ELSEVIER

Biophysical Chemistry 87 (2000) 1–13

Biophysical Chemistry

www.elsevier.nl/locate/bpc

Analysis of protein self-association under conditions of the thermodynamic non-ideality

Joachim Behlke*, Otto Ristau

Max Delbrück Center for Molecular Medicine, Robert Rössle Strasse 10, D-13092 Berlin, Germany

Received 7 March 2000; received in revised form 4 May 2000; accepted 30 May 2000

Abstract

Analysis of protein–protein interactions in highly concentrated solutions requires a consideration of the non-ideality in such solutions which is expressed by the virial coefficients. Different equations are presented to estimate effects of the thermodynamic non-ideality on the macromolecular interaction of self-associating proteins in sedimentation equilibrium experiments. Usually the influence of thermodynamic non-ideal behavior are described by concentration power series. The convergence of such power series is limited at high solute concentration. When expressing the thermodynamic non-ideality by an activity power series this disadvantage can be minimized. The developed centrifuge equations are the basis for a global analysis to estimate equilibrium constants and the corresponding thermodynamic activities of the reactants. Based on fit analysis of synthetic concentration profiles it was established that marked deviations from the expected association constants are observed for proteins with strong association forces between solute molecules. Considerable differences were also observed in weakly interacting systems. This was due to the excluded volume of the protein which is similar in magnitude to the binding constant. For interactions with moderate affinities values extremely close to the true binding values were obtained, as confirmed by experimental results with concanavalin A.

Keywords: Sedimentation equilibrium; Virial coefficient; Excluded volume; Macromolecular interactions; Association constants; Concanavalin A

* Corresponding author. Tel.: +49-30-9406-2205; fax: +49-30-9406-2802.
E-mail address: behlke@mdc-berlin.de (J. Behlke).

0301-4622/00/$ - see front matter
PII: S0301-4622(00)00173-3

1. Introduction

The complex interaction between identical and non-identical macromolecules are increasingly studied in the analytical ultracentrifuge using the technique of sedimentation equilibrium (see Laue, Stafford [1], Rivas et al. [2], Behlke, Ristau [3]). Most previous experiments employed dilute solutions of macromolecules but, depending on the speed at sedimentation equilibrium, the concentration near the cell base can be increased 10-fold. In such highly concentrated solutions there are non-ideality effects, expressed by the virial coefficients, that must be considered when analyzing macromolecular interactions. Especially by using the high speed technique [4] a considerable concentration increase at the cell base can be observed. Furthermore, this increase is of advantage for estimation of higher virial coefficients which was the aim of this study.

Recently, we studied the thermodynamic non-ideality of monomeric lysozyme and ovalbumin expressed by the second virial coefficient [5]. This parameter was derived from sedimentation equilibrium experiments with loading concentrations in the grams per liter range and a speed that results in at least a fivefold concentration increase near the cell bottom. For these proteins the second virial coefficient is mainly influenced by the excluded volume and net charge. Since net charge can prevent self-associations, such charge effects can be minimized by addition of salt which decreases repulsive forces and reduces the virial coefficient allowing associations to occur [5].

Simultaneous consideration of the association constant and virial coefficient necessary for self-association of proteins under non-ideal conditions has been discussed by Hill and Chen [6]. These authors have used statistical thermodynamics to provide a thermodynamically precise description of aggregates (associates) in osmotic solution.

The expression for the osmotic pressure or the chemical potential is given by a concentration power series where the binding constants are included in the coefficients. This means, the binding forces themselves act as a source of non-ideality. However, this concept leads to two difficulties. (i) For high association constants the power series do not converge. (ii) The centrifuge equations derived for such approach contain only one exponential function. Usually one exponential function is employed for each size of associate [7].

When only considering a monomer–dimer equilibrium, the first difficulty can be eliminated by substituting an analytical function for the binding constant K_2 in the concentration power series, if the binding constant and the virial coefficients are not directly coupled by multiplication (such K values are here named 'pure binding constants'). As will be shown, this approach can also be applied to higher order associates. Removal of the pure binding constants from the concentration power series enables us to simultaneously transform the single-exponential term into a multi-exponential term expression.

In order to verify these considerations we have analyzed the self-association behavior of concanavalin A (Con A), a protein that forms dimers in solution which, at alkaline pH, are in equilibrium with tetramers [8]. In studies by Senear and Teller [9]; Huet [10]; and Huet and Claverie [11], association constants for this oligomerization reaction were estimated without consideration of virial effects. In order to calculate whether this is justified we have repeated sedimentation equilibrium experiments on Con A solutions in phosphate buffer containing different electrolytes. The aim of the present study was to determine to what extent association constants are modified by the virial coefficients. We also show how both of these parameters can be simultaneously calculated. This method requires the generation of concentration profiles and the convergence examination of the developed fit functions. Based on these theoretical considerations it was concluded that in solutions of moderate protein concentrations and moderate self-association, the association constants were not noticeably influenced by the virial coefficients. Deviations from expected binding constants were observed for highly concentrated solutions. In very dilute protein solutions no significant virial effects were observed. However, when the binding constants were the same order of magnitude as the excluded volume, strong deviations from the expected values occurred. These results indicated

that very weak as well as very strong binding constants can cause non-ideality.

