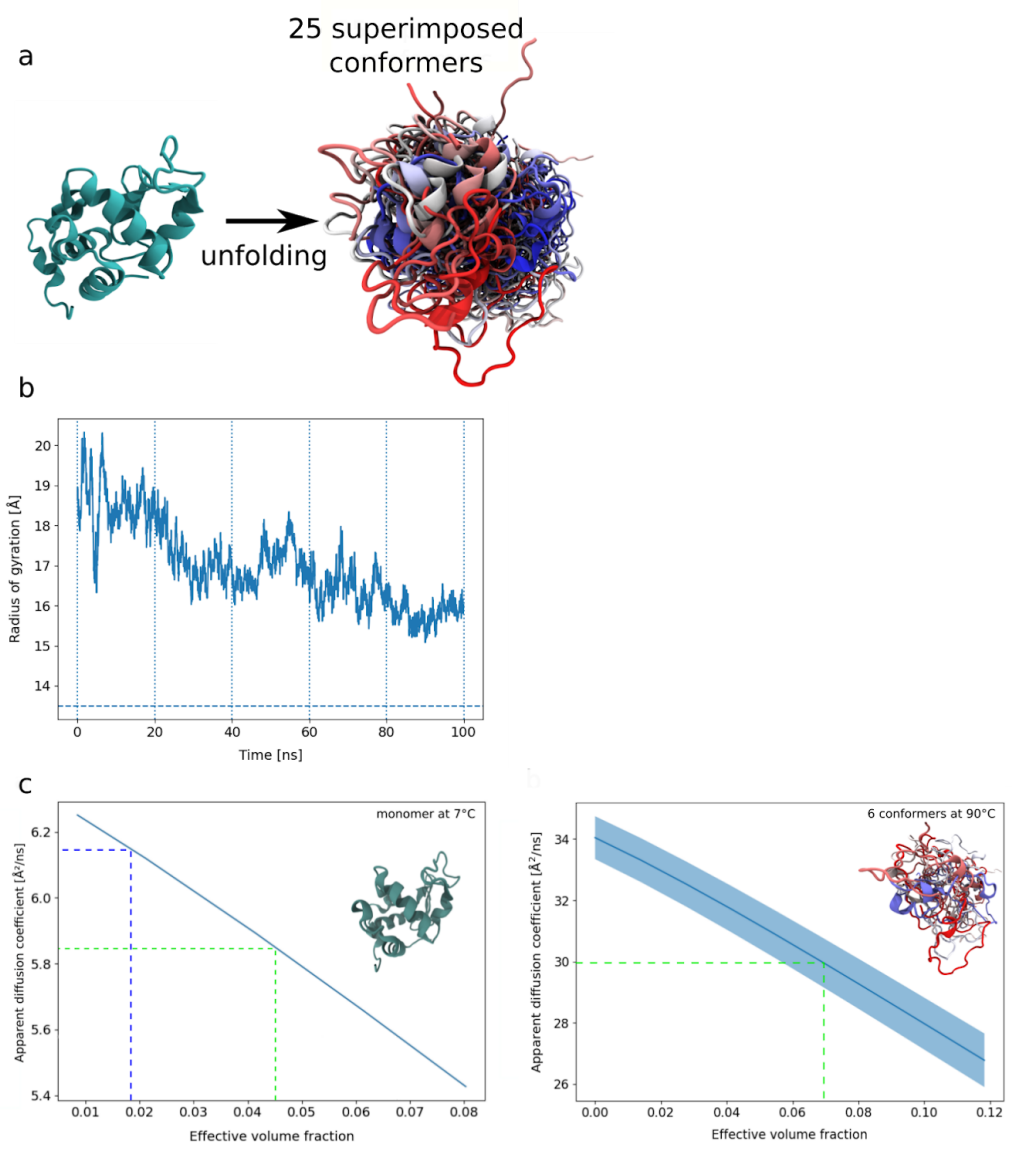

Figure S5. Determination of theoretical apparent center-of-mass diffusion coefficients for monomeric lysozyme at 7 a 90°C. a The native lysozyme structure (PDB 3IJV) was used to compute the apparent center-of-mass diffusion coefficient (as described in Methods, sections 1.5 and 1.6) for the monomer as a function of the protein effective volume fraction at 7° and was subsequently unfolded as explained in section 1.5 during MD simulation. Twenty-five superimposed conformers, extracted at a definite interval of 100 frames along the 2500 frames simulation trajectory, are shown.. **b** Radius of gyration for each conformer during the 100 ns equilibration. The horizontal blue line gives the radius of gyration for the native lysozyme. The vertical dotted lines represent the frames that were used to compute the theoretical diffusion coefficient (as described in Methods, sections 1.5 and 1.6). **c - left** Theoretical center-of-mass diffusion coefficient for the native monomeric lysozyme as a function of the volume fraction. The green dashed line indicates the diffusion coefficient corresponding to the volume fraction that was used in the neutron scattering experiment and the blue dashed line indicates the