Pharmaceutical Research, Vol. 23, No. 6, June 2006 (© 2006)
DOI: 10.1007/s11095-006-0018-y

Research Paper

Protein Structural Conformation and Not Second Virial Coefficient Relates to Long-Term Irreversible Aggregation of a Monoclonal Antibody and Ovalbumin in Solution

Harminder Bajaj,[1] Vikas K. Sharma,[1,4] Advait Badkar,[2] David Zeng,[2] Sandeep Nema,[2] and Devendra S. Kalonia[1,3,5]

Received October 31, 2005; accepted January 17, 2006

Purpose. The purpose of the study was to investigate the relationship of the second virial coefficient, B_{22}, to the extent of irreversible protein aggregation upon storage.
Methods. A monoclonal antibody and ovalbumin were incubated at 37°C (3 months) under various solution conditions to monitor the extent of aggregation. The B_{22} values of these proteins were determined under similar solution conditions by a modified method of flow-mode static light scattering. The conformation of these proteins was studied using circular dichroism (CD) spectroscopy and second-derivative Fourier transform infrared spectroscopy.
Results. Both proteins readily aggregated at pH 4.0 (no aggregation observed at pH 7.4); the extent of aggregation varied with the ionic strength and the presence of cosolutes (sucrose, glycine, and Tween 80). Debye plots of the monoclonal antibody showed moderate attractive interactions at pH 7.4, whereas, at pH 4.0, nonlinear plots were obtained, indicating self-association. CD studies showed partially unfolded structure of antibody at pH 4.0 compared with that at pH 7.4. In the case of ovalbumin, similar B_{22} values were obtained in all solution conditions irrespective of whether the protein aggregated or not. CD studies of ovalbumin indicated the presence of a fraction of completely unfolded as well as partially unfolded species at pH 4.0 compared with that at pH 7.4.
Conclusions. The formation of a structurally altered state is a must for irreversible aggregation to proceed. Because this aggregation-prone species could be an unfolded species present in a small fraction compared with that of the native state or it could be a partially unfolded state whose net interactions are not significantly different compared with those of the native state, yet the structural changes are sufficient to lead to long-term aggregation, it is unlikely that B_{22} will correlate with long-term aggregation.

KEY WORDS: antibody formulation; monoclonal antibody; protein aggregation; protein stability; second viral coefficient.

INTRODUCTION

Irreversible protein aggregation is a critical issue during protein purification, formulation, production, and storage (1,2). From a therapeutic point of view, these aggregates could have compromised activity and often lead to enhanced immunogenicity (3). It is generally believed that irreversible aggregation proceeds via a nonnative (4–6), expanded aggregation-prone conformational state (7,8). This perturbation is followed by formation of the aggregated state, which has a lower free energy than the native conformation of the protein and is composed of nonnative states of the protein. The kinetics of protein aggregation is often described by the classical Lumry-Eyring equation (9)

$$N \Leftrightarrow U \rightarrow A \tag{1}$$

where N represents the native state, U represents the perturbed state, and A represents the aggregated state. Whereas the first step is an equilibrium step, the second step is governed by the kinetics of aggregation of the aggregation-prone state. Examples exist to indicate that the aggregation-prone state could be slightly perturbed or near native state (10), molten globule-like (11–14), a partially unfolded intermediate (15–18), or even the completely unfolded state (19).

The complexities in the mechanism of aggregation are attributed to the complex structural properties of proteins and the complex behavior of different proteins under varied solution conditions toward unfolding and aggregation. Thus, prediction of long-term aggregation stability from accelerated stability studies has largely remained a challenge to protein scientists. This is further complicated by a change in

[1] Department of Pharmaceutical Sciences, University of Connecticut, Storrs, Connecticut 06269, USA.
[2] Pfizer Biologics, St. Louis, Missouri 63017, USA.
[3] U-3092, School of Pharmacy, University of Connecticut, Storrs, Connecticut 06269, USA.
[4] Present address: Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.
[5] To whom correspondence should be addressed. (e-mail: kalonia@uconn.edu)

the protein structural conformation and properties as a function of temperature, often leading to a non-Arrhenius behavior in the kinetics of protein aggregation (20).

Although the role of conformational stability in protein aggregation has been well documented (21), recently, there has been interest in the role of colloidal interactions in protein aggregation under conditions that favor native state (22–29). The aggregation mediated through a structurally altered state is also a result of colloidal interactions, where intermolecular hydrophobic interactions between partially or completely unfolded states are overwhelmingly strong over electrostatic repulsions caused by the exposure of otherwise buried hydrophobic groups. Whether it is aggregation under conditions that favor native state or under conditions where structurally altered species is also present, protein–protein interactions will govern the extent and kinetics of aggregation. Hence, it is reasonable to assume that an understanding of protein–protein interactions can provide important insights into the mechanism of irreversible protein aggregation.

One of the experimentally accessible parameter used to study protein–protein interactions is the second virial coefficient, B_{22}. B_{22} appears in the virial expansion of the osmotic pressure term to represent first deviation from ideality in dilute colloidal solutions (30). Positive B_{22} values indicate net repulsive protein–protein interactions, and negative B_{22} values indicate net attractive protein–protein interactions. Through a rigorous theoretical statistical thermodynamic analysis (31), it has been shown that B_{22} can be represented in terms of the excluded volume term and the interaction free energy, ΔG, taking into account the steric, electrostatic, hydrophobic, and short-range interactions along with the structural and functional anisotropy of the protein molecules (32)

$$B'_{22} = -\frac{1}{16\pi^2}\int_{\Omega}^{2\pi}\left[\frac{1}{3}r_c^3 - \int_{r_c}^{\infty}\left(e^{-\Delta G/kT} - 1\right)r_{12}^2 \mathrm{d}r_{12}\right]\mathrm{d}\Omega \quad (2)$$

where r_c is the center-to-center distance on contact for a particular orientation, ΔG is the interaction free energy, r_{12} is the intermolecular center-to-center distance, and Ω represents all possible orientations. The first term in the brackets represents the excluded volume contribution, which is always positive, and the second term, which represents the energetic interactions, leads to positive or negative B_{22} values depending on repulsive or attractive interactions. Thus, experimental work (33–35) and theoretical analysis (31,32) strongly suggest that B_{22} can serve as an important parameter to represent net protein–protein interactions.

Wilson *et al.* (36–38) have established the usefulness of B_{22} in predicting the phase behavior of proteins in solution and conditions where protein crystallization is likely to occur. Recently, a few reports (22,26,29,39) have emerged on studies involving B_{22} and aggregation; however, it still remains unclear whether a correlation exists between B_{22} values and irreversible aggregation of proteins. For example, Chi *et al.* (39) have shown using recombinant human granulocyte colony-stimulating factor that negative B_{22} values correlate with aggregation under conditions where native protein conformation is favored, but aggregation was also observed in solution conditions yielding positive B_{22} values. This was attributed to the overwhelming contribution of dimers instead of the aggregation-prone unfolded monomer species to the overall B_{22}.

The accurate and reliable determination of B_{22} values of proteins in aqueous solutions by the method of batch-mode static light scattering is often affected by the presence of higher-order aggregates and dust particles in a solution. We have recently developed (40) a chromatography-based method to determine B_{22} values of proteins in aqueous solutions that is based on simultaneous measurement of scattered light intensity and protein concentration following the elution of the protein through a size-exclusion column. This method utilizes a custom-designed dual-detector cell in conjunction with size-exclusion chromatography (SEC) that measures the intensity of the scattered light through a 90° light scattering detector and the concentration through a UV detector at the same time. A single protein injection is sufficient to generate a range of protein concentrations and corresponding light scattering intensities in the eluting peak. The concentration and scattering data are then used to generate Debye plot using the Debye's light scattering equation

$$\frac{Kc}{R_\theta} = \frac{1}{M} + 2B_{22}c \quad (3)$$

where R_θ is the excess Rayleigh's ratio of the protein in a solution of concentration c and M is the weight average molecular weight of the protein. K is the optical constant and is defined as

$$K = \frac{4\pi^2 n^2 (\mathrm{d}n/\mathrm{d}c)^2}{N_A \lambda_o^4} \quad (4)$$

where n is the solvent refractive index, $\mathrm{d}n/\mathrm{d}c$ is the refractive index increment, λ is the wavelength of the incident light, and N_A is the Avogadro's number. Thus, B_{22} is obtained from the slope of the linear Debye plot (Kc/R_θ *vs.* c). We have shown that this method provides reliable and accurate estimates of B_{22} using such proteins as lysozyme and chymotrypsinogen. The advantages of using SEC is that the aggregates, dust particles, and low molecular weight cosolutes are separated from the monomeric peak, and hence, the virial coefficient obtained represents that of the pure monomer. Recently, we have been able to apply this technique to quantitatively characterize protein self-association, thus making it a very powerful technique for protein characterization (Sharma *et al.*, unpublished data). Hence, from a single protein injection, irreversible aggregates could be separated, and at the same time, if the protein undergoes reversible self-association, the association constants can be estimated. Because the association results in nonlinear Debye plots, a simple inspection of the Debye plot will indicate the protein behavior in a solution.