2. Theory

2.1. Statistical thermodynamics of multicomponent solution

The McMillan–Mayer solution theory [12] yields the following concentration power series [Eq. (1)] for the osmotic pressure of a solution consisting of four solutes. In this equation the B values are the statistically defined virial coefficients, and the concentrations (c_n) are given in mol/l.

$$\begin{aligned}\frac{\pi}{RT} = {} & c_1 + c_2 + B_{20}c_1^2 + B_{30}c_1^3 + B_{40}c_1^4 + B_{02}c_2^2 \\ & + B_{11}c_1c_2 + B_{21}c_1^2c_2 + B_{50}c_1^5 + B_{12}c_1c_2^2 \\ & + B_{31}c_1^3c_2 + B_{201}c_1^2c_3 + B_{1001}c_1c_4 \\ & + B_{011}c_2c_3 \ldots \end{aligned} \quad (1)$$

When assuming chemical equilibria between the components it is important initially to construct an activity series with small b values as coefficients. Only in this way can the equilibria be precisely considered with respect to activities (z_n) as measures of the concentrations and results in binding constants being in components of the virial series [Eq. (5)].

$$\begin{aligned}\frac{\pi}{RT} = {} & z_1 + z_2 + z_3 + z_4 + b_{20}z_1^2 + b_{30}z_1^3 + b_{40}z_1^4 \\ & + b_{02}z_2^2 + b_{11}z_1z_2 + b_{21}z_1^2z_2 + b_{50}z_1^5 \\ & + b_{12}z_1z_2^2 + b_{31}z_1^3z_2 + b_{201}z_1^2z_3 + b_{1001}z_1z_4 \\ & + b_{011}z_2z_3 \ldots \end{aligned} \quad (2)$$

In both functions the first position at the indices of (B)bs refers to the number of molecules of the first component involved in an associate. The second position reflects the number of involved molecules of the second component, etc. The relationships between small and capital bs can be derived by transferring the concentrations of components c_n of series 1 into the corresponding activities z_n. According to statistical thermodynamics [6] the following relation is valid for the total concentration of components:

$$c_n = z_n \frac{\partial(\pi/RT)}{\partial z_n}$$

$$\begin{aligned} c_1 = {} & z_1 + 2b_{20}z_1^2 + 3b_{30}z_1^3 + 4b_{40}z_1^4 + b_{11}z_1z_2 \\ & + 2b_{21}z_1^2z_2 + 5b_{50}z_1^5 + b_{12}z_1z_2^2 + 3b_{31}z_1^3z_2 \\ & + 2b_{201}z_1^2z_3 + b_{1001}z_1z_4 \ldots \end{aligned}$$

$$\begin{aligned} c_2 = {} & z_2 + 2b_{02}z_2^2 + b_{11}z_1z_2 + b_{21}z_1^2z_2 + 2b_{12}z_1z_2^2 \\ & + b_{31}z_1^3z_2 + b_{011}z_2z_3 \ldots \end{aligned}$$

$$c_3 = z_3 + b_{201}z_1^2z_3 + b_{1001}z_1z_4 + b_{011}z_2z_3 \ldots$$

$$c_4 = z_4 + b_{1001}z_1z_4 \ldots \quad (3)$$

Substitution of the identities in Eq. (3) into Eq. (1) leads to the relationships between small and capital bs by equating coefficients of like powers of z with Eq. (2).

$$\begin{aligned} & b_{20} = -B_{20}\ ,\ 2b_{30} = 4B_{20}^2 - B_{30}, \\ & \qquad 3b_{40} = -16B_{20}^3 + 9B_{20}B_{30} - B_{40} \\ & b_{02} = -B_{02},\ b_{11} = -B_{11}, \\ & \qquad 2b_{21} = 4B_{11}B_{20} + B_{11}^2 - B_{21} \\ & 2b_{12} = 4B_{11}B_{02} + B_{11}^2 - B_{12}, \\ & \qquad 24b_{50} = -6B_{50} + 400B_{20}^4 - 360B_{20}^2B_{30} \\ & \qquad\quad + 64B_{20}B_{40} + 27B_{30}^2 \\ & 6b_{31} = -2B_{31} + 3B_{21}B_{11} + 9B_{30}B_{11} + 12B_{20}B_{21} \\ & \qquad - B_{11}^3 - 12B_{20}B_{11}^2 - 36B_{20}^2B_{11} \end{aligned} \quad (4)$$

2.2. Equilibria between components

Assuming chemical equilibria between the components according to $z_n = K_n z_1^n$ from Eq. (2) one obtains:

$$\frac{\pi}{RT} = z_1 + (b_{20} + K_2) z_1^2 + (b_{30} + b_{11} K_2) z_1^3 + \left(b_{40} + b_{21} K_2 + b_{02} K_2^2\right) z_1^4 + \left(b_{50} + K5 + b_{12} K_2^2 + b_{31} K_2 + b_{201} K_3 + b_{1001} K_4 + b_{011} K_2 K_3\right) z_1^5 ... \quad (5)$$

or in an abbreviated form:

$$\frac{\pi}{RT} = z_1 + b_2 z_1^2 + b_3 z_1^3 + b_4 z_1^4 + b_5 z_1^5 .... \quad (6)$$

For the total concentration $c = c_1 + 2c_2 + 3c_3 + 4c_4$ one obtains in analogous fashion to Eq. (3)

$$c = z_1 + 2b_2 z_1^2 + 3b_3 z_1^3 + 4b_4 z_1^4 + 5b_5 z_1^5 \quad (7)$$

The one-component activity series (small bs) is now changed into a concentration series (capital Bs). This is done by inverting series [Eq. (7)] relative to z_1 and incorporating the result in Eq. (6) followed by equating coefficients of like concentration powers with a corresponding concentration series for π/RT and the capital Bs as coefficients resulting in:

$$B_2 = -b_2, \; B_3 = -2b_3 + 4b_2^2,$$
$$B_4 = -3b_4 - 20b_2^3 + 18b_2 b_3$$
$$B_5 = -4b_5 - 144b_2^2 b_3 + 18b_3^2 + 32b_2 b_4 + 112b_2^4 \quad (8)$$

Because the small bs in Eqs. (6)–(8) correspond to the coefficients in Eq. (5) it can be defined:

$$\frac{\pi}{RT} = c + (B_{20} - K_2)c^2 + \left[B_{30} - 2K_2(4B_{20} - B_{11}) + 4K_2^2 - 2K_3\right]c^3 + \left[B_{40} + K_2\left(12B_{11}B_{20} - 24B_{20}^2 - 9B_{30} - 1.5B_{11}^2 + 1.5B_{21}\right) + 3K^2(20B_{20} - 6B_{11} + B_{02}) - 3K_3(6B_{11} - B_{101}) - 20K_2^3 - 3K_4\right]c^4 .. \quad (9)$$

Neglecting molecular interactions by setting all capital Bs $= 0$, and expanding until c^5 can be derived yields:

$$\frac{\pi}{RT} = c\left(1 - K_2 c + 4K_2^2 c^2 - 20K_2^3 c^3 + 112K_2^4 c^4 - 672K_2^5 c^5 ...\right) \quad (10)$$

The convergence properties of this concentration series are clearly unsatisfactory, particularly for sedimentation equilibrium experiments with strongly increasing concentrations near the cell base. A partial improvement can be obtained by an analytical treatment of the thermodynamically ideal system [all $(B)b$s are zero].

