Biophysical Chemistry, 37 (1990) 121–140
Elsevier

BIOCHE 01471

A simple model for solvation in mixed solvents

Applications to the stabilization and destabilization of macromolecular structures

John A. Schellman

Department of Chemistry and Institute of Molecular Biology, University of Oregon, Eugene, OR 97403, U.S.A.

Received 11 January 1990
Accepted 8 February 1990

Solvation; Binding model; Structure stabilization

The properties of a simple model for solvation in mixed solvents are explored in this paper. The model is based on the supposition that solvent replacement is a simple one-for-one substitution reaction at macromolecular sites which are independent of one another. This leads to a new form for the binding polynomial in which all terms are associated with ligand interchange rather than ligand addition. The principal solvent acts as one of the ligands. Thermodynamic analysis then shows that thermodynamic binding (i.e., selective interaction) depends on the properties of $K'-1$, whereas stoichiometric binding (site occupation) depends on K'. K' is a 'practical' interchange equilibrium constant given by $(f_3/f_1)K$, where K is the true equilibrium constant for the interchange of components 3 and 1 on the site and f_3 and f_1 denote their respective activity coefficients on the mole fraction scale. Values of K' less than unity lead to negative selective interaction. It is selective interaction and not occupation number which determines the thermodynamic effects of solvation. When $K'>100$ on the mole fraction scale or $K'>2$ on the molality scale (in water), the differences between stoichiometric binding and selective interaction become less than 1%. The theory of this paper is therefore necessary only for very weak binding constants. When $K'-1$ is small, large concentrations of the added solvent component are required to produce a thermodynamic effect. Under these circumstances the isotherms for the selective interaction and for the excess (or transfer) free energy are strongly dependent on the behavior of the activity coefficients of both solvent components. Two classes of behavior are described depending on whether the components display positive or negative deviations from Raoult's law. Examples which are discussed are aqueous solutions of urea and guanidinium chloride for positive deviations and of sucrose and glucose for negative deviations. Examination of the few studies which have been reported in the literature shows that most of the qualitative features of the stabilization of proteins by sugars and their destabilization by urea and guanidinium chloride are faithfully represented with the model. This includes maxima in the free energy of stabilization and destabilization, decreased and zero selective interaction at high concentrations, etc. These phenomena had no prior explanation. Deficiencies in the model as a representation of solvation in aqueous solution are discussed in the appendix.

1. Introduction

It is a pleasure to be a contributor to this dedicatory special issue for Jeffries Wyman. My experiences with his writings and scientific influence began in an unusual way. The topic of my doctoral thesis with Walter Kauzmann was the theory of dielectric properties and one of the earliest scientific papers that I read was Wyman's seminal paper on the dielectric properties of associated liquids [1] which led to the theories of Onsager [2] and Kirkwood [3]. It was only many years later that I became acquainted with his papers on linkage and allosteric interactions which

Correspondence address: J.A. Schellman, Department of Chemistry and Institute of Molecular Biology, University of Oregon, Eugene, OR 97403, U.S.A.

have been the central themes of his life's work. Concepts that he introduced have been very germane to my own recent work and for this reason I am in the agreeable position of contributing an article directly related to linkage relations and binding polynomials, two of the major ideas which he introduced into the field of molecular interactions in solution. It is also fitting that the material for this article was recently presented orally at the celebration of the achievements of Stanley J. Gill, one of Wyman's principal collaborators in recent years (September 1989 at Boulder, CO).

The topic of this paper is the solvation of a very dilute component in the presence of a mixed solvent. We define a mixed solvent as any system which contains a principal solvent and any other component in high concentration. For example, concentrated solutions of sugars, denaturing agents, alcohols and the like are considered to have two solvent species: 1, the principal solvent, usually water; and 3, the subsidiary solvent. The main purpose is to determine the thermodynamic effect of adding component 3 to the principal solvent. There are numerous applications in the fields of macromolecular chemistry and chromatography, but special emphasis will be given to the problems of protein denaturation and stabilization.

The study is based on a model in which solvation is considered to be a simple molecular occupation of independent sites. In this respect it resembles the independent-site binding model which leads to a Langmuir isotherm, $\bar{\nu} = \sum_j K_j A_3/(1 + K_j A_3)$, where K_j is the intrinsic binding constant for a ligand and A_3 its activity in solution. The main difference is that components 1 and 3 are considered to be in competition. This leads to a modified binding polynomial which contains the activities of both solvent components. It is a binding polynomial for substitution rather than simple ligand attachment. Solvation by components 1 and 3 is linked in the general Wyman sense [4], since occupation by 3 excludes occupation by 1 and by the fact that the activities of 1 and 3 in solution are related by the Gibbs-Duhem equation.

This model provides an oversimplified description of solvation and some of its limitations are discussed in section 6. Nevertheless, like the independent-site binding model, it may serve as a basic prototype for the solvation process upon which refinements may be based. As is shown in section 7, some of the more perplexing results of protein denaturation and stabilization studies are automatic consequences of this representation of the solvation process.

2. The binding model

We consider the solvation properties of a molecule, P, which is normally a macromolecule. It is assumed that the surface of P can be divided into a number of sites which can be occupied by solvent molecules. The sites are assumed to be independent of one another. The surface of the molecule presented to the solvent depends on conformation, so the nature and number of sites is conformation dependent. In the absence of an added third component, all of the sites are occupied by the principal solvent, 1. The addition of a component 3 causes the partial replacement of 1 at the binding sites. We express this as the exchange equilibrium

$$b:1 + 3 = b:3 + 1 \qquad K_b = \frac{[b:3]A_3}{[b:1]A_1} \tag{1}$$

where the colon indicates a location for solvent molecules at site b; $[b:1]$ and $[b:3]$ are the concentrations of site b occupied by 1 and 3, respectively; and A_1 and A_3 denote the activities of 1 and 3 on the mole fraction scale. Mole fractions are used because it is desirable to have the solvent and solute 3 on a similar footing. 1 and 3 are not necessarily molecules of the same size. Possible types of substitution are depicted in fig. 1.

The chemical potential of P can be written as

$$\mu_2 = \mu_2^\circ + RT \ln[\mathrm{P}] + RT\beta_2^\circ \tag{2}$$

where [P] is the concentration of P, usually on a molal or molar basis, and $RT\beta_2^\circ$ the excess free energy. [6,7] In the limit of very low concentration of P where interactions between the macromolecules can be neglected (these profoundly affect the solvation of P), the excess free energy is a function

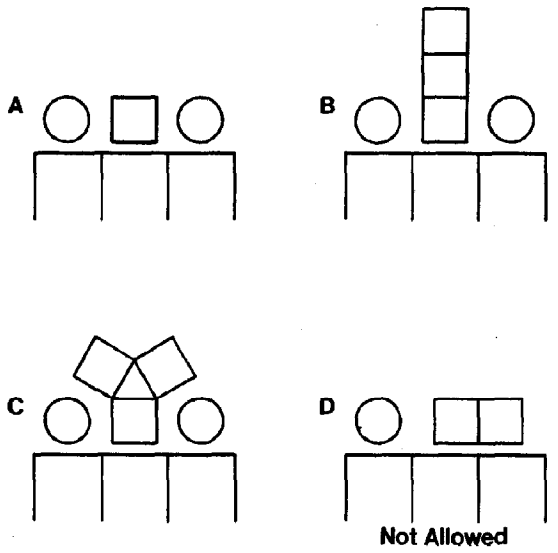

Fig. 1. Cartoons of binding events related to the substitution model. The circles are component 1, squares and aggregates of squares are component 3. (A) Molecules of equal size; (B) substituting group at end; (C) multivalent group or branched molecule (this could be a guanidinium ion); (D) coverage of more than one site, which is not allowed by the model.

only of the concentration or activity of the third component. The condition of dilute component 2 is indicated by the superscript ° on β.

For strong stoichiometric binding β_2° is given by $-\ln \Sigma$ [8], where Σ is the usual binding polynomial $1 + K_1A_3 + K_2A_3^2 + \cdots$. For weak selective interaction β_2° is normally given as a power series in m_3, the molality of 3. The first form is inappropriate for weak selective interactions because it equates occupancy of a site to binding at the site. In concentrated solutions solutes occupy sites with high probability even in the absence of favorable interactions. The second form is general but does not converge properly for stoichiometric binding which is represented by sigmoid curves which saturate at relatively low m_3. It was shown in a previous paper [8] that the deficiencies of the binding polynomial method for selective interactions could be removed by placing the solvent 1 and the solute 3 on an equal footing in constructing the polynomial. This is easily done for the independent site model.

Since the sites are independent, each site contributes separately to the excess free energy so that

$$-\beta_2^\circ = \ln \Sigma = \sum_{b=1}^{n} \ln \Sigma_b \tag{3}$$

where Σ_b is the binding polynomial for site b. Normally, the site binding polynomial for the reaction of eq. 1 is simply $(1 + K_bA_3)$. The purpose, however, is to represent solvation by 3 as a replacement of the principal solvent, and not as a ligand attachment. This problem was considered previously and it was shown by both thermodynamic and statistical mechanical arguments [8] that the site binding polynomial for this situation is given by

$$\Sigma_b = A_1 + K_bA_3 \tag{4}$$

so that

$$\beta_2^\circ = -\sum_{b=1}^{n} \ln(A_1 + K_bA_3) \tag{5}$$

All of the properties of thermodynamic interest can be derived from β_2°, if it is known as a function of the concentration of 3, temperature, pressure, etc. [9], and some of these will be developed in later sections. This is in fact the purpose of this model. Though oversimplified it gives an explicit formula for β_2° which covers cases from negative selective interaction to strong site binding and will permit us later on to give a quantitative definition to the terms strong and weak binding and to obtain explicit formulas for the selective interaction.

3. Thermodynamic binding: selective interaction

Selective interaction is a purely thermodynamic measure of the relative interaction of a solute and the principal solvent with a third species. It can be measured quantitatively with no assumptions about the underlying molecular events [10]. As is shown in eq. 6 below, it is the quantity that is *the relevant quantity* if one is interested in the change in chemical potential of a macromolecule which is induced by adding denaturants, stabilizers, acids, bases, substrates, etc. When the interaction is

strong, it is identical to the usual molecular definitions of binding. When interactions are weak or negative (relative to the principal solvent) it can differ significantly from the molecular description.