In the present report, we have investigated the relationship between irreversible aggregation upon storage and B_{22} values to gain a better understanding on the role of B_{22} in its ability to relate to the tendency of a protein to undergo aggregation. We utilized two different proteins for this investigation, a monoclonal antibody (M_w = 144 kDa, pI = 7–9) and ovalbumin (M_w = 44 kDa, pI = 4.5). Both proteins were subjected to storage stability under conditions that led to aggregation as well as those where these proteins were completely resistant to aggregation. B_{22} values were obtained

under similar conditions using the methodology recently developed by us (40). We have shown that the second virial coefficient by itself does not correlate with the extent of aggregation; however, the Debye plots, used to obtain B_{22} values, provide important insights into the aggregation mechanism in conjunction with the spectroscopic studies.

MATERIALS AND METHODS

Materials

All buffer components and chemical reagents used in the present studies were of the highest purity grade, obtained from commercial sources, and used without further purification. The monoclonal antibody was donated generously by Pfizer Biologics (St. Louis, MO, USA) and was supplied as a 11.3 mg/mL solution in a 20 mM acetate buffer (pH 5.5) containing 140 mM NaCl and 0.02% w/v polysorbate 20. The monoclonal antibody is an IgG2 with kappa light chains and a M_w of 144 kDa. Ovalbumin, L-glycine, sucrose, and Tween 80 were obtained from Sigma Chemicals Co. (St. Louis, MO, USA). Double-distilled water that was filtered through a 0.1-μm polycarbonate membrane filter was used for the preparation of the protein solutions and the mobile phase.

Methods

The following solution conditions were used for determination of B_{22} and for aggregation studies: pH 4.0 (μ = 0.04 M and 0.3M), pH 7.4 (μ = 0.3 M), and pH 5.4 (μ = 0.3 M), 0.1% w/v Tween 80 (pH 4.0, μ = 0.3 M), 0.3 M glycine (pH 4.0, μ = 0.3 M), and 10% w/v sucrose (pH 4.0, μ = 0.3 M). The ionic strength of all solutions was adjusted using NaCl. The concentration of the buffer was fixed to 10 mM for the studies; acetate buffer was used for all studies performed at pH 4 and 5.4, whereas sodium phosphate salts were used to make buffers at pH 7.4. The final pH of all solutions was measured using a Piccoloplus Hi-1295 digital pH meter (Fisher Scientific, Pittsburgh, PA, USA) and adjusted to the desired pH using either 1.0 N NaOH or 1.0 N HCl. Because the antibody was supplied as a solution, it was dialyzed against appropriate buffers through ultradialysis. For the purpose of dialysis, the ionic strength of the solution was kept low (0.04 M) to prevent aggregation and was later adjusted using concentrated NaCl solutions. The ovalbumin solutions were prepared directly by dissolving the protein into the appropriate buffer. The pH of all solutions was readjusted following complete solubilization of the protein. Protein concentration was determined by measuring absorbance at 280 nm and using an $E_{1\%}$ value of 14.0 for antibody and 7.0 for ovalbumin.

Aggregation Studies

Stability studies were carried out to monitor irreversible physical aggregation of the antibody and ovalbumin upon storage at 37°C for a period of 3 months. Protein samples were prepared (7.5 mg/mL) in appropriate buffers, and 1 mL of the protein solution in each buffer in a 1.5-mL polypropylene microcentrifuge tube was kept (in triplicate) in a controlled temperature incubator at 37°C. Samples were withdrawn at 1-month intervals and were analyzed for monomer content using SEC. To this end, the microcentrifuge tubes were first centrifuged at 10,000 rpm for 15 min. Following this, 150 μL of the protein solution from the supernatant was injected into a SEC column (YMC-pack Diol-200, DL20S05-3008WT column, 200-Å pore size, 5-μm bead size, and 30 × 0.8 cm column dimensions; YMC, Kyoto, Japan). A Precision Detectors PD 2000 system encasing a dual-source dual-detector cell with a 90° light scattering detector along with the UV detector was used for the molecular weight analysis of the amount of monomer present in the injected sample (see "Determination of second virial coefficient"). The monomer content of the samples kept for stability, evaluated from the area under the monomer peak, was compared with that of the initial samples (i.e., at $t = 0$) to obtain the percent monomer loss. Note that the aggregates formed were physical aggregates and not covalent aggregates, as these dissolved completely in 6 M guanidine hydrochloride solution.

Determination of Second Virial Coefficient Values

The B_{22} values of the antibody and ovalbumin in various solutions were determined at 37°C. As described above, B_{22} values were obtained from the slope of the linear Debye plots, wherever applicable. The Debye plots were obtained using the method recently developed by us and as described previously (40). This method utilizes a Precision Detectors PD 2000 detection system that encases a specially designed dual-source dual-detector cell for simultaneous detection of protein concentration and scattered light intensity as the protein elutes from a SEC column. The detection system was connected to a Spectra Physics P4000 pump in conjunction with a Rheodyne 7725 manual injector with a 200-μL injection loop. A flow rate of 1.0 mL/min was used for all studies. Briefly, 150 μL of the protein solution of a known concentration (10 mg/mL, unless otherwise specified) was injected into a SEC column, and the eluting protein was simultaneously detected for scattered light intensity and concentration using a dual-detector cell consisting of a concentration detector (UV) and a 90° light scattering detector. This cell has a volume of 10 μL, and the scattering volume is 0.01 μL. The path length for UV measurements is 3 mm. A YMC-pack Diol-200, DL20S05-3008WT column was used for both proteins.

Following elution from the column, the chromatograms obtained from the UV detector and the light scattering detector were analyzed to generate the Debye plot. A range of concentrations and corresponding scattered light intensities, which correspond to the data points on the latter half of the peak, were obtained from a single protein injection. Each data point is then converted to Rayleigh's ratio, R_θ (light scattering detector), and concentration (UV detector) as described below.

The molecular weight of the protein sample in dilute solutions and for polarized light is related to the intensity of the scattered light from the sample, through the following equation:

$$M_w = \frac{N_A \lambda_o^4 R^2 I_s}{4\pi^2 \sin^2 \phi c (dn/dc)^2 n^2 I_o} \quad (5)$$

where N_A is the Avogadro's number, λ_o is the wavelength of the incident radiation, R is the distance of the sample from the detector, I_s is the intensity of the scattered light, I_o is the intensity of the incident light, c is the concentration of protein sample, dn/dc is the refractive index increment of protein solution, ϕ is the angle between the plane of the incident polarized light and the scattering detector, and n is the refractive index of the solvent. Collecting all the constants and instrument parameters into an overall light scattering instrument constant, A_{90}, Eq. (5) can be written as

$$M_{\mathrm{w}} = \frac{I_s}{A_{90}c(\mathrm{d}n/\mathrm{d}c)^2} \tag{6}$$

where

$$A_{90} = \frac{I_o 4\pi^2 n^2}{N_A \lambda_o^4 R^2} \tag{7}$$

Because the intensity of the incident radiation, I_o, and the distance between the sample and detector, R, are fixed, the ratio of these two parameters can be obtained by rearranging the above equation and is represented as K_1, i.e.,

$$\frac{R^2}{I_o} = \frac{4\pi^2 n^2}{N_A \lambda_o^4 A_{90}} = K_1 \tag{8}$$

Hence, K_1 can be simply obtained from the instrument constant A_{90}, the wavelength of the incident light (685 nm), and the refractive index of the solution. Rayleigh's ratio at 90° scattering angle is defined as

$$R_{90} = \frac{I_s R^2}{I_o} \tag{9}$$

Combining Eqs. (8) and (9), Rayleigh's ratio can now be expressed as

$$R_{90} = K_1 I_s \tag{10}$$

Eq. (10) is used to obtain Rayleigh's ratio of a given data point on the light scattering chromatogram once the instrument has been calibrated using bovine serum albumin (BSA) as the standard (see "Calibration").

The concentration for each corresponding data point on the UV chromatogram was estimated from the UV signal intensity. In the present instrument configuration, the UV chromatogram represented the intensity of the transmitted light. Hence, the concentration of the injected protein at each data point was estimated using the following equation:

$$c_{(g/ml)} = \log\left(\frac{I_{100\%\mathrm{T}} - I_{0\%\mathrm{T}}}{I_a - I_{0\%\mathrm{T}}}\right) \cdot 10 \Big/ (E_{1\%} b) \tag{11}$$

where c is the concentration of the protein, $I_{100\%\ \mathrm{T}}$ is the intensity of the UV signal at the baseline, $I_{0\%\ \mathrm{T}}$ is the signal of the UV detector in off-mode, I_a is the UV signal at a given time point on the chromatogram, $E_{1\%}$ is the extinction coefficient of 1% protein solution, and b is the path length of the UV cell (3 mm).

Once the R_θ values and the corresponding concentrations are obtained for data at each time point on the chromatogram, Debye plot is then generated using the Debye equation [Eq. (3)]. For estimation of parameter K in Eq. (3), the dn/dc values were obtained from the differential refractive index (DRI) detector connected in-line with the dual-detector cell, following calibration of this detector using a standard of known dn/dc (see "Calibration").

