ELSEVIER

Biophysical Chemistry 101–102 (2002) 99–111

Biophysical Chemistry

www.elsevier.com/locate/bpc

Thermodynamic binding and site occupancy in the light of the Schellman exchange concept

Serge N. Timasheff*

Biochemistry Department, MS 009, 415 South Street, Waltham, MA 02453-2728, USA

Abstract

An analysis of Schellman's treatment of preferential interactions is presented, as viewed by a laboratory practitioner of the art. Starting with an intuitive description of what binding is in terms of the distribution of molecules of water and of a weakly interacting ligand (co-solvent), Schellman proceeded to a rigorous thermodynamic definition in which he showed that classical, dialysis equilibrium, binding is a purely thermodynamic quantity. Putting water and the co-solvent on an equivalent footing, he showed that the classical binding treatment is inadequate for weakly interacting systems, in which the replacement of water by ligand and exclusion of co-solvent are symmetrical concepts. Analyzing specifically the simple model of a single independent site, Schellman demonstrated how a positive binding constant can give rise to a measured negative binding stoichiometry. He showed that the origin of the complicated binding isotherms is the non-idealities of water and co-solvent, and went further to analyze critically the effect of site heterogeneity on the ligand concentration dependencies of site occupancy, preferential binding and the thermodynamic quantities, enthalpy, entropy and Gibbs free energy. This exposition of the Schellman treatment is accompanied by illustrations drawn from the experimental results obtained in this author's laboratory.

Keywords: Hydration; Preferential binding; Preferential hydration; Co-solvents; Water–ligand exchange

1. Introduction

It is with particular pleasure that I welcome this opportunity to participate in this issue dedicated to John Schellman. While our laboratory has been reporting experimentally measured values of equilibrium binding, i.e. preferential binding, for close to three decades, we were initially puzzled by our finding that a number of protein stabilizing agents resulted in measured negative stoichiometries of binding. The first such results obtained with aqueous sucrose solutions [1] led us to the realization that the protective action of this sugar must be due to the fact that it must give even more negative binding stoichiometries with unfolded protein. Nevertheless, the relation between the negative binding stoichiometries and binding equilibrium constants escaped us until John Schellman in 1987 [2] started publishing his theoretical papers in which he elegantly developed the pertinent multicomponent thermodynamics that gave a lucid explanation of how it is that a positive equilibrium constant can lead to negative stoichiometries being measured in equilibrium binding experiments [2–

*Tel.: +1-781-736-2369; fax: +1-781-736-2349.
E-mail address: timasheff@brandeis.edu (S.N. Timasheff).

0301-4622/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0301-4622(02)00188-6

4] and how the variation of activity coefficients can modulate not only the value, but actually the sign of the measured binding [3,4]. It is in the light of this personal perspective that I shall attempt to describe the logical development of Schellman's concept of exchange [2–4] and how it permitted to sharpen the focus in the analysis and the understanding of experimental observations by laboratory practitioners such as this author.

The interaction of proteins with weakly acting ligands has been a subject of interest for a good part of a century, in view of their use as protein denaturants (e.g. urea, guanidine hydrochloride), stabilizers (e.g. sucrose, glycerol), salting out and crystallizing agents (e.g. Na_2SO_4, polyethylene glycol, i.e. PEG). The way to the understanding of the underlying mechanisms of interaction was opened with the development some 50 years ago of the theory of multicomponent solution thermodynamics [5,6]. This made possible to submit the interactions of three component systems (protein–water–co-solvent) to rigorous quantitative analysis in terms of the preferential (or selective) interaction parameter, $(\partial m_3/\partial m_2)_{T,P,\mu_3}$, which Schellman has identified as thermodynamic binding[1] [3,7]. This parameter is identical with the binding measured by dialysis equilibrium or similar techniques, $\bar{v}$ in the Scatchard notation [8], Γ_{23} in the Casassa and Eisenberg notation [9] used by Schellman.

In contrast to strongly binding ligands, such as enzyme effectors, which act at low concentration ($<10^{-3}$ M), it is characteristic of these weakly acting ligands that, in order to exert their action, they must be used at high concentration (typically 0.5–10 M). In fact, these ligands may occupy as much as a 30–40% fraction of the solvent volume. As a consequence, they are being referred to as co-solvents. The requirement of a high concentration excludes the formation of stable complexes at specific sites on the protein. Such a mechanism would imply binding constants <1 M^{-1}, which contradicts stability of complexation. The measured binding must be, then, a consequence of the simultaneous weak interactions of many molecules with diverse loci distributed over a broad area of a protein surface.

[1] In what follows we have adopted the Scatchard notation [8] that component 1 is water, component 2 is protein and component 3 is co-solvent; m_i is the molal concentration of component i and μ_i is the chemical potential of component i. Since all operations are performed at constant temperature and pressure, we will drop the subscripts T and P from the partial derivatives. Furthermore, it is to be understood that all that follows applies to single, isolated protein molecules, a state that can be attained generally by extrapolation of results to zero protein concentration.

2. The puzzle

Dialysis equilibrium measurements of the binding of these kinds of molecules to proteins have resulted in complex patterns of the dependence of stoichiometry on ligand concentration; both positive and negative values of Γ_{23} have been reported ([10] and references therein). For example, guanidine hydrochloride gives predominantly positive values of Γ_{23}, while sugars, polyols and amino acids give negative values, and in the case of urea the sign of Γ_{23} varies with protein [11,12] and actually may change with an increase in urea concentration. At 6 M concentration, the measured binding of guanidine hydrochloride to a number of unfolded proteins is zero or close to zero [13,14]. In three component analysis, positive binding stoichiometries are defined as preferential binding, while negative stoichiometries are defined as preferential exclusion. How to analyze such results?

Classical binding theory states that for n equivalent sites, the binding isotherm is [15]:

$$v = n \frac{K_b C_3}{1 + K_b C_3} \quad (1)$$

where $K_b = [PL]/[P][L]$ is the binding constant for the binding equilibrium, $P + L \rightleftharpoons PL$, and C_3 is the ligand concentration (or activity) in corresponding units. The free energy of binding, ΔG_b^0, is given by:

$$\Delta G_b^0 = -nRT \ln(1 + K_b[L]) \quad (2)$$

The desired parameters, K_b and n, are usually obtained by curve fitting, such as the classical Scatchard plot [16].