2.3. Restriction to two solute components in chemical equilibrium

By initially considering only two components the osmotic pressure expressed in activity form means:

$$\frac{\pi}{RT} = z_1 + K_2 z_1^2 \quad (11)$$

The total concentration c results in:

$$c = z_1 + 2K_2 z_1^2 \quad (12)$$

Rearranging Eq. (12) yields:

$$z_1 = \frac{1}{4K_2}\left[-1 + \sqrt{1 + 8K_2 c}\right] \quad (13)$$

Substitution of z_1 Eq. (13) into Eq. (11) leads to:

$$\frac{\pi}{RT} = \frac{1}{4K_2}\left[-1 + \sqrt{1 + 8K_2 c}\right] + \frac{1}{16K_2}\left[1 - 2\sqrt{1 + 8K_2 c} + 1 + 8K_2 c\right]$$

$$\frac{\pi}{RT} = \frac{c}{2} - \frac{1}{8K_2} + \frac{1}{8K_2}\sqrt{1+8K_2c} \tag{14}$$

By developing Eq. (14) into a power series based on K_2, one obtains the complete series in Eq. (10). Therefore, by means of Eq. (14) we can substitute the pure K_2 terms in series in Eq. (9). The remaining part of the terms that describe the non-ideal molecular interactions can be considered as an additive term. Eq. (9) can be changed to:

$$\begin{aligned}\frac{\pi}{RT} = {} & \frac{1}{2}c + \frac{1}{8K_2}\left(\sqrt{1+8K_2c}-1\right) + B_{20}c^2 \\ & + [B_{30} - 2K_2(4B_{20} - B_{11})]c^3 \\ & + \Big[B_{40} + K_2\left(12B_{11}B_{20} - 24B_{20}^2 - 9B_{30}\right. \\ & \left. - 1.5B_{11}^2 + 1.5B_{21}\right) + 3K_2^2 \\ & \times (20B_{20} - 6B_{11} + B_{02})\Big]c^4\end{aligned} \tag{15}$$

This additive concentration power series describes the thermodynamic non-ideality converge much better than Eq. (9) and with restriction $K_3 = K_4 = 0$. This is due to the separation of the pure K_2 values. Therefore, it is important to transfer the improved convergence properties for the osmotic pressure into the chemical potential equation.

For the osmotic equilibrium the following thermodynamic relation is valid:

$$n_1\left(\frac{\partial\mu_1}{\partial c}\right)_{T,\mu_0} = V_S\left(\frac{\partial\pi}{\partial c}\right)_{T,\mu_0} \tag{16}$$

Here μ_0 means the chemical potential of the solvent and V_S the solvent volume (not solution volume) per n_1 mole solute. The decrease of the pure solvent volume related to the solution volume by the solute concentration can be approximately determined by using Eq. (17) for calculation of the partial chemical potential (see Ross and Minton [13], Behlke and Ristau [5]).

$$\left(\frac{\partial\mu_1}{\partial c}\right)_{T,P} = \frac{(1-\bar{v}_m c)}{c}\left(\frac{\partial\pi}{\partial c}\right)_{T,P} \tag{17}$$

In this equation $\bar{v}_m$ is the partial molar volume of the solute and c the concentration in mole solute per liter solution. Prerequisites for this correction term are the incompressibility and concentration independence of the partial specific volumes of solvent and solute [5]. Furthermore, it is to assumed the partial molar volume of homologous associates is identical with that of monomers. By accepting these corrections and assumptions, the difference between the equilibrium conditions $\mu_0 = \text{constant}$ (osmotic equilibrium) and $P = \text{constant}$ (constant pressure) disappears. Formulation of the chemical potential in the corrected molar-B_n mode is now equivalent to the molal-C_n mode.

Usually the term $(1-\bar{v}_m c)$ is omitted when considering the behavior of dilute solutions. For the further development of our point we will also momentarily eliminate this term. The reason is that an analogous term with the same assumptions approximately describes the change of the solution density in the centrifuge cell. Therefore, this term is eliminated after insertion of Eq. (17) into Eq. (19).

With this simplification [omitting the factor $(1-\bar{v}_m c)$] of Eq. (17) it follows:

$$\begin{aligned}\frac{1}{RT}\left(\frac{\mathrm{d}\mu_1}{\mathrm{d}c}\right)_{T,P} = {} & \frac{1}{2} + \frac{1}{2\sqrt{1+8K_2c}} + 2B_{20}c^2 \\ & + 3(B_{30} - 2K_2(4B_{20} - B_{11}))c^2 \\ & + 4\Big[B_{40} + K_2 \\ & \times\left(12B_{11}B_{20} - 24B_{20}^2 - 9B_{30}\right. \\ & \left. - 1.5B_{11}^2 + 1.5B_{21}\right) \\ & + 3K_2^2(20B_{20} - 6B_{11} + B_{02})\Big]c^3\end{aligned} \tag{18}$$

The centrifuge equation according to Williams et al. [14] means:

$$M(1-\bar{v}\rho_0)(1-\bar{v}_m c)\omega^2 r = \left(\frac{\partial\mu_1}{\partial c}\right)_{T,P}\frac{\mathrm{d}c}{\mathrm{d}r} \tag{19}$$