In the present studies, linear Debye plots were analyzed according to Eq. (3). For Debye plots obtained otherwise, the Debye equation was modified as described in "Results," "Discussion," and Appendix.

Calibration

The calibration of the instrument was carried out to determine the constant A_{90} for the determination of R_θ and the DRI constant, defined as B, to determine the dn/dc of a given protein. For this purpose, BSA was used as the standard. One hundred microliters of a 2 mg/mL BSA solution in pH 7.4 was injected into a TSK3000SWXL size-exclusion column. A dn/dc of 0.167 and molecular weight of 66,000 were used to calculate calibration constants from the monomer peak of BSA. Under these conditions, the following calibration constants were obtained using the Precision Analyze software: $K_{90} = (B/A_{90}) = 4569.8$ and $B = 54{,}618.1$. A_{90} is then obtained by dividing B with K_{90}. Once the DRI constant, B, is obtained, the dn/dc of any given protein for a given solution condition is determined using the following equation:

$$\mathrm{d}n/\mathrm{d}c = \frac{RI_{sample}}{BA_{inj}} \tag{12}$$

where RI_{sample} is the area of the RI chromatogram and A_{inj} is the amount of protein injected (in micrograms).

Circular Dichroism Studies

Circular dichroism (CD) measurements were carried out using a Jasco-710 spectropolarimeter. The far-UV CD studies were carried out in a 0.05-cm path length cell using a protein concentration of 0.25 mg/mL. The spectra were collected at a scan speed of 20 nm/min from 190 to 260 nm. The near-UV CD studies were carried out in a 1.0-cm path length cell using a protein concentration of 1.0 mg/mL. The spectra were collected at a scan speed of 20 nm/min from 240 to 310 nm. Each scan was a result of five accumulations to increase signal-to-noise ratio. All scans were normalized for concentration and number of amino acid residues by converting the obtained ellipticities to mean residue ellipticities. Secondary structure content was estimated using the CONTINLL analysis program provided with the CDPro software suite (41).

Fourier Transform Infrared Spectroscopy Studies

Fourier transform infrared (FTIR) studies were carried out on the proteins at pH 4.0 (μ = 0.3 M) and that of the aggregated protein to compare the secondary structure of the protein in solution and following aggregation. The spectra were recorded in the region of 4000 to 400 cm^{-1} on a Nicolet Magna 560 FTIR spectrometer (Nicolet, Inc., Madison, WI, USA) equipped with a nitrogen-cooled deuterated triglycine sulfate detector at a resolution of 4 cm^{-1}. The spectrometer

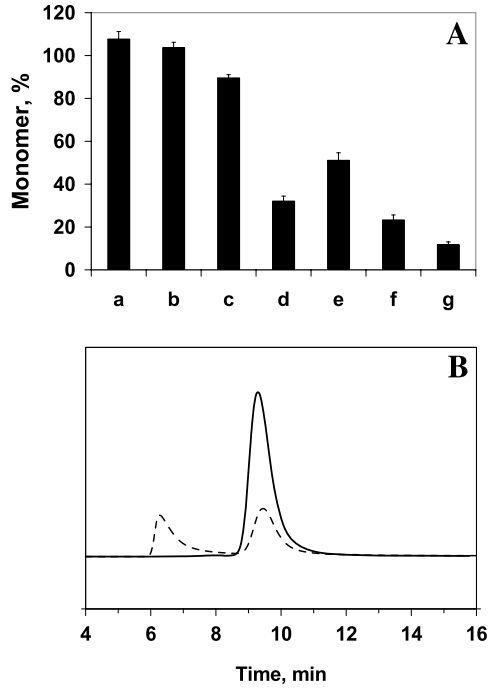


Fig. 1. (A) Percent monomer of the monoclonal antibody remaining after storage at 37°C for a period of 3 months as determined by size-exclusion chromatography (SEC). a, pH 7.4 (μ = 0.3 M); b, pH 5.4 (μ = 0.3 M); c, pH 4.0 (μ = 0.04 M); d, pH 4.0 (μ = 0.3 M); e, 10% w/v sucrose (pH 4.0, μ = 0.3 M); f, 0.3 M glycine (pH 4.0, μ = 0.3 M); g, 0.1% w/v Tween 80 (pH 4.0, μ = 0.3 M). (B) SEC chromatograms of the monoclonal antibody at $t = 0$ (solid line) and after storage at 37°C (pH 4.0, μ = 0.3 M) for a period of 3 months (dashed line).

was continuously purged with dry nitrogen. The protein solution with and without aggregates was injected into a CaF_2 window cell (6-μm spacer) (42). The absorbance of the buffer was subtracted from that of the protein to obtain the protein FTIR spectra. The criterion used was to obtain a straight baseline in the 1800- to 2200-cm^{-1} region (43). The subtracted spectra were smoothed using a nine-point smoothing function, and the second derivatives of these scans were obtained in the 1750- to 1550-cm^{-1} region (amide I). The derivatized spectra were baseline-subtracted and, finally, area-normalized to unit areas for relative comparison.

RESULTS

Monoclonal Antibody

Aggregation Studies

Figure 1A shows the percent monomer remaining of the antibody after incubation at 37°C for a period of 3 months under various solution conditions. The antibody did not aggregate at pH 7.4 or 5.4 (μ = 0.3 M); however, aggregation to different extents was observed at pH 4.0 for all solution conditions. The extent of aggregation increased with an increase in the NaCl concentration at pH 4.0, 10 mM acetate buffer (compare 0.04 to 0.3 M, ionic strength), presumably because at higher ionic strength, effective shielding of charges takes place, leading to the enhanced hydrophobic interactions over weakened electrostatic repulsions. At pH 4.0, μ = 0.3 M, only sucrose exhibited partial protective effect, whereas the extent of aggregation further increased with Tween 80 and glycine. Both insoluble and soluble aggregates formed at pH 4.0, as indicated by the visible turbidity in the incubated samples and the presence of high molecular weight species evident from the SEC chromatograms of the aggregated sample (Fig. 1B).

Second Virial Coefficient Studies

We next examined the B_{22} values of this protein under similar solution conditions used for aggregation studies. Figure 2 shows the Debye plots of the antibody at pH 7.4 (no aggregation observed) and pH 4.0 (aggregation observed) at μ = 0.3 M. At pH 7.4, a linear Debye plot is observed with a negative B_{22} value ($B_{22} = -8.7 \times 10^{-5}$). Although the negative B_{22} value indicates net attractive interactions between the protein molecules, this attraction is not enough to cause aggregation. In fact, this value is much less as compared with the values reportedly required for the crystallization of proteins ($\sim 1 \times 10^{-4}$ to 8×10^{-4}), often referred to as the crystallization slot (36).

At pH 4.0, interestingly, a nonlinear upward-curving Debye plot is obtained. This type of nonlinearity in Debye plot is typical of a self-associating system and has been reported previously by several authors (44,45) and recently by us (Sharma *et al.*, unpublished data). Thus, these data indicate that the antibody exhibits reversible self-association under the solution conditions where it also undergoes aggregation upon storage, i.e., at pH 4.0 (μ = 0.3 M). To further confirm the fact that conditions which promote irreversible aggregation also result in protein self-association in the initial solution for this antibody, we obtained Debye plots in all solution conditions where aggregation was observed upon storage. The Debye plots of the antibody in such solution conditions are shown in Fig. 3A–D for pH 4.0, μ = 0.3 M, 0.3 M glycine (pH 4.0, μ = 0.3 M), 0.1% w/v Tween 80 (pH 4.0, μ = 0.3 M), and 10% w/v sucrose (pH 4.0, μ = 0.3 M), respectively. Clearly, curved Debye plots were obtained in all cases, indicating that the antibody self-associates in these solution conditions.

The nonlinear Debye plots were analyzed using the modified Debye's equation as described in Appendix [Eq. (20)] to obtain the association constants, K_{dim}, monomer molecular weight, and the second virial coefficient. Table I summarizes

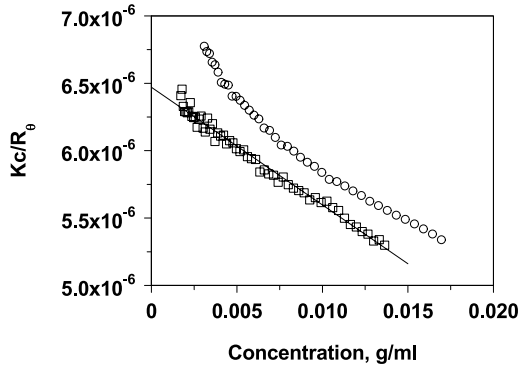


Fig. 2. Debye plots of the monoclonal antibody at pH 7.4 (squares) and pH 4.0 (circles) at μ = 0.3 M. The line is obtained after linear regression of the data points using Eq. (3).

