A Contribution to the Theory of Preferential Interaction Coefficients

J. Michael Schurr,* David P. Rangel,* and Sergio R. Aragon†
*Department of Chemistry, University of Washington, Seattle, Washington 98195-1700; and †Department of Chemistry/Biochemistry, San Francisco State University, San Francisco, California 94132

ABSTRACT A simple and complete derivation of the relation between concentration-based preferential interaction coefficients and integrals over the relevant pair correlation functions is presented for the first time. Certain omissions from the original treatment of pair correlation functions in multicomponent thermodynamics are also addressed. Connections between these concentration-based quantities and the more common molality-based preferential interaction coefficients are also derived. The pair correlation functions and preferential interaction coefficients of both solvent (water) and cosolvent (osmolyte) in the neighborhood of a macromolecule contain contributions from short-range repulsions and generic long-range attractions originating from the macromolecule, as well as from osmolyte-solvent exchange reactions beyond the macromolecular surface. These contributions are evaluated via a heuristic analysis that leads to simple insightful expressions for the preferential interaction coefficients in terms of the volumes excluded to the centers of the water and osmolyte molecules and a sum over the contributions of exchanging sites in the surrounding solution. The preferential interaction coefficients are predicted to exhibit the experimentally observed dependence on osmolyte concentration. Molality-based preferential interaction coefficients that were reported for seven different osmolytes interacting with bovine serum albumin are analyzed using the this formulation together with geometrical parameters reckoned from the crystal structure of human serum albumin. In all cases, the excluded volume contribution, which is the volume excluded to osmolyte centers minus that excluded to water centers in units of $\bar{V}_1$, exceeds in magnitude the contribution of the exchange reactions. Under the assumption that the exchange contribution is dominated by sites in the first surface-contiguous layer, the ratio of the average exchange constant to its neutral random value is determined for each osmolyte. These ratios all lie in the range 1.0 ± 0.15, which indicates rather slight deviations from random occupation near the macromolecular surface. Finally, a mechanism is proposed whereby the chemical identity of an osmolyte might be concealed from partially ordered multilayers of water in clefts, grooves, and pits, and its consequences are noted.

INTRODUCTION

The effects of weakly interacting osmolytes on the conformational equilibria and ligand binding reactions of biological macromolecules have been studied intensively over the past two decades (1–4). A major objective in many cases was to ascertain the difference between the number of water molecules ''associated'' with the products of a particular reaction on one hand and the corresponding number ''associated'' with its reactants on the other. The precise meaning, or interpretation, of the numbers of ''associated'' waters and the differences therein remains a subject of discussion and some debate (5–9). This general approach to studying changes in ''associated'' waters has come to be known as the osmotic stress method. In the case of a solution, consisting of water (solvent, component 1), dilute macromolecules (components 2_J, $J = 1, \ldots M$), and neutral osmolyte (cosolvent, component 3), the osmotic stress method yields the slope $(\partial \ln K/\partial \ln a_1)_{T,P,c_{2J}}$, where K is the equilibrium constant for the reaction when written so as to take no account of either water or osmolyte, a_1 is the activity of the water, and c_{2_J} denotes the concentrations of each kind of macromolecule. This slope is extrapolated to the limit of infinite dilution, $c_{2_J}^{\infty} \to 0$. The difference in ''associated'' waters between products and reactants is sometimes taken to be the aforementioned slope,

$$\begin{aligned}\Delta\Gamma_1 &\equiv (\partial \ln K/\partial \ln a_1)_{T,P,c_{2_J}^{\infty}} \\ &= -\Big(\sum_p \nu_p \Big(\partial \mu_p^0/\partial \mu_1\Big)_{T,P,c_{2J}^{\infty}} - \sum_r \nu_r \big(\partial \mu_r^0/\partial \mu_1\big)_{T,P,c_{2J}^{\infty}}\Big) \\ &= \sum_p \nu_p \Gamma_1(p) - \sum_r \nu_r \Gamma_1(r), \end{aligned} \quad (1)$$

where the index p or r denotes macromolecular products or reactants, respectively, μ_p^0 and μ_r^0 denote the respective standard state chemical potentials, ν_p and ν_r denote the respective stoichiometric coefficients of the reaction under consideration, and

$$\Gamma_1(2_J) \equiv -\Big(\partial \mu_{2_J}^0/\partial \mu_1\Big)_{T,P,c_{2J}^{\infty}}. \quad (2)$$