diffusion coefficient we would obtain for the lysozyme monomer at the volume fraction used in the DLS experiment. **c - right** Theoretical apparent diffusion coefficient as a function of the effective volume fraction for the unfolded lysozyme monomer at 90°C averaged over six conformers (blue line). The shaded blue area represents the standard deviation from the average value. The green dashed line indicates the diffusion coefficient corresponding to the volume fraction that was used in the neutron experiment. Due to the unfolding, the volume fraction occupied by the protein is increased relative to the native protein.

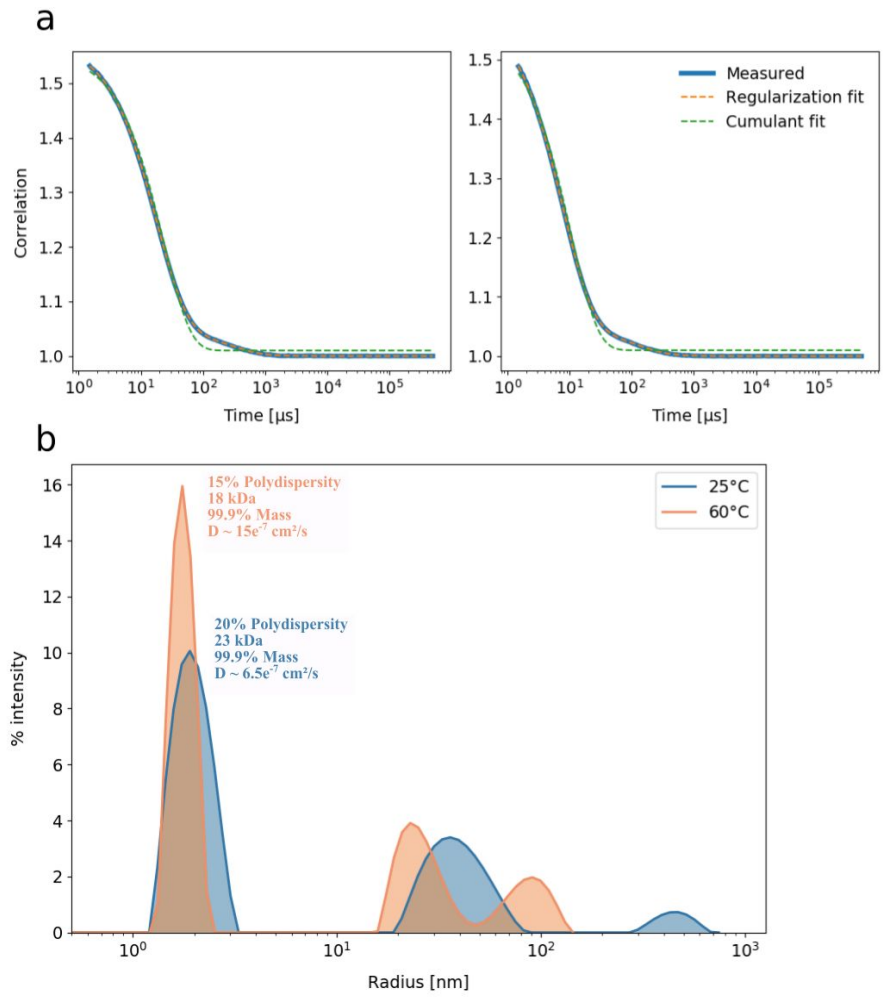

Figure S6. Dynamic light scattering measurement of lysozyme solution suggests the presence of oligomers. a Measured DLS signal (blue line) and fit result - using cumulant fit in green dashed line or regularization fit in orange dashed line as provided in the software (see section 1.7) - for a lysozyme solution at 25 °C (left panel) and 60 °C (right panel). **b** Radius distribution for both temperatures - 25°C in blue and 60°C in orange - represented using the percentage of intensity. Parameters obtained from the software analysis are given for the first peak at each temperature.

