ABSTRACT

Title of Thesis: NATURE OF MESOSCOPIC AGGREGATES IN SOLUTIONS OF LYSOZYME

Shakiba Nikfarjam, Master of Science, 2018

Thesis Directed By: Prof. Taylor J. Woehl

Department of Chemical and Biomolecular Engineering

Prof. Mikhail Anisimov

Department of Chemical and Biomolecular Engineering

Institute for Physical Science and Technology

An anomalous class of mesoscopic aggregates have previously been observed in solutions of lysozyme. These aggregates are thought to play an important role in nucleation of protein crystals and ordered protein aggregates, like amyloid fibers. Mesoscopic aggregates are currently thought to be in thermodynamic equilibrium with the protein solution, where transient oligomers of partially unfolded lysozyme monomers are thought to be the formation source of these aggregates. However, there is little experimental evidence to back up this proposed formation mechanism and thermodynamic behavior. Specifically, the effects of temperature on these aggregates and their thermodynamic reversibility have not been systematically tested. In this thesis, we investigate the equilibrium nature and the formation source of mesoscopic aggregates in solutions of model protein, lysozyme. We tested the effects of temperature on aggregate size and concentration and the aggregate reversibility after removal by systematic filtration. We used light and x-ray scattering and chromatography to experimentally characterize the aggregates during this study. Our findings indicate that mesoscopic aggregates are minimally sensitive to temperature changes and do not reform after removal by filtration. Together, these results indicate that mesoscopic aggregates are not in thermodynamic equilibrium with protein monomers or oligomers in solution. Overall, our experimental results contrast the current accepted formation mechanism of these mesoscopic aggregates and suggest they instead form due to contaminants present in solution or a sub-population of partially unfolded proteins.

NATURE OF MESOSCOPIC AGGREGATES IN SOLUTIONS OF LYSOZYME

By

Shakiba Nikfarjam

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2018

Advisory Committee:
Professor Taylor J Woehl, Chair
Professor Mikhail Anisimov
Srinivasa R Raghavan
Amy J Karlsson

© Copyright by

Shakiba Nikfarjam

2018

Dedication

This thesis is dedicated to my parents for all their support and love.

Acknowledgement

I would like to my advisors Dr. Mikhail Anisimov and Dr. Taylor Woehl first for their believing in me and second for their constant support, guidance and patience during this project. Teaching me how to believe in myself, my abilities and knowledge is the most precious thing you thought me.

I would like to thank Ms. Elena Jouravleva for her endless help in all stages of my research, sharing he knowledge of DLS with me and her support until the last day the experiments took.

I would also like to thank Professor Marc Taraban for his contribution and suggestions which enhanced the quality of this study.

I would also like to thank Professor Amy J. Karlsson for her guidance and generosity in sharing her lab instruments with us during this study.

I would also like to thank Sayanee Adhikari and Mahdi Ghorbani for participating in this research and helping with performing long and precise experiments.

Table of Contents

List of Figures

List of Abbreviations

ConA	Concanavalin A
DLS	dynamic light scattering
HEPES	(*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic)
PBS	Phosphate-buffered saline
PDDF	pair distance distribution function
PES	polyethersulfone
PI	Isoelectric point
SAXS	small angle x-ray scattering
SEC	size exclusion chromatography

Chapter 1: Introduction

1.1. Background

1.1.1. Importance of Protein Aggregation

Aggregation phenomena in protein formulations is generally undesirable because it alters the original properties of the solution. The issue of protein aggregation is important for the biomanufacturing of protein therapeutics, both in development and commercialization stages (Wang & Roberts, 2018). Both reversible self-association and irreversible aggregation matter in development of therapeutics in two ways. First, highly concentrated proteins have a propensity to form aggregates, which negatively impacts therapeutic effectiveness and quality of therapeutics, negatively impacting patients' health (Dekel et al., 2017; Wang & Roberts, 2018). Large numbers of aggregates or formation of interlinked networks increases the viscosity of therapeutics, which is undesirable for manufacturing and administration to patients (Wang & Roberts, 2018; Pathak et al., 2013). The immune response of a protein-based drug has been shown to be exacerbated due to the existence of protein aggregates (Joubert et al., 2016; Filipe et al., 2012; Freitag et al., 2015; Uchino et al.,2017). Biomanufacturing is not the only circumstance where aggregate formation causes undesired consequences. Amyloid diseases, such as Alzheimer's Disease and amyloidosis, and sickle cell anemia are also linked to protein aggregation *in vivo* (Alam et al., 2017).

1.1.2. Colloidal Aggregation

Colloids experience a combination of attractive and repulsive forces; the balance of these forces determines the stability of the colloidal system (Kovalchuk & Starov, 2012). Aggregation of colloidal particles occurs when attractive colloidal forces overcome repulsive forces. Derjaguin-Landau-Verwey-Overbeek (DLVO) theory provides the physical basis for understanding colloidal interactions, which includes Van der Waals attraction and electrostatic repulsion. Additional forces may influence aggregation behavior of proteins, such as hydrophobic forces and more specific biomolecular interactions.

In proteins, aggregation falls into two different categories: native state aggregation and non-native state (Kotch, 2015). Self-association under certain experimental conditions, such as pH, ionic strength, temperature, and protein concentration, occurs between native state proteins ('folded clusters', Figure 1) (Kotch, 2015; Kramer et al., 2012). Surface charge, hydrophobicity and potential to form α-helixes and β-sheets are among the physical aspects of proteins that affect native state self-association (Lauer et al., 2012; Roberts, 2014). Due to their amphiphilic nature, some proteins such as ConA can form reversible tetramers near neutral pH which disassociate by a change in pH from ~7.5 to ~5 (Silvers & Myers, 2013).

Unlike native aggregation, non-native aggregation is accompanied by partial unfolding of the protein monomers. The Lumry-Eyring model describes protein aggregation as a two-step process: (partial) unfolding of the monomers is a reversible equilibrium reaction that occurs first, while aggregation of unfolded monomers

proceeds far from equilibrium *via* kinetic processes (Li & Roberts, 2010). Partial unfolding reveals the hydrophobic core of the protein monomers, which induces attractive hydrophobic forces leading to aggregation (Barnett et al., 2016). Hydrogen bonds can further stabilize protein aggregates (Kotch, 2015).

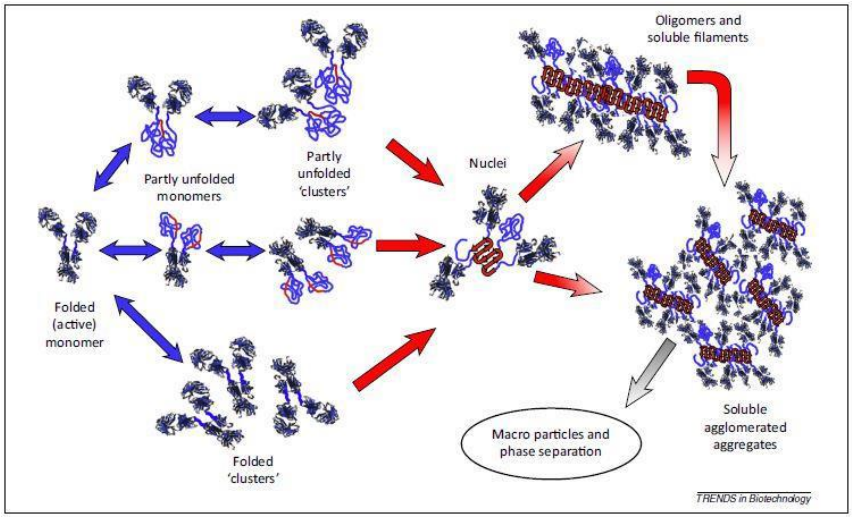

Figure.1.1. Possible pathways for non-native aggregation for multidomain proteins. 'Hot spot' sequences which tend to form strong irreversible inter-protein contacts stabilizing the aggregates, are denoted by red arrows. Double arrows denote effectively reversible steps. Single arrows denote irreversible steps. Figure and caption reproduced with permission from (Roberts, 2014).

1.1.3. Kinetics of Protein Aggregation

Proteins are complex compared to polymers and traditional colloids, but basic physical and chemical principles can be applied to understand their aggregation behavior. Thermodynamically, folding of a polypeptide chain is an organizing event requiring removal of the water molecules from residues forming the core of the folded state. During folding, a decrease in enthalpy (ΔH) compensates the energy cost of the water removal, which creates an entropic penalty (ΔS) (Equation 1.1). In folded state,

the free energy (ΔG) is 5-20 kcal/mol more negative compared to the unfolded state (Chi et al., 2003; Wang & Roberts, 2010; Dill et al., 2008).

$$\Delta G = \Delta H - T\Delta S \quad (1.1)$$

Based on the observed aggregation kinetics, a wide variety of mechanisms have been suggested for justifying the formation of aggregates, each of which starts with partial unfolding of the native state. To name these models: Monomer conversion, oligomeric intermediates, nucleation in protein folding and domain swapping (Wang and Roberts, 2010).

Figure 1.1 illustrates the relationship between self-association, aggregation, folding, and unfolding. Depending of the conditions that promote aggregation, the pathway can begin with either folded or unfolded proteins. Experimental and theoretical research has indicated that in many cases of the aggregate formation in proteins, reversible oligomers consisting of ten or less subunits of native or partially unfolded monomers can act as a precursor (*i.e.* a nucleus) to aggregation (Wang & Roberts, 2010). The population of these oligomers is typically too small to be quantified (Young & Roberts, 2007). Monomers located inside the oligomers interact through anisotropic colloidal forces. The dissociation equilibrium constant for these oligomers is typically higher than those for larger multimers (Wei & Roberts,.2010). This process leads to terminal formation of soluble self-associated protein oligomers, but does not lead to formation of larger irreversible protein aggregates. The concentration of each species in folded oligomers is thermodynamically controlled. The population of these

species is only dependent of concentration of the total protein at constant temperature, pressure and solvent conditions (Young & Roberts, 2009).

1.1.4. Lysozyme as A Model System

Lysozyme is a small single domain enzyme that is among one of the most well-characterized proteins due to its wide availability and ability to act as a model protein for many systems. Lysozyme shows interactions similar to those of monoclonal antibodies, which makes it a good model protein with simpler structure and shape (Dharmaraj et al., 2016). Lysozyme has a net positive surface charge; however, due to non-uniformities of the surface charge and hydrophobicity, lysozyme can have a combination of attractive short-range interactions and long-range repulsion (Dharmaraj et al., 2016).

The primary structure of lysozyme consists of a single polypeptide chain containing 129 amino acids and molecular weight of 14.4 kDa (Abeyrathne at al.,2014). Lysozyme has an elliptical shape with dimensions of 3.0 x 3.3 x 4.5 nm (Whitford, 2016). The isoelectric point of the lysozyme is at pH = 11.35, which is considered a high value relative to other proteins (Vorontsova et al., 2015). This high pKa gives lysozyme a positive surface charge at neutral pH.

1.2. Summary of The Thesis

The outline of the thesis is as follows. Chapter 2 describes the theory behind light and x-ray scattering techniques utilized to characterize protein aggregates in this thesis.

We present a theoretical treatment of dynamic light scattering (DLS), beginning with a description of self-diffusion in colloidal systems. We then present a theoretical treatment of small angle x-ray scattering (SAXS), including its use for measuring particle size and interparticle interactions. In Chapter 3 we investigate the cause of formation, effects of temperature, and reversibility of mesoscopic aggregates in concentrated lysozyme solutions. We systematically varied solution temperature and the size of filter used on the protein solution and utilize DLS and SAXS to characterize the size and relative concentration of mesoscopic aggregates. After removal of aggregates by filtration, we stimulate reformation of the aggregates by increasing the temperature, addition of denaturants, and sonication. Lastly, we utilize size exclusion chromatography (SEC) to test a current theory that dimers are the source of the protein aggregate formation. Finally, in Chapter 4, we summarize the thesis and present several avenues for future research on this topic. These include investigating the source of mesoscopic aggregates and their composition.

In this thesis, we will answer the following research questions:

- Are mesoscopic protein aggregates in thermodynamic equilibrium with monomers or oligomers in solution?
- What is the effect of temperature on the size and concentration of mesoscopic lysozyme aggregates?
- Will mesoscopic aggregates reform if removed from solution?
- Do oligomers form in concentrated solutions of lysozyme?