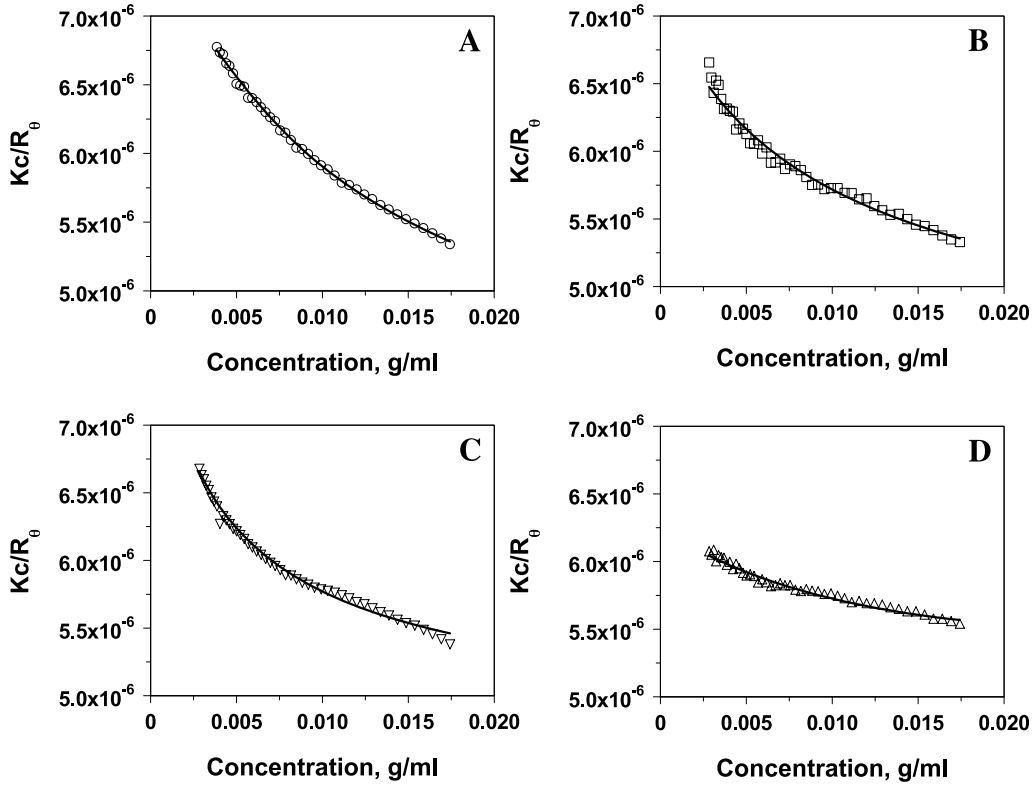


Fig. 3. Debye plots of the monoclonal antibody at pH 4.0 (μ = 0.3 M). (A) No excipient; (B) 0.3 M glycine; (C) 0.1% w/v Tween 80; (D) 10% w/v sucrose. The markers represent the experimental data, and the lines are obtained after fitting Eq. (20) to the data. The values of the parameters K_{dim}, M_w, and B or B_{22} are shown in Table I.

the results of the analysis for all solution conditions. Note that linear Debye plot was obtained for pH 5.4 (μ = 0.3 M), where no aggregation was observed upon storage, whereas a nonlinear Debye plot was obtained for pH 4.0 (μ = 0.04 M), where aggregation was observed. K_{dim} values obtained were of the order of 10^3, indicating moderately strong binding. In the case of sucrose, a lower K_{dim} value was obtained at pH 4.0 (μ = 0.3 M), indicating that sucrose inhibited the self-association of the antibody. Rather low values of B were obtained from the fitting of the Debye plots, indicating that the nonideality plays only a small role in the association of the antibody molecules. In our analysis, the molecular weight was kept as a floating parameter because it allowed better fitting to the data. This, however, also introduces error in the estimation, as molecular weight is calculated from the inverse of the intercept of the fitted curve. Furthermore, because the Debye plot also depends strongly on the dn/dc of the protein, error in dn/dc translates into the estimation of molecular

Table I. Values of the Parameters Obtained by Analysis of the Linear [Eq. (3)] and Nonlinear [Eq. (20)] Debye Plots of the Monoclonal Antibody and Ovalbumin for Various Solution Conditions

Solution condition	K_{dim} (M^{-1})	M_m (Da)	$B \times 10^4$ (mol mL/g^2)[a]
Monoclonal antibody			
pH 7.4, μ = 0.3 M	Monomer	142,000 (6000)[b]	−0.9 (0.06)
pH 5.4, μ = 0.3 M	Monomer	140,000 (9000)	−0.8 (0.1)
pH 4.0, μ = 0.04 M	9.03 (0.8) × 10^3	130,000 (8000)	−0.07 (0.01)
pH 4.0, μ = 0.3 M	7.76 (0.6) × 10^3	135,000 (6000)	−0.1 (0.04)
0.3 M glycine (pH 4.0, μ = 0.3 M)	4.23 (0.5) × 10^3	130,000 (7000)	−0.02 (0.01)
0.1% Tween 80 (pH 4.0, μ = 0.3 M)	3.19 (0.9) × 10^3	140,000 (5000)	−0.05 (0.01)
10% sucrose (pH 4.0, μ = 0.3 M)	7.54 (0.4) × 10^2	140,000 (7000)	−0.2 (0.08)
Ovalbumin			
pH 7.4, μ = 0.3 M	Monomer	43,000 (3000)	−1.01 (0.05)
pH 4.0, μ = 0.3 M	monomer	44,000 (2000)	−0.8 (0.07)
pH 4.0, μ = 0.3 M	Monomer	42,000 (2000)	−1.13 (0.07)
10% sucrose (pH 4.0, μ = 0.3 M)	Monomer	43,000 (2000)	−1.14 (0.03)
0.1% Tween 80 (pH 4.0, μ = 0.3 M)	Monomer	45,000 (2000)	−1.2 (0.06)

[a] The numbers in parentheses represent the standard deviation (n = 3).
[b] B represents the deviation from ideality. In the case of a linear plot where only monomer species is present, B is equal to B_{22}. In the case where the plot is nonlinear (presence of associating species), B represents overall nonideality.

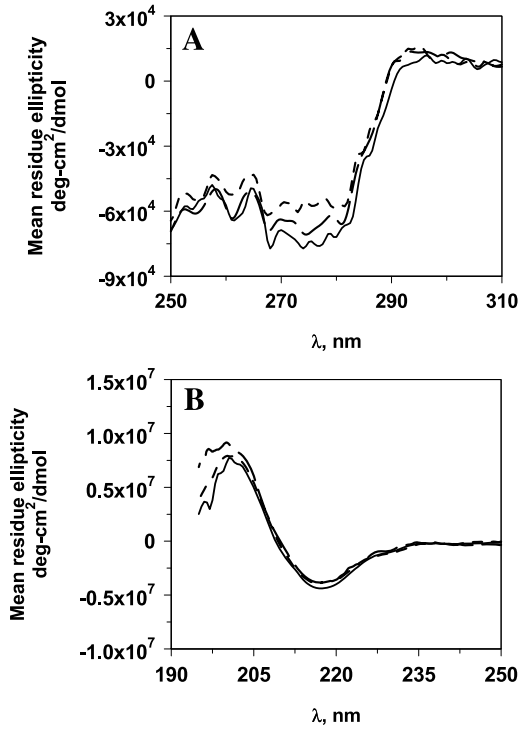


Fig. 4. Near-UV (A) and far-UV (B) circular dichroism (CD) spectra of the monoclonal antibody at pH 7.4 (solid line), pH 4.0 (small dash), and 10% w/v sucrose, pH 4.0 (large dash). The ionic strength of all solutions was adjusted to 0.3 M with NaCl.

weight. Despite this, we have observed that the calculated molecular weight is always obtained within ±10% of the true molecular weight. Overall, these studies showed that in solution conditions where the antibody aggregated upon storage, it exhibited self-association in the initial solutions, whereas linear Debye plots with a negative slope were obtained in solution conditions where no aggregation was observed.

CD Spectroscopy

To obtain better understanding of the mechanism of aggregation and self-association, we obtained the far-UV CD spectra (indicative of secondary structure conformation) and the near UV-CD spectra (indicative of tertiary structure conformation) of the antibody under various solution conditions at 37°C in the freshly prepared solutions. The shape of the spectra indicates a primarily β-sheet structure (minimum at 217 nm), typical of that present in immunoglobulin (46). No significant change is observed in the far-UV CD spectra of the antibody at pH 4.0, μ = 0.3 M, (aggregation observed), compared with that obtained at pH 7.4, μ = 0.3 M (no aggregation observed) (Fig. 4). Analysis of the secondary structure elements using CONTINLL software (41) estimated 49% total β-sheet content at pH 7.4 and 48% total β-sheet content at pH 4.0.