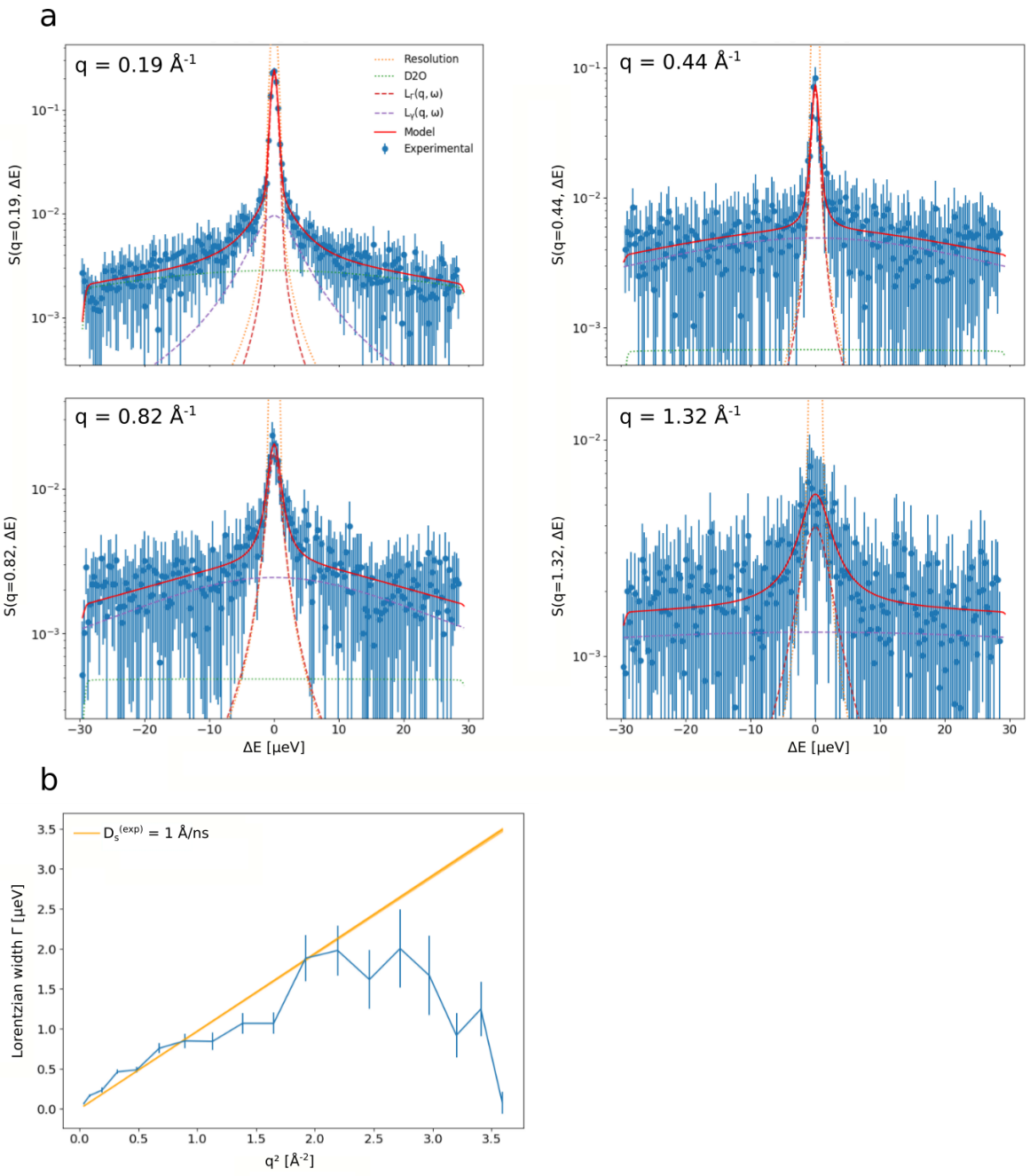

Figure S7. Determination of the center-of-mass diffusion coefficient after aggregation with full QENS. a QENS spectra at four different momentum transfer q values measured at 90°C 6.5 hours after aggregation was triggered (blue circles, errors represented with blue vertical bars) and fit results using the momentum transfer-wise procedure (sections 1.9 and 1.10). The dotted orange line represents the resolution function the green dotted line the D_2O signal, the dashed red line the Lorentzian describing center-of-mass diffusion, the dashed purple line the Lorentzian describing the protein-internal dynamics, and the red solid line the resulting model. **b** The width of the Lorentzian describing center-of-mass diffusion (blue line with error bars) is plotted against q^2 and fitted using the relation $\Gamma(q) = D_s^{(exp)} q^2$ (orange line with shaded area to represent the standard deviation) to obtain the indicated apparent diffusion coefficient $D_s^{(exp)}$.

