ELSEVIER

Biophysical Chemistry 76 (1999) 13–23

Biophysical Chemistry

Analysis of the thermodynamic non-ideality of proteins by sedimentation equilibrium experiments

Joachim Behlke*, Otto Ristau

Max Delbrück Center for Molecular Medicine, Robert Rössle Str. 10, D-13122 Berlin, Germany

Received 17 June 1998; received in revised form 13 October 1998; accepted 13 October 1998

Abstract

This paper presents a modified method to determine experimentally the second virial coefficient of protein solutions by sedimentation equilibrium experiments. The improvement is based on the possibility of fitting simultaneously up to seven radial concentration distribution curves of solutions with different loading concentrations. The possibility of precise determination of the second virial coefficient allows estimation of the net charge and the excluded volume of a monomeric protein. Application of the method is demonstrated for lysozyme and ovalbumin. In 0.1 M sodium acetate buffer, pH 4.5, the second virial coefficient of hen egg white lysozyme amounts to 24 ± 1 ml/g. Analysis based on spherical particle theory yield an excluded volume of 3.5 ml/g and a charge dependent value of 20.5 ml/g which is induced by a net charge number of 14.1 ± 1. Under low salt conditions self-association processes on lysozyme are unfavorable due to electrostatic repulsion. To overcome these repulsive contributions, either a shift to neutral pH or addition of at least 2% NaCl is necessary. In this way the charge dependent contribution decreases below the value responsible for the excluded volume and allows crystallization of the protein. Similar effects can be observed with ovalbumin. The high virial coefficient observed at pH 8.5 is induced by the high net charge number of 27 ± 1.

Keywords: Sedimentation equilibrium; Virial coefficient; Excluded volume; Net charge; Lysozyme; Ovalbumin

1. Introduction

The chemical potential of a solute in a given solvent can be measured by the osmotic pressure, the scattering or sedimentation behavior [1]. These techniques allow one to determine the molecular mass of the solute by extrapolation of the measured parameters to infinite dilution. In solutions of non-associating solute at finite concentrations, the parameters are given by an ascending series of positive powers of solute concentration. These coefficients of the power series that influence the measurement and can be derived from the experiments in addition to the molecular mass are called virial coefficients (B).

The development of highly sensitive optics [2,3]

* Corresponding author. Tel.: +49 30 94062205; fax: +49 30 94062802; e-mail: behlke@mdc-berlin.de

0301-4622/99/$ - see front matter
PII:S0301-4622(98)00212-9

for analytical ultracentrifuges enables us to use more dilute biopolymer solutions well below 1 g/l that demonstrate an apparent ideal behavior. Deviations from ideality as indicated by the appearance of virial coefficients, have to be expected mainly due to steric exclusion or charge. This is obvious when analyzing weak association events that become clear at higher concentrations [4–10]. Charge induced effects can be observed in acidic protein solutions where the net charge can increase dramatically [11,12]. Charge-related repulsion can be reduced by adding neutral salts [1].

Recently, the application of the osmotic second virial coefficient was proposed to identify favorable crystallization conditions [13,14]. In general negative values for the second virial coefficient were determined by static light scattering in order to find the condition for protein crystallization. Normally for the second virial coefficient of proteins we have to expect positive values [10]. This is due to the excluded volume and Donnan effect that dominate in comparison to the attractive forces such as van der Waals interactions. Negative values for the second virial coefficient as observed by George and Wilson [13] or Rosenbaum and Zukowski [14] may be accounted for by equilibria describing self-association. Reduced scattering data obtained in such experiments correspond to the reversed apparent molecular masses of sedimentation equilibrium runs. The interpretation of these ultracentrifuge data require a set of equilibrium constants to describe the interaction of monomers with different oligomers. As demonstrated by several authors [4,5] the value of B or BM_1 varies with the association model chosen for analysis. Obviously, the estimated negative values for B obtained by using light scattering indicate reliable conditions for protein crystallization and do not allow one to get information about the excluded volume or the net charge of the protein. Likewise, this also concerns efforts to obtain virial coefficients from the slope of concentration-dependent molecular mass data plots derived by sedimentation equilibrium experiments. Holladay and Sophianopoulus [15] have tried to use a better model function considering equilibrium constants too but the results obtained are mostly statistically ascertained. An alternative method to characterize such solute self-associations derived by sedimentation equilibrium runs is based on the use of the so-called omega function. According to Jacobson and Winzor [16] it allows to estimate unequivocally the thermodynamic activity of monomers as a function of total solute concentration throughout the equilibrium distribution. However, in spite of the attempts of Morris and Ralston [17,18] to improve the extrapolation of the omega function to zero solute concentration, the accuracy with regard to determination of the monomer mass M_1 is limited. The same problem is observed in conventional curve fitting procedures too. Obviously the value M_1 has a certain influence on the second virial coefficient. Therefore, it seems advisable to introduce the theoretical molecular mass and not to estimate this value and the virial coefficient simultaneously.

