Monolayer Hydration Governs Nonideality in Osmotic Pressure of Protein Solutions

M. A. Yousef
Argonne National Laboratory, Bioscience Division, Argonne, IL 60439

R. Datta
Dept. of Chemical Engineering, Worcester Polytechnic Institute, Worcester, MA 01609

V. G. J. Rodgers
Dept. of Chemical and Biochemical Engineering, and Center for Biocatalysis and Bioprocessing, University of Iowa, Iowa City, IA 52242

Strong nonideality observed in the osmotic pressure of concentrated globular proteins in aqueous media of moderate salt concentrations has long been associated with protein–protein interaction. As a consequence, virial expansions, based on the McMillan–Mayer theory, have been used to extract the physical phenomena observed. Our recent articles showed that a free-solvent model assuming hydration and salt binding are dominate factors for nonlinearity, successfully modeled BSA (67 kDa) and IgG (155 kDa) protein aqueous solutions at moderate salt concentrations. Similar findings for lysozyme (HEL, 14 kDa) and ovalbumin (45 kDa) are reported here. More significantly, the independently calculated hydration values regressed from the osmotic pressure data were compared with the solvent accessible surface areas of each protein investigated. The results showed that the hydration values determined from the free-solvent model are remarkably a monolayer equivalent of water.

Introduction

The osmotic pressure of the globular proteins such as bovine serum albumin (BSA) and immuno gamma globulin (IgG) in aqueous solutions has been found to exhibit significant non-ideality, particularly at high mass concentrations (Yousef et al., 1998a,b). Often this behavior is attributed to protein–protein interactions and, for dilute solutions, is accounted for by a virial expansion on the concentration variable. However, the virial expansion model has not generally been able to successfully predict the physical behavior of the osmotic pressure of concentrated globular proteins. The hydration model, or free-solvent model, on the other hand, has been able to represent both the behavior and the magnitude of concentrated globular protein solutions in aqueous media under certain conditions. This report further shows that this hydration is well described by a monolayer hydration model based on the surface of the proteins used.

Correspondence concerning this article should be addressed to V. G. J. Rodgers.

Virial expansion model

The virial expansion approach has its origins in an analogy of the dilute liquid theory with imperfect gas theory. It has a statistical mechanical basis for a binary system, and treats the solvent and its constituents as a continuum, and assumes that their interactions with the protein molecules are ignored (McMillan and Mayer, 1945). The main argument of the virial expansion is that the osmotic pressure can be expressed as a function in solute concentration using a Taylor series expansion. Moreover, the virial coefficients have been related by McMillan and Mayer to the average force potential or the potential energy between the solute molecules, that is, protein–protein interactions (McMillan and Mayer, 1945; Hill, 1986). Using statistical mechanics, it has been shown that osmotic pressure can be expressed as an integral series expansion of concentration for dilute solutions (McMillan and Mayer, 1945; Hill, 1986).

There are two major assumptions for the virial expansion in addition to the fact that it was developed for a binary sys-

tem. First, while the virial expansion in activity is generally accepted to be rigorous, the expansion in many concentration variables, such as number density (molecules per volume), ρ, is valid only in the limit as $\rho \to 0$ (Hill, 1986). This assumption is needed in order to make the calculations of the virial coefficients via the potential of mean force tractable, since the model assumes that the nonidealities are the result of solute–solute interaction due to short-range and screened long-range forces acting on the solutes.

The dilute solution assumption can be invalid in another way for concentrated proteins. For dilute protein solutions, the mole fraction of protein is $O(10^{-4})$ and this would be a reasonable approximation. However, it is well recognized that hydration and microion binding is substantial and can be between $O(10^{-3})$ and $O(10^{-4})$. For concentrated protein solutions, this would increase the concentration of the water-bound species to well outside the range of a dilute assumption. This leads to the second major assumption in applying the viral expansion model to concentrated protein solutions: The model treats the solvent and its constituents as an unperturbed continuum. This assumption is not reasonable for concentrated proteins, since the water interacting in hydration is uniquely different in chemical potential from that in bulk. Moreover, this quantity of hydrated water can be within the same order of magnitude as that in the bulk.

Despite these potential restrictions, Vilker et al. (1981) did an extensive evaluation on the applicability of the virial expansion model for concentrated BSA solutions. They investigated the contributions to the intermolecular potential function for charge–charge, charge–dipole, dipole–dipole, charge fluctuation, charge-induced dipole, dipole-induced dipole, and dispersion interactions. They concluded that the dominant pair potentials were the repulsive charge–charge interactions and the attractive dispersion interactions (similar to DLVO theory) and used a prolate ellipsoid model to describe the excluded volume of BSA. The resulting predicted osmotic pressures were of the same order of magnitude as the measured osmotic pressure, but the model did not fit the data well. The authors followed with an extensive analysis of their results, including an evaluation of the significance of increasing the order of the virial expansion and analysis of the effect of varied excluded volume due to hydration within the structure of the virial expansion-based model. They could not find any conclusive factors that the might explain the deviation between the virial expansion model and the data.

Free-solvent model

Recently Yousef et al. (1998a,b) used a free-solvent model to investigate the nonidealities of the osmotic pressure of concentrated protein solutions. In moderate salt concentrations, it was found that the hydration and ion binding to the protein and not solute–solute interaction dominate the nonideal behavior for the osmotic pressure of globular protein solutions.