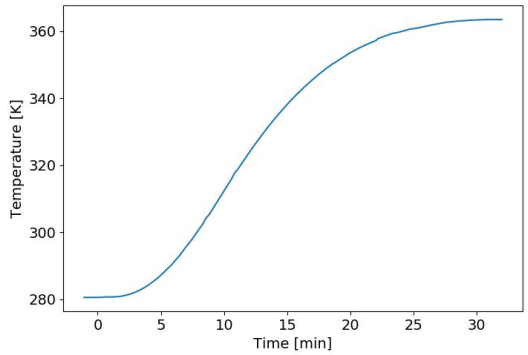

Figure S8. Temperature increase - measured on the sample temperature sensor of the instrument - triggering the aggregation process used during the neutron scattering experiment on IN16B at the ILL (section 1.9).

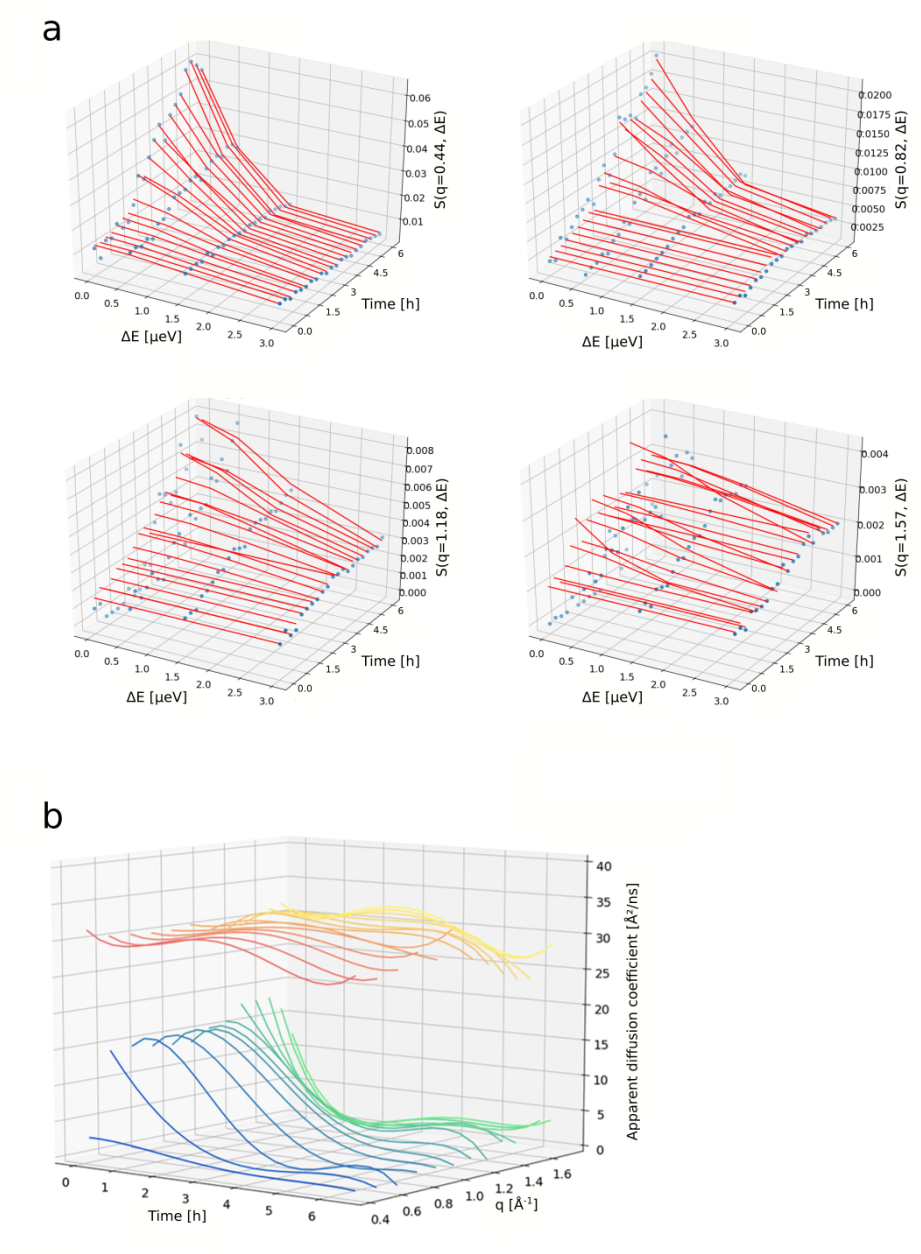

Figure S9. E/IFWS data fitted using a momentum transfer-wise procedure. a Normalized experimental E/IFWS spectra for each scan recorded during aggregation (blues circles) and fitted model (eq. 5 in section 1.10; red lines). **b** Apparent diffusion coefficient from E/IFWS data fitting for center-of-mass (cold colors) and protein-internal dynamics (warm colors) as a function of time and momentum transfer *q* values, as obtained from individual fits of eq. 5 in section 1.10. A