Chapter 2: Fundamentals of Light and X-ray Scattering

The two predominant experimental techniques used to investigate protein aggregation in this thesis are dynamic light scattering (DLS) and small angle x-ray scattering (SAXS). Here we review the fundamental principles of DLS and SAXS. We describe instrument setup, fundamental underlying scattering physics, thermodynamics of these related scattering methods, and the information provided by each technique.

2.1. Dynamic Light Scattering (DLS)

2.1.1. Basics of Light Scattering

In a quiescent fluid solution, the thermal motion of solvent molecules leads to their random collision with dissolved solute molecules or colloids, which causes random displacement. This process, known as Brownian motion, can be embodied by a balance between thermal and viscous forces on the solute molecules *via* the Stokes-Einstein equation for the self-diffusion coefficient (D). Assuming spherical non-interacting particles the hydrodynamic radius, R_h, is related to the self-diffusion coefficient by Stokes-Einstein equation (Russel et l., 1991):

$$R_h = \frac{\mathrm{k}_B T}{6\pi\eta D} \tag{2.1}$$

Where, k_B is Boltzmann's constant, T is the temperature, and η is the kinematic viscosity of the solvent.

As evidenced in equation (2.1), a characteristic of Brownian motion is that small molecules move faster than large ones due to decreased viscous hindrance.

A small fraction of light is scattered from a molecular or colloidal solution when it is irradiated with visible monochromatic light of high spatial and temporal coherence. The intensity of this scattered light fluctuates in a time-dependent manner due to the continuously changing distances between particles, which leads either to constructive or destructive interference. DLS takes advantage of the particle size dependence of the scattered light fluctuations to gain information pertaining to size of particles. As discussed above, large particles diffuse slowly due to large viscous hinderance and thus cause less rapid fluctuations in the scattered light intensity, compared to small fast moving particles (Stetefeld et al., 2016) (Figure 2.1).

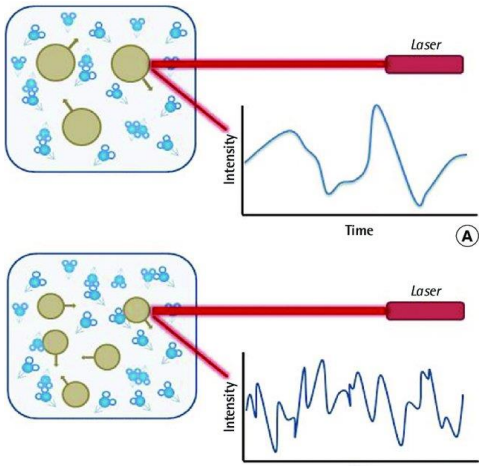

Figure 2.1. Hypothetical fluctuation of scattering intensity of larger particles and smaller particles. Figure and capture reproduced with permission from (Li and Baron)

The autocorrelation function of the scattered-light intensity can be used to obtain the particle diffusion coefficient and size by correlating the intensity fluctuations of scattered light with respect to time. A digital autocorrelator is used to extract the

correlations in intensity fluctuations related to the diffusion behavior of colloidal particles. The autocorrelation function, $G_2(\tau)$, is defined using a comparison of the intensity $I(t)$ of the scattered light at a time t with that at some later time $(t+\tau)$: (Berne &Pecora, 1976; Kensal et al., 2006),

$$G_2(\tau) = \langle I(t)I(t+\tau)\rangle, \tag{2.2}$$

where τ is the lag time between two time points. The normalized autocorrelation function can be written as:

$$g_2(\tau) = \frac{\langle I(t)I(t+\tau)\rangle}{\langle I(t)\rangle^2}. \tag{2.3}$$

The braces in equations 2.2 and 2.3 denote averaging of properties over time. For the simplest case of spherical monodisperse particles in a fluid, the autocorrelation function has a single characteristic decay time τ_c.

$$g_2(\tau) = 1 + e^{-2\Gamma\tau} \tag{2.4}$$

The decay rate, Γ, is proportional to the diffusion coefficient of the particles and is the inverse of the decay time, τ_c, where

$$\Gamma = Dq^2. \tag{2.5}$$

Here $q = \left(\frac{4\pi n}{\lambda}\right)\left(\frac{\theta}{2}\right)$ is the wave vector, n is refractive index, λ is the laser wavelength, and θ is the scattering angle (Harding, 1999). For spherical particles or solutes, the corresponding particle size can be determined by converting the diffusion coefficient to hydrodynamic radius using equation 2.1 and the known solution viscosity.

In our experiments, the heterodyne technique (which employs dual frequencies) is used, which is suitable for small intensities. Equation (2.5) relating the decay rate Γ to diffusion coefficient D is applicable for heterodyne experiments. Homodyne measurement is another technique that extracts information of the frequency of an oscillating signal by comparing that signal with a standard oscillation, and is suitable for large intensities (e.g., near critical point of a fluid or for colloid systems). For homodyne spectrum, the relation between Γ and D is:

$$\Gamma = 2Dq^2 \tag{2.6}$$

The autocorrelation function $g_2(\tau)$ of the scattered light as a function of the decay time τ_c can be represented by (Burchard 1983):

$$g_2(\tau) = b[1 + \varepsilon \exp\left(-\frac{\tau}{\tau_c}\right)] \tag{2.7}$$

Where $\frac{\tau}{\tau_c}$ is equal to $2Dq^2\tau$.

The baseline (accidental) correlation level b is proportional to the total scattered light intensity (I). Figure 2.2 shows a schematic of the correlation function. The coefficient ε depends on the amount of stray light and the apertures in the system and is an instrument dependent parameter.

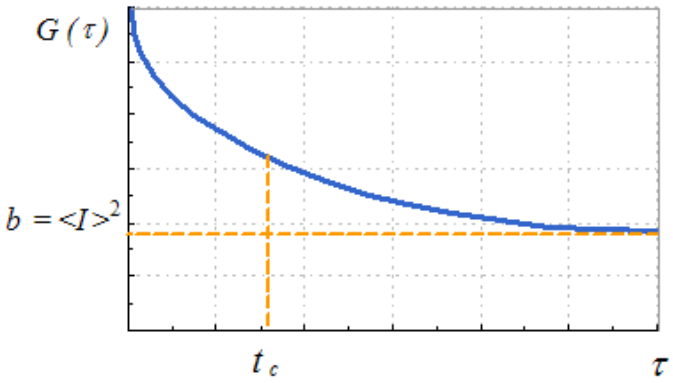

Figure 2.2. Schematic autocorrelation function of the intensity of scattered light. Figure and caption reproduced with permission from (Yudin et al, 1997).

2.1.2. Light Scattering Instrument

A DLS instrument is composed of a laser light source, a sample cell, a detector placed at a fixed or variable angle, a photomultiplier amplifying the signal, and a correlator (Figure 2.3). A multi-angle dynamic and static light scattering instrument is used for the experiments in this thesis (Photocore Complex). The laser source is a TEC stabilized diode laser 650 nm with 30 mW of power. The scattering angle was fixed at $\theta = 30^{\circ}$ and the duration of the measurement is 1 hour for each sample (120 repetitions of 30 second frames).

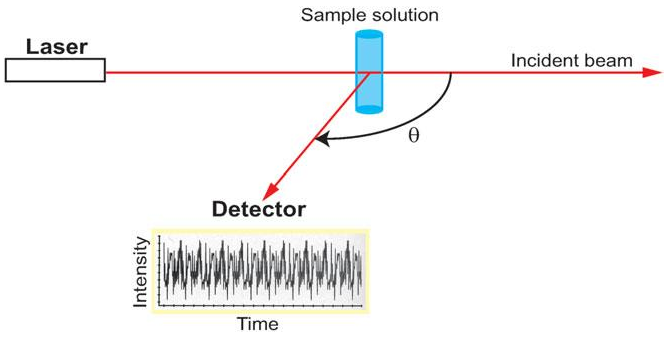

Figure 2.3. Schematic representation of the optical setup of a DLS system.

2.2. Small Angle X-ray Scattering (SAXS)

2.2.1. Background

SAXS is an analytical method for characterizing the nanostructure of complex liquids and solids containing nanosized particles or domains. In SAXS, a monochromatic, collimated x-ray beam is transmitted through a thin sample (1 – 100 μm). The x-ray beam diffracts and scatters off particles or domains to produce a diffraction pattern. The x-ray beam in 1 mm in diameter, thus SAXS yields an ensemble averaged structure of the sample being measured. SAXS is highly accurate, non-destructive, and requires a relatively thin sample (Schnablegger, 2013). The most significant application of SAXS is linked with the studies of the angular dependence of scattering. Scattering t the very small-angle range is strongly influenced by the size of domains in the samples, while scattering to larger angles is dependent on the shape of particles (Kensal et al., 2006). The Thomson formula for scattering from a single electron is the fundamental relation for SAXS and can be expressed as:

$$I_e = \frac{r_0^2}{r^2}\frac{1+cos\,2\theta}{2}\,I_0, \tag{2.8}$$

where I_0 is the intensity of upcoming x-rays, 2θ is the angle of observation, I_e is the intensity of scattered x-rays, r denotes to the distance between the electron and the detector, and r_0 is the electron radius $\frac{e^2}{mc^2} = 2.817 * 10^{-15}\,m$ (Weiss, 2016).

According to Debye (Debye 1915), scattering from an ensemble of atoms with known atomic positions yields the scattering intensity as a function of wave vector:

$$I(q) = \sum_i \sum_j \frac{f_i f_j sinqR_{ij}}{qR_{ij}} \tag{2.9}$$

The above summation goes over all atoms constituting the scattering object. Where R_{ij} is the distance between the atomic pairs *i* and *j*, *q* is wave vector. f_i is called scattering length of the atom and *I* describes the amount of scattered intensity by that atom (Chaudhuri et al., 2017).

In SAXS, distances are measured based on the wavelength of the x-ray source. The SAXS scattering pattern is usually shown as a function of *q*, the scattering vector:

$$q = \frac{4\pi}{\lambda}\sin(\phi) \tag{2.10}$$

Information about the size, shape, and structure of the sample by calculation of radius of gyration can be extracted from Guinier, Fourier, and Porod region in the SAXS profile. By fitting a line to the natural log of the intensity as a function of q^2(scattering vector) in the Guinier region, the radius of gyration (R_g) can be calculated. R_g is affected by the existence of aggregated particles and polydispersity, and errors in subtraction of background or buffer (Boldon et al., 2015).

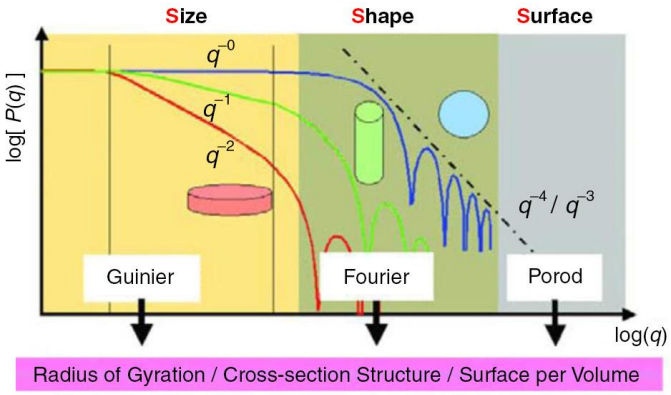

Figure 2.4. Regions of SAXS profile. Figure and caption reproduced from. Figure and caption reproduced with permission from Boldon et al., 2015).

Figure 2.4 depicts the three regions of the SAXS profile. The Guinier region contains information giving the radius of gyration, while information regarding the surface per volume of particles can be extracted from the Porod region. Using a Fourier-transformation in the Fourier region of the SAXS profile, shape of the particles can be approximated.

2.2.1.1. The Form Factor

The square of the sum of all wave amplitudes from each particle/domain un the sample at the detector leads to an interference or scattering pattern. When the pattern oscillations are indicative of the shape and size of the particles, it is known as "form factor, $P(q)$". The SAXS pattern corresponds to the form factor of a single particle when the particles are monodisperse, and sample is dilute enough that particles do not interact, *i.e.* $I(q) = P(q)$.

