molecular pharmaceutics

pubs.acs.org/molecularpharmaceutics

Article

Hydration-Induced Structural Changes in the Solid State of Protein: A SAXS/WAXS Study on Lysozyme

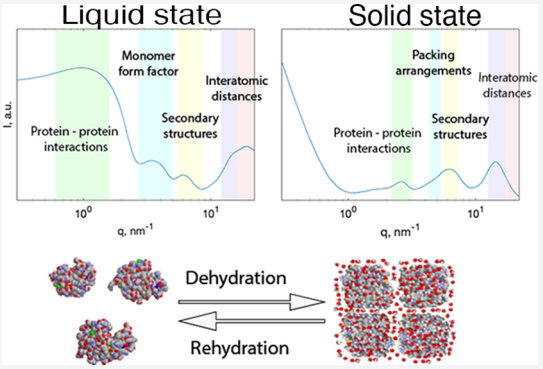
Tuan Phan-Xuan,* Ekaterina Bogdanova, Anna Millqvist Fureby, Jonas Fransson, Ann E. Terry, and Vitaly Kocherbitov*

Cite This: *Mol. Pharmaceutics* 2020, 17, 3246−3258

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The stability of biologically produced pharmaceuticals is the limiting factor to various applications, which can be improved by formulation in solid-state forms, mostly via lyophilization. Knowledge about the protein structure at the molecular level in the solid state and its transition upon rehydration is however scarce, and yet it most likely affects the physical and chemical stability of the biological drug. In this work, synchrotron small- and wide-angle X-ray scattering (SWAXS) are used to characterize the structure of a model protein, lysozyme, in the solid state and its structural transition upon rehydration to the liquid state. The results show that the protein undergoes distortion upon drying to adopt structures that can continuously fill the space to remove the protein−air interface that may be formed upon dehydration. Above a hydration threshold of 35 wt %, the native structure of the protein is recovered. The evolution of SWAXS peaks as a function of water content in a broad range of concentrations is discussed in relation to the structural changes in the protein. The findings presented here can be used for the design and optimization of solid-state formulations of proteins with improved stability.



KEYWORDS: *solid-state protein, biotherapeutics, hydration, dehydration, small- and wide-angle X-ray scattering, distorted structure*

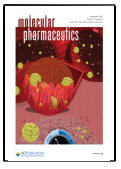
1. INTRODUCTION

Protein drugs (biotherapeutics, biologics) have become an increasingly important class of drugs for the treatment of many diseases.[1] Biologics are typically used as liquid formulations for injection but also as topical creams for local administration. However, the physical and chemical degradation of biologics in aqueous solution is an impediment to their use and their shelf life, thus limiting the wider adoption of these treatments. To improve stability, protein products are usually dried to a solid state by, for example, lyophilization (freeze-drying), commonly used in industrial manufacturing of these drugs, which entraps the protein in a glassy sugar matrix. This generally causes slower degradation kinetics of the protein and thus the shelf life may be increased to 2−3 years. However, lyophilization does not completely stop degradation and can also have an effect on the native structure and activity of the protein upon resolubilization. One of the possible reasons for this is that a small amount of water remains in the sample after lyophilization.[2−5] Thus, in this study, we will examine the role of water and water content on the stability of dried proteins. As controlling the degree of dehydration via lyophilization is difficult to achieve, we start with dry, lyophilized protein powders and rehydrate them to the required water content.

Proteins in the dehydrated state possess different properties compared to proteins in solution. When strongly dehydrated, proteins may not only reversibly or irreversibly lose their biochemical function but also undergo changes in dynamics (as evidenced by changes in the glass transition) and structure. Secondary structure changes can be monitored using vibrational spectroscopy, which has shown that the α-helix content decreases while the amount of β-sheet increases upon dehydration.[6−8] For many globular proteins, including lysozyme, these changes are reversible.

The dehydration-induced structural transitions in proteins can be explained by their inability to continuously fill the space with the native globular shapes in the absence of water. The inability to fill the space is equivalent to the formation of voids, which would result in a drastic rise of the Gibbs energy of the system due to high surface tension of the protein−air interface. To prevent this, the protein molecules adopt structures that can continuously fill the space.[8,9]

Received: March 31, 2020
Revised: August 4, 2020
Accepted: August 5, 2020
Published: August 5, 2020



https://dx.doi.org/10.1021/acs.molpharmaceut.0c00351
Mol. Pharmaceutics 2020, 17, 3246−3258

Small-angle scattering (X-ray and neutron) is a powerful technique to study the native size and shape of proteins in solution.[10] Typically, these experiments are conducted at dilute (<10 mg/mL) protein concentrations, although at higher protein concentrations, some information can be gained about protein−protein interactions and aggregation of proteins. From the scattering of a dilute protein solution, several parameters can be directly obtained providing information about the size, oligomeric state, and overall shape of the molecule. In combination with theoretical simulation, it is possible to determine a three-dimensional (3D) structure of the molecules.[11]

In more concentrated liquid solutions, the protein might behave differently due to the interplay between short-range attraction and long-range repulsion interactions. Stradner and co-workers[12] have reported the existence of a low-angle interference peak in small-angle X-ray and neutron scattering (SAXS and SANS) for concentrated solutions of lysozyme at low ionic strength and close to the physiological pH. They found that the position of this interference peak was essentially independent of the protein concentration and attributed these unexpected results to the presence of equilibrium clusters of individual protein monomers. In contrast to these reported interpretations, Shukla and co-authors[13] repeated the same experiment using SAXS and SANS and found that the correlation peak has a concentration dependence, which might be due to increased repulsive interactions with increasing protein concentration. They concluded that these data do not reveal any equilibrium clusters, therefore.

As is mentioned above, proteins at dehydrated conditions may have a different conformation when compared to proteins at a high water content. These differences could also be connected to the protein stability because of the formation of the glassy state upon strong dehydration. There are few studies on the structural transition in the solid state of protein using scattering techniques: for example, SANS has been carried out on dried formulations in the presence of excipients[14−16] and also SAXS.[17,18] However, no SAXS and wide-angle X-ray scattering (WAXS) studies discussing the change in protein morphology under dehydrated conditions or dehydration-induced intramolecular structural changes are found in the literature.

Experimental X-ray scattering studies of proteins at low water contents, including solid-state forms, are associated with many practical difficulties. In particular, strong scattering from the surface of particles can hide structural information that could be obtained at small angles (low q-values). Furthermore, traditional methods of background subtraction used for the liquid state cannot be used because of, among other factors, the low amount or absence of the liquid medium (liquid water or buffer) and the difficulty of completely filling of the capillary with a dry, powder sample. Despite these complications, in this work, we will present a systematic study on the structural changes of lysozyme from dehydrated to fully hydrated conditions at room temperature using SAXS and WAXS to shed some light on the change of the morphology of the protein at different water contents. Lysozyme was chosen as a model protein because of its well-known structure in the fully hydrated native condition as well as its well-known thermodynamic behavior at low water contents[19] and its reversible behavior during hydration and dehydration. Moreover, it can be used as a potential biotherapeutic agent for cancer[20] and antimicrobial treatments.[21]

2. MATERIALS AND METHODS

2.1. Materials. Lysozyme from chicken egg white (CAS number 12650−88−3, Lot SLBL 7146V, 100% purity) was purchased from Sigma-Aldrich and used as received. Another batch of lysozyme (Lot SLBZ 2146) with a lower purity (96%) needed dialysis and lyophilization, as detailed below, to remove the impurities. Milli-Q purified water was used for all experiments.

Dialysis. Lysozyme powder (Lot SLBZ 2146) was dissolved in MiliQ water to reach a stock solution of 4 wt % protein at 20 °C. After all powder had dissolved, the stock solution was filtered through a 0.2 μm Acrodisc syringe filter to remove large-size aggregates before transferring to an Amicon ultracentrifugal filter tube (Merck) with 3 kDa cutoff and a maximum initial sample volume of 15 mL. Water was exchanged several times (8−10 times) using centrifugation (Becker, 4000g, 30 min). The solution was then adjusted to 2 wt % protein prior to lyophilization.

Lyophilization. The aqueous protein solutions $c_{protein}$ = 20 mg/mL were lyophilized in a 4 mL glass vial (liquid filled volume = 2 mL) in an Epsilon 2−6D LSCplus (Martin Christ) freeze-dryer. During the freezing step, the shelf temperature and cooling rate were set at −45 °C and 0.2 °C/min, respectively, and held at −45 °C for 2 h. The primary drying was done at −35 °C and 0.1 mbar chamber pressure for 10 h. During the secondary drying, the temperature was raised to −20 °C at a heating rate of 0.2 °C/min and kept at −20 °C for 5 h at 0.01 mbar. The temperature was then raised to 25 °C with a ramp rate of 1 °C/min and held for 6 h. At the end of the freeze-drying cycle, the chamber was filled with dry nitrogen, and the vials were sealed and stored at −20 °C. This process produces a solid cake with no collapse and a typical moisture content of less than 3%.