Savitzki-Golay filter was used in the time dimension for better visualization (window length of 1 hour and polynomial order of 3).

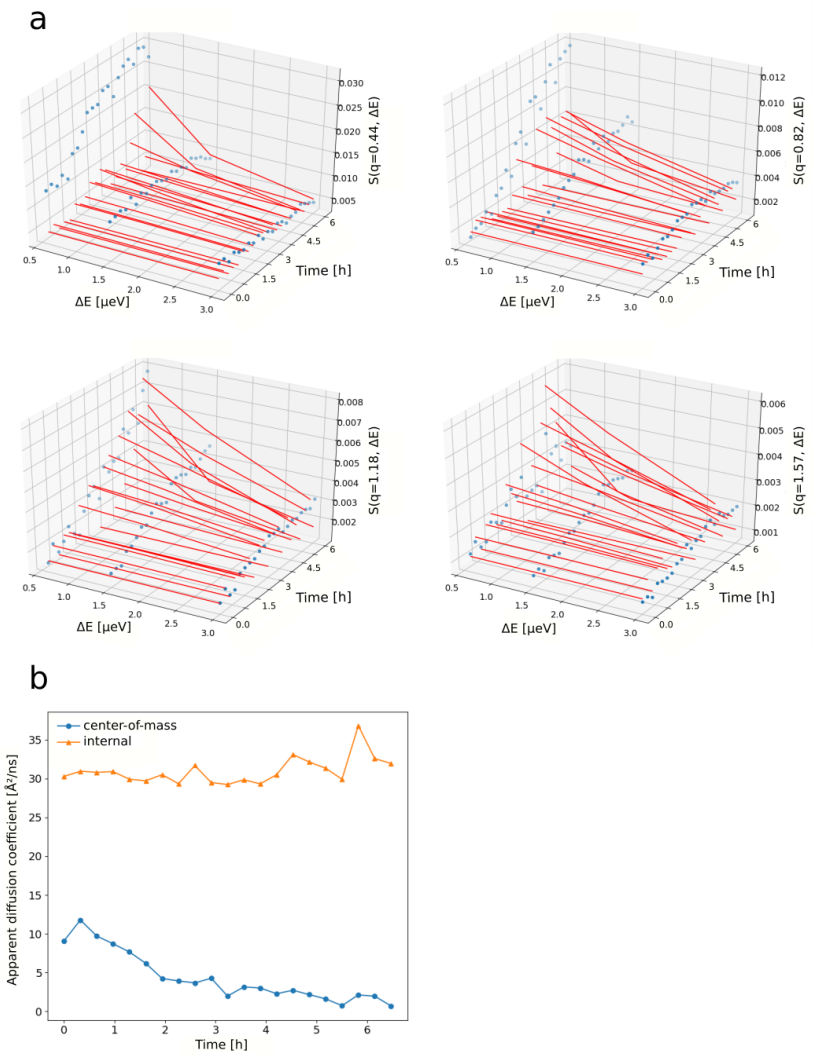

Figure S10. IFWS data fitted using a global fitting procedure. a Normalized experimental IFWS (elastic data excluded) spectra for each scan recorded during aggregation (blues circles) and fitted model (eq. 5 in section 1.10; red lines). **b** Apparent diffusion coefficients from IFWS data (i.e. excluding the EFWS data) fitting for center-of-mass (blue circles) and protein-internal (orange triangles) dynamics as a function of time and momentum transfer q values, as obtained from eq. 5 in section 1.10. The apparent diffusion coefficients obtained using E/IFWS are in reasonable agreement (Figure 3).