2.2.1.2. The Structure Factor

In densely packed particle systems, the scattering due to interparticle and intraparticle structure have similar orders of magnitude and the SAXS pattern contains information from both types of scattering. This SAXS pattern therefore has contributions from the form factor and the structure factor, which encompasses scattering due to interparticle structures, viz. $I(q) = p(q)S(q)$, where $S(q)$ is the structure factor. The structure factor encompasses x-ray scattering due to both short and long-ranged order in materials. The characteristic equation describing a SAXS pattern is as follows:

$$\Delta I(q) = K \cdot P(q) \cdot S(q) \tag{2.11}$$

Where *K* is a constant dependent on contrast (difference between electron densities of the discontinuous and continuous phases), illuminated sample volume and concentration among others (Schnablegger., 2013).

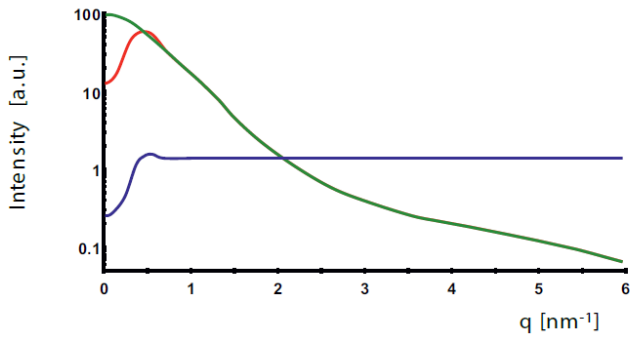

Figure 2.5. The red line is the SAXS profile of a concentrated arbitrary particle dispersion, which is the product of the form factor (green) and structure factor (blue) lines. Figure and caption reproduced from (Schnablegger., 2013).

2.2.2. Data Interpretation

2.2.2.1.Resolution

The detectable size range in SAXS experiments is limited by the sample to detector distance, wavelength of x-ray used, and detector noise. These values ($q_{\min}$ and $q_{\max}$) determine particle characteristics between D_{min} and D_{max}. $D_{min} = \frac{\pi}{q_{max}}$ and $D_{max} = \frac{\pi}{q_{min}}$ (Schnablegger., 2013).

2.2.2.2.Radius of Gyration

Sample properties like molecular mass, radius of gyration, hydrated particle volume and polydispersity can be measured with SAXS. All parameters are extracted from model fits of the processed SAXS data. The initial and the most basic model fit is a Guinier analysis, which is used to extract scattering intensity at $q = 0, i.e. I(0)$, and the radius of gyration, R_g (Guinier, 1939).

$$I(q) = I(0) \exp\left(-\frac{1}{3} R_g^2 q^2\right) \tag{2.12}$$

This is done by plotting the SAXS data as a Guinier plot, which linearizes the SAXS data at low q by plotting $ln\{I(q)\}$ versus q^2. $I(0)$ and R_g can be simply extracted from the y-axis intercept and slope, respectively (Figure 2.6).

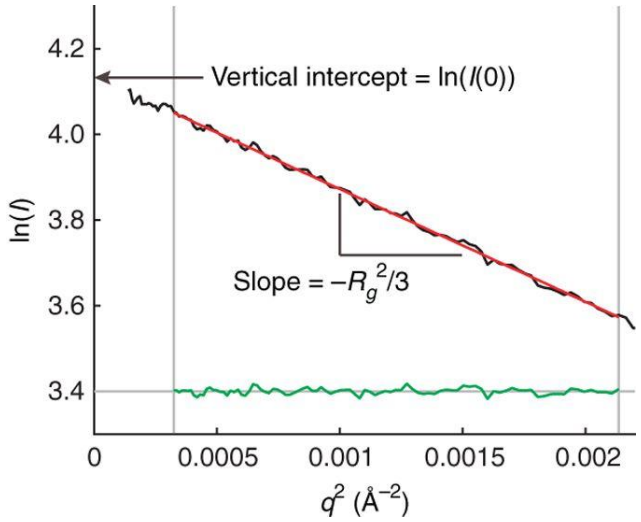

Figure 2.6. A line (red) is fitted to the low-q region of a Guinier plot (black), such that the maximum q to be included in the fit is $q < 1.3/ R_g$. The linearity of the fitted region is determined by the flatness of the residuals (green). R_g is derived from the slope, and $I(0)$ is derived from the vertical intercept. Figure and caption reproduced with permission from (Skou et al., 2014).

The lower limit of Guinier region is restricted by instrument noise, while the upper limit is determined by $q_{max} < 1.3/R_g$ (Haydyn et al., 2010).

2.2.2.3.Pair Distance Distribution Function

To gain more information about structure of particles in the system the pair distance distribution function (PDDF) $p(r)$ can be determined. The PDDF is the distribution of intraparticle distances. $d_{\max}$ is the maximum diameter in the particle. This function includes information about the shape of molecules. The $I(q)$ and $P(r)$ are related by (Debye, 1915; Liu at al., 2012; Koch et al., 2003, Weiss, 2016):

$$I(q) = \int_0^{d_{max}} \frac{P(r)\sin(qr)}{qr} dr \tag{2.13}$$

Inverse transformation leads to:

$$p(r) = 4\pi \int_0^{\infty} qrI(q)\sin(qr)\, dq \tag{2.14}$$

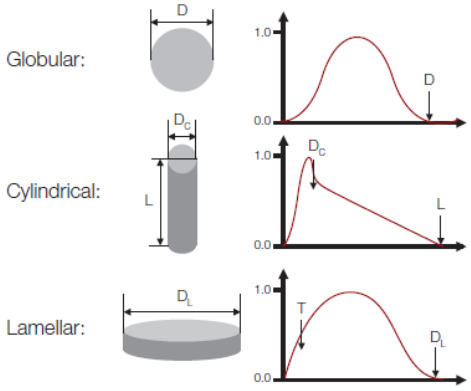

Figure 2.7. pair distance distribution of particles, globular particles can be identified from the bell-shaped PDDF, figure 2.7. reproduced with permission from (Schnablegger et al., 2013).

2.2.2.4.Interparticle Interactions

The structure factor $S(q)$ includes information about colloidal scale interparticle forces. The Coulomb interaction and the Van-der-Waals interactions are some examples of forces existing between particles in a solution. In a dilute dispersion of particles, we can neglect the effect of multiparticle scattering is negligible and the structure factor mainly arises from two-particle scattering interference, according to Zimm's analysis of dilute macromolecule dispersions (Zimm, 1948). The mathematical form of structure factor reflecting the dependence on both solute concentration and interaction strength, is: $S(q,c) = 1 - \xi P(q)$ where, ξ depends on the particle concentration *c*, the osmotic second virial coefficient A_2, and molecular weight of scatterers. A positive value of A_2 is indicative of repulsive interaction, while negative values indicate interparticle attraction. We can rewrite $S(q,c)$ by substituting $\xi = 2A_2cM_w$. The osmotic second virial coefficient A_2, and second virial coefficient B_2 (obtained from statistical mechanics) are directly dependent. By taking an integral measure of the pair interaction potential $u(r)$, between particles in solution, B_2 can be calculated.

$$B_2 = A_2 v_c \left(\frac{M_w^2}{N_A}\right) = -2\pi \int \left(e^{-\frac{u(r)}{KT}} - 1\right) r^2 dr \qquad (2.17)$$

Where N_A is Avogadro's number, r is the center-to-center particle separation, and v_c is the ratio of the solvent and protein molar volumes (Zamora & Zukoski, 1996; Davis, 1996).

At low q or the Guinier region, the attractive and repulsive colloidal interactions can result in deviations from the particle form factor, $P(q)$. A modified Zimm analysis written to extract A_2 from SAXS:

$$\frac{1}{M_w.c} = \frac{K}{I_0(0,q)} - 2A_2 \tag{2.18}$$

Where K is an optical constant (Zimm, 1948; Zamora & Zukoski, 1996; Davis, 1996; Saunders & Korgel, 2004).

2.2.3. Instrument

SAXS instrument is comprised of an x-ray source, a collimation system, sample holder, beam stop and a detection system. The radiations from the sample particles goes to the detector at different angles. The collimation system is responsible for making the beam narrow and defining the zero-angle position. The beam stop protects the detector from intensive incident beam (Schnablegger., 2013).

Here to run our SAXS experiments, we used a Xenocs Xeuss small angle x-ray scattering system. The system is equipped with a 5 Meter slide system with CuKα (wavelength = 1.5406 Å) sealed 30W tube high brightness micro-focus source, Parallel beam optics and scatterless slits with automatic alignment, 300K Dectris Pliatus detector for small angle scattering with a minimum q = 0.0045$Å^{-1}$, 100K Dectris Pilatus detector for wide angle scattering (up to about 45°2θ), and Linham stage controlling temperature from about -100°C to 250°C.

Chapter 3: Characterization of Mesoscopic Aggregates in Lysozyme Solutions

Study of protein stability in solutions is important for better understanding the pathogenesis of diseases caused by abnormal protein folding and aggregation. Lysozyme is known to form mesoscopic aggregates (30-100 nm radius) in concentrated solutions (>30 mg/mL), however the origin and thermodynamic status of these aggregates remain unclear. In this work we have investigated the effects of concentration, filtration, and temperature on the sizes and relative amount of mesoscale aggregates in solutions of lysozyme. We have used dynamic light scattering (DLS), small–angle X-ray scattering (SAXS), and size exclusion chromatography (SEC).

Mesoscopic protein aggregates were commonly thought to be in equilibrium with protein monomers in solution, resulting from a reversible self-assembly of the monomers. We instead show that systematic filtration through 20 nm pore size filters completely removes the aggregates from solution. The aggregates do not reemerge. Without filtering, the relative number of monomers decreases with increasing solution temperature, indicating formation of more aggregates. SEC has been used to search for the presence of lysozyme dimers, which have been previously hypothesized to be related to the formation of mesoscopic aggregates, however SEC has not detected dimers in solutions of filtered or unfiltered lysozyme. Taken together, our results strongly suggest that the mesoscopic aggregates in lysozyme are not caused by reversible self-assembly of lysozyme monomers and are not an intrinsic property of

lysozyme monomers in their native state. We hypothesize that the lysozyme aggregation is likely due to some impurities in lysozyme introduced during purification or lyophilization and/or to traces of misfolded lysozyme.

3.1. Overview and Introduction

Aggregation of peptides and proteins has gained much attention due its implication in the pathogenesis in a variety of neurodegenerative disorders (Siddiqi et al., 2017). Proteins undergo various intermediates during folding before reaching the biologically active or native conformation. In some cases, these intermediates lead to misfolded conformations, possibly resulting from genetic mutations (Ellis et al., 2006). Since the energy barrier separating native and non-native states of proteins is small, proteins can misfold under stress conditions (Dobson, 2003, Jahn et al., 2005). Stresses such as heat, pH, shear, surfactants, and denaturants are capable of driving proteins to form partially unfolded intermediates leading to aggregation (Siddiqi et al., 2017). However, proteins in their native conformation will form condensates in which crystals, fibrils, or dense liquids are found (Vekilov et al., 2017). Some proteins, in particular lysozyme, form mesoscopic (~100 nm) aggregates in a macroscopically homogenous phase where most of the protein molecules should be folded (Pan et al., 2010, Li et al., 2012, Gliko et al., 2007, Vorontsova et al., 2015, Safari et al., 2017, Byington et al., 2018). The aggregates are hypothesized to be amorphous protein-rich clusters, however, there is no direct observation of their structure. Based on a hypothesized

structure of aggregates of partially unfolded lysozyme in the presence of urea, Pan et al. estimated the concentration of lysozyme inside the clusters to be at least ~50 wt % (Pan et al., 2010). In case of lysozyme as a model protein, the diameter of these mesoscopic clusters ranges from 100 to 200 nm, where each contains $10^4 - 10^5$ protein molecules. Compared to total soluble protein, the mesoscopic clusters constitute about $10^{-5} - 10^{-3}$ number fraction, a very small fraction of total soluble protein in the solution (Safari et al., 2017). It was also believed that mesoscopic clusters of lysozyme are reversible and in constant exchange with protein monomers in the host solution (Safari et al., 2017; Vorontsova et al., 2015).