$\Gamma_1(2_J)$ and the symmetrically defined $\Gamma_3(2_J) = -(\partial \mu_{2J}^0/\partial \mu_3)_{T,P,c_{2_J}^{\infty}}$ are concentration-based ''preferential interaction coefficients'', which characterize the variation of that part of μ_{2J} that does not depend upon c_{2J} with either μ_1 or μ_3, respectively.

Alternative preferential interaction coefficients are defined in connection with equilibrium dialysis experiments and are usually molality based. The molalities of species 1, 2, and 3 are denoted by, respectively, $m_1 = 55.6$, m_2, and m_3. Two

Submitted December 1, 2004, and accepted for publication July 1, 2005.

Address reprint requests to J. Michael Schurr, Tel.: 206-543-6681; Fax: 206-685-8665; E-mail: schurr@chem.washington.edu.

0006-3495/05/10/2258/19 $2.00

doi: 10.1529/biophysj.104.057331

common molality-based preferential interaction coefficients are: $\Gamma^{\mathrm{m}}_{\mu_3} = (\partial m_3/\partial m_2)_{\mathrm{T,P},\mu_3}$ and $\Gamma^{\mathrm{m}}_{\mu_1,\mu_3} = (\partial m_3/\partial m_2)_{\mathrm{T},\mu_1,\mu_3}$, where the index J denoting the macromolecular conformation has been suppressed. Although relations between these and other molality-based preferential interaction coefficients have been intensively investigated, the connections between molality-based and concentration-based preferential interaction coefficients, like $\Gamma_1(2_{\mathrm{J}})$, have received less attention. Clever and intuitive thermodynamic approaches indicate that for any given macromolecular species 2,

$$\Gamma^{\mathrm{m}}_{\mu_1,\mu_3} = N_{32} - (c_3/c_1)N_{12}, \tag{3}$$

where N_{12} and N_{32} denote the total number of water and osmolyte molecules, respectively, in a domain of sufficient size surrounding a single isolated macromolecule, and c_1 and c_3 denote the respective bulk concentrations in an exterior domain, no part of which is near any macromolecule (1–9). $\Gamma^{\mathrm{m}}_{\mu_1,\mu_3}$ can be regarded as the excess number of osmolyte molecules in the vicinity of the macromolecule above the quantity that would be expected from the number of water molecules in that region and the bulk concentration ratio, c_3/c_1.

Although the analysis below indicates that Eq. 3 is correct, the rigor of the thermodynamic approaches used to derive it is debatable. For example, the neglect of the osmotic pressure due to the macromolecule within its local domain is justifiable only for a domain of very great size, yet in many cases that domain was assumed to extend no more than one or two hydration layers beyond the macromolecule. The likely resolution of this paradoxical circumstance is noted briefly below.

