Pharmaceutical Research, Vol. 24, No. 11, November 2007 (© 2007)
DOI: 10.1007/s11095-007-9345-x

Research Paper

A High-Throughput Method for Detection of Protein Self-Association and Second Virial Coefficient Using Size-Exclusion Chromatography Through Simultaneous Measurement of Concentration and Scattered Light Intensity

Harminder Bajaj,[1,3] Vikas K. Sharma,[1,2] and Devendra S. Kalonia[1,4]

Received February 20, 2007; accepted May 11, 2007; published online June 19, 2007

Purpose. To characterize protein self-association along with second virial coefficient (a measure of solution nonideality) using size-exclusion chromatography (SEC) utilizing a novel flow cell that is capable of simultaneously measuring protein concentration and scattered light intensity.
Methods. β-lactoglobulin A (βLg), known to exhibit NaCl-dependent monomer–dimer equilibrium at pH 3.0, was used as the model protein. A range of concentrations and corresponding scattered light intensities, obtained in the eluting peak from a single protein injection, in different solution conditions, were used to generate the Debye plots (Kc/R_θ vs c). The Debye light scattering equation was modified to include the monomer–dimer equilibrium model and the second virial coefficient to analyze the data obtained.
Results. Debye plots of βLg, while linear at pH 2.3, 0 M NaCl (pure monomer) and at pH 3.0, 1 M NaCl (pure dimer), showed curvature at pH 3.0, for varying NaCl concentrations (0.02–0.5 M). The curvature was indicative of the association behavior of this protein. The modified Debye light scattering equation, when fit onto the nonlinear Debye plots, yielded apparent K_a values ranging from 10^2 to 10^5 M^{-1} under various solution conditions. The apparent K_a values obtained from this method followed similar trend to those reported in literature.
Conclusions. SEC combined with simultaneous detection of scattered light intensity and concentration provides a rapid means of detection of protein self-association. The short duration of sample detection and analysis combined with SEC makes this method a useful tool for high-throughput characterization of protein association during early stages of protein formulation.

KEY WORDS: light-scattering (static); proteins; self-association; size exclusion chromatography; virial coefficient.

INTRODUCTION

The phenomenon of protein self-association is essential for several physiological and biochemical processes at the cellular and molecular level (1–4). Understanding protein self-association is crucial as often the monomeric and the oligomeric states have different biological and physicochemical properties and the interactions involved in self-association can be easily perturbed by solution properties, i.e. pH (5,6) ionic strength, temperature (7) and co-solutes (8,9). Recently, additional interest has been generated on studying protein self-association at relatively high concentration, for example, in the presence of crowding agents as sugars and nonionic polymers, to understand the effect of molecular crowding on proteins in cells (10,11). Similarly, reversible protein self-association is a critical issue in development of high concentration protein solutions, especially of antibodies for therapeutic use (12,13). From a protein formulation point of view, it is desirable to characterize behavior of a protein molecule for its tendency to undergo self-association early on in the development phase and often under several different solution conditions for rational formulation design. The fast emergence of proteomics based pharmaceutical biotechnology industry, hence, demands that protein association behavior be characterized rapidly, possibly utilizing a high-throughput based method.

Analytical equilibrium sedimentation (14–19) remains the method of choice to quantitatively study protein self-association. Although AUC is presumably the most accurate method to characterize self-association (and straightforward for simple monomer–dimer systems), because of the somewhat longer times and often complex analysis, it is seldom

[1] Department of Pharmaceutical Sciences, University of Connecticut, Storrs, Connecticut 06269, USA.
[2] *Present Address*: Genentech, Inc., 1 DNA Way, South San Francisco, California 94030, USA.
[3] *Present Address*: Bayer Biologicals, 800 Dwight Way, Berkeley, California 94710, USA.
[4] To whom correspondence should be addressed. (e-mail: kalonia@uconn.edu)

used as a tool for preformulation analysis and precludes its use as a high-throughput method. Recently, chromatography-based new techniques have emerged for characterization of protein association. These include self-interaction chromatography (20) and frontal gel chromatography (21,22). Although these techniques offer novel methods of evaluating protein self-association, they are a few concerns. For example, SIC requires prior immobilization of the protein on the column, which is often not desirable. Similarly, the determination of equilibrium constants in frontal gel chromatography is based on indirect measurement of molecular weights, which further depends on the magnitude of elution volume in a chromatogram (22). Another method based on light scattering has been proposed recently, which utilizes dual syringe pump method for rapid characterization of protein association (23). However, caution must be exercised in preparing solutions for use of light scattering technique in aqueous solutions. For example, a pure sample is required that must be free from other scattering impurities. Furthermore, in this particular method, the tubing volume from the two pumps must be exactly same (no time lag) for accurate analysis.

One of the earlier methods to characterize protein self-association involved batch-mode static light scattering (24,25). However, this method did not gain enough popularity due to the problems encountered in conducting light scattering experiments in aqueous solutions and due to the evolution of the sedimentation equilibrium methodology. We have recently developed a chromatography-based method to determine second virial coefficient of proteins (26) 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 was used recently to investigate the relationship of second virial coefficient with aggregation of a monoclonal antibody (27).

The method utilizes a custom-designed dual-source dual-detector cell that measures the intensity of the scattered light through a 90° light scattering detector and concentration through a UV detector at the same time in conjunction with size exclusion chromatography (SEC) (Appendix III). Because of the use of a SEC column, the irreversible aggregates and other interfering solutes do not contribute to the solute peak of interest and 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 is then used to generate Debye plot using the Debye's light scattering equation,

$$\frac{Kc}{R_\theta} = \frac{1}{M} + 2B_{22}c \tag{1}$$

where, R_θ is the excess Rayleigh's ratio of the protein in solution of concentration c, M is the weight average molecular weight of the protein and B_{22} is the second virial coefficient. K is the optical constant and is defined as,

$$K = \frac{4\pi^2 n^2 (dn/dc)^2}{N_A \lambda_o^4} \tag{2}$$

where, n is the solvent refractive index, dn/dc is the refractive index increment, λ is the wavelength of the incident light and N_A is the Avogadro's number.

It should be noted that the accuracy and precision of the Debye equation to characterize protein behavior in solution is based on the fact that the diameter of most proteins is smaller than $\lambda/20$, thus making these proteins behave like isotropic point-scatterers. For example, using a wavelength of 800 nm, one can conveniently use a single angle to study proteins of the order of ~40 nm in diameter corresponding to a trimer of an antibody (~12 nm diameter for a single monomeric unit) or even an aggregate of the order of 20-mer of a smaller protein as β-lactoglobulin (~2 nm diameter for a single monomeric unit). Under these circumstances, scattering is independent of the angle of scattering and observation at a single angle, mostly 90° is accurate for the data analysis. Thus, measurements at multiple angles to extrapolate data at a zero angle may not be necessary for most proteins of interest. Hence the statement that the data obtained by the dual-source dual-detector cell using single or two angles is less precise and questionable is irrelevant as long as the conditions for the isotropic point-scatterers are being met (28). Nevertheless, the concept of dual-detector, in principle, can be easily extended to multiple detectors (involving multiple angles) for high molecular weight species.

In this report, we have investigated the utilization of SEC in conjunction with the dual-source dual-detector cell to study self-association of a model protein, β-lactoglobulin (βLg). We show that the SEC-based method recently developed by us to obtain second virial coefficient of proteins could be utilized to simultaneously characterize nonideality and self-association of a model protein, βLg and apparent association constants can be obtained under different solution conditions. The term "apparent" is used in the present study to define the association constant for a nonequilibrium reversible self-association compared to that for a "true" equilibrium.

βLg (29–32) has been characterized extensively for its self-association behavior and is routinely used as calibration standard in sedimentation equilibrium studies. βLg exhibits salt-dependent monomer–dimer equilibrium at acidic pH. At low ionic strength and low pH (pH 2.3), βLg exists as pure monomer (M_w=18.4 KDa), whereas at high ionic strengths (~1 M) and moderately high pH (pH 3.0), it exists primarily as a dimer (29). At intermediate solution conditions βLg exhibits various levels of monomer–dimer equilibrium depending on pH and ionic strength. The association constants of βLg self-association under various solution conditions are widely reported in literature.

MATERIALS AND METHODS

All buffer components and chemical reagents used in the present studies were of highest purity grade, obtained from commercial sources, and used without further purification. Bovine β-lactoglobulin A (L7880, purity 99% by polyacrylamide electrophoresis as per manufacturer's certificate of analysis) was obtained from Sigma Chemicals Co. (St. Louis, MO). Double distilled water that was filtered through a 0.1 μm polycarbonate membrane filter was used for preparation of the mobile phase and protein solutions. A 20 mM Glycine–HCl buffer was used for all studies (prepared by dissolving glycine in double distilled water and adjusting the pH to 3.0 by using 1.0 N HCl). The final pH of all solutions was measured

using a Piccoloplus Hi-1295 digital pH meter (Fisher Scientific, Pittsburgh, PA).