The factor $(1-\bar{v}_m c)$ on the left side describes the density modification of the solution [15]. This

correction originally developed for single solute systems is approximately valid also for homologous association systems with the above mentioned prerequisites. As shown in Eq. (17) an analogue factor converts the ratio mole solvent to mole solute, thus the factor reduces when inserting Eq. (18) into Eq. (19). This corresponds with the earlier conclusion by Wills and Winzor [16] and Behlke and Ristau [5]. After subsequent integration one obtains:

$$c = c_0 \exp\left(\frac{M(1-\bar{v}\rho_0)\omega^2}{2RT}\left(r^2 - r_0^2\right) + \ln\left(\frac{1+\sqrt{1+8K_2c}}{1+\sqrt{1+8K_2c_0}} \right) - \ln(\gamma_Z) \right) \quad (20)$$

with the virial coefficients series

$$\begin{aligned} \ln(\gamma_Z) = {} & 2B_{20}(c-c_0) + \frac{3\left(c^2-c_0^2\right)}{2} \\ & \times [B_{30} - 2K_2(4B_{20} - B_{11})] \\ & + \frac{4\left(c^3-c_0^3\right)}{3} \\ & \times \Big[B_{40} + K_2\left(12B_{11}B_{20} - 24B_{20}^2 - 9B_{30}\right. \\ & \left. - 1.5B_{110}^2 + 1.5B_{21}\right) \\ & + 3K_2^2(B_{02} + 20B_{20} - 6B_{11})\Big] \end{aligned} \quad (21)$$

For Eq. (20) one can write:

$$c = \frac{1+\sqrt{1+8K_2c}}{2} c_i \exp\left(\frac{M(1-\bar{v}\rho_0)\omega^2}{2RT} \times \left(r^2 - r_0^2\right) - \ln(\gamma_Z) \right) \quad (22)$$

and with the abbreviation

$$E = \exp\left(\frac{M(1-\bar{v}\rho_0)\omega^2}{2RT}\left(r^2 - r_0^2\right) - \ln(\gamma_Z) \right) \quad (23)$$

and rearrangement of Eq. (22) we obtain:

$$2c - c_i E = c_i E\sqrt{1+8K_2c} \quad (24)$$

Squaring of both sides

$$4c^2 - 4cc_iE + c_i^2E^2 = c_i^2E^2(1+8K_2c) \quad (25)$$

leads to a squared equation (the solution $c = 0$ was excluded)

$$c = c_iE + 2K_2c_i^2E^2 \quad (26)$$

Substitution of Eq. (23) into Eq. (26) leads to the following relation:

$$\begin{aligned} c = {} & c_i \exp\left(\frac{M(1-\bar{v}\rho_0)\omega^2}{2RT}\left(r^2 - r_0^2\right) - \ln(\gamma_Z) \right) + 2K_2c_i^2 \exp\left(\frac{2M(1-\bar{v}\rho_0)\omega^2}{2RT} \times \left(r^2 - r_0^2\right) - 2\ln(\gamma_Z) \right) \end{aligned} \quad (27)$$

When omitting the expression for the activity coefficients, $\ln(\gamma_Z)$, Eq. (27) is the classical centrifuge equation for a monomer dimer equilibrium without non-ideal interaction of the molecules. Here the integration constant c_i corresponds to the monomer concentration.

$$c_i = \frac{2c_0}{1+\sqrt{1+8K_2c_0}} = c_{\text{monomer}} \quad (28)$$

The consideration of non-ideality by means of the activity coefficients $\ln(\gamma_Z)$ is related to the total concentration $c = c_M + 2c_D$. The concentrations of monomers and dimers themselves are not thermodynamic quantities. In order to obtain the monomer concentration we need the apparent equilibrium constant which can be calculated from the true equilibrium constant according to Hill and Chen [5].

$$\begin{aligned} K_{\text{app}} = \frac{K_2}{1 + (4B_{20} - B_{11})c_0 + [2K_2(2B_{11} - 4B_{20} - B_{02}) + 8B_{20}^2 + 3B_{30} + 0.5B_{11}^2 - 0.5B_{21} - 4B_{11}B_{20}]c_0^2 ..} \end{aligned} \quad (29)$$

Assuming the chosen reference radius is identical with the meniscus position, only small differences between the apparent and true equilibrium constant are observed. Generally, this is due to the very low concentration c_0 at the meniscus.

Eq. (27) is the single equation that is generally used for centrifugal analysis of each solute species (monomer, dimer) for ideal systems. This equation reflects a rigorous derivation based on statistical thermodynamics.

2.4. Multisolute components at chemical equilibrium

The above mentioned approach cannot be applied to more complex associates without a modification. Eqs. (22) and (23) can be rearranged to

$$\frac{2c}{1+\sqrt{1+8K_2c}} = c_iE \tag{30}$$

For Eq. (30) we can then write:

$$cF(c) = c_iE \tag{31}$$

The function F is defined as a power series in c with the association constants as their coefficients. Substitution of Eq. (31) in the usual centrifuge equation for a monomer–dimer system Eq. (32)

$$c = c_iE + 2K_2c_i^2E^2 \tag{32}$$

yields

$$c = cF + 2K_2c^2F^2 \quad \text{or} \quad 1 = F + 2K_2cF^2 \tag{33}$$

F can be calculated as a concentration power series by using the method of uncertain coefficients

$$F = 1 - 2K_2c + 8(K_2c)^2 - 40(K_2c)^3 + 224(K_2c)^4 - 1344(K_2c)^5 \ldots\ldots \tag{34}$$

This expression is accurate the Taylor series of the function Eq. (35) [see also Eq. (30)]

$$F = \frac{2}{1+\sqrt{1+8K_2c}} \tag{35}$$

This approach can be applied to more complex associates. Considering a monomer–dimer–trimer system (mass balance):

$$c = c_iE + 2K_2c_i^2E^2 + 3K_3c_i^3E^3 \tag{36}$$

Substitution of Eq. (31) for c_iE yields

$$1 = F + 2K_2cF^2 + 3K_3c^2F^3 \tag{37}$$

For F one obtains the following power series in c

$$F = 1 - 2K_2c + \left(8K_2^2 - 3K_3\right)c^2 - \left(40K_2^3 - 30K_2K_3\right)c^3 + \left(224K_2^4 - 252K_2^2K_3 + 27K_3^2\right)c^4 \ldots. \tag{38}$$

