

Correlation between the osmotic second virial coefficient and solubility for equine serum albumin and ovalbumin

Kristen Demoruelle, Bin Guo, Shangming Kao, Heather M. McDonald, Dragan B. Nikic, Steven C. Holman and W. William Wilson*

Department of Chemistry, Mississippi State University, Box 9573, Mississippi State, MS 39762, USA. E-mail: bilwil@ra.msstate.edu

The Haas – Drenth – Wilson (HDW) (Haas *et al.*, 1999) theoretical model was used to correlate osmotic second virial coefficient (B) values with solubility (S) values for equine serum albumin (ESA) and ovalbumin for corresponding solution conditions. The best fit from the theoretical model was compared to experimental S *versus* B data. B values were experimentally measured using static light scattering. Solubilities of ESA were estimated using a sitting drop method. When the experimental data for S *versus* B were plotted, an excellent fit for ESA was obtained according to the HDW model. The results showed that the coordination number (z) in the crystal lattice was 6, and the adjustable parameter (A) was 0.072. For ovalbumin, previously reported solubility data in aqueous ammonium sulfate solutions were utilized. The solubility data for ovalbumin were correlated with the measured B values obtained in our laboratory. The resulting best fit from the HDW model showed that $z = 6$ and $A = 0.084$.

Keywords: equine serum albumin; ovalbumin; second virial coefficient; solubility

1. Introduction

The osmotic second virial coefficient (B) is a dilute solution thermodynamic parameter that reflects the extent of protein – protein interactions in a given solvent condition and has proven to be an effective predictor for several aspects of protein crystal growth. For example, the “crystallization slot” (George *et al.*, 1994) relates favorable crystallizing conditions for proteins to B values in the range of $(\sim -1 \text{ to } -8) \times 10^{-4} \text{ mol}\cdot\text{ml}\cdot\text{g}^{-2}$. This range of B values corresponds to protein – protein interactions described as weakly to moderately attractive and is consistent with ordered nucleation and subsequent post-nucleation growth. The concept of a crystallization slot has been verified (George *et al.*, 1994; Pjura *et al.*, 2000; Hitscherich *et al.*, 2000; Neal *et al.*, 1999) for a variety of water soluble and membrane proteins. More recently, it was pointed out that there is a correlation between the solubility (S) of a protein in aqueous solution and B of the solution (George *et al.*, 1997; Guo *et al.*, 1999). A theoretical explanation of this relation was given by Haas *et al.* (1999).

In this paper, B values for ESA and ovalbumin at different ammonium sulfate (AS) concentrations were obtained using static light scattering (SLS) methods. Solubility values for ESA were obtained using a sitting drop method (Nikic, 2001), while solubility data from the literature (Judge *et al.*, 1996; Sorensen *et al.*, 1915) was utilized for ovalbumin. It is shown for both ESA and ovalbumin that B and S are strongly correlated according to the HDW theoretical model. These results verify the practical implication that a dilute solution property (B) mimics phase behavior (S) for proteins, and that B measurements alone can be used to predict solubility behaviors.

2. Experimental

2.1. Reagents

The sources of the reagents used and purities were as follows: ESA 99% (Sigma Chemical Co., A3434) and ovalbumin 99% (Sigma Chemical Co., A5503), glacial acetic acid 99.7% (Fisher Scientific, A38), sulfuric acid 95 - 98% (Fisher Scientific, A484), sodium hydroxide 97% (Fisher Scientific, S318), and ammonium sulfate 99.7% (Fisher Scientific, A702). The water was distilled then deionized using a QUANTUM EX Ultrapure Organex Cartridge (Millipore).