Preparation of Rehydrated Lysozyme. Lysozyme was first dried in a desiccator for 48 h using molecular sieves (type 3Å) as sorbents. The amount of residual water in the powder after drying was determined using thermal gravimetry analysis (Q500, TA Instrument). Dry lysozyme powder of 20−100 mg was rehydrated to the desired water content to produce samples covering the hydration range from 2.5 to 99.6 wt % water content.

Specifically, the rehydration of lysozyme was carried out as follows. Samples with low water contents were prepared by spreading appropriate amounts (ca. 100 mg) of dried lysozyme on the cavity of a weighed glass microscopy slide. The slide was placed in a desiccator with a vapor atmosphere from various saturated salt solutions (corresponding water activity is given in parentheses): $Mg(NO_3)_2$ (0.33), NaCl (0.75), KCl (0.85), KNO_3 (0.95), and K_2SO_4 (0.97) at ambient temperature for a period of 5−7 days until the sample's mass reached stable values. The amount of water uptake was determined by measuring the total mass of the powder and the slide before and at the end of the incubation (Figure S1). The powder was then transferred to glass capillaries (Hilgenberg GmbH, Germany) of diameter of 1.5 mm (transfer time about 5−10 min). To compensate for possible water loss or gain during the transfer to capillaries, all of the capillaries were then put back to desiccators with saturated salt solutions for 2−3 days to ensure that the samples returned to equilibrium. Capillaries with hydrated lysozyme were sealed with vacuum grease and then by a layer of nail polish (neither the grease nor the nail polish came into contact with the lysozyme and the grease

https://dx.doi.org/10.1021/acs.molpharmaceut.0c00351
Mol. Pharmaceutics 2020, 17, 3246−3258

prevents the solvent from the nail polish interacting with the sample) and left to equilibrate at room temperature. To ensure the uniformity of the water distribution in lysozyme particles of different sizes, we allowed the samples to equilibrate for a week after vapor sorption. With this procedure, samples with water contents up to 32 wt % were obtained.

For water concentrations of 33−50 wt % in which the protein is in the form of a paste, the samples were prepared in differential scanning calorimetry (DSC) aluminum pans by weighing appropriate amounts (ca. 10 mg) of dried lysozyme, adding the required amounts of liquid water and immediately sealing the DSC pan by crimping.

For water contents above 55 wt % in which the samples are in a liquid solution state, they were prepared one day before the experiment and kept at 5 °C to avoid degradation.

2.2. Simultaneous SAXS/WAXS Measurements. Two SAXS setups have been used, one of which was a standard SAXS for liquid samples and the other one of which was used for SAXS/WAXS measurements.

Small-angle X-ray scattering measurements for aqueous dilute samples were performed at the Swing beamline at Soleil Synchrotron (France) using an Aviex CCD detector. The sample-to-detector distance was selected to cover a q-range from 0.05 to 6.83 nm^{-1}, with $q = 4\pi \sin\theta/\lambda$, where 2θ is the scattering angle, and λ = 1.03 Å is the X-ray wavelength. Samples of 2 and 4 g/L lysozyme were flowed through a capillary cell[22] to keep constant the sample thickness and the parasitic scattering arising from the walls of the capillary.

Small- and wide-angle X-ray scattering (SAXS/WAXS) experiments with concentrated liquid and solid samples were carried out at the NCD-SWEET beamline at ALBA synchrotron (Spain) using an X-ray wavelength of 1 Å. Two-dimensional (2D) SAXS and WAXS images were recorded in a SAXS detector PILATUS 1M (Dectris) located at a sample−detector distance of 2.69 m and a WAXS detector LX255-HS (Rayonix) at 0.12 m sample−detector distance, respectively. The scattering vector, q, was calibrated with a silver behenate sample. Reported scattering profiles $I(q)$ were obtained by radially averaging 2D SAXS images using the data reduction scripts at the beamline. With this setup, a broad range of q-values—q = 0.007−5.6 nm^{-1} for SAXS and q = 5−85 nm^{-1} for WAXS—was covered. The effect of radiation damage was also checked before the measurement by exposing the diluted solution samples in a series of very short time intervals for a continuous exposure to the X-ray beam (2 min). The experimental exposure time and the number of frames were then selected for all samples to assure that there was no sign of aggregation or structural changes due to radiation damage or heating caused by X-ray.

Data Processing. A set of MATLAB scripts was created to perform different steps in the data processing.

SAXS and WAXS data merging was performed by adjusting the intensity of the entire WAXS data so that it matched the SAXS data in the overlapping q-range.

The scattering from an empty sample container (glass capillary or aluminum pan) was subtracted from the scattered intensities of the sample. Since the transmitted X-ray intensity values could not be collected due to technical issues with the diode on the beam stop, the data were normalized by the incoming beam intensity via the following equation:

$$I_s = I_{sample} - I_{empty} \times coef \quad (1)$$

where I_s is the subtracted intensity, I_{sample} is the scattered intensity of the sample, I_{empty} is the scattered intensity from the empty capillary/pan, and coef is an empirical coefficient to compensate for variations in capillary/pan thicknesses and powder filling density.

To keep the scattering patterns consistent for all samples from dried to liquid states, the buffer subtraction is not performed. Water is only subtracted for a very diluted solution of lysozyme to determine the dimensions of the protein monomer.

Peak deconvolution: The data were plotted on the log I vs q scale and fitted with a combination of several Gaussian components and a linear background via the equation

$$y = a \times e^{-(x-b)^2/2c^2} \quad (2)$$

where y is the subtracted intensity, x is the q-value, a is the peak height, b is the peak position, and c is a fitting parameter related to the full width at half-maximum (FWHM), $c = \frac{FWHM}{\sqrt{2 \times \log 2}}$.

Most of the scattering curves were fitted by applying the script for the range of q-values 2−22 nm^{-1}. Some data obtained from the high-water content samples (above 60 wt % water) were divided into two separate ranges for a better selection of the baseline.

3. RESULTS

3.1. Overview of the Phase Behavior of the Lysozyme−Water System. Upon rehydration of the dried lysozyme powder, different states of the samples can be observed visually depending on the amount of water. For water content below 35 wt %, the samples are in the solid powder state. Between 35 and 55 wt %, it becomes a soft paste and we also observe a two-phase system consisting of a solution and a soft gel of lysozyme. Above 55 wt % water, the samples appear to be clear aqueous solutions.

3.2. SAXS/WAXS Results. A SAXS experiment on the lysozyme in dilute solution (xH_2O = 99.6 wt %) was performed to confirm the size of the lysozyme molecule at these conditions (Figure 1). The data were fitted with a sum of two models: an ellipsoid model for the high q-value regime (q > 2 nm^{-1}) and a Guinier−Porod model for the lower q-value regime. The fitting of the data with an ellipsoidal model yields dimensions (half-axis) of 2.72 nm (±0.008 nm) × 1.55 nm (±0.003 nm), in reasonable agreement with published results for lysozyme.[23] The volume of lysozyme determined from these data is 27.37 (±0.187) nm^3. This value is higher than the value obtained via the density of lysozyme of 1.4 g/cm^3 ($V_{lys} = \frac{M_{lys}}{N_A \times \rho_{lys}}$ = 16.96 nm^3). The discrepancy probably arises from the fact that the first hydration shell of the protein has a higher density than the bulk water,[24] which apparently increases the protein volume when probed by X-ray scattering. It should also be noticed that even at this highly diluted condition, a small fraction of lysozyme aggregates was observed, which contribute to the low q upturn. A detailed analysis of the fitting can be found in the Supporting Information (SI) (Figure S2). A generated scattered intensity of lysozyme from PDB using Crysol was also plotted in the same graph and showed a good agreement with the experimental data in the middle q-range. We, however, note an obvious limitation of the ellipsoidal model since it is not

https://dx.doi.org/10.1021/acs.molpharmaceut.0c00351
Mol. Pharmaceutics 2020, 17, 3246−3258

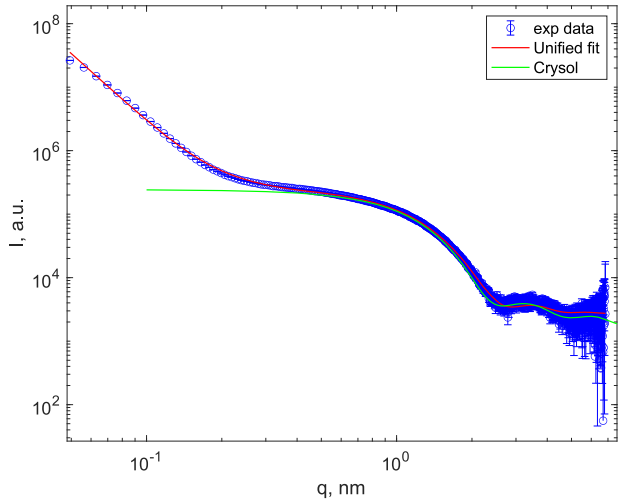


Figure 1. SAXS pattern of lysozyme in the dilute aqueous solution $C_{protein}$ = 4 g/L. The solid red line is the fit to SAXS data using a sum of a Guinier−Porod and an ellipsoidal model. The green line is the generated intensity from the lysozyme structure in PDB using Crysol.

able to describe the internal structure of the protein molecule and the intricate details of its surface. As a consequence, there is a clear difference between the ellipsoidal model curve and the Crysol-generated intensity at high q-values.