Solution Conditions

All measurements were carried out at 20°C. The association behavior of βLg was studied at pH 2.3 using a 20 mM Glycine–HCl buffer (No NaCl) and at pH 3.0 (similar buffer) in the presence of NaCl (0.02–1 M). The mobile phase buffer was used to prepare protein solutions. The concentrations of the final protein solutions were determined using an $E_{1\%}$ of 9.3 cm^{-1} for βLg.

Chromatography

The Debye plots of the proteins using the dual-detector cell in conjunction with SEC were generated using a Precision Detectors's PD 2000 (Northampton, MA) detection system modified according to the method described previously (26). 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 except when the effect of flow rate was studied, where a flow rate of 0.3 ml/min was used. Briefly, 150 µl of the protein solution of a known concentration (15 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. A 100 mW laser at 685 nm was used as the incident light for light scattering. 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 (200 Å pore size, 5 µm bead size and 30×0.8 cm column dimensions) from YMC (Kyoto, Japan) was used.

Data analysis

Data analysis was carried out as described earlier (26). 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 on the chromatogram is a result of a moving-average of the signal collected for duration of 0.5 s. Note that the data points near the baseline with a signal-to-noise ratio of less than 10 were excluded from the data analysis. Similarly, the data points near the top 5% of the peak were also truncated, since these data points produced artifacts in the Debye plot because of the way the instrument collects and averages signals near the peak to obtain each data point. Each data point is then converted to Rayleigh's ratio, R_θ (light scattering detector) and concentration (UV detector) as described before and summarized in Appendix I (26). For each solution condition, dn/dc values were obtained according to the procedure described in Appendix I.

In the present studies, linear Debye plots were analyzed according to Eq. 1. For Debye plots obtained otherwise, the Debye equation was modified as described in the Results and Discussion section.

RESULTS AND DISCUSSION

Debye Plots of β-lactoglobulin

Figure 1A shows the UV chromatograms and Fig. 1B shows the Debye plots of βLg at pH 2.3 (No NaCl) and at pH 3.0 (1.0 M NaCl), respectively. At pH 2.3, the main peak in the SEC chromatograms corresponds to the average monomer molecular weight of ~18,000 Da (calculated over the whole main peak, using light scattering and concentration data, Fig. 1A) and the small peak eluting before the monomer peak corresponds to βLg dimer (~2.0%). A broader monomer peak is obtained due to the absence of salt and enhanced protein–column interactions. Under these solution conditions, a linear Debye plot is obtained with a positive slope corresponding to the positive virial coefficient value (B_{22}=8.5×10^{-4}). As expected, the intercept corresponds to the inverse of the molecular weight of the monomer of βLg (Table I). The positive virial coefficient is indicative of net

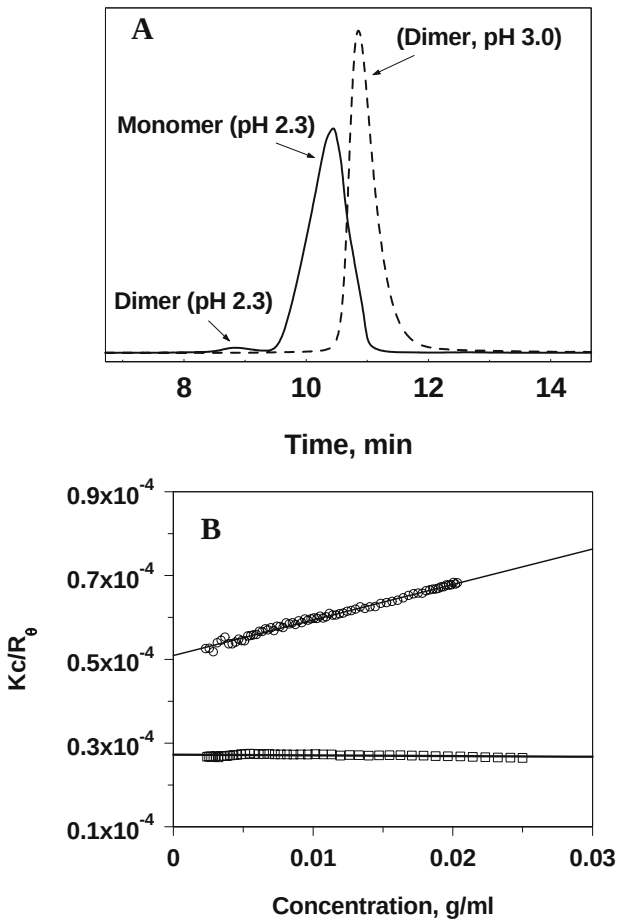

Fig. 1. **A** UV Chromatograms of β-lactoglobulin A following elution from an SEC column at pH 2.3 , 0 M NaCl (*solid line*) and at pH 3.0, 1.0 M NaCl (*dash line*). **B** Debye plots of β-lactoglobulin A at pH 2.3, 0 M NaCl (*empty circles*) and at pH 3.0, 1.0 M NaCl (*empty squares*). The lines are generated through linear regression of the data points using Equation 1.

Table I. Values of the Parameters Obtained by Analysis of the Linear Debye Plots (Eq. 1) and Nonlinear Debye Plots (Eq. 14) of β-lactoglobulin A for Various Solution Conditions

Solution Condition (20°C)	*dn/dc* (ml/g)	Molecular Weight (Da)	K_a (M^{-1})	$B\times10^4$ (mol ml/g^2)[f]
No NaCl, pH 2.3	0.162	18250 (±450)	–[a]	8.5 (±0.5)
0.02 M NaCl, pH 3.0	0.156	–[c]	1.44 (±0.50, 1.05) $\times10^2$	1.7 (±0.2, 0.3)
		19,100 (±105, 500)[d]	1.14 (±0.20, 1.00) $\times10^2$	–[e]
0.05 M NaCl, pH 3.0	0.155	–	1.12 (±0.10, 0.32) $\times10^3$	0.8 (±0.1, 0.1)
		19,900 (±150, 320)	8.48 (±0.08, 1.50) $\times10^2$	–
0.1 M NaCl, pH 3.0	0.152	–	3.90 (±0.10, 0.63) $\times10^3$	−0.3 (±0.1, 0.1)
		18,000 (±95, 430)	4.10 (±0.06, 0.52) $\times10^3$	–
0.2 M NaCl, pH 3.0	0.150	–	1.23 (±0.10, 0.13) $\times10^4$	−0.3 (±0.1, 0.1)
		17,500 (±230, 450)	1.70 (±0.05, 0.20) $\times10^4$	–
0.5 M NaCl, pH 3.0	0.150	–	1.08 (±0.18, 0.32) $\times10^5$	−0.1 (±0.1, 0.1)
		18,300 (±350, 500)	1.08 (±0.07, 0.30) $\times10^5$	–
1.0 M NaCl, pH 3.0	0.145	37,400 (±940)	–[b]	−0.6 (±0.1, 0.1)

[a] Primarily monomer present
[b] Primarily dimer present; analysis was done using Eq. 1. All other solution conditions were analyzed using Eq. 14.
[c] For each solution condition, the upper row represents the values of parameters obtained by keeping the molecular weight fixed at 18,400 Da, and the lower row represents the values of the parameters by floating the molecular weight.
[d] The first number in the parenthesis represents the maximum statistical uncertainty (goodness-of-fit) observed in the fitted value for a single experiment (1 SD) within 95% confidence interval and the second number represents the standard deviation in the estimation of the parameter from experiments in triplicate. In the case of pure monomer and dimer, the number in parenthesis is simply the standard deviation from experiments in triplicate.
[e] The *B* values obtained by fitting the data using Eq. 14 with floating molecular weight were not significantly different than those obtained with fixed molecular weight
[f] *B* represents the deviation from ideality. In case of a linear plot, where only monomer species is present, *B* is equal to B_{22}. In case where the plot is nonlinear (presence of associating species), *B* represents overall nonideality.

repulsive interactions between the protein molecules and hence relates to the tendency of this protein to remain in a monomeric state under these solution conditions.

At pH 3.0, 1.0 M NaCl, a single peak corresponding to an average molecular weight of ~36,000 Da is observed (Fig. 1A). It should be noted that, the pure dimer actually elutes later (10.85 min) as compared to the monomer (10.4 min). This is attributed to the effect of solution conditions on the retention time of these species as there is no NaCl present at pH 2.3, whereas, 1 M NaCl is used at pH 3.0. At pH 3.0, a linear plot with a negative slope is observed ($B_{22}=-6\times10^{-6}$). The molecular weight obtained from the inverse of intercept corresponds to that of the βLg dimer (Table I) a near zero (small negative) slope (Fig. 1B) suggests weakly attractive protein–protein interactions. The data presented above also emphasizes the fact that SEC alone may not be sufficient to characterize proteins for molecular weight determination as the retention times could be dramatically affected by solution conditions.