Yamazaki et al. have claimed that these aggregates are amorphous and potentially serve as nucleation sites for protein crystallization in solution when the solution condition is far away from liquid-liquid coexistence region and well above the crystallization line (Pan et al., 2010, Yamazaki et al., 2016). Moreover, others have claimed that the volume fraction of clusters is related to thermodynamics of solutions, while the size of clusters is explained by protein complexes (Pan et al., 2010). The complexes are thought to be weakly-bond dimers, trimers, or tetramers of lysozyme (Li at al., 2015, Byington et al., 2016 and 2018, Vorontsova et al., 2015).

However, there is little evidence that lysozyme mesoscopic clusters are formed by reversible self-assembly of lysozyme monomers. In this work we have tested the hypothesis of the reversible nature of mesoscopic aggregation in solutions of lysozyme. We have shown that systematic filtration through 20 nm filter pores completely removes the aggregates from solution. The aggregates do not reemerge. Our results

suggest that the mesoscopic aggregates in lysozyme are not caused by reversible self-assembly of lysozyme monomers in their native (folded) state, and instead they are formed due to irreversible aggregation of contaminated or partially unfolded proteins.

3.2. Materials and Methods

3.2.1. Materials

Lyophilized powder of lysozyme ($\boldsymbol{M_w = 14\,kDa}$) was obtained from three sources: ThermoFisher Scientific (20,000 units/mg solid), Sigma Aldrich (~100000 units/mg solid), and MP Biomedical (crystallized powder, ≥20,000 units/mg solid). HEPES buffer (Sodium salt of 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, >99%) was supplied as a solid powder from VWR. Syringe filters with PES (polyethersolfune) membrane and pore sizes 0.1 and 0.22 μm were purchased from VWR while Whatman Anotop 25 filters with 0.02 μm pore size were purchased from Sigma Aldrich. Lysozyme powder was dissolved by gently swirling in 20 mM HEPES buffer with pH adjusted to 7.8 by addition of 12 M HCl.

3.2.2. Sample Preparation

Protein concentration in the samples ranged from 5 to 60 mg/mL. After the sample preparation the solutions appeared cloudy due to the presence of insoluble macroscopic aggregates. Then these samples were filtered through 0.22 μm PES filter following a previously established procedure [2, 4, 5]. For further investigation of the nature of lysozyme aggregates, we subsequently filtered the solutions through 0.1 **μ**m

and 0.02 μm filters. To minimize the applied shear stress during the fine filtration through 0.02 μm pore size filters, an RHV low flow pump was used. Protein concentration was measured with a ThermoFisher Nanodrop 2000 UV spectrometer, using absorption at 280 nm and the extinction coefficient of lysozyme (36000 $M^{-1}cm^{-1}$). A Neslab RTE-100 Refrigerated Bath Circulator was used to adjust the temperature for DLS measurements. Each sample was held at a target temperature for 1 hour for thermal equilibration. For SAXS and SEC measurement, the buffer was degassed by high speed stirring under vacuum prior to addition of protein. SAXS measurements were taken of each solution using a flow cell. The SEC elution buffer (PBS) was filtered and degassed. Samples were dialyzed in degassed filtered HEPES buffer for 48 hours prior to SAXS measurement.

3.3. Results

Detection of lysozyme monomers and aggregates. The objective of this study was to understand the origin and nature of mesoscopic aggregates present in highly concentrated (> 30 mg/mL) solutions of hen egg white lysozyme. DLS was used to determine the relaxation modes associated with the lysozyme monomers and mesoscopic aggregates. We measured the decay rates of these modes in solutions with lysozyme concentrations ranging from 4 to 60 mg/mL under 5 different scattering angles from 30 to 135°. In concentrated solution (c > 30 mg/mL) we commonly

observed three relaxation modes with the corresponding rates, $\boldsymbol{\Gamma_m}$, $\boldsymbol{\Gamma_{c1}}$, and $\boldsymbol{\Gamma_{c2}}$, hence the DLS correlation function can be approximated as

$$g_2(\tau) - 1 = (A_m \exp(-\Gamma_m) + A_{c1} \exp(-\Gamma_{c1}) + A_{c2} \exp(-\Gamma_{c2}))^2 \qquad (3.1)$$

First, we tested whether the observed relaxation modes are diffusive. If the relaxation mode is associated with Brownian diffusion, the decay rate should be proportional to square of the light scattering wave number as shown in equation (2.6). Indeed, as shown in Fig. 1, all observed decay rates vary linearly with q^2.

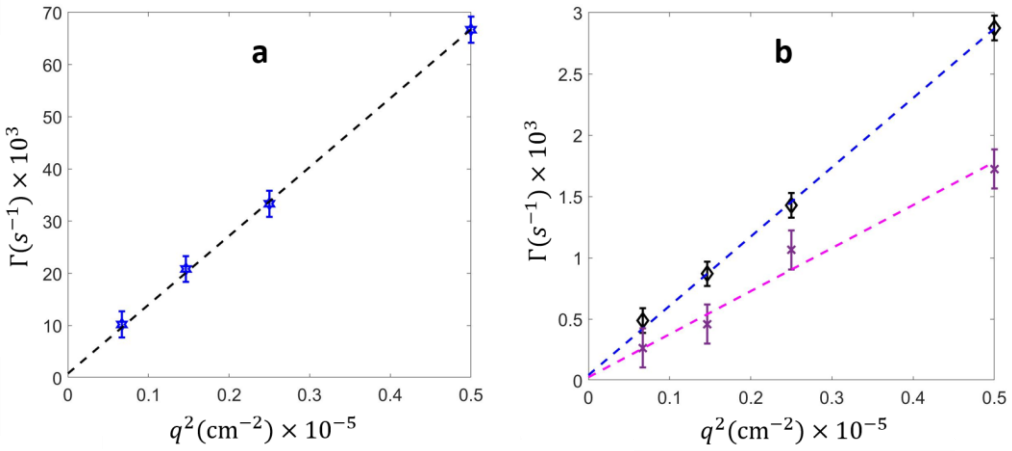

Figure. 3.1. Decay rates observed in a 60 mg/mL solution of lysozyme at 25 ◦C as a function of the square of the wave vector. (a) The diffusion rate of lysozyme monomers. (b) The diffusion rate of lysozyme aggregates: black diamonds are for smaller aggregates and purple crosses are for the larger aggregates.

The size of lysozyme monomers in dilute solution. The measurements of the monomer diffusion coefficient in dilute solutions show that the hydrodynamic radius of lysozyme (R_(h,m)=1.85±0.05 nm) calculated from the Stokes-Einstein relation

(2.1), within experimental errors, does not depend on concentration in the range between 4 to 8 mg/mL. This indicates that in this concentration range that we measure the actual size of undisturbed lysozyme molecules that do not mutually interact with neighboring molecules. However, at higher concentrations the apparent hydrodynamic radius gradually decreases with the increase of concentration (Figure 3.2), a well-known phenomenon in physical chemistry of polymer solutions (Zheng et al., 2018). If the solution is not dilute, the protein monomers interact, and the size calculated from the Stokes-Einstein relation (2.1) cannot be interpreted as hydrodynamic radius of individual protein molecules. The concentration that separates the dilute and semi-dilute regimes, is indicated in Figure 2.1 by a dashed vertical line.

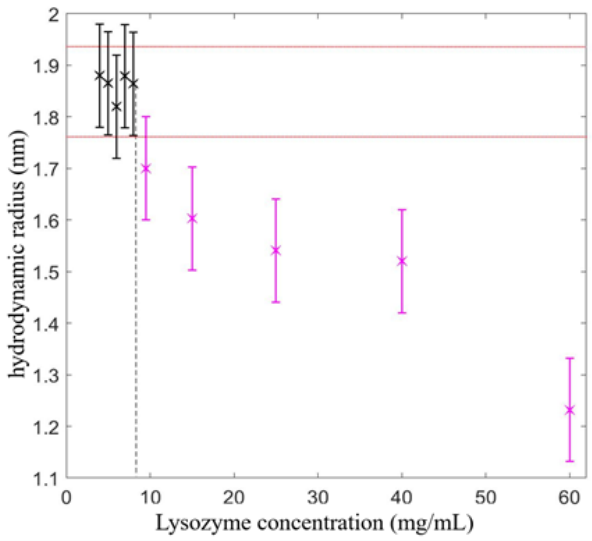

Figure 3.2. Effective hydrodynamic radius of native lysozyme monomers as a function of concentration at 25 ◦C. The concentration ~9 mg/mL corresponds to the transition from dilute to semi-dilute regimes. Red dotted lines indicate the corridor for experimentally obtained values for lysozyme hydrodynamic radius in different solution conditions (Parmar et al., 2009).

We have also measured the radius of gyration, R_g, of lysozyme monomers with SAXS and compared it with the hydrodynamic radius of monomers obtained with DLS.

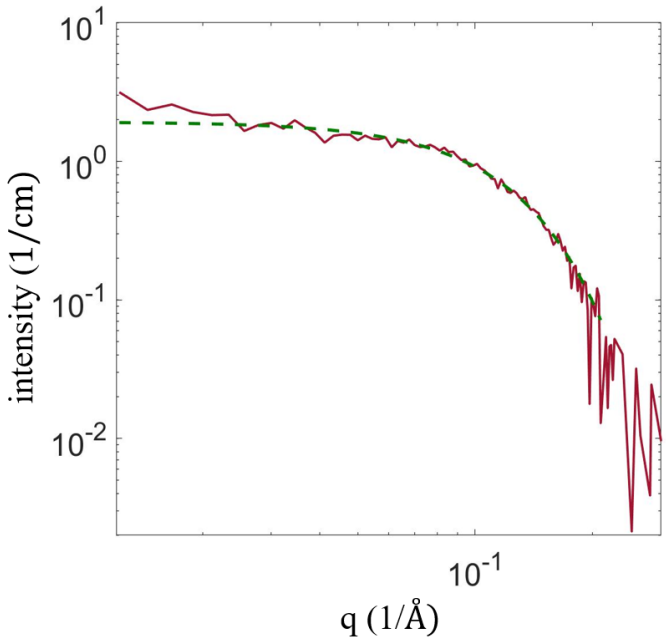

Figure 3.3. SAXS scattering intensity as a function of the wave number in the solution of 6 mg/mL lysozyme at 25 ℃. The green dashed line is a Guinier fit according to eq 2.12.

The radius of gyration determined from SAXS was $R_g = 1.49$ nm (± 0.01). For a globular protein molecule, a ratio $R_g/R_h \sim 0.775$ is expected [14]. This ratio obtained from our measurements is 0.8 ± 0.03, which is in agreement with the expected value.

Second virial coefficient in solutions of lysozyme. We have investigated two-body interparticle interactions using a modified Zimm analysis and SAXS as described in section 2.2.2.4 to calculate osmotic second virial coefficient of lysozyme monomers under our experimental conditions (T = 25 ℃, 20 mM HEPES buffer and pH 7.8). We

prepared a series of dilute lysozyme solutions with concentrations between 1 – 6 mg/mL, each filtered with 0.22 µm syringe filters. SAXS measurements were taken of each solution using a flow cell. After background subtraction to remove the contribution from the buffer, we performed a Guinier analysis to determine the scattering at $q = 0$, I_0. Based on equation 2.17, plotting $\frac{1}{M_w c}$ against $\frac{1}{I(0,c)}$ gives the second osmotic virial coefficient as the y-intercept of this plot (Figure 3.4). We measured a positive value of $A_2 = 110 * 10^{-5}$ $\left[\frac{cm^3 mol}{g^2}\right]$ for the second osmotic virial coefficient, which indicates presence of repulsive interactions between lysozyme monomers.

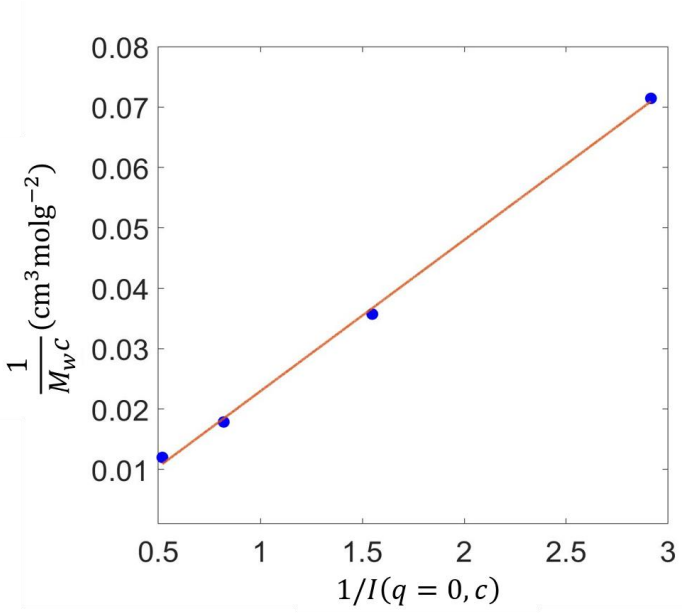

Figure. 3.4. Zimm plot obtained from a Guinier analysis of the SAXS data for dilute solutions of lysozyme at 25 ℃. The osmotic second virial coefficient, A_2, was extracted from the y-intercept of the red line *via* the Zimm equation.