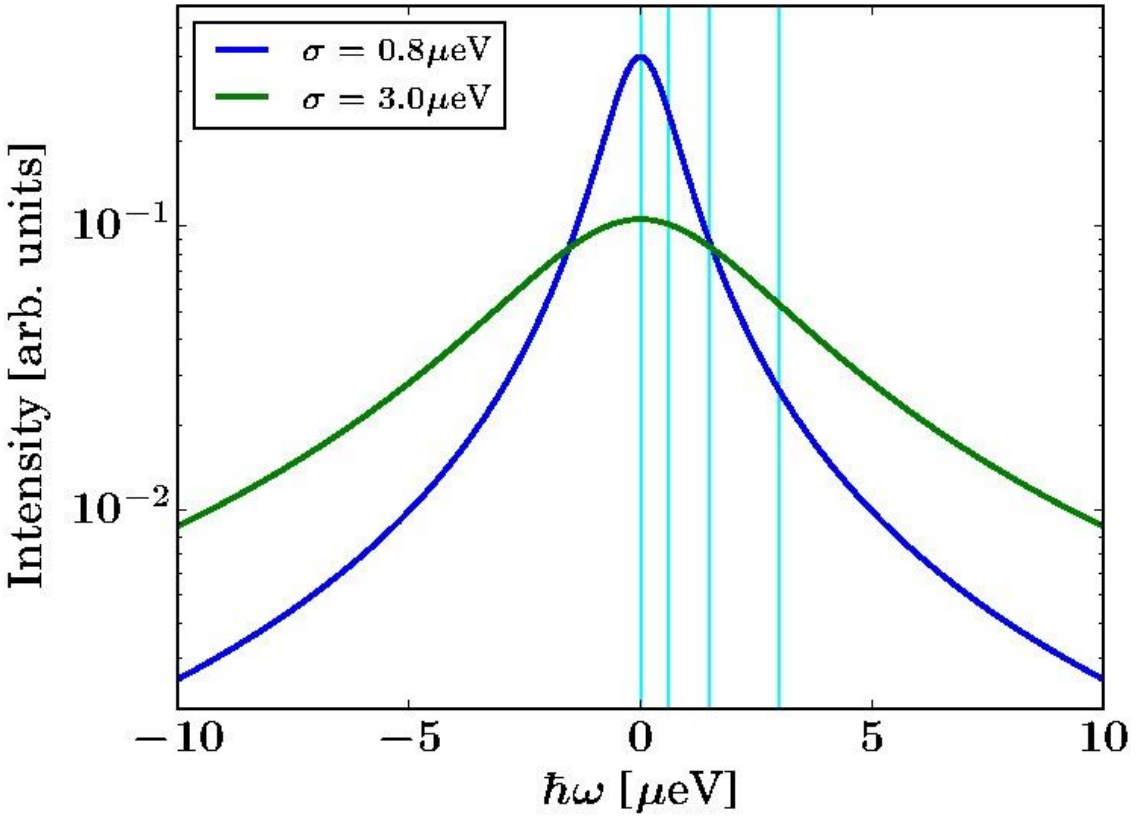

Figure S11: Schematic illustration of the concept of elastic and inelastic fixed window scans (E/IFWS): Two Lorentzian functions with linewidths (HWHW) of σ=0.8 and 3.0 μeV, respectively, are displayed as intensity versus energy transfer, corresponding to simple full QENS spectra (without convolution with the energy resolution function). The vertical lines mark the elastic (EFWS) and three inelastic positions (IFWS) where the intensities were recorded during the neutron backscattering experiment. The intercepts of the vertical lines with the two Lorentzians help to illustrate how the measured intensities evolve at different energy offsets as the sample evolves dynamically: For instance, the formation of protein aggregates corresponds to a narrowing of the Lorentzians. (Cf. section 1.9)

REFERENCES

(1) Kieffer, J.; Wright, J. P. PyFAI: A Python Library for High Performance Azimuthal Integration on GPU. *Powder Diffr.* **2013**, *28* (S2), S339–S350. https://doi.org/10.1017/S0885715613000924.