Generally the effects of non-ideality on the molecular mass are rather small and require for their analysis very accurate concentration distribution curves at sedimentation equilibrium. Furthermore, to obtain accurate results for B, the number of fitted parameters has to be as small as possible. We have developed a method based on the analysis of sedimentation equilibrium experiments with up to seven protein solutions containing different loading concentrations. Details of this proposed procedure that allows to consider not only the loading but also the actual meniscus concentration in the distribution curves for determination of B for lysozyme and ovalbumin are presented here. From the data the net charge number of a protein under the current solvent conditions as well as the excluded volume can be derived.

2. Theory

To investigate the influence of charge on the activity coefficient it is necessary to add electrolytes to the protein solution. Rigorously this system consists of three components: protein, water and buffer or salt. At present no general method is available for complete description of a three component system. Therefore it is usual to apply the two component system and to take into the

bargain that some influence of the third component on the activity coefficient of the protein may occur. This influence may be small because of the very low molecular weight with respect to the protein.

To describe the radial concentration distribution of a macromolecular two-component system at sedimentation equilibrium, Williams et al. [19] have derived the following differential equation where the concentration c is given as a weight concentration:

$$\frac{M(1-\bar{\nu}\rho)\omega^2 rc}{RT} = \left[1 + c\left(\frac{\partial \ln \gamma_2}{\partial c}\right)_{T,P}\right]\left(\frac{\mathrm{d}c}{\mathrm{d}r}\right) \quad (1)$$

Here M is the molecular mass of the solute, $(1-\bar{\nu}\rho)$ the buoyancy term, r the radius position, R the gas constant, T the absolute temperature and γ_2 the activity coefficient of the solute. On the prerequisite that no volume change occurs during the alteration of concentration and that both solute and solvent are non-compressible for the density ρ, one can obtain the following relation [20]:

$$\rho = \rho_1 + (1-\bar{\nu}\rho_1)c \quad (2)$$

Eq. (2) indicates that the density of the solution is a function of the solvent density, the concentration and partial specific volume of the solute. In this way the alteration of solution density is considered by the rise in the concentration near the cell base. When inserting Eq. (2) into Eq. (1) we obtain the following expression (Eq. (3)).

$$\frac{M(1-\bar{\nu}\rho_1)(1-\bar{\nu}c)\omega^2 rc}{RT} = \left[1 + c\left(\frac{\partial \ln \gamma_2}{\partial c}\right)_{T,P}\right]\left(\frac{\mathrm{d}c}{\mathrm{d}r}\right) \quad (3)$$

In sedimentation equilibrium runs, the concentrations near the cell base achieve considerable values. Therefore, it is necessary to find a thermodynamically exact solution of the differential equation. So, a higher significance must be attached to the activity coefficients.

For the chemical potential of the solvent Eq. (4) is valid, especially based on the statistical mechanical calculations of McMillan and Mayer [21] that were also established for high concentrations [1].

$$\mu_1 = F_1^0 - RTV_1^0(c_2 + B_1^* c_2^2 + B_2^* c_2^3 + B_3^* c_2^4 \dots) \quad (4)$$

Here B^* are the molar virial coefficients, V_1^0 the molar volume of the solvent and c_2 the concentration of the solute in mol/l. The chemical potential of the solute μ_2 can be derived from Eq. (4) by means of the Gibbs–Duhem equation.

$$\left(\frac{\partial \mu_2}{\partial c_2}\right)_{T,P} = -\frac{n_1}{n_2}\left(\frac{\partial \mu_1}{\partial c_2}\right)_{T,P} \quad (5)$$

The ratio of molar numbers corresponds to the ratio of the molar concentrations. To determine the concentration c_1 of the solvent we can use the general relation (6) for non-compressible two-component systems (similar as for Eq. (2)).

$$c_1 V_1^0 + c_2 V_2^0 = 1 \quad \text{or} \quad c_1 = \frac{1 - c_2 V_2^0}{V_1^0} \quad (6)$$

V_2^0 is the molar volume of the solute. When substituting in Eq. (5) the molar numbers by the molar concentrations c_1 and c_2 from Eqs. (4)–(6) we obtain the following expression:

$$\left(\frac{\partial \mu_2}{\partial c_2}\right)_{T,P} = RT\left(\frac{1 - c_2 V_2^0}{c_2} \times (1 + 2B_1^* c_2 + 3B_2^* c_2^2 + 4B_3^* c_2^3 \dots)\right) \quad (7)$$

The integration yields Eq. (8).

$$\mu_2 = F_2 + RT\left(\ln(c_2) + c_2(2B_1^* - V_2^0) + c_2^2\left(\frac{3}{2}B_2^* - B_1^* V_2^0\right) + c^3\left(\frac{4}{3}B_3^* - B_2^* V_2^0\right)\dots\right) \quad (8)$$

Using the weight concentrations in g/l one obtains Eq. (9).

$$\mu_2 = \mu_2^0 + RT\left(\ln(c) + c(2B_1 - \bar{\nu}) + c^2\left(\frac{3}{2}B_2 - B_1\bar{\nu}\right) + c^3\left(\frac{4}{3}B_3 - B_2\bar{\nu}\right)\dots\right) \tag{9}$$

with $c = c_2 M_2$ and $B_i = B_i^*/M_2^i$.