3.1. Non-ionic substances

We define selective interaction by the quantity, Γ'_{23}, read as 'the selective interaction of 3 with 2' and given by [6]

$$\Gamma'_{23} \equiv \left(\frac{\partial m_3}{\partial m_2}\right)_{T,\mu_3} = -\left(\frac{\partial \mu_2}{\partial \mu_3}\right)_{T,m_2} = -\left(\frac{\partial \beta_2^o}{\partial \ln A_3}\right)_{T,m_2} \quad (6)$$

It represents the number of molecules of 3 which must be added to a solution to restore its chemical potential, when a molecule of component 2, the macromolecule, has been added to the solution. It can be shown to be the excess number of molecules of 3 which are present in the solvent shell of the macromolecule relative to the principal solvent [9]. Γ_{23} can be positive or negative. There are three common ways of defining selective interaction. They arise because of the choice of variables in describing thermodynamic systems. Eisenberg [7] has discussed these possibilities and introduced the prime and double prime notations to distinguish them. The appendix shows the relationship amongst the three measures. For a macromolecule the differences are small and we may often assume that $\Gamma_{23} \cong \Gamma'_{23} \cong \Gamma''_{23}$ (see the appendix for details). We will be using Γ'_{23}, but the prime will be dropped in later sections.

The second equality in eq. 6 is a general relationship which plays the important role of relating selective interactions with changes in free energy; the last is a consequence of the fact that β_2^o is the only part of μ_2 which depends on A_3. To simplify notation it will be understood in the following that there are N binding sites and the symbols Γ'_{b3} and Γ_{b1} will be used for the selective interaction at site b of components 3 and 1, respectively.

From eqs 3 and 6

$$\Gamma'_{23} = \sum_{b=1}^{n} \Gamma'_{b3}; \quad \Gamma'_{b3} = \left(\frac{\partial \ln \Sigma_b}{\partial \ln A_3}\right)_{T,m_2} \quad (7)$$

where Γ'_{b3} is the selective interaction of 3 associated with site b. This is not the same as the number of molecules of 3 occupying site b! The total number of molecules selectively bound to all sites of the macromolecule is given by $\sum_b \Gamma'_{b3}$.

In the application of eq. 6 or 7, it must be recalled that the activities A_1 and A_3 are not independent of one another, but are related by the Gibbs-Duhem equation at constant T and P:

$$\chi_1 \, d \ln A_1 + \chi_2 \, d \ln A_2 + \chi_3 \, d \ln A_3 = 0$$

where χ represents mole fraction. In the limit of very small concentration of the macromolecule, $\chi_2 \, d \ln A_2$ will be negligible so that

$$\left(\frac{\partial A_1}{\partial \ln A_3}\right)^o_{T,P} = -\frac{\chi_3}{\chi_1} A_1 \quad (8)$$

Combining eqs 4, 6 and 8, we obtain

$$\Gamma'_{b3} = \frac{-\frac{\chi_3 A_1}{\chi_1} + K_b A_3}{A_1 + K_b A_3}.$$

To simplify this relation we eliminate activities by introducing activity coefficients on the mole fraction scale, $A_i = f_i \chi_i$, $i = 1, 3$. Note that because of the formal equivalence of component 1 and component 3 as solvent, the nonideality of component 1 is measured by an activity on the mole fraction scale instead of the usual osmotic coefficient.

$$\Gamma_{b3} = \frac{(K_b f_3 - f_i)\chi_3}{f_1 \chi_1 + K_b f_3 \chi_3} = \frac{(K'_b - 1)\chi_3}{1 + (K'_b - 1)\chi_3} \quad (9)$$

This is the mole fraction formula for the selective interaction of component 3 on site b. The form on the right was obtained by dividing the numerator and denominator by f_1 and eliminating χ_1 by means of the relation $\chi_1 + \chi_3 = 1$, which is valid at low concentrations of component 2. The mole fraction activity coefficients were then introduced into the equilibrium constant by defining $K'_b = f_3 K_b / f_1$. This quantity is often called the 'practical' equilibrium constant. It is not a constant, since it depends on concentration as well as T and P, but may be used with concentrations rather than activities.

The discussion and equations are simplified if mole ratios are used as concentration units rather than mole fractions or molalities. We define r_3 as the ratio of the number of moles of 3 to that of 1 in the bulk solvent, i.e., the 'free' molecules of these components. Similarly, r_{b3} is the ratio of the average number of moles of 3 bound to site b (θ_3) to that of 1 (θ_1), i.e., $r_{b3} = \theta_3/\theta_1 = \theta_3/(1-\theta_3)$. With these definitions we may rewrite eq. 1 as

$$r_{b3} = K_b' r_3 \tag{10}$$

Note that ratios of concentrations are independent of concentration units. Concentrations in mole ratios are quickly converted to molalities since $r_3 = m_3/m_1$, where $m_1 = 1000/M_1$ is the constant molality of the principal solvent. K_b' provides a comparison of the mean composition on the site with the composition of the solvent. For K_b' less than, greater than, or equal to unity one is dealing with negative selective interaction, positive selective interaction and no selective interaction, respectively.

The molal version of eq. 9 may be obtained by dividing its middle expression by X_1, substituting $r_3 = \chi_3/\chi_1$, and rearranging:

$$\Gamma_{b3}' = \frac{(K_b'-1)m_3}{m_1 + K_b' m_3} = \frac{(K_b'-1)r_3}{1 + K_b' r_3} \tag{11}$$

The equilibrium constants K and K' are defined on the mole fraction or mole ratio basis. For practical calculations it is useful to convert to a constant on a molality basis. Dividing the numerator and denominator of the second form of eq. 11 by m_1, the number of moles of principal solvent per kg, gives

$$\Gamma_{b3}' = \frac{(K_b'' - 1/m_1)m_3}{1 + K_b'' m_3} \tag{12}$$

where $K'' = K'/m_1 = K(f_3/f_1 m_1)$. Recall that f_3 and f_1 are activity coefficients on the mole fraction scale. This is a formula that is most useful for practical calculations. It shows that on the molal scale, the effective binding constant must be compared with the very small number $1/m_1$ which, for example, is 0.018 for water. The significance of this will be discussed in section 3.2.

The selective interaction of component 1, defined as $\Gamma_{b1}' = (r_1'/r_2')_{T,\mu_1}$, can be derived in a fashion similar to the analysis given in eqs 7–11, with the result

$$\Gamma_{b1}' = \frac{-(K_b'-1)\chi_1}{\chi_1 + K_b'\chi_3} \tag{13}$$

The primes on r (again Eisenberg's notation [7]) indicate that the mole ratios are now based on component 3, i.e., $r_1' = n_1/n_3$, $r_2' = n_2/n_3$. Converting to other concentration units

$$\Gamma_{b1} = \frac{-(K_b'-1)m_1}{m_1 + K_b' m_3}\,(\text{molality}) = \left(\frac{-(K_b'-1)}{1 + K_b' r_3}\right)(\text{mole ratio}) \tag{14}$$

In the present analysis not much use will be made of Γ_{1b}, but formulas are presented because it is often the selective interaction parameter that is reported, normally in the form of 'preferential hydration'. Preferential hydration is measured directly in aqueous isopiestic determinations and is usually given in mass units as $\xi_1' = (\partial w_1/\partial w_2)_{T,\mu_1}$, where w_1 and w_2 are the concentrations of 1 and 2 in g per g of component 3. Similarly, $\Gamma_{21}' = (\partial r_1'/\partial r_2')_{T,\mu_1}$ where r' is the mole ratio per mole of component 3. These unusual concentration units give ξ_1 and Γ_{21}' some nonintuitive properties. They are related by $\Gamma_{21}' = (M_2/M_1)\xi_1$.

It should be noted that Γ_{23}' and Γ_{21} do not have the properties of ordinary binding fractions. They do not sum to unity and their ratio is given by $(\Gamma_{23}'/\Gamma_{21}') = -r_3$. The latter formula is a general result which is independent of model [7]. It clearly applies to the binding at an independent site b and serves as a check on the thermodynamic consistency of the results.

3.2. *Ionic substances*

If component 3 is ionic and shares a common ion with the macromolecule, the selective interaction has a Donnan component. The formulas which cover this case are discussed in detail in the appendix. We require eq. A11

$$\Gamma_{23}' = \frac{-|Z|}{2 + m_3\beta_{33}} - \left(\frac{\partial \beta_2}{\partial \ln A_3}\right)_{T,P,m_2} \tag{A11}$$

where Z is the charge on the macromolecule, the first term on the right is the Donnan effect corrected for nonideality, and the second term on the right is identical to that in eq. 6 for uncharged molecules.

We write the binding polynomial for site b as

$$\Sigma_b = A_1 + K_+A_+ + K_-A_- \tag{14}$$

where A_+ and A_-, and K_+ and K_- are the activities and the association constants of the cation and anion, respectively. Since the ionic activities are not known we transform to mean ionic activities and write

$$K_+A_+ + K_-A_- = K_bA_\pm$$

This is essentially a definition of K_b. For symmetrical electrolytes $A_+A_- = A_\pm^2 = A_3$ and for an ideal, symmetrical electrolyte,

$$A_+ = A_- = \chi_+ = \chi_- = \chi_3 \equiv \frac{m_3}{m_1 + 2m_3}$$

in the limit of zero concentration of component 2. Note that in this definition of the mole fraction of an ionic component, component 3 is considered to contribute two particles to the total number of molecules (see ref. 10 for a discussion of this convention). Thus, for the ideal, symmetrical case, $K_+ + K_- = K_b$, and K_b represents the *sum* of the association constants for the anion and cation. Thermodynamics will not distinguish the binding of the anion and cation. Though the investigator may have prejudices concerning the ion which is involved (for example, the guanidinium ion in guanidinium chloride) the thermodynamic relations maintain the full set of possibilities. This is important, since both ions have been shown to have important effects on the interactions with macromolecules [5,11,12].