The structure of lysozyme at different levels of hydration, ranging from a few percents of water in powder to solution, was investigated using simultaneous SAXS and WAXS. The scattering data cover a q-range from 0.1 to 30 nm^{-1}, in which the intra- and intermolecular correlations of length scale ranging from 0.2 to 60 nm could be probed.

An overview of the scattering patterns of lysozyme at different water contents is shown in Figure 2. Most of the structural information in the scattering data appears in the form of correlation peaks, of which the q-value of maximum intensity together with the corresponding correlation distance is noted in Table 1. The nature of the peaks will be described in the discussion part.

At water contents below 14 wt %, where the samples are still a powder, the scattering curves show a power law dependence of the intensity at q < 1 nm^{-1} with a slope of −4 (Figure 2a). The presence of the so-called Porod slope can be explained by the scattering from the solid protein−air interface of lysozyme powder particles. At q > 1 nm^{-1}, various correlation peaks can be clearly identified: at q = 2.5 nm^{-1}, which corresponds to a

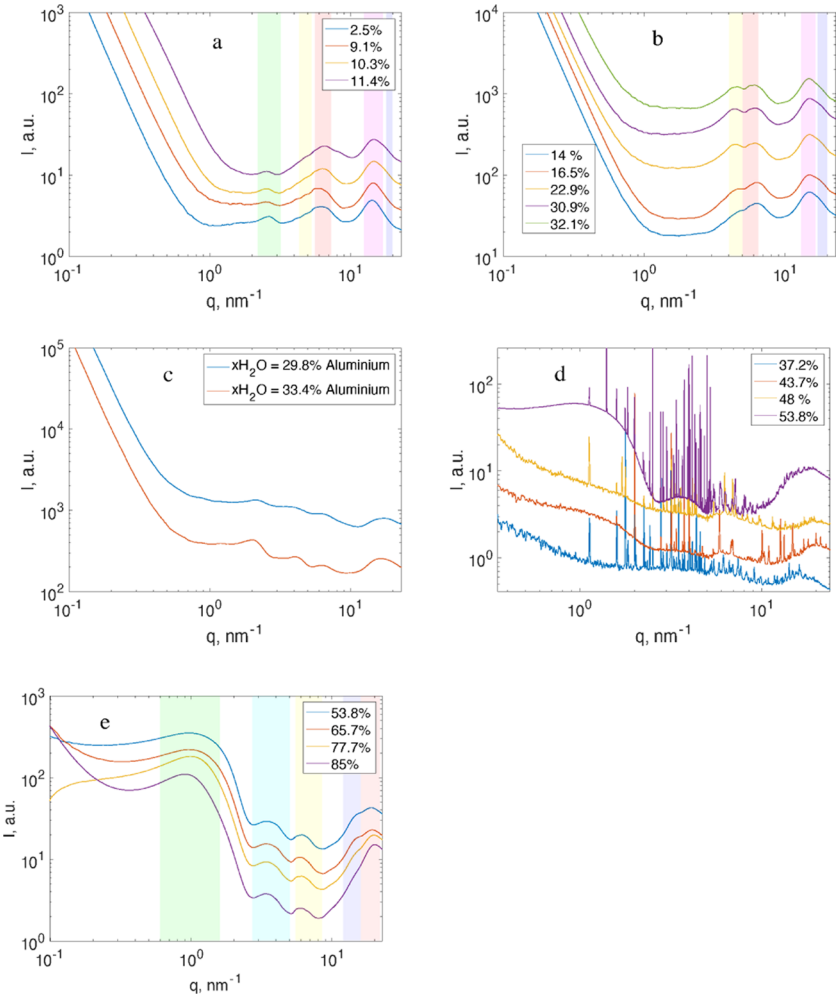


Figure 2. Hydration-level dependence of scattered intensity $I(q)$ of lysozyme at 25 °C. Samples shown in (a), (b), and (e) were studied in capillaries, while samples shown in panels (c) and (d) (water contents from 29 to 48 wt %) were studied in aluminum pans. For clarity, each SAXS curve is plotted with an offset in the vertical axis.

https://dx.doi.org/10.1021/acs.molpharmaceut.0c00351
Mol. Pharmaceutics 2020, 17, 3246−3258

Table 1. Assignment of SAXS/WAXS Peaks Obtained from Lysozyme Samples with Different Water Contents at 25 °C

peak position q (nm^{-1})	distance d (nm)	structural origin	water contents
0.90−0.97	6.30−7.00	protein−protein interaction in liquid solution	xH_2O > 50 wt %
2.09−2.61	2.41−3.01	intermolecular correlation; the d-value proportional to protein−protein distance	xH_2O < 35 wt %
3.00−4.00	1.57−2.09	first submaxima of the form factor of lysozyme in solution	xH_2O > 50 wt %
4.10−5.00	1.26−1.53	half of the protein−protein distance	xH_2O < 35 wt %, especially 14−35 wt %
5.50−5.80	1.08−1.14	second peak of the form factor including a hydration shell (low intensity)	all water contents
6.00−6.20	1.01−1.05	interactions between α helixes (similar to the lysozyme PDB structure)	all water contents
7.00−12.00	0.52−0.90	secondary structures	all water contents
14.00−15.00	0.42−0.45	interatomic distance (hydrophobic groups)	all water contents
17.00−18.50	0.34−0.37	interatomic distance (water and hydrogen-bonded atoms)	all water contents

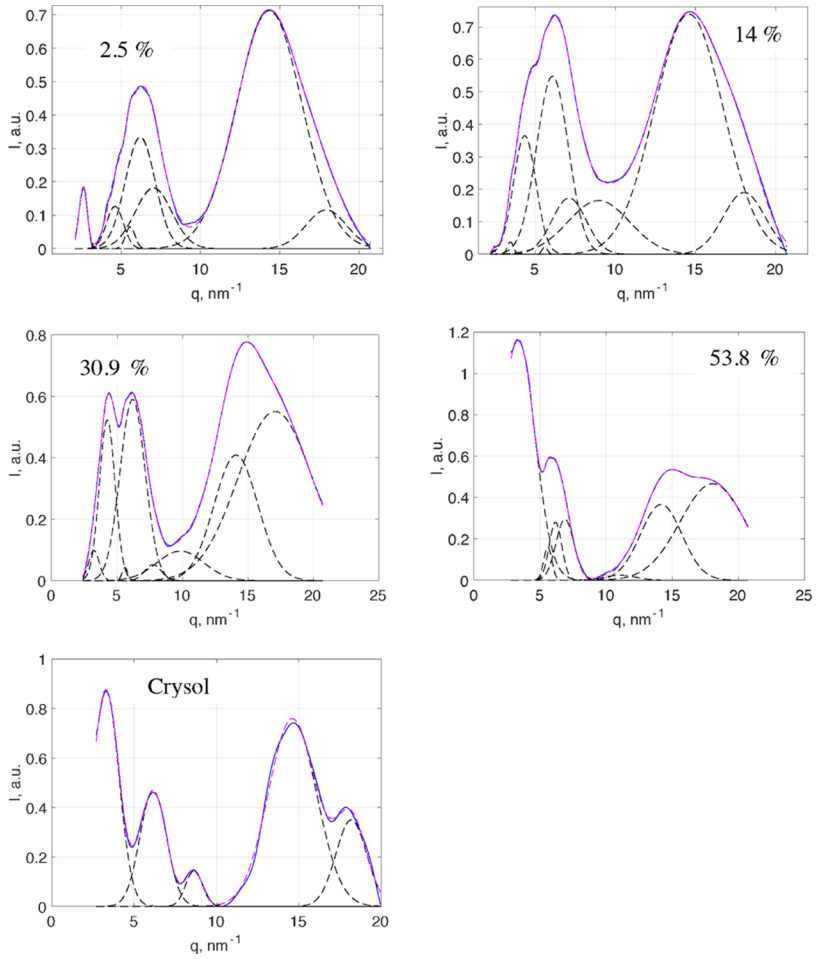


Figure 3. Gaussian deconvolution of the experimental scattering data for different water contents and theoretical data generated by Crysol from PDB.

distance of d = 2.5 nm, q = 6.1 nm^{-1} (d = 1 nm), and q = 14 nm^{-1} (d = 0.45 nm). The peak around q = 2.5 nm^{-1} slightly shifts from 2.6 to 2.4 nm^{-1} with increasing water content, while the peak at q = 14 nm^{-1} becomes broader with the appearance of a shoulder on the right-hand side of the peak position. The peak at q = 6.1 nm^{-1} shows concentration independence for this hydration regime.