Eq. (9) can be written also in the following manner:

$$\mu_2 = \mu_2^0 + RT\ln(c \cdot \gamma_2) \tag{10}$$

with the activity coefficient γ in the logarithmic form:

$$\ln(\gamma) = c(2B_1 - \bar{\nu}) + c^2\left(\frac{3}{2}B_2 - B_1\bar{\nu}\right) + c^3\left(\frac{4}{3}B_3 - B_2\bar{\nu}\right) \tag{10a}$$

given in l/g.

The additional terms that contain the partial specific volume of the solute result from the correct determination of the molar number ratio in the Gibbs–Duhem equation. When inserting the activity coefficient from Eq. (10) into Eq. (3) one obtains:

$$\frac{M(1-\bar{\nu}\rho_0)(1-\bar{\nu}c)\omega^2 rc}{RT} = \left[1 + c(2B_1 - \bar{\nu}) + 2c^2\left(\frac{3}{2}B_2 - B_1\bar{\nu}\right) + 3c^3\left(\frac{4}{3}B_3 - B_2\bar{\nu}\right)\dots\right]\left(\frac{\mathrm{d}c}{\mathrm{d}r}\right) \tag{11}$$

or after division of the right side by $(1 - \bar{\nu}c)$, i.e. cancelling of the corresponding multiplication in Eq. (7), the following expression yields:

$$\frac{M(1-\bar{\nu}\rho_0)\omega^2 rc}{RT} = \left[1 + 2B_1 c + 3B_2 c^2 + 4B_3 c^3\dots\right]\left(\frac{\mathrm{d}c}{\mathrm{d}r}\right) \tag{12}$$

After separation of the variables and integration of Eq. (12) from r_0 to r or c_0 to c_r one obtains the following equation truncated beyond the third virial coefficient:

$$\frac{M(1-\bar{\nu}\rho_0)\omega^2(r^2 - r_0^2)}{2RT} = \ln\left(\frac{c_r}{c_0}\right) + 2B_1(c_r - c_0) + \frac{3}{2}B_2(c_r^2 - c_0^2)\dots \tag{13}$$

Eq. (13) in exponential form results:

$$c_r = c_0\exp\left(\frac{M(1-\bar{\nu}\rho_0)\omega^2(r^2 - r_0^2)}{2RT} - 2B_1(c_r - c_0) - \frac{3}{2}B_2(c_r^2 - c_0^2)\dots\right) \tag{14}$$

The correct treatment of the mole fraction in the Gibbs–Duhem equation describes the decrease of the molar number of solvent with increasing concentration of solute. This fact which is usually neglected, leads to a surprising result. The alteration of solvent density and therefore alteration of the buoyancy term with rising concentration of the solute towards the cell base is eliminated from the centrifuge equation. In this way the virial coefficients used in our equations are comparable with those derived by statistical mechanical considerations. The same result was obtained by Wills and Winzor [22] who used molal concentrations. In contrast to the molarity, the molality directly represents the molar ratio of solute/solvent.

For determination of the virial coefficients it is useful to record several concentration profiles with different loading concentrations at equal speed and identical column height. An eight-hole rotor can be used for such kind of experiments. For ideal monodisperse solutions one obtains several parallel concentration distribution curves. The concentration c_0 at the reference radius r_0 is exactly proportional to the current loading con-

centration. This is particularly relevant because the number of estimated parameters can be reduced.

The improved accuracy of the analytical ultracentrifuge (XL-A) to simultaneously measure a large number of concentration profiles enables us to reliably determine the non-ideality of globular proteins. The virial coefficients determined in this manner are thermodynamically correct ones and agree well with those derived by McMillan and Mayer [21]. They can be interpreted by the theory of statistical mechanics.

3. Materials and methods

Stock solutions of crystallized hen egg lysozyme (Sigma, St. Louis, USA) in 0.1 M sodium acetate, pH 4.5, or 0.1 M potassium phosphate, pH 7.0, were diluted with the same buffer containing different amounts of NaCl as indicated. To remove aggregates the protein solutions were centrifuged for 15 min at 10 000 rev./min on a table centrifuge and the protein concentration was subsequently determined using an absorbance of 2.65 for 1 g/l at 280 nm [23]. In each cell ~ 300 μl of protein solution with loading concentrations among 1.5 and 4.5 g/l were filled in seven double sector cells with a path length of 12 mm. Ovalbumin (Sigma) was dissolved in 0.05 M potassium phosphate buffer, pH 8.5, and treated as above. The ovalbumin concentration was determined spectroscopically using an absorbance of 0.716 for 1 g/l and 280 nm [23].

The sedimentation equilibrium experiments were carried out by means of an eight-hole rotor at 10°C using an analytical ultracentrifuge XL-A (Beckman, Palo Alto, CA). To obtain the absorbance coefficient under centrifugation conditions, the absorption of the cells with different loading concentrations was recorded at 300 nm for lysozyme or 291 nm for ovalbumin during the beginning centrifugation phase at 3000 rev./min. An average coefficient was calculated from the recorded values between r_m and r_b considering the loading concentrations. This value allowed us to determine the actual concentration of the profiles at sedimentation equilibrium that was reached after 2 h at 24 000 rev./min (overspeed) and 45–50 h equilibrium speed at 20 000 rev./min.