The near-UV CD spectra show significant differences in the structural conformation of the antibody, especially in the region of 268–285 nm, indicating changes in the environment surrounding primarily tyrosines as well as that of tryptophans. A more intense spectrum at pH 7.4 in this wavelength region compared with that at pH 4.0 indicates a more compact protein structure at the former pH in the vicinity of these amino acids. The fact that subtle changes in protein tertiary structure are observed at pH 4.0 compared with pH 7.4, whereas no significant changes are observed in the far UV region, indicates that the protein attains a partially unfolded structure at pH 4.0 (47,48). In other words, the changes in near-UV CD spectra are not a result of the presence of a small fraction of unfolded protein but are attributed to subtle changes in protein structure itself. In the presence of sucrose at pH 4.0 (μ = 0.3 M), a near-UV CD spectrum intermediary to that observed at pH 7.4 and 4.0 (no sucrose) is observed, indicating that sucrose shifts the equilibrium toward the native species albeit not to a complete extent. Note that DSC studies did not provide any additional information (data not shown) on the unfolding of the antibody because the onset of unfolding temperature was observed above 40°C at both pH. A thermodynamic analysis of the unfolding of antibody was not feasible because of its rapid aggregation at high temperatures (more than 60°C).

FTIR Spectroscopy

We next examined the secondary structure of the antibody in the aggregates compared with that present in the solution at pH 4.0, μ = 0.3 M. Figure 5 shows the area-normalized second derivative FTIR spectra of this protein at pH 4.0 (μ = 0.3 M) compared with those of the insoluble aggregates obtained after incubation at 37°C for a period of 3 months at pH 4.0 (μ = 0.3 M). At pH 4.0 in the initial solution, an intense peak is observed at 1637 cm^{-1}, attributed to the primarily β-sheet structure present in this antibody (41). The satellite peaks at 1688 and 1670 cm^{-1} relate to the β-sheets as well and the turns, respectively, present in the secondary structure of this antibody. Upon aggregation, a broad peak is observed with a loss in the intensity of the peak at 1637 cm^{-1} compared with that at pH 7.4, indicating loss in the specific secondary structure of the antibody. The peak position at 1637 cm^{-1}, however, indicates that the aggregates largely possess a β-sheet structure. These results show that the protein secondary structure is significantly perturbed upon aggregation, pointing to the fact that the aggregation process presumably involves the nonnative species of the antibody.

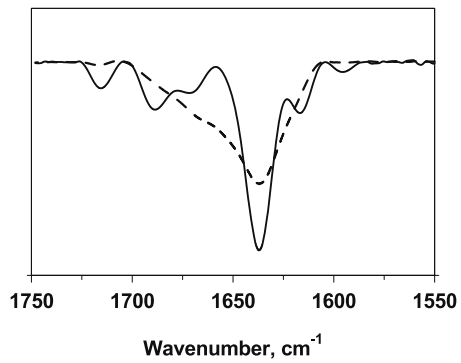


Fig. 5. Area-normalized second-derivative Fourier transform infrared spectra of the monoclonal antibody at pH 4.0 (μ = 0.3 M) in a solution at t = 0 (solid line) and upon aggregation at 37°C after storage for a period of 3 months (dashed line).

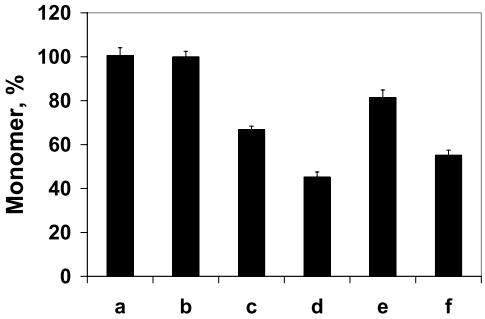


Fig. 6. Percent monomer of ovalbumin remaining after storage at 37°C for a period of 3 months as determined by SEC. a, pH 7.4 (μ = 0.3 M); b, pH 5.4 (μ = 0.3 M); c, pH 4.0 (μ = 0.04 M); d, pH 4.0 (μ = 0.3 M); e, 10% w/v sucrose (pH 4.0, μ = 0.3 M); f, 0.1% w/v Tween 80 (pH 4.0, μ = 0.3 M).

Ovalbumin

Aggregation Studies

Figure 6 shows the percent monomer remaining of ovalbumin following incubation at 37°C for a period of 3 months. Evidently, no aggregation is observed at pH 7.4 or 5.4, whereas significant aggregation of this protein was observed at pH 4.0 under all solution conditions. Similar to the results obtained with the monoclonal antibody, sucrose provided only partial protective effect at pH 4.0 against aggregation. Additionally, the extent of aggregation increased with an increase in the NaCl concentration, pointing to the fact that electrostatic repulsions are weakened at higher ionic strength. Visual inspection of the samples and SEC confirmed that ovalbumin also formed soluble as well as insoluble aggregates (data not shown).

Second Virial Coefficient Studies

Figure 7 shows the Debye plots of ovalbumin at conditions where no aggregation was observed (pH 7.4, μ = 0.3 M) and at conditions where aggregation was observed [pH 4.0, μ = 0.3 M; 0.3 M glycine (pH 4.0, μ = 0.3 M); 10% w/v sucrose (pH 4.0, μ = 0.3 M), respectively]. Linear Debye plots with similar negative slopes were obtained in all solutions (Table I). This was strikingly in contrast with the results obtained with the monoclonal antibody, where the Debye plots showed a different behavior in the solutions where aggregation was observed compared with those where no aggregation was observed. In the case of ovalbumin, the B_{22} values indicated attractive interactions in all solution conditions with no evidence of self-association. The attractive interactions did not lead to aggregation at pH 7.4 but caused significant aggregation at pH 4.0, although the magnitude of the attractions was similar, as shown by the magnitude of virial coefficients in these two solution conditions. Furthermore, ovalbumin did not exhibit self-association, as seen in the case of the monoclonal antibody.

CD Spectroscopy

To investigate the role of the structural conformation on ovalbumin aggregation, we examined the far-UV and near-UV CD spectra at pH 4.0 and 7.4. The far-UV CD spectrum of ovalbumin at pH 7.4 showed minima at 208 and 222 nm, typical of primarily alpha-helical structure present in this protein (Fig. 8). As the solution pH was lowered to 4.0, the intensity of the spectra decreased, indicating a loss in the secondary structure of ovalbumin. Analysis of the CD spectra using CONTINLL software indicated 35% alpha helical and 20% β-sheet content at pH 7.4 compared with 23% alpha helical and 16% β-sheet content at pH 4.0. Correspondingly, the near-UV CD spectrum at pH 4.0 showed significant changes in the wavelength region below 295 nm, indicating alterations in the tertiary structure of ovalbumin at this pH (Fig. 8). Furthermore, the changes in tertiary structure are far more drastic compared with those observed in the secondary structure. For example, at pH 7.4, a positive near-UV CD spectrum is observed, whereas, at pH 4.0, a primarily negative near-UV CD spectrum close to the zero value is observed. It is difficult to assess based on these observations whether these changes are caused by the presence of partially unfolded state of ovalbumin or the presence of a fraction of completely unfolded state or both. Nevertheless, it is evident that significant changes in the structure of ovalbumin take place at pH 4.0 compared with that at pH 7.4. The CD data indicate that the aggregation of ovalbumin at pH 4.0 is a result of structural alterations in the proteins that presumably expose the hydrophobic groups. Similar to that observed in the case of antibody, sucrose shifts the equilibrium toward the native form of ovalbumin, although not to a complete extent, which also correlates with the ability of sucrose to provide partial protection against aggregation at pH 4.0 (μ = 0.3 M).

DISCUSSION

Irreversible physical aggregation is a major impediment toward development of a successful aqueous protein formulation. Whereas several theories and models have been proposed to understand the mechanism of physical aggregation, successful prediction of the aggregation behavior remains a daunting task to the protein formulation scientist. Although the role of conformational stability is understood to play an important role toward aggregation, recently, a few reports have emerged on the role of colloidal stability as

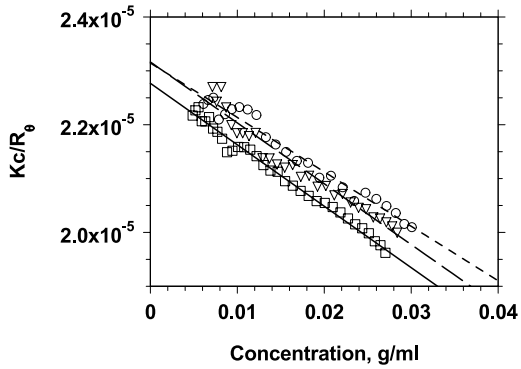


Fig. 7. Debye plots of ovalbumin at pH 7.4 (circles), pH 4.0, no sucrose (squares), and 10% w/v sucrose, pH 4.0 (triangles), at μ = 0.3 M. The lines are generated by linear regression of the data points using Eq. (3): pH 7.4 (small dash), pH 4.0, no sucrose (solid line), and 10% w/v sucrose, pH 4.0 (large dash).