With increasing water content from 14 to 35 wt %, the lysozyme samples are still a solid powder. Figure 2b represents the scattering patterns of lysozyme in this regime of hydration. At low q-values, the intensity has the same power law dependence on q as at lower water contents. However, at higher q-values, the peak at q = 2.5 nm^{-1} is absent and a new peak at q = 4.5 nm^{-1} appears, which shifts to lower q-values with increasing water content. It should be noted that the measurements on the lysozyme in the aluminum pan at different water contents from 16.9 to 34 wt % water show the presence of the peak at q = 2.5 nm^{-1} (Figure 2c). The peaks at q = 6.1 nm^{-1} and q = 14 nm^{-1} are also observed in this hydration regime in all sample holders. In addition, a new peak at q = 18 nm^{-1} starts to develop.

Figure 2d shows the scattering patterns of lysozyme in the range 35−55 wt % water in which the lysozyme samples become a soft paste at water contents below 40 wt % and a two-phase system of a soft paste and a solution of lysozyme above 40 wt % water. In this regime, the samples are very

https://dx.doi.org/10.1021/acs.molpharmaceut.0c00351
Mol. Pharmaceutics 2020, 17, 3246−3258

Table 2. Peak Positions Obtained by Fitting the Scattering Data in Figure 3 Using Gaussian Deconvolution

	q-range (nm^{-1})								
water (%)	2−3	3−4	4−5	5.3−5.8	6−6.2	6.5−9.0	9−12	13−15	17−19
2.5	2.60	3.54	4.60	5.57	6.20	7.04	9.46	14.34	17.93
14	2.45	3.46	4.35	5.60	6.08	7.12	8.97	14.60	18.05
30.9	2.10	3.24	4.26	5.61	6.20	7.76	9.85	14.10	17.09
53.8		3.33		5.58	6.16	6.87	11.04	14.62	18.07
Crysol		3.28			6.15	8.64		14.62	18.19

sticky, which makes the transfer to a capillary for SAXS/WAXS measurements difficult. Therefore, the samples are prepared in aluminum pans and sealed to prevent water loss. Figure 2d represents the scattering patterns in this hydration regime. Numerous Bragg peaks dominate the q-range from q = 1 nm^{-1} in the SAXS range up to the WAXS range, which are due to the lysozyme being partly crystalline at this level of hydration. As a result, the structural information of lysozyme in the form of correlation peaks cannot be identified in this range. Nevertheless, as a "baseline" under the Bragg peaks, we could observe a gradual recovery of the first submaxima of the lysozyme form factor around q = 3.5 nm^{-1} with increasing water content.

In the regime of water content above 55 wt %, the samples are clear aqueous solutions. The scattering patterns in this hydration regime (Figure 2e) show a broad peaklike feature at q = 1 nm^{-1} followed by another peak at q = 3.5 nm^{-1} in the SAXS range. The peak at 3.5 nm^{-1} originates from the first maxima of the monomer form factor of lysozyme in the native state and is also found in the scattering pattern from the dilute lysozyme solution, as well as from the PDB structure of lysozyme (see Figure 1). At larger q-values in the WAXS range, where the internal structures of lysozyme are probed, similar correlation peaks at q = 6.1 nm^{-1}, q = 14 nm^{-1}, and q = 18 nm^{-1} are found, as are observed in the lower hydration regime. The peak at q = 18 nm^{-1} is known to arise from the water and hydrogen-bonded atoms of the protein.[25] At the highest water content of 85 wt %, the water peak dominates the scattering in this q-range, and the peak at q = 14 nm^{-1} is not observed.

3.3. Gaussian Deconvolution of the Scattering Data.

To obtain more accurate peak positions, a peak deconvolution procedure is used to fit multiple Gaussian peaks to the scattering data, as well as a term for a linear baseline, as described in the Methods section. For most of the data, the q limit is fixed from 2 to 21 nm^{-1} for the deconvolution procedure. In some cases where the determination of a linear baseline is complicated for the whole q-range, we have separated the q-values for the deconvolution into two ranges, q_1 = 2 to 8 nm^{-1} and q_2 = 8 to 22 nm^{-1}. Examples of the Gaussian deconvolution for lysozyme at different water contents together with Crysol data and the deconvoluted peak positions are given in Figure 3 and Table 2, respectively. More complete data can be found in Table S1 (SI). The deconvolution of the theoretical data in a q-range of 2.7−20 nm^{-1} yielded five different peaks: at q = 6.15 nm^{-1}, at q = 14 nm^{-1}, at q = 18 nm^{-1}, and q around 3−4 and 8−12 nm^{-1}. These peaks are all present in the experimental data using the deconvolution. Apart from the peaks that are inherited from the lysozyme structure, the peaks arising at q = 2−3 and 4−5 nm^{-1} originated from the arrangement of lysozyme molecules.

The absolute values of the intensity of the peaks are not quantitatively considered here because of the difficulties associated with accurate background subtraction for solid samples, as explained in the previous sections. On the other hand, as the peak in the q-range 6.0−6.2 nm^{-1} is present in all data, it is possible to normalize the intensity to the intensity at q = 6.0−6.2 nm^{-1} to monitor the changes of peaks with changing water contents. In Figure 4, scattering data with the applied linear baseline for the deconvolution are plotted in a log-lin representation of the normalized intensity as a function of q-values for different water contents.

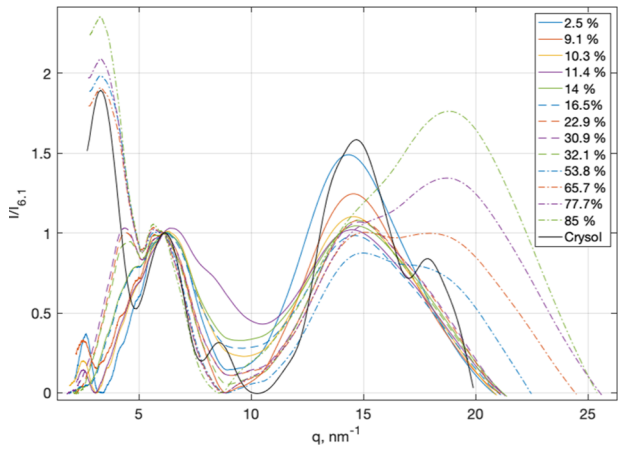


Figure 4. Hydration-level dependence of the normalized scattering intensity $\log(I(q)/I(6.1))$ of lysozyme at the water content indicated in the legend. A similar representation of the protein form factor generated by Crysol is plotted for comparison.

The evolution of the peak positions as a function of water contents is shown in Figure 5. Figure 5a shows a linear decrease in the q-values of the peak position in the q-range of 2−3 nm^{-1} with increasing water content for both samples measured in capillaries (blue symbol) and in aluminum pans (red symbol). This correlation peak appears only at water content below 35 wt % when the protein samples are still in a powder form. Above 11 wt % water, a steeper decrease in the q-values of the peak is observed for samples in aluminum pans. It should also be noticed that in the samples measured in the capillaries, the intensity of the peak in this q-range is very small and often could not be obtained by deconvolution due to the dominance of adjacent peaks. On the contrary, for samples measured in aluminum pans, the peak is clearly observed.

In Figure 5b, the evolution of the peak positions in the q-range of 3−4 nm^{-1} is shown. A general trend of a decrease in the q-values of the peak positions could be identified in two different water regimes at below 35 wt % and above 55 wt %. The gap in the range of water content 35−55 wt % is due to the fact that the lysozyme is at least partly crystallized at this hydration level, and as a result, the scattering signal is dominated by the intense Bragg peaks from lysozyme crystals.

https://dx.doi.org/10.1021/acs.molpharmaceut.0c00351
Mol. Pharmaceutics 2020, 17, 3246−3258