The buffer for ESA was 0.1 M NaAc/HAc pH 5.6. The buffer was prepared by adding 6.0 g of glacial acetic acid to ~900 ml of water, titrating to pH 5.6 with 0.1 M NaOH, transferring to a 1 l volumetric flask, and filling to the mark with water. Buffer solutions containing ammonium sulfate (AS) were prepared by adding appropriate amounts of AS to 0.1M NaAc/HAc buffer then adjusting the pH to 5.6 using 0.1 M NaOH. ESA concentrations were measured spectrophotometrically at 276 nm using $A_{1\text{cm}}^{1\%} = 5.4$, according to the method of Hartree (1972).

For ovalbumin, aqueous AS solutions were prepared by weighing 100g of H₂O, adding a weighed amount of AS, and adjusting to the desired pH using 0.1 M sulfuric acid. Ovalbumin concentrations were determined spectrophotometrically at 280nm using $A_{1\text{cm}}^{1\%} = 7.0$ (Judge *et al.*, 1996).

2.2. Purification

The purity of commercial ESA was checked using a Tosohaas TSK gel G3000SWXL analytical size exclusion column, 7.8 mm x 30 cm, with a mobile phase of 0.1 M sodium sulfate plus 0.1 M sodium phosphate at pH 6.7. The elution data was collected using a Precision Detectors PD2000 Light Scattering Detector and a Waters 410 Differential Refractometer. The light scattering chromatographic profile for the commercial ESA sample is shown in Figure 1, and the weight-average molecular weights were determined to be 63 kDa for peak A (monomer) and 125 kDa for peak B (dimer).

Preparative size exclusion chromatography was performed on ESA using a Hi load 16/60 Superdex 200 HR (16 mm x 620 mm) size exclusion column (Pharmacia Biotech) connected to a BioCAD SPRINT Perfusion Chromatography System and Gilson FC 205 Fraction Collector. The mobile phase was 0.1 M NaAc/HAc buffer at pH 5.6 (the crystallization buffer), and good separation of monomer and dimer was achieved when the flow rate was 1.5 ml/min and 0.9 ml of ESA at 40 mg/ml was injected. The ESA monomer fraction was collected then concentrated using an Amicon 8010 concentrator with a DIAFLO ultrafiltration membrane YM5. The molecular size homogeneity of the concentrated ESA monomer solution was checked by dynamic light scattering (DLS) using a DynaPro 99 from Protein Solutions. The diffusion coefficient resulting from the DLS experiment gave estimates of $M \sim 64 \text{ kDa}$, $R_h = 3.8 \text{ nm}$, and polydispersity index = 0.02.

The purity of commercial ovalbumin was checked using a Tosohaas TSK gel G2000SWXL analytical size exclusion column, 7.8 mm x 30 cm, with a mobile phase of 0.1 M sodium sulfate plus 0.1 M sodium phosphate at pH 6.7. The elution data was collected using a Precision Detectors PD2000 Light Scattering Detector and a Waters 410 Differential Refractometer. The light scattering chromatographic profile for the commercial ovalbumin sample is shown in Figure 2, and the weight-average molecular weights were determined to be 45 kDa for peak A (monomer) and 91 kDa for peak B (dimer).

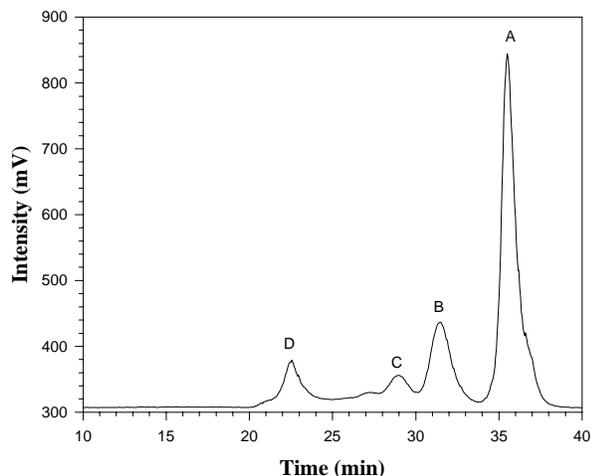


Figure 1 PD2000 light scattering detector intensity at 90° versus time. Size exclusion chromatogram of ESA from Sigma using a TosoHaas G3000SWXL column with eluent 0.1 M Na₂SO₄ + 0.1 M Na phosphate buffer pH 6.7. A: ESA monomer (63 kDa), B: ESA dimer (125 kDa), C: ESA trimer (~190 kDa), D: high molecular weight impurities.