This can be recognized if we examine the origin of the ideal model for osmotic pressure. For a two-component osmometer, let the side of the chamber containing the proteins, solvent, and diffusible ions be denoted as compartment II (inside compartment) and the compartment containing only the solvent and diffusible ions be denoted as I (outside compartment). Then, the osmotic pressure for an ideal system can be expressed in terms of the solvent mol fraction via the van Laar equation (van Laar, 1894)

$$\pi = \frac{\mu_1^{\mathrm{I}} - \mu_1^{\mathrm{II}}}{\bar{V}_1} = \frac{RT}{\bar{V}_1} \ln \frac{x_1^{\mathrm{II}}}{x_1^{\mathrm{I}}} \tag{1}$$

where μ_1 is the chemical potential for the solvent. This model is borne from the Maxwell–Boltzmann distribution law, which also has a statistical mechanical basis. It is therefore subject to the assumptions that govern the Maxwell–Boltzmann model to represent ideality. In particular, the system comprises a large number of species that are submerged in a noninteracting medium and that have no attractive interaction between them. Then x_1 is the fraction of these molecules having a specific free energy per molecule, μ_1. The form of Eq. 1 is the result of determining the maximum probability distribution of species with a specific energy within the system (Glasstone, 1946).

This model for ideality might apply to species such as proteins in aqueous media provided the basic assumptions are not violated. This is particularly true for the requirement that the molecules remain nonattracting. However, as stated earlier, globular proteins are charged species with electrostatic interactions. However, in aqueous media where the ionic strength is sufficiently high (> 0.1 M salt), electrostatic interactions between the solute macromolecules are screened, reducing the attractive interaction between the solutes by a factor of e^{-kr}, where κ^{-1} is the Debye length and r is the distance between solutes (Hill, 1986; Tombs and Peacocke, 1974). Note that this requirement also constrains the applicability of the rigorous virial expansion model.

But even when the proteins are nonattracting, Eq. 1 cannot strictly hold for proteins in aqueous media. This is because, in principle, any solute introduced into a solvent will have interactions with neighboring solvent molecules, resulting in solvation, or hydration, when an aqueous solution. These neighboring solvent molecules undergo an alteration in their free energy to interact with the entering molecule (Israelachvili, 1992).

Although the amount of water that interacts with the protein has been extensively studied, precise values are elusive and reported values are generally a function of the method used (Rupley and Careri, 1991). However, from both water–^{17}O magnetic resonance studies (Otting, 1991) and mathematical modeling that minimizes the energy required for hydration orientation (Sedykh and Sedykh, 1967), the hydration zone was estimated to be about two layers of water. Based on the hydration-force distance, this extends to approximately 10 Å (Israelachvili and Marra, 1986; Claesson et al., 1989). This is equivalent to about 1 g-H_2O/g-protein for globular proteins (Rupley and Careri, 1991). The first layer of water is about 0.3 g-H_2O/g-protein. However, this number depends on the protein size, shape, and its charge distribution, and may be slightly higher for solutions containing salts. For a 15-kDa molecular weight protein this is about 250 molecules of water. The remaining water is associated with cooperative interaction and completion of the amount of water that surrounds the protein. Because of surface roughness, a 15-kDa

protein needs at least 600 water molecules to complete a monolayer (Richards, 1977). Nevertheless, all water involved is estimated to be $O(1)$ g-H_2O/g-protein.

This solute–solvent interaction would be sufficient to eliminate the applicability of an idealized model as described by Eq. 1 for representing proteins in solution. However, we can construct a macromolecule that is composed of the protein and all interactive solvent species that still satisfies the restrictions for idealization, provided solute–solute interactions are eliminated. In this case, to maintain consistency with the Maxwell–Boltzmann distribution law, the solvated molecule, accompanying solvent molecules, and interacting coions and counterions are grouped and treated as a single macromolecule that is in solution with the remaining solvent molecules that have the original free energy of the solvent. This concept, known as the free-solvent model, captures the key physical phenomena contributing to the observed nonlinearity in measured osmotic pressure with respect to solute mass concentration. The result is a system of these hydrated macromolecules in a solvent with a reduced number of solvent molecules. This modeling concept is not new and was used as early as 1916 by Frazer and Myrick (1916) to analyze the nonidealities in aqueous sucrose solutions. Figure 1 shows this model.

The free-solvent mol fraction can be determined by first assuming that the solution is made up of n distinct species and species 1 is the solvent, species 2 through p are the proteins, and species $p+1$ through n are the remaining diffusible solvent components. Then the initial total mols of the solution in compartment II is $N^{\mathrm{II}} = \sum_{i=1}^{n} N_i^{\mathrm{II}}$, where i is representative of each of the n species. The final total mols of free solvent (diffusible salts and water) in chamber II is $N_*^{\mathrm{II}} = N^{\mathrm{II}} - \sum_{\substack{i=1 \\ i\neq 2,\, p}}^{n} \sum_{j=2}^{p} \nu_{ij} N_j^{\mathrm{II}} - \sum_{j=2}^{p} N_j^{\mathrm{II}}$, where N_j^{II} denotes the mols of protein j in solution and ν_{ij} is the net number of mols of solution component i that is interacting with protein j to make up to new solvent–interacting protein. Defining our osmometer compartments as described earlier, it can be shown that the mol fraction of the free solvent, component 1, in compartment II (protein present) is then

$$x_1^{\mathrm{II}} = \frac{N_1^{\mathrm{II}} - \sum_{j=2}^{p} \nu_{1j} N_j^{\mathrm{II}}}{N_*^{\mathrm{II}} + \sum_{j=2}^{p} N_j^{\mathrm{II}}} \tag{2}$$

and

$$x_1^{\mathrm{I}} = \frac{N_1^{\mathrm{I}}}{N^{\mathrm{I}}} \tag{3}$$

is the mol fraction of the solvent in the chamber.