To determine the second and third virial coefficient with simultaneous calculation of the meniscus (reference) concentration from the loading concentration, it is useful to consider the differential Eq. (12). The average (loading) concentration c_E is obtained by direct integration and consideration of Eq. (16), (see also Fujita, [20]):

$$c_E = \frac{c_b - c_m + B_1(c_b^2 - c_m^2) + B_2(c_b^3 - c_m^3)}{M_r(r_b^2 - r_m^2)} \tag{15}$$

with

$$c_E = \frac{2}{r_b^2 - r_m^2}\int_{r_m}^{r_b} c_r r \mathrm{d}r \tag{16}$$

r_b, r_m = bottom − , meniscus radius

M_r is the reduced molecular mass.

$$M_r = \frac{M(1 - \bar{\nu}\rho_0)\omega^2}{2RT} \tag{17}$$

When equating reference and meniscus radius from the exponential Eq. (14) we obtain the following relation between concentration at the radius of the bottom and meniscus:

$$c_m = c_b \exp\Big(-M_r(r_b^2 - r_m^2) + 2B_1(c_b - c_m) + \frac{2}{3}B_2(c_b^2 - c_m^2)\Big). \tag{18}$$

Eq. (15) in a favorable form for the iteration reads:

$$c_b = C_E M_r(r_b^2 - r_m^2) + c_m - B_1(c_b^2 - c_m^2) - B_2(c_b^3 - c_m^3). \tag{19}$$

Using Eqs. (18),(19) we are now able to calculate iteratively the concentration at the meniscus (reference position) directly from the loading concentration c_E and the radius of the meniscus and bottom.

To calculate the start concentration c_m for further iteration the following relation is suitable:

$$c_m = \frac{c_E M_r (r_b^2 - r_m^2)}{\exp\left(M_r (r_b^2 - r_m^2)\right) - 1}. \quad (20)$$

In this equation the virial coefficients are held at zero. For the starting concentration at the cell base, only the second virial coefficient B_1 is considered. From Eq. (19) it follows:

$$c_b = \frac{1}{2B_1} \left(\sqrt{1 + 4B_1^2 c_m^2 + 4B_1\left[c_E M_r (r_b^2 - r_m^2) + c_m\right]} - 1\right). \quad (21)$$

Using these starting values one can calculate the values of c_m and c_b by repeated iterations of Eqs. (18),(19).

The derivations of c_r in Eq. (14) according to the estimated parameters M_r, B_1 and B_2 can be obtained analytically. We developed a computer program VIRIAL written in Turbo Pascal and running under MS-DOS 6.0. It is able to fit simultaneously up to seven concentration profiles from an eight-hole rotor to the model function Eq. (14). A common base line for the seven profiles was fitted as additional parameter.

To determine the uniform loading concentration c_E, we use a numerical integration of the experimental profiles and divide by the extinction coefficient that likewise is obtained by the XL-A at low speed (3000 rev./min). This was done by averaging of the chosen sections of absorbance records and division by the loading concentration. The numerical integration of the concentration profiles was achieved by fitting the profile sections chosen graphically at the monitor to an auxiliary function with subsequent integration. As auxiliary function Eq. (22) containing seven parameters (p_i) is reliable.

$$c_r = p_1 \exp\left(p_2 (r^2 - r_b^2)\right) + p_3 \exp\left(p_4 (r^2 - r_b^2)\right) + p_5 r^8 + p_6 r^3 + p_7 \quad (22)$$

The fitting procedure is restricted so that only the constant term p_7 is allowed to adopt negative values. By integration of Eq. (22) according to Eq. (16) an average concentration of the chosen section is determined. This procedure has the advantage that only the average concentration of the chosen profile sections selected manually by mouse and cross wires have to be used for calculation of the reference concentration. It is often difficult to find the exact bottom radius because deviations from the expected curves by artefacts are also observed. On the other hand, this part of the profile contributes considerably to the total concentration. Because the auxiliary function (22) contains only seven parameters that have to be estimated, concentration profiles with noise can be accurately fitted and integrated. The influence of a base line on the magnitude of the uniform loading concentration determined by numerical integration is taken into consideration by the fit program VIRIAL.

When analyzing experiments far from the isoelectric point of a protein a charge-dependent term in addition to the excluded volume part of the second virial coefficient arises. Under the assumption that the charges of macromolecules are spread evenly over its surface, Wills and Winzor [22] have calculated the charges based on the Debye Hückel theory. To compute this we must apply the potential function of the ion in the bulk solution derived by the Debye Hückel theory [1] as described by Eq. (23).

$$\varphi = \frac{q}{(1 + R_M \kappa) D} \cdot \frac{\exp(-r\kappa + R_M \kappa)}{r} \quad (23)$$

Here, κ is the inverse Debye Hückel radius, r the distance from the midpoint of the macroion, R_M the radius of the macroion, q the net charge ($Z \cdot e$) of the ion and D the dielectric constant. The first term in this equation (except D) reflects the effective charge and the other part describes the distance dependency of the potential.