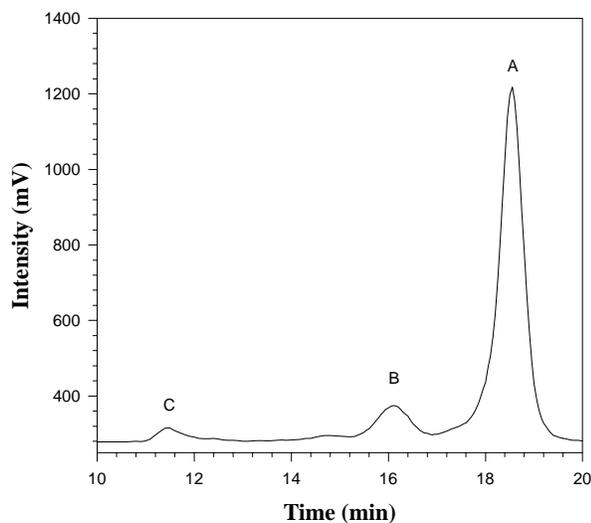


Figure 2 PD2000 light scattering detector intensity at 90° versus time. Size exclusion chromatogram of ovalbumin from Sigma using a TosoHaas G2000SWXL column with eluent 0.1 M Na₂SO₄ + 0.1 M Na phosphate buffer pH 6.7. A: ovalbumin monomer (45 kDa), B: ovalbumin dimer (91 kDa), C: high molecular weight impurities.

Commercial ovalbumin was purified using a Superdex 75 HR (16mm x 160 mm) size exclusion column (Pharmacia Biotech) with a mobile phase of 0.1 M AS solution at pH 4.9. Good separation of monomer and dimer fractions was achieved when the flow rate was 0.5 ml/min, the ovalbumin stock concentration was 50 mg/ml, and the injection volume was 0.25 ml. The ovalbumin monomer was collected then concentrated using an Amicon 8010 concentrator with a DIAFLO ultrafiltration membrane YM5. The molecular size homogeneity of ovalbumin was checked by DLS. The diffusion coefficient resulting from the DLS experiment gave estimates of $M \sim 44.3$ kDa, $R_h = 2.8$ nm, and polydispersity index = 0.02.

The concentrated monomer fractions of ESA and ovalbumin were stored in their elution buffer at 4°C. Integrity of the samples was verified by size exclusion chromatography before subsequent experiments were performed.

2.3. Measurement of second virial coefficients

The analytical method used to determine B was SLS. This method requires that the intensity of light scattered by a protein solution in excess of background scattering due to solvent and stray light be measured as a function of the protein concentration. The working equation used to analyze the SLS data is that given by Kratochvil (1987):

$$\frac{Kc}{R_{90}} = \frac{1}{M} + 2Bc, \quad (1)$$

where K is the optical constant ($\text{cm}^2 \cdot \text{mol} \cdot \text{g}^{-2}$),

$$K = \frac{4\pi^2 n_0^2 (dn/dc)^2}{N_A \lambda^4} \quad (2)$$

c is the protein concentration ($\text{g} \cdot \text{cm}^{-3}$), M is the molecular weight of the protein ($\text{g} \cdot \text{mol}^{-1}$), B is the second virial coefficient ($\text{mol} \cdot \text{ml} \cdot \text{g}^{-2}$), n_0 is the solvent refractive index, N_A is Avogadro's number (mol^{-1}), dn/dc is the refractive index increment ($\text{cm}^3 \cdot \text{g}^{-1}$), λ is the wavelength (cm) of the incident light in a vacuum, and R_{90} is the excess Rayleigh factor (cm^{-1}) at angle 90°.