Size of mesoscopic aggregates. Mesoscopic aggregates were observed in solutions with lysozyme concentration ranging from 30 to 60 mg/mL. Figure 3.5 (a and b) shows the autocorrelation functions at 25 °C and 35 °C for a 60 mg/mL lysozyme solution filtered through a 0.22 μm syringe filter. The correlation function is not a single-exponential. A distribution analysis reveals three diffusion modes: a fast mode corresponding to diffusion of protein monomers and two slower modes belonging to mesoscopic aggregates with radii of ~30 nm and ~100 nm. Even though the mesoscopic aggregates strongly contribute to the intensity of scarred light, their number in the scattering volume is much smaller than the number of monomers. Therefore, the Stokes-Einstein relation can still be used to obtain the hydrodynamic radius of aggregates because concentrated solutions of monomers can be considered as dilute with respect to mesoscopic aggregates. Increasing temperature led to a slight decrease in number of monomers relative to amount of aggregates. Figures 3.5c and 3.5d shows the correlation function in the same solution after subsequent filtration through a 0.1 μm pore size filter. The correlation function shows a significant change in the shape at longer delay times (Figure 3.5c) due to removal of mesoscopic aggregates with the sizes larger than 0.1 μm (Figure. 3.5d). Correspondingly, the amplitude associated with the relative amount of the larger mesoscopic aggregates decreased. Incidentally, in both cases, increasing the temperature causes formation of more aggregates at expense of lysozyme monomers.

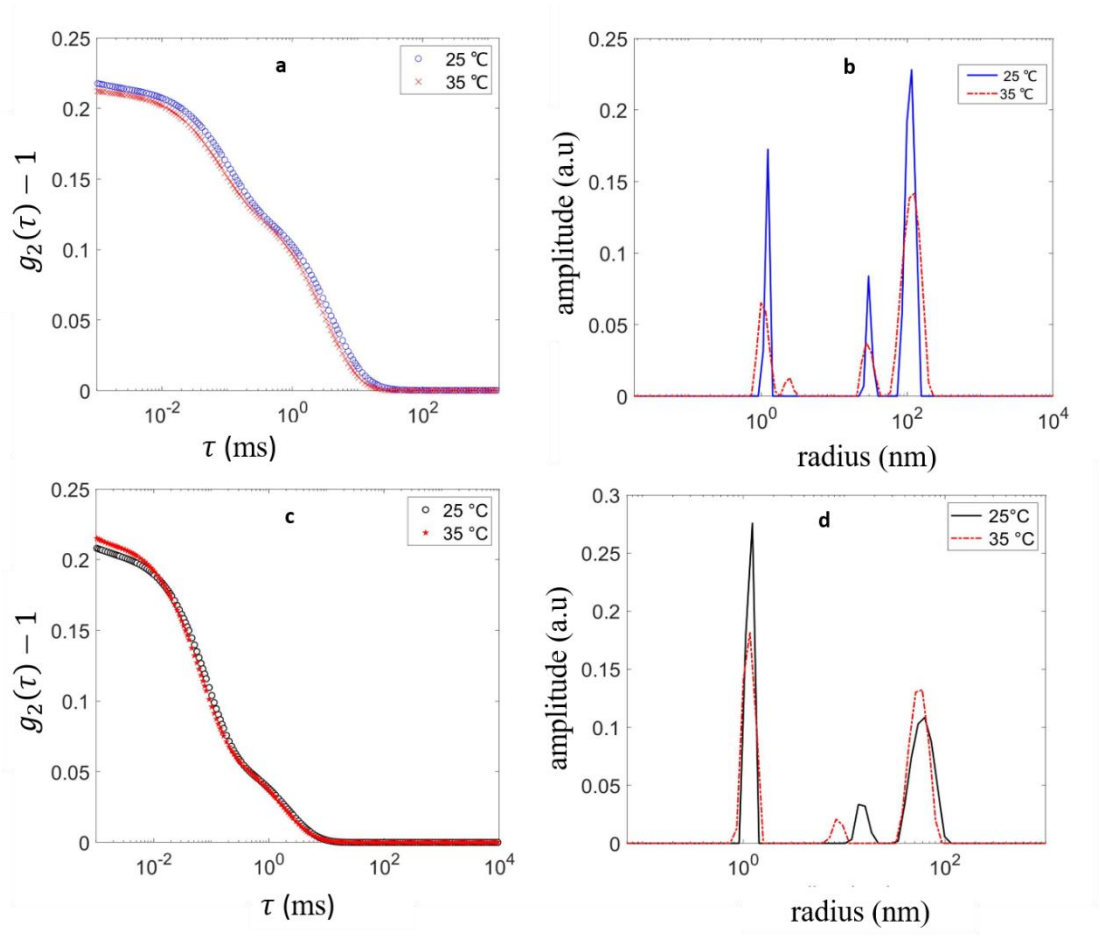

Figure 3.5. DLS correlation function (left) and particle size distribution (right) obtained for a 60 mg/mL lysozyme solution filtered through a 0.22 µm pore size filter (a,b) and 0.1 µm pore size filter (c,d).

We have tested lysozyme solutions purchased from three different vendors and found that the amount of the mesoscopic aggregates is strongly source dependent. The distribution of monomers and aggregates present in these solutions is shown in Figure 3.6. Although the sizes of the aggregates for all three samples are similar (30 -100 nm), the relative population of monomers and aggregates is different. The lysozyme

purchased from Thermofisher showed the smallest amount of aggregates relative to monomers, compared to the other vendors.

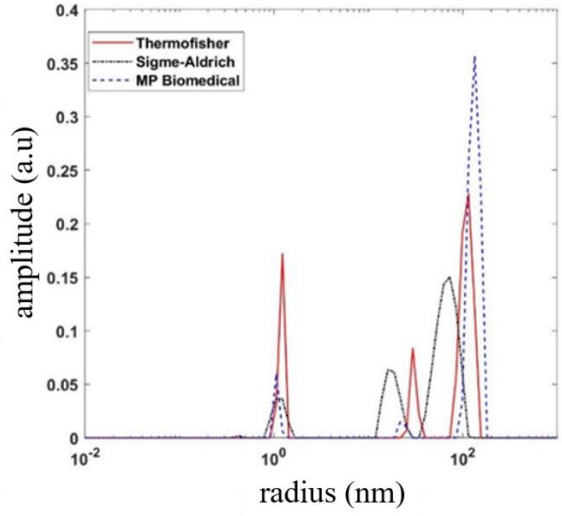

Figure. 3.6. Intensity distribution of lysozyme monomers and aggregates in 60 mg/mL solutions, with lysozyme acquired from different sources.

Lysozyme solutions prepared from different batches supplied by the same vendor exhibited different distributions of aggregates, as demonstrated in Figure 3.7.

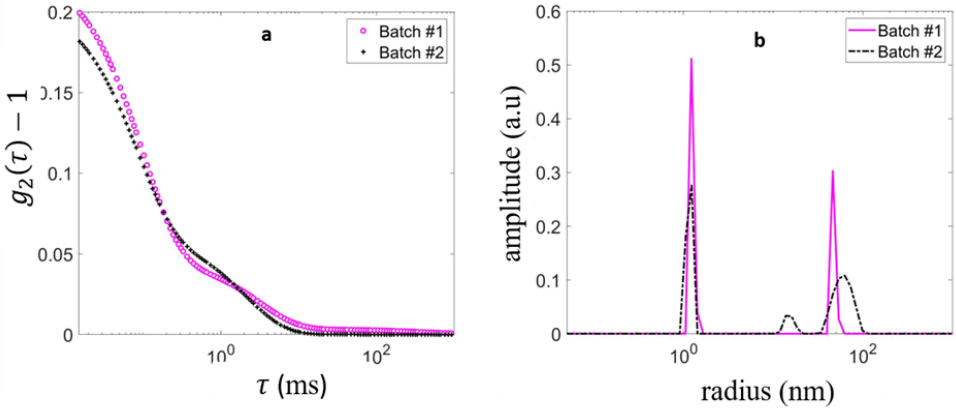

Figure. 3.7. DLS correlation function (g_2) obtained for a 60 mg/mL lysozyme solution for two different batches of the same brand (Thermofisher).

Figure 3.8 shows the DLS correlation functions for a 60 mg/mL solution filtered through 0.02, 0.1, and 0.22 µm syringe filters. Analysis of the correlation function for a fine-filtered sample (0.02 µm) yielded a particle size distribution with a single peak corresponding to monomers. Elevating the temperature had no effect on the correlation function or fitted particle size distribution.

A significant feature of these mesoscopic aggregates is that they are completely removed from the solution by subsequent fine filtration through 0.02 µm pore size filters with use of an RHV low flow pump. Even more importantly, the aggregates did not reemerge even after several days of incubation in the refrigerator. This fact strongly indicates the irreversible nature of the mesoscopic aggregates in solutions of lysozyme.

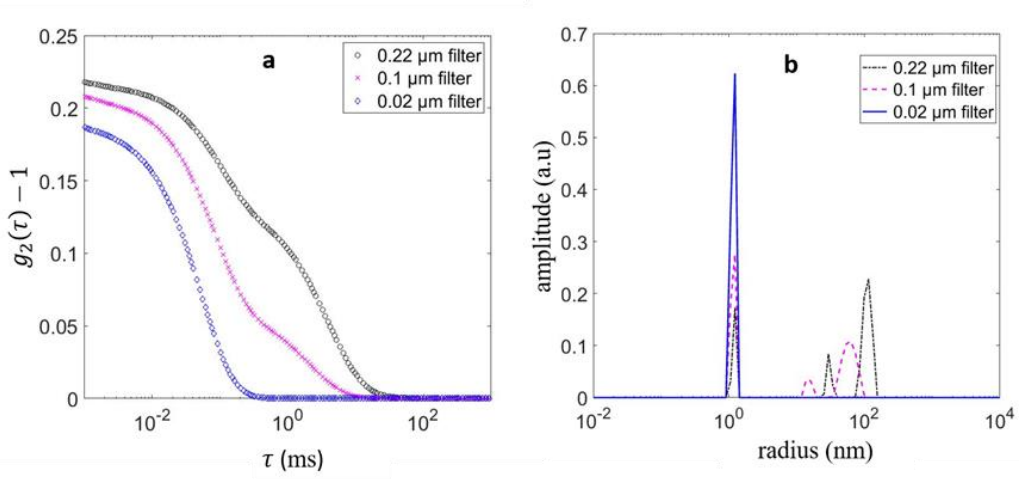

Figure 3.8. Effect of filtration on mesoscopic aggregates in a solution of 60 mg/mL lysozyme filtered through a 0.02, 0.1, and 0.22 µm pore size filter

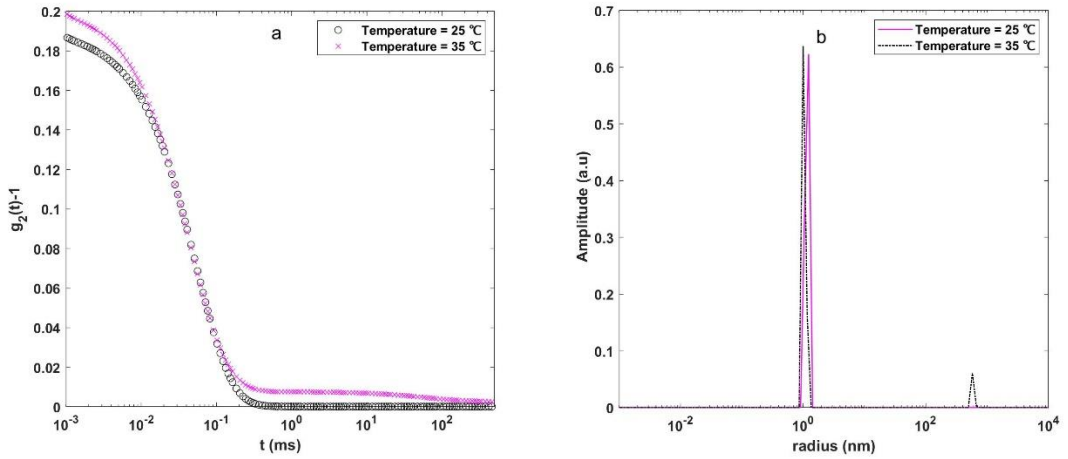

Figure. 3.9. (a) Intensity correlation function (g_2) of scattered light from a solution of 60 mg/mL lysozyme filtered through a 0.02 µm pore size filter, in a 20 mM HEPES Sodium buffer pH 7.8 in two temperatures of 25 ℃ and 35℃. Only one shoulder (single exponential) appeared in intensity correlation function indicative of only one diffusive mode. (b) Particle size distribution shows only monomers of lysozyme.