Therefore, the interaction energy between two macroions can be described by Eq. (24).

$$\mu = \frac{q^2}{(1 + R_M \kappa)^2 D} \cdot \frac{\exp(-r\kappa + R_M \kappa)}{r} \quad (24)$$

When substituting this relation in the function of McMillan and Mayer [21] and developing a Taylor series up to the first term we obtain the charge dependent part of the second virial coefficient B_1^q (Eq. (25)) with k, the Boltzmann constant.

$$B_1^q = -2\pi N_A \int_R^\infty \left(\exp\left(-\frac{\mu}{kT}\right) - 1\right) r^2 \mathrm{d}r \qquad (25)$$

The total amount of the second virial coefficient given in the dimension of a weight concentration can be obtained by Eq. (26) as pointed out by Wills and Winzor [22].

$$B_1 = \frac{16\pi}{3M} N_A R_M^3 + \frac{Z^2[1 + 2\kappa R_M]}{4MI(1 + \kappa R_M)^2} \qquad (26)$$

Here I denotes the ionic strength. For a known hydrodynamic radius of the macroion or the excluded volume, the net charge number Z can be derived directly from B_1. For the excluded volume the simple spheric shape was adapted that means the same approximation as used in the Debye Hückel theory.

4. Results

4.1. Lysozyme

4.1.1. Experiments in acidic solution and at low ionic strength

Fig. 1 represents seven radial concentration distribution curves of lysozyme dissolved in 0.1 M sodium acetate buffer, pH 4.5, obtained from one sedimentation equilibrium run. The curves were fitted simultaneously by Eq. (14) to obtain the second virial coefficient B_1. To reduce the number of degrees of freedom, the molecular mass of monomeric lysozyme with 14 316 Da [24] was held constant. For the fitting procedure only those curve sections were used that fitted the profile satisfactorily. Sometimes small parts of the deviating data in the meniscus region based on convection or aggregates near the cell base were removed. The actual concentrations have to be determined by means of the integration procedure.

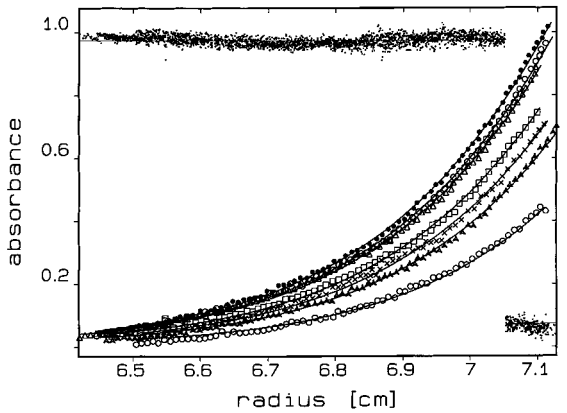

Fig. 1. Radial concentration distribution curves of lysozyme dissolved in 0.1 M acetate buffer, pH 4.5, at sedimentation equilibrium reached after 50 h at 20 000 rev./min and 10°C. The absorption was recorded at 300 nm and fitted using Eq. (14) taking into account a molecular mass $M_1 = 14\,316$ Da for the monomeric protein. Some data of small curve sections with irregularities near the meniscus or the cell base (aggregates) were omitted to increase the accuracy of the fit. According to the integration of the areas under the profiles the real concentrations of the curves used for the calculation amount from top to bottom to 2.67 g/l; 2.56 g/l; 2.32 g/l; 1.93 gl; 1.82 g/l; 1.62 g/l and 1.10 g/l.

The virial coefficient derived from the seven curves by using the program VIRIAL is $B_1 = 24 \pm 1$ ml/g. This value exceeds the amount of the excluded volume nearly fivefold and reflects charge-induced repulsion forces in the solute that can only partially be screened by the electrolytes with an ionic strength of 0.1. Attempts to estimate also the third virial coefficient have not been successful. The values obtained are very small (about 0.1 ml/g) and uncertain with errors larger than the estimate. Therefore, all further examinations of non-ideality under the conditions used in the experiments were related to the second virial coefficient.

4.1.2. Influence of ionic strength on the second virial coefficient

In further experiments, the ionic strengths of the acidic protein solutions at pH 4.5 were increased by addition of NaCl. This clearly leads to a reduction of the second virial coefficient (Fig. 2). The data obtained nearly approach the amount of the excluded volume of lysozyme that is calcu-