Treatment in such manner of equilibrium binding data for weakly interacting co-solvents leads

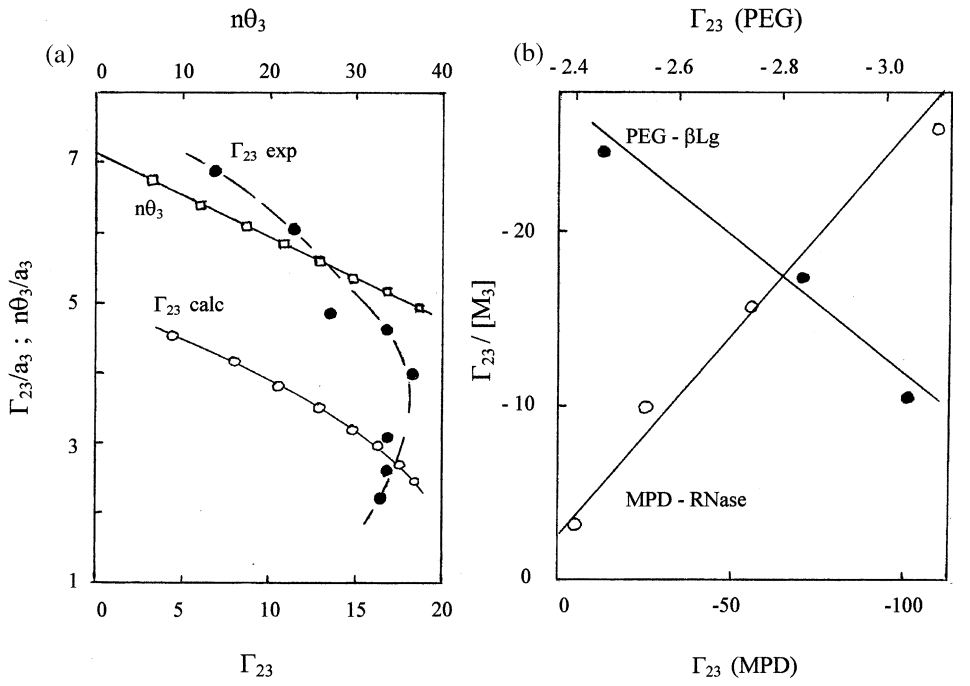

Fig. 1. Scatchard plots of preferential binding (Γ_{23}) of co-solvents. *Part a.* The interaction of urea with lysozyme, pH. 7.7. Filled circles (●): experimental results; squares (□): site occupancy calculated from the calorimetric titration of Makhatadze and Privalov [37]; open circles (○): preferential binding calculated from the calorimetric titration data. For details, see text. *Part b.* Filled circles (●): preferential binding of PEG-1000 to β-lactoglobulin; open circles (○): preferential binding of MPD to ribonuclease A.

to puzzling results. Typical examples are shown in Fig. 1 in the form of Scatchard plots. Fig. 1a depicts the binding of urea to lysozyme at pH 7.7 (Xie and Timasheff, submitted to Biophys. Chem.), where Γ_{23} is positive at all urea concentrations. The experimental data, plotted as filled circles, give the weird pattern in which the plot reverses sense with increasing urea concentration. On the other hand, plots of two systems with negative experimental values of Γ_{23} give reasonable straight lines. Their analysis, however, leads to the bizarre results that the binding of 2-methyl-2,4-pentanediol (MPD) to RNase A [17] is characterized by a positive binding constant, $K_b = 24.4$ M^{-1} and a negative number of sites, $n = -3.7$, while that of PEG-1000 to β-lactoglobulin [18] is characterized by a positive number of sites, $n = 11.3$, but a negative binding constant, $K_b = -0.23$ M^{-1}. In fact, von Hippel et al. in 1973 [19] were the first to point out that classical treatment of negative binding leads to the absurd conclusion that there must be negative binding parameters. What, then, is the proper way of analyzing the measured binding for such weakly interacting systems?

3. The Schellman paradox

The fundamentals of the problem have been stated clearly in the Schellman paradox. Let us follow Schellman's reasoning [2,7]. The problem is presented schematically in Fig. 2, in which a protein molecule is immersed into a mixture of two solvents, say water and a co-solvent. By definition, the protein is totally indifferent to whether it is in contact with molecules of one or the other solvent component. Therefore, at any site on the protein surface, it interacts with identical free energy with water or co-solvent, so that the probability of component 3 occupying any site *X*

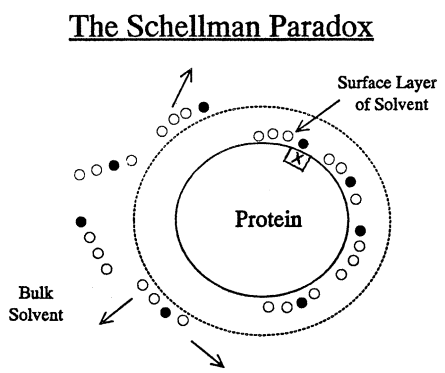

Fig. 2. The Schellman paradox. Dialysis equilibrium of a protein molecule dissolved in water–co-solvent mixture, in which it is indifferent to the component with which it makes contacts. Mixed solvent composition in the bag (surface layer of solvent) and outside the bag (bulk solvent); water molecules (○), co-solvent molecules (●). *X* marks any site, selected at random, on the protein molecule, at which it makes contact with solvent component molecules. For details, see text.

is proportional to its volume fraction in the solvent. This means that the solvent composition on the protein surface is identical with that of the bulk solvent. Therefore, if we were to place a dialysis membrane around the protein and measure the binding at equilibrium, we would obtain the result that the binding, $\Gamma_{23}=0$, because, operationally, $\Gamma_{23}=[(\text{conc 3 in bag})-(\text{conc 3 in bulk})]/(\text{Protein conc})$. On the other hand, ligand (co-solvent) molecules are making contact with the protein surface. This means interaction, i.e. sites on the protein surface are statistically occupied by the ligand, and this occupancy changes directly with solvent composition. Formal analysis in terms of the classical binding theory (Eq. (1)) leads to a non-vanishing binding constant, K_b, and a finite free energy of binding (Eq. (2)) [7]. But this contradicts the dialysis equilibrium result which shows no binding. Obviously, classical treatment fails when the interactions are weak. What is the resolution of this problem?