Recently several articles appeared in which $\Gamma_1(2)$, or the equivalent $\Gamma_3(2) = -(c_3/c_1)\Gamma_1(2)$, was expressed in terms of the so-called Kirkwood-Buff integrals (10), $G_{12} \equiv \int d^3\mathbf{r}(g_{12}(r) - 1)$ and $G_{32} \equiv \int d^3\mathbf{r}(g_{32}(r) - 1)$, where $g_{12}(r)$ and $g_{32}(r)$ are the pair correlation functions, which are described in greater detail below (11–14). The derivation of the main relation followed an unusually circuitous, piecewise, and technically demanding route that took place over three different articles and a book that collectively spanned 26 years (11,15–17). Chitra and Smith combined two relations that appeared earlier in Ben-Naim's book (17), namely his Eq. 6.7.49 for $(\partial \ln c_3/\partial \ln a_3)_{\mathrm{T,P},c_2\to 0}$ and Eq. 6.17.16 for $(\partial \mu_2/\partial \ln c_3)_{\mathrm{T,P},c_2\to 0}$, to obtain the final expression for $\Gamma_3(2)$. The Eq. 6.7.49 was explicitly derived in Ben-Naim's book, but the derivation of the much more difficult Eq. 6.17.16 was simply described as quite lengthy and omitted entirely. In fact, the first stage of that proof was presented in his 1975 article (15), and the second stage was presented in his 1988 article (16). Unfortunately, neither Chitra and Smith (11) nor Ben-Naim (17) referenced directly those earlier articles, from which the entire proof could be assembled. Chitra and Smith (11) demonstrated the approximate validity of their expression for $\Gamma_1(2)$ by molecular dynamics simulations of both the pair correlation functions and the free energies of insertion of different small species 2 into aqueous solutions over a wide range of concentrations of various cosolvents. Shimizu (13) suggested a way to obtain the separate G_{12} and G_{32} from the measured $\Gamma_1(2)$ and $\bar{V}_2$, where $\bar{V}_2$ is the partial molecular volume. He employed a relation between $\bar{V}_2$ and G_{12} and G_{32} that was also first presented in Ben-Naim's book (17) (Eq. 6.17.22), but the derivation, described as quite lengthy, was also omitted entirely. Again, a two-stage proof of the relevant relation can be found in the same two earlier articles (15,16). Shimizu (12) also extended his idea to determine the changes, ΔG_{32} and ΔG_{12}, accompanying a reaction of species 2 from the measured $\Delta\Gamma_1(2)$ and $\Delta\bar{V}_2$, which was assumed to be the entire ΔV associated with the reaction. Shimizu and Smith (14) examined the differences between the effects of crowders, such as polyethylene glycol, and small osmolytes, such as glycerol, that stabilize native protein structures, on the separate G_{12} and G_{23}. Schellman (18) undertook a related analysis in terms of the cross-second virial coefficients (B_{23}).

The initial objective of this study is to provide a complete and much simpler derivation of the relevant expression for $\Gamma_1(2)$ directly from the results of Kirkwood and Buff (10), as well as some important details that are missing from their original treatment of multicomponent thermodynamics. Such details include the choice of origin of the coordinate frame in a highly deformable macromolecule, its manifestation in the pair correlation functions, the invariance of the integrals of $g_{\alpha\beta}(r) - 1$ to that choice, a complete definition of the pair correlation function in the classical grand ensemble, and a derivation of the partial molecular volume. This derivation of $\Gamma_1(2)$ follows a considerably more direct line than the Ben-Naim-Chitra-Smith development, and is technically much simpler. All of the results of Kirkwood and Buff that are needed to derive $\Gamma_1(2)$ were rederived and found to be correct. In addition, a short proof of Ben-Naim's expression for $\bar{V}_2$ is provided in Appendix D.

Connections between this concentration-based $\Gamma_1(2)$ and the molality-based $\Gamma^{\mathrm{m}}_{\mu_3}$ and $\Gamma^{\mathrm{m}}_{\mu_1,\mu_3}$ are derived via thermodynamic arguments that make use of certain expressions of Anderson et al. (19,20), which were also verified by rederivation.

The main objective of this study is to clarify the meaning(s) of the $\Gamma_1(2)$ and $\Gamma_3(2)$, and especially to relate them to more familiar quantities such as excluded volumes and equilibrium constants for osmolyte-solvent exchange in the region surrounding the macromolecule (21–26). Although this development is more heuristic than rigorous, useful predictions and significant insights emerge. As an example, the experimental $\Gamma^{\mathrm{m}}_{\mu_1,\mu_3}$ data for seven different osmolytes interacting with bovine serum albumin (BSA) are analyzed using this formulation in conjunction with geometrical parameters reckoned from the crystal structure of human serum albumin (HSA). The separate excluded volume and exchange contributions are evaluated. Under the assumption that only the surface-contiguous layer of osmolyte sites is

important, the ratio of the average exchange constant to its neutral random value is obtained in each case.