Assuming no solute–solute interaction, this free-solvent mol fraction has been used in Eq. 1 to successfully relate osmotic pressure to BSA and IgG solute concentrations (Yousef et al., 1998a,b). Equations 2 and 3 allow for correction of additional microions in the solvent chamber due to satisfaction of the electroneutrality condition; however, if appreciable change in net mols occurs, this must also be considered. For the case in which a solution contains a single monovalent salt and a single protein, the resulting model for osmotic pressure becomes

$$\pi \approx \frac{RT}{\bar{V}_1} \ln\left\{ \frac{\left(N_1^{\mathrm{II}} + (1 - \nu_{12} - \nu_{32}) N_2^{\mathrm{II}} + N_3^{\mathrm{II}}\right) N_1^{\mathrm{I}}}{\left(N_1^{\mathrm{II}} - \nu_{12} N_2^{\mathrm{II}}\right) N^{\mathrm{I}}} \right\} \tag{4}$$

It should be recognized that both microion interaction and hydration can be determined independent of osmotic pressure data. Thus, in principle, no fitted parameters are required. However, the sensitivity of the osmotic pressure to hydration is

$$\frac{\partial \pi}{\partial \nu_{12}} = \frac{RT}{\bar{V}} \left(\frac{x_2^{\mathrm{II}}}{x_1^{\mathrm{II}}} - (1+\alpha) x_2^{\mathrm{II}} \right) = \frac{RT}{\bar{V}} \frac{\left(x_2^{\mathrm{II}}\right)^2}{x_1^{\mathrm{II}}} \approx \frac{RT}{\bar{V}} \left(x_2^{\mathrm{II}}\right)^2 \tag{5}$$

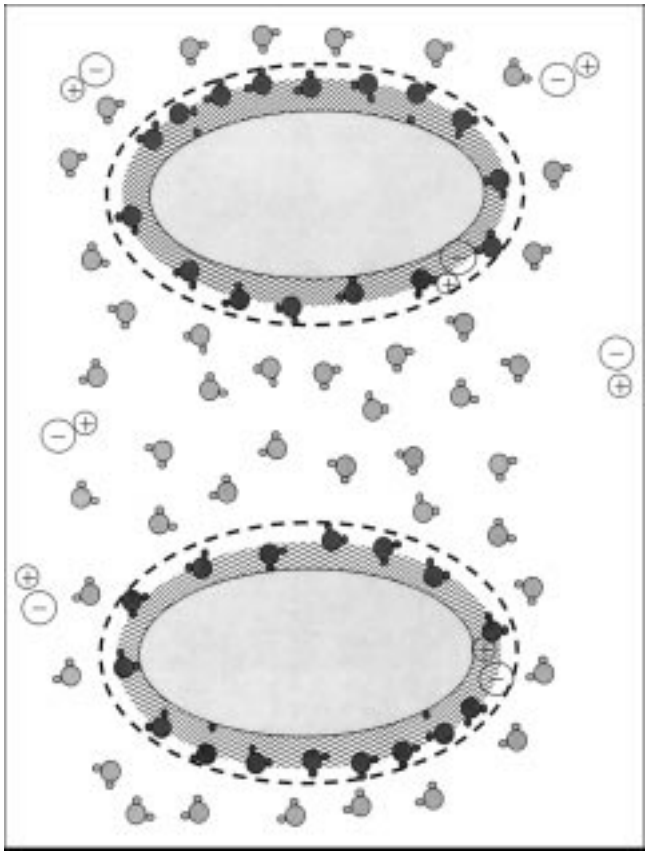

Figure 1. Hydrated proteins in solution.

Dash lines, 2-D representation of hydration region; species outside of dash lines, part of solvent; species inside of dash lines, protein and ions and water influencing the protein; wavy area, Debye length from charged protein. The free-solvent model treats all of the materials within dashed line as a single macromolecule. This model is assumed to be valid for identifying nonidealities when the range of electrostatic interactions (as determined by the Debye length) do not extend significantly into the region of other neighboring proteins.

where α is the ratio of ν_{32}/ν_{12}. For dilute solutions, $x_1^{II} \rightarrow 1/(1+\alpha)$ and the osmotic pressure is insensitive to the value of ν_{12}. However, at high concentrations of protein (hydrated), the osmotic pressure is $O(RT/\bar{V})$, which can be as high as 2×10^4 psi/(g-H_2O/g-protein) (Yousef et al., 1998b). This implies that the hydration number can be sensitive to ±0.0001 g-H_2O/g-protein or just a few molecules of water in predicting the osmotic pressure within ±1 psi at high concentrations. Therefore, because of the lack of precise hydration measurements and the sensitive dependency of the model on hydration, ν_{12} is regressed to minimize the sum of the error squared between the predicted and measured osmotic pressure.

The objective of this study is to find a general relationship between the regressed hydration values obtained from the free-solvent model. In addition to BSA and IgG, the osmotic pressure of two additional globular proteins, ovalbumin (OV) and lysozyme (HEL), in aqueous media are determined. The osmotic pressures of ovalbumin are measured in solutions of 0.15 M NaCl. The osmotic pressures for hen egg lysozyme (HEL) are measured in solutions of 0.15 M NaCl and 0.15 M KCl up to the highest concentration possible. All solutions were at 7.0 pH.