When the SLS data is cast according to Equation (1), a linear plot is obtained with the slope equal to 2B. Scattering measurements were performed using a Right Angle Laser Light Scattering (RALLS) Detector Model 600 ($\lambda = 670$ nm) from Viscotek. Protein solutions of several known concentrations (~1 - 5 mg/ml) were injected directly into the 10 μl detection cell using a syringe. An inline solvent filter (Upchurch A-314) combined with 0.2 and 0.1 μm pore size Nucleopore polycarbonate membrane filters were used to remove dust particles from the solution that would cause spurious intensity fluctuations.

2.4. Solubility estimates for ESA

S values were estimated in 0.1M NaAc/HAc buffer at pH 5.6 and 22°C as a function of [AS] using a sitting drop method. For each solution condition, two rows of six 20 μl droplets were prepared, the first row with lower and the second row with higher initial ESA concentrations. The droplets were equilibrated in Costar plates at 22°C against 1 ml of reservoir solution containing buffer and the desired AS concentration. After 3 days, the mother liquor from two droplets at each solution condition was collected and centrifuged (Marathon 21000R, Fisher Scientific) at 22°C for ~10 minutes at 11,000 rpm to remove residual crystalline material. A measured volume of the supernatant was carefully removed and diluted to 120 μl with 0.1 M NaAc pH 5.6. The diluted ESA solution was placed in a 100 μl micro-cuvet and the UV absorbance at 276 nm was used to determine the protein concentration. The procedure was repeated at 7 days and 14 days.

2.5. Haas - Drenth - Wilson theoretical model

In a previous paper (Guo *et al.*, 1999), it was concluded that there was an empirical correlation between S and B as evidenced by data from lysozyme and ovalbumin. Using the data for lysozyme, Haas, Drenth, and Wilson (1999) derived a molecular-based theoretical expression (Equation 3) relating S and B.

$$B = \left(\frac{4}{M\rho} \right) \left[1 - A \left\{ \left(\frac{S}{10^3 \rho m} \right)^{-\left(\frac{2}{z}\right)} - 1 \right\} \right] \quad (3)$$

where B is the second virial coefficient, M is the molecular weight of the protein, ρ is the density of the protein ($\sim 1.36 \text{ g}\cdot\text{cm}^{-3}$), A is the fitting parameter, S is the solubility (mg/ml), $m = M / 18\rho$ (the number of water molecules that can be placed in the volume of one protein molecule), and z is the coordination number. The HDW theory models the interaction potential between protein molecules by using a square well potential and accounts for the restricted range of protein interactions (distance and orientation) using adjustable parameters A and z. The fitting parameter, A, depends on a combination of the anisotropy, p ($p = 1$ for isotropic interactions and $p < 1$ for anisotropic interactions), and the range of interactions between protein molecules. The coordination number, z, is the number of nearest neighbor contacts in the crystal lattice, alternatively interpreted by some as the number of macro-bonds in the crystal lattice.

3. Results

3.1. ESA

3.1.1. SLS. Figure 3 shows the Kc/R_{90} versus c plots for ESA in solutions of different AS concentrations. B values were calculated from the slopes of these plots and are presented in Figure 4. For [AS] less than $\sim 1 \text{ M}$, the B values are slightly positive. In a thermodynamic sense, this means the ESA molecules are experiencing a net repulsion among themselves, a condition not conducive to crystallization. As the concentration of AS increases beyond $\sim 1 \text{ M}$, the B values become progressively more negative. The solution conditions with negative B values correspond to net attraction among the ESA molecules and are consistent with B values within the crystallization slot. Average B values for the various solution conditions are reported in Table 1.

3.1.2. Solubility estimates for ESA. A sitting drop method was used to estimate the solubility of ESA at 22°C in 0.1 M NaAc/HAc buffer pH 5.6 for four different molarities of AS; 1.75, 2.00, 2.20, and 2.50.