Selective interaction at site b contributes

$$\beta_{b3} = -\ln(A_1 + K_bA_\pm) \tag{15}$$

to the excess free energy of the molecule. To obtain the formula for the selective interaction itself, this must be differentiated with respect to $\ln A_3$ in the limit of low concentration of component 2, with due regard to the Gibbs-Duhem relation. For a symmetrical ionic compound we have the two relations, eq. 8 and $(\partial A_\pm/\partial \ln A_3) = A_\pm/2$. Substituting these into eq. A11

$$\Gamma'_{23} = \frac{-|Z|}{2 + m_3\beta_{33}} + \sum_b \frac{(K_b/2)A_\pm - \dfrac{\chi_3 A_1}{\chi_1}}{A_1 + K_bA_\pm}$$

This formula is simplified by the same steps as those which lead to eq. 9, with the result

$$\Gamma'_{23} = \frac{-|Z|}{2 + m_3\beta_{33}} + \sum_b \frac{(K'_b/2 - 1)\chi_3}{1 + (K'_b - 1)\chi_3}$$

(mole fraction basis)

$$= \frac{-|Z|}{2 + m_3\beta_{33}} + \sum_b \frac{(K'_b/2 - 1)r_3}{1 + K'r_3}$$

(mole ratio basis)

$$= \frac{-|Z|}{2 + m_3\beta_{33}} + \sum_b \frac{(K'_b/2 - 1)m_3}{m_1 + K'm_3}$$

(molality basis)

$$= \frac{-|Z|}{2 + m_3\beta_{33}} + \sum_b \frac{(K''_b/2 - 1/m_1)m_3}{1 + K''_bm_3}$$

(molality basis, molal K'') (16)

It is seen that the main differences between the ionic and nonionic cases is the factor of 2 in the numerator of the rightmost term of eq. 16 and the appearance of the Donnan effect term. The factor of 2 could be incorporated into K' and K'' but this would make the distinction between the ionic and nonionic cases less clear. For example, it will often be the case that one ion only is effectively bound to the site. We assume it is the cation. Then $K_b = K_+$. If the factor of 2 were incorporated in K_b, then K_b would have half the value of the real association constant.

At first sight it might appear that the Donnan term would need correction for the change in charge which occurs on the binding of ions but this is not true; this effect is already included in eq. 16. We show this with a simple example. Suppose that $Z = +10$ and an anion is strongly bound, $K \gg 1$. Then the species in solution is $(PA)^{9+}A_9^-$ where A is the singly charged anion. In the ideal limit the selective interaction would then arise only from the Donnan effect of the nine net

charges on the macromolecule and would be given by $\Gamma_{b3} = -4.5$. Eq. 16, ignoring corrections for nonideality, gives $-5 + 1/2 = -4.5$. Suppose instead that cation, C, is strongly bound. Then the species in solution would be $(PC)^{11+}A_{11}^-$. Consequently, the Donnan contribution would be $-11/2$ to which one must add unity for the additional mole of CA which is now part of the macromolecular complex. This gives $\Gamma_{b3} = -4.5$, which is what is calculated from eq. 16. Thus, regardless of the Donnan effect, Γ_{b3} does not depend on the charge of the ion that is bound and eq. 16 gives the correct result.

4. Site occupation

In this section we will be comparing selective interaction at a site with elementary binding theory, which expresses binding in terms of site occupancy. Since there is no Donnan contribution to site occupancy, the comparison will be made with the selective interaction of uncharged substances. *To make the comparison with ionic substances, the Donnan effect contribution to Γ_{23} must first be subtracted and the occupancy must be summed over ionic species to conform with eq. 16.*

Occupancy at a site can be calculated in the usual way from the equilibrium relation, eq. 10. The average occupancy of site b is given by

$$\theta_{3b} = \frac{r_{3b}}{1 + r_{3b}} \tag{17}$$

and from eq. 10,

$$\theta_{3b} = \frac{K_b' \chi_3}{\chi_1 + K_b' \chi_3} \quad \text{(mole fraction)}$$

$$= \frac{K_b' r_3}{1 + K_b' r_3} \quad \text{(mole ratio)}$$

$$= \frac{K_b'' m_3}{1 + K_b'' m_3} \quad \text{(molality)} \tag{18}$$

Apart from the χ_1 in the first formula, these look like the standard results. It should be recalled that K' is dependent on the activity coefficient of component 1. Activity coefficients are often ignored in this type of analysis but this is not permissible with weak selective interaction. By comparing eqs 11 and 18 we see that for the independent site model, selective interaction at a site and average occupancy of that site are related by the simple formula

$$\Gamma_{b3} = \left(\frac{K' - 1}{K'}\right)\theta_{3b} \tag{19}$$

or on the molal scale by

$$\Gamma_{b3} = \left(\frac{K'' - 1/m_1}{K''}\right)\theta_{3b} \tag{20}$$

This formula provides us with the answer to the question: 'When is a binding constant large enough that the distinction between selective interaction and simple binding can be ignored?' If a 1% error can be tolerated, then on the molal scale K'' should be $100/m_1$ or larger. For water this is $K'' \geq 2$. A binding constant this small is rarely reported except in the problem we are addressing, the solvation of macromolecules. Yet, even a binding constant this small represents a very considerable segregation of the solvent, since the relative concentration of component 3 on the site is 100-times that in the bulk solvent.

The association constants which one normally encounters usually range from 100 to very large numbers (in molar or molal units). Consequently, the distinction between site occupancy and solvent replacement does not arise. Associations of this kind normally saturate at very low concentrations of component 3 and have their thermodynamic consequences at very low concentrations. On the other hand, when high concentrations (1–10m) are required to produce a thermodynamic effect, the possibility of K'' values less than 1 (K' less than 50 in water) must be considered. For example, the unfolded chain of a typical small protein has contact with the order of 1000 more molecules of water than the folded state. Considering the nature of polypeptide chains and the fact that reagents like urea and guanidine are capable of interacting favorably with both polar and hydrophobic groups, it seems far more reasonable to expect that a large fraction of these water molecules are sites for possible replacement with denaturant molecules rather than 10–20 special sites. The latter is often the conclusion of a naive interpretation of binding models.

We can now compare our solvation model with others that have been used in the interpretation of data. Eisenberg [7] has made use of a model for the interaction of salts and denaturants with DNA and proteins. It is assumed that the quantities of component 1 and component 3 in the solvation shell of the macromolecule remain constant over a range of concentrations. The equivalent of this assumption in the present context is that a fraction of the sites F_1 bind component 1 strongly and component 3 negligibly ($K' = 0$) and that a fraction of sites F_3 bind 3 strongly and 1 negligibly ($K' = \infty$). Then, from eqs 11 and 14, and summing over all sites

$$\Gamma_{23} = n(-F_1 r_3 + F_3)$$
$$\Gamma_{21} = n(F_1 - F_3/r_3) \qquad (21)$$

The first term in these equations arises from setting $K = 0$, the second from $K = \infty$. With this model Γ_{23} is linear in r_3 (and therefore m_3) while Γ_{21} is linear in $1/r_3$ ($1/m_3$). The extrapolation of plots of Γ_{21} vs $1/r_3$ permits the estimation of the hydration, nF_1, of the macromolecule in the presence of component 3. One of the difficulties with this model is that it implies a very strong interaction with component 3 at low concentrations until the constant level of solvation by component 3 is reached. This strong interaction would be easy to detect but has so far not been observed by Timasheff and his co-workers who have worked with many biological macromolecules and many solvent components.

Inoue and Timasheff [13] have made use of another model, based on the concept that the solvation shell of a macromolecule may be represented as the removal of a quantity of the solvent components 1 and 3 from the pool of free components in the solution. In the absence of component 3 the solvation is described as the fraction of hydration, A_1, i.e., g of component 1, normally water, bound per g of macromolecule. This is a quantity which has been estimated for many biological macromolecules. In the presence of component 3 quantities of both components, A_1 and A_3 are bound to the macromolecule. They show generally that this assumption leads to the relation $A_3 = \xi_3 + w_3 A_1$. ξ_3 is the selective interaction on a mass rather than mole basis ($\xi_3 = (M_3/M_2)\Gamma_{23}$) and w_3 is the concentration as a mass ratio g_3/g_1. The site binding model assumes that molecules are either bound or free and so must lead to an equivalent relation. It is easily derived. From eqs 11 and 14 we can split the selective interaction into two terms, giving

$$\Gamma_{23} = \sum_b \left(\frac{K_b' r_3}{1 + K' r_3} - \frac{r_3}{1 + K_b' r_3} \right)$$
$$= n(\bar{\theta}_3 - r_3 \bar{\theta}_1) \quad (\xi_3 = A_3 - w_3 A_1) \qquad (22a)$$
$$\Gamma_{21} = \sum_b \left(\frac{-K_b'}{1 + K_b' r_3} + \frac{1}{1 + K_b' r_3} \right)$$
$$= n(-\bar{\theta}_3/r_3 + \bar{\theta}_1) \quad (\xi_1 = A_1 - A_3/w_3) \qquad (22b)$$

The relations on the right are the equivalents in the mass units of Inoue and Timasheff. In the utilization of these formulas Timasheff and his co-workers assume that A_1 is a constant whereas in the present analysis it is dependent on the concentration of component 3.

We have then three models of the solvation layer: one in which both components are constant, one in which the principal solvent is constant and one in which both the principal and secondary solvent are variable but are related by an exchange equilibrium constant. Considering the heterogeneous nature of biopolymers and the great variety of interactions (which must include not only weak and strong interactions in the sense of the present paper, but also ion condensation, occupation of cavities, long-range interactions, etc.), none of these models can be wholly representative of the true physical situation. Nevertheless, the independent site model generates most of the properties of the other two models discussed above and has a useful flexibility of interpretation. Deficiencies of the model will be discussed below in section 6.