Figure 2A shows the Debye plot of βLg at pH 3.0, 0.1 M NaCl (expanded in Fig. 2B) compared to that at pH 2.3, 0.0 M NaCl and pH 3.0, 1.0 M NaCl. A distinct upward curvature is observed in the Debye plot in the presence of 0.1 M NaCl at pH 3.0. The slope of the Debye plot decreases with an increase in the concentration of βLg. It is reported in literature that βLg exhibits monomer–dimer association equilibrium under these solution conditions (29). The curvature observed in this solution condition is typical of self-associating systems, and has been previously reported in batch-mode light scattering as well as in sedimentation equilibrium studies (33). This form of curvature has also been reported in osmotic pressure studies of a mixture of proteins; for example, serum albumin and γ-globulin, where no self-association behavior exists (34). In this instance, the curvature arises because of a change in the average virial coefficient of the system (usually written as a z-average), for example as a function of NaCl concentration, due to different extent of effect of NaCl on the virial coefficient of each protein. However, since dimerization of a protein does not change the charge-to-mass ratio of the protein, the contribution of nonideality term to the curvature is generally assumed to be minimal and only shows up at high enough concentrations where no further self-association is observed. Hence, the change in slope, as the concentration increases, is due to an increase in the weight-average molecular weight of the protein because of dimerization (a faster rate of increase at lower concentrations than at higher concentration) (33).

Figure 2C shows the Debye plots for βLg at pH 3.0 at 0.05, 0.1 and 0.2 M NaCl. Similar to the Debye plot at 0.1 M NaCl, curved Debye plots are also observed at 0.05 and 0.2 M NaCl pointing to the fact that βLg exhibits monomer–dimer association equilibrium under these conditions as well. However, the shape and the position of the Debye plots are distinctly different for the different NaCl concentrations studies, which probably relates to the different association constants under these different solution conditions.

Figure 3A shows the chromatograms obtained for βLg at pH 3.0 (0.02 M NaCl) following three separate injections and Fig. 3B shows the resulting Debye plots. These figures illustrate the reproducibility of the data and the ability of the technique to minimize scatter in the analysis of the data. Note that the small peak observed at 9.5 min is due to the presence of irreversible soluble aggregate in the sample and was also observed in solution condition where βLg is present as a pure monomer (pH 2.3, 0 M NaCl). To further assess that the dimer peak comprises of irreversible aggregates and

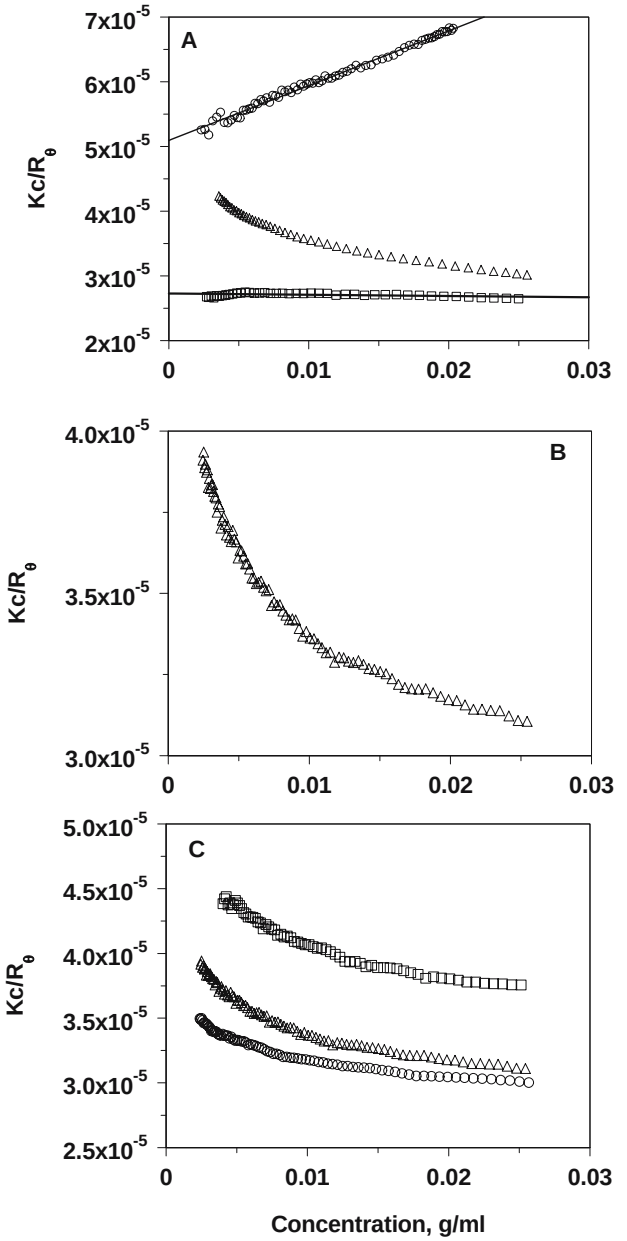

Fig. 2. **A** Debye plots of β-lactoglobulin A at pH 2.3, 0 M NaCl (*empty circles*), pH 3.0, 0.1 M NaCl (*empty triangles*) and at pH 3.0, 1.0 M NaCl (*empty squares*). **B** Expanded view of the Debye plot of β-lactoglobulin A at pH 3.0, 0.1 M NaCl. **C** Debye plots of β-lactoglobulin A at pH 3.0 for 0.05 M (*empty squares*), 0.1 M (*empty triangles*), and 0.2 M (*empty circles*) NaCl concentrations.

did not include the dimer involved in the association equilibrium, the chromatograms for different solution conditions were compared. Figure 3C shows the LS chromatograms at pH 3.0 for 0.02, 0.05 and 0.1 M NaCl. As observed, the maximum intensity of the dimer peak is similar in all conditions indicating that the dimer concentration is similar and is not affected by difference in association behavior. This indicated that the dimer peak observed is primarily composed of irreversible aggregates. Furthermore, a higher intensity of the main peak, as the solution NaCl concentration is increased, indicated increase in the formation of dimer involved in equilibrium, presumably corresponding to a higher association constant.

It is evident from this data that linear Debye plots are obtained under conditions where βLg exists primarily as a monomer and as a dimer, whereas, curved Debye plots are obtained under conditions where βLg exhibits monomer–dimer association equilibrium. The next step was to build up a model to analyze these curved Debye plots in an attempt to retrieve the association constants for monomer–dimer equilibrium.

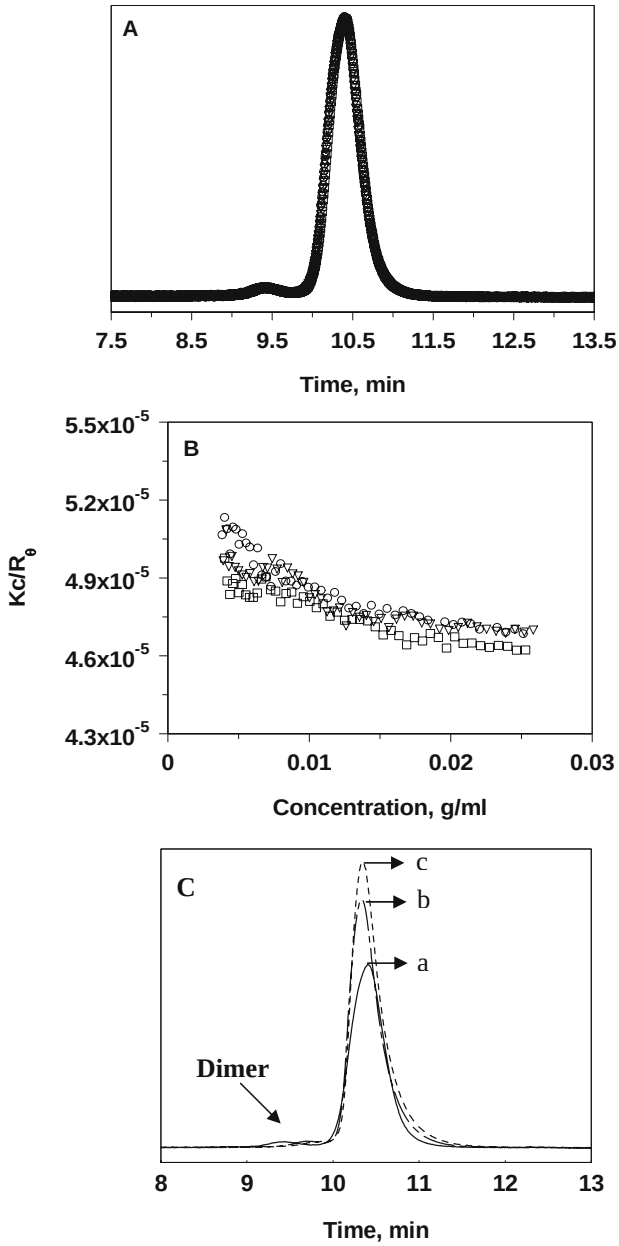

Fig. 3. **A** Chromatograms generated by the 90° light scattering detector for β-lactoglobulin A eluting through an SEC column following three separate injections (pH 3.0, 0.02 M NaCl). **B** Debye plots of β-lactoglobulin A obtained using the latter half of the main peak from the chromatograms shown in Fig. 3A (*see text for details*). **C** Chromatograms generated by the 90° light scattering detector for β-lactoglobulin A eluting through an SEC column at pH 3.0 for **a** 0.02 M NaCl, **b** 0.05 M NaCl, and **c** 0.1 M NaCl.