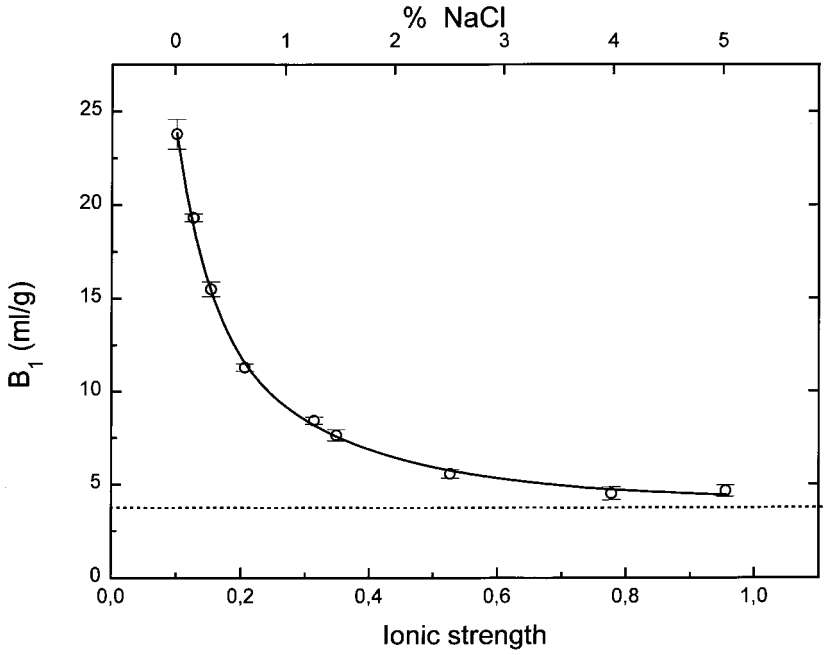

Fig. 2. Dependence of the second virial coefficient B_1 of lysozyme dissolved in 0.1 M acetate buffer, pH 4.5, on the ionic strength (lower abscissa) obtained by addition of NaCl (upper abscissa). The data were fitted using Eq. (26) (full line) resulting in a net charge number of 14.1 ± 1. The excluded volume (dotted line) amounts to 3.57 ml/g.

lated by the first part of Eq. (26) to be 3.57 ml/g when considering a $R_M = 1.72$ nm [25]. Furthermore, when all data of B_1 were fit as a function of the ionic strength on the basis of Eq. (26) the net charge of lysozyme at pH 4.5 was estimated at 14.1 ± 1.

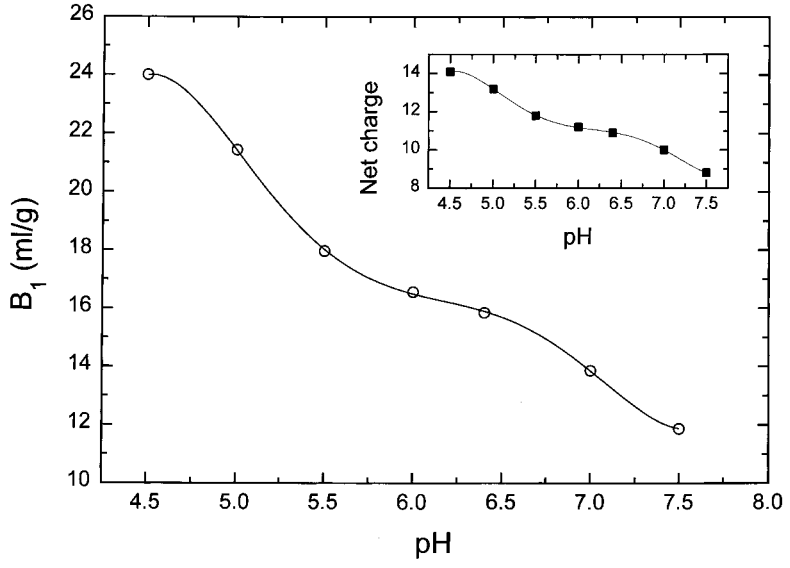

Fig. 3. Plots of the second virial coefficient B_1 and the net charge Z (insert) of lysozyme in 0.1 M Na-acetate buffer vs. pH. Conditions as given in Fig. 1.

4.1.3. Influence of pH on the virial coefficient

As expected, the high virial coefficient of lysozyme in weak acidic solution was reduced by increasing pH values (Fig. 3). This was due to a decrease of the net charge by moving the pH to the weak alkaline milieu. The values of B_1 and Z do not diminish regularly, but allow to recognize some steps reflecting the deprotonization of carboxylic and histidine groups. A further decrease of the second virial coefficient down to the value of the excluded volume at pH 9.32, the pI of lysozyme, has to be expected. However, under these conditions the solubility of the protein is strongly reduced.

4.2. Ovalbumin

With ovalbumin a further protein was investigated with respect to non-ideal behavior of the solution. Fig. 4 shows the radial concentration distribution curves of this protein at the sedimentation equilibrium. This experiment was carried out in 0.05 M potassium phosphate buffer, pH 8.5, at an ionic strength of 0.148. Because the pI of this protein was found to be 5.19, a larger virial coefficient is expected. When fitting the concentration distribution curves by Eq. (14) considering a molar mass of 42 970 for the monomeric protein the second virial coefficient B_1 was found to be 15.8 ± 1 ml/g. Assuming that ovalbumin has a hydrodynamic radius of 2.5 nm the excluded volume part of the virial coefficient is 3.7 ml/g. Therefore, the charge dependent term of the second virial coefficient results in a net charge number of $Z = 27 \pm 1$. Consequently, ovalbumin in phosphate buffer, pH 8.5, is a polyanion with a high negative charge. The large charge-dependent contribution over the excluded volume part of the virial coefficient suggests that the charges are less screened by the buffer ions.