Finally, a mechanism is proposed whereby the chemical identity of the osmolyte may be concealed from partially ordered hydration multilayers in clefts, grooves, and pits, and its consequences are briefly noted.

A DERIVATION OF $\Gamma_1(2)$

Let us consider a system comprising ν different molecular species, $\alpha, \beta, \ldots \eta$, at constant T, V. In this case, when each species j undergoes a change of dN_j mol,

$$d\mu_\beta = \sum_{j=1}^{\nu}\left(\frac{\partial\mu_\beta}{\partial N_j}\right)_{T,V,N_{\gamma\neq j}} dN_j = \sum_{j=1} M_{\beta j} dN_j, \qquad (4)$$

where

$$M_{\alpha\beta} \equiv \left(\frac{\partial\mu_\alpha}{\partial N_\beta}\right)_{T,V,N_{\gamma\neq\beta}} = \frac{1}{V}\left(\frac{\partial\mu_\alpha}{\partial c_\beta}\right)_{T,V,c_{\gamma\neq\beta}}. \qquad (5)$$

Thus, the column vector containing the ν different $d\mu_k$ is related to the ν different dN_j by the matrix relation $d\boldsymbol{\mu} = \mathbf{M}\,\mathbf{dN}$, where the elements of $\mathbf{M}$ are given by Eq. 5. Inversion of this matrix relation gives $\mathbf{dN} = \mathbf{M}^{-1}d\boldsymbol{\mu}$, or

$$dN_\alpha = \sum_{\beta=1}^{\eta}\left(\frac{\partial N_\alpha}{\partial\mu_\beta}\right)_{T,V,\mu_{\gamma\neq\beta}} d\mu_\beta = \sum_{\beta=1}^{\eta}(\mathbf{M}^{-1})_{\alpha\beta}d\mu_\beta, \qquad (6)$$

where

$$(\mathbf{M}^{-1})_{\alpha\beta} = \left(\frac{\partial N_\alpha}{\partial\mu_\beta}\right)_{T,V,\mu_{\gamma\neq\beta}}. \qquad (7)$$

Kirkwood and Buff (10) established that the $(\mathbf{M}^{-1})_{\alpha\beta}$ in Eq. 7 are directly related to integrals of the relevant pair correlation functions,

$$B_{\alpha\beta} \equiv c_\alpha\delta_{\alpha\beta} + c_\alpha c_\beta\int_0^\infty d^3\mathbf{r}(g_{\alpha\beta}(r)-1) = (kT/V)(\mathbf{M}^{-1})_{\alpha\beta}, \qquad (8)$$

where $g_{\alpha\beta}(r)$ is the $\alpha\beta$-pair correlation function, or radial distribution function, and $r = |\mathbf{r}_1 - \mathbf{r}_2|$ is the distance between the arbitrarily chosen central atom of an α-molecule at $\mathbf{r}_1$ and that of a β-molecule at $\mathbf{r}_2$, as indicated in Appendix A. A complete definition of $g_{\alpha\beta}(r_{12})$ in the grand ensemble (27) is given in Eq. A1 in Appendix A. It must be emphasized that $g_{\alpha\beta}(r_{12})$ pertains to no atoms other than the arbitrarily chosen central atom of each molecule, and will in general depend upon that choice. Because the relations presented here derive ultimately from fluctuations in the numbers of molecules in a volume V that is large enough to contain on average a great many molecules of each kind, those relations must be independent of the choice of central atom. It may be concluded from Eq. 8 that integrals of the $g_{\alpha\beta}(r) - 1$ over the volume V, or at least from 0 out to a distance where $g_{\alpha\beta}(r)$ has converged to 1.0, are independent of the choice of central atom. The grand ensemble used to derive Eq. 8 can itself be derived by considering that the volume V is a tiny fraction of an enormously larger supersystem with a fixed number of molecules (27).