5. Ideal isotherms

If components 1 and 3 may be assumed to form an ideal solution, the situation is considerably simplified because the selective interaction relations can be interpreted without data on activity coefficients as a function of concentration. The

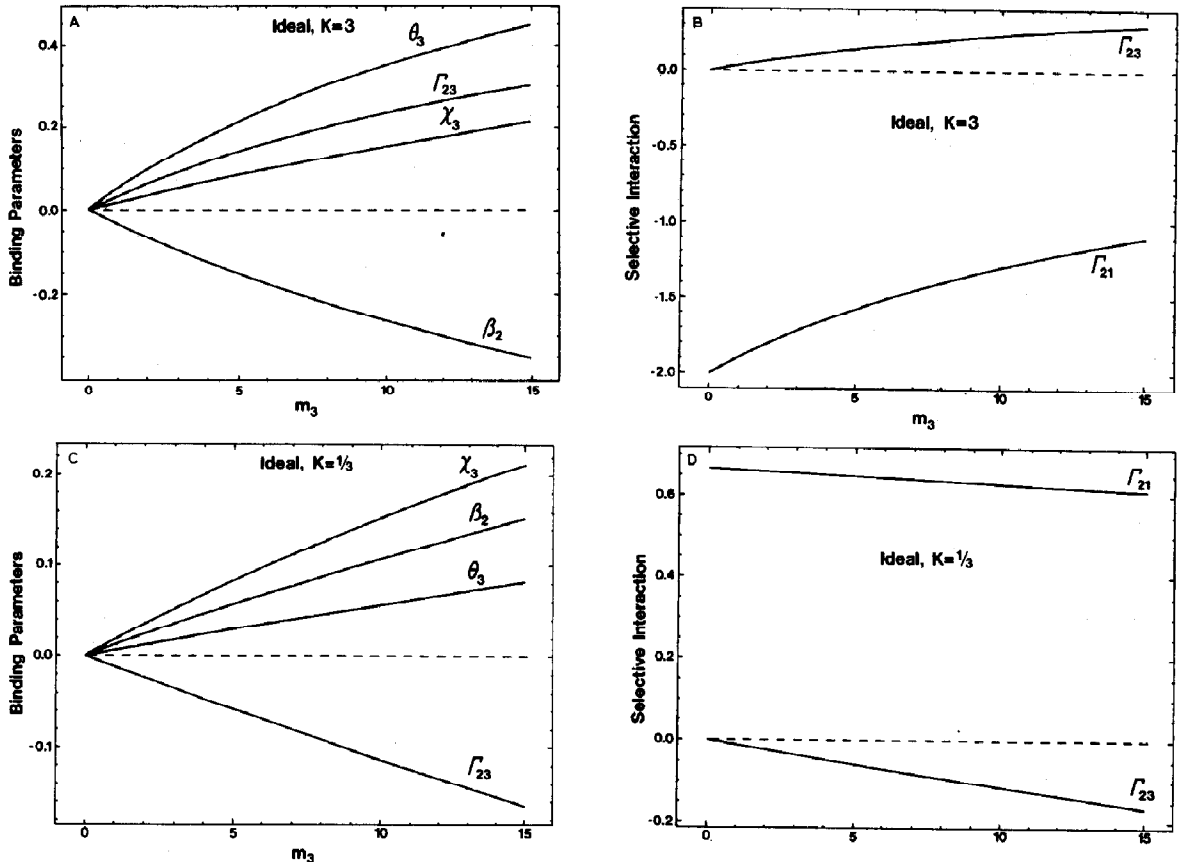

Fig. 2. Ideal isotherms calculated for a site in a solution which obeys Raoult's law for components 1 and 3. (A) $K = 3$: plot of occupation θ_3, selective interaction Γ_{b3}, mole fraction χ_3, and excess free energy β_2. (B) $K = 3$: comparison of Γ_{b3} and Γ_{b1}. (C) $K = 1/3$: otherwise the same as A. (D) $K = 1/3$: otherwise the same as B.

operative relations are eqs 11 and 16 with $K' = K$. As in the previous section, the relations for an unionized component 3 will be given. For ionic compounds the results must be adjusted for the Donnan contribution and for the possible binding of anions or cations (the factor of 2 in eqs 16). When K is large (greater than 2 on the molal scale, greater than 100 on the mole fraction scale) these relations are the standard independent site binding isotherms. In this case, for a given site the selective interaction and the site occupancy are essentially equal: $\Gamma_{b3} \cong \theta_{3b}$; $\Gamma_{b1} \cong 1 - \theta_{3b}$.

Γ_{b3} and Γ_{b1} begin to deviate very significantly from site occupancies for $K < 10$. Fig. 2 shows calculated curves for two symmetric values of K: $K = 3$ and $K = 1/3$. We will use these curves together with the formulas to deduce a few generalities.

5.1. Case 1, $K > 1$

Here we have positive selective interaction for component 3 and negative selective interaction for the principal solvent. The upper part of fig. 2 shows that in this case Γ_{b3} is greater than the mole fraction but less than site occupancy. As K approaches unity from above, the site occupancy curve approaches χ_3, and Γ_{b3} goes to zero for all concentrations. For $1 < K < 2$, Γ_{b3} lies below the mole fraction curve. For $K > 2$ the curves for Γ_{b3} and χ_3 eventually cross but it is only for $2.0 < K$

< 2.2 that this crossover occurs at a practical concentration between 1 and $10m$. In all cases the contribution of a site to Γ_{b3} lies between 0 and 1.

When $K > 1$, Γ_{b1} is negative and bears little relation to a site occupancy. As the concentration of 3 goes to zero, Γ_{b1} approaches $1 - K$ (eq. 14). Since K may be large, Γ_{b1} may assume large negative values of magnitude much greater than unity. For example, if $K = 10^4$, then, in the limit of low m_3, $\Gamma_{b1} \cong -10^4$. This seemingly bizarre result, where a single binding site affects thousands of solvent molecules, can be understood via a simple example. Suppose that $K = 10^4$ and conditions are such that the site is on the average half occupied by a molecule of 3; $\theta_{3b} = \theta_{1b} = 0.5$. Two molecules of component 2 are added to the solution so that on the average one molecule has its site filled with a molecule of 1 and the other with a molecule of 3. The solution has now been depleted by one molecule of each of the solvent components. However, if the ratio of the number of molecules on the site, r_b, is unity, then by eq. 10 there are 10^4 molecules of component 1 per molecule of component 3 in the solution. Therefore, to restore the original concentration ratio, $10^4 - 1$ molecules of component 1 must be eliminated. The large number is not associated with binding but with the restoration of concentration ratios. *Though it is highly exaggerated when K is large, this effect is always present in the selective interaction process.*

At high concentrations of 3, Γ_{b1} tends to zero. This limit is of little practical importance in aqueous solution where r_3 rarely exceeds 0.2–0.3. The exception is when components 1 and 3 are infinitely miscible and the full range of composition is of interest.

5.2. Case 2, $K < 1$

When $K < 1$, the order of the curves for Γ_{b3} is changed since 3 molecules are now selectively repelled from the site. The fractional occupation is now less than the mole fraction in the bulk solution. Γ_{b3} is negative and becomes larger in magnitude as r_3 becomes larger. Γ_{b3}, in fact, behaves in the opposite fashion to the fractional occupation; as the fractional occupation increases, the negative selective interaction increases in magnitude. As $K \to 0$ from below the fractional occupation approaches the mole fraction curve as a limit while Γ_{b3} rises to zero at all concentrations. On the other hand, Γ_{b1} appears to be almost constant. This is easily understood. For weak binding and low concentration $-\Gamma_{b3}$ will in general be proportional to r_3 as can be seen in fig. 2C. But since $\Gamma_{b1} = -\Gamma_{b3}/r_3$ this means that Γ_{b1} is essentially constant. This is presumably the origin of many reports that the hydration of biological macromolecules is independent of added solutes. It also demonstrates that Γ_{b3} (or more generally the selective interaction of the least concentrated solvent component) is far more sensitive to binding events at a site than Γ_{b1}.

The preceding discussion might make it appear that there is an asymmetry between components 1 and 3, but this is not the case. It appears so only because the discussion has been aimed toward solutions in which the mole ratio of component 3 to component 1 does not exceed unity, which is true for most aqueous solutions of interest. Considered over the complete range of composition, there is symmetry. For example as $r_3 \to 0$, $\Gamma_{b1} \to 1 - K$; as $1/r_3 \to 0$, $\Gamma_{b3} \to 1 - 1/K$.

6. Nonideality, critique of the model

Binding equilibria are normally studied with values of K'' in the 10^2–10^6 range and half saturation occurs at a concentration near $1/K''$ molal. If the binding components are the only solutes, the solutions may usually be considered to be ideally dilute. If, on the other hand, there are other components at high concentration which give activity coefficients different from unity, the change in conditions caused by the addition of 10^{-6}–10^{-2} molal of the binding species normally has little effect on activity coefficients which can be considered as constants and incorporated into the equilibrium constant. This is presumably the justification for the widespread use of concentrations rather than activities in practical calculations.

For weak interactions such as those involving the denaturation or stabilization of proteins in solution, the situation is quite different. Concentrations vary over wide ranges, often up to 10

molar or molal and activity coefficients vary considerably. As a result K' (or K'') can be a relatively strong function of concentration. Since with nonsymmetric reference systems, the activity coefficient of component 1 normally decreases when that of component 3 increases [10] (and vice versa), the variation of K' (or K'') with concentration can be considerably larger than that of the activity coefficients themselves, since it depends on the factor f_3/f_1.

The variation of $K'-1$ with concentration is illustrated in fig. 3 for three different types of solutions. These are: (1) urea for which the solvent water shows positive deviations from ideality; (2) sucrose for which the deviations are negative; and (3) guanidinium chloride, an ionic compound with positive deviations.

For urea and guanidinium chloride, values of K greater than unity have been chosen because positive selective interaction is required to account for the denaturing properties of these substances, whereas values less than unity have been chosen for sucrose in accord with the data of Lee and Timasheff [14] which shows that this substance has a negative selective interaction that leads to protein stabilization. In constructing fig. 3, arbitrary values of K were selected for representation, but the activities were obtained from careful thermodynamic studies of the respective solutions of urea [15], sucrose [16], and guanidinium chloride [17].

Fig. 3 shows that in moderately concentrated solutions ($> 0.5m$), $K'-1$ always deviates rather strongly from $K-1$. The most striking effect is where $K'-1$ changes sign. For K between 1 and 1.3 (not shown), the curves for urea cross the abscissa, going from positive selective interaction to negative selective interaction *as the concentration is increased*. Conversely, the curve for sucrose at $K=1/3$ starts with negative selective interaction and switches to positive selective interaction above $5.5m$. As can be seen from the graphs, slightly higher values of K will crossover at lower molalities.