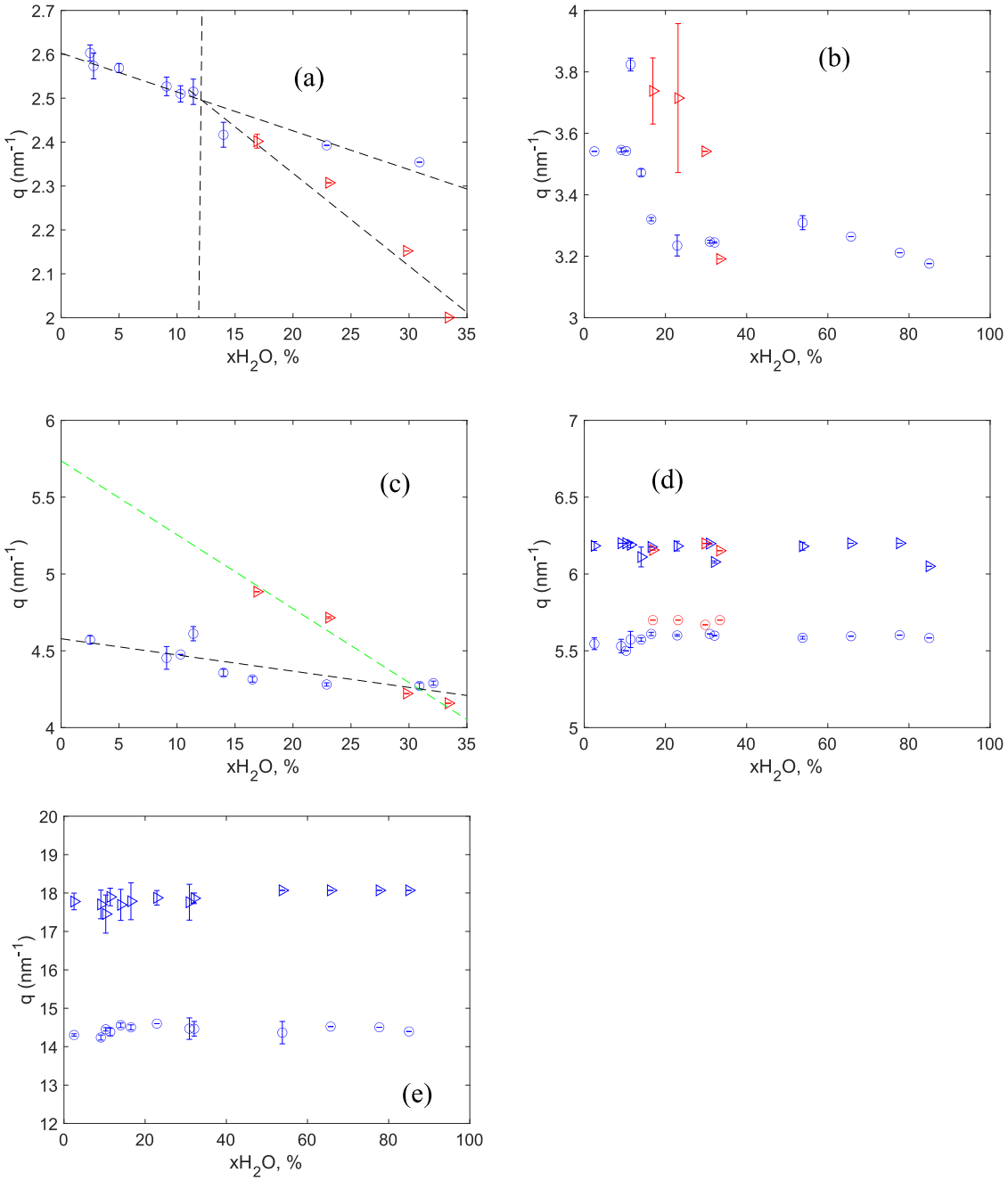


Figure 5. Evolution of the peak positions after the Gaussian deconvolution of scattering data of samples measured in capillaries (blue symbols) and in aluminum pans (red symbols) in the q-ranges 2−3 nm^{-1} (a), 3.0−4.0 nm^{-1} (b), 4.0−5.0 nm^{-1} (c), 5.5−7.0 nm^{-1} (d), 14−15 nm^{-1}, and 17−19 nm^{-1} (e). Error bars are determined as described in Section 3.3.

The evolution of the peak positions in the q-range of 4−5 nm^{-1} is shown in Figure 5c. Similar to Figure 5a, a linear decrease is observed in the q-values of the peak position with increasing water content. The results from samples measured in aluminum pans show a steeper decrease than those obtained from samples in capillaries.

Figure 5d represents the results obtained by the deconvolution of the scattering peaks at q = 5.5 nm^{-1} and q = 6.5 nm^{-1}. These two peaks do not change for all of the water contents and are independent of the type of sample containers.

Figure 5e represents the evolution of the peaks at q = 14 nm^{-1} and q = 18 nm^{-1}. These peaks are present at all water contents and show a slight increase of q-values on increasing the water content. From the deconvolution, a low-intensity peak in the q-range of 8−10 nm^{-1} (SI), which could be related to the secondary structure of lysozyme, was obtained.

4. DISCUSSION

4.1. SAXS/WAXS Peak Positions and Structural Organization of Lysozyme on Different Length Scales at Different Water Contents. The positions of SAXS/WAXS peaks at different water contents provide valuable information about the evolution of the protein structure upon hydration/dehydration. The peaks registered in our experiments correspond to a broad range of characteristic length scales: from those exceeding protein sizes down to typical

https://dx.doi.org/10.1021/acs.molpharmaceut.0c00351
Mol. Pharmaceutics 2020, 17, 3246−3258

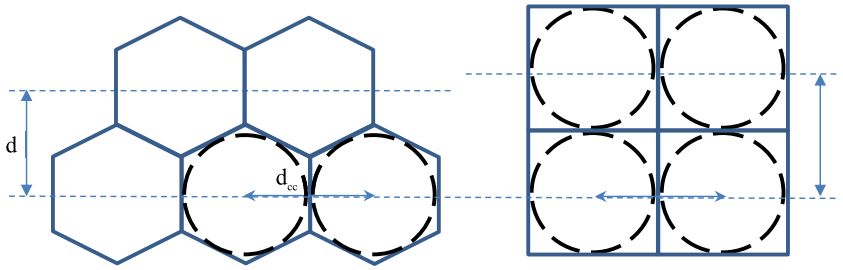


Figure 6. Schematic of the possible shape of distorted protein molecules and imaginary spheres. The horizontal arrows indicate the protein−protein distance d_{cc}, while the vertical arrows in the hexagonal structure are the interplanar distance d, which corresponds to the peak around q = 2.5 nm^{-1}.

interatomic distances (see Table 1). Based on these characteristic lengths and also on the literature data, we suggest that the peaks seen in the SAXS range (q-values below 5 nm^{-1}) arise from scattering from the overall shape of the protein molecules and from protein−protein correlations. The peaks in the WAXS range (q-values above 5 nm^{-1}) arise from the intramolecular structure of the protein, including the secondary structure and interatomic correlations. Below, we will discuss further details, starting from the evolution of the overall shape and intermolecular arrangement of proteins as dependent on water content; then, we will discuss the effect of hydration on the intramolecular structure of proteins based on the properties of WAXS peaks.

4.2. Evolution of the Shape and Intermolecular Arrangement of Lysozyme Molecules as a Function of Hydration. *4.2.1. Nanostructure of Solid Lysozyme Powder at Water Content Below 35 wt % (Peaks at q = 2.6 and 4.5 nm^{-1}).* In the dried powder and at low water content, the correlation peak at q-values around 2.5 nm^{-1} (Figure 2a) is clearly seen. A peak in a similar q-range has been reported before by SANS in lyophilized formulations of lysozyme as well as mAb formulations and was characterized as a protein−protein interaction peak.[14−16] For a better understanding of the intermolecular structure corresponding to this peak, it is instructive to consider it in relation to the 4.5 nm^{-1} peak that has a similar dependence on water content (see Figure 5a,c) and thus also arises from protein−protein interactions. In this section, we will consider these peak positions in the dry limit and in the following section evaluate their dependence on water content. The d-value ($d = 2\pi/q$) corresponding to the 2.6 nm^{-1} peak is around 2.41 nm, which roughly corresponds to the size of the lysozyme molecule, while the d-value of the 4.59 nm^{-1} peak (1.37 nm) is closer to half of the lysozyme molecule size. The ratio of the d-values mentioned above is 0.87, which is very close to $\sqrt{3}/2$. Such a ratio is typical for systems exhibiting hexagonal arrangements; see, for example, Figure 6 (left-hand side). In this 3D arrangement, hexagonal layers can form, for example, hexagonal close packed (HCP) and face-centered cubic (FCC) structures, which provide the highest possible packing densities. According to the schematic in Figure 6, the distance between the protein molecule centers d_{cc} is 2.78 nm, corresponding to the 2.6 nm^{-1} peak divided by 0.866. We note, however, that the arrangement described here corresponds only to the short-range order and the long-range order is absent at these conditions.

In some samples (mostly in experiments with Al pans and at slightly higher water contents above 14%), the ratio of the q-values of the two peaks was close to 2 (see Figure 5a,c). This suggests that two different structures could be formed depending on the history of the sample and the method of its preparation. An example of the structure where the destructive interference would not eliminate the peak corresponding to the center-to-center distance is shown in Figure 6 (right-hand side). The arrangement of the molecules in the amorphous sample can be between the structures shown in this figure and can also change with concentration.

Below, we will try to further relate the protein−protein correlation distance to the expected shape and 3D packing of the molecules.

Protein molecules in solution can move due to thermal fluctuations (Brownian motion), and the protein solutions can macroscopically flow when external forces are applied. However, upon gradual dehydration, when the volume fraction of the molecules increases above certain critical values, they are no longer freely moving, and the system starts to behave as an amorphous solid. At these conditions, the molecules are sometimes referred to as "arrested".[26] For lysozyme, it was recently reported that the protein concentration of the arrest transition is 35 vol % (which corresponds to water content of about 57 wt %).[27] The lysozyme molecule still adopts a size and shape close to the native lysozyme (observed in dilute solutions) at these conditions. Upon stronger dehydration, as has been shown using other techniques such as Fourier transform infrared (FTIR)[6,7] and Raman techniques,[8] the protein molecules assume different conformations than in liquid. Moreover, freeze-dried proteins absorb 2 orders of magnitude less nitrogen than water,[28] which strongly suggests that they adjust their shapes to fill the space. These shapes are, however, difficult to establish experimentally. Due to extreme crowding and the absence of a clear solvent−protein interface, the form factors of the molecules do not contribute to the experimentally observed scattering intensity. The only structural information available is the correlation peaks' positions.