(2) Ortega, A.; Amorós, D.; García de la Torre, J. Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models. *Biophys. J.* **2011**, *101* (4), 892–8. https://doi.org/10.1016/j.bpj.2011.06.046.

(3) Booth, D. R.; Sunde, M.; Bellotti, V.; Robinson, C. V.; Hutchinson, W. L.; Fraser, P. E.; Hawkins, P. N.; Dobson, C. M.; Radford, S. E.; Blake, C. C. F.; Pepys, M. B. Instability, Unfolding and Aggregation of Human Lysozyme Variants Underlying Amyloid Fibrillogenesis. *Nature* **1997**, *385* (6619), 787–793. https://doi.org/10.1038/385787a0.

(4) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

(5) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802. https://doi.org/10.1002/jcc.20289.

(6) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **2001**, *105* (43), 9954–9960. https://doi.org/10.1021/jp003020w.

(7) Huang, J.; MacKerell, A. D.; Jr. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34* (25), 2135–45. https://doi.org/10.1002/jcc.23354.

(8) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant Pressure Molecular Dynamics Algorithms. *J. Chem. Phys.* **1994**, *101* (5), 4177–4189. https://doi.org/10.1063/1.467468.

(9) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys.* **1995**, *103* (11), 4613–4621. https://doi.org/10.1063/1.470648.

(10) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126* (1), 014101. https://doi.org/10.1063/1.2408420.

(11) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327–341. https://doi.org/10.1016/0021-9991(77)90098-5.

(12) Grubmüller, H.; Heller, H.; Windemuth, A.; Schulten, K. Generalized Verlet Algorithm for Efficient Molecular Dynamics Simulations with Long-Range Interactions. *Mol. Simul.* **1991**, *6* (1–3), 121–142. https://doi.org/10.1080/08927029108022142.

(13) Tuckerman, M.; Berne, B. J.; Martyna, G. J. Reversible Multiple Time Scale Molecular Dynamics. *J. Chem. Phys.* **1992**, *97* (3), 1990–2001. https://doi.org/10.1063/1.463137.

(14) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577–8593. https://doi.org/10.1063/1.470117.

(15) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of PKa Values. *J. Chem. Theory Comput.* **2011**, *7* (7), 2284–2295. https://doi.org/10.1021/ct200133y.

(16) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical PKa Predictions. *J. Chem. Theory Comput.* **2011**, *7* (2), 525–537. https://doi.org/10.1021/ct100578z.

(17) Meersman, F.; Atilgan, C.; Miles, A. J.; Bader, R.; Shang, W.; Matagne, A.; Wallace, B. A.; Koch, M. H. J. Consistent Picture of the Reversible Thermal Unfolding of Hen Egg-White Lysozyme from Experiment and Molecular Dynamics. *Biophys. J.* **2010**, *99* (7), 2255–2263. https://doi.org/10.1016/j.bpj.2010.07.060.