4. Binding is exchange

As shown by Schellman, the resolution of this paradox is found in the Principle of Exchange [2–4,7]. Let us consider again the premises of the hypothetical system discussed above. It is clear from Fig. 2 that, in aqueous medium, the occupancy of a site by any ligand molecule requires the replacement of water. Now, the system being defined as indifferent to whether contact is with water or co-solvent, the departure of water from the given site is accompanied by a free energy change equal to and opposite in sign to that of its occupancy by the replacing ligand. Hence, the measurement of $\Delta G_b^0=0$. These considerations demonstrate that the classical binding equilibrium, $P+L\rightleftharpoons PL$, is incomplete and that water must be introduced explicitly into the stoichiometry:

$$P\cdot nH_2O+L\rightleftharpoons PL+nH_2O \tag{3}$$

Then, the binding constant, K_b, measured by dialysis equilibrium, is an exchange constant, K_{ex} [2–4,7]:

$$K_b=K_L/K_W\equiv K_{ex} \tag{4}$$

where K_L and K_W are the binding constants of ligand (co-solvent) and water, respectively, in the hypothetical equilibria of their binding to an empty (dry site) on the protein in aqueous medium, clearly physically unattainable situations. Within the exchange concept, the measured binding free energy for a ligand, ΔG_b^0, is [20]:

$$\Delta G_b^0=\Delta G_L^0-\Delta G_w^0 \tag{5}$$

where ΔG_L^0, and ΔG_w^0, are the free energies of binding for the hypothetical equilibria of the binding of ligand and water to dry protein in an aqueous medium. Eq. (5) shows that the sign and magnitude of the measured free energy of binding, ΔG_b^0, is a function of the relative affinities of ligand and water for the site on the protein. Equality of ΔG_L^0, and ΔG_w^0, corresponds to the neutral situation depicted in Fig. 2.

When ΔG_L^0, is more negative than ΔG_w^0, i.e. $K_L>K_W$, there is an excess of ligand (co-solvent) on the protein surface relative to its concentration in the bulk solvent. This defines the term preferential binding [9] of component 3 (ligand). The opposite situation, an insufficiency of ligand, defines the term preferential exclusion [9] of component 3 (negative values of Γ_{23}). Insufficiency of ligand means perforce excess of water, which defines the term preferential hydration. The epithet

preferential, therefore, means that the protein has higher affinity for that particular solvent component than for the other one.

5. Preferential binding is a purely thermodynamic quantity

Schellman, who refers to preferential binding also as 'selective binding' states that [3] '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. 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, substrates, etc. When the interaction is strong, it is identical to the usual molecular definitions of binding. When the interactions are weak or negative it can differ significantly form the molecular description.' This leads to the identification of the parameter $(\partial m_3/\partial m_2)_{\mu_3}$ as thermodynamic binding [3,7]. A simple algebraic rearrangement of this binding parameter gives:

$$\Gamma_{23}=\left(\frac{\partial m_3}{\partial m_2}\right)_{\mu_3}=-\left(\frac{\partial \mu_3}{\partial m_2}\right)_{m_3}\Big/\left(\frac{\partial \mu_3}{\partial m_3}\right)_{m_2} \tag{6}$$

and application of the Euler criterion leads to:

$$\Gamma_{23}=\left(\frac{\partial m_3}{\partial m_2}\right)_{\mu_3}=-\left(\frac{\partial \mu_2}{\partial \mu_3}\right)_{m_2} \tag{7}$$

It is clear, therefore, that Γ_{23} is a purely thermodynamic quantity: it is the mutual perturbation of the chemical potentials of protein and co-solvent, which means that 'a molecule does not have to be in contact with a macromolecule to be bound to it and that binding can be positive or negative' [7]. On the molecular level, thermodynamic binding, $(\partial m_3/\partial m_2)_{\mu_3}$, represents the number of molecules of ligand which must be added to a solution or subtracted from it to restore its chemical potential when a molecule of protein (macromolecule) has been added to the binary solvent (water–component 3) [21]. Once again, there is no requirement of molecular contact. The total free energy of interaction, $\Delta\mu_{2,\mathrm{tr}}$, which is the free energy of transfer of protein from water into the binary solvent, is obtained by integration of Eq. (6) [22]:

$$\begin{aligned}\Delta\mu_{2,\mathrm{tr}}&=\int_0^{m_3}\left(\partial\mu_2/\partial m_3\right)_{m_2}\mathrm{d}m_3\\&=-\int_0^{m_3}\left[\left(\partial m_3/\partial m_2\right)_{\mu_3}\left(\partial\mu_3/\partial m_3\right)_{m_2}\right]\mathrm{d}m_3\end{aligned} \tag{8}$$

6. Equivalence of solvent components

The concept of exchange in terms of Eq. (3) and Eq. (5) has been long understood [20]. Its significance in defining binding stoichiometry, however, was first demonstrated by Schellman in 1987 [2]. In his analysis, Schellman started with the realization that, when dealing with a co-solvent, the two solvent components are equivalent. He put them on equal footing by treating the system in terms of mole fraction concentration units [3,4]. This key intellectual development led to the understanding that interactions with water and co-solvent are symmetrical, i.e. binding and exclusion are symmetrical concepts. Schellman states [4]: 'the selective interaction formalism is completely symmetric with respect to components 1 and 3; the apparent asymmetry of the preferential binding equations comes from the fact that we have chosen component 1 as the principal solvent component and are, therefore, studying the binding of component 3 in this medium.' A corollary to this is: had we chosen component 3 as the principal solvent we would have been measuring the binding of water to the protein in the same medium.