Table 1 B and S data for ESA in 0.1 M NaAc/HAc, pH 5.6.

Crystallizing agent	T ($^\circ\text{C}$)	$B \times 10^4 \text{ (mol}\cdot\text{ml}\cdot\text{g}^{-2})$	Solubility (mg/ml)
None	22	2.0 ± 0.3	N/A
0.5 M AS	5	2.9 ± 0.3	N/A
0.5 M AS	22	1.1 ± 0.4	N/A
0.5 M AS	35	-0.4 ± 0.7	N/A
1.0 M AS	22	0.1 ± 0.8	N/A
1.5 M AS	22	-0.1 ± 0.4	N/A
1.75 M AS	22	-1.0 ± 0.3	35.7 ± 1.5
2.0 M AS	22	-1.7 ± 0.1	15.0 ± 0.7
2.20 M AS	22	-2.4 ± 0.3	6.3 ± 0.1
2.25 M AS	22	-2.8 ± 0.3	N/A
2.5 M AS	22	-4.2 ± 0.3	1.6 ± 0.3

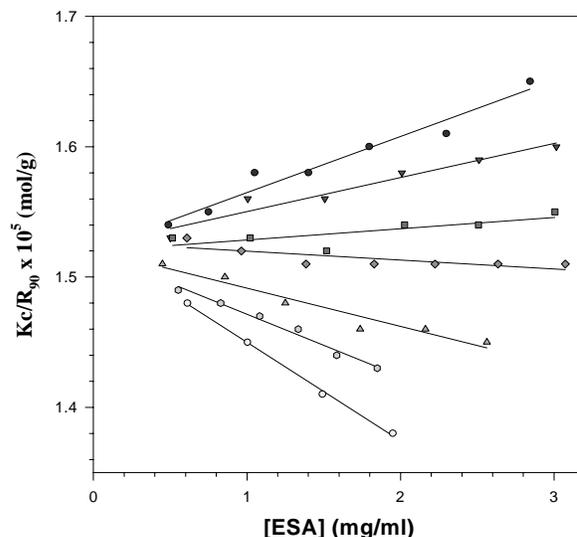


Figure 3 Kc/R_{90} versus [ESA] monomer at various [AS], pH 5.6, 22°C . [AS] (M): ● 0.0, ▼ 0.5, ■ 1.0, ◆ 1.5, ▲ 2.0, ○ 2.25, ○ 2.5

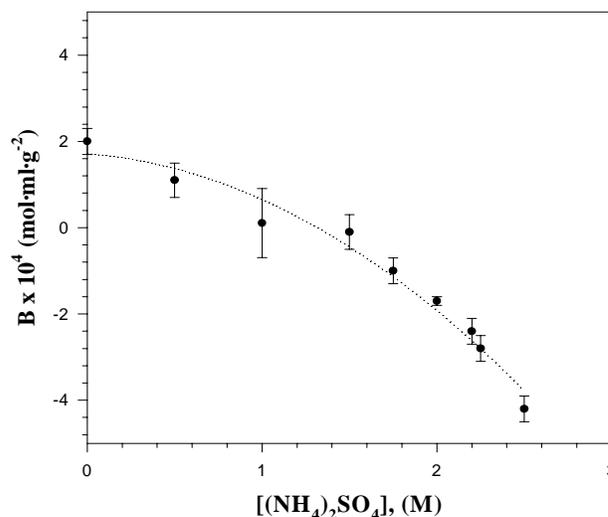


Figure 4 B versus [AS] for ESA monomer in 0.1 M NaAc/HAc, pH 5.6, 22°C .