(18) Millero, F. J.; Ward, G. K.; Chetirkin, P. Partial Specific Volume, Expansibility, Compressibility, and Heat Capacity of Aqueous Lysozyme Solutions. *J. Biol. Chem.* **1976**, *251* (13), 4001–4004.
(19) Tokuyama, M.; Oppenheim, I. Dynamics of Hard Sphere Suspensions. *Phys. Rev. E - Stat. Nonlinear Soft Matter Phys.* **1994**, *50* (1), 1–16. https://doi.org/10.1016/0378-4371(90)90213-C.
(20) Banchio, A. J.; Nägele, G. Short-Time Transport Properties in Dense Suspensions: From Neutral to Charge-Stabilized Colloidal Spheres. *J. Chem. Phys.* **2008**, *128* (10). https://doi.org/10.1063/1.2868773.
(21) 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* (29), 11815–20. https://doi.org/10.1073/pnas.1107287108.
(22) 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. https://doi.org/10.1051/epjconf/20158302005.
(23) Cho, C. H.; Urquidi, J.; Singh, S.; Robinson, G. W. Thermal Offset Viscosities of Liquid H2O, D2O, and T2O. *J. Phys. Chem. B* **1999**, *103* (11), 1991–1994. https://doi.org/10.1021/jp9842953.
(24) Frick, B. The Neutron Backscattering Spectrometer IN16 at ILL—High Energy Resolution with High Intensity and Excellent Signal-to-Noise Ratio. *Neutron News* **2002**, *13* (2), 15–22. https://doi.org/10.1080/10448630208218478.
(25) Frick, B.; Mamontov, E.; Eijck, L. van; Seydel, T. Recent Backscattering Instrument Developments at the ILL and SNS. *Z. Für Phys. Chem.* **2010**, *224* (1–2), 33–60. https://doi.org/10.1524/zpch.2010.6091.
(26) Frick, B.; Combet, J.; van Eijck, L. New Possibilities with Inelastic Fixed Window Scans and Linear Motor Doppler Drives on High Resolution Neutron Backscattering Spectrometers. *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip.* **2012**, *669*, 7–13. https://doi.org/10.1016/J.NIMA.2011.11.090.
(27) Uehara, K.; Takeshita, H.; Kotaka, H. Hydrogen Gas Generation in the Wet Cutting of Aluminum and Its Alloys. *J. Mater. Process. Technol.* **2002**, *127* (2), 174–177. https://doi.org/10.1016/S0924-0136(02)00121-8.
(28) Grimaldo, M.; Roosen-Runge, F.; Hennig, M.; Zanini, F.; Zhang, F.; Jalarvo, N.; Zamponi, M.; Schreiber, F.; Seydel, T. Hierarchical Molecular Dynamics of Bovine Serum Albumin in Concentrated Aqueous Solution below and above Thermal Denaturation. *Phys. Chem. Chem. Phys.* **2015**, *17* (6), 4645–4655. https://doi.org/10.1039/C4CP04944F.
(29) Grimaldo, M.; Lopez, H.; Beck, C.; Roosen-Runge, F.; Moulin, M.; Devos, J. M.; Laux, V.; Härtlein, M.; Da Vela, S.; Schweins, R.; Mariani, A.; Zhang, F.; Barrat, J.-L.; Oettel, M.; Forsyth, V. T.; Seydel, T.; Schreiber, F. Protein Short-Time Diffusion in a Naturally Crowded Environment. *J. Phys. Chem. Lett.* **2019**, *10* (8), 1709–1715. https://doi.org/10.1021/acs.jpclett.9b00345.
(30) Grimaldo, M.; Roosen-Runge, F.; Zhang, F.; Schreiber, F.; Seydel, T. Dynamics of Proteins in Solution. *Q. Rev. Biophys.* **2019**, *52*. https://doi.org/10.1017/S0033583519000027.
(31) Arnold, O.; Bilheux, J. C.; Borreguero, J. M.; Buts, A.; Campbell, S. I.; Chapon, L.; Doucet, M.; Draper, N.; Ferraz Leal, R.; Gigg, M. A.; Lynch, V. E.; Markvardsen, A.; Mikkelson, D. J.; Mikkelson, R. L.; Miller, R.; Palmen, K.; Parker, P.; Passos, G.; Perring, T. G.; Peterson, P. F.; Ren, S.; Reuter, M. A.; Savici, A. T.; Taylor, J. W.; Taylor, R. J.; Tolchenov, R.; Zhou, W.; Zikovsky, J. Mantid—Data Analysis and Visualization Package for Neutron Scattering and μ SR Experiments. *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip.* **2014**, *764*, 156–166. https://doi.org/10.1016/J.NIMA.2014.07.029.
(32) *NIST Handbook of Mathematical Functions*; Olver, F. W. J., National Institute of Standards and Technology (U.S.), Eds.; Cambridge University Press : NIST: Cambridge ; New York, 2010.
(33) Wales, D. J.; Doye, J. P. K. Global Optimization by Basin-Hopping and the Lowest Energy

Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *J. Phys. Chem. A* **1997**, *101* (28), 5111–5116. https://doi.org/10.1021/jp970984n.

(34) van de Weert, M.; Andersen, M. B.; Frokjaer, S. Complex Coacervation of Lysozyme and Heparin: Complex Characterization and Protein Stability. *Pharm. Res.* **2004**, *21* (12), 2354–2359. https://doi.org/10.1007/s11095-004-7689-z.

(35) Krebs, M. R. H.; Devlin, G. L.; Donald, A. M. Protein Particulates: Another Generic Form of Protein Aggregation? *Biophys. J.* **2007**, *92* (4), 1336–1342.

(36) Eanes, E. D.; Glenner, G. G. X-Ray Diffraction Studies on Amyloid Filaments. *J. Histochem. Cytochem.* **1968**, *16* (11), 673–677. https://doi.org/10.1177/16.11.673.