To get the correct dimension for the usual activity series the natural logarithm of F is necessary:

$$\ln(F) = -2K_2c + \left(6K_2^2 - 3K_3\right)c^2 - \left(\frac{80}{3}K_2^3 - 24K_2K_3\right)c^3 + \left(140K_2^4 - 180K_2^2K_3 + \frac{45}{2}K_3^2\right)c^4 \tag{39}$$

With respect to the pure K_2 and K_3 terms and their products this is the same series as derived by Hill and Chen [5]. After removing all these pure K-terms one obtains the following series for the activity coefficients of a monomer–dimer–trimer system:

$$\ln(\gamma) = 2B_{20}(c - c_0) + \frac{3\left(c^2 - c_0^2\right)}{2} \times [B_{30} - 2K_2(4B_{20} - B_{11})] + \frac{4\left(c^3 - c_0^3\right)}{3} \times \left[B_{40} + K_2\left(12B_{11}B_{20} - 24B_{20}^2 - 9B_{30}\right.\right.$$

$$-1.5B_{110}^2 + 1.5B_{21}\Big)$$
$$+3K_2^2(B_{02} + 20B_{20} - 6B_{11})$$
$$-3K_3(6B_{11} - B_{101})] \quad (40)$$

Compared with the monomer–dimer system [Eq. (21)], only a K_3 expression is added to the last term. The multi-exponential solution now reads:

$$c = c_i\exp\left(\frac{M(1-\bar{v}\rho_0)\omega^2}{2RT} - \ln(\gamma)\right) + 2K_2c_i^2\exp\left(\frac{2M(1-\bar{v}\rho_0)\omega^2}{2RT} - 2\ln(\gamma)\right) + 3K_3c_i^3\exp\left(\frac{3M(1-\bar{v}\rho_0)\omega^2}{2RT} - 3\ln(\gamma)\right) \quad (41)$$

In principle, one can convert the usual multi-exponential into a single-exponential centrifuge equation. The multi exponential centrifuge equation presented in this paper is the thermodynamically exact variant of the equation derived by Johnson et al. [7].

The possibility to remove the pure association constants from the generalized virial coefficients series according to Hill and Chen [6] is based on the fact, that the non-ideality being described only by additive terms to the ideal variant. As in the classical multi-exponential equation, in the generalized single-exponential centrifuge equation the mass balance of all associates is involved. This balance is done by an activity power series which reduces to the classical associate series if all virial coefficients (the Bs) are zero.

A further possibility to describe concentration profiles is provided by Eq. (7). This function describes the total concentration by an activity power series. A consideration of the osmotic pressure presented by a concentration power series is thus not necessary. The activity in statistical thermodynamics is defined for a concentration scale related to volume solution (dilute solutions) [6].

$$\pi = RT\ f(z) \text{ and } c = z\left(\frac{\partial f}{\partial z}\right) \quad (42)$$

Therefore, the same correction term for higher concentrations is used in the Gibbs–Duham equation as in Eq. (17).

$$\left(\frac{\partial\mu_1}{\partial c}\right)_{P,T} = \frac{(1-\bar{v}_m c)}{c}\left(\frac{\partial\pi}{\partial c}\right)_{P,T} = \frac{(1-\bar{v}_m c)}{c}\left[\left(\frac{\partial\pi}{\partial z}\right)\frac{\partial z}{\partial c}\right]_{P,T} = \frac{RT(1-\bar{v}_m c)}{z\left(\frac{\partial f}{\partial z}\right)_{P,T}}\left[\left(\frac{\partial f}{\partial z}\right)\frac{\partial z}{\partial c}\right]_{P,T}$$
$$\left(\frac{\partial\mu_1}{\partial c}\right)_{P,T} = RT\frac{(1-\bar{v}_m c)}{z}\left(\frac{\partial z}{\partial c}\right)_{P,T} \quad (43)$$

After substitution in the centrifuge equation [Eq. (19)] and integration the result is:

$$\mu_1 = \mu_1^0 + RT\ln\ z \quad (44)$$

The correction term $(1-\bar{v}_m c)$ is again eliminated. For the dependence of z on the centrifugal force an exponential expression, the so-called ψ function [17] is obtained:

$$z = z_0\exp\left(\frac{M(1-\bar{v}\rho_0)\omega^2\left(r^2 - r_0^2\right)}{2RT}\right) \quad (45)$$

Substitution of the equality of z from Eq. (45) into Eq. (7) and replacing small bs with capital Bs results in:

$$c = z + 2(K_2 - B_{20})z^2 + 3\left(K_3 + 2B_{20}^2 - 0.5B_{30} - B_{11}K_2\right)z^3 + 4\Big(K_4 + 3B_{20}B_{30} - \frac{16}{3}B_{20}^3 - \frac{B_{40}}{3} + K_2\left(2B_{11}B_{20} + 0.5B_{11}^2 - 0.5B_{21}\right) - K_2^2B_{02} - K_3B_{101}\Big)z^4 \quad (46)$$

This direct approach was first suggested by Wills et al. [18].

When all Bs are zero the classical multi-exponential centrifuge equations for ideal systems are obtained. On the other hand, if all association

constants are zero a multi-exponential equation for a single solute system is obtained.

By changing to weight concentrations (w), the following relations are derived:

$$K_n = \frac{M^{n-1}}{n} K_n^w,\ c = \frac{w}{M},\ z = \frac{z_w}{M}$$

$$B_{n,m,l} = M^{n+m+l-1} \times B_{n,m,l}^w \tag{47}$$

$B_{n,m,l}^w$ and K_n^w denote the virial coefficients or binding constants related to weight concentration. The powers of the molecular mass M presented in Eq. (47) reduces by substitution in the centrifuge equation.