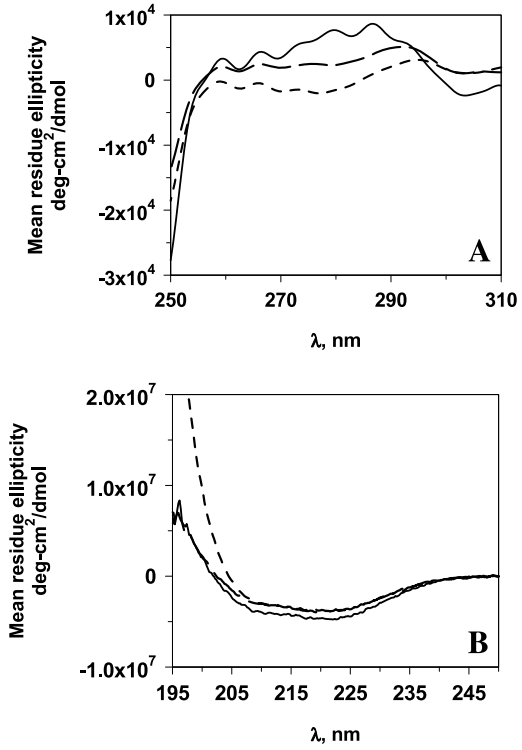


Fig. 8. Near-UV (A) and far-UV (B) CD spectra of ovalbumin at pH 7.4 (solid line), pH 4.0 (small dash), and 10% w/v sucrose, pH 4.0 (large dash). The ionic strength of all solutions was adjusted to 0.3 M with NaCl.

measured through the second virial coefficient toward causing aggregation, especially under solution conditions that primarily favor native conformation. A detailed study by Chi *et al.* on the balancing act of conformational stability *vs.* colloidal stability indicated that solution conditions govern as to which process becomes rate limiting. Thus, using human granulocyte growth-stimulating factor, these authors have demonstrated that under certain conditions where the $\Delta G_{\text{unfolding}}$ was comparable, a different extent of aggregation was observed, and this was attributed to the colloidal stability as the rate-limiting factor. However, under few conditions, neither colloidal stability nor conformational stability could relate to aggregation. This was attributed to interference from irreversible dimers, which prevented true estimation of the B_{22} value of the aggregation-prone monomeric species.

In the present study, we have used an improved method, recently developed by us, to measure the B_{22} values of proteins that separate the monomeric species from irreversible aggregates and other interfering cosolutes and that also provide quantitative characterization of self-association behavior of the protein under study. We studied aggregation of two proteins, a monoclonal antibody (a multidomain monomeric protein at physiological conditions) and ovalbumin (a globular monomeric protein), and investigated the relationship between the extent of aggregation and the B_{22} values, along with the spectroscopic studies, to understand the structural conformation. The goal of these studies was to investigate the role of colloidal stability toward aggregation. Aggregation, whether of the unfolded molecules or of the native molecules, is, in principle, a colloidal interaction and is determined by the net outcome of the competing electrostatic, van der Waals, and hydrophobic interactions [as explained by the extended Derjaguin-Landau-Verwey-Overbeek (DLVO) theory]. We hypothesized that B_{22}, a parameter that represents the net attractive or repulsive solute–solute interactions in the solution, could relate to the behavior of a given protein for its tendency to undergo irreversible aggregation.

Aggregation studies performed on the monoclonal antibody showed that this protein is resistant to physical aggregation at pH 7.4 and 5.4 but undergoes significant aggregation at pH 4.0 under all solution conditions (Fig. 1A). The p*I* of this antibody lies in the region of pH 7–9, indicating that the protein will have a net positive charge at pH 4.0. Hence, the additional charge at pH 4.0, in principle, should have provided enough electrostatic repulsion to prevent aggregation. Evidently, it is not so, indicating that this aggregation is not a result of the direct interaction between the native molecules. Additionally, the extent of aggregation is lower at $\mu = 0.04$ M compared with that at $\mu = 0.3$ M, indicating that the net positive charge does provide protection against aggregation in the absence of enough shielding by salt ions.

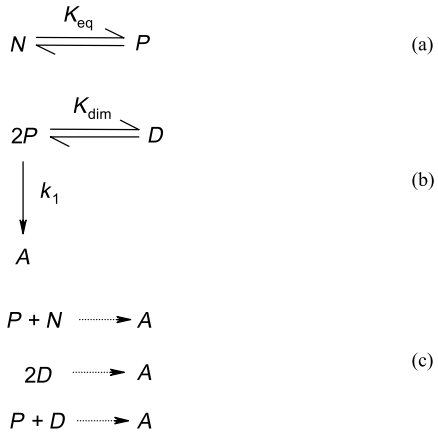
At pH 7.4, a linear Debye plot with a negative slope was obtained, indicating net attractive interactions (Fig. 2); however, these attractions were not sufficient to lead to aggregation because the protein was stable up to a period of 3 months at 37°C. Because the protein aggregated at pH 4.0 ($\mu = 0.3$ M), according to our hypothesis, a more negative B_{22} value was expected. However, a nonlinear Debye plot was obtained under these conditions. This shape of Debye plot results because of the self-association of a protein, pointing to the fact that the monoclonal antibody undergoes reversible self-association at pH 4.0. A qualitative analysis of the nonlinear Debye plot using a modified Debye equation (Appendix) yielded a K_{dim} value of 7.24×10^3 M^{-1} at pH 4.0 ($\mu = 0.3$ M). Because the monomer–dimer model fits well to the data, it is concluded that this antibody undergoes reversible monomer–dimer self-association. Using this value of K_{dim} and the equations for monomer and dimer concentration in conjunction with the total protein concentration (Appendix), we obtained that even at 20 mg/mL antibody concentration, the majority of the species present in the solution are monomeric (~75%).

The question arises: "Are the native species involved in the process of self association?" In general, if a protein self-associates in its native form at a pH away from p*I*, then it would also exhibit self-association or be present in a multimeric form at p*I*, as attractive forces are stronger at p*I* because of a lack of net charge on the protein. Thus, in the case of antibody, if the native species are involved in the self-association, then the association would also be expected at pH 7.4, where the net charge is close to zero, as compared with pH 4.0 (net positive charge), resulting in more attractive interactions. That neither aggregation nor association is observed at pH 7.4, whereas both phenomena occur at pH 4.0, and the protein is present as a monomer at pH 7.4, points to the fact that the native species of this antibody presumably are not involved in the self-association at pH 4.0.

Note that the K_{dim} values are similar at $\mu = 0.04$ M and $\mu = 0.3$ M (Table I), whereas significant difference in the extent of aggregation was observed at these two salt concentrations, indicating that electrostatic interactions do

not play a major role in the self-association. In addition, a lower K_{dim} value is observed in the presence of sucrose, and also, a lesser extent of aggregation is observed compared with that in its absence at pH 4.0 (μ = 0.3 M). Sucrose is known to increase the $\Delta G_{unfolding}$ and hence preserve the native state of proteins via its preferential exclusion effect. In the present case, an increase in the population of the native species at pH 4.0 would lead to a lesser extent of aggregation and a lower K_{dim} value, if the two phenomena indeed proceed via a nonnative state. Because that is indeed the case, it can be concluded that both protein aggregation and association proceed via a nonnative state of the antibody. The self-association of unfolded states of proteins has also been observed previously, e.g., in the case of bovine human growth hormone (49).

Far-UV and near-UV CD studies showed subtle changes in the tertiary structure of antibody with no significant differences in the secondary structure (Fig. 4), confirming that this antibody forms a partially folded state at pH 4.0, whose structure is close to the native form. These studies also confirm that sucrose indeed provides partial reversal to the native form of the antibody. However, even a small extent of unfolding could lead to the exposure of enough hydrophobic groups, making the protein aggregation-prone or association-prone. Note that the changes in the near-UV CD of antibody are attributed to the structural changes caused by the partial unfolding and not to the presence of dimer, because at the concentrations used for CD studies (0.5–2 mg/mL) and based on the K_{dim} values, the antibody will be primarily present as a monomer. Based on these studies, the following simplified reaction scheme is proposed for the aggregation and self-association of the antibody:

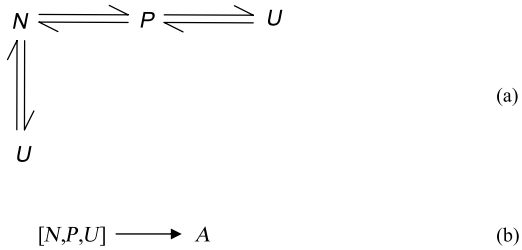


where N is the native species, P is the partially unfolded species, D is the dimer, A is the aggregated species, K_{eq} is the equilibrium constant of protein unfolding, K_{dim} is the protein association constant, and k_1 is the rate constant of aggregation. It should be noted that the unfolded species could also interact with the native species and the dimer to undergo aggregation. In addition, the dimer could by itself also participate toward forming irreversible aggregates (Scheme I, reaction c). Evidently, these studies and the above-proposed model point to the complexities that could be involved in the process of aggregation of this antibody.