Ignoring the subnanometer details, the result of the deformation of the lysozyme molecules can be compared with Voronoi tessellation of the 3D space. The resulting partition will, however, be dependent on the positions of the centers (generating points). The information obtained from the correlation peaks from scattering experiments describes only the distances between these centers and the interplanar distances but not the particular details of the geometrical arrangements. However, for practical applications such as in pharmaceuticals, protein conformation is the most valuable information. Hence, we will try to evaluate possible molecular conformation based on the data on the local arrangement of the molecules in the lattice structure. Here, we will test several ordered and disordered arrangements and deduce the resulting

https://dx.doi.org/10.1021/acs.molpharmaceut.0c00351
Mol. Pharmaceutics 2020, 17, 3246−3258

distance between the centers. As a mathematical abstraction, we draw imaginary spheres inside the structures obtained after distortion (see Figure 6). Since the packing of equally sized spheres is a well-known mathematical problem, we will use them for testing possible options for the packing of protein molecules. For lysozyme, this is an approximation since lysozyme has an ellipsoidal shape in dilute solutions. On the other hand, in the absence of water, the new adapted conformation most probably does not have the same aspect ratio as the native shape in solution due to the properties and spatial arrangement of α and β domains of the lysozyme molecule.

The volume fraction, φ, occupied by a sphere is

$$\varphi = \frac{V_{\text{sphere}}}{V_{\text{total}}} \tag{3}$$

where V_{total} is the total volume (continuously filled with lysozyme molecules). This parameter could be determined by dividing the molecular mass by the density of the protein (we will use the value of 1.4 g/cm^3). Then, using the formula for the volume of a sphere, one can write

$$\varphi = \frac{V_{\text{sphere}}}{V_{\text{total}}} = \frac{N_A \times \rho \times \pi \times d_{cc}^3}{M_{\text{lys}} \times 6} \tag{4}$$

From this, it follows that the sphere diameter and also the protein−protein distance is

$$d_{cc} = \sqrt[3]{\frac{\varphi \times M_{\text{lys}} \times 6}{N_A \times \rho \times \pi}} \tag{5}$$

Using eq 5, one can calculate the distances corresponding to different packing geometries, if the volume fraction of the packing is known. One should stress, however, that this volume fraction does not represent the concentration of the protein, but rather a mathematical abstraction for testing different packing arrangements.

In Table 3, we illustrate different options for the packing of the lysozyme molecules and the correlation distances

Table 3. Center-to-Center Distance between Solid Spheres after Deformation

packing	φ	d_{cc} (nm)
cubic	0.5236	2.5693
hexagonal close packing and face-centered cubic	0.74048	2.8840
random close packing	0.64	2.7472
random loose packing	0.55	2.6118

corresponding to them. In all cases, the calculated distances are in reasonable agreement with the protein−protein distance observed experimentally. The best agreement is seen in the cases of HCP, FCC, and random close packing. This indicates that the short-range order in lysozyme at low water contents can be similar to local arrangements in those three structures. One should however keep in mind that a discrepancy may arise from the fact that the lysozyme protein has an ellipsoidal shape in dilute solutions, which may be partly retained in the solid state. It can be expected that the "stretching-out" of the cells obtained by tessellation of the space (keeping their volumes constant) will decrease the shortest protein−protein distances.

Dependence of the Intermolecular Structure on the Water Content. Upon being exposed to water vapor at controlled relative humidity, the dried powder of lysozyme absorbs water and undergoes swelling as water enters the space between the protein molecules. It is important to realize that the protein structure is not opened by hydration or dehydration, but to some extent, the molecules can change their conformations while still being folded. The amount of water uptake has an important effect on the peak position and the corresponding interprotein distances. Indeed, the positions of the correlation peaks progressively shift to lower q-values, which correspond to an increase of the protein−protein distance by 0.3 nm, as observed in Figure 7a (blue data points).

It is well known that addition of water leads to a glass transition of the protein.[29] For lysozyme, it has been reported that the glass transition occurs at a water content of 11 wt %.[30] It has also been reported that lysozyme undergoes a glass transition in a broad range of water contents from 11 to 20 wt % at 25 °C instead of a stepwise transition, as observed in synthetic polymers and thermally denatured proteins.[31] The authors also reported a critical water concentration of 35 wt % at which the properties of the system clearly change upon slow dehydration. Later, it was shown that this concentration corresponds to the minimum water content required for filling the space with natively shaped lysozyme molecules.[8]

As the water content increases, the distance between the centers of the protein molecules increases too. We will assume that the swelling of the protein sample is isotropic and in general can be described via the following equation:

$$d_{cc}^3 = a_0 + a_1 r_w \tag{6}$$

where d_{cc} is the distance between two neighboring protein molecules, $r_w = m_w/m_p$ is the water-to-protein mass ratio, and a_0 and a_1 are coefficients of linear regression. The d^3 values are proportional to the cube of the center-to-center distance: $d^3 = d_{cc}^3 \times 3^{3/2}/8$, and their dependence on r_w is shown in Figure 7b. To discuss the meaning of the experimentally observed linear regression coefficients, we will use the same approach as described above, i.e., assume that the interprotein distance can be described by the distance between the centers of imaginary spheres drawn inside the distorted molecules. In the case when both water and the protein are present, the volume fraction of spheres can be expressed as

$$\varphi_{\text{sph}} = \frac{V_{\text{sph}}}{V_p + V_w} = \frac{d_{cc}^3}{6}\pi\left(\frac{1}{\frac{m_p}{\rho_p} + \frac{m_w}{\rho_w}}\right) \tag{7}$$

Rearranging this equation, one obtains

$$\frac{d_{cc}^3\pi}{6} = \varphi_{\text{sph}} m_p\left(\frac{1}{\rho_p} + \frac{r_w}{\rho_w}\right) \tag{8}$$

or

$$d_{cc}^3 = \frac{6\varphi_{\text{sph}} V_p}{\pi}(1 + \bar{\rho}_p\, r_w) \tag{9}$$

where $\bar{\rho}_p$ is the density of protein relative to the density of water. In terms of d, this can be expressed as follows:

$$d^3 = \frac{3.897 \times \varphi_{\text{sph}} V_p}{\pi}(1 + \bar{\rho}_p\, r_w) \tag{10}$$

https://dx.doi.org/10.1021/acs.molpharmaceut.0c00351
Mol. Pharmaceutics 2020, 17, 3246−3258

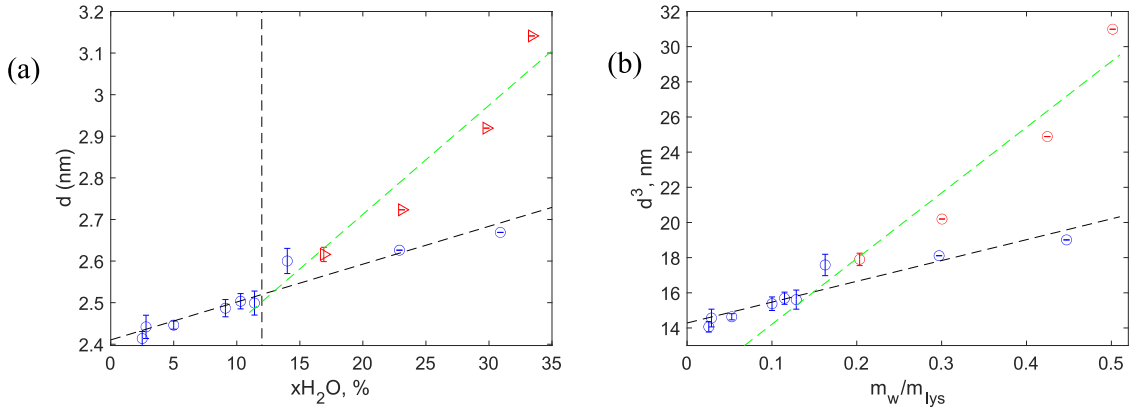


Figure 7. (a) Evolution of the correlation distance as a function of the water content for samples measured in glass capillaries (blue symbols) and aluminum pans (red symbols). (b) Three-dimensional swelling of lysozyme in the solid state. The lines are linear fits to the data.

Using the dependence presented in Figure 7b (black line), one obtains the values of the intercept with the y-axis and slopes of 14.3 and 11.8 nm^3, respectively. When using the trend with higher d^3 values (green line), the intercept and the slope are 10.5 and 37.4 nm^3, respectively. If one extrapolates the water content to zero, one can estimate from eq 10 the volume fraction of spheres φ_{sph}, which is 0.68 for the black dependence. This fraction lies between the values for random close packing and hcp/fcc, in agreement with the discussion above.

According to eqs 9 and 10, the ratio of the slope of the linear dependence to the intercept of the y-axis should be equal to the relative density of the protein. This ratio, calculated from the slope obtained from the black line in Figure 7b, gives a value of relative density of 0.83, while the steeper trend (green line) is 3.57. It is easy to see that in the first case the relative density at the higher water content is underestimated, while in the second case it is overestimated. This corresponds to too low and too high swelling (compared to the expected isotropic 3D swelling behavior), respectively. This fact shows complexity of the structure of the dehydrated lysozyme and once again suggests that the peaks observed here reflect only a local short-range order in the system. The total structure might be thought of as a superposition of two or more local structures that exhibit different swelling properties.