The pair correlation function has the following physical meaning. If the chosen central atom of a molecule of kind α is located at $\mathbf{r}_1$, then $c_\beta g_{\alpha\beta}(r)$ is the probability per unit volume of finding the chosen central atom of a molecule of kind β at $\mathbf{r}_2$, such that $r = |\mathbf{r}_1 - \mathbf{r}_2|$. A completely random disposition of β-molecules in the vicinity of α corresponds to $g_{\alpha\beta}(r) = 1.0$. In general, $g_{\alpha\beta}(r)$ is the factor by which the purely random probability per unit volume (i.e., c_β) must be multiplied to reckon the actual probability per unit volume of finding a β-molecule at distance r from an α-molecule. The pair correlation functions are by definition symmetric, so $g_{\alpha\beta}(r) = g_{\beta\alpha}(r)$, and also $B_{\alpha\beta} = B_{\beta\alpha}$. We shall later regard $c_\beta g_{\alpha\beta}(r)$ as the rotationally averaged mean density of centers of β-molecules at a distance r from the center of an α-molecule.

The matrix relation in Eq. 8 can be written as $\mathbf{B} = (kT/V)\mathbf{M}^{-1}$, which can be inverted to give $\mathbf{M} = (kT/V)\mathbf{B}^{-1}$, and

$$\left(\frac{\partial\mu_\alpha}{\partial c_\beta}\right)_{T,V,c_{\gamma\neq\beta}} = kT(B^{-1})_{\alpha\beta} = kT\frac{|B|_{\alpha\beta}}{|B|}, \qquad (9)$$

where $|B|_{\alpha\beta}$ is the cofactor of $B_{\alpha\beta}$ (i.e., $(-1)^{\alpha+\beta}$ times the determinant of the matrix obtained by striking out the αth row and the βth column) and $|B|$ denotes the determinant of $\mathbf{B}$ (10).

For the particular case of a three-component system held at constant T and V, the chemical potential $\mu_2(T,c_1,c_2,c_3)$ depends on all three concentrations, so

$$d\mu_2 = \left(\frac{\partial\mu_2}{\partial c_1}\right)_{c_2,c_3} dc_1 + \left(\frac{\partial\mu_2}{\partial c_2}\right)_{c_1,c_3} dc_2 + \left(\frac{\partial\mu_2}{\partial c_3}\right)_{c_1,c_2} dc_3. \qquad (10)$$

The constant T subscript is suppressed in Eqs. 10–16 below. When c_2 is held constant, then $d\mu_2 = d\mu_2^0$, and it follows from Eq. 10 that

$$\left(\frac{\partial\mu_2}{\partial\mu_1}\right)_{P,c_2} = \left(\frac{\partial\mu_2^0}{\partial\mu_1}\right)_{P,c_2} = \left(\frac{\partial\mu_2}{\partial c_1}\right)_{c_2,c_3}\left(\frac{\partial c_1}{\partial\mu_1}\right)_{P,c_2} + \left(\frac{\partial\mu_2}{\partial c_3}\right)_{c_1,c_2}\left(\frac{\partial c_3}{\partial\mu_1}\right)_{P,c_2}. \qquad (11)$$

An equation analogous to Eq. 10 holds for $d\mu_1$, from which it follows that

$$\left(\frac{\partial\mu_1}{\partial\mu_1}\right)_{P,c_2} = 1 = \left(\frac{\partial\mu_1}{\partial c_1}\right)_{c_2,c_3}\left(\frac{\partial c_1}{\partial\mu_1}\right)_{P,c_2} + \left(\frac{\partial\mu_1}{\partial c_3}\right)_{c_1,c_2}\left(\frac{\partial c_3}{\partial\mu_1}\right)_{P,c_2}. \qquad (12)$$