We fit the DLS data from the 20 nm filtered sample using a single exponential decay to further demonstrate that only monomers exist in solution. Figure 3.10 shows that monomodal fitting (equation 2.7) model detects only a single species of lysozyme monomers in solutions filtered through 0.02 µm pore size filters.

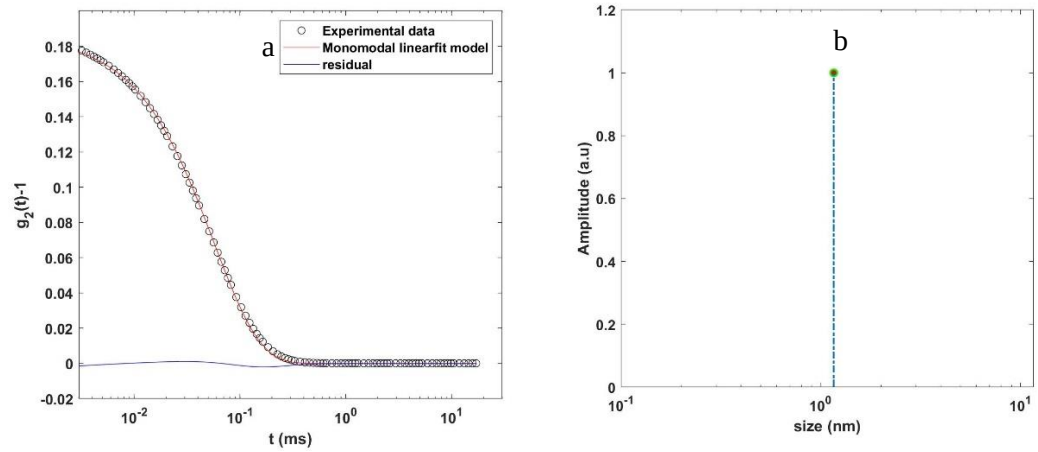

Figure 3.10. (a) binomial and (c) monomodal exponential fit for the sample completely filtered through 0.02 um pore size filter. As it shown in (a) and (b) both binomial and monomodal fits shows the lack of presence of aggregates.

Effect of Stress on filtered lysozyme solutions. We applied various external perturbations, including heating, sonication, and denaturing agents, to test whether the mesoscopic aggregates would reemerge in solutions after being removed by fine filtration through 0.02 μm filter. Sonication did not induce aggregation. Figure 3.11 shows the effect of heating on a filtered sample initially containing only monomers. We observed that heating the solution below the denaturation temperature of lysozyme (~69℃) did not cause formation of the mesoscopic aggregates. However, heating the sample to 80 ℃ and keeping the sample at this temperature for 2 hours, followed by overnight incubation at room temperature led to formation of small aggregates (~10 nm).

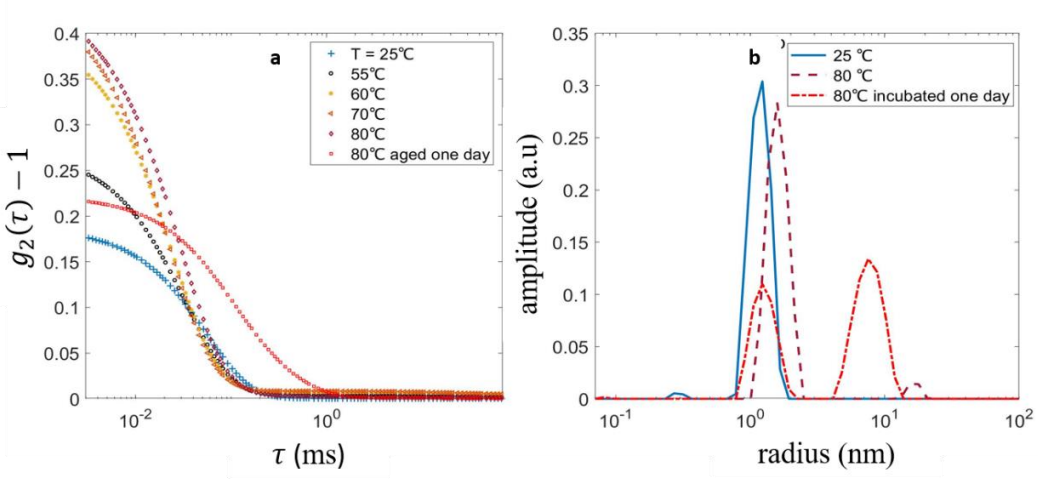

Figure. 3.11. Effects of heating on DLS correlation functions (a) and on the particle size distribution (b) in 60 mg/mL solution of lysozyme.

Figures 3.12 shows the effect of addition of isopropanol as a protein denaturant. The addition of 5 wt% isopropanol caused formation of large protein aggregates with sizes of >200 nm, which were distinctly different from mesoscopic aggregates observed in unfiltered lysozyme solutions.

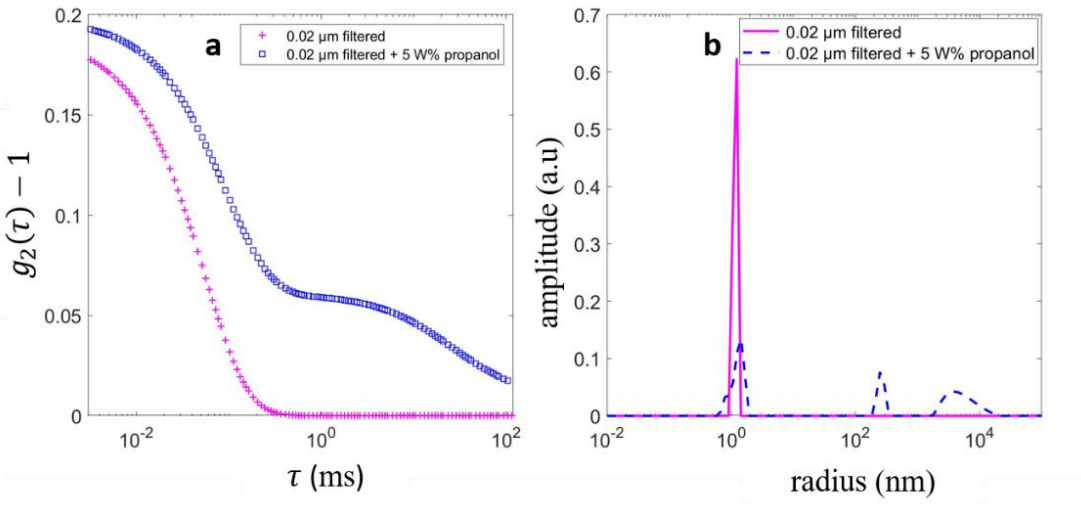

Figure 3.12. Formation of aggregates due to addition of a denaturant to a fine-filtered 60 mg/mL lysozyme solution; (a) is the DLS correlation function and (b) is the particle size distribution.

Size Exclusion Chromatography. To test for the presence of lysozyme oligomers (dimers, trimers, etc.) in the filtered lysozyme solutions, we performed SEC on a gel filtration column with a molecular weight cut-off of 650 kDa (Bio-Rad NGC Chromatography System and Enrich SEC 650 size exclusion high resolution column). Figure 3.13 shows calibration data for this SEC column. We used the Bio-Rad size exclusion standards which was a mixture of thyroglobulin (670 kDa), bovine γ-

globulin (158 kDa), chicken ovalbumin (44 kDa), equine myoglobin (17 kDa), and vitamin B12 (1.35 kDa). The calibration curve was constructed using the elusion volumes of the calibrant material.

Figure 3.14 shows SEC data for lysozyme samples filtered through 0.1 and 0.02 µm pore size filters. Each plot shows a single peak corresponding to the lysozyme monomer; no oligomers or larger aggregates were observed. The mesoscopic aggregates are not observed in the SEC data because they are too large to be observed by the SEC column with a 650 kDa cutoff. The lysozyme eluted later than expected based on the calibration curve, which is possibly due to interaction of the protein with the resin in the column.

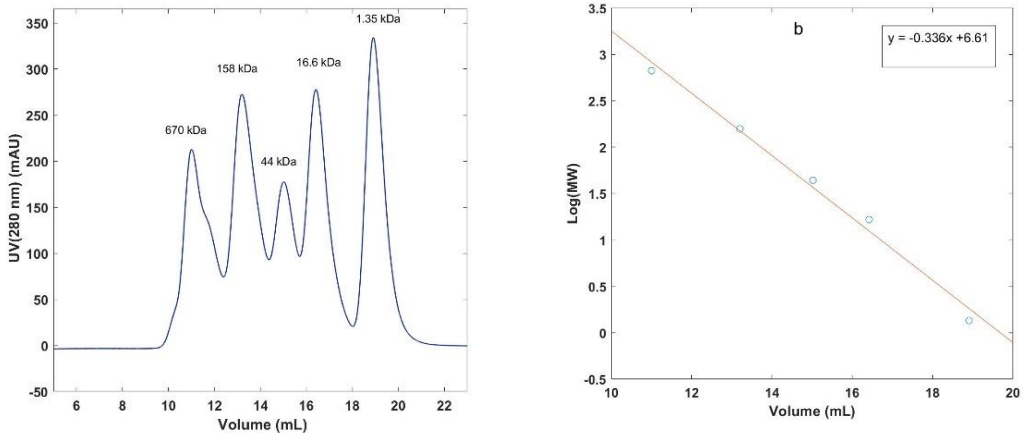

Figure 3.13. (a) Gel filtration SEC standards and (b) MW calibration curve.

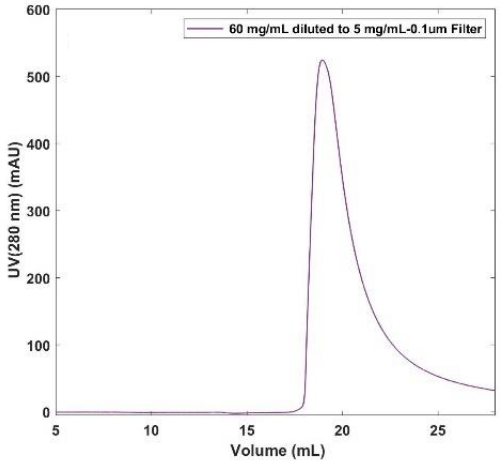

Figure 3.14. SEC chromatogram (a) Sample filtered by a 0.1 µm filter (b) Sample filtered by a 0.02 µm filter. Both results are indicative of monomers as the only species present in filtered samples, in the limit of detection of the column.

3.4. Discussion

Taken together, our results on mesoscopic aggregates in solutions of lysozyme suggest that the aggregation phenomenon in such solutions away from the conditions of denaturation is not an intrinsic property of biologically active (folded) lysozyme monomers. We believe that the main reason for the formation of mesoscopic aggregates in solutions that were not subjected to fine-filtration is due to the presence of some contaminants, most likely including some unfolded or partially misfolded monomers introduced during purification or lyophilization of protein. Other types of proteins or contaminants could also be present during extraction of lysozyme from raw hen egg whites (Roberts et al., 2018).