4.2.2. Nanostructure of Lysozyme at Water Contents above 35 wt % (Peak at q = 3.5 nm^{-1}). In the q-range of 3−4 nm^{-1}, the scattering patterns in the dilute aqueous solution showed a broad peak with a maximum around 3.5 nm^{-1} (Figure 1). As mentioned in Section 3.1, the peak originates from the form factor of the lysozyme, i.e., the data can be partly fitted to a prolate ellipsoid with a radius of 2.72 (±0.008) × 1.55 (±0.003) × 1.55 (±0.003) nm.

As the water content is increased, the peak position slightly shifts to lower q-values (from 3.4 to 3.3 nm^{-1}; see Figure 5b). This small change might be due to the variation in side-chain conformations, while the general shape of the native protein remains the same. Indeed, in more concentrated solutions, the side chains are in a more confined environment, where more folded conformations (or interpenetration) of side chains are expected, giving rise to effectively more compact structures.

4.3. Effect of Hydration on the Intramolecular Structure of Lysozyme Based on WAXS Data. *4.3.1. Evolution of the Peak Position in the Range of q = 5.5−7.5 nm^{-1}.* In the range of q = 5.5−7.5 nm^{-1}, three different peaks can be identified and appear to be present for all of the water contents (Figure 5d), namely, at q-values of 5.5−5.7, 6.0−6.2, and 6.5−7.5 nm^{-1}.

As mentioned earlier in the Gaussian deconvolution section, the peak in the q-range 6.0−6.2 nm^{-1} is present in all experimental data as well as in the Crysol data. It is clear that this peak is due to the protein structure. Hirai et al.[32] simulated scattering curves using the Debye equation and showed that this peak originated only from the signal of the α-helixes, i.e., it represented the correlation distance between α-helixes. In another study, Makowski et al.[33] used the atomic coordinates of a large set of protein domains to predict the WAXS patterns. They have also reported that the peak position at d = 1 nm (q = 6 nm^{-1}) characterizes the distance between α-helixes.

It is a well-established fact that upon dehydration the amount of α-helixes in lysozyme decreases.[8] Still, the peak at 6.1 nm^{-1} is clearly seen in the entire range of concentrations, including the dry state, but with reduced intensity (area under the peak). This supports the idea that the α-helix content only decreases but does not disappear upon dehydration.

The peak in the q-range 5.5−5.7 nm^{-1} is present in all of the concentration ranges. The peak position does not change with changing water content (Figure 5d); however, an increase in the peak intensity (normalized to the intensity of the peak at q = 6.1 nm^{-1}) is observed (Figure 4). The fact that this peak does not appear in the Crysol data could be explained possibly by the contribution from the hydration shell. As the water molecule at the border layer is denser than the bulk water (typically 1.05−1.25 times),[24] the electron density difference between the two densities of water could give rise to such a peak. The inclusion of the hydration shell in the analysis of the lysozyme crystal structure could show a much better agreement between theoretical calculated and experimental data. The signal from the hydration shell also showed an additional small peak at q around 5.7 nm^{-1}.[3434]

The third peak obtained by deconvolution in the q-range 6.5−7.5 nm^{-1} is also present in all of the concentration ranges. This peak, together with the low-intensity peak in the q-range 9−12 nm^{-1}, could be related to different secondary structures in the protein molecule. As mentioned earlier, the secondary structures change dramatically on changing the water contents as studied by Raman and FTIR techniques. However, it is not easy to observe the change in the case of SAXS/WAXS as the position and the intensity of the peak greatly depend on the baseline selection.

https://dx.doi.org/10.1021/acs.molpharmaceut.0c00351
Mol. Pharmaceutics 2020, 17, 3246−3258

4.3.2. Evolution of the Peak Position in the Range of q = 14−15 nm^{-1}. The peak at q = 14 nm^{-1} probes the structure in the length scale of 0.45 nm. This correlation distance could be related to the interatomic distance of the lysozyme between, for example, polypeptide groups and hydrophobic functional groups.[35]

The q-range = 14−15 nm^{-1} covers different structures in the length scale of 0.45−0.52 nm. This length scale characterizes the strand-to-strand correlation distance in β-sheets at d = 0.47 nm as well as the C-to-N distance between C═O and N−H groups of polypeptide main chains that are present in both α-helixes and β-sheets. The interactions between hydrophobic groups (CH_3, CH_2, and CH) should also result in the same separation. Since these features are always present in the protein independent of the hydration, the peak position is almost invariant with the water content (see Figure 2d).

4.3.3. Evolution of the Peak Position in the Range of q = 17−19 nm^{-1}. In this q-range, the peak around 18 nm^{-1} represents the hydrogen-bonded atoms in water and the protein.[36] In pure water, this peak is, however, at 20 nm^{-1}, corresponding to slightly shorter distances than in the protein solution.

On moving from dry powder to dilute solution, we observe a slight increase in the q-value of the peak position. This could be explained by the fact that H-bonds that are formed in bulk water on average are shorter than intraprotein H-bonds. The formation of the first hydration layer, which has a higher density than the bulk water,[37] would not affect the general trend.

4.4. Structural Changes of Lysozyme at Different Water Contents and Their Implications for Protein Stability. The evolution of the peak positions and intensities as a function of the hydration level is a valuable source of information that, in combination with other methods, reveals the structure of the protein−water system at different hydration levels. The data presented here show that below 35 wt % water, the lysozyme molecules are distorted to fulfill the space-filling requirements. The distortion of the overall structure in the dry conditions, together with other stresses during/upon dehydration, reversibly changes the native structure of the protein. The evolution of the protein−protein correlation peak position shows a 3D swelling upon hydration, which occurs even in the glassy state (below 10 wt % water) when the dynamics of the system are expected to be very slow.

In the concentration range of 35−55 wt % water, lysozyme undergoes spontaneous hydration-induced crystallization even if the dry material is fully amorphous. At higher water contents, the protein is in the native state (as seen from the comparison between the simulated scattering pattern from PDB and the experimental SAXS/WAXS data), but protein−protein interactions still have an effect on the scattering pattern evolution as a function of water content. The fact that lysozyme has different structures at different water contents has a direct effect on their thermal stability. Indeed, in excess water, unfolding of the protein is a cooperative process in which the native and unfolded states are in equilibrium.[38] Under the physical stresses caused by distortion and dehydration processes, the equilibrium will shift because the stresses affect the native and denatured structures differently.

To obtain a detailed picture of the hydration effects on protein stability, in situ time-resolved temperature SWAXS studies of proteins with different hydration levels would be needed. These studies are currently being performed by our group and will be published as a separate study.

SWAXS appears to be an effective method for monitoring the structural changes associated with the overall shape of the protein molecules. It would be interesting to use this method to investigate other proteins at different hydration states that exhibit poorer stability than lysozyme to establish a method to track the degradation pathway of solid-state proteins.

The monitoring of the changes in secondary protein structures using SWAXS is, however, a challenge that requires better procedures for background subtraction, matching data from different detector positions and scattering intensity deconvolution. It should be noticed that the secondary structures of lysozyme undergo dramatic changes as reported using Raman and FTIR techniques.[2,8] From the data presented here, one can conclude that the onset of the changes in secondary structures coincides with the onset of the overall structural changes as observed by SAXS/WAXS.

Studies of the swelling of protein at low water contents using SWAXS should be combined with studies of the thermodynamic behavior of the protein. In this way, the relationship between structural changes and the glass-transition behavior of the protein−water system can be established.

5. CONCLUSIONS

Simultaneous SAXS and WAXS have been used to characterize the structure of lysozyme in the solid state after freeze-drying and their structural changes upon rehydration to an aqueous solution. We have demonstrated that the protein molecules undergo deformation to pack closely together at low water contents. The native structure of the protein is replaced by a distorted structure, which allows the protein molecules to continuously fill the space that may be formed during freeze-drying. During rehydration, the material starts to absorb water and swelling of the system occurs. A change in the swelling behavior around 11 wt % water is observed, which is the expected onset of structural and dynamic transitions, i.e., a certain amount of water is needed to activate these transitions. When the uptake of water reaches 35 wt %, the protein molecules start to crystalize, which implies that the native structure of the protein is recovered. These changes in protein morphology may play a crucial role in determining a protein's thermal stability, shifting the denaturation conditions compared to the native state, or causing nonequilibrium effects in the denaturation processes. Although the WAXS data have not yet provided definitive evidence for the secondary structural changes due to both low resolution and the necessary averaging of different structural patterns, various secondary structural features of the protein could be identified and can serve as a reference to compare with other techniques such as vibrational spectroscopy or NMR.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.molpharmaceut.0c00351.