Even stronger changes in binding properties with concentration are observed with guanidinium chloride. The Donnan contribution, given by the first term of eq. 16, has been omitted from the

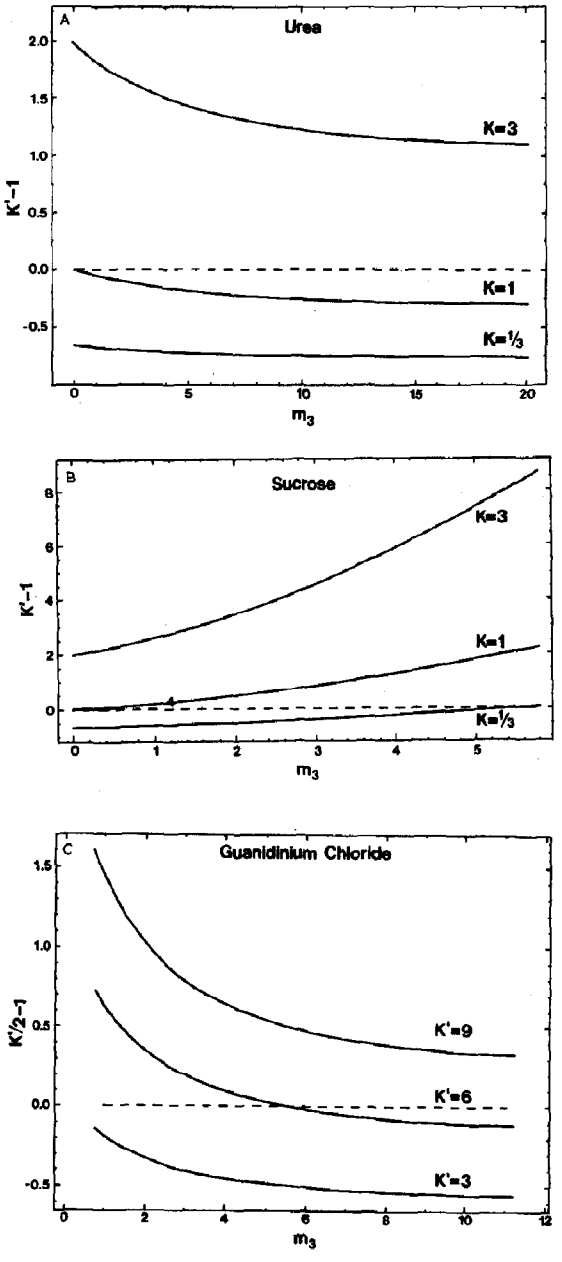

Fig. 3. The effect of nonideality on the apparent binding parameter $K'-1$ (or $(K'/2)-1$ for guanidinium chloride). The curves are based on assumed values of K and the experimental activities of the aqueous solutions. (A) Urea, data of Ellerton and Dunlop [15]; (B) sucrose, data of Lee and Timasheff [14]; (C) guanidinium chloride, data of Schrier and Schrier [17]. Note that varying the activities can produce selective interaction even when $K=1$, and that the sign of the selective interaction can change as a function of concentration.

graphs of fig. 3C. This term varies only slowly with molality of component 3. Since it vanishes at zero charge on the macromolecule, the curves of fig. 3C can be taken as those appropriate for the binding to a site of a neutral macromolecule. Note how the binding constant is attenuated by the various factors in $(K'/2) - 1$. We use $K = 6$ as an example and suppose that this value results from a binding constant of 5 for the guanidinium ion and unity for the chloride ion, i.e., no selective interaction for the chloride. The division by 2 is needed because it is the mean of the two binding events that is required. Guanidinium chloride solutions are highly non-ideal and the factor $f_{\pm}/f_1$ is less than 0.6 even at $0.75m$, the lowest concentration investigated by Schrier and Schrier [17]. Finally unity must be subtracted to account for the displacement of a water molecule. The net result is that a guanidinium binding constant of 5 is reduced to an effective value of 0.73 at this concentration and switches to negative selective interaction above $5.5m$.

There is evidence that the selective interaction of denatured proteins with guanidinium chloride follows the anomalous pattern described above, namely, that at high concentrations the selective interaction decreases with increasing concentration, leading to an eventual crossing to negative selective interaction. This will be discussed in section 7.

One of the more interesting aspects of figs 3 and 4 is that systems with $K = 1$, i.e., no selective interaction on an activity basis, acquire negative interaction for urea and guanidinium chloride and positive interaction for sucrose as the concentration increases. The origins of this effect lie not in the binding event, but in the bulk solution. It is implicitly assumed in the model described in section 2 that the chemical potential of the bound molecules is not affected by changes in the composition of the solution. The chemical potential of the molecules in solution, on the other hand, undergo changes represented by the $\ln f_3$ or β terms. This results in an imbalance between the average composition on the site and the composition of the solution, i.e., there is selective interaction. Since molecules occupying a site on the macromolecule are also in contact with the solution, it is to be expected that their chemical potential is also affected by solution composition, albeit to a lesser degree. Since, however, the effect of solution composition on molecules bound to macromolecules is an unknown territory, it is difficult to incorporate into a simple model. One might try to handle it by assuming an excess free energy, dependent on m_3, for a molecule bound to the site. This is equivalent to assuming an activity coefficient for bound molecules, which is dependent on solution concentrations. The simplest model is one in which the excess free energy of a molecule on the site is proportional to the excess free energy in solution, i.e., $\beta'_{bj} = d\beta_{bj}$, $j = 1, 3$, where β'_{bj} is the excess free energy of a molecule of j on a site, β_{bj} its excess free energy in solution and d a proportionality constant. If the proportionality constant is the same for both components 1 and 3, then it is easy to show that the factor f_3/f_1 in the previous relations must be replaced by $(f_3/f_1)^{1-d}$. The result is an attenuation of the nonideality effects shown in fig. 3. If $d = 0$, the bound molecules are not affected by the solution composition. If $d = 1$, the effect on bound molecules is the same as on molecules in solution and nonideality effects cancel.

Such an attenuation factor may be useful eventually, if experiment shows that K', evaluated as Kf_3/f_1, exaggerates the effects of solution activities on the selective interactions, but for the purposes of the present paper d will be assumed to be zero.

The discussion of fig. 3 has brought out one possible deficiency of the model described in section 2, namely, that it may exaggerate the effect of solution activities on the binding event. There are two other aspects of the model which should be considered whenever it is used for the interpretation of experiments. The first is that only one-to-one replacements of molecules on sites are postulated. Molecules like urea, sucrose and the guanidinium ion are larger than water molecules and, depending on binding geometry, can replace more than one water molecule (fig. 1). The second is the assumption of site independence. High-resolution diffraction experiments have revealed fixed geometric clusters of water molecules on the surface of proteins [18] and nucleic acids [19]

which could behave cooperatively if one or more of the water molecules is replaced by another ligand. These objections can be met by developing the solvent surface layer of the protein as a virial series in $\bar{\theta}_3$, in the same way that solution phases are developed in series of powers of the concentration. This will be done in a later publication, but the parameters of this type of representation do not have the clarity and simplicity of interpretation of the independent site binding model.

7. Examples

In this section model calculations will be performed for four systems: aqueous solutions of urea, sucrose, guanidinium chloride and glucose. These have been selected because they have been intensively studied and have established interactions with proteins. Urea and guanidinium salts are the most common denaturing reagents for proteins, whereas sucrose and glucose by contrast are stabilizing agents [14]. It is a thermodynamic consequence of the theory of denaturation that positive selective interaction leads to the unfolding of proteins (the exposure of interior residues to the solvent lowers the free energy), while the protective action of sucrose and glucose is associated with negative selective interaction. This is sometimes expressed by saying that proteins have preferential hydration in sugar solutions (see, however, the discussion of figs 4 and 6 below).

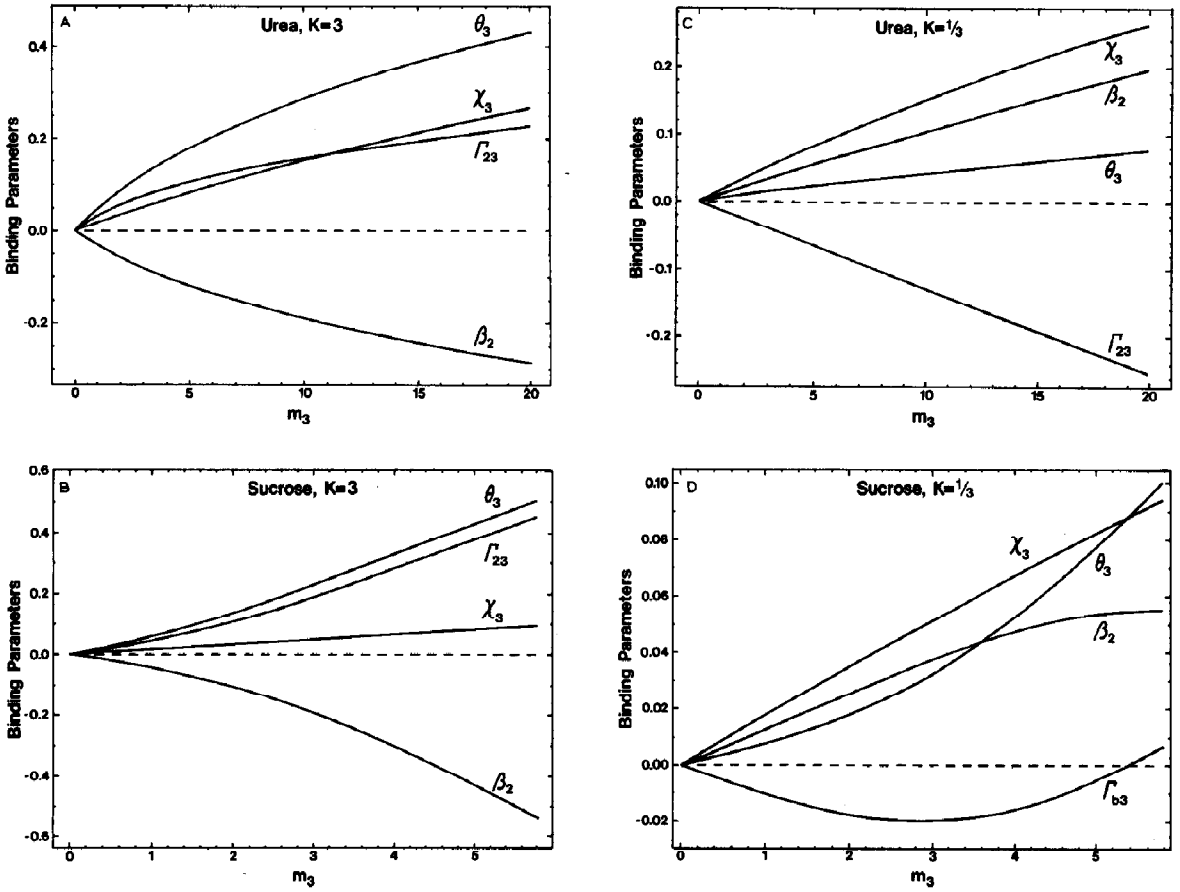

Fig. 4. The effect of K and solution properties on the binding parameters. Urea-water is taken as an example of a solution showing positive deviations from ideality, sucrose-water as an example of negative deviations. See fig. 2 for notation.