Hydration and Ion Binding for Ovalbumin and Lysozyme

The free-solvent model, although providing a physical representation of general protein–solvent interaction, hinges on a comparison of an independently determined hydration number and/or ion binding for verification. Alternately, the hydration and/or ion binding may be regressed on in fitting the model to the osmotic pressure data. In this case, it is paramount that the resulting values are consistent with physical phenomena. As mentioned earlier, ^{17}O–NMR analysis considers both the first hydration layer and cooperative interaction of adjacent water. Using this method, the hydration for globular proteins is on the order of 1 g-water/g-protein. For ovalbumin, Sedykh and Sedykh (1967) used a whole-number linear-programming model to minimize/maximize the energy required to orient the water of hydration by the protein hydrophilic groups. They were able to obtain the minimum and maximum number of water molecules required to hydrate the one protein molecule. For ovalbumin, they found that the more ordered (firmly bound) water in the first hydration layer ranged from 0.066 to 0.228 g-H_2O/g-ovalbumin and the less ordered water (loosely bound) ranged between 0.226 and 0.716 g-H_2O/g-ovalbumin (Sedykh and Sedykh, 1967). This gives an overall range between 0.292 and 1.044 g-H_2O/g-ovalbumin. Bull and Breese (1968, 1970, 1976) studied the hydration of ovalbumin and the binding of NaCl and other chloride salts to ovalbumin. In the absence of salt and at a relative humidity of 0.92 and 25°C, they reported a hydration value of 0.295-g H_2O/g-ovalbumin (1968).

Yang and Rupley (1979) used a drop calorimeter to determine that the hydration of lysozyme was approximately 0.38, which corresponds to no more than a monolayer of water about the protein (Rupley and Careri, 1991). However, it is expected that hydration determined by osmotic pressure would result in a larger cooperative interaction contribution (Rupley and Careri, 1991).

The chloride ion binding to lysozyme has not been measured directly, to our knowledge. However, two estimates have been reported. Curtis et al. (1998) estimated the chloride ion binding to lysozyme to be approximately 2.0 mol Cl^-/mol-lysozyme in 0.1 M KCl. This was determined indirectly using LALLS and regression of a potential of mean-force model (DLVO theory with uniformly charged spherical proteins and water modeled as a dielectric continuum) with an osmotic second virial coefficient. Later, Kuehner et al. (1999) estimated the chloride ion binding to be around 4.0 mol Cl^-/mol-lysozyme. They used the hydrogen-ion titration curve for lysozyme in 0.1 M KCl and a general molecular–thermodynamic model for ion binding developed by Fraije and Lyklema (1991) and Fraije et al. (1991a,b).

Experimental Method

The osmotic pressure was measured in the osmometer described elsewhere (Yousef et al., 1998a). The solution chamber is charged first with the protein solution and the valves are closed. The solvent chamber is then filled with the buffer or saline solution and left open to the atmosphere. The pressure reading stabilizes in 5–6 h.

For the ovalbumin solutions, the saline solution was prepared by dissolving the proper amount of NaCl (Sigma Cat. No. S-7653) in one liter of deionized water to produce 0.15 M or 0.5 M solution. A weighted amount of ovalbumin powder (Sigma Cat. No. A-5503), was dissolved in a known amount of saline solution. The solution was then stirred and stored in the refrigerator at 4°C for one or two days until it completely dissolved.

For lysozyme, the saline solution was prepared by dissolving the proper amount of NaCl (Sigma Cat. No. S-7653) or KCl (Fisher Scientific, No. P217-500) in one liter of deionized water to produce 0.15 M saline solution. A weighted amount of lysozyme powder (Sigma, Cat. No. A-5503) was dissolved in a known amount of solution.

The pH of the protein solutions and the solvents were adjusted before each run to 7.0 by adding aliquots of 0.1 N NaOH or 0.1 N HCl. The protein solutions were vigorously mixed using a vortex mixer (Vortex Genie, model no. G-560, Fisher Scientific, St. Louis, MO) during the pH adjustment to prevent local denaturation of the protein. The pH measurements were conducted using a gel-filled combination electrode with epoxy body (Model E6-B BIO/CAL, Lab Research Products, Inc., Lincoln, NE) and pH/ion meter (Accument 950, Fisher Scientific). The electrode was cleaned and calibrated before each measurement. After each run, the pH of both the protein solution and the solvent were checked and found to be within the error of the pH electrode (±0.2 pH units).

The amount of acid or base used to adjust the pH, between 10 and 100 μL, was considered part of the solvent, and was taken into account when the concentration of the protein solution was determined.

The concentration of the protein solution, in gram protein per kg solvent, was determined directly from the known amounts of dry protein and solvent used to make the solution. On the other hand, the concentration of the solution in grams protein per liter of solution, w_2, was calculated by di-

viding the weight of the dry protein, wt_p, by the volume of the solution in liters. The volume of the solution, in milliliters, was calculated using the specific volume of dry protein, $\bar{\nu}_p =$ 0.75 cm^3/g for ovalbumin, $\bar{\nu}_p = 0.688$ cm^3/g for lysozyme, and that of the solvent, $\bar{\nu}_s = 1.00$ cm^3/g, that is, $V = \bar{\nu}_p wt_p + \bar{\nu}_s wt_s$. Here, wt_p and wt_s are the weight, in grams, of the dry protein and the solvent, respectively (von Hippel, 1989). The concentration of different samples of protein solutions were checked using a spectrometer (DU-64 spectrophotometer, Beckman, Fullerton, CA); the extinction coefficient of ovalbumin used is 7.01 at 280 nm (Pace et al., 1995; Gill and von Hippel, 1989). The difference between the calculated and the measured concentration was no greater than 5.0%.

After testing several solutions, a maximum difference of 1.0% was found between the concentration of protein solution before the run and that after the run. There was no protein detected in the solvent side.