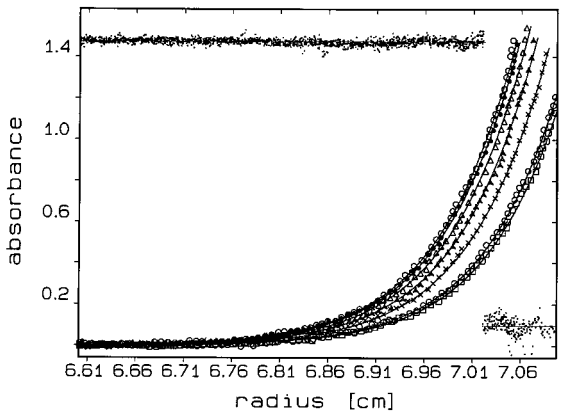

Fig. 4. Radial concentration distribution curves of ovalbumin dissolved in 0.05 M potassium phosphate buffer, pH 8.5. The profiles were recorded at sedimentation equilibrium reached after 46 h at 20 000 rev./min at a wavelength of $\lambda = 291$ nm. The curves were fitted using Eq. (14) and a molecular mass $M_1 = 42\,970$ Da. The real concentrations of the curve sections involved in the fitting procedure vary between 1.41 g/l (upper curve) and 1.03 g/l (lower curve).

5. Discussion

We have presented a method that allows determination of the second virial coefficient of monomeric proteins with relatively high precision. The accuracy of this method is based on the possibility of simultaneously fitting several radial concentration distributions in those sections where aggregates or other irregularities of the curves can be excluded. Furthermore, a proposed integration procedure allows us to determine the actual protein concentration involved in the fit. To reduce the error in calculation of B, it is advisable not to estimate the molecular mass but to hold this value constant. Using this method, the second virial coefficient was determined with deviations smaller than 5%. This situation allows us to derive some information about the excluded volume and the net charge of a protein and to demonstrate how the induced repulsive forces can be reduced by a shift in pH or screened by neutral salts added to the solution. This information is essential and may play an important role in protein crystallization.

The second virial coefficient has been studied in detail using lysozyme. At pH of 5.0 or lower, this protein exists in a monomeric state [26] and is able to crystallize well. As a basic protein (pI = 9.3), lysozyme possesses a relatively high net charge in the dilute acidic buffer as indicated by the large positive virial coefficient. By means of the Debye Hückel theory, plotting the B values against the ionic strength, we calculated an aver-

age net charge of 14.1 ± 1. This value is consistent with the expected one considering 18 positively charged residues of arginine, lysine and histidine or the nine aspartic and glutamic acidic groups, half of which should be protonated at this pH. Furthermore, as demonstrated by Sophianopoulus and Van Holde [26] lysozyme monomers contain an acidic ionizable group with a high $pK_a = 6.2$ which is possibly induced by the proximity of two carboxyl residues. Under these circumstances the charge of only a few positive amino acids is compensated by some negative carboxyl groups. The high net charge in lysozyme induces large repulsive forces that cause a high solubility as demonstrated by Howard et al. [27] and prevent formation of oligomers.

When adding small amounts of NaCl (or other neutral salts) to the protein solution, the B_1 values decrease dramatically up to salt concentrations of about 2% and then decrease gradually to data corresponding to the excluded volume. This value was calculated to 3.5 ml/g using Eq. (26) using a hydrodynamic radius of lysozyme of 1.72 nm [25]. The excluded volume includes the water shell surrounding the protein and is therefore somewhat larger than the value of 2.92 ml/g derived from the fourfold partial specific volume ($\bar{\nu} = 0.730$ ml/g, [28]). Rigorously, lysozyme has to be considered not as a sphere, but as a triaxial ellipsoid with axial ratios of 1.3 to 1.5 [29]. For such molecules Rallison and Harding [30] have proposed a slightly increased excluded volume compared with that of a sphere.

When salt is added to the protein solution (solvent/macromolecule) it becomes a three-component system that leads to a preferential binding to the macromolecule. Hamabata and von Hippel [31] concluded that these electrolytes bind mostly to the amide groups of proteins. However, 'inert' ions (e.g. Cl^-, Na^+, K^+) exhibit relative binding constants close to zero [31]. Pundah and Eisenberg [32] have analyzed the preferential binding of NaCl on bovine serum albumin to be 0.012 g/g. The confirmation of this low value was the reason for the formal treatment of the NaCl-containing protein solution as a two-component system. The excellent fit of the second virial coefficient vs. the ionic strength (Fig. 2) which describes the screening of the charges demonstrates indirectly that the preferential binding of NaCl by lysozyme under the chosen condition should be negligibly small.

The second virial coefficient of ovalbumin at pH 8.5 was determined to $B_1 = 15.1$ ml/g. As expected, this value exceeds that of 11.2 ml/g for isoelectric ovalbumin ($pI = 5.19$) derived by Jacobson and Winzor [16]. Because the charge-dependent term of the virial coefficient at the isoelectric point should be negligible in comparison with the excluded volume, the value obtained by Jacobson and Winzor [16] seems to be too high. As possible reasons for this deviation the lower solubility of the protein at the isoelectric point could be discussed as well as the missing control whether the loading concentration is identical with the real one at sedimentation equilibrium in their experiment. Additionally, these authors have used a molecular mass of 45 000 Da, a value clearly higher than the theoretical one of 42 970, derived from the amino acid composition including one bound *N*-acetylglucosamine [33].