These molarities correspond to initial [ESA] of 55 and 60 mg/ml, 30 and 40 mg/ml, 35 and 40 mg/ml, and 20 and 25 mg/ml, respectively. Resulting solubility estimates are reported in Table 1. Figure 5 plots [ESA] versus time and shows convergence at a value taken to be an estimate of the solubility for a final [AS] of 2.2 M. In this case, the higher initial [ESA] is 40 mg/ml and the lower initial [ESA] is 35 mg/ml. The final [ESA] converges at 6.3 mg/ml.

3.1.3. Haas – Drenth - Wilson fit. When S versus B values were plotted, an excellent fit was obtained according to the HDW model (Figure 6) over the entire range of the experimental data. Available S values for ESA range from 35.7 to 1.62 mg/ml, and corresponding B values range from $(-1.0$ to $-4.2) \times 10^{-4} \text{ mol}\cdot\text{ml}\cdot\text{g}^{-2}$. The resulting best fit shows according to the HDW model with $M = 65.5 \text{ kDa}$ that the coordination number (z) for ESA in the crystal lattice is six, and the adjustable parameter (A) is 0.072. This value for A corresponds to anisotropies of

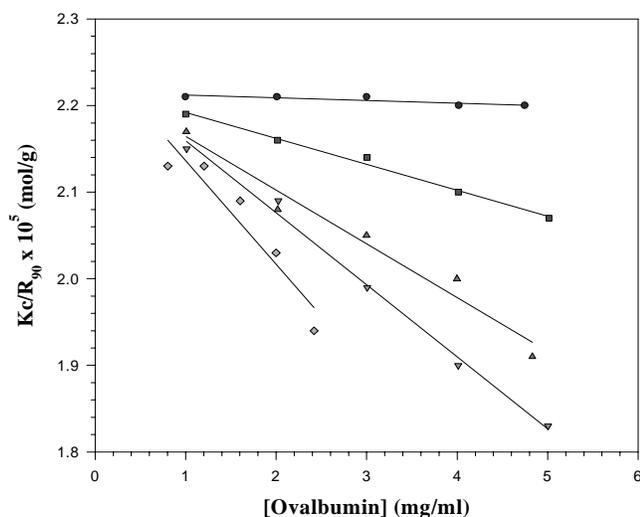


Figure 5 [ESA] versus time for sitting drop solubility estimates for ESA monomer in 0.1M NaAc/HAc, 2.2 M final [AS], pH 5.6, 22°C.

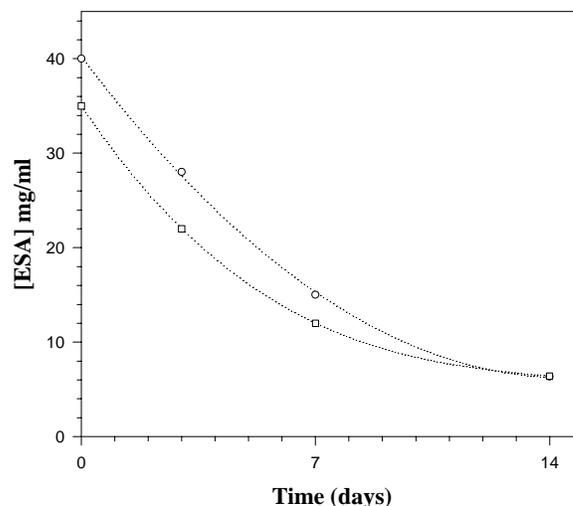


Figure 7 Kc/R_{90} versus [ovalbumin] monomer at various [AS], pH 4.9, and 30°C.