The change in $c_1(T,P,c_2,c_3)$ at constant T,P,c_2 is

$$dc_1 = \left(\frac{\partial c_1}{\partial c_3}\right)_{P,c_2} dc_3. \tag{13}$$

It is shown in Appendix B that $(\partial c_1/\partial c_3)_{P,c_2} = -\bar{V}_3/\bar{V}_1$, where $\bar{V}_j$ denotes the partial molecular volume (m^3/molecule). Then Eq. 13 yields

$$\left(\frac{\partial c_1}{\partial \mu_1}\right)_{P,c_2} = -(\bar{V}_3/\bar{V}_1)\left(\frac{\partial c_3}{\partial \mu_1}\right)_{P,c_2}. \tag{14}$$

After substituting Eq. 14 into Eq. 12 and rearranging one finds

$$\left(\frac{\partial c_3}{\partial \mu_1}\right)_{P,c_2} = 1/\left(-(\bar{V}_3/\bar{V}_1)\left(\frac{\partial \mu_1}{\partial c_1}\right)_{c_2,c_3} + \left(\frac{\partial \mu_1}{\partial c_3}\right)_{c_1,c_2}\right). \tag{15}$$

After substituting Eqs. 14 and 15 into Eq. 11, and Eq. 11 into Eq. 2, there results

$$\begin{aligned}\Gamma_1(2) = (-)\frac{\left[-(\bar{V}_3/\bar{V}_1)\left(\frac{\partial \mu_2}{\partial c_1}\right)_{c_2^\infty,c_3} + \left(\frac{\partial \mu_2}{\partial c_3}\right)_{c_1,c_2^\infty}\right]}{\left[-(\bar{V}_3/\bar{V}_1)\left(\frac{\partial \mu_1}{\partial c_1}\right)_{c_2^\infty,c_3} + \left(\frac{\partial \mu_1}{\partial c_3}\right)_{c_1,c_2^\infty}\right]} \\ = -\frac{\left[-(\bar{V}_3/\bar{V}_1)|B|^\infty_{21} + |B|^\infty_{23}\right]}{\left[-(\bar{V}_3/\bar{V}_1)|B|^\infty_{11} + |B|^\infty_{13}\right]}.\end{aligned} \tag{16}$$

Equation 9 was used to obtain the second line of Eq. 16 from the first, and the superscript ∞ on the $|B|_{\alpha\beta}$ indicates that they are to be evaluated in the limit $c_2^\infty \to 0$. The $\bar{V}_3/\bar{V}_1$ must be evaluated in the same limit.

The right-hand side of Eq. 16 is partially evaluated by leaving the $\bar{V}_3/\bar{V}_1$ factors in place, but expanding the $|B|^\infty_{\alpha\beta}$ in terms of elements of the three-component B-matrix, $B_{\alpha\beta} = c_\alpha\delta_{\alpha\beta} + c_\alpha c_\beta G_{\alpha\beta}$, where

$$G_{\alpha\beta} \equiv \int_0^\infty d^3\mathbf{r}(g_{\alpha\beta}(r) - 1). \tag{17}$$

Every term in both the numerator and denominator of the right-hand side of Eq. 16 contains at least one factor of c_2, which can be divided out. Any remaining terms that still contain a factor of c_2 will vanish in the limit $c_2^\infty \to 0$, and are therefore omitted. After effecting some factorization and cancellation, the result can be expressed as

$$\Gamma_1(2) = c_1G_{12} - c_3G_{32}\left(\frac{B_{11} + (\bar{V}_3/\bar{V}_1)B_{13}}{B_{31} + (\bar{V}_3/\bar{V}_1)B_{33}}\right). \tag{18}$$