Application of Free-Solvent Model to Globular Proteins

The measured osmotic pressures of ovalbumin and lysozyme and the other referenced proteins are shown in Figure 2. The maximum HEL concentration that could be attained prior to precipitation was approximately 130 g/L in NaCl and 210 g/L in KCl. The osmotic pressures of HEL in both solutions are nearly identical for the same concentration. The results of using the free-solvent model to correct the mol fraction of water in the bulk are shown in Figure 3. The hydration and salt-binding values used to generate Figure 3 are summarized in Table 1. The hydration values for

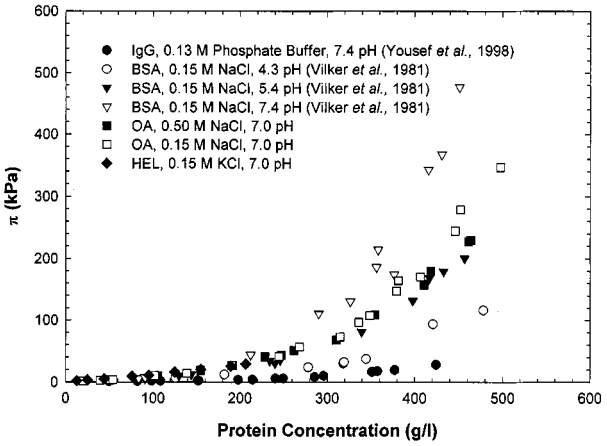

Figure 2. Measured osmotic pressure vs. protein concentration.

HEL, lysozyme; OA, ovalbumin; BSA, bovine serum albumin; IgG, immuno gamma globulin. The IgG solution includes a phosphate buffer.

BSA (66 kDa) and IgG (155 kDa) were determined by us earlier (Yousef et al., 1998a,b).

The hydration of ovalbumin was characterized by the single parameter ν_{12}, and it was assumed that the ions bound to the protein and the counterions in the hydration zone were represented by the single parameter, ν_{32}. Both parameters were best fit to Eq. 4 using nonlinear regression (TableCurve 2D, Jandel Scientific, San Rafael, CA). The resulting values

Table 1. Hydration and Salt Binding Values

Protein (MW kDa)	SASA (Å^2) (PDB ID)	Solution Properties Salt Conc. (pH)	κ^{-1} (Å)	ν_{12} Regressed Hydration $\left(\frac{\text{Mol } H_2O}{\text{Mol protein}}\right)$	ν_{12} Regressed Hydration $\left(\frac{g H_2O}{g \text{ protein}}\right)$	ν_{32} $\left(\frac{\text{Mol salt}}{\text{Mol protein}}\right)$	References
Lysozyme (14.5)	6463 (4LYZ)	0.15 M KCl 7.0	7.9	583 ± 6.4	0.724 ± 0.008	2*	* Curtis et al. (1998)
		0.15 M KCl	7.9	591 ± 2.4	0.734 ± 0.003	2*	* Curtis et al. (1998)
		0.15 M KCl 7.0	7.9	1,300 ± 8.0	1.614 ± 0.010	4**	** Kuehner et al. (1999)
		0.15 M KCl	7.9	1,284 ± 5.6	1.595 ± 0.007	4**	** Kuehner et al. (1999)
Ovalbumin (45)	15702 (1OVA)	0.15 M NaCl, 7.0	7.9	2,150 ± 100	0.86 ± 0.04	4.08 ± 0.43	
		0.15 M NaCl, 7.0	4.4	2,225 ± 100	0.89 ± 0.04	18.87 ± 1.03	
BSA (66)	27495 (1AO6)	0.15 M NaCl, 4.5	8.3†	4,081 ± 23	1.113 ± 0.0063	11.59	Yousef et al. (1998b)
		0.15 M NaCl, 5.4	8.2†	4,169 ± 22	1.137 ± 0.0059	10.62	Yousef et al. (1998b)
		0.15 M NaCl, 7.4	8.2†	4,316 ± 18	1.177 ± 0.0050	8.81	Yousef et al. (1998b)
IgG (155)	63478 (1IGT)	0.13 M Ph. Buff. 7.4	8.2	9,558 ± 353	1.11 ± 0.041	24.3 ± 0.97	Yousef et al. (1998a)

†The Debye lengths are based on the ionic strength corrected for the presence of protein.

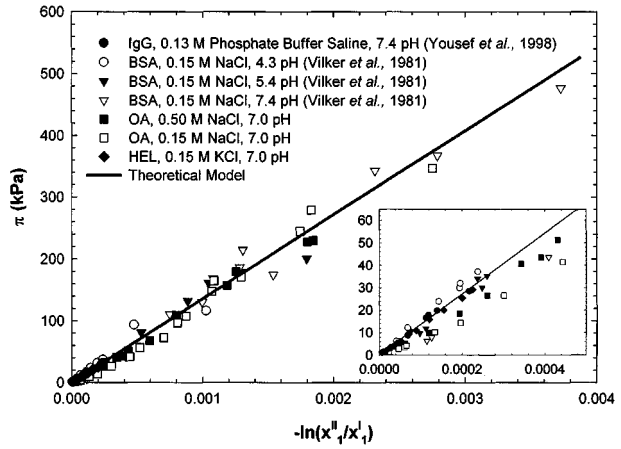

Figure 3. Measured osmotic pressure vs. natural log of the bulk water mol fraction ratio after correcting for hydration and ion binding.

HEL, lysozyme; OA, ovalbumin; BSA, bovine serum albumin; IgG, immuno gamma globulin. The IgG solution includes a phosphate buffer. Solid line, the solution to Eq. 1.