The principles of equivalence and symmetry lead to the realization that, while preferential binding of ligand means preferential exclusion of water, equivalently preferential exclusion of ligand means preferential binding of water, Γ_{21}. The symmetry between the two is expressed by (in molal concentration units):

$$m_1\Gamma_{23}=-m_3\Gamma_{21} \tag{9}$$

7. Relation between exchange constant and binding stoichiometry

The question of exchange has been treated by Schellman in a series of important theoretical

papers [2–4,7]. Working within the framework of an independent single site binding model, Schellman derived the equations for the free energy and stoichiometry of binding within the exchange concept. He used two approaches: first, a qualitative, intuitive approach with the grand partition function, and, second, a rigorous thermodynamic derivation in terms of multicomponent theory [2]. The two give identical results. For a single independent site, *i*, the free energy of binding is:

$$\Delta G^0_{b,i} = \Delta G^0_{ex,i} = -RT \ln(A_1 + K^i_{ex} A_3) \quad (10)$$

where $K_{ex} = K_L / K_W$ and A_1 and A_3 are the water and co-solvent activities in mole fraction units. Introduction of the activity coefficients of water and co-solvent in mole fraction units, f_1 and f_3, and summation over all sites[2] result in the transfer free energy of the protein from water into the mixed solvent [3,4]:

$$\sum \Delta G^0_{ex,i} = -RT \sum_i \left\{ \ln f_1 + \ln \left[1 + \left(K^i_{ex} \frac{f_3}{f_1} - 1 \right) X_3 \right] \right\} = \Delta \mu_{2,tr} \quad (11)$$

where X_3 is the mole fraction of the co-solvent.

The identity of this quantity with the transfer free energy given by Eq. (8) permits the calculation of this quantity from experimental data. Setting $K_{ex}(f_3/f_1) = K'$ and $K_{ex}[f_3/(f_1 m_1)] = K''$, apparent (practical) equilibrium constants in mole fraction and molal concentration units, respectively, gives the result that, for a single independent site [2–4]:

$$\Gamma_{23} = \frac{(K' - 1)X_3}{1 + (K' - 1)X_3} = \frac{(K'' - 1/m_1)m_3}{1 + (K'' m_3)} \quad (12)$$

Eq. (12) shows that, if $K' = 1$, the thermodynamically measured binding, $\Gamma_{23} = 0$; when $K' > 1$, the measured binding is positive; when $K' < 1$, the measured binding is negative. In developing this model, Schellman has cautioned that 'a solvation model based on independent sites is not very realistic and is mainly used as a framework for thinking about the solvation process' [7]. Nevertheless, Eq. (12) is very instructive, as it permits us to understand how a positive binding (exchange) constant can lead to measured negative binding stoichiometries, Γ_{23}. This fact is not intuitively evident and originally was a source of puzzlement to this author.

The definitions of the practical equilibrium constants, K' and K'', shows that they are not true constants but vary with solvent composition because of non-ideality. The factor f_3/f_1 can attain significant values because the concentrations of component 3 are high. Furthermore, the dependence of f_3/f_1 on co-solvent concentration can assume a variety of shapes, e.g. monotonic increase or decrease, maxima or minima; its numerical value can be greater than or smaller than one. Schellman [3,4] has calculated typical co-solvent concentration dependencies of the various interaction parameters using known activity coefficients. All have been seen with real systems. For example, Reisler and Eisenberg [23] have found that the preferential interaction of unfolded serum albumin with guanidine hydrochloride extrapolates to zero at 7 M and presumably becomes negative above that concentration. The measured preferential binding of sorbitol to native ribonuclease A at 48 °C, pH 5.5, is negative below approximately 1.5 m sorbitol and becomes positive above that concentration [24]. This reflects a variation of the transfer free energy, $\Delta\mu_{2,tr}$, which is increasingly positive until it reaches a maximal value at approximately 1.5 m co-solvent, above which it decreases and tends toward zero at 3.5 m. On the other hand, the dependence of Γ_{23} on the concentration of $MgCl_2$ in its interaction with β-lactoglobulin [25] is increasingly negative below 1 M $MgCl_2$, at which point it reverses the trend and becomes positive above 2.5 M. The corresponding bell-shaped dependence of $\Delta\mu_{2,tr}$ is again increasingly positive below 2.5 M $MgCl_2$, at which point it attains a maximum and starts decreasing. These examples demonstrate that measured zero binding does not indicate an absence of interaction. In fact, Γ_{23} being the derivative of $\Delta\mu_{2,tr}$ (Eqs. (6) and (8)), the point of $\Gamma_{23} = 0$ corresponds to a maximum or

[2] As shown above, a site need not represent a definite physical locus on the protein surface, and the molecule in question need not make contact with the protein.

minimum in the ligand concentration dependence of the free energy of interaction of the ligand with the protein (maximal repulsion or attraction). Schellman cautions that single-point determinations of preferential interaction cannot express the effect of the interaction on a protein, because they cannot give the free energy change caused by the added co-solvent [3]. Hence, the requirement of establishing complete concentration dependent isotherms.

8. Site occupancy and exchange

The above discussion of preferential binding, as expressed by Eq. (11) and Eq. (12), clearly indicate that the two solvent components participate on equal footing in the interaction process [2,3]. Is there a way to determine unequivocally the participation of each component in this process? For a single independent site, the fractional occupancy of the site by component 3, θ_3, is given by [3]:

$$\theta_3 = \frac{K' X_3}{X_1 + K' X_3} = \frac{K'' m_3}{1 + K'' m_3} \tag{13}$$

As Schellman comments, while Eq. (13) is identical in form with the classical binding isotherm (Eq. (1)), it is not a true binding isotherm: K' and K'' are not simple equilibrium constants for the binding of component 3; they are exchange constants with water. The fraction of the time that the site is not occupied by component 3, it is occupied by water, i.e. $\theta_1 = (1 - \theta_3)$. If Eq. (12) is split into two terms, we obtain [3]:

$$\Gamma_{23} = \frac{K'' m_3}{1 + K'' m_3} - \frac{m_3/m_1}{1 + K'' m_3} \tag{14}$$

For n identical independent sites this leads to

$$\Gamma_{23} = n\left(\theta_3 - \frac{m_3}{m_1}\theta_1\right) \tag{15}$$

where $n\theta_3$ is number of sites occupied by co-solvent and $n\theta_1$ of those occupied by water.

In order to show the magnitudes of the various parameters involved in preferential binding, let us subject to analysis in terms of the Schellman illustrative model of identical independent sites actual measurements performed on two systems. The first is the dialysis equilibrium binding of 30% (5.81 m) glycerol to β-lactoglobulin [26]. The measured preferential binding is: $\Gamma_{23} = -14.7$ glycerol molecules per molecule of protein. Let us calculate K'' with Eq. (12). For this calculation, for the sake of illustration, let us assign to n, the total number of sites, the number of waters of hydration calculated by the n.m.r. method of Kuntz [27][3], with the naive assumption that this represents the number of whole water molecules in contact with the protein surface in pure aqueous medium. This gives $n = 300$. Application of Eq. (12) results in $K'' = 0.0092\ \mathrm{m}^{-1}$, i.e. a positive exchange constant. But $(K'' - 1/m_1) = -0.0088$ which explains the measured negative binding stoichiometry, and reflects the positive free energy change per site (Eq. (11)), $\Delta G^0_{\mathrm{ex},i} = +27$ cal mol^{-1}. The site occupancy, calculated with Eq. (13), gives $\theta_3 = 0.051$ and $n\theta_3 = 15.3$. This shows that protein–co-solvent contacts are made, even though the measured preferential binding is negative and corresponds to a large preferential hydration, $\Gamma_{21} = 140$ water molecules per molecule of protein (Eq. (9)).