3. Material and methods

Con A (Sigma, St Louis, MO, USA) was dissolved in 1% NH_4HCO_3, and incubated at 37°C for 16 h according to Cunningham et al. [19]. After removing the precipitate by centrifugation the supernate was exhaustedly dialyzed against water and finally 0.14 M K-phosphate buffer, pH 7.5, containing 0.2 mM $CaCl_2$. Protein concentration was determined using an extinction coefficient of 1.37 for a 1 mg/ml solution at 280 nm [20].

Sedimentation equilibrium runs were performed with an XL-A ultracentrifuge (Beckman Instruments, Palo Alto, CA, USA). Experiments were carried out in conventional double sector cells using an eight-hole rotor. Approximately 300 μl protein graduated in concentrations of 1–2 mg/ml in seven cells were centrifuged for 2 h at 18 000 rev./min (overspeed) followed by 48–50 h equilibrium speed at 14 000 rev./min and 10°C. The radial concentration distributions were recorded at 298 nm and analyzed by global fitting procedure.

To study the influence of the non-ideality on the binding constant, mostly exact radial concentration profiles were generated. This was due by developing an additional term to Eq. (46). The additional term was:

$$\begin{aligned}
&+5\Big[K_5 - B_{1001}K_4 + \big(2B_{20}B_{101} + 1/2B_{101}^2 \\
&\quad - 1/2B_{201}\big)K_3 - B_{011}K_2K_3 \\
&\quad + \big(2B_{02}B_{11} + 1/2B_{11}^2 - 1/2B_{12}\big)K_2^2 \\
&\quad + (0.5B_{21}B_{11} + 1.5B_{30}B_{11} + 2B_{21}B_{20} \\
&\quad - 1/6B_{11}^3 - 2B_{11}^2B_{20} - 6B_{11}B_{20}^2 - 1/3B_{31}\big) \\
&\quad \times K_2 - 0.25B_{50} + 50/3B_{20}^4 - 15B_{20}^2B_{30} \\
&\quad + 8/3B_{20}B_{40} + 9/8B_{30}^2\Big]z^5
\end{aligned} \tag{48}$$

For the simulation only monomers and dimers were considered ($K_3 = K_4 = K_5 = 0$). For this simulation the molecules were considered as hard core bodies. The excluded volumes were calculated according to Boublik et al. [21–23]. The dimers were considered as prolate spherocylinders with a volume twice that of the monomer spheres. The following coefficients were obtained.

$$B_{11} = 1.5B_{20},\ B_{30} = 0.625B_{20}^2$$

$$B_{02} = 2.167B_{20},\ B_{40} = 0.287B_{20}^3$$

$$B_{12} = 1.727B_{20}^2,\ B_{31} = 0.4844B_{20}^3$$

$$B_{21} = 1.0035B_{20}^2,\ B_{50} = 0.1103B_{20}^4 \tag{49}$$

The last two coefficients B_{31} and B_{50} were only required for the additional term [Eq. (48)]. For calculation of the virial coefficients of the monomers (hard spheres), the exact formulas were used.

$$B_{20} = 4V,\ B_{30} = \frac{5}{8}B_{20}^2,$$

$$B_{40} = \left(\frac{2707}{4480} + \frac{438\sqrt{2} - 4131\arccos(1/3)}{4480\pi}\right) \times B_{20}^3 \approx 0.286950B_{20}^3 \tag{50}$$

Since there is no accurate analytical equation for B_{50} the following value was used:

$$B_{50} = (0.110252 \pm 0.000001) B_{20}^4 \quad (51)$$

For the third cross virial coefficients of prolate spherocylinders the following approximation has been employed [23]:

$$B_{i,j,k} = \frac{1}{3}\Big(V_i V_j + V_i V_k + V_j V_k + V_i(R_j S_k + R_k S_j) + V_j(R_i S_k + R_k S_i) + V_k(R_i S_j + R_j S_i) + \frac{1}{4\pi} G_{i,j,k} \Big) \quad (52)$$

with

$$G_{i,j,k} = \frac{2\pi}{3}\left(\left(R_i^2 + \frac{S_i}{4\pi} \right) S_j S_k + \left(R_j^2 + \frac{S_j}{4\pi} \right) S_i S_k + \left(R_k^2 + \frac{S_k}{4\pi} \right) S_i S_j \right) \quad (53)$$

For the fourth cross virial coefficients an analogous but only roughly approximate equation is available. For details see Eq. (9) in the paper of Boublik [21]. In these equations V means the volume, S the surface area and R the mean curvature integral of the bodies divided by 4π.

A program was written which allowed us to fit concentration profiles with the new Eq. (46) as well as with the classical Eq. (41). In order to reduce the number of estimated parameters, the meniscus activities (at reference position) were directly calculated (iteratively) from the loading concentrations in an analogous way to the procedure published in Behlke and Ristau [5]. The loading concentration was obtained both by direct integration of Eq. (46) (important for the calculation of the meniscus activities) and by numeric integration of the experimental or simulated concentration profiles. This was done by fitting an auxiliary function to the profiles with subsequent integration. The following auxiliary function was used

$$c_r = p_1 \exp\left(p_2 \left(r^2 - r_b^2 \right) \right) + p_3 \exp\left(p_4 \left(r^2 - r_b^2 \right) \right) + p_5 \exp\left(p_6 \left(r^2 - r_b^2 \right) \right) + p_7 r + p_8 \quad (54)$$

The fitting procedure is restricted so that only the parameters p_7 and p_8 can adopt negative values.

The criterion for a sufficient accuracy of Eq. (46) was the exact fit of the simulated concentration profiles was judged provided that all parameters were held constant at the theoretical values.

4. Results

4.1. General remarks

Determination of thermodynamically exact binding constants requires employment of the virial series in powers of concentration. To use such a series there must be convergence. The single-exponential equation has very bad convergence properties because the binding constants are directly involved in the concentration power series. In contrast, the classical multi-exponential functions are more useful because of separation of the pure binding constants from the virial series. Moreover, it is necessary to consider that there are also the multiplication products of virial coefficients and binding constants that occur in the concentration series. Therefore, the convergence for large binding constants is also insufficient. As will be shown later the more direct procedure expressed in Eq. (46) offers the best convergence properties, perhaps because of the absence of concentration powers in this function.