The aggregation of ovalbumin showed similar dependence on the pH and ionic strength as that observed in the case of the antibody. Whereas no aggregation was observed at pH 7.4 and 5.4, ovalbumin readily aggregated at pH 4.0 in all solution conditions. Because the aggregation studies were carried out close to the pI of ovalbumin (pI ~4.5), note that the difference in the extent of aggregation between 0.3 M (45% monomer remaining) and 0.04 M (66% monomer remaining) is less as compared with that observed in the case of the antibody (32% monomer remaining at μ = 0.3 M and 89% monomer remaining at μ = 0.04 M). Thus, solution ionic strength plays an important role in the aggregation of these two molecules. Also, in the case of ovalbumin, sucrose provides a strong protective effect against aggregation at pH 4.0 (μ = 0.3 M), however, not to a complete extent (Fig. 6). This effect of sucrose is attributed to its ability to shift the equilibrium toward the native state of the protein.

Linear Debye plots of ovalbumin with similar negative slopes were obtained for all solution conditions studied irrespective of whether aggregation was observed or not (Fig. 7; Table I). Although the B_{22} values were negative, indicating net attractive interactions, these values did not explain why aggregation proceeded in one pH (pH 4.0) and not in the other (pH 7.4).

Spectroscopic studies using far-UV and near-UV CD showed loss in the secondary structure (a net 16% loss in the alpha helical and β-sheet content) as well as in the tertiary structure of ovalbumin (Fig. 8). These changes could be attributed to either the partial unfolding of the protein or the presence of fraction of unfolded species or both. Nevertheless, it is reasonable to believe that aggregation of ovalbumin proceeds through the ensemble of these unfolded species. Sucrose would shift the equilibrium toward the native state, and this is confirmed by CD studies, as the loss in secondary and tertiary structures is not to an extent as that observed in the absence of sucrose at pH 4.0 (μ = 0.3 M). Note that although significant changes in the protein structure were observed at pH 4.0 compared with that at pH 7.4, attributed to either a small fraction of unfolded molecules or to the partial unfolding of the protein, similar B_{22} values were obtained at these two pH. This indicates that the estimation of B_{22} values is not sensitive to the presence of a small fraction of unfolded molecules or subtle changes in the entire population of the molecules. Clearly, in the case of ovalbumin, spectroscopic studies provide a better indication of its tendency to undergo aggregation than the second virial coefficient values. Based on these observations, the aggregation scheme of ovalbumin could be written as:



where N represents the native state, P represents the partially unfolded state, U represents the unfolded state, and the term $[N,P,U]$ represents the reaction between various states to form the aggregated state A.

CONCLUSIONS

It is concluded from the present studies that the second virial coefficient, B_{22}, a parameter that represents protein–protein interaction in a solution, did not relate to the aggregation of proteins, a monoclonal antibody, and ovalbumin. Although our technique to determine the B_{22} values excludes contributions from aggregates or cosolutes, it cannot exclude contribution from unfolded or partially unfolded species, which is the aggregation-prone species in the case of antibody and ovalbumin as well as in several other aggregation cases reported in literature. Because even a small fraction of this species is sufficient to lead to protein aggregation, the kinetics of aggregation being governed by the concentration of the aggregation-prone species, the characterization of this species is more critical than the determination of the overall protein–protein interactions. B_{22} represents net interactions of all the species present in a solution and is often a weighted result of contributions from all species present in a solution. Hence, under the circumstances where the native protein is the predominant species, small changes in the population of the unfolded species will not affect the B_{22} value but can significantly affect the extent of long-term aggregation. Furthermore, in the case of partial unfolding of the protein, subtle changes in protein structure may not affect the net second virial coefficient values but could lead to long-term aggregation as a result of specific hydrophobic interactions. However, Debye plots (used for determination of B_{22} values) provide critical information about protein behavior in a solution, as seen in the case of the antibody where evidence of self-association was obtained under conditions leading to aggregation. Note that the ability of B_{22} to predict protein solubility stems from the fact that solubility involves native state of proteins and B_{22} values also present interactions of the native state. However, because physical irreversible aggregation rarely involves the native species in the initial step, it is highly unlikely that B_{22} will correlate to aggregation. Nevertheless, spectroscopic studies along with the Debye plots can provide important insights into the mechanism of physical irreversible aggregation.

APPENDIX

Model for Self-Association and Data Analysis of Nonlinear Debye Plots

A monomer–dimer equilibrium is written as

$$M + M \overset{K_{\text{dim}}}{\Leftrightarrow} D \tag{13}$$

where the association constant K_{dim} is defined as

$$K_{\text{dim}} = \frac{[c_{\text{d}}]}{[c_{\text{m}}]^2} \tag{14}$$

where $[c_{\text{d}}]$ is the molar concentration of the dimer and $[c_{\text{m}}]$ is the molar concentration of the monomer. The total molar concentration $[c_{\text{t}}]$ of the protein can be written in terms of the monomer concentration as

$$[c_{\text{t}}] = [c_{\text{m}}] + 2[c_{\text{d}}] \tag{15}$$

Combining Eqs. (13) and (14), solving the resulting quadratic equation for positive solution of $[c_{\text{m}}]$ and $[c_{\text{d}}]$, and converting molar concentration to grams per milliliter, the monomer and dimer concentrations can be written as

$$c_{\text{monomer}} = \frac{-1 + (1 + 8000K_{\text{dim}}c_{\text{t}}/M_{\text{m}})^{1/2}}{4000K_{\text{dim}}/M_{\text{m}}} \tag{16}$$

$$c_{\text{dimer}} = \frac{1 + 4000K_{\text{dim}}c_{\text{t}}/M_{\text{m}} - (1 + 8000K_{\text{dim}}c_{\text{t}}/M_{\text{m}})^{1/2}}{4000K_{\text{dim}}/M_{\text{m}}} \tag{17}$$

For an associating system, the Debye equation is written as

$$\frac{Kc_{\text{t}}}{R_{90}} = \left(\frac{1}{M_{\text{av}}} + Bc_t\right) \tag{18}$$

where M_{av} is the weight average molecular weight of all the species present in the solution. Note that B_{22} has been substituted with the term B to represent the nonideality arising from monomer–monomer, monomer–dimer, and dimer–dimer interactions. For an associating system, the change in the chemical potential of the solvent with solute concentration is written as (50)

$$\frac{\partial \mu_1}{\partial c_{\text{t}}} = \frac{\partial c_{\text{m}}}{\partial c_{\text{t}} \cdot M_{\text{m}}} + \frac{\partial c_{\text{d}}}{\partial c_{\text{t}} \cdot 2M_{\text{m}}} + \frac{\partial (Bc_{\text{t}}^2)}{\partial c_{\text{t}}} \tag{19}$$

Substituting for c_{m} and c_{d} from Eqs. (16) and (17), taking partial derivatives, and using the result in the derivation of the Rayleigh's light scattering equation, the following Debye equation is obtained:

$$\frac{Kc_{\text{t}}}{R_{90}} = \frac{(1 + 8000K_{\text{dim}}c_{\text{t}}/M_{\text{m}})^{1/2} + 1}{2(1 + 8000K_{\text{dim}}c_{\text{t}}/M_{\text{m}})^{1/2}M_{\text{m}}} + Bc_{\text{t}} \tag{20}$$

Equation (20) was used to fit the curved Debye plots for the parameters K_{dim}, M_{m}, and B.

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support and donation of the monoclonal antibody from Pfizer Biologics (St. Louis, MO, USA) and financial support from the National Science Foundation Industry/University Cooperative Research Center for Pharmaceutical Processing (http://www.ipph.purdue.edu/~nsf/aboutCPPR.html).

REFERENCES

1. J. L. Cleland, M. F. Powell, and S. J. Shire. The development of stable protein formulations: a close look at protein aggregation, deamidation, and oxidation. *Crit. Rev. Ther. Drug Carr. Syst.* **10**:307–377 (1993).
2. J. Brange. Physical stability of proteins. In S. Frokhaer and L. Hovgaard (eds.), *Pharmaceutical Formulation Development of Peptides and Proteins*, Taylor & Francis, London, 2000, pp. 89–112.
3. H. Schellekens. Factors influencing the immunogenicity of therapeutic proteins. *Nephrol. Dial. Transplant.* **20**:3–9 (2005).