It remains to evaluate the factor in parentheses on the right-hand side of Eq. 18. An expression for $\bar{V}_\alpha$ was presented by Kirkwood and Buff (10) without explicit derivation. That derivation is sketched briefly in Appendix C and the result is given in Eq. C6. Note that the denominator of Eq. C6 is independent of α, and cancels out of the ratio, $\bar{V}_3/\bar{V}_1$. An important simplification is that $\bar{V}_3/\bar{V}_1$ applies in the limit $c_2^\infty \to 0$, which leaves just a two-component (2×2), rather than a three-component (3×3) B-matrix, so the indicated cofactors become just elements of the two-component B-matrix. In fact, Eq. C6 gives the simple expressions, $\bar{V}_3 = (-c_1B_{13} + c_3B_{11})/D$ and $\bar{V}_1 = (c_1B_{33} - c_3B_{31})/D$, where D is the denominator, which cancels out of $\bar{V}_3/\bar{V}_1$. After performing straightforward algebra, invoking the symmetry, $B_{\alpha\beta} = B_{\beta\alpha}$, and omitting any canceling terms, the entire factor in parentheses reduces to c_1/c_3, and Eq. 18 becomes simply

$$\Gamma_1(2) = c_1G_{12} - (c_1/c_3)c_3G_{32}. \tag{19}$$

INTERPRETATION AND DISCUSSION

From the definition of G_{12} in Eq. 17, it is clear that c_1G_{12} is the excess number of 1-molecules in the vicinity of a 2-molecule beyond what would be expected from a random disposition of 1-molecules. An analogous meaning holds for c_3G_{32}. Although the c_1G_{12} and c_3G_{32} in Eq. 19 are explicitly excess numbers, rather than the total numbers of molecules in a domain surrounding the 2-molecule, Eq. 19 for $\Gamma_1(2)$ can be written in a form that is completely analogous to Eq. 3 for $\Gamma^m_{\mu_1,\mu_3}$, as will be seen.

The pair correlation functions $g_{12}(r)$ and $g_{32}(r)$ must converge to the value 1.0 at large distances. Typically, for small osmolytes in a solution of dilute macromolecules, this occurs within, at most, a few nanometers beyond the maximum extension of the macromolecule (species 2). Thus, the upper limit of the integral in G_{12} or G_{32} can be reduced from ∞ to R, where R is any value sufficiently great that both $g_{12}(r)$ and $g_{32}(r)$ have converged to 1.0. Then Eq. 19 can be written as

$$\begin{aligned}\Gamma_1(2) &= c_1\left(\int_0^R d^3\mathbf{r}(g_{12}(r) - 1) - \int_0^R d^3\mathbf{r}(g_{32}(r) - 1)\right) \\ &= c_1\left(\int_0^R d^3\mathbf{r}g_{12}(r) - \int_0^R d^3\mathbf{r}g_{32}(r)\right) \\ &= N_{12} - (c_1/c_3)N_{32},\end{aligned} \tag{20}$$

where $N_{\alpha 2} \equiv c_\alpha \int_0^R d^3\mathbf{r}g_{\alpha 2}(r)$ is the number of α-molecules within a sphere of radius R around the 2-molecule. The relevant criterion for the minimum size, R_{min}, of the domain surrounding the macromolecule is clearly the convergence of the relevant pair correlation functions to 1.0 at all $r \geq R_{min}$. Because standard osmolytes are typically at least a few times larger than water, species 3 is typically excluded by the macromolecule from a larger volume than is species 1. Consequently, $g_{32}(r)$ cannot possibly converge to 1.0 within the volume defined by the centers of 1-molecules in the first hydration shell, and the minimum domain size generally must involve more water molecules than those in the first hydration shell in order for Eq. 20 to be valid.

Equation 20 is rigorously valid for a finite domain size of radius $R \geq R_{min}$, even though no account was taken of the osmotic pressure due to the macromolecule. This is likely

a consequence of allowing the domain boundary to move with the macromolecule, so that it can never be contacted by the macromolecule and never experience its contribution to the osmotic pressure inside the macromolecular domain.