[AS] (g/100gH₂O): ● 22.0, ■ 23.6, ▲ 24.0, ▼ 24.5, ◆ 27.1

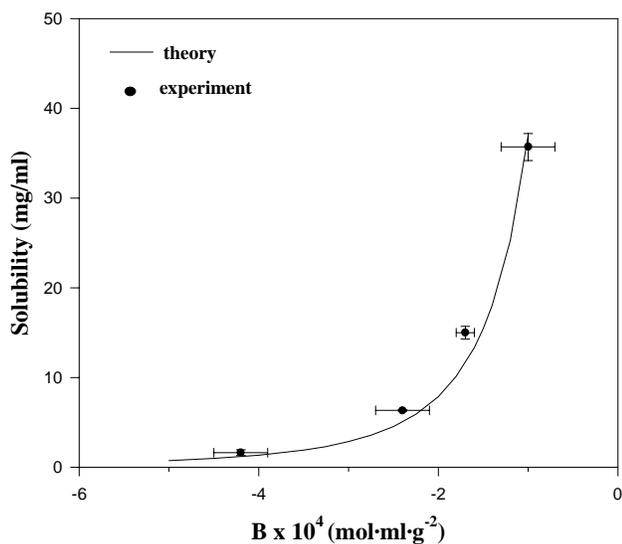


Figure 6 HDW best fit for ESA with $M = 65.5$ kDa, $\rho = 1.36$ g·cm⁻³, $z = 6$, and $A = 0.072$.

0.89 and 0.58 when the range of interactions are 2 Å and 3 Å, respectively (Haas *et al.*, 1999).

3.2. Ovalbumin

The Kc/R_{90} versus c plots of ovalbumin in different concentrations of AS are shown in Figure 7. Values for B were calculated from the slopes of these plots, and average values obtained from experimental repetitions are reported in Table 2 along with the average deviations. Measurements for B were taken at various pH's, temperatures, and [AS]. When S versus B values were plotted, a reasonably good fit was obtained according to the HDW model (Figure 8). S values range from 46.6 to 3.4 mg/ml, and corresponding B values range from $(-0.5$ to $-4.5) \times 10^{-4}$ mol·ml·g⁻². The resulting best fit shows according to the HDW model that with $M = 45$ kDa (Judge, 1996) z is six and $A = 0.084$. This value for A corresponds to anisotropies of 0.75 and 0.49 when the range of interactions are 2 Å and 3 Å, respectively (Haas *et al.*, 1999).

Table 2 B and S data for ovalbumin without buffer.

Crystallizing agent	pH	T (°C)	$B \times 10^4$ (mol·ml·g ⁻²)	Solubility (mg/ml)
22.0g AS / 100 g H ₂ O	4.9	30	$-0.5 \pm 0.6^*$	$46.4 \pm 2.3 \dagger$
23.6g AS / 100 g H ₂ O	4.9	30	-1.9 ± 0.4	$17.1 \pm 0.9 \dagger$
23.7g AS / 100 g H ₂ O	4.8	18	-3.0 ± 0.4	$18.8 \pm 0.8 \ddagger$
23.9g AS / 100 g H ₂ O	5.1	30	-1.6 ± 0.1	$31.0 \pm 1.6 \dagger$
24.0g AS / 100 g H ₂ O	4.9	29	-2.3 ± 0.4	$15.3 \pm 0.9 \ddagger$
24.0g AS / 100 g H ₂ O	4.9	30	$-3.3 \pm 0.5^{**}$	$15.0 \pm 0.8 \dagger$
24.5g AS / 100 g H ₂ O	4.9	30	$-3.5 \pm 0.4^{**}$	$10.3 \pm 0.5 \dagger$
25.8g AS / 100 g H ₂ O	4.8	18	-4.2 ± 0.2	$6.9 \pm 0.2 \ddagger$
25.9g AS / 100 g H ₂ O	5.3	18	-1.9 ± 0.2	$25.5 \pm 1.3 \ddagger$
27.1g AS / 100 g H ₂ O	5.0	18	-3.7 ± 0.1	$4.6 \pm 0.2 \ddagger$
27.6g AS / 100 g H ₂ O	4.9	30	$-4.5 \pm 0.3^{**}$	$3.4 \pm 0.2 \dagger$

* Previously published (Guo *et al.*, 1999) as $B = 0.2 \times 10^{-4}$ (mol·ml·g⁻²). Value was revised after further examination of the data.