The mesoscopic aggregates are kinetically stable over long times. Increasing temperature resulted in decrease in the relative concentration of protein monomers in solution (Figure 3.5b), possibly because of condensation of monomers onto the mesoscopic aggregates due to increased hydrophobic forces, partial monomer unfolding, or formation of molten globule states. Mesoscopic aggregates could serve as 'nucleation sites' that reduce the free energy barrier for partial unfolding of lysozyme monomers. However, as it is indicated in Figure 3.8 after filtration of solutions with the smallest pore size (0.02 μm) syringe filters, the aggregates separate out from the solutions. Aggregates of the same size and relative amount did not reform after filtration after subjecting lysozyme to different perturbations, including heating and addition of denaturant.

Our studies also demonstrated that lysozyme solutions made of proteins coming from different sources (distinct companies or batches form same companies) showed a variety of aggregates with different sizes and relative concentrations. This further supports our hypothesis that mesoscopic aggregates exist due to some form of "impurity" in solution, which apparently varied in concentration in different batches and original sources of protein

Our results elucidate previous interpretations regarding the formation and thermodynamics of mesoscopic lysozyme aggregates. Some prior investigators have suggested that the mesoscopic aggregates are reversible, with monomers and oligomers being in constant exchange with the host solution. Partial unfolding of monomers and

formation of transient dimers have been proposed as mechanisms for the formation of the aggregates (Pan et al., 2010, Li at al., 2015, Byington et al., 2016 and 2018, Vorontsova et al., 2015). However, our studies demonstrate that the aggregates are formed irreversibly, as they did not reform after being removed by filtration, while the overall concentration of protein was not changed in the detection limits of absorption measurement at 280 nm. The presence of various protein species with variable molecular weights of 6, 18, and 29 kDa in commercial lysozyme has been confirmed by previous studies (Thomas et al., 1996, Parmar et al., 2007).

Our results indicate that mesoscopic clusters in solutions of lysozyme are not formed by reversible self-assembly of lysozyme monomers in their native state. It means that the suspension of mesoscopic aggregates in solution of lysozyme monomers is not a thermodynamically stable "mesophase" (Pan et al., 2010), but rather a long-lived kinetically stabilized colloid (Shchukin., et al., 2001). A positive, relatively large value of the second osmotic virial coefficient (Fig. 4) and the fact that lysozyme monomers carry a positive surface charge at pH < 11 (Parmar et al., 2009) support the assertion that lysozyme molecules in their native, undisturbed state do not tend to aggregate.

Lysozyme dimers have been claimed to exist in commercial lysozyme, however, Parmar et al. argued (Parmar et al., 2007) that the dimers could not be associated with formation of mesoscopic aggregates because the dimers remained in solution even after dissociation of these aggregates by addition of sodium dodecyl sulfate. Moreover, after being separated from the solution the aggregates dissociated by sodium dodecyl sulfate

only to monomers (Parmar et al., 2007). To test the hypothesis that lysozyme dimers are responsible for the mesoscopic aggregates (Vekilov e al., 2018), we performed SEC on a gel filtration column with a molecular weight cut-off of 650 kDa (Bio-Rad NGC Chromatography System and Enrich SEC 650 size exclusion high resolution column). Figure 3.15 shows SEC data for lysozyme samples filtered through a 0.02 µm pore size filter at two concentrations. Each plot shows a single peak corresponding to the lysozyme monomer; no oligomers or larger aggregates were observed. The observation of lysozyme dimers reported in previous works might be explained by the difference in production and purification used for commercial lysozyme.

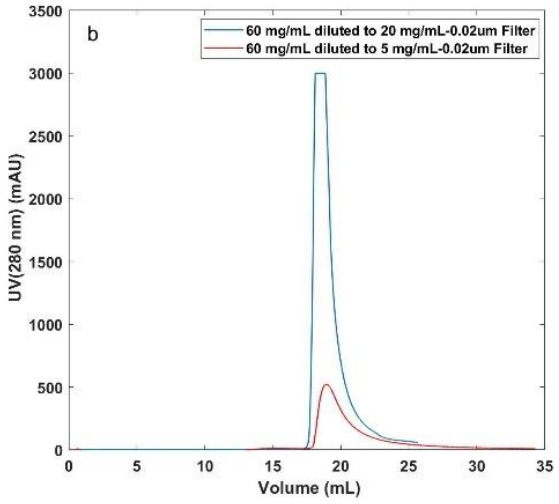

Figure. 3.15. SEC chromatogram of two different lysozyme solutions filtered through a 0.02 µm filter. Both results indicate that within the limit of SEC detection (670 – 1.35 kDa) the monomers are the only species present in the samples.

There is another intriguing direction of the mesoscopic aggregation in lysozyme solutions, which could be considered for the future study. In protein solutions there is

always a certain probability of the existence of short-lived folding intermediates caused by thermal fluctuations. If such fluctuation-induced intermediates could contribute to the formation of mesoscopic aggregates, the aggregates, after being removed by fine filtration, would eventually reemerge after sufficient structural fluctuations have developed.

The formation of mesoscopic aggregates in solutions of lysozyme resembles kinetically stable mesoscale inhomogeneities observed in aqueous solutions of non-ionic hydrotropes, such as tertiary butanol (Subramanian et al., 2014, Rak & Sedlak 2019), which are attributed to mesoscale solubilization of a hydrophobic impurity. On the other hand, our findings encourage revisiting the problem of mesoscopic aggregation in aqueous solutions of high molecular weight poly(ethylene)oxide (PEO) (Zheng et al., 2018), which is commonly interpreted as self-assembly of amphiphilic PEO macromolecules.

3.5. Conclusion

In summary, we reached to a new understanding of the nature and origin of mesoscopic aggregates found in high concentration solutions of lysozyme. Measurements of particle size distribution in the lysozyme solutions using DLS after systematic filtrations show no mesoscopic aggregates. Despite attempts to force the formation of mesoscopic aggregates in filtered lysozyme solutions with various stresses, aggregates with similar characteristics (size and relative population) did not

appear in the solutions. After confirming the absence of lysozyme oligomers in the lysozyme solutions using SEC and existence of repulsive interparticle interactions using SAXS, we conclude that mesoscopic aggregates are not in thermodynamic equilibrium with protein monomers or oligomers in solution.

Chapter 4: Summary and Future Directions

4.1. Key Findings

The study investigated the nature, source of formation, and reversibility of mesoscopic aggregates present in concentrated solutions of lysozyme. The concluding statements are as follows:

- We described the fundamental physical principles behind DLS and SAXS, which are the main techniques used for detection and characterization of aggregates in solutions.
- We demonstrated that mesoscopic aggregates could be removed irreversibly from the lysozyme solutions with systematic filtration down to 20 nm filter pore size.
- We demonstrated temperature elevation below the melting point of lysozyme increased the relative population of mesoscopic aggregates prior to the filtration, while after the filtration increased temperatures did not stimulate the formation of the mesoscopic aggregates.
- We determined the interactions between the lysozyme monomers are repulsive based on the positive value of the osmotic second virial coefficient measure by SAXS.
- We detected monomers as the only low molecular weight species in the lysozyme solutions, regardless of filtration.

- We provide a revised mechanism for the nature and formation source of mesoscopic aggregates, which are shown to be irreversible aggregates that are formed by contaminants in lyophilized proteins.

4.2. Suggestions for Future Work

4.2.1. Other Methods for Determining the Nature of Mesoscopic Aggregates

Based on our results, the mesoscopic aggregates were removed irreversibly from lysozyme solutions. We tested stresses such as heat and denaturants to force reformation of mesoscopic aggregates in filtered solutions, but were not successful in forming similar aggregates. We propose several future experiments to determine the formation source:

- Lyophilization of lysozyme may be a cause of mesoscopic aggregates. Lyophilizing filtered solutions could be used to test whether this process produces a small sub-population of partially unfolded proteins that are precursors for mesoscopic aggregates.
- Agitation of protein solutions can lead to aggregation. Stirring lysozyme solutions at several hundred rotations per minute would form aggregates.
- To investigate the composition of the aggregates, electron microscopy can be used for taking images of them and doing elemental analysis. This would enable determining whether aggregates are organic or inorganic in nature.
- Using SEC or other separation techniques, aggregates can be separated from solutions and be analyzed on their own.

4.2.2. Mesoscopic Aggregates in Other Proteins

A large body of work have been conducted on aggregation phenomena in proteins such as lysozyme, hemoglobin, and lumazine; however, the study of mesoscopic aggregation is mostly concentrated on the lysozyme. To gain a better understanding of the mesoscopic aggregation, similar studies should be conducted other proteins. Many proteins are thought to form similar mesoscopic aggregates. Another possible comparison would be to express or purify the hen egg white lysozyme directly in the lab and compare the lab-made solutions with commercially produced sources of lysozyme.

4.2.3. Alternative Methods for Detection of Oligomers

As mentioned in previous chapters, mesoscopic aggregates have been proposed to be in equilibrium with oligomers present in solution. The lifetime of the oligomers is claimed to be very short, making it difficult to detect them with common experimental techniques. SAXS is a powerful tool for detection of polydispersity in colloidal dispersions. With precise data interpretation, SAXS is likely to help with detection of the transient oligomers of lysozyme. Fluorescence anisotropy and fluorescence correlation spectroscopy are other possible technique for detection of oligomers of fluorescent molecule conjugated lysozyme.

Bibliography

Abeyrathne, E. D. N. S., Lee, H. Y., Ahn, D. U. (2014). "Sequential separation of lysozyme, ovomucin, ovotransferrin, and ovalbumin from egg white," *Poult Sci*, vol. 93, no. 4, pp. 1001–1009, Apr. 2014.

Alam, P., Siddiqi, K., Chturvedi, S. K., Khan, R. H. (2017). "Protein aggregation: From background to inhibition strategies," *International Journal of Biological Macromolecules*, vol. 103, pp. 208–219, Oct. 2017.

Andrews, J. M., Roberts, C. J. (2007). "A Lumry−Eyring Nucleated Polymerization Model of Protein Aggregation Kinetics: 1. Aggregation with Pre-Equilibrated Unfolding," *J. Phys. Chem. B*, vol. 111, no. 27, pp. 7897–7913, Jul. 2007.

Berne, B. J., Pecora, R. Dynamic light scattering. With applications to chemistry, biology, and physics., by Berne, B. J.; Pecora, R. New York, NY (USA): Wiley-Interscience, 8 + 376 p.

Boldon, L., Laliberte, F., Liu, L. (2015). "Review of the fundamental theories behind small angle X-ray scattering, molecular dynamics simulations, and relevant integrated application," *Nano Rev*, vol. 6, Feb. 2015.

Burchard, W. (1983), Light scattering from polymers. Berlin Heidelberg: Springer; 1983. Static and dynamic light scattering from branched polymers and biopolymers; pp. 1–124.

Chi, E. Y., Krishnan, S., Randolph, T. W., Carpenter, J. F. (2003). "Physical Stability of Proteins in Aqueous Solution: Mechanism and Driving Forces in Nonnative Protein Aggregation," *Pharm Res*, vol. 20, no. 9, pp. 1325–1336, Sep. 2003.

Davis, H. T., Winkelmann, J. (1997). "Statistical mechanics of phases, interfaces, and thin films, VCH publishers, New York 1996, ISBN 1-56081-513-2, 712 Seiten, 267 Abb., Preis: DM 135,00," *Berichte der Bunsengesellschaft für physikalische Chemie*, vol. 101, no. 3, pp. 641–642, Mar. 1997.

Debye, P. (1915). "Zerstreuung von Röntgenstrahlen," *Annalen der Physik*, vol. 351, no. 6, pp. 809–823, Jan. 1915.

Dekel, Y. *et al.* (2017). "Formation of multimeric antibodies for self-delivery of active monomers," *Drug Delivery*, vol. 24, no. 1, pp. 199–208, Jan. 2017.

Dharmaraj, V. L., Godfrin, P. D., Liu, Y., Hudson, S. D. (2016). “Rheology of clustering protein solutions,” *Biomicrofluidics*, vol. 10, no. 4, p. 043509, Jul. 2016.

Dill, K. A., Ozkan, S. B., Shell, M. S., Weikl, T. R. (2008). “The Protein Folding Problem,” *Annu Rev Biophys*, vol. 37, pp. 289–316, Jun. 2008.

Dobson, C. M. (2003). “Protein folding and misfolding,” *Nature; London*, vol. 426, no. 6968, pp. 884–90, Dec. 2003.