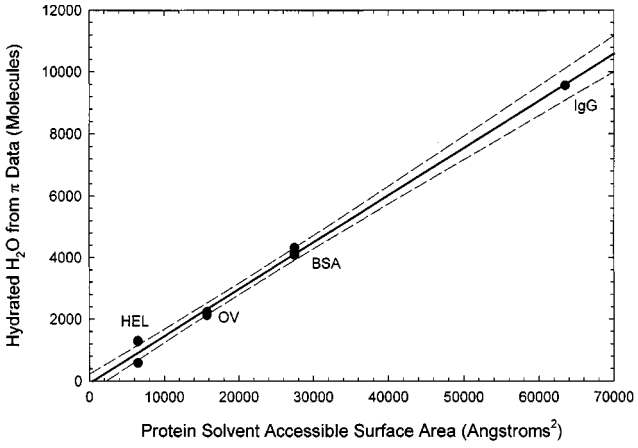

Figure 4. Solvent accessible surface area (SASA) vs. regressed hydration from fit of the free-solvent model to osmotic pressure data.

Line represents the linearly regressed best-fit relationship between water molecules and surface area (slope of 15.2 ± 0.5 molecules of water per nm^2 of surface area with a $R^2 = 0.991$). Remarkably this is parallel to the inverse of the projected surface area occupied by water molecules (16.2 molecules of water per nm^2 of surface area). Dashed lines correspond to the 95% confidence range of this fit.

for ν_{12} and ν_{32} are shown in Table 1. The degrees-of-freedom-adjusted R^2 value was 0.978. This, coupled with the small standard error for each variable, indicates a strong goodness-of-fit. The regressed hydration number, ν_{12}, is in agreement with the numbers obtained from ^{17}O–NMR studies (approximately one g-H_2O/g-globular protein) and falls within the range theoretically determined by Sedykh and Sedykh (1967) (0.292–1.044 g-H_2O/g-ovalbumin) and the same order of magnitude as Bull and Breese (1968, 1970, 1976) (0.295 g-H_2O/g-ovalbumin).

The limited concentration range for lysozyme resulted in an osmotic pressure curve that lacked the appropriate deviation from linear behavior to provide a strong goodness-of-fit for Eq. 4 when regressing on both ν_{12} and ν_{32}. Therefore, the estimated salt-binding information was used to fix ν_{32} (Curtis et al., 1998; Kuehner et al., 1999). Using the salt-binding results determined by Curtis et al. (1998), and assuming that there are equal amounts of anions and cations in the vicinity of the protein, Eq. 4 results in having only one fitted parameter, ν_{12}. The resulting regressed hydration values in 0.15 M NaCl and 0.15 M KCl are listed in Table 1. The hydration numbers obtained are 0.724 ± 0.008 g-H_2O/g-lysozyme in 0.15 M NaCl and 0.734 ± 0.003 g-H_2O/g-lysozyme in 0.15 M KCl. The hydration numbers obtained in this way are the same for either solution at the corresponding salt-binding value, and the predicted osmotic pressure for both solutions is nearly the same. Alternatively, when the salt binding suggested by Kuehner et al. (1999) is used, (4.0 mols-salt/mol-lysozyme) the hydration numbers are 1.614 ± 0.010 g-H_2O/g-lysozyme and 1.595 ± 0.007 g-H_2O/g-lysozyme, respectively (Table 1).

Monolayer Hydration for Globular Proteins

Table 1 also shows the hydration values determined for all the globular proteins studied. A general trend can be seen for an increase in the hydrated mols of water as the protein molecular weight increases. To better understand the relationship of the hydration to protein size, the solvent accessible surface area (SASA) for a solvent probe with a radius of 1.4 Å (the spherical model radius of water) was determined for each protein. Using the Research Collaboratory for Structural Bioinformatics (RCSB) (Berman et al., 2000), the protein data base identification (PDB ID) for each protein (or structurally similar protein) was determined. These data provide the unique structure for each globular protein. For lysozyme, ovalbumin, albumin, and immuno gamma globulin, the PDB IDs used were 4LYZ, 1OVA, 1AO6, and 1IGT, respectively (Diamond, 1974; Stein et al., 1991; Carter and Ho, 1994; Harris et al., 1997). The SASA for each structure was determined using GETAREA 1.1 with a 1.4-Å probe radius (Fraczkiewicz and Braun, 1998). Table 1 shows the respective surface areas for each protein.

Figure 4 shows the results of comparing the hydration or interactive water determined from the hydration model with the independently calculated SASA of the proteins. Assuming water to be represented as a spherical molecule with a radius of 1.4 Å, the regressed hydration values remarkably correspond to an equivalent monolayer of water in all cases except that of lysozyme with the hydration value determined from the salt-binding data provided by Kuehner et al. (1999). The best-fit linear regression gives a slope of 15.2 ± 0.5 molecules of water per nm^2 of surface area, with a $R^2 = 0.991$. This is very nearly equivalent to the inverse of the projected surface area occupied by water molecules (16.2 molecules of water per nm^2 of surface area). These results imply that the hydration values determined from the free-solvent model represent an equivalent monolayer of water. This is extraordinary considering that the hydration and SASA are all independently determined. Note that the free-solvent model only makes a distinction between molecules that are not of the free energy, such as that of bulk water. Thus it is plausible

for the water considered in the hydration layer for the free-solvent model to have different degrees of protein interaction (Richards, 1977; Creighton, 1993).