The virial coefficient of lysozyme remains positive under the conditions of the experiments. Under the influence of neutral salts, the charge-induced repulsive forces are short compared to the diameter of the protein molecules. This allows the formation of oligomers or nuclei which are a prerequisite for crystallization. As pointed out by Riès-Kautt and Ducruix [34], this takes place in the presence of 1.8% NaCl or more. Under these conditions, we could demonstrate that the charge-dependent term of the second virial coefficient was smaller than the part responsible for the excluded volume (see Fig. 2).

In the pH range 6.0–8.0 where lysozyme dimerization can occur at higher protein concentrations, negative apparent B values could be expected. However, the protein crystallization was found to be least favorable [35]. This suggests the negative apparent virial coefficients to be an indicator for suitable crystallization conditions but not be considered as an implicit predictor as recently proposed by George and Wilson [13] or Rosenbaum and Zukowski [14].

Acknowledgements

The skilful technical assistance of Mrs Bärbel Bödner is gratefully acknowledged. The investigations were supported by the Deutsche Forschungsgemeinschaft (HE 1318/15-1).

References

[1] C. Tanford, Physical Chemistry of Macromolecules, Wiley, New York, 1991.
[2] S. Hanlon, K. Lamers, G. Lauterbach, R. Johnson, H.K. Schachman, Arch. Biochem. Biophys. 99 (1962) 157.
[3] H.K. Schachman, L. Gropper, S. Hanlon, F. Putney, Arch. Biochem. Biophys. 99 (1962) 175.
[4] E.T. Adams, Jr., Biochemistry 4 (1964) 1646.
[5] E.T. Adams, Jr., D.L. Filmer, Biochemistry 5 (1965) 2971.
[6] P.D. Jeffrey, J.H. Coates, Biochemistry 5 (1966) 489.
[7] D.A. Albright, J.W. Williams, Biochemistry 7 (1968) 67.
[8] E.T. Adams, Jr., M.S. Lewis, Biochemistry 7 (1968) 1044.
[9] D.K. Hancock, J.W. Williams, Biochemistry 8 (1969) 2598.
[10] R.C. Deonier, J.W. Williams, Biochemistry 9 (1970) 4260.
[11] S. Szuchet, D.A. Yphantis, Biochemistry 12 (1973) 5115.
[12] S. Szuchet, D.A. Yphantis, Arch. Biochem. Biophys. 173 (1976) 495.
[13] A. George, W.W. Wilson, Acta Cryst. D 50 (1994) 361.
[14] D.F. Rosenbaum, C.F. Zukowski, J. Cryst. Growth 169 (1996) 752.
[15] L.A. Holladay, A.K. Sophianopoulus, Anal. Biochem. 57 (1974) 506.
[16] M.P. Jacobson, D.J. Winzor, Biophys. Chem. 45 (1992) 119.
[17] M. Morris, G.B. Ralston, Biophys. Chem. 23 (1985) 49.
[18] M. Morris, G.B. Ralston, Biochemistry 28 (1989) 8561.
[19] J.W. Williams, K.E. Van Holde, R.L. Baldwin, H. Fujita, Chem. Rev. 58 (1958) 715.
[20] H. Fujita, Mathematical Theory of Sedimentation Analysis, Academic Press, New York, 1962.
[21] W.G. McMillan, J.E. Mayer, J. Chem. Phys. 13 (1945) 276.
[22] P.R. Wills, D.J. Winzor, in: S.E. Harding, A.J. Rowe, J.C. Horton (Eds.), Analytical Ultracentrifugation in Biochemistry and Polymer Science, Royal Society, Cambridge, UK, 1992, p. 311.
[23] S.C. Gill, P.H. von Hippel, Anal. Biochem. 182 (1989) 319.
[24] R.E. Canfield, J. Biol. Chem. 238 (1963) 2698.
[25] V.E. Mikol, E. Hirsch, R. Giegé, J. Mol. Biol. 213 (1990) 187.
[26] A.J. Sophianopoulus, K.E. Van Holde, J. Biol. Chem. 239 (1964) 2516.
[27] S.B. Howard, P.J. Twigg, J.K. Baird, E.J. Meehan, J. Cryst. Growth 90 (1988) 94.
[28] A. Schausberger, I. Pilz, Makromol. Chem. 178 (1977) 211.
[29] K. Harata, Acta Cryst. D 49 (1993) 497.
[30] J.M. Rallison, S.E. Harding, J. Colloid Interface Sci. 103 (1985) 284.
[31] A. Hambata, P.H. von Hippel, Biochemistry 12 (1973) 1264.
[32] S. Pundak, H. Eisenberg, Eur. J. Biochem. 118 (1981) 463.
[33] P.E. Stein, A.G.W. Leslie, J.T. Finch, R.W. Carrel, J. Mol. Biol. 221 (1991) 941.
[34] M.M. Riès-Kautt, A.F. Ducruix, J. Cryst. Growth 110 (1991) 20.
[35] A.J. Sophianopoulus, J. Biol. Chem. 244 (1969) 3188.