4.2. Fitting of theoretical concentration profiles

The influence of virial coefficients on estimation of association constants depends on both the concentration as well as the magnitude of the association constant. To roughly distinguish between these two influences we used the quantity $\delta = (K_2^w - B_{20}^w)w$. For values of δ between 0.1 and 1 the influence of the non-ideality of the equilibrium constants was rather small (see Fig. 1). For δ values greater than ≈ 2 the correct values of both the association constant and the virial coefficient could be simultaneously estimated. However, the value of δ decreased directly depending on the amount of noise. In con-

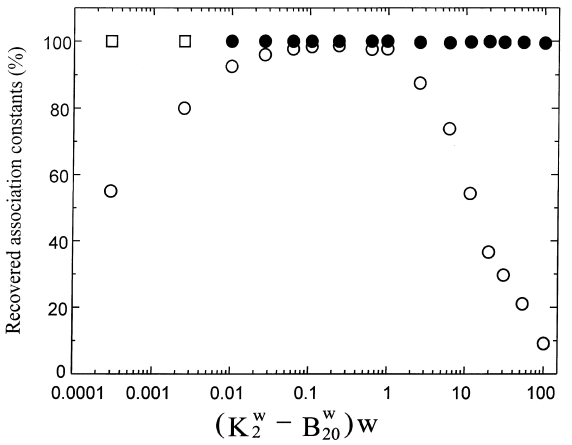

Fig. 1. Recovered association constants (%) depending on the loading concentration (expressed by the term $\delta = (K_2^w - B_{20}^w)$ estimated by our fit program based on Eq. (46). Fits were done with (●) and without (○) simultaneous consideration of the thermodynamic non-ideality. (□) The theoretical concentration profiles with the corresponding association constants were calculated taking into account virial coefficients based on the excluded volume according to Boublik et al. [21–23]. To get high accuracy of the theoretical curves simulation was done using an improved Eq. (46) by including Eq. (48).

trast, for δ smaller than ≈ 0.1, especially near the value of B_{20}, neither virial coefficients nor the association constants could be accurately estimated. Generally the estimates were too low. In this case the virial coefficients (mostly the second) should be known and incorporated into the program.

The classical Eqs. (27) and (41) for systems where association occurs have more limited validity. Thus, these equations are useful only for δ values smaller than 1. The same restrictions with respect to the influence of non-ideality are valid as mentioned above.

In all cases exact estimation of all virial coefficients was not possible. Estimation of the higher coefficients are possible, only if the ratios defined in Eq. (49) are held constant.

4.3. Analysis of experimental curves of Con A

Sedimentation equilibrium experiments were carried out on Con A in 0.14 M K-phosphate buffer, pH 7.5, containing 0.2 mM $CaCl_2$. To reduce charge effects on the virial coefficient increasing amounts of NaCl were added. The profiles were fitted with Eq. (46). With increasing salt concentration the estimated dimer–tetramer association constants slightly increased and remained constant at 0.2 M NaCl. Fig. 2 shows concentration distribution profiles of Con A under the various conditions. The association constant without consideration of the non-ideality was (9.0 ± 0.6) l/g. Simultaneous estimation of accurate binding constant and virial coefficients was impossible. This was also true for the constant ratios among the different virial coefficients as defined in Eq. (49). Considering the theoretical virial coefficient B_{20} and the ratios defined in Eq. (49), the binding constant increased up to (9.5 ± 0.6) g/l. This indicated for the chosen conditions the influence of the non-ideality on the binding constant was relatively small. This was in accord with results in Fig. 1, since the loading concentration in this experiment was approximately $w = 0.2$ g/l yielding a δ value of 2. The virial coefficient $B_{20} = 0.00035$ g/l used for the calculation is somewhat higher than four times the value [24] of the partial specific volume (0.74 ml/g).

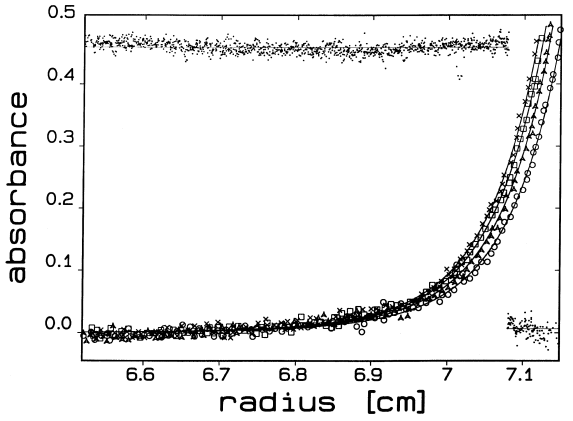

Fig. 2. Radial concentration distribution curves of Con A dissolved in 0.14 M K-phosphate buffer, pH 7.5, containing 0.2 mM $CaCl_2$ and 0.2 M NaCl. The profiles were recorded at 14 000 rev./min at a wavelength of $\lambda = 298$ nm. The curves were fitted using Eq. (46).The measured concentration of the curve sections used for the fitting procedure was approximately 0.2 g/l. From the seven curves used for fit only each second one is presented. Residues are given in 1.5-fold amplification related to the concentration profiles.

5. Discussion

In principle, the problem of estimating binding constants under conditions of non-ideality was solved by Hill and Chen [6]. However, the single-exponential solution derived by Hill and Chen is problematic with respect to its bad convergence for proteins at moderate concentrations and with moderate association constants. This occurs because chemical forces are taken as part of the non-ideality. Here we demonstrated the possibility to transfer the single-exponential function into the classical multi-exponential function thus signifying applicability of the thermodynamically exact form of the equations of Johnson et al. [7]. However, the convergence properties of these equations are not always sufficient because of products obtained by multiplying large binding constants with virial coefficients. Furthermore, the increasing concentration near the cell bottom can result in exceeding the convergence limit of the virial expansion series. Because of these difficulties it is preferable to describe the concentration by an activity series as is usually done by statistical thermodynamics. This approach was first discussed by the Winzor group [18]. We have extended their equation up to the fifth power in activity. The dependence of the activity distribution on the gravity field during the sedimentation equilibrium is well known and it can be described by an exponential function, the so-called ψ-function. Recently a similar procedure has been published by Winzor et al. [17].