Amount of water uptake after incubating dried powder of lysozyme in different saturated salts (Figure S1), fitting results of the diluted solution of lysozyme (Figure S2), Gaussian deconvolution of SWAXS data at different water contents (Table S1), comparison of the SAXS

https://dx.doi.org/10.1021/acs.molpharmaceut.0c00351
Mol. Pharmaceutics 2020, 17, 3246−3258

pattern of lysozyme in crystalline and diluted regimes (Figure S3) (PDF)

AUTHOR INFORMATION

Corresponding Authors

[#]Tuan Phan-Xuan − *Biomedical Science and Biofilms Research Center for Biointerfaces, Malmö University, 214 32 Malmö, Sweden; Max IV Laboratory, Lund University, 224 84 Lund, Sweden;* orcid.org/0000-0003-1091-1468; Email: tuan.phan-xuan@maxiv.lu.se

Vitaly Kocherbitov − *Biomedical Science and Biofilms Research Center for Biointerfaces, Malmö University, 214 32 Malmö, Sweden;* orcid.org/0000-0002-9852-5440; Email: vitaly.kocherbitov@mau.se

Authors

Ekaterina Bogdanova − *Biomedical Science and Biofilms Research Center for Biointerfaces, Malmö University, 214 32 Malmö, Sweden*

Anna Millqvist Fureby − *RISE Research Institutes of Sweden, 114 86 Stockholm, Sweden*

Jonas Fransson − *SOBI Swedish Orphan Biovitrum, 112 76 Stockholm, Sweden*

Ann E. Terry − *Max IV Laboratory, Lund University, 224 84 Lund, Sweden*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.molpharmaceut.0c00351

Author Contributions

The manuscript was written through the contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

[#]Present Address: (T.P.-X.) Max IV Laboratory, Lund University, Fotongatan 2, 22484 Lund, Sweden.

ACKNOWLEDGMENTS

SOLEIL synchrotron (beamline SWING) and ALBA synchrotron (beamline NCD-SWEET) are thanked for allocation of the SAXS (proposal number 20180524) and SAXS/WAXS beamtimes (proposal numbers 2018092983 and 201902356), respectively. This research was funded by the Swedish Governmental Agency for Innovation Systems (VINNOVA) and was carried out within the competence center, NextBioForm. Dr. Christopher Söderberg (RISE) is thanked for useful discussions. Mr. Shuai Bai (Lund University) is thanked for freeze-drying lysozyme.

REFERENCES

(1) Walsh, G. Biopharmaceutical benchmarks 2018. *Nat. Biotechnol.* **2018**, *36*, 1136−1145.

(2) Costantino, H. R.; Langer, R.; Klibanov, A. M. Moisture-Induced Aggregation of Lyophilized Insulin. *Pharm. Res.* **1994**, *11*, 21−29.

(3) Flores-Fernández, G. M.; Sola, R. J.; Griebenow, K. The relation between moisture-induced aggregation and structural changes in lyophilized insulin. *J. Pharm. Pharmacol.* **2009**, *61*, 1555−1561.

(4) Langer, R.; Klibanov, A. M. In Moisture-induced aggregation of lyophilized proteins, *Abstracts of Papers of the American Chemical Society*, 1996; pp 25−Btec.

(5) Liu, W. R.; Langer, R.; Klibanov, A. M. Moisture-Induced Aggregation of Lyophilized Proteins in the Solid-State. *Biotechnol. Bioeng.* **1991**, *37*, 177−184.

(6) Costantino, H. R.; et al. Fourier-Transform Infrared Spectroscopic Investigation of Protein Stability in the Lyophilized Form. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1995**, *1253*, 69−74.

(7) Griebenow, K.; Klibanov, A. M. Lyophilization-Induced Reversible Changes in the Secondary Structure of Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 10969−10976.

(8) Kocherbitov, V.; et al. Hydration of lysozyme studied by Raman spectroscopy. *J. Phys. Chem. B* **2013**, *117*, 4981−4992.

(9) Kocherbitov, V.; Arnebrant, T. Hydration of lysozyme: the protein-protein interface and the enthalpy-entropy compensation. *Langmuir* **2010**, *26*, 3918−3922.

(10) Mertens, H. D. T.; Svergun, D. I. Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J. Struct. Biol.* **2010**, *172*, 128−141.

(11) Yang, S.; et al. Multidomain assembled states of Hck tyrosine kinase in solution. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 15757−15762.

(12) Stradner, A.; et al. Equilibrium cluster formation in concentrated protein solutions and colloids. *Nature* **2004**, *432*, 492−495.

(13) Shukla, A.; et al. Absence of equilibrium cluster phase in concentrated lysozyme solutions. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 5075−5080.

(14) Koshari, S. H. S.; et al. In Situ Characterization of the Microstructural Evolution of Biopharmaceutical Solid-State Formulations with Implications for Protein Stability. *Mol. Pharmaceutics* **2019**, *16*, 173−183.

(15) Castellanos, M. M.; McAuley, A.; Curtis, J. E. Investigating Structure and Dynamics of Proteins in Amorphous Phases Using Neutron Scattering. *Comput. Struct. Biotechnol. J.* **2017**, *15*, 117−130.

(16) Curtis, J. E.; et al. Small-angle neutron scattering study of protein crowding in liquid and solid phases: lysozyme in aqueous solution, frozen solution, and carbohydrate powders. *J. Phys. Chem. B* **2012**, *116*, 9653−9667.

(17) Wahl, V.; et al. The influence of residual water on the solid-state properties of freeze-dried fibrinogen. *Eur. J. Pharm. Biopharm.* **2015**, *91*, 1−8.

(18) Giuffrida, S.; et al. SAXS study on myoglobin embedded in amorphous saccharide matrices. *Eur. Phys. J. E* **2011**, *34*, No. 87.

(19) Kocherbitov, V.; Arnebrant, T.; Soderman, O. Lysozyme-water interactions studied by sorption calorimetry. *J. Phys. Chem. B* **2004**, *108*, 19036−19042.

(20) Das, S.; Banerjee, S.; Gupta, J. D. Experimental Evaluation of Preventive and Therapeutic Potentials of Lysozyme. *Chemotherapy* **1992**, *38*, 350−357.

(21) Massa, M. H. et al. Middle Ear and Eustachian Tube Mucosal Immunology. In *Mucosal Immunology*; Academic Press, 2015; pp 1923−1942.

(22) David, G.; Perez, J. Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J. Appl. Crystallogr.* **2009**, *42*, 892−900.

(23) Matthews, B. W.; Remington, S. J. The three dimensional structure of the lysozyme from bacteriophage T4. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 4178−4182.

(24) Svergun, D. I.; et al. Protein hydration in solution: Experimental observation by x-ray and neutron scattering. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2267−2272.

(25) Levitt, M.; Sharon, R. Accurate Simulation of Protein Dynamics in Solution. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7557−7561.

(26) van Hecke, M. Jamming of soft particles: geometry, mechanics, scaling and isostaticity. *J. Phys.: Condens. Matter* **2010**, *22*, No. 033101.

(27) Bergman, M. J.; et al. Experimental Evidence for a Cluster Glass Transition in Concentrated Lysozyme Solutions. *J. Phys. Chem. B* **2019**, *123*, 2432−2438.

(28) Hageman, M. J.; Possert, P. L.; Bauer, J. M. Prediction and Characterization of the Water Sorption Isotherm for Bovine Somatotropin. *J. Agric. Food Chem.* **1992**, *40*, 342−347.
(29) Angell, C. A. Formation of glasses from liquids and biopolymers. *Science* **1995**, *267*, 1924−35.
(30) Gregory, R. B. Protein Hydration and Glass Transition Behavior. In *Protein-Solvent Interactions*; Dekker: New York, 1995; pp 191−264.
(31) Kocherbitov, V.; Arnebrant, T. Hydration of thermally denatured lysozyme studied by sorption calorimetry and differential scanning calorimetry. *J. Phys. Chem. B* **2006**, *110*, 10144−10150.
(32) Hirai, M.; et al. Structural hierarchy of several proteins observed by wide-angle solution scattering. *J. Synchrotron Radiat.* **2002**, *9*, 202−205.
(33) Makowski, L.; et al. Characterization of protein fold by wide-angle X-ray solution scattering. *J. Mol. Biol.* **2008**, *383*, 731−44.
(34) Svergun, I. D. et al. *Monodisperse Systems, in Small Angle X-Ray and Neutron Scattering from Solutions of Biological Macromolecules*; Oxford University Press: New York, USA, 2020; pp 125−127.
(35) Corey, R. B.; Donohue, J. Interatomic Distances and Bond Angles in the Polypeptide Chain of Proteins. *J. Am. Chem. Soc.* **1950**, *72*, 2899−2900.
(36) Persson, F.; Soderhjelm, P.; Halle, B. The geometry of protein hydration. *J. Chem. Phys.* **2018**, *148*, No. 215101.
(37) Merzel, F.; Smith, J. C. Is the first hydration shell of lysozyme of higher density than bulk water? *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5378−5383.
(38) Privalov, P. L. Stability of proteins: small globular proteins. *Adv. Protein Chem.* **1979**, *33*, 167−241.

https://dx.doi.org/10.1021/acs.molpharmaceut.0c00351
Mol. Pharmaceutics 2020, 17, 3246−3258