Model for Self-Association and Data Analysis

A monomer–dimer equilibrium is written as

$$M + M \overset{K_a}{\Leftrightarrow} D \tag{3}$$

Where the association constant K_a is defined as,

$$K_a = \frac{[c_d]}{[c_m]^2} \tag{4}$$

Where $[c_d]$ is the molar concentration of the dimer and $[c_m]$ is the molar concentration of the monomer. The total molar concentration, $[c_t]$, of the protein can be written in terms of the monomer concentration as,

$$[c_t] = [c_m] + 2[c_d] \tag{5}$$

Combining Eqs. 4 and 5, following quadratic equation is obtained,

$$2K_a[c_m]^2 + [c_m] - [c_t] = 0 \tag{6}$$

Solving for positive solution of $[c_m]$,

$$[c_m] = \frac{-1 + (1 + 8K_a[c_t])^{1/2}}{4K_a} \tag{7}$$

Converting molar concentration to g/ml using the expression $[c_m]=c_m/M_m$, where $[c_m]$ is molar concentration and c_m is concentration in g/ml, the monomer and dimer concentration can be written as,

$$c_{monomer} = \frac{-1 + (1 + 8000K_a c_t/M_m)^{1/2}}{4000K_a/M_m} \tag{8}$$

$$c_{dimer} = \frac{1 + 4000K_a c_t/M_m - (1 + 8000K_a c_t/M_m)^{1/2}}{4000K_a/M_m} \tag{9}$$

Note that the terms 4,000 and 8,000 in Eqs. 8 and 9 instead of 4 and 8 (Eq. 7) are used to maintain consistency between the volume units in K_a (liters/mole) and concentration (g/ml).

For an associating system, the Debye equation (Eq. 1) is modified to include the weight-average molecular weight of all the species present in the solution and the term B_{22} is replaced with the term B to represent general two-body interactions between all macromolecular species (for example, monomer–monomer, dimer–dimer, and monomer–dimer in the case of a monomer–dimer association equilibrium). The modified Debye equation is written as,

$$\frac{Kc_t}{R_\theta} = \left(\frac{1}{M_{av}} + B_{av}c_t\right) \tag{10}$$

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_{av} to represent the average nonideality arising from monomer–monomer, monomer–dimer and dimer–dimer interactions. Considering that there is negligible contribution of nonideality towards the curvature, our first approximation is to assume the term B to be zero. This assumption is routinely used in the analysis of the sedimentation equilibrium data (11). The Debye equation is then simplified to,

$$\frac{Kc}{R_\theta} = \frac{1}{M_{av}} \tag{11}$$

For an associating system, the change in the chemical potential of the solvent with solute concentration is written as (See Appendix II for details) (35),

$$\frac{\partial \mu_1}{\partial c_t} = \frac{\partial c_m}{\partial c_t \bullet M_m} + \frac{\partial c_d}{\partial c_t \bullet 2M_m} + \frac{\partial \left(B_{av}c_t^2\right)}{\partial c_t} \tag{12}$$

Once again, assuming B_{av}=0 in Eq. 12, substituting for c_m and c_d from Eq. 8 and 9, taking partial derivatives and using the result in the derivation of the Rayleigh's light scattering equation, the following Debye equation is obtained (see Appendix II for details),

$$\frac{Kc_t}{R_\theta} = \frac{1}{M_{av}} = \frac{(1 + 8K_a c_t/M_m)^{1/2} + 1}{2(1 + 8K_a c_t/M_m)^{1/2} M_m} \tag{13}$$

Eq. 13 is the modified Debye equation that was used to fit the nonlinear Debye plots for the parameters K_a using a monomer molecular weight of 18.4 kDa for βLg. The fitting was carried out by nonlinear least squares regression using Scientist software from Micromath (St. Louis, MO). Figure 4A shows the fit of Eq. 13 to the nonlinear Debye plot obtained for βLg at pH 3.0 and 0.05 M NaCl solution concentration. The equation does not fit well to the curved Debye plot as evident from the distribution of the residuals (Fig. 4A, inset), which shows that the data at low concentrations are under-predicted and that at higher concentrations are over-predicted by the equation. Thus, the association constant, K_a, does not allow accurate fitting of the data, pointing to the fact that the nonideality term could also contribute to the curvature.

In view of this, the nonideality term was included into the Debye Equation, i.e. Eq. 10 was used instead of Eq. 11. Following the same derivation procedure as described above using Eq. 12 with inclusion of the B term (See Appendix II for details), following Debye equation is obtained.

$$\frac{Kc_t}{R_\theta} = \frac{(1 + 8K_a c_t/M_m)^{1/2} + 1}{2(1 + 8K_a c_t/M_m)^{1/2} M_m} + Bc_t \tag{14}$$

Note that this is similar to Eq. 13 with an additional term that represents the first deviation from ideality. As discussed earlier, the term B represents the overall nonideality (See Appendix II) of arising from solute–solute interactions (monomer–monomer, monomer–dimer, dimer–dimer) in solution. Eq. 14 was now used to fit the curved Debye plots for the parameters K_a and B, using a value of 18.4 kDa for M_m.

Figure 4B, shows the fit of Eq. 14 to the nonlinear Debye plots obtained for βLg at pH 3.0 for 0.05 M NaCl solution concentration. Clearly, this equation fits well to the data (compare Fig. 4B to Fig. 4A), as also evident from the distribution of the residuals (inset). The sinusoidal-type pattern in the residuals results because of slight waviness of the data points in the Debye plot. Hence, the correction of the nonideality term is necessary for an accurate fit of the data. In fact, this was observed with all the solution conditions investigated in the present study. Figure 4C and 4D further

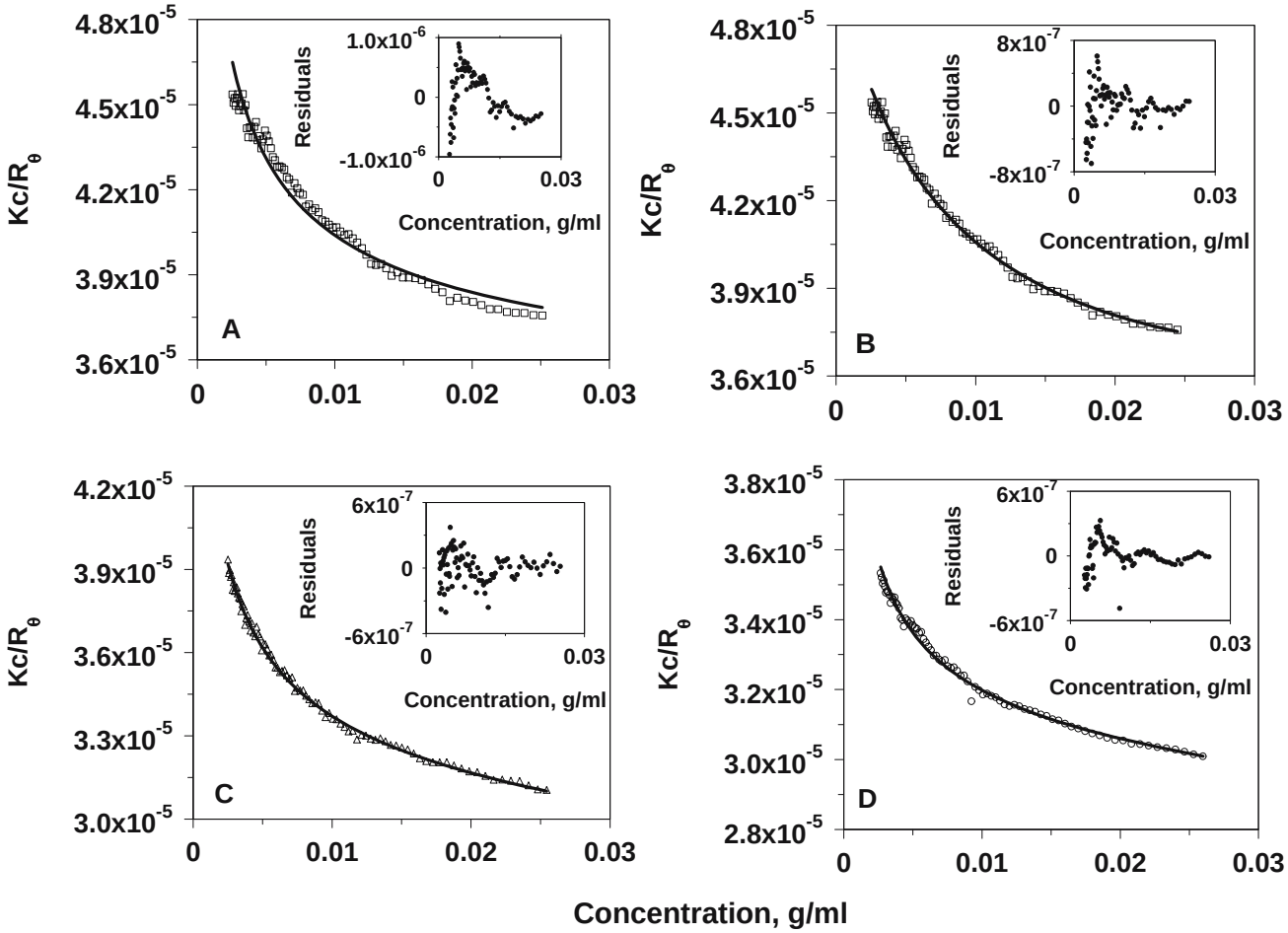

Fig. 4. Analysis of the nonlinear Debye plots of **β**-lactoglobulin A by nonlinear least squares regression. The markers represent the experimental data and the lines represent the fit to the experimental data using Eq. 13 at pH 3.0 for **A** 0.05 M, and using Eq. 14 at pH 3.0 for **B** 0.05 M, **C** 0.1 M, and **D** 0.2 M NaCl concentrations. The *insets* show the distribution of residuals between the experimental data and the theoretical fit.

