New Approaches to Investigating the Self-Association and Colloidal Stability of Protein Pharmaceuticals at High Concentrations

John S. Philo* and N. Karl Maluf

Alliance Protein Laboratories, San Diego, CA

Introduction

It is well known that self-association or non-specific molecular interactions in protein products at high concentrations can significantly impact solution viscosity, which in turn affects drug delivery and product manufacturability. It is also essential to minimize aggregation and particle formation (to have colloidal stability).

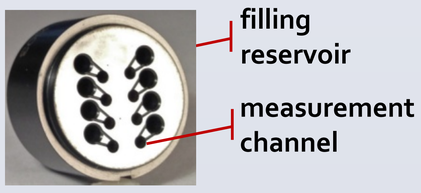
Aggregation and particle formation, solubility, and/or viscosity are often highly correlated with the solution second virial coefficient, B_{22}, which quantifies (to first order) the non-specific attractive or repulsive interactions between protein molecules (the solution "non-ideality").[1,2,3] However B_{22} is typically measured at concentrations below 10-20 mg/mL ("semi-dilute" solutions), and it is unclear whether this single parameter is truly sufficient to describe the behavior at formulated concentrations of ≥100 mg/mL which are now common.

Unfortunately experimental methods to probe weak interactions for samples at ~100 mg/mL and for measuring virial coefficients are rather limited, and can require prohibitive volumes of sample. The lack of primary methods using low volumes and with reasonable throughput has led many labs to use indirect methods such as measuring the concentration dependence of the diffusion coefficient via DLS,[2] methods which do not directly monitor the solution thermodynamic non-ideality or give B_{22}. Some methods such as self-interaction chromatography or nanoparticle plasmon spectroscopy require binding the protein to surfaces,[1,4] and thus may not reflect the true interactions of proteins free in solution. Further, it is not clear whether these different experimental probes are truly measuring the same thing---B_{22} values are not always equivalent, even when from first-principles methods such as static light scattering and sedimentation equilibrium.[5]

Here we illustrate a recently-described[6] new approach using sedimentation equilibrium (SE-AUC) on samples of very low volume (~20 μL), and demonstrate that it works at concentrations up to at least 120 mg/mL. The small sample volume allows equilibration to occur in only a few hours, increasing throughput, and together with the use of refractive index (RI) detection and low rotor speeds it keeps the concentration gradients across the SE cell within a workable range despite the high protein concentrations.

With this approach it should be possible to measure up to 64 samples per day (2 runs of 32 samples), at concentrations up to at least 200 mg/mL, and to do so with a first-principles method that directly gives the true osmotic second virial coefficient B_{22}.

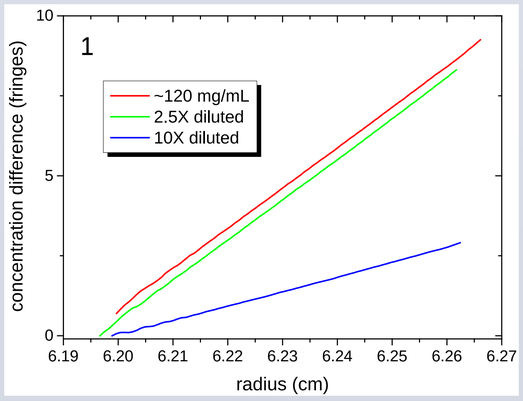
Experimental



These experiments were done using "8-channel" SE-AUC centerpieces, which are currently seldom used. Each cell holds 4 sample/solvent pairs. The 20 μL sample is initially loaded into the larger diameter filling reservoir, and then when the rotor is spun it flows through a capillary into the small cylindrical measurement channel to produce a sample with a column height (radial extent) of about 0.9 mm.

The trick that allows the acquisition of data at very high protein concentrations is to spin at very low rotor speeds (typically 4,000 to 8,000 rpm) so that at equilibrium the concentration difference across the sample stays very small (only a small fraction of the loading concentration). This keeps the refractive index gradient within the maximum range that can be followed by the Rayleigh interference optical system.

Figure 1 (right) shows some raw data for a moderately sized (20-70 kDa) aggregate-free protein loaded at ~120 mg/mL, 2.5X diluted, and 10X diluted, measured at 6,000 rpm. Due to the low rotor speed (1) the concentration varies linearly across the cell rather than exponentially, (2) the maximum concentration difference across the cell is quite small (<2.5% in this case), and (3) the concentration at the mid-point of the cell remains essentially at the loading concentration.



Data interpretation

Note that if the solution molar mass is constant with concentration, the concentration gradient at equilibrium will be proportional to concentration. Strikingly, in this case the concentration gradients (Fig. 1) are almost independent of concentration from ~50 to 120 mg/mL due to the strong repulsive solution non-ideality.

Under these conditions the apparent weight-average molar mass at each loading concentration, M_{app} can be determined from the concentration gradient at the column mid-point, $(\partial c/\partial r)_{\mathrm{mid}}$, the loading concentration, c_0, and the radius at the column mid-point, r_{mid}, using the following equation:

$$\frac{M_{app}(1-\bar{\nu}\rho)\omega^2}{\mathrm{RT}} = \frac{(\partial c/\partial r)_{\mathrm{mid}}}{r_{\mathrm{mid}} c_0}$$

where R is the gas constant, T is the temperature (Kelvin), ω is the rotor angular frequency, $\bar{\nu}$ is the protein partial specific volume, and ρ is the solvent density.

As can be seen in Fig. 2 (right), in this case the apparent molar mass (shown as a ratio relative to the native state sequence mass) decreases strongly with concentration rather than increasing due to reversible self-association. Indeed near 120 mg/mL M_{app} is nearly 4-fold lower than at low concentration.