The second example will be drawn from the system of Fig. 1a, lysozyme–urea. Let us take the point at 5 M urea (6.44 m). The measured preferential binding is 18.1 urea molecules per molecule of protein. Let us again assign n in the same way, $n = 290$. Combination of these values gives $K'' = 0.029\ \mathrm{m}^{-1}$, $K'' - (1/m_1) = 0.011$, $\theta_3 = 0.157$ and $n\theta_3 = 45.5$. The free energy of binding per site is negative and small, $\Delta G^0_{\mathrm{ex},i} = -36$ cal mol^{-1}. Comparison of the two examples shows that in both cases the free energy of interaction per site is very small, so that the exchange constant is small. But it is its value relative to m_1^{-1} that determines whether the interaction will be one of preferential binding or preferential exclusion [2,3,7]. In both cases, a significant number of co-solvent molecules occupy sites on the protein. The results of these calculations should not be taken as representing the true molecular situation. Mindful of the Schellman caveat [7], they can serve,

[3] Any other assigned values of n would give different numerical results without changing the conclusions.

nevertheless, as illustrations of the kind of small numbers that, through their accumulation, generate the complex interaction patterns that we observe.

9. General site occupancy

The Schellman expressions (Eqs. (12) and (13)) describe the exchange between water and co-solvent at a site. Experimental measurements of preferential interactions have shown that, with certain co-solvents, component 3 is essentially, if not totally, excluded from protein. A striking example is the interaction of aqueous glycine betaine with bovine serum albumin [28]. In such a case, $\theta_3 \sim 0$ and Eq. (13) becomes indeterminate. Eq. (13) can be generalized in terms of solely occupancies by co-solvent and water without reference to exchange [29–32]. Then, if B_1 and B_3 are the effective numbers of molecules of water and co-solvent that occupy sites on the protein.

$$\Gamma_{23} = B_3 - \frac{m_3}{m_1} B_1 \tag{16}$$

The parameters B_1 and B_3 do not have a real physical meaning [29] in the sense of being actual whole molecules of components 1 and 3 in contact with protein, but they are only a summation of the perturbations of the chemical potentials of solvent component molecules by the protein whether they make contact with it or not [7]. Conceptually, B_1 and B_3 are the numbers of water and co-solvent molecules that are contained within a putative hydration layer with the composition of the bulk solvent to which have been added (or from which subtracted) the numbers of solvent component molecules involved in preferential interaction, Γ_{23} [33]. They are not thermodynamic quantities [4]. Furthermore, since the only measurable quantity is Γ_{23}, B_3 and B_1 pairs can assume an infinite number of values which can be fractional as long as they satisfy Eq. (16) [29].

Eq. (16) can be expanded to encompass both exchangeable sites, expressed by Eq. (15) and non-exchangeable sites, i.e. sites at which $K_{ex} \sim 0$ [31–34]. All occupancies by component 3 are considered to occur at exchangeable sites. Then [35,36]:

$$\Gamma_{23} = \sum_{i=1}^{n^{ex}} \left(\theta_3^i - \frac{m_3}{m_1} \theta_1^i \right) - \frac{m_3}{m_1} B_1^{Nex} = B_3 - \frac{m_3}{m_1} (B_1^{ex} + B_1^{Nex}) \tag{17}$$

10. Binding isotherms, exchange and site occupancy

To analyze the interactions of proteins with the denaturants, urea and guanidine hydrochloride, Makhatadze and Privalov [37] in an important study carried out careful calorimetric titration measurements. In their analysis, they assumed that the measured heat was proportional to the number of denaturant molecules bound to the protein. Scatchard plots of their data for unfolded proteins resulted in straight lines from which they deduced values of an equilibrium constant, K_{M-P}, and of a number of identical independent binding sites, n. From this and denaturation data, they calculated values of n for native proteins. Their claim that this approach afforded a better way of measuring thermodynamic binding than dialysis equilibrium, coupled with high precision densimetry, was in error. As Schellman and Gassner [38] have pointed out, the M and P treatment was in terms of classical binding theory (Eq. (1)) and, therefore, neglected exchange with water mandated by Eq. (3). By Eq. (13), we see that their plot [37] which neglected water activity [38] gave apparent values of K_{ex} in molal activity units, and n, the apparent number of exchangeable sites. The binding stoichiometry which they reported at any concentration was actually apparent site occupancy, $n\theta_3$, of exchangeable sites, and not the thermodynamic binding, Γ_{23} [38].

In Fig. 1a we have compared the results of this analysis of the calorimetric measurements with the thermodynamic binding, Γ_{23}, measured by dialysis equilibrium. The straight line defined by the squares represent the Scatchard plot of the Makhatadze and Privalov [37] data for the interaction of urea with native lysozyme, using their assigned parameters, $n = 119$ and $K = 0.06$ (molal activity)$^{-1}$. The expected values of Γ_{23} were calculated with Eq. (12) using the same parameters. The results of the calculation are presented in

the form of a Scatchard plot in Fig. 1a (Γ_{23}calc). Comparison with the experimental values (Γ_{23}exp) shows a total lack of similarity between the two plots.