demonstrate utilization of Eq. 14 for the analysis of the data of βLg at pH 3.0 for 0.1 and 0.2 M NaCl solution concentrations. Evidently, the equation fits well to the data. Note that, although Eq. 13 did not fit well to the data (Fig. 4A), the theoretical line still lies close to the data points. Thus, the major contribution to the curvature still arises from the association of monomers to form dimers and the contribution of the nonideality term although small, is important to attain an accurate fit of the data. In general, it was observed, that inclusion of the B term slightly reduced the K_a value. For example, a K_a value of 1.64×10^3 M^{-1} was obtained at pH 3.0, 0.05 M NaCl (Figs. 4A and 6), when B was not included in the model as against a value of 1.12×10^3 M^{-1}, when B was included (Figs. 4B and 6). Overall the data indicated that even though the dependence of K_a was small on B, inclusion of B was still necessary to obtain a good fit of the data.

The fact that the term B plays an important role in obtaining accurate fit to the data indicates the presence of nonideality in this particular system. The values of B obtained from the current analysis decrease as the pH is raised (for example, pH 2.3 and pH 3.0) and as the NaCl concentration is increased (pH 3.0) and even becomes negative at high NaCl concentrations. As reported previously and also observed in these studies, an increase in pH and solution NaCl concentration results in an increase in the population of dimers in solution. Theoretically, based on an increase in the dimer radius, the B value, calculated from simple hard sphere contributions, would be eight times more for a dimer–dimer interaction than for a monomer–monomer interaction. However, clearly, the B values actually decrease with an increase in dimer population. This shows that interactions other than excluded-volume contributions as the charge–charge interactions (presumably opposite charge interactions) play an important role in contributing towards the nonideality in the present case. However, it should be noted that the present study does not allow deconvoluting the contribution of individual monomer–monomer, monomer–dimer, and dimer–dimer interactions to the net B term.

The fitting yields the apparent K_a values and the B values, which are summarized in Table I for various solution conditions used to study βLg. Evidently, this method could track the K_a values over three orders of magnitude (10^2–10^5 M^{-1}). Note that the molecular weight of the βLg was used as a fixed parameter; however, it can also be used as a

floating parameter. We observed that when M_m was used as a floating parameter, the molecular weight obtained always lay within 10% of that reported in literature and did not significantly affect the apparent K_a values (Table I). This shows that the present method can also be used without prior accurate knowledge of the protein molecular weight.

Assessment of Equilibrium: Effect of Initial Solution Protein Concentration and Flow Rate

In SEC, while irreversible aggregates can be resolved by choosing appropriate column and solution conditions, the behavior of reversibly associating species would depend on the equilibrium and the interaction of these species with column, thus affecting the profile of the eluting peaks (36–38). Three cases of equilibrium could exist; slow equilibrium between the two species, intermediate equilibrium and rapid equilibrium. In case of slow equilibrium, the column will separate the associating species present in solution within the time frame of the experiment and these species will elute into different peaks. For this equilibrium, the ratio of the monomer-to-dimer peak areas would change for different concentrations injected because of different proportions of the two species present in the injection solution as defined by K_a. In the case of rapidly equilibrating associating species, the associating species will rapidly re-equilibrate following elution from column and the dimer and the monomer will not be resolved resulting in a peak with a polydisperse molecular weight distribution. For the intermediate case, one would expect some separation as well as asymmetry in the eluting peak depending on K_a and initial concentration.

The three different types of cases could be distinguished by injecting solutions of different initial protein concentrations and different flow rates. For example, for a rapidly equilibrating species, different flow rates will not affect the distribution of the associating species, whereas, the flow rate will have a distinct effect in the case of slowly equilibrating species. Hence, by varying initial protein concentration and flow rate, it would be feasible to study the type of equilibrium and whether equilibrium conditions are maintained during the time of study.

To test whether equilibrium conditions were maintained, the effect of initial protein concentration injected into the column and the flow rate on the estimation of apparent K_a values was investigated. Figure 5 shows the effect of initial protein concentration on the Debye plots at pH 3.0 (0.1 M NaCl). As observed, parallel Debye plots, which lie in close proximity to each other, were obtained for the two protein concentrations evaluated (15 mg/ml and 7.5 mg/ml). Although, exact overlap of the plots was not obtained, analysis indicated only slight differences in the apparent K_a values obtained (Table II) pointing to the fact that conditions close to equilibrium were maintained under these solution conditions.

Varying the flow rate slightly affected the apparent K_a values depending on the solution NaCl concentration. At low solution NaCl concentration (0.05 M NaCl), corresponding to a relatively lower apparent K_a value, flow rate did not appreciably affect the estimated K_a values (Table II). This is consistent with the effect of concentration on apparent K_a values, since at 0.1 M NaCl concentration, as discussed above, K_a values were not significantly affected. At higher solution NaCl concentration (0.5 M NaCl), corresponding to a relatively higher apparent K_a value, a lower flow rate resulted in a moderate increase in the estimation of the K_a value. This indicated that presumably equilibrium between the monomer and dimer species of βLg was affected slightly by the change in flow rate at 0.5 M NaCl solution concentration. However, note that the difference is only moderate indicating that the conditions were maintained near the "true" equilibrium of the system.

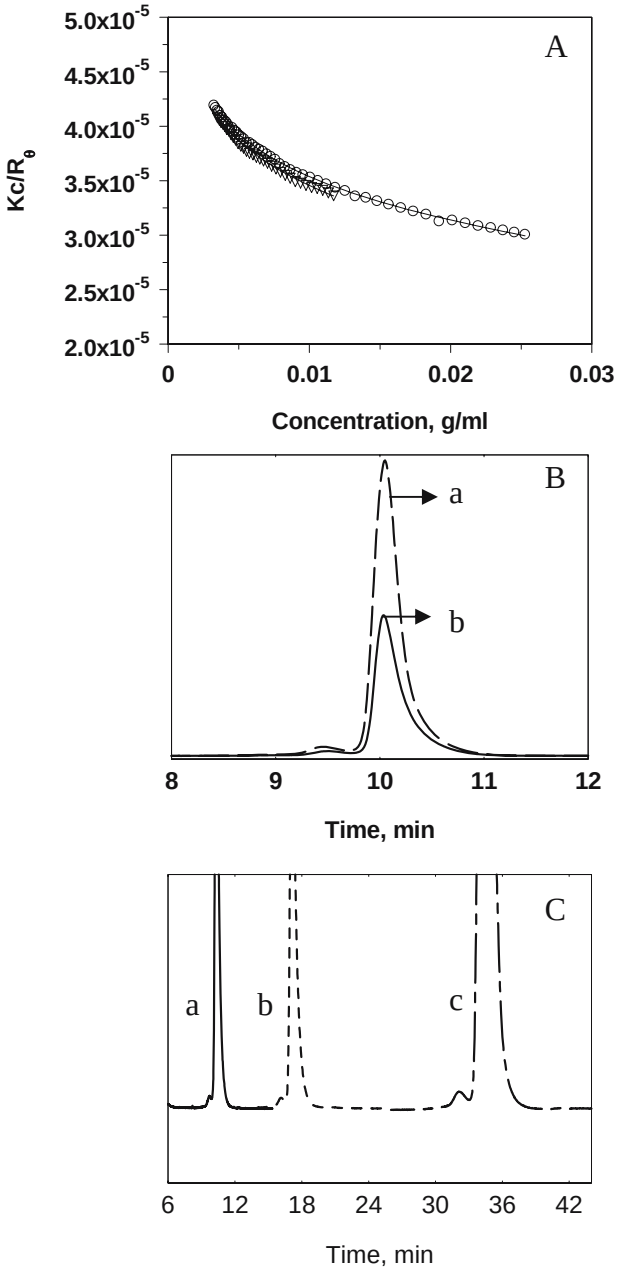

Fig. 5. **A** Effect of initial protein concentration on the Debye plots of β-lactoglobulin at pH 3.0, 0.1 M NaCl. (*empty circles*) 15 mg/ml and (*empty triangles*) 7.5 mg/ml. The markers represent the experimental data and the lines represent the fit to the experimental data using Eq. 14. **B** UV chromatograms of β-lactoglobulin A eluting through an SEC column at pH 3.0, 0.1 M NaCl for **A** 15 mg/ml and **B** 7.5 mg/ml. **C** UV chromatograms of β-lactoglobulin A eluting through an SEC column at pH 3.0, 0.05 M NaCl for **A** 1.0 ml/min, **B** 0.6 ml/min and **C** 0.3 ml/min).