In terms of lysozyme, it is not clear whether the salt-binding data provided by Kuehner et al. (1999) or Curtis et al. (1998) is more accurate. However, if the salt binding of Kuehner et al. (1999) is determined to be more accurate, then the hydration number corresponds to two monolayers of water instead of one monolayer. While this hydration value is still within reasonable agreement, it is important to recognize that lysozyme is considered to have much more attractive interaction than the other proteins studied. Using the convention of the second virial coefficient, a negative second virial coefficient is an indication that less repulsive forces are dominant in such solutions and that the solution conditions are favorable for proteins to precipitate or crystallize (Tanford, 1961). Yousef (2000) found the second virial coefficient for lysozyme to be $-0.84 \pm 0.25 \times 10^{-7}$ L-mol/g^2 and $-1.37 \pm 0.49 \times 10^{-7}$ L-mol/g^2 for 0.15 M NaCl and KCl solutions, respectively. However, the virial coefficients for the other proteins studied were $4.94 \pm 3.27 \times 10^{-9}$ L-mol/g^2 for IgG (Yousef et al., 1998a), 1.063×10^{-7} L-mol/g^2, 1.114×10^{-7} L-mol/g^2, and 1.338×10^{-7} L-mol/g^2 for BSA in 4.5, 5.4, and 7.4 pH solutions (Vilker et al., 1981), and $1.79 \pm 0.08 \times 10^{-7}$ L-mol/g^2 and $1.96 \pm 0.09 \times 10^{-7}$ L-mol/g^2 for ovalbumin at 0.15 and 0.5 M NaCl, respectively (Yousef, 2000). Thus, the deviation of the calculated hydration number from a monolayer in the case using the salt-binding values of Kuehner et al. (1999) may be due to protein–protein interaction, which is not considered in the free-solvent model. Nevertheless, the close agreement of the hydration value indicates that the protein–solvent interactions are dominant, even in this case.

Conclusions

The free-solvent model was used to analyze the osmotic pressure of lysozyme at 25°C and pH 7.0 in the presence of 0.15 M NaCl or 0.15 M KCl and ovalbumin at pH 7.0 and 0.15 M NaCl. The results for these proteins combined with studies for BSA and IgG demonstrated that the hydration obtained from the free-solvent model was found to be one equivalent monolayer of water.

Given this, these results imply that, under the correct conditions, the osmotic-pressure data related to the free-solvent model provides the relative amount of water whose chemical potential significantly deviates from that in the bulk solvent. The chemical potential is fundamentally the most important parameter defining a species, and this may be one of the most complete methods for determining hydration under these circumstances. However, the high sensitivity of the hydration number to the osmotic pressure data and constraints in the method make this approach difficult to use to predict osmotic pressure at this time. These results also suggest that the free-solvent model may capture the dominant physical contribution to the observed nonidealities of the osmotic pressure of concentrated globular proteins in moderate ionic-strength aqueous solutions. It should be expected that even in solutions containing significantly solute–solute interaction, the hydration and salt-binding contributions cannot be ignored for concentrated solutions even though the free-solvent model cannot fully describe the osmotic pressure behavior (Yousef et al., 2001). Thus, for highly concentrated globular proteins, this approach appears to offers a more physically realistic starting point for models of solution nonideality than solute–solute-based interaction models in further understanding the thermodynamic properties of concentrated proteins in moderate salt solutions.

Acknowledgments

The USDA provided financial support through the Iowa Biotechnology Byproducts Consortium. We thank Kenneth P. (Kip) Murphy (University of Iowa), Kiszhake Soman (University of Texas Medical Branch), and Robin Curtis (UCLA) for their valuable discussions.

Notation

N_*^{II} = final available total solvent mols
N^j = initial total mols of solvent in compartment j
N_i^j = initial total mols of solvent species i in compartment j
n = number of solvent species
p = number of protein species
R = ideal gas constant
r = surface-to-surface distance between the solutes
T = temperature
$\bar{V}_i$ = specific molar volume
x_i = mol fraction

Greek letters

$\alpha = \nu_{32}/\nu_{12}$
κ = inverse of Debye length
μ_i = chemical potential of species i
π = osmotic pressure
ν_{ij} = net number of mols of solvent component i that is interacting with protein j

Superscripts and subscripts

I = compartment I
II = compartment II
1 = solvent
2 = protein
3 = salt

Literature Cited

Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, "The Protein Data Bank," *Nucleic Acids Res.*, **28**, 235 (2000).

Bull, H. B., and K. Breese, "Protein Hydration. I. Binding Sites," *Arch. Biochem. Biophys.*, **128**, 488 (1968).

Bull, H. B., and K. Breese, "Water and Solute Binding by Proteins. 1. Electrolytes," *Arch. Biochem. Biophys.*, **137**, 299 (1970).

Bull, H. B., and K. Breese, "Binding of Water and Electrolytes to Proteins. An Equilibrium Dialysis Study," *Biopolymers*, **15**, 1573 (1976).

Carter, D. C., and J. X. Ho, "Structure of Serum-Albumin," *Adv. Protein Chem.*, **45**, 153 (1994).

Claesson, P. M., T. Arnebrant, B. Bergenstaahl, and T. Nylander, *J. Colloid Interface Sci.*, **130**, 457 (1989).

Creighton, T. E., *Proteins. Structures and Molecular Properties*, 2nd ed., Freeman, New York (1993).

Curtis, R. A., J. M. Prausnitz, and H. W. Blanch, "Protein–Protein and Protein–Salt Interactions in Aqueous Protein Solutions Containing Concentrated Electrolytes," *Biotechnol. Bioeng.*, **57**(1), 11 (1998).

Diamond, R., "Real-Space Refinement of the Structure of Hen Egg-White Lysozyme," *J. Mol. Biol.*, **82**, 371 (1974).

Fraczkiewicz, R., and W. Braun, "Exact and Efficient Analytical Calculation of the Accessible Surface Areas and Their Gradients for Macromolecules," *J. Comput. Chem.*, **19**, 319 (1998).