A more extended analysis was carried out on the RNase A–urea system at pH 7.7 in which the protein remains native below 5.5 M urea. A striking feature of the dialysis equilibrium results is that the experimental values of Γ_{23} are negative below 3.4 M urea and becomes positive above that co-solvent concentration (Prakash et al., in preparation). From their calorimetric measurements, Makhatadze and Privalov [37] have deduced the parameters $n=122$ total identical independent binding sites and $K_{ex}=0.06$ (molal activity)$^{-1}$. Using these parameters, their site occupancy values ($n\theta_3$) were calculated and from this the expected thermodynamic binding at exchangeable sites, Γ_{23}exch. Fig. 3 again shows total lack of agreement with the measured values, Γ_{23}exp. One possible explanation could be the presence of non-exchangeable sites on the protein molecules which would not be detected by site contact measuring techniques, but would make a negative contribution to Γ_{23} [23,34,35]. These are given by $\Gamma_{23}\text{Nex}=\Gamma_{23}\text{exp}-\Gamma_{23}\text{exch}$. The results, plotted in Fig. 3, show that this parameter attains significant negative values of the same order of magnitude as Γ_{23}exch, which suggest that the small experimental values of the thermodynamic binding are the result of cancellation between Γ_{23}exch and Γ_{23}Nex. Combination of the measured Γ_{23} with the $n\theta_3$ values calculated with the parameters of $M+P$ gives, by Eq. (16), the expected values of B_1. The results of this calculation were that the number of contributing water molecules decreases from 545 in 1 M urea to 200 in 5 M urea. Such a variation in interacting water molecules appears most unrealistic. What are the reasons for these discrepancies?

11. Heterogeneity, non-ideality, cooperativity

A critical analysis of the current status of preferential interaction theory and practice has been given by Schellman in two important papers [38,39]. The model used in the development of the understanding of weak interactions of solvent components with proteins is far from the real situation. Schellman [7] has cautioned that it is only a framework for thinking about the solvation process. As such, it has been highly successful due to its simplicity and lucidity. Its exact application to real systems meets with difficulties, as demonstrated by the examples of Fig. 1 and Fig. 3. Schellman states that 'there are assumptions in the model which are not in accord with reality. These are the assumptions of denumerable sites, which are independent, and on which a one-for-one exchange of solvent species can take place' [39]. First, the one-to-one replacement of Eq. (11) and Eq. (12) contradicts the fact that most co-solvent molecules are larger than water [3]. Second, simple multiplication by n, the total number of sites, assumes that the protein surface is homogeneous and that the sites are independent. In reality, the protein surface is highly heterogeneous. It can be described as a mosaic of chemically diverse loci, polar, non-polar aromatic and aliphatic, positively and negatively charged, exposed peptide groups, each of which has its affinity pattern for any co-solvent. At any solvent composition, co-solvent molecules will be facing loci to which they are attracted, from which they are repelled, or toward which they are indifferent ($\Delta\mu_{2,tr}$ negative, positive or zero). Such a pattern of interactions of urea and some osmolytes with different amino acid residues and peptide groups has been well documented [40–42]. A measurement of Γ_{23} gives an overall interaction summed over all the sites and its analysis in terms of n tractable sites is naive [38]. A similar insufficiency arises from the assignment of independence to sites. As pointed out by Schellman [2,38,39], it is much more likely that there is a continuum of sites, that nearest ligands interact with each other, that water molecules form clusters, and the arrangement of solvent components around a given protein molecule fluctuates with time and varies from one protein molecule to another at any instant in time [39]. Furthermore, assignment of the total number of sites, n, requires assumptions, for example the identification of n as the effective number of waters of hydration measured by an auxiliary technique or an analysis of site occupancy detected by, say, calorimetric titration. Schellman [39] states

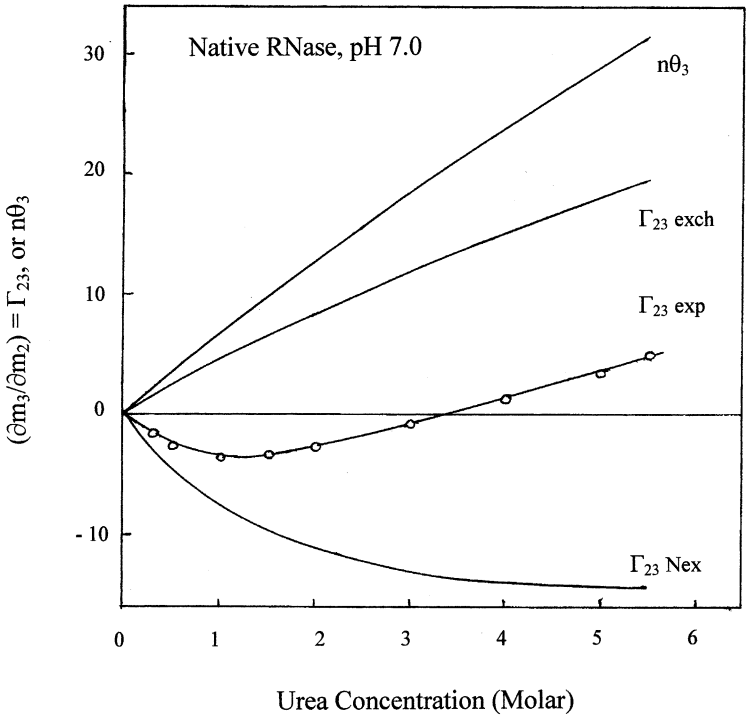

Fig. 3. Interactions of urea with ribonuclease A at pH 7.0. Open circles (○): experimentally measured preferential binding, Γ_{23}. $n\theta_3$: site occupancy calculated from the calorimetric titration data of Makhatadze and Privalov [37]. Γ_{23}exch: expected preferential binding calculated from the calorimetric data with Eq. (12). Γ_{23}Nex: difference between the experimental data and Γ_{23}exch, which is assigned to non-exchangeable sites. For details, see text.

that 'the occupancy number cannot normally be measured experimentally'.

With all of the above cautions in mind, an obvious question is: why did the calorimetric titration measurements give excellent Scatchard plots defined by single values of n and K? Since the assumptions were the same as those made by Schellman in his development of the exchange equilibrium it would seem at first that this vindicates the practical use of the simple model. The answers were given by Schellman [38,39] who analyzed the relation between enthalpy and binding and the effect of heterogeneity on the shape of the binding isotherm.