Senear and Teller [9] have estimated the average molecular mass of Con A by the point by point method at different loading concentrations and pH. Based on these data the calculated association constants were analyzed as a Wyman linked function by using truncated van't Hoff temperature dependence. From the published truncated van't Hoff function we derived a binding constant of $K_w = 9.9$ g/l for $T = 283$ K and pH = 7.5. This result was in good agreement with our finding, indicating the low influence of virial coefficients on moderate association constants. The possible direct influence of the salt cosolute on the binding constant was not taken into account. In contrast to the effects of non-electrolytes [25] no experiences concerning electrolytes are available. As demonstrated in Fig. 1 a stronger influence of non-ideality can be expected for protein associates with considerably higher association constants. Unfortunately such proteins were not available for analysis. However, we suggest when analyzing interaction constants in extremely high protein concentrations the proposed functions can be advantageous.

Acknowledgements

We are grateful to Ms Bärbel Bödner for technical assistance and Dr Howard Etlinger for reading the manuscript. This work was supported by a grant of the Deutsche Forschungsgemeinschaft HE 1318/15-1/2.

References

[1] T.M. Laue, W.F. Stafford, Modern applications of analytical ultracentrifugation, Annu. Rev. Biophys. Biomol. Struct. 28 (1999) 75–100.
[2] G. Rivas, W. Stafford, A.P. Minton, Characterization of heterologous protein protein interactions using analytical ultracentrifugation, Methods — A Companion to Methods in Enzymology 19 (1999) 194–212.
[3] J. Behlke, O. Ristau, Analysis of interacting biopolymer systems by analytical ultracentrifugation, Eur. Biophys. J. 25 (1997) 325–332.
[4] D.A. Yphantis, Equilibrium ultracentrifugation, Biochemistry 3 (1964) 297–317.
[5] J. Behlke, O. Ristau, Analysis of the thermodynamic non-ideality of proteins by sedimentation equilibrium experiments, Biophys. Chem. 76 (1999) 13–23.
[6] T.L. Hill, Y.D. Chen, Theory of aggregation in solution. I. General equations and application to the stacking of bases, nucleotides, etc, Biopolymers 12 (1973) 1285–1312.
[7] M.L. Johnson, J.J. Coreira, D.A. Yphantis, H.R. Halvorson, Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques, Biophys. J. 36 (1981) 575–588.
[8] G.H. McKenzie, W.H. Sawyer, L.W. Nichol, The molecular weight and stability of concanavalin A, Biochim. Biophys. Acta 263 (1972) 283–293.
[9] D.F. Senear, D.C. Teller, Thermodynamics of concanavalin A dimer–tetramer self-association: sedimentation equilibrium studies, Biochemistry 20 (1981) 3076–3083.
[10] M. Huet, Factors affecting the molecular structure and the agglutinating ability of concanavalin A and other lectins, Eur. J. Biochem. 59 (1975) 627–632.

[11] M. Huet, J.M. Claverie, Sedimentation studies of the reversible dimer–tetramer transition kinetics of concanavalin A, Biochemistry 17 (1978) 236–241.
[12] W.G. McMillan, J.E. Mayer, The statistical thermodynamics of multicomponent systems, J. Chem. Phys. 13 (1945) 276–305.
[13] P. Ross, A.P. Minton, Analysis of non-ideal behavior in concentrated hemoglobin solutions, J. Mol. Biol. 112 (1977) 437–452.
[14] J.W. Williams, K.E. van Holde, R.L. Baldwin, H. Fujita, The theory of sedimentation analysis, Chem. Rev. 58 (1958) 715–806.
[15] H. Fujita, Mathematical Theory of Sedimentation Analysis, Academic Press, New York, 1962, p. 238.
[16] 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, pp. 311–330.
[17] D.J. Winzor, M.P. Jacobsen, P.R. Wills, Allowance for the thermodynamic nonideality in the analysis of sedimentation equilibrium distributions reflecting complex formation between dissimilar reactants, Progr. Colloid Polym. Sci. 113 (1999) 69–75.
[18] P.R. Wills, M.P. Jacobson, D.J. Winzor, Direct analysis of solute self-association by sedimentation equilibrium, Biopolymers 38 (1996) 119–130.
[19] B.A. Cunningham, J.L. Wang, M.N. Pflumm, G.M. Edelman, Isolation and proteolytic cleavage of the intact subunit of concanavalin A, Biochemistry 11 (1972) 3233–3239.
[20] J. Yariv, A.J. Kalb, A. Levitzki, The interaction of concanavalin A with methyl α-D-glucopyranoside, Biochim. Biophys. Acta 165 (1968) 303–305.
[21] T. Boublik, Equation of state of hard convex body fluids, Mol. Phys. 42 (1981) 209–216.
[22] T. Boublik, I. Nezbeda, P–V–T behaviour of hard body fluids, Theory and Experiment, Collections Czechoslovak Chem. Commun. 51 (1986) 2301–2432.
[23] M. Sindelka, T. Boublik, The third cross virial coefficient of hard convex bodies, Mol. Phys. 96 (1999) 243–247.
[24] J.M. Rallison, S.E. Harding, Excluded volume for pairs of triaxial ellipsoids at dominant Brownian-motion, J. Colloid Interface Sci. 103 (1985) 284–289.
[25] P.R. Wills, M.P. Jacobsen, D.J. Winzor, Direct analysis of sedimentation equilibrium distributions reflecting macromolecular interactions, Progr. Colloid Polym. Sci. 107 (1997) 1–10.