The figures of this section are a partial exploration of the dependence of selective interaction on the strength of the interaction as measured by the solvent-exchange constant K and the solution properties of the three reagents. As discussed in section 6, the interaction with this class of substances is sufficiently weak that high concentrations must be used, and the lack of ideality plays an important part in determining the shape of the isotherm. In the calculations K is taken as a parameter for which values are assumed, and the activity coefficients are calculated from the measurements of Ellerton and Dunlop [15], Robinson and Stokes [16], Schrier and Schrier [17] and P.N. Henrion cited in ref. 20, respectively, for urea, sucrose, guanidinium chloride and glucose. The lack of smoothness of some of the curves results from the use of the spaced experimental points of the original investigation.

7.1. Urea and sucrose

The calculated selective interaction properties of urea and sucrose are shown in figs 4 and 5. These are set up to contrast the behavior of urea with positive deviations from ideality with those of sucrose which has negative deviations. The cases for $K > 1$ and $K < 1$ are shown for both substances, though it is usually the case that $K > 1$ for urea and $K < 1$ for sucrose when protein solutions are considered. In fig. 4 curves for χ_3 are included as a reference so that one can compare the composition on the site, θ_{b3}, and the selective interaction at a site, Γ_{b3}, with the composition in

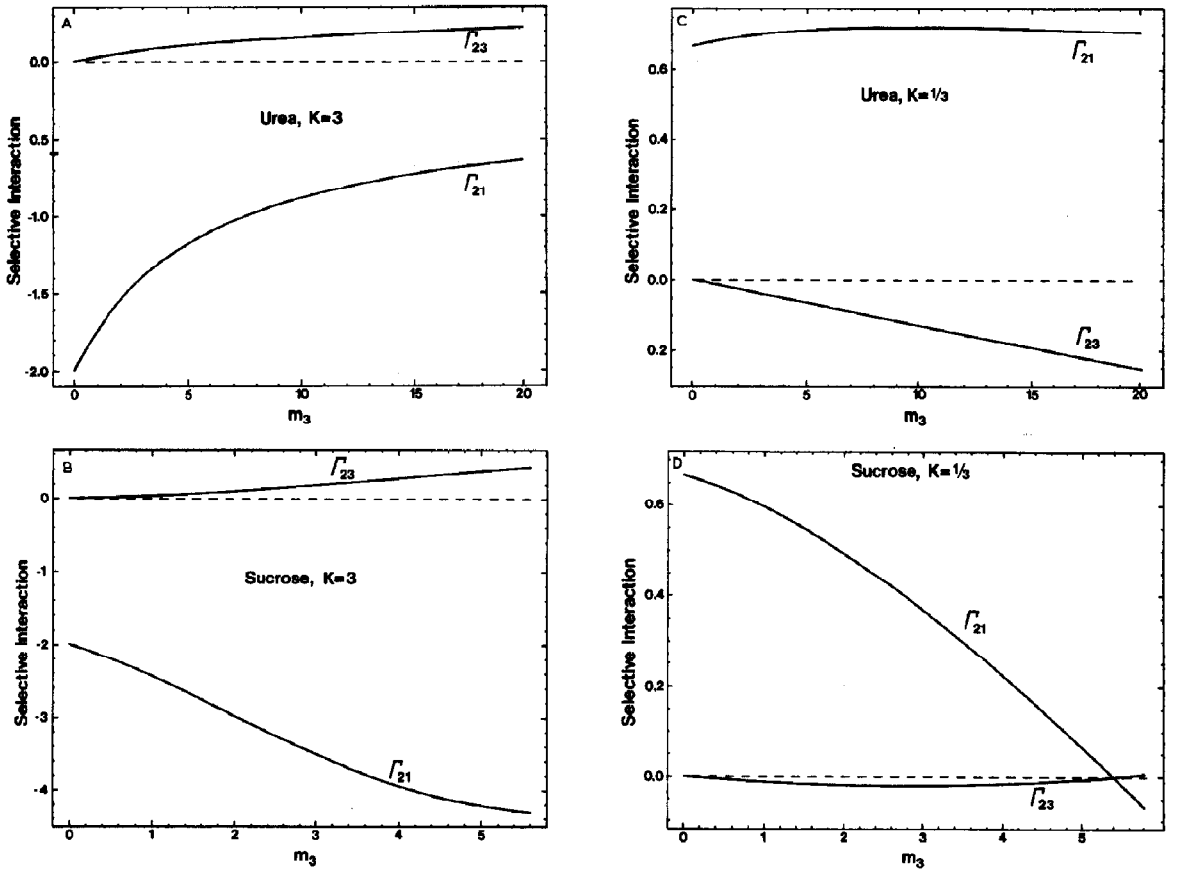

Fig. 5. The preferential hydration of water, Γ_{b1}, in solutions of urea and sucrose. Γ_{b3} is included to permit comparisons with fig. 4.

solution. Values are also provided for the excess free energy, β_2, which can be calculated from either of the formulas [21]

$$\beta_2 = \ln A_1 + \ln(1 + K' r_3)$$
$$= \ln f_1 + \ln(1 + (K' - 1)\chi_3) \qquad (23)$$

This figure demonstrates the strong effect of the activity properties of the solvent mixture on the thermodynamic properties of component 2. For most of the curves the curvatures associated with solutes like urea are opposite to those for solutes like sucrose. This affects not only the shapes of the curves but also the magnitude of the effects. We list some of the more interesting features of these curves. χ_3 is identical in the figures and serves as a reference.

(1) Though the selected values of K represent a 3 : 1 preference of one solvent component or the other, the equilibrium constants are quite small when converted to the molal scale. For example, $K = 3$ implies a molal constant of about 0.06. Significant solvent selection can be concealed in very small values of K.

(2) θ_{b3} is always greater than Γ_{b3}. This occurs because Γ_{b3} operates against the background ('Contrast' in Eisenberg's terminology [7]) of the solution composition while θ_{b3} does not.

(3) $\theta_{b3} > \chi_3$ for $\Gamma_{b3} > 0$ and $\theta_{b3} < \chi_3$ for $\Gamma_{b3} < 0$. This arises directly from the definition of selective interaction.

(4) For equivalent situations the curvature of the urea curves (positive deviation) is generally opposite in sign to the sucrose curves (negative deviation).

(5) Because of this curvature the selective interaction can go through an extremum and change sign. This occurs for urea when $K > 1$ and for sucrose when $K < 1$, which are the usual cases of interest (see also guanidinium chloride in fig. 6).

(6) β_2 usually has the opposite sign to Γ_{b3} but not necessarily. From eq. 6 it is clear that β_2 is the integral of Γ_{b3} with respect to the activity of 3. As such, it goes through a maximum or a minimum when $\Gamma_{b3} = 0$. β_2 and Γ_{b3} have the same sign either below or above this extremum. *In general, one cannot tell the effect of the interaction on the protein or other biopolymer from one measurement*

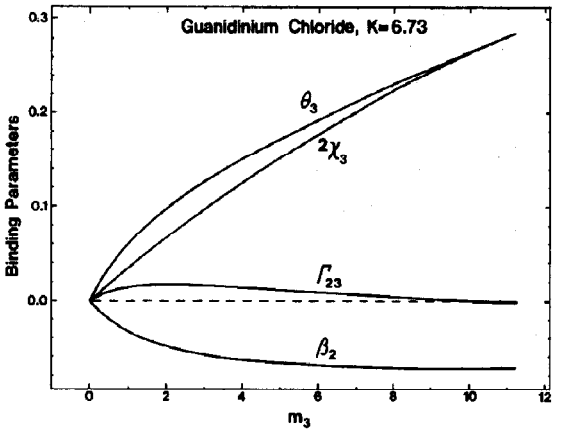

Fig. 6. The selective interaction properties of aqueous solutions of guanidinium chloride. Note that K represents the sum of the selective interaction constants of the two ions and that the conventional mole fraction has been multiplied by 2 to give the particle mole fraction. K has been assigned the value of 6.73 to simulate observed phenomena. See text.

of the selective interaction. Single-point determinations of selective interaction have little value in determining the free energy change caused by an added solute when the binding is weak enough to require the present analysis.

(7) β_2 has a component which depends on the activity of the principal solvent regardless of the selective interaction process. This can be seen from the second form of eq. 23. Suppose under certain circumstances $K' - 1 = 0$, so that there is no selective interaction. Then $\beta_2 = -\ln f_1$. This contribution is not negligible compared with weak interactions. It has the following physical origin. We assume that component 1 is water. The standard free energy of the site (where $\beta_2 = 0$) has been taken as the free energy of a hydrated site in pure water, i.e., $f_1 = 1$. Its free energy of hydration in a concentrated solution of component 3 is altered. This arises not because there is a difference in the hydrated site, but because there is a difference in the free energy of the water in the solution (changes in free energy at the site are certainly possible, but they are not contained in the model; See sections 2 and 6). A practical example is the following. We compare the free energy of unfolding of a protein in water and in a concentrated solution of urea. We assume site b to be essentially hydrated in both cases. The free

energy of hydration is different because the water molecule originates from pure water in the first case ($A_1 = 1$) and from a concentrated solution in the other ($A_1 > 1$). If this aspect of the model turns out to be correct for real systems, then all solvated sites contribute to the free energy of unfolding whether there is selective interaction or not.