Table II. Effect of Initial β-lactoglobulin Solution Concentration and Flow Rate on the Estimation of Association Constants at pH 3.0

Condition				
Initial protein concentration (1.0 ml/min flow rate, 0.1 M NaCl)	K_a (M^{-1})		$B\times10^4$ (mol ml/g^2)	
15 mg/ml	4.10 (±0.06, 0.52)[a] $\times10^3$		−0.3 (±0.1, 0.1)	
7.5 mg/ml	4.40 (±0.15, 0.63) $\times10^3$		−0.4 (±0.1, 0.1)	
	0.05 M NaCl		0.5 M NaCl	
Flow rate (15 mg/ml protein concentration)	$K_a\times10^{-2}$ (M^{-1})	$B\times10^4$ (mol ml/g^2)	$K_a\times10^{-5}$ (M^{-1})	$B\times10^4$ (mol ml/g^2)
1.0 ml/min	8.48 (±0.08, 1.50)	0.8 (±0.1, 0.1)	1.08 (±0.07, 0.30)	−0.1 (±0.1, 0.1)
0.6 ml/min	10.8 (±0.12, 1.50)	0.6 (±0.1, 0.2)	3.40 (±0.10, 0.45)	−0.08 (±0.1, 0.1)
0.3 ml/min	11.3 (±0.16, 1.80)	0.7 (±0.1, 0.2)	6.42 (±0.09, 0.40)	−0.1 (±0.1, 0.1)

[a] The first number in the parenthesis represents the maximum statistical uncertainty (goodness-of-fit) observed in the fitted value for a single experiment (1 SD) within 95% confidence interval and the second number represents the standard deviation in the estimation of the parameter from experiments in triplicate.

The presence of equilibrium under varying conditions of initial protein concentration and flow rate was further assessed by evaluating the percent main peak and percent dimer peak in the chromatograms. Nonequilibrium or slow equilibrium conditions will affect the distribution of the percent monomer and could also alter the shape of the main peak. Figure 5B and 5C show the LS chromatograms for varying initial protein concentration at pH 3.0 (0.1 M NaCl) and varying flow rate at pH 3.0 (0.05 M NaCl), respectively. As seen in Fig. 5B, the chromatogram for higher initial protein concentration (15 mg/ml) is similar in shape and profile to that of the low protein concentration (7.5 mg/ml), In fact, the percent dimer and percent main peak have similar values (2.0% dimer) indicating presence of fast equilibrium in these solution conditions. Similar results were also obtained for the chromatograms obtained with different flow rates. Note that while the resolution of the dimer is affected by the flow rate (highest at low flow rate of 0.3 ml/min), the overall, profile of the chromatogram and the percent main peak and percent dimer (2.0–2.3%) is similar across different flow rates. These data clearly affirm that conditions close to equilibrium were maintained in the current solution conditions evaluated. These data also further corroborate the fact that the dimer peak observed is primarily comprised of irreversible aggregates. In addition, no significant difference in the B values was observed across different flow rates and initial protein concentrations under these solution conditions further asserting presence of equilibrium (Table II).

The above data demonstrates the strength of this technique to assess whether equilibrium is maintained during these studies as investigated by varying initial concentration and flow rate. It should be noted that although true equilibrium may not be achieved for certain solution conditions and proteins, a simple analysis based on the shape of Debye plots from a single injection along with concentration and flow rate effects makes this technique a powerful tool to rapidly discern between associating and nonassociating species with minimum sample and time requirements. In the event that the data indicates nonequilibrium, subsequent detailed equilibrium sedimentation studies could be performed to obtain true binding constants.

Comparison with Literature Values

Figure 6 compare the values of apparent K_a obtained using the present method with those reported in literature for βLg (29). As evident, the values reported by us follow the same trend as those reported in literature, thus, validating the use of this method for determining equilibrium association constants of protein self-association. The somewhat lower values of K_a were attributed to the following factors. First, the inclusion of B value in the model lowers the K_a values obtained through fitting of the data. This is because both B and K_a values contribute to the negative slopes in the Debye plots. This is also illustrated in Fig. 6, that is, if the fitting of the data is carried out without incorporating the B value (as has been done in literature for βLg), higher values of K_a are obtained as compared to those when B value is used for fitting. Second, the use of SEC removes higher order aggregates, e.g., soluble irreversible dimers that are already present in the original material, and

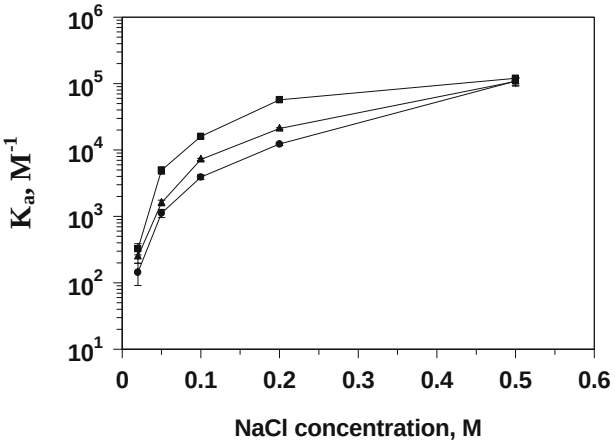

Fig. 6. K_a values for NaCl-dependent self-association of β-lactoglobulin A at pH 3.0 obtained by the method used in the present study (from Table I) using Eq. 14 (*closed circles*) and Eq. 13 (*closed triangles*) compared to those reported in literature (*closed squares*) (29). The *lines* are guide to the eye. Note that Eq. 14 incorporates the second virial coefficient, B, whereas, Eq. 13 does not.

could contribute towards K_a values if it remains in the sample under study. Finally, a nonequilibrium state compared to a "true" equilibrium state could also contribute to the slightly lower apparent "K_a" values. The similarity in the trend, however, demonstrates the application of this technique to screen different conditions for protein self-association behavior. It should be noted that we have evaluated a simple association model to assert the validity of this method. Currently, studies are underway to use this methodology and to develop models for more complicated self-association systems.

CONCLUSIONS

Rapid detection of (typically less than 1 h) of protein self-association behavior along with estimation of solution non-ideality (as determined through second virial coefficient) can be achieved using the present method that is based on simultaneous measurement of protein concentration and scattered light intensity in flow-mode in conjunction with size-exclusion chromatography. The dependence of scattered light intensity on protein concentration shows a typical curvature for an associating system in the Debye plots that can be analyzed to yield apparent association constants for relative scaling of different solution conditions. This method offers several advantages such as no interference from irreversible aggregates or dust particles and most importantly, amenability to high-throughput screening due to the use of automated HPLC-based method that can be used early on in protein formulation development with low sample requirements.

ACKNOWLEDGEMENTS

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

APPENDIX I

Molecular weight of the protein sample in dilute solutions and for polarized light is related to intensity of the scattered light from the sample through the following equation (26),

$$\mathrm{Mw} = \frac{N_A\,{}^4_o R^2 I_s}{4\,{}^2 \sin^2\phi c\left(dn/dc\right)^2 n^2 I_o} \tag{A1}$$

Where, N_A is the Avogadro's number, λ 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. 1 can be written as,

$$\mathrm{Mw} = \frac{I_s}{A_{90} c\left(dn/dc\right)^2} \tag{A2}$$

where,

$$A_{90}\frac{I_o 4\pi^2 n^2}{N_A \lambda_o^4 R^2} \tag{A3}$$

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

$$\frac{R^2}{I_o} = \frac{4\pi^2 n^2}{N_A \lambda_o^4 A_{90}} = K_1 \tag{A4}$$

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

$$R_\theta = \frac{I_s R^2}{I_o} \tag{A5}$$

Combining Eq. 4 and 5, Rayleigh's ratio can now be expressed as,

$$R_\theta = K_1 I_s \tag{A6}$$

Eq. A6 is used to obtain Rayleigh's ratio of a given data point on the LS 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,

$$c_{(\mathrm{g/ml})} = \log\left(\frac{I_{100\%\mathrm{T}} - I_{0\%\mathrm{T}}}{I_a - I_{0\%\mathrm{T}}}\right) \bullet 10/(E_{1\%} b) \tag{A7}$$

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. For estimation of parameter K in Eq. 1, the dn/dc values were obtained from the differential refractive index detector connected in line with the dual-detector cell, following calibration of this detector using a standard of known dn/dc (see calibration).

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. 100 µl 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 66000 was used to calculate calibration constants from the monomer peak of

BSA. Under these conditions, following calibration constants were obtained using the Precision Analyze software: $K_{90}=(B/A_{90})=4{,}569.8$ 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:

$$dn/dc = \frac{RI_{\text{sample}}}{B * A_{\text{inj}}} \tag{A8}$$

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

APPENDIX II

The chemical potential of the solvent in a two component system as a function of solute concentration is written as (35)

$$\mu_1 - G_1^0 = -RTV_1^0\left[\frac{c_2}{M_2} + Bc_2^2 + Cc_2^3 + \ldots\right] \tag{A9}$$

where, μ_1 is the chemical potential of the solvent, G_1^0 is the molar free energy of pure solvent, R is the gas constant, T is temperature, V_1^0 is the molar volume of the solvent, c_2 is solute concentration (g/ml), M_2 is the molecular weight of the solute, B and C are the second and the third virial coefficients, respectively.