4. C. J. Roberts. Kinetics of irreversible protein aggregation: analysis of extended Lumry-Eyring models and implications for predicting protein shelf life. *J. Phys. Chem. B.* **107**:1194–1207 (2003).
5. M. Sukumar, S. M. Storms, and M. R. De Felippis. Non-native intermediate conformational states of human growth hormone in the presence of organic solvents. *Pharm. Res.* **22**:789–796 (2005).
6. R. Wetzel. Mutations and off-pathway aggregation of proteins. *Trends Biotechnol.* **12**:193–198 (1994).
7. S. Krishnan, E. Y. Chi, J. N. Webb, B. S. Chang, D. Shan, M. Goldenberg, M. C. Manning, T. W. Randolph, and J. F. Carpenter. Aggregation of granulocyte colony stimulating factor under physiological conditions: characterization and thermodynamic inhibition. *Biochemistry* **41**:6422–6431 (2002).
8. B. S. Kendrick, J. F. Carpenter, J. L. Cleland, and T. W. Randolph. A transient expansion of the native state precedes aggregation of recombinant human interferon-γ. *Proc. Natl. Acad. Sci. USA* **95**:14142–14146 (1998).
9. R. Lumry and H. Eyring. Conformation changes of proteins. *J. Phys. Chem.* **58**:110–120 (1954).
10. S. W. Raso, J. Abel, J. M. Barnes, K. M. Maloney, G. Pipes, M. J. Treuheit, J. King, and D. N. Brems. Aggregation of granulocyte-colony stimulating factor *in vitro* involves a conformationally altered monomeric state. *Protein Sci.* **14**:2246–2257 (2005).
11. J. L. Cleland and D. I. C. Wang. Cosolvent effects on refolding and aggregation. *A.C.S. Symp. Ser.* **516**:151–166 (1993).
12. A. Kjellsson, I. Sethson, and B.-H. Jonsson. Hydrogen exchange in a large 29 kD protein and characterization of molten globule aggregation by NMR. *Biochemistry* **42**:363–374 (2003).
13. K. Boren, H. Grankvist, P. Hammarstrom, and U. Carlsson. Reshaping the folding energy landscape by chloride salt: impact on molten-globule formation and aggregation behavior of carbonic anhydrase. *FEBS Lett.* **566**:95–99 (2004).
14. L. A. Kueltzo and C. R. Middaugh. Structural characterization of bovine granulocyte colony stimulating factor: effect of temperature and pH. *J. Pharm. Sci.* **92**:1793–1804 (2003).
15. V. K. Sharma and D. S. Kalonia. Temperature- and pH-induced multiple partially unfolded states of recombinant human interferon-α2a: possible implications in protein stability. *Pharm. Res.* **20**:1721–1729 (2003).
16. D. Bulone, V. Martorana, and P. L. San Biagio. Effects of intermediates on aggregation of native bovine serum albumin. *Biophys. Chem.* **91**:61–69 (2001).
17. M. Karlsson, L.-G. Martensson, P. Olofsson, and U. Carlsson. Circumnavigating misfolding traps in the energy landscape through protein engineering: suppression of molten globule and aggregation in carbonic anhydrase. *Biochemistry* **43**:6803–6807 (2004).
18. V. N. Uversky, J. Li, and A. L. Fink. Evidence for a partially folded intermediate in alpha-synuclein fibril formation. *J. Biol. Chem.* **276**:10737–10744 (2001).
19. S. Yoshioka, Y. Aso, K.-I. Izutsu, and S. Kojima. Is stability prediction possible for protein drugs? Denaturation kinetics of β-galactosidase in solution. *Pharm. Res.* **11**:1721–1725 (1994).
20. C. J. Roberts, R. T. Darrington, and M. B. Whitley. Irreversible aggregation of recombinant bovine granulocyte-colony stimulating factor (BG-CSF) and implications for predicting protein shelf life. *J. Pharm. Sci.* **92**:1095–1111 (2003).
21. L. R. De Young, A. L. Fink, and K. A. Dill. Aggregation of globular proteins. *Acc. Chem. Res.* **26**:614 – 620 (1993).
22. W. Liu, T. Cellmer, D. Keerl, J. M. Prausnitz, and H. W. Blanch. Interactions of lysozyme in guanidinium chloride solutions from static and dynamic light-scattering measurements. *Biotechnol. Bioeng.* **90**:482 – 490 (2005).
23. R. N. Keener and R. A. Hart. Impact of conformational and colloidal stability on ultrafiltration–diafiltration. Abstracts of Papers, 229th ACS National Meeting, San Diego, CA, USA, March 13–17, 2005. BIOT-271 (2005).
24. J. Zhang and X. Y. Liu. Effect of protein–protein interactions on protein aggregation kinetics. *J. Chem. Phys.* **119**:10972–10976 (2003).
25. E. Y. Chi, S. Krishnan, B. S. Kendrick, B. S. Chang, J. F. Carpenter, and T. W. Randolph. Roles of conformational stability and colloidal stability in the aggregation of recombinant human granulocyte colony-stimulating factor. *Protein Sci.* **12**: 903–913 (2003).
26. J. G. S. Ho, A. P. J. Middelberg, P. Ramage, and H. P. Kocher. The likelihood of aggregation during protein renaturation can be assessed using the second virial coefficient. *Protein Sci.* **12**: 708–716 (2003).
27. J. G. S. Ho and A. P. J. Middelberg. Aggregation during refolding is dependent on the second virial coefficient. Abstracts of Papers, 225th ACS National Meeting, New Orleans, LA, United States, March 23–27, 2003. BIOT-035 (2003).
28. J. G. S. Ho, A. P. J. Middelberg, P. Ramage, and H. P. Kocher. The likelihood of aggregation during protein renaturation can be assessed using the second virial coefficient. *Protein Sci.* **12**: 708–716 (2003).
29. S.-L. Huang, F.-Y. Lin, and C.-P. Yang. Microcalorimetric studies of the effects on the interactions of human recombinant interferon-α2a. *Eur. J. Pharm. Sci.* **24**:545–552 (2005).
30. B. H. Zimm. Application of the methods of molecular distribution to solutions of large molecules. *J. Chem. Phys.* **14**:164–179 (1946).
31. W. G. McMillan Jr. and J. E. Mayer. The statistical thermodynamics of multicomponent systems. *J. Chem. Phys.* **13**:276–305 (1945).
32. B. L Neal, D. Asthagiri, and A. M. Lenhoff. Molecular origins of osmotic second virial coefficients of proteins. *Biophys. J.* **75**:2469–2477 (1998).
33. T. L. Hill. Adsorption on proteins and interactions between protein molecules in solution. *J. Chem. Phys.* **21**:1395–1396 (1953).
34. G. Scatchard, A. Gee, and J. Weeks. Physical chemistry of protein solutions. Vi. The osmotic pressures of mixtures of human serum albumin and γ-globulins in aqueous sodium chloride. *J. Phys. Chem.* **58**:783–787 (1954).
35. S. Lapanje and C. Tanford. Proteins as random coils. IV. Osmotic pressures, second virial coefficients, and unperturbed dimensions in 6 M guanidine hydrochloride. *J. Am. Chem. Soc.* **89**:5030–5033 (1967).
36. A. George, Y. Chiang, B. Guo, A. Arabshahi, Z. Cai, and W. W. Wilson. Second virial coefficient as predictor in protein crystal growth. *Methods Enzymol.* **276**:100–110 (1997).
37. C. Haas, J. Drenth, and W. W. Wilson. Relation between the solubility of proteins in aqueous solutions and the second virial coefficient of the solution. *J. Phys. Chem. B* **103**:2808–2811 (1999).
38. B. Guo, S. Kao, H. McDonald, A. Asanov, L. L. Combs, and W. W. Wilson. Correlation of second virial coefficients and solubilities useful in protein crystal growth. *J. Cryst. Growth* **196**: 424 – 433 (1999).
39. E. Chi, Y. S. Krishnan, B. S. Kendrick, B. S. Chang, J. F. Carpenter, and T. W. Randolph. Roles of conformational stability and colloidal stability in the aggregation of recombinant human granulocyte colony-stimulating factor. *Protein Sci.* **12**:903–913 (2003).
40. H. Bajaj, V. K. Sharma, and D. S. Kalonia. Determination of second virial coefficient of proteins using a dual-detector cell for simultaneous measurement of scattered light intensity and concentration in SEC-HPLC. *Biophys. J.* **87**:4048–4055 (2004).
41. N. Sreerama. *CDPro: A Software Package for Analyzing Protein Spectra*. http://lamar.colostate.edu/~sreeram/CDpro/index.html.
42. J. D. Andya, C. C. Hsu, and S. J. Shire. Mechanisms of aggregate formation and carbohydrate excipient stabilization of lyophilized humanized monoclonal antibody formulations. *AAPS PharmSci.* **5** (2003).
43. A. Dong, S. J. Prestrelski, S. D. Allison, and J. F. Carpenter. Infrared spectroscopic studies of lyophilization- and temperature-induced protein aggregation. *J. Pharm. Sci.* **84**:415 – 424 (1995).
44. P. Doty, M. Gellert, and B. Rabinovitch. The association of insulin. I. Preliminary investigations. *J. Am. Chem. Soc.* **74**:2065 –2069 (1952).
45. R. Townend and S. N. Timasheff. Molecular interactions in β-lactoglobulin. III. Light-scattering investigation of the stoichiometry of the association between pH 3.7 and 5.2. *J. Am. Chem. Soc.* **82**:3168 – 3174 (1960).

46. A. W. P. Vermeer and W. Norde. The thermal stability of immunoglobulin: unfolding and aggregation of a multi-domain protein. *Biophys. J.* **78**:394–404 (2000).
47. O. B. Ptitsyn. Molten globule and protein folding. *Adv. Protein Chem.* **47**:83–229 (1995).
48. A. L. Fink. Compact intermediate states in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **24**:495–522 (1995).
49. D. N. Brems, S. M. Plaisted, E. W. Kauffman, and H. A. Havel. Characterization of an associated equilibrium folding intermediate of bovine growth hormone. *Biochemistry* **25**:6539–6543 (1986).
50. C. Tanford. Thermodynamics. In C. Tanford (ed.), *Physical Chemistry of Macromolecules*, Wiley, New York, 1961, pp. 202–204.