(8) Note that for $K = 1/3$, sucrose (or any other substance with this type of solution activity) has an excess free energy and selective interaction much smaller in magnitude than those of urea. In contrast, the magnitudes of these quantities for $K = 3$ are larger for sucrose than for urea, though here the effect is smaller. These trends have their origin in the fact that f_3/f_1 is less than unity for urea and greater than unity for sucrose.

Fig. 5 provides a comparison of the selective interaction of a site with the principal solvent and with component 3. The systems are the same as in fig. 4. The main interest of these curves is that they provide a description of the preferential hydration in aqueous solution, which is often measured and discussed. In addition they demonstrate the limiting way that $\Gamma_{b1} \to -(K' - 1)$ and $m_3 \to 0$. Macromolecules will usually possess numerous heterogeneous sites for weak interactions. In the limit of $m_3 \to 0$, the total selective interaction will be

$$\Gamma_{21}(0) = \sum_b \Gamma_{b1}(0) = -\Sigma(K' - 1)$$
$$= -n(K'_{av} - 1) \qquad (24)$$

where K'_{av} is the average of the site exchange constants and n the number of sites. Thus, if one knows, or can make an estimate of the number of sites, it is possible to evaluate K'_{av}. The same information is available from the limiting slope of Γ_{b3} vs r_3 or m_3.

7.2. *Guanidinium chloride*

The selective interaction properties of guanidinium chloride are shown in fig. 6 for $K > 1$. The value of K was chosen to match the crossover of Γ_{b3} to negative values with observations made on ribonuclease A (see below). Qualitatively, the curves are similar to those for urea (because ($f_3/f_1 < 1$) but the effects are more extreme because of greater deviations from ideality. In this case there are some experimental data with which the predictions of the model can be compared. The following studies all refer to the selective interaction of guanidinium chloride with proteins.

In 1967, Hade and Tanford [22] studied the selective interaction of a series of proteins in 6 M guanidinium chloride (10.6m). Selective interaction was small but positive in most cases, but was zero for ribonuclease. This result was totally unexpected since at that time the denaturation of ribonuclease was one of the best established biophysical reactions and it was universally believed that the unfolding was promoted by the favorable interaction of the unfolded form with guanidinium ions.

In 1969, Reisler and Eisenberg [23] studied the selective interaction of unfolded rabbit muscle aldolase as a function of concentration using the densimetric method and found that above 3 M (3.8m) it decreased as the concentration of guanidinium was increased and extrapolated to zero at about 6.5 M (12.4m).

In 1974, Lee and Timasheff [24] studied the selective interaction of 12 proteins at 6 M (10.6m) using the densimetric method and confirmed Hade and Tanford's results for ribonuclease and a number of other proteins. Unfolded lima bean trypsin inhibitor was also shown to have essentially zero selective interaction with guanidinium chloride at 6 M.

In 1977, Reisler et al. [25] studied unfolded bovine serum albumin as a function of concentration and found that selective interaction decreased with increasing concentration above 3 M and extrapolated to zero at about 7 M (14.2m).

We see in fact that the experimental results do match qualitatively those predicted from the model in two main aspects: the decrease in selective interaction at high concentrations and the ultimate passage through zero. We see further that there is no anomaly in the fact that zero selective interaction can lead to the unfolding of proteins. It is the free energy of interaction, not the selective interaction, that determines the stability of a system in mixed solvents. Comparison of the

curves for Γ_{b3} and β_2 for urea in fig. 4 and guanidinium chloride in fig. 6 shows that β_2 reaches a minimum when $\Gamma_{b3} = 0$. Thus the condition that $\Gamma_{b3} = 0$ (for sites exposed as a result of the unfolding) is the *optimum* condition for the use of a denaturing agent. This arises because β_2 is the integral of Γ_{b3} with respect to concentration and therefore has an extremum when Γ_{b3} is zero.

The results of the groups of Tanford, Eisenberg and Timasheff cited above cannot be directly applied to the problem of the unfolding of proteins in solutions of guanidinium chloride, because they pertain to the unfolded protein molecules only. Solvent denaturation depends on the difference in interaction free energy of the folded and unfolded forms, and this can only be obtained by studying selective interaction over the entire range of concentration including the transition of the protein. T. Lilley (personal communication) has undertaken this type of investigation.

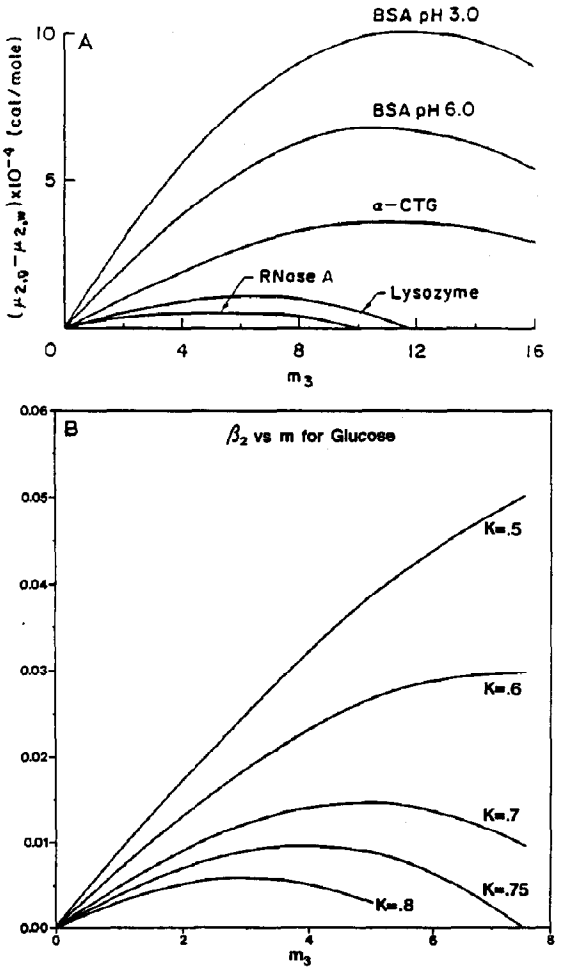

Fig. 7. Excess free energies for proteins in aqueous glucose solutions. (A) The excess free energy of several proteins in glucose solutions [26]. (B) The excess free energy β_2 in glucose solutions calculated for several values of K.

7.3. Glucose

The selective interaction of glucose with several proteins has been studied by Arakawa and Timasheff [26] who have also integrated eq. 6 to obtain the excess free energy of these proteins in glucose solutions as a function of molality. Their results are displayed in fig. 7A. These curves are extrapolations of the analytic form of β_2 above the experimental data which did not go higher than $3m$. To model these results β_2 was calculated using eq. 23 and the activity data of Henrion [20] for the selected values of K shown in fig. 7B. These curves comprise the full range of accessible data and are not extrapolated. We see substantial agreement between the experiments and the independent site binding model. For both the experimental and theoretical curves the excess free energy tends toward a maximum and for weaker interaction can change sign at higher glucose concentrations. The correlation is only within a scaling factor since fig. 7A contains data for real proteins and fig. 7B is the theory for an effective site. A quantitative comparison will require estimates of the number and nature of the sites and a study of the effect of site heterogeneity.

8. Conclusion

The solvation model described in this paper leads to a description of the effects of mixed solvents on macromolecules which agrees with experimental results on the denaturation and stabilization of proteins and other molecules. This includes phenomena such as denaturation under conditions where there is no selective interaction, the decreasing selective interaction of guanidinium salts and sucrose at high concentrations, and max-

ima or minima in the free energy of stabilization or destabilization which have been considered to be quite puzzling. Quantitative comparisons can only be made if selective interaction studies are performed over the full range of concentrations. This type of data is now available for a number of stabilizing substances thanks to Timasheff and his collaborators, but is lacking for denaturing agents, especially at low concentration. The likely next steps toward progress in this field are: (1) the collection of more data; (2) studies on model substances (proteins, especially, have very heterogeneous surfaces); and (3) theoretical considerations of heterogeneity and cooperativity.

Appendix: Three measures of selective interaction

Methods developed by Casassa and Eisenberg [6] will be used to obtain formulas for the selective interaction of an ionic component with a polyelectrolyte. Component 3 is a univalent ionic substance sharing a common ion with component 2. The chemical potentials are

$$\mu_2 = \mu_2^\circ + RT \ln A_2 = \mu_2^\circ + RT \ln m_2 + |Z| RT \ln(m_3 + |Z| m_2) + \beta_2 \quad \text{(A1)}$$

$$\mu_3 = \mu_3^\circ + RT \ln A_3 = \mu_3^\circ + RT \ln m_3 + RT \ln(m_3 + |Z| m_2) + \beta_3 \quad \text{(A2)}$$

where $|Z|$ is the magnitude of the average charge on the polyion of component 2. At constant T

$$d\mu_1 = \bar{V}_1 \, dP + RTa_{12} \, dm_2 + RTa_{13} \, dm_3 \quad \text{(A3)}$$

$$d\mu_2 = \bar{V}_2 \, dP + RTa_{22} \, dm_2 + RTa_{23} \, dm_3 \quad \text{(A4)}$$

$$d\mu_3 = \bar{V}_3 \, dP + RTa_{32} \, dm_2 + RTa_{33} \, dm_3 \quad \text{(A5)}$$

with $a_{jk} = (\partial\mu_j/\partial m_k)_{T,P,m}$ and $a_{jk} = a_{kj}$. From ref. 7

$$a_{22} = \frac{1}{m_2} + \frac{|Z|^2}{m_3 + |Z| m_2} + \beta_{22}$$

$$a_{23} = a_{32} = \frac{|Z|}{m_3 + |Z| m_2} + \beta_{23}$$

$$a_{33} = \frac{1}{m_3 + |Z| m_2} + \frac{1}{m_3} + \beta_{33}$$

We now consider the three closely related forms of the selective interaction parameter, which have been discussed by Eisenberg.