Analysis of the enthalpy of binding gave the result that the 'the relative enthalpy of interaction of a macromolecule with a co-solvent depends directly on occupancy, rather than on net preference' [39]. Consistently with this, Makhatadze and Privalov [37] have reported that the directly measured enthalpy changes are much larger than those obtained from binding measurements: calorimetric enthalpy is related to site occupancy, $n\theta_3$; equilibrium binding experiments measure preferential binding, Γ_{23}, which, by definition, is much smaller than occupancy (Eq. (15)). Development of the site occupancy relation for a heterogeneous system led Schellman [39] to the result that the 'binding isotherm' reduced to the form of Eq. (13), i.e. its form is identical with that of a single site or of a number of identical sites. The obtained parameters n and K, however, are only apparent values, n_{app} being smaller than the actual number of sites, n, and K_{app} being greater than the average values of the binding constants [39]. Under such circumstances, 'what is obtained are the parameters for an identical-site system which behaves like the system under consideration' [38]. It is 'a represen-

tation of the data and not a model for the system' [39]. Schellman carried out simulation calculations of occupancy for a heterogeneous system [39] in which the actual binding constants varied by a factor of 10. He compared the exact isotherm with one in which the constant is the proper average of the actual constants. The two calculations gave essentially identical isotherms. This demonstrated that conformity of an isotherm to a homogeneous multisite system may, in fact, conceal a high degree of heterogeneity. The calculated binding parameters, then, are only apparent values.

As reminded above, a protein presents a highly chemically heterogeneous surface to a solvent. Therefore, the measurement of identical sites may be surprising. The Schellman analysis may well give an explanation for the experimentally measured homogeneous isotherms [37]. The values of the reported parameters, n and K, are only descriptions of the data, but do not represent the real situation. Nevertheless, such procedures are useful in that they permit at least qualitative thinking about the process. The conclusion that n and K are not real physical entities explains also the total divergence between the preferential interaction calculated with these parameters, Γ_{23}exch of Fig. 3, and the experimentally measured values. Schellman [38] showed that, contrary to site occupancy, the use of average parameter leads to a preferential binding isotherm that diverges from the one calculated for the heterogeneous system. The expected difference, however, is much smaller than that observed (Fig. 3). The causes for this must be sought in some of the other factors referred to above, such as cooperativity between sites or clustering of ligand molecules.

12. Conclusions

While John Schellman's interest in protein denaturation by solvents goes back almost 50 years, with the publication of his early fundamental treatment [43], he turned his interest to the concept of preferential interactions some 20 years later [21,22]. In 1978, Schellman [21] started his analysis by establishing a definition of binding. First, he defined binding by an intuitive process which consists in the counting of the number of solvent molecules (water and co-solvent) in a general hypothetical internal (local) domain within an 'accessibility contour' where a macromolecule can influence solvent component molecules and comparing the resulting solvent composition with that of the bulk solvent which is not influenced by the macromolecule. The excess of molecules of component 3 (ligand) within the internal domain was defined as the binding. This corresponds to the experimental approach. He then used the fluctuation theory [5,6] to calculate the excess of component 3 in the neighborhood of a macromolecule without recourse to any model, which led directly to the thermodynamic definition of binding, expressed by Eq. (7), and permitted to encompass into the concept of binding ligand molecules which interact with protein molecules at a distance through long range effects [21]. This thermodynamic definition of binding provided a resolution of the puzzle that denaturation by urea or guanidine hydrochloride is typically characterized by transition curves that correspond to a change in binding of the denaturant of $\sim 10-15$ additional molecules 'bound' during the course of the unfolding. From the accompanying free energy change of ~ -10 kcal mol^{-1}, one would have expected the transition to occur at approximately 1 M denaturant. Yet, it requires 6–8 M denaturant for unfolding to occur. Schellman's definition that the measured change is in preferential binding, $\Delta\Gamma_{23}$, and not in occupancy [21] solved this dilemma, because the actual number of denaturant molecules involved (change in occupancy) may be more than one order of magnitude greater than $\Delta\Gamma_{23}$.

In 1987, Schellman [2] introduced specifically the concept of exchange into the treatment of binding. As described in the preceding text, this led to the explanation of how positive binding constants can generate negative binding stoichiometries and to the demonstration of the origin of the complex shapes of binding isotherms as residing in the variation of solvent component activity coefficients. As Schellman has stated on several occasions: these are only beginnings and the development of more advanced models and more detailed theory must await the accumulation of new knowledge. Such progress is being made. One may cite the work of Bolen et al. [42,44] who are

generating data on the transfer free energies of amino acid residues into various osmolyte systems as part of their studies of osmophobicity, i.e. the basis on which nature selects agents to stabilize cellular integrity and function. On the other hand, Record et al. have extended the analysis of the preferential binding of electrolytes [33] and are probing more complex models of interaction and ligand clustering [28,32,33,45]. Developments such as these are welcome, because they should permit to extend the edifice for which Schellman has laid much of the foundation.

References

[1] J.C. Lee, S.N. Timasheff, The stabilization of proteins by sucrose, J. Biol. Chem. 256 (1981) 7193–7201.

[2] J.A. Schellman, Selective binding and solvent denaturation, Biopolymers 26 (1987) 549–559.

[3] J.A. Schellman, A simple model for solvation in mixed solvents. Applications to the stabilization and destabilization of macromolecular structures, Biophys. Chem. 37 (1990) 121–140.

[4] J.A. Schellman, The relation between the free energy of interaction and binding, Biophys. Chem. 45 (1993) 273–279.

[5] J.G. Kirkwood, R.J. Goldberg, Light scattering arising from composition fluctuations in multi-component systems, J. Chem. Phys. 18 (1950) 54–57.

[6] W.H. Stockmayer, Light scattering in multi-component systems, J. Chem. Phys. 18 (1950) 58–61.

[7] J.A. Schellman, The thermodynamic stability of proteins, Ann. Rev. Biophys. Biophys. Chem. 16 (1987) 115–137.

[8] G. Scatchard, Physical chemistry of protein solutions. I. Derivation of the equations for the osmotic pressure, J. Am. Chem. Soc. 68 (1946) 2315–2319.

[9] E.F. Casassa, H. Eisenberg, Thermodynamic analysis of multicomponent solutions, Adv. Protein Chem. 19 (1964) 287–395.

[10] S.N. Timasheff, Control of protein stability and reactions by weakly interacting cosolvents: the simplicity of the complicated, Adv. Protein Chem. 51 (1998) 355–432.

[11] E. Zerovnic, S. Lapanje, Interactions of myoglobin water urea and some alkylureas. I. Solvation in urea and alkylurea solutions, Biophys. Chem. 24 (1986) 53–59.

[12] N. Poklar, S. Lapanje, Solvation of β-lactoglobulin in alkylurea solutions, Biophys. Chem. 42 (1992) 283–290.