Ellis, R. J., Minton, A. P. (2006). “Protein aggregation in crowded environments,” *Biol. Chem.*, vol. 387, no. 5, pp. 485–497, May 2006.

Filipe, V., Jiskoot, Basmeleh, W., A. H., Halim, A., Schellekens, H., Brinks, V. (2012). “Immunogenicity of different stressed IgG monoclonal antibody formulations in immune tolerant transgenic mice,” *MAbs*, vol. 4, no. 6, pp. 740–752, Dec. 2012.

Freitag, A. J. *et al.* (2015). “Investigation of the immunogenicity of different types of aggregates of a murine monoclonal antibody in mice,” *Pharm. Res.*, vol. 32, no. 2, pp. 430–444, Feb. 2015.

Guinier, A. (1939). “La diffraction des rayons X aux très petits angles: application à l’étude de phénomènes ultramicroscopiques,” *Ann. Phys.*, vol. 11, no. 12, pp. 161–237, 1939.

Harding, SE. (1999). Protein hydrodynamics. In: Allen G, editor. Protein: a comprehensive treatise, vol 2. Greenwich, CT: JAI Press; 1999. pp. 271–305.

Hotze, E. M., Phenrat, T., Lowry, G. V. (2010). “Nanoparticle Aggregation: Challenges to Understanding Transport and Reactivity in the Environment,” *Journal of Environment Quality*, vol. 39, no. 6, p. 1909, 2010.

Jahn, T. R., Radford, S. E. (2005). “The Yin and Yang of protein folding,” *The FEBS Journal*, vol. 272, no. 23, pp. 5962–5970, Dec. 2005.

Jiménez, M., Rivas, G., Minton, A. P. (2007). “Quantitative characterization of weak self-association in concentrated solutions of immunoglobulin G via the measurement of sedimentation equilibrium and osmotic pressure,” *Biochemistry*, vol. 46, no. 28, pp. 8373–8378, Jul. 2007.

Joubert, M. K. *et al.* (2016). "Use of In Vitro Assays to Assess Immunogenicity Risk of Antibody-Based Biotherapeutics," *PLOS ONE*, vol. 11, no. 8, p. e0159328, Aug. 2016.

Koch, M. H. J., Vachette, P., Svergun, D. I. (2003). "Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution," *Quarterly Reviews of Biophysics*, vol. 36, no. 2, pp. 147–227, May 2003.

Koch, T. J. (2015). "Aggregation Propensity: Characterization of Monoclonal Antibody Stability," Oct. 2015.

Koppel, D. E. (1972). "Analysis of Macromolecular Polydispersity in Intensity Correlation Spectroscopy: The Method of Cumulants," Journal of Chemical Physics, vol. 57, pp. 4814–4820, Dec. 1972.

Kovalchuk, N. M., Starov, V. M. (2012). "Aggregation in colloidal suspensions: Effect of colloidal forces and hydrodynamic interactions," *Advances in Colloid and Interface Science*, vol. 179–182, pp. 99–106, Nov. 2012.

Kramer, R. M., Shende, V. R., Motl, N., Pace, C. N., Scholtz, J. M. (2012). "Toward a Molecular Understanding of Protein Solubility: Increased Negative Surface Charge Correlates with Increased Solubility," *Biophysical Journal*, vol. 102, no. 8, pp. 1907–1915, Apr. 2012.

Lauer, M., Agrawal, N. J., Chennamsetty, N., Egodage, K., Helk, B., Trout, B. L. (2012). "Developability index: a rapid in silico tool for the screening of antibody aggregation propensity," *J Pharm Sci*, vol. 101, no. 1, pp. 102–115, Jan. 2012.

Li and Baron, "Dynamic Light Scattering." [Online]. Available: https://cnx.org/contents/P8mNrZNN@2/Dynamic-Light-Scattering. [Accessed: 05-Dec-2018].

Li, Y., Lubchenko, V., Vorontsova, Filobelo, M. A., L., Vekilov, P. G. (2012). "Ostwald-Like Ripening of the Anomalous Mesoscopic Clusters in Protein Solutions," *J. Phys. Chem. B*, vol. 116, no. 35, pp. 10657–10664, Sep. 2012.

Li, Y., Roberts, C. J. (2009). "A Lumry-Eyring Nucleated-Polymerization (LENP) Model of Protein Aggregation Kinetics 2. Competing Growth via Condensation- and Chain-Polymerization," *J Phys Chem B*, vol. 113, no. 19, pp. 7020–7032, May 2009.

Liu, H., Zwart, P. H. (2012). “Determining pair distance distribution function from SAXS data using parametric functionals,” *Journal of Structural Biology*, vol. 180, no. 1, pp. 226–234, Oct. 2012.

Lü, C., Yang, B. (2009). “High refractive index organic–inorganic nanocomposites: design, synthesis and application,” *Journal of Materials Chemistry*, vol. 19, no. 19, pp. 2884–2901, 2009.

Pan, W., Vekilov, P. G., Lubchenko, V. (2010). “Origin of Anomalous Mesoscopic Phases in Protein Solutions,” *J. Phys. Chem. B*, vol. 114, no. 22, pp. 7620–7630, Jun. 2010.

Parmar, A. S., Gottschall, P. E., Muschol, M. (2007). “Pre-assembled clusters distort crystal nucleation kinetics in supersaturated lysozyme solutions,” *Biophysical Chemistry*, vol. 129, no. 2, pp. 224–234, Sep. 2007.

Price,W. S., Tsuchiya, F., Arata, Y. (1999). “Lysozyme Aggregation and Solution Properties Studied Using PGSE NMR Diffusion Measurements,” *Journal of the American Chemical Society*, vol. 121, no. 49, pp. 11503–11512, Dec. 1999.

Roberts, C. J. (2006). “Nonnative Protein Aggregation,” in *Misbehaving Proteins: Protein (Mis)Folding, Aggregation, and Stability*, R. M. Murphy and A. M. Tsai, Eds. New York, NY: Springer New York, 2006, pp. 17–46.

Roberts, C. J. (2007). “Non-native protein aggregation kinetics,” *Biotechnol. Bioeng.*, vol. 98, no. 5, pp. 927–938, Dec. 2007.

Roberts, C. J. (2014). “Therapeutic protein aggregation: mechanisms, design, and control,” *Trends Biotechnol.*, vol. 32, no. 7, pp. 372–380, Jul. 2014.

Russel, W. B., Saville, D. A., Schowalter, W. R. (1991). *Colloidal Dispersions*. Cambridge University Press, 1991.

Safari, M. S., Byington, M. C., Conrad, J. C., Vekilov, P. G. (2017). “Polymorphism of Lysozyme Condensates,” *J. Phys. Chem. B*, vol. 121, no. 39, pp. 9091–9101, Oct. 2017.

Saluja, A., Kalonia, D. S. (2008). “Nature and consequences of protein-protein interactions in high protein concentration solutions,” *Int J Pharm*, vol. 358, no. 1–2, pp. 1–15, Jun. 2008.

Saunders, A. E., Korgel, B. A. (2004). “Second Virial Coefficient Measurements of Dilute Gold Nanocrystal Dispersions Using Small-Angle X-ray Scattering,” *J. Phys. Chem. B*, vol. 108, no. 43, pp. 16732–16738, Oct. 2004.

Schnablegger, H., Singh, Y. (2013). “The SAXS Guide, Getting acquainted with the principles 4th edition,” p. 14.

Siegert, A. J. F. (1943). *On the fluctuations in signals returned by many independently moving scatterers*. Cambridge, Mass.: Radiation Laboratory, Massachusetts Institute of Technology, 1943.

Silvers, T. R., Myers, J. K. (2013). “Osmolyte Effects on the Self-Association of Concanavalin A: Testing Theoretical Models,” *Biochemistry*, vol. 52, no. 51, pp. 9367–9374, Dec. 2013.

Skou, S., Gillilan, R. E., Ando, N. (2014). “Synchrotron-based small-angle X-ray scattering of proteins in solution,” *Nature Protocols*, vol. 9, no. 7, pp. 1727–1739, Jul. 2014.

Stetefeld, J., McKenna, S. A., Patel, T. R. (2016). “Dynamic light scattering: a practical guide and applications in biomedical sciences,” *Biophys Rev*, vol. 8, no. 4, pp. 409–427, Oct. 2016.

Tande, B. M., Wagner, N. J., Mackay, M. E., Hawker, C. J., Jeong, M. (2001). “Viscosimetric, Hydrodynamic, and Conformational Properties of Dendrimers and Dendrons,” *Macromolecules*, vol. 34, no. 24, pp. 8580–8585, Nov. 2001.

Uchino, T., Miyazaki, Y., Yamazaki, T., Kagawa, Y. (2017). “Immunogenicity of protein aggregates of a monoclonal antibody generated by forced shaking stress with siliconized and nonsiliconized syringes in BALB/c mice,” *J. Pharm. Pharmacol.*, vol. 69, no. 10, pp. 1341–1351, Oct. 2017.

Vekilov, P. G. *et al.* (2018). “Weakly-bound Dimers that Underlie the Crystal Nucleation Precursors in Lysozyme Solutions,” *bioRxiv*, p. 275222, Mar. 2018.

Vorontsova, Chan, M. A., Lubchenko, H. Y., V., Vekilov, P. G. (2015). “Lack of Dependence of the Sizes of the Mesoscopic Protein Clusters on Electrostatics,” *Biophysical Journal*, vol. 109, no. 9, pp. 1959–1968, Nov. 2015.

Wang, W. C., Roberts, J. (2018). “Protein aggregation – Mechanisms, detection, and control,” *International Journal of Pharmaceutics*, vol. 550, no. 1, pp. 251–268, Oct. 2018.

Weiss, Thomas M. (2016). “introduction to small angle x-ray scattering” BioSAXS workshop 2016; Stanford University

Weiss, W. F., Young, T. M., Roberts, C. J. (2009). “Principles, approaches, and challenges for predicting protein aggregation rates and shelf life,” *J Pharm Sci*, vol. 98, no. 4, pp. 1246–1277, Apr. 2009.

Yamazaki, T. *et al.* (2017). “Two types of amorphous protein particles facilitate crystal nucleation,” *PNAS*, vol. 114, no. 9, pp. 2154–2159, Feb. 2017.

Young, T. M., Roberts, C. J. (2007). “A quasichemical approach for protein-cluster free energies in dilute solution,” *The Journal of Chemical Physics*, vol. 127, no. 16, p. 165101, Oct. 2007.

Young, T. M., Roberts, C. J. (2009). “Structure and thermodynamics of colloidal protein cluster formation: Comparison of square-well and simple dipolar models,” *J. Chem. Phys.*, vol. 131, no. 12, p. 125104, Sep. 2009.

Yudin, I. K., Nikolaenko, G. L., Kosov, V. I., Agayan, V. A., Anisimov, M. A., Sengers, J. V. (1997). “A compact photon-correlation spectrometer for research and education,” Int J Thermophys, vol. 18, no. 5, pp. 1237–1248, Sep. 1997.

Zamora, P. C., Zukoski, C. F. (1996). “Interactions and Phase Behavior of Nanosized Particles,” *Langmuir*, vol. 12, no. 15, pp. 3541–3547, Jan. 1996.

Zimm, B. H. (1948). “The Scattering of Light and the Radial Distribution Function of High Polymer Solutions,” *J. Chem. Phys.*, vol. 16, no. 12, pp. 1093–1099, Dec. 1948.

B. R. Pauw, “How to do a perfect SAXS measurement,” p. 50.

Thomas, B. R., P. G. Vekilov, and F. Rosenberger,1996. “Heterogeneity determination and purification of commercial hen egg-white lysozyme,” *Acta Crystallogr. D Biol. Crystallogr.*, vol. 52, no. Pt 4, pp. 776–784, Jul.

Subramanian, D, Jeffery B. Klauda, Peter J. Collings, and Mikhail A. Anisimov (2014).” Mesoscale Phenomena in Ternary Solutions of Tertiary Butyl Alcohol, Water, and Propylene Oxide” *J. Phys. Chem. B, 118* (22), 5994-6006

Rak, D. and M. Sedlák, 2019. “On the Mesoscale Solubility in Liquid Solutions and Mixtures,” *The Journal of Physical Chemistry B*, vol. 123, no. 6, pp. 1365–1374, Feb..