For an associating species, where total solute concentration c_2 or $c_t=c_m+c_d$, the contributions of the two solutes are added and Eq. A9 is written as,

$$\mu_1 - G_1^0 = -RTV_1^0\left[\frac{c_m}{M_m} + \frac{c_d}{2M_m} + B_{av}c_2^2 + \ldots\right] \tag{A10}$$

Where, M_m is the molecular weight of the monomer. The term B_{av} represents the z-average second virial coefficient (defined by the quadratic mixing rule) (39) and for a monomer–dimer associating system is defined as,

$$B_{av} = B_m x_m^2 + B_d x_d^2 + B_{md} x_m x_d \tag{A11}$$

where, B_m, B_d, B_{md}, are the second virial coefficient values for the monomer–monomer, dimer–dimer, and monomer–dimer interactions, respectively, and x_m and x_d represent the mole fractions of the monomer and dimer, respectively. Defining, $x_m=[c_m]/[c_t]$; $x_d=[c_d]/[c_t]$; $[c_t]=c_t/M_m$, $[c_m]=c_m/M_m$; and $[c_d]=c_d/2M_m$, where, $[c_m]$ and c_m represent the molar concentration and g/L concentration, respectively, Eq. A11 can be written as,

$$B_{av}c_t^2 = B_m c_m^2 + \frac{B_d c_d^2}{4} + B_{md}c_m c_d \tag{A12}$$

Using the definitions of c_m and c_d from Eq 8 and 9, respectively, Eq. A12 can be written as,

$$\begin{aligned} B_{av}c_t^2 = {} & B_m\left(\frac{M_m}{4K_a}\right)^2 (D-1)^2 \\ & + \frac{B_d\left(\frac{M_m}{4K_a}\right)^2\left(\frac{4K_a c_t}{M_m} + 1 - D\right)^2}{4} \\ & + B_{md}\left(\frac{M_m}{4K_a}\right)^2 (D-1)\left(\frac{4K_a c_t}{M_m} + 1 - D\right) \end{aligned} \tag{A13}$$

Where,

$$D = \left(1 + \frac{8K_a c_t}{M_m}\right)^{1/2} \tag{A14}$$

The change in the solvent chemical potential as a function of the total solute concentration can be represented in the form of a partial differential equation of Eq. A10,

$$-\frac{1}{RTV_1^0}\frac{\partial\mu_1}{\partial c_t} = \frac{\partial c_m}{\partial c_t \bullet M_m} + \frac{\partial c_d}{\partial c_t \bullet 2M_m} + \frac{\partial\left(B_{av}c_t^2\right)}{\partial c_t} \tag{A15}$$

or

$$-\frac{1}{N_A kTV_1^0}\frac{\partial\mu_1}{\partial c_t} = \frac{1}{M_m}\left(\frac{\partial c_m}{\partial c_t} + \frac{1}{2}\frac{\partial c_d}{\partial c_t}\right) + \frac{\partial\left(B_{av}c_t^2\right)}{\partial c_t} \tag{A16}$$

The left term in Eq. A16 appears in the derivation of the light scattering equation for Rayleigh's scattering as (35),

$$\frac{i_s}{I_0} = \frac{2\pi^2 n^2 (dn/dc)^2\left(1+\cos^2\theta\right)c_t}{\lambda^4 r^2[-(1/V_1 kT)(\partial\mu_1/\partial c_t)]} \tag{A17}$$

Defining R_θ as,

$$R_\theta = \frac{i_s R^2}{I_o(1+\cos^2\theta)} \tag{A18}$$

and combining with Eq. A16, Eq. A17 can be written as

$$\frac{Kc_t}{R_\theta} = \frac{1}{M_m}\left(\frac{\partial c_m}{\partial c_t} + \frac{1}{2}\frac{\partial c_d}{\partial c_t}\right) + \frac{\partial\left(B_{av}c_t^2\right)}{\partial c_t} \tag{A19}$$

where, K is defined by Eq. 2 in the main text. Using Eqs. 8 and 9 for c_m and c_d, respectively, and Eq. A13 for the term $B_{av}c_t^2$, the following partial derivatives are obtained,

$$\frac{\partial c_m}{\partial c_t} = \frac{1}{\left(1 + 8K_a c_t/M_m\right)^{1/2}} \tag{A20}$$

$$\frac{\partial c_d}{\partial c_t} = 1 - \frac{1}{\left(1 + 8K_a c_t/M_m\right)^{1/2}} \tag{A21}$$

$$\begin{aligned} \frac{\partial\left(B_{av}c_t^2\right)}{\partial c_t} = {} & \left(\frac{M_m}{8K_a}\right)\left(\frac{D-1}{D}\right) \\ & \left[4B_m + B_d\left(\frac{4K_a c_t}{M_m} + 1 - D\right) + 4B_{md}\left(\frac{4K_a c_t}{M_m(D-1)} + D - 2\right)\right] \end{aligned} \tag{A22}$$

where, D is defined by Eq. A14. Substituting Eqs. A20–A22 in Eq. A19, and upon simplification the following expression is obtained,

$$\frac{Kc_t}{R_\theta} = \frac{(1+8K_a c_t/M_m)^{1/2}+1}{2(1+8K_a c_t/M_m)^{1/2}M_m} + \left(\frac{M_m}{8K_a}\right)\left(\frac{D-1}{D}\right) \left[4B_m + B_d\left(\frac{4K_a c_t}{M_m}+1-D\right) + 4B_{md}\left(\frac{4K_a c_t}{M_m(D-1)}+D-2\right)\right] \quad \text{(A23)}$$

Note that this expression involves five parameters to be fitted, namely, K_a, M_m, B_m, B_d, and B_{md}. A fitting routine using five parameters would easily fit a nonlinear curve, however, is prone to introduce more error in the estimation of the value of the parameter involved. To simplify the above equation and to improve reliability of the parameters obtained, we use a single parameter B instead of B_m, B_d, and B_{md}, to reflect the overall nonideality of the system. This approximation is often used to analyze the sedimentation equilibrium data for associating species involving nonideality (40–42). This simplification improves the fitting analysis, yet provides an estimate of nonideality present in the system. Following equation is obtained upon this simplification,

$$\frac{Kc_t}{R_\theta} = \frac{(1+8K_a c_t/M_m)^{1/2}+1}{2(1+8K_a c_t/M_m)^{1/2}M_m} + Bc_t \quad \text{(A24)}$$

APPENDIX III

The dual-detector dual-source cell for simultaneous measurement of light scattering intensity and protein concentration is illustrated in Fig. 7 (43). The flow-cell includes a sample cell, 1, defining an interior volume, 2, a first side occupied by a sample inlet, 3, and a second side occupied by a sample outlet, 4. The sample cell may be termed a "flow-through" cell that is configured to receive a continuous flow of protein solution from a SEC column (not shown) at its first side via the sample inlet, 3, and to continually discharge such flow at its second side via the sample outlet, 4.

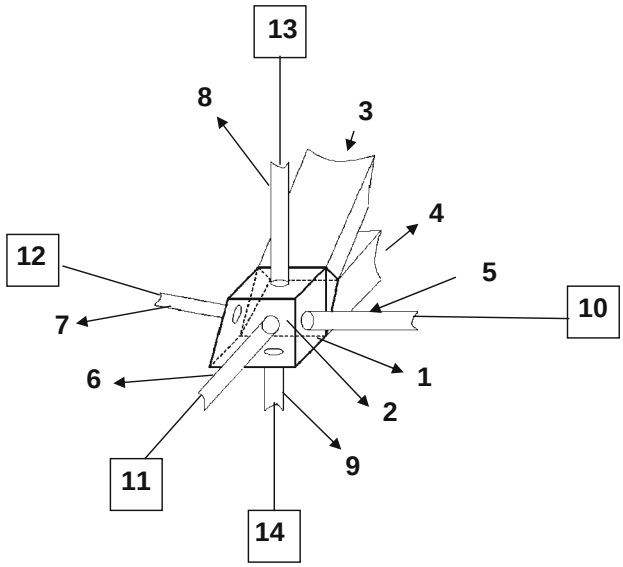

Fig. 7. A schematic of the dual-source, dual-detector multi-port cell that allows for simultaneous measurement of scattered light intensity at 90° and protein concentration through UV detection. See Appendix III for details.

Ports 5, 6, 7, 8, and 9, shown schematically in Fig. 7 occupy respective third, fourth, fifth, sixth, and seventh sides of the cell for permitting light and/or light beams to pass into the interior volume, 2, of the cell, into and through a solution passing through the cell (e.g., a protein solution), and/or outward of the interior volume, 2, of the cell for measurement of scattered light and/or concentration (e.g., protein concentration). A laser light source 10 is interoperably coupled to the cell via the port 5 for directing a laser light beam (not shown) into the cell, a detector 11 is interoperably coupled to the cell via the port 6 for receiving and detecting 90° scattered light from the laser light beam emerging from the cell, and a detector 12 is interoperably connected to the cell via the port 7 for receiving and detecting 15° scattered light from the laser light beam emerging from the cell.