Fraaije, J. G. E. M., and J. Lyklema, "Thermodynamics of Ion Binding by Proteins—Phenomenological Linkage Relations for Binding

of Electrolyte and Interpretation by Double-Layer Theory," *J. Biophys. Chem.*, **39**, 31 (1991).
Fraaoke, J. G. E. M., R. M. Murris, W. Norde, and J. Lyklema, "Interfacial Thermodynamics of Protein Adsorption, Ion Coadsorption and Ion Binding in Solution: 1. Phenomenological Linkage Relations for Ion-Exchange in Lysozyme Chromatography and Titration in Solution," *J. Biophys. Chem.*, **40**, 303 (1991a).
Fraaije, J. G. E. M., W. Norde, and J. Lyklema, "Interfacial Thermodynamics of Protein Adsorption, Ion Coadsorption and Ion Binding in Solution: 2. Model Interpretation of Ion-Exchange in Lysozyme Chromatography," *J. Biophys. Chem.*, **40**, 317 (1991b).
Frazer, J. C., and R. T. Myrick, "The Osmotic Pressure of Sucrose Solution at 30°," *J. Amer. Chem. Soc.*, **38**, 1907 (1916).
Gill, S. C., and P. H. von Hippel, "Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data," *Analytical BioChem.*, **182**, 319 (1989).
Glasstone, S., *Textbook of Physical Chemistry*, 2nd ed., Van Nostrand, Toronto, Ont., Canada (1946).
Harris, L. J., S. B. Larson, K. W. Hasel, and A. McPherson, "Refined Structure of an Intact IgG2a Monoclonal Antibody," *Biochemistry*, **36**, 1581 (1997).
Hill, T. L., *An Introduction to Statistical Thermodynamics*, Dover, New York (1986).
Israelachvili, J., *Intermolecular & Surface Forces*, 2nd ed., Academic Press, London (1992).
Israelachvili, J., and J. Marra, *Methods in Enzymology*, Vol. 127, L. Packer, ed., Academic Press, Orlando, FL, p. 353 (1986).
Kuehner, D. E., J. Engmann, F. Fergg, M. Wernick, H. W. Blanch, and J. M. Prausnitz, "Lysozyme Net Charge and Ion Binding in Concentrated Aqueous Electrolyte Solutions," *J. Phys. Chem. B*, **103**(8), 1368 (1999).
McMillan, W. G., and J. E. Mayer, "The Statistical Thermodynamics of Multicomponent Systems," *J. Chem. Phys.*, **13**(7), 276 (1945).
Otting, G., E. Liepinsh, and K. Wüthridh, "Protein Hydration in Aqueous Solution," *Science*, **254**, 974 (1991).
Pace, N. C., F. Vajdos, L. Fee, G. Grimsley, and T. Gray, "How to Measure and Predict the Molar Absorption Coefficient of a Protein," *Protein Science*, **4**, 2411 (1995).
Richards, F. M., "Area, Volume, Packing, and Protein Structure," *Ann. Rev. Biophys. Bioeng.*, **6**, 151 (1977).
Rupley, J. A., and G. Careri, "Protein Hydration and Function," *Adv. Protein Chem.*, **41**, 38 (1991).
Sedykh, L. G., and N. V. Sedykh, "Mathematical Models of the Phenomenon of Hydration of Biopolymers in Solutions," *Biofizika*, **12**(5), 936 (1967).
Stein, P. E., A. G. Leslie, J. T. Finch, and R. W. Carrell, "Crystal-Structure of Uncleaved Ovalbumin At 1.95 Å Resolution," *J. Mol. Biol.*, **221**, 941 (1991).
Tanford, C., *Physical Chemistry of Macromolecules*, Wiley, New York (1961).
Tombs, M. P., and A. R. Peacocke, *The Osmotic Pressure of Biological Macromolecules*, Clarendon Press, Oxford (1974).
van Laar, J. J., "Über die genauen Formeln für die Anderungen der Löslichkeit, für Gefrierpunkts- und Siedepunkts-änderungen, und für die Lösungs- und Verdünnungs-wärmen bei in Lösung dissocierten Körpern," *Z. Phys. Chem.*, **15**, 457 (1894).
Vilker, V. L., C. K. Colton, and K. A. Smith, "The Osmotic Pressure of Concentrated Protein Solutions: Effect of Concentration and pH in Saline Solutions of Bovine Serum Albumin," *J. Colloid Interface Sci.*, **79**, 548 (1981).
Yang, P., and J. A. Rupley, "Protein-Water Interactions. Heat Capacity of the Lysozyme-Water System," *Biochemistry*, **18**, 2654 (1979).
Yousef, M. A., R. Datta, and V. G. J. Rodgers, "Free-Solvent Model of Osmotic Pressure Revisited: Application to Concentrated IgG Solution at Physiological Conditions," *J. Colloid Interface Sci.*, **197**, 108 (1998a).
Yousef, M. A., R. Datta, and V. G. J. Rodgers, "Understanding Non-Idealities of the Osmotic Pressure of Concentrated Bovine Serum Albumin," *J. Colloid Interface Sci.*, **207**, 273 (1998b).
Yousef, M. A., R. Datta, and V. G. J. Rodgers, "Confirmation of Free Solvent Model Assumptions in Predicting the Osmotic Pressure of Concentrated Globular Proteins," *J. Colloid Interface Sci.*, **243** (2001).
Yousef, M. A., "The Effect of Molecular Interactions on the Osmotic Pressure and Diffusion Coefficients of Protein Solutions," PhD Thesis, Univ. of Iowa, Ames (2000).

Manuscript received Aug. 3, 2001, and revision received Dec. 4, 2001.