[13] E. Hade, C. Tanford, Isopiestic compositions as a measure of preferential interactions of macromolecules in two-component solvents. Application to proteins in concentrated aqueous cesium chloride and guanidine hydrochloride, J. Am. Chem. Soc. 89 (1967) 5034–5040.

[14] J. Lee, S.N. Timasheff, Partial specific volumes and interactions with solvent components of proteins in guanidine hydrochloride, Biochemistry 13 (1974) 257–265.

[15] C. Tanford, Physical Chemistry of Macromolecules, Wiley, New York, 1961.

[16] G. Scatchard, J.S. Coleman, A.L. Shen, Physical chemistry of protein solutions. VII. The binding of some small anions to serum albumin, J. Am. Chem. Soc. 79 (1957) 12–20.

[17] E.P. Pittz, S.N. Timasheff, Interaction of ribonuclease A with aqueous 2-methyl-2,4-pentanediol at pH 5.8, Biochemistry 17 (1978) 615–623.

[18] T. Arakawa, S.N. Timasheff, Mechanism of poly (ethylene glycol) interaction with proteins, Biochemistry 24 (1985) 6756–6762.

[19] P.H. von Hippel, V. Peticolas, L. Schack, L. Karlson, Model studies on the effects of neutral salts on the conformational stability of biological macromolecules. I. Ion binding to polyacrylamide and polystyrene columns, Biochemistry 12 (1973) 1256–1264.

[20] M.J. Kronman, S.N. Timasheff, Light scattering investigation of ordering effects in silicotungstic acid solutions, J. Phys. Chem. 63 (1959) 629–633.

[21] J.A. Schellman, Solvent denaturation, Biopolymers 17 (1978) 1305–1322.

[22] J.A. Schellman, Macromolecular binding, Biopolymers 14 (1975) 999–1018.

[23] E. Reisler, H. Eisenberg, Interpretation of equilibrium sedimentation measurements of proteins in guanidine hydrochloride solutions. Partial volumes, density increments, and the molecular weight of the subunits of rabbit muscle aldolase, Biochemistry 8 (1969) 4572–4578.

[24] G. Xie, S.N. Timasheff, Mechanism of the stabilization of ribonuclease A by sorbitol: preferential hydration is greater for the denatured than the native protein, Protein Sci. 6 (1997) 211–221.

[25] T. Arakawa, R. Bhat, S.N. Timasheff, Why preferential hydration does not always stabilize the native structure of globular proteins, Biochemistry 29 (1990) 1924–1931.

[26] K. Gekko, S.N. Timasheff, Mechanism of protein stabilization by glycerol: preferential hydration in glycerol–water mixtures, Biochemistry 20 (1981) 4667–4676.

[27] I.D. Kuntz, Hydration of macromolecules. III. Hydration of polypeptides, J. Am. Chem. Soc. 93 (1971) 514–516.

[28] E.S. Courtenay, M.W. Capp, C.F. Anderson, M.T. Record, Vapor pressure osmometry studies of osmolyte-protein interactions: implications for the action of osmoprotectants in vivo and for the interpretation of 'osmotic stress' experiments in vitro, Biochemistry 39 (2000) 4455–4471.

[29] C. Tanford, Extension of the theory of linked functions to incorporate the effects of protein hydration, J. Mol. Biol. 39 (1969) 539–544.
[30] H. Inoue, S.N. Timasheff, Preferential and absolute interactions of solvent components with proteins in mixed solvent systems, Biopolymers 11 (1972) 737–743.
[31] E. Reisler, Y. Haik, H. Eisenberg, Bovine serum albumin in aqueous guanidine hydrochloride solutions. Preferential and absolute interactions and comparison with other systems, Biochemistry 16 (1977) 197–203.
[32] M.T. Record, C.F. Anderson, Anderson, Interpretations of preferential interaction coefficients of nonelectrolytes and of electrolyte ions in terms of a two-domain model, Biophys. J. 68 (1995) 786–794.
[33] M.T. Record, W. Zhang, C.F. Anderson, Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: a practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects to salts, Adv. Protein Chem. 51 (1998) 281–353.
[34] D.W. Kupke, Density and volume change measurements, in: S.J. Leach (Ed.), Physical Principles and Techniques of Protein Chemistry, Part C, Academic Press, NY, 1973, pp. 1–75.
[35] S.N. Timasheff, Water as ligand: preferential binding and exclusion of denaturants in protein unfolding, Biochemistry 31 (1992) 9857–9864.
[36] S.N. Timasheff, The control of protein stability and association by weak interactions with water: how do solvents affect these processes?, Ann. Rev. Biophys. Biomol. Struct. 22 (1993) 67–97.
[37] G.L. Makhatadze, P. Privalov, Protein interactions with urea and guanidinium chloride. A calorimetric study, J. Mol. Biol. 226 (1992) 491–505.
[38] J.A. Schellman, N.C. Gassner, The enthalpy of transfer of unfolded proteins into solutions of urea and guanidinium chloride, Biophys. Chem. 59 (1996) 259–275.
[39] J.A. Schellman, The thermodynamics of solvent exchange, Biopolymers 34 (1994) 1015–1026.
[40] C. Tanford, Protein denaturation. Part C. Theoretical models for the mechanism of denaturation, Adv. Protein Chem. 24 (1970) 1–95.
[41] C.N. Pace, The stability of globular proteins, Crit. Rev. Biochem. 3 (1975) 1–43.
[42] Y. Liu, D.W. Bolen, The peptide backbone plays a dominant role in protein stabilization by naturally occurring osmolytes, Biochemistry 34 (1995) 12884–12891.
[43] J.A. Schellman, Stability of hydrogen bonded peptide structures in aqueous solution, C.R. Trav. Lab. Carlsberg Ser. Chim. 29 (1955) 230–259.
[44] D.W. Bolen, I.V. Baskakov, The osmophobic effect: natural selection of a thermodynamic force in protein folding, J. Mol. Biol. 310 (2001) 955–963.
[45] E.S. Courtenay, M.W. Capp, R.M. Saecker, M.T. Record, Thermodynamic analysis of interactions between denaturants and protein surface exposed on unfolding: interpretation of urea and guanidinium chloride *m*-values and their correlation with changes in accessible surface area (ASA) using preferential interaction coefficients and the local-bulk domain model, Proteins: Struct. Funct. Genet. 41 (S4) (2000) 72–85.