An ultraviolet (UV) light source 13 is interoperably connected to the cell via the port 8 for directing a UV light beam into the cell, and a detector 14 is interoperably connected to the cell via the port 9 for receiving and detecting unscattered UV light from the UV light beam emerging from the cell. Ports 8 and 9 are aligned with each other and oriented at approximately a 90° angle to the laser light beam (not shown) directed by the laser light source, 10, into the cell so as to minimize the potential for interference between and among the detectors 11, 12, 14. The laser and UV light sources 10, 13 and the detectors 11, 12, 14 are all shown schematically for the sake of convenience and simplicity.

In the present set up, a laser light source was employed that produced collimated light at a wavelength of 685 nm. In addition, a fiber-optic cable type UV light source hosting a deuterium lamp and manufactured by MiniDATA UV (Analytical Instrument Systems, Flemington, NJ), and a detector for detection of transmitted UV light at 280 nm, were employed. The cell volume in the experimental system was 10 μl and the scattering volume was 0.01 μl. The path length for UV measurements was 3 mm.

REFERENCES

1. F. McNally. Capturing a ring of samurai. *Nat. Cell Biol.* **2**:E4–E7 (2000).
2. A. Weiss and J. Schlessinger. Switching signals on or off by receptor dimerization. *Cell* **94**:277–280 (1998).
3. N. J. Marianayagam, M. Sunde, and J. M. Matthews. The power of two: Protein dimerization in biology. *Trends. Biochem. Sci.* **29**:618–625 (2004).
4. G. Milligan. Oligomerization of G-protein-coupled receptors. *J. Cell. Sci.* **114**:1265–1271 (2001).
5. E. J. Parkinson, M. B. Morris, and S. Bastiras. Acid denaturation of recombinant porcine growth hormone: formation and self-association of folding intermediates. *Biochemistry* **39**:12345–12354 (2000).
6. S. N. Timasheff and K. C. Aune. Dimerization of a-chymotrypsin. I. pH dependence in the acid region. *Biochemistry* **10**:1609–1617 (1971).
7. S. N. Timasheff, K. C. Aune, and L. C. Goldsmith. Dimerization of a-chymotrypsin. II. Ionic strength and temperature dependence. *Biochemistry* **10**:1617–1622 (1971).

8. K. Snoussi and B. Halle. Protein self-association induced by macromolecular crowding: a quantitative analysis by magnetic relaxation dispersion. *Biophys. J.* **88**:2855–2866 (2005).
9. C. N. Patel, S. M. Noble, G. T. Weatherly, A. Tripathy, D. J. Winzor, and G. J. Pielak. Effects of molecular crowding by saccharides on a-chymotrypsin dimerization. *Protein Sci.* **11**:997–1003 (2002).
10. N. Kozer and G. Schreiber. Effect of crowding on protein–protein association rates: fundamental differences between low and high mass crowding agents. *J. Mol. Biol.* **336**:763–774 (2004).
11. L. W. Nichol, A. G. Ogston, and P. R. Wills. Effect of inert polymers on protein self-association. *FEBS Lett.* **126**:18–20 (1981).
12. S. J. Shire, Z. Shahrokh, and J. Liu. Challenges in the development of high protein concentration formulations. *J. Pharm. Sci.* **93**:1390–1402 (2004).
13. J. Liu, M. D. H. Nguyen, J. D. Andya, and S. J. Shire. Reversible self-association increases the viscosity of a concentrated monoclonal antibody in aqueous solution. *J. Pharm. Sci.* **94**:1928–1940 (2005).
14. J. L. Cole. Characterization of human cytomegalovirus protease dimerization by analytical centrifugation. *Biochemistry* **35**:15601–15610 (1996).
15. R. C. Williams Jr and D. A. Yphantis. Ultracentrifugation. *Encycl. Polym. Sci. Technol.* **14**:97–116 (1971).
16. A. P. Minton. Alternative strategies for the characterization of associations in multicomponent solutions via measurement of sedimentation equilibrium. *Prog. Colloid & Polym. Sci.* **107**:11–19 (1997).
17. P. Hensley. Defining the structure and stability of macromolecular assemblies in solution: the re-emergence of analytical ultracentrifugation as a practical tool. *Structure* **4**:367–373 (1996).
18. A.P. Minton. Quantitative characterization of reversible macromolecular associations via sedimentation equilibrium: an introduction. *Exp. Mol. Med.* **32**:1–5 (2000).
19. J. Lebowitz, M. S. Lewis, and P. Schuck. Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci.* **11**:2067–2079 (2002).
20. S. Y. Patro and T. M. Przybycien. Self-interaction chromatography: a tool for the study of protein–protein interactions in bioprocessing environments. *Biotech. Bioeng.* **52**:193–203 (1996).
21. S. Beeckmans. Chromatographic methods to study protein–protein interactions. *Methods* **19**:278–305 (1999).
22. D. J. Winzor and H. A. Scheraga. Chemically reacting systems on sephadex. II. Molecular weights of monomers in rapid association equilibrium. *J. Phys. Chem.* **68**:338–343 (1964).
23. A. K. Attri and A. P. Minton. New methods for measuring macromolecular interactions in solution via static light scattering: basic methodology and application to nonassociating and self-associating proteins. *Anal. Biochem.* **337**:103–110 (2005).
24. P. Doty, M. Gellert, and B. Rabinovitch. The association of insulin. I. Preliminary investigations. *J. Am. Chem. Soc.* **74**:2065–2069 (1952).
25. 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).
26. 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).
27. H. Bajaj, V. K. Sharma, A. Badkar, S. Nema, D. Zeng, and D. S. Kalonia. Protein structural conformation and not second virial coefficient relates to the long-term irreversible aggregation of a monoclonal antibody and ovalbumin in solution. *Pharm. Res.* **23**:1382–1394 (2006).
28. A. K. Attri and A. P. Minton. Composition gradient static light scattering: a new technique for rapid detection and quantitative characterization of reversible macromolecular hetero-associations in solution. *Anal. Biochem.* **346**:132–138 (2005).
29. K. Sakurai, M. Oobatake, and Y. Goto. Salt-dependent monomer–dimer equilibrium of bovine β-lactoglobulin at pH 3. *Protein Sci.* **10**:2325–2335 (2001).
30. G. Baldini, S. Beretta, G. Chirico, H. Franz, E. Maccioni, P. Mariani, and F. Spinozzi. Salt-induced association of β-lactoglobulin by light and X-ray scattering. *Macromolecules* **32**:6128–6138 (1999).
31. L. H. Tang and E. T. Adams Jr. Sedimentation equilibrium in reacting systems. VII. Temperature-dependent self-association of β-lactoglobulin A at pH 2.46. *Arch. Biochem. Biophys.* **157**:520–530 (1973).
32. S. Uhrinova, M. H. Smith, G. B. Jameson, D. Uhrin, L. Sawyer, and P. N. Barlow. Structural changes accompanying pH-induced dissociation of the beta-lactoglobulin dimer. *Biochemistry* **39**:3565–3574 (2000).
33. H. K. Eisenberg. Monographs on physical biochemistry: biological macromolecules and polyelectrolytes in solution. In H. K. Eisenberg (ed.), *Biological Macromolecules and Polyelectrolytes in Solution*, Oxford University Press, London, 1976, pp. 134–140.
34. G. Scatchard, A. Gee, and J. Weeks. Physical chemistry of protein solutions. VI. The osmotic pressures of mixtures of human serum albumin and g-globulins in aqueous sodium chloride. *J. Phys. Chem.* **58**:783–787 (1954).
35. C. Tanford. Thermodynamics. In C. Tanford (ed.), *Physical Chemistry of Macromolecules*, Wiley, New York, 1961, pp. 202–204.
36. K. Prochazka, T. Mandak, B. Bednar, J. Trnena, and Z. Tuzar. Behavior of reversibly associating systems in size exclusion chromatography. Interpretation of experimental data based on theoretical model. *J. Liquid. Chrom.* **13**:1765–1783 (1990).
37. G. K. Ackers and T. E. Thompson. Determination of stoichiometry and equilibrium constants for reversibly associating systems by molecular sieve chromatography. *Proc. Natl. Acad. Sci. U.S.A.* **53**:342–349 (1965).
38. T. W. Patapoff, R. J. Mrsny, and W. A. Lee. The application of size exclusion chromatography and computer simulation to study the thermodynamic and kinetic parameters for short-lived dissociable protein aggregates. *Anal. Biochem.* **212**:71–78 (1993).
39. K. R. Hall and G. A. Iglesia-Silva. Quadratic mixing rules for equations of state. Origins and relationships to the virial expansions. *Fluid Phase Equilib.* **91**:67–76 (1993).
40. H. Kim, R. C. Deonier, and J. W. Williams. The investigations of self-association reactions by equilibrium ultracentrifugation. *Chem. Rev.* **77**:659–690 (1977).
41. J. L. Cole. Characterization of human cytomegalovirus protease dimerization by analytical centrifugation. *Biochemistry* **35**:15601–15610 (1996).
42. J. E. Godfrey, T. H. Eickbush, and E. N. Moudrianakis. Reversible association of calf thymus histones to form the symmetrical octamer (H2AH2BH3H4)2: A case of mixed associating system. *Biochemistry* **19**:1339–1346 (1980).
43. US Patent pending. Filed 2005.