Qualitatively Fig. 2 tells us these protein molecules exhibit strong repulsive interactions, but how do we quantitate that? The simplest model is the first-order virial expansion:

$$\frac{1}{M_{app}} = \frac{1}{M_0} + 2B_{22}c \qquad [1]$$

where the second virial coefficient B_{22} measures the strength of the net repulsion (positive values) or attraction (negative values).

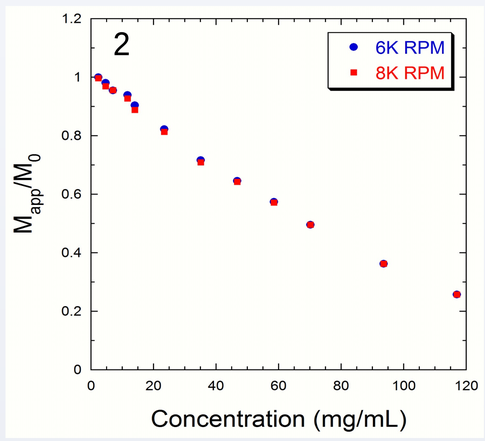
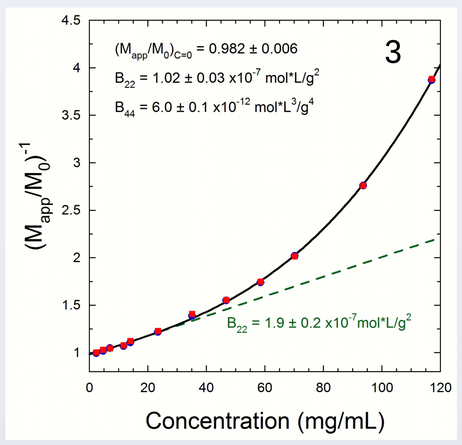


Figure 3 (right) re-plots the data from Fig. 2 in a form where Eq. 1 predicts a straight line whose slope gives B_{22}. In this case the data at the lower concentrations are reasonably approximated as a straight line, but over the whole range the plot is clearly strongly curved upward.



We have fitted the data over the full range of concentrations to a third-order virial expansion:

$$\frac{1}{M_{app}} = \frac{1}{M_0} + 2B_{22}c + 3B_{33}c^3 + 4B_{44}c^4 \qquad [2]$$

This model does provide a good fit (solid curve), but for these data the value of B_{33} happens to be zero (within experimental error), and thus it was held fixed at zero.

The B_{22} value from this fit is ~60% larger than the value predicted based solely on excluded volume (hard spheres). That could indicate some residual electrostatic repulsion (although these samples are at physiological ionic strength), or perhaps it is simply due to an extended conformation.

Discussion

These results illustrate two important points:

1. Extrapolating the results from the data below 20 mg/mL would grossly under-estimate the strength of the non-ideality at higher concentrations.
2. Ignoring the higher-order virial coefficients (assuming they are zero) and using only the data below 20 mg/mL would lead to a significant over-estimate for B_{22} (1.9 *vs.* 1.0 x 10^{-7} mol*L/g^2). That is, because the higher-order terms are so large very precise data at low concentrations would be needed to get a correct linear extrapolation.

Advantages & drawbacks of this approach

Advantages:

- low material requirements (20 μL samples)
- moderate throughput (up to 64 samples/day)
- trivial raw data analysis (measure slope)
- no interference from UV-absorbing excipients
- free-solution method (protein is not bound to a surface)
- directly measures true osmotic virial coefficients, even at high excipient concentrations (which static light scattering does not[5])

Drawbacks:

- for samples which reversibly associate above ~30 mg/mL (ones with high M_{app} values) it may not be possible to spin slow enough to reach the highest concentrations
- dilution errors, or protein loss on cell surfaces, directly affect M_{app} value accuracy
- getting precise data for concentrations below ~5 mg/mL may require additional data at higher rotor speeds or from larger sample volumes
- requires knowing either the native molar mass at low c, or buffer density + protein $\bar{\nu}$
- expensive equipment requiring skilled operator

References

(1) Valente, J. J., Payne, R. W., Manning, M. C., Wilson, W. W., and Henry, C. S. (2005). Colloidal behavior of proteins: effects of the second virial coefficient on solubility, crystallization and aggregation of proteins in aqueous solution. *Curr. Pharm. Biotechnol. 6,* 427-436.

(2) Garidel, P., Blume, A., and Wagner, M. (2014). Prediction of colloidal stability of high concentration protein formulations. *Pharm. Dev. Technol.*, epub ahead of print

(3) Saito, S., Hasegawa, J., Kobayashi, N., Kishi, N., Uchiyama, S., and Fukui, K. (2011). Behavior of monoclonal antibodies: Relation between the second virial coefficient B(2) at low concentrations and aggregation propensity and viscosity at high concentrations. *Pharm. Res. 29*, 397-410.

(4) Liu, Y., Caffry, I., Wu, J., Geng, S. B., Jain, T., Sun, T., Reid, F., Cao, Y., Estep, P., Yu, Y., Vasquez, M., Tessier, P. M., and Xu, Y. (2014). High-throughput screening for developability during early-stage antibody discovery using self-interaction nanoparticle spectroscopy. *MAbs 6,* 483-492.

(5) Winzor, D. J., Deszczynski, M., Harding, S. E., and Wills, P. R. (2007). Nonequivalence of second virial coefficients from sedimentation equilibrium and static light scattering studies of protein solutions. *Biophys. Chem. 128,* 46-55.

(6) Berkowitz, S. A. and Philo, J. S. (2014). Characterizing biopharmaceuticals using analytical ultracentrifugation. In: *Biophysical characterization of proteins in developing pharmaceuticals.* D.J. Houde and S.A. Berkowitz, eds. Elsevier, Amsterdam, pp. 211-260.