A1. Case 1: Constant μ_3 and P (Γ'_{23})

These conditions arise in the theory of fluctuations [27,28] and in experimental situations such as electrochemical cells [7] where the chemical potential of component 3 is constant but not the activity of the principal solvent. The selective interaction for this case is given by $\Gamma'_{23} = (\partial m_3/\partial m_2)_{T,P,\mu_3}$

$$\Gamma'_{23} = -\frac{a_{32}}{a_{33}} = -\frac{\dfrac{|Z|}{m_3 + |Z| m_2} + \beta_{23}}{\dfrac{1}{m_3 + |Z| m_2} + \dfrac{1}{m_3} + \beta_{33}}$$

$$\approx -\frac{|Z| + m_3\beta_{23}}{2 + m_3\beta_{33}} \quad (\text{for } m_2 \to 0) \quad \text{(A6)}$$

The form at the right is the limit for the low concentrations of the macromolecule which are of interest in this paper. Note that for ideal solutions the formula reduces to the Donnan value $-|Z|/2$.

A2. Constant μ_1 and P (Γ''_{23})

These are the conditions for isopiestic measurements [10]. From eq. A3

$$\Gamma''_{23} = \left(\frac{\partial m_3}{\partial m_2}\right)_{T,P,\mu_1} = -\frac{a_{12}}{a_{13}} \quad \text{(A7)}$$

To develop this relation we use the Gibbs-Duhem equation at constant T and P,

$$m_1 \, d\mu_1 + m_2 \, d\mu_2 + m_3 \, d\mu_3 = 0.$$

Substituting from eqs A3–A5,

$$(m_1a_{12} + m_2a_{22} + m_3a_{32}) \, dm_2 + (m_1a_{13} + m_2a_{23} + m_3a_{33}) \, dm_3 = 0$$

where m_1 is the number of moles of component 1 per kg of component 1. Equating the quantities in parentheses to zero we find

$$\Gamma''_{23} = -\left[1 + m_2\left(\frac{|Z|^2}{m_3 + |Z| m_2} + \beta_{22}\right) + m_3\left(\frac{|Z|}{m_3 + |Z| m_2} + \beta_{23}\right)\right]$$

$$\times\left[m_2\left(\frac{|Z|}{m_3+|Z|m_2}+\beta_{23}\right)+1\right.$$
$$\left.+m_3\left(\frac{1}{m_3+|Z|m_2}+\beta_{33}\right)\right]^{-1}$$
$$\approx -\frac{1+|Z|+m_3\beta_{23}}{2+m_3\beta_{23}}=\Gamma'_{23}-\frac{1}{2+m_3\beta_{33}}$$
$$(\text{for } m_2\to 0) \qquad \text{(A8)}$$

The last term is usually small as discussed by Hade and Tanford [22].

A3. Constant μ_1 and μ_3 (Γ_{23})

This condition is appropriate for dialysis experiments. P cannot remain constant when m_2 is varied under these conditions (osmotic pressure). From eq. A5

$$\Gamma_{23}=\left(\frac{\partial m_3}{\partial m_2}\right)_{T,\mu_1,\mu_3}=-\frac{a_{23}}{a_{33}}-\frac{\bar{V}_3}{RTa_{33}}\left(\frac{\partial P}{\partial m_2}\right)_{T,\mu_1,\mu_3}$$
$$\approx \Gamma'_{23}-\frac{\phi_3}{2+m_3\beta_{33}} \qquad \text{(A9)}$$

Or, alternatively, from eq. A3

$$\Gamma_{23}=\left(\frac{\partial m_3}{\partial m_2}\right)_{T,\mu_1,\mu_3}=-\frac{a_{12}}{a_{13}}-\frac{\bar{V}_1}{RTa_{13}}\left(\frac{\partial P}{\partial m_2}\right)_{T,\mu_1,\mu_3}$$
$$\approx \Gamma''_{23}+\frac{\phi_1}{2+m_3\beta_{33}} \quad (\text{for } m_2\to 0) \qquad \text{(A10)}$$

The quantities ϕ_3 and ϕ_1 in eqs A9 and A10 are the volume fractions of 3 and 1 calculated from partial molar volumes. Since volume fractions cannot exceed unity, the rightmost terms of eqs A7 and A8 are usually small and at most approach unity. Selective interaction is difficult to measure with high precision. It is usually admissible to assume that $\Gamma_{23}\cong\Gamma'_{23}\cong\Gamma''_{23}$. The correction terms can be calculated if partial molar volumes and activities are known for the solvent system.

We note that the independent variables in this appendix have been the molalities of the components whereas they were the activities in the main body of the text. Translation from one set to the other is relatively simple. We note that

$$\left(\frac{\partial}{\partial \ln A_3}\right)_{T,P,m_2}=\left(\frac{\partial m_3}{\partial \ln A_3}\right)_{T,P,m_2}\left(\frac{\partial}{\partial m_3}\right)_{T,P,m_2}$$
$$=\frac{1}{a_{33}}\left(\frac{\partial}{\partial m_3}\right)_{T,P,m_2}$$

so that $\beta_{23}/a_{33}=(\partial\beta_2/\partial \ln A_3)$. Consequently, eq. A6 can be written in the form

$$\Gamma'_{23}=\frac{-|Z|}{2+m_3\beta_{33}}-\left(\frac{\partial\beta_2}{\partial \ln A_3}\right)_{T,P,m_2} \qquad \text{(A11)}$$

This is the formula that will be used for the calculation of the selective interaction of ionic substances. The term on the right has been used for nonionic substances in previous papers. Reduction to this form is accomplished by putting $Z=0$. The integer 2 in the denominator of the first term arises from the assumption that component 3 is a 1:1 electrolyte. Relations for dealing with other ionic valencies can be obtained by starting with the correct expression in eq. A2 and repeating the subsequent derivation.

Note added in proof

Recently, Nilsson et al. [29] have also extended the solvent binding model, but in a different way. Their approach is to take into account the interactions amongst the macromolecule and the two solvents by a mean field calculation.

Acknowledgements

This research was supported by NSF Grant 8609113 and NIH Grant GM20195. The author has benefited from discussions with Robert Baldwin, Henryk Eisenberg, Stanley Gill and Serge Timasheff on the solvation properties of proteins and nucleic acids and with Wayne Bolen and Nick Pace on the thermodynamic behavior of proteins in denaturing agents. The manuscript was considerably improved by the criticisms and suggestions of Charlotte Schellman.

References

1 J. Wyman, J. Am. Chem. Soc. 58 (1936) 1482.
2 L. Onsager, J. Am. Chem. Soc. 58 (1936) 1486.
3 J.G. Kirkwood, J. Chem. Phys. 7 (1939) 911.
4 J. Wyman, J. Mol. Biol. 11 (1965) 631.
5 C.N. Pace and G.R. Grimsley, Biochemistry 27 (1988) 3242.

6 E.F. Casassa and H. Eisenberg, Adv. Protein Chem. 19 (1964) 287.
7 H. Eisenberg, Biological macromolecules and polyelectrolytes in solution (Clarendon, Oxford, 1976).
8 J.A. Schellman, Biopolymers 14 (1975) 999.
9 J.A. Schellman, Biopolymers 17 (1978) 1305.
10 R.A. Robinson and R.H. Stokes, Electrolyte solutions (Butterworth, London, 1955).
11 P.H. von Hippel and K. Wong, J. Biol. Chem. 240 (1965) 3909.
12 T. Arakawa and S.N. Timasheff, Biochemistry 23 (1984) 5924.
13 H. Inoue and S.N. Timasheff, Biopolymers 11 (1972) 737.
14 J.C. Lee and S.N. Timasheff, J. Biol. Chem. 256 (1981) 7193.
15 H.D. Ellerton and P.J. Dunlop, J. Phys. Chem. 70 (1966) 1831.
16 R.A. Robinson and R.H. Stokes, J. Phys. Chem. 65 (1961) 1954.
17 M.Y. Schrier and E.E. Schrier, J. Chem. Eng. Data. 22 (1977) 73.
18 M.M. Teeter, Proc. Natl. Acad. Sci. U.S.A. 81 (1984) 6014.
19 M.L. Kopka, A.V. Fratina, H.R. Drew and R.B. Dickerson, J. Mol. Biol. 163 (1983) 129.
20 R.H. Stokes and R.A. Robinson, J. Phys. Chem. 70 (1966) 2126.
21 J.A. Schellman, Biopolymers 26 (1987) 549.
22 E.P.K. Hade and C. Tanford, J. Am. Chem. Soc. 89 (1967) 5034.
23 E. Reisler and H. Eisenberg, Biochemistry 8 (1969) 4572.
24 J.C. Lee and S.N. Timasheff, Biochemistry 13 (1974) 257.
25 E. Reisler, Y. Haik and H. Eisenberg, Biochemistry 16 (1977) 197.
26 T. Arakawa and S.N. Timasheff, Biochemistry 21 (1982) 6536.
27 J.G. Kirkwood and R.J. Goldberg, J. Chem. Phys. 18 (1950) 54.
28 W.H. Stockmayer, J. Chem. Phys. 18 (1950) 58.
29 S. Nilsson, L. Picullel and M. Malmsten, J. Phys. Chem. (1990) in the press.

Glossary

1 – the principal solvent.
2 – the macromolecular component.
3 – an added third component.
A_i, f_i – activity and activity coefficient on the mole fraction scale with the ideal dilute solution as reference system; i.e., f_1, $f_3 \rightarrow 1$ as $\chi_1 \rightarrow 1$.
K_b – equilibrium constant for occupancy at site b on mole fraction scale.
K' – practical equilibrium constant, equal to $f_3 K_b / f_1$, mole fraction scale.
K'' – practical equilibrium constant on molal scale.
m – molality, m_1 – the (constant) molality of the principal solvent = $1000/M_1$.
N – number of binding sites.
r_3 – mole ratio of 3 in solution = $m_3/m_1 = \chi_3/\chi_1$.
r_{3b} – mole ratio of 3 on site b.
w_3 – mass ratio concentration of component 3.
Z – charge on a macromolecule.
β_2° – free energy of 2 in dilute solution in units of RT.
Γ_{2j} – selective interaction of component j with component 2. Primes and double primes on Γ indicate differences in the definition of selective interaction which are usually immaterial. See appendix for explicit relations.
Γ_{bj} – selective interaction of component j with site b.
θ_{b3}, θ_{b1} – fraction of occupation of site b by 3 and 1 respectively.
ξ_j – mass ratio selective interaction of component j.
χ – mole fraction.