** Previously published (Guo *et al.*, 1999)

† (Judge *et al.*, 1996)

‡ (Sorensen *et al.*, 1915)

4. Conclusions

ESA and ovalbumin monomer fractions were separated from oligomers present in the commercial products for use in SLS measurements. ESA solubility values were estimated using a sitting-drop method. S and B data for the two proteins were subjected to the HDW theoretical model. In each case, the DHW model sufficiently described the trend of the S versus B data. This work further supports the finding that B measurements can be used to quantitatively estimate S behavior of proteins and provide an attractive alternative to the often tedious and irreproducible direct measurements of S .

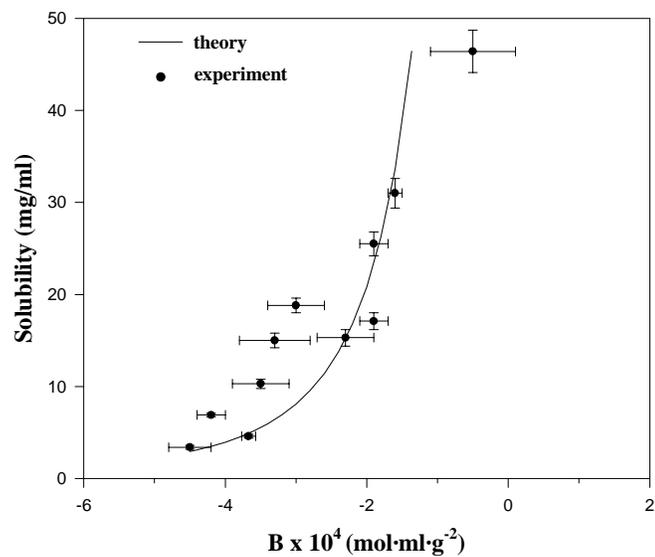


Figure 8 HDW best fit for ovalbumin with $M=45$ kDa, $\rho=1.36$ g·cm⁻³, $z=6$, and $A=0.084$.

References

- George, A. & Wilson, W. W. (1994). *Acta Cryst.* **D50**, 361-364.
- George, A., Chiang, Y., Guo, B., Arabshahi, Z., Cai & Wilson, W. W. (1997). *Methods Enzymol.* **276** 100-110.
- Guo, B., Kao, S., McDonald, H., Asanov, A., Combs, L. L. & Wilson, W. W. (1999). *J. Cryst. Growth*, **196**, 424-433.
- Haas, C., Drenth, J. & Wilson, W. W. (1999). *J. Phys. Chem. B*, **103**(14), 2808-2811.
- Hartree, E. F. (1972). *Anal. Biochem.* **48** 422-427.
- Hitscherich, Jr., C., Kaplan, J., Allaman, M., Wiencek, J. & Loll, P. J. (2000). *Protein Sci.* **9**, 1559-1566.
- Judge, R. A., Johns, M. A. & White, E. T. (1996). *J. Chem. Eng. Data*, **41** 422-424.
- Kratochvil, P. (1987). *Classical Light Scattering from Polymer Solutions*, Elsevier, Amsterdam, p. 24.
- Neal, B. L., Asthagiri, D., Velve, O. D., Lenhoff, A. M. & Kaler, E. W. (1999). *J. Cryst. Growth*, **196** 377-387.
- Nikic, D. B. (2001). *Solubility Measurements of Equine Serum Albumin and Sweet Protein Thaumatin I*, Ph.D. Dissertation, Mississippi State University, p. 53-58.
- Pjura, P. E., Lenhoff, A. M., Leonard, S. A. & Gittis, A. G. (2000). *J. Mol. Biol.* **300**, 235-239.
- Sorensen, S. P. L., Hoyrup, M. (1915-1917). *Studies on Proteins. IV. Comptes-Rendus des travaux du Laboratoire de Carlsberg*, **12**, 213-261.