The preferential interaction coefficient can also be written in the simple form

$$\Gamma_1(2) = c_1\left(\int_0^R d^3\mathbf{r}(g_{12}(r) - g_{32}(r))\right), \qquad (21)$$

which is most useful for our analysis. Corresponding expressions for $\Gamma_3(2)$ can also be obtained simply by replacing the index 1 by 3 and vice versa in Eqs. 2 and 3 and 19–21, which is permitted by the evident symmetry of the theory in regard to $1 \leftrightarrow 3$ interchange. It follows from Eqs. 19–21 that

$$\Gamma_3(2) = -(c_3/c_1)\Gamma_1(2) = c_3\int_0^R d^3\mathbf{r}(g_{32}(r) - g_{12}(r)). \qquad (22)$$

The right-hand side of Eq. 22 is just $N_{32} - (c_3/c_1)N_{12}$, which matches the right-hand side of Eq. 3. Furthermore, it is shown by thermodynamic arguments in Appendix E that in the limit, $c_2, m_2 \to 0$,

$$\Gamma_3(2) = \Gamma^{\mathrm{m}}_{\mu_1,\mu_3} = \Gamma^{\mathrm{m}}_{\mu_3}\left[1 - c_3\bar{V}_3\left(1 - \left(\Gamma^{\mathrm{m}}_{\mu_1}/\Gamma^{\mathrm{m}}_{\mu_3}\right)\right)\right], \qquad (23)$$

where $\Gamma^{\mathrm{m}}_{\mu_1} \equiv (\partial m_3/\partial m_2)_{\mathrm{T,P},\mu_1}$ (c.f. Eqs. E16 and E19). The relations in Eq. 23 were obtained by assuming that $\bar{V}_1$, $\bar{V}_2$, and $\bar{V}_3$ remain constant, independent of c_3 and c_2. This should be a rather good approximation, when $c_2 \to 0$ and $c_3\bar{V}_3 \ll 1.0$, which correspond to prevailing conditions in many studies. $\Gamma^{\mathrm{m}}_{\mu_1}$ is obtained via vapor pressure osmometry, and $\Gamma^{\mathrm{m}}_{\mu_1,\mu_3}$ is measured by equilibrium dialysis. At typically low osmolyte volume fractions ($c_3\bar{V}_3 \ll 1.0$), $\Gamma^{\mathrm{m}}_{\mu_3}$ is quite close to $\Gamma^{\mathrm{m}}_{\mu_1,\mu_3}$, but $\Gamma^{\mathrm{m}}_{\mu_1}$ is rather different (5,19,20). In any case, most experimental work has reported $\Gamma^{\mathrm{m}}_{\mu_3}$ or $\Gamma^{\mathrm{m}}_{\mu_1,\mu_3}$ or both. Equation 23 thus provides the principal connections between these theoretical expressions for $\Gamma_3(2)$ (or $\Gamma_1(2)$) in terms of pair correlation functions and the experimentally measured quantities.

We note that this $\Gamma_3(2)$ cannot be simply expressed as $(\partial c_3/\partial c_2)_{\mathrm{T,P},\mu_3}$, because there is no Maxwell relation equating $(\partial\mu_3/\partial c_2)_{\mathrm{T,P},c_3}$ to $(\partial\mu_2/\partial c_3)_{\mathrm{T,P},c_2}$. Moreover, $\Gamma_3(2)$ is also not equivalent to $(\partial c_3/\partial c_2)_{\mathrm{T},\mu_1,\mu_3}$, because direct evaluation of the latter in terms of N_{12} and N_{32} (9) yielded a result that is not equivalent to the right-hand side of Eq. 22.

Radial distribution functions of multicomponent systems have not yet been treated rigorously and analytically, and no suitable approximate formulation in terms of basic quantities, such as excluded volumes and exchange constants for specific sites, was presented previously. Heuristic approximate evaluations of various contributions to $c_1g_{12}(r)$, $c_3g_{32}(r)$, $\Gamma_1(2)$, and $\Gamma_3(2)$ are presented in the following section.

HEURISTIC EVALUATION OF $\Gamma_1(2)$ AND $\Gamma_3(2)$

In general, both repulsive exclusion forces and attractive binding forces contribute simultaneously to $\Gamma_1(2)$ and $\Gamma_3(2)$. These contributions are evaluated approximately below. Comparisons with the models adopted by other workers will be discussed after this model is developed.

Repulsive exclusion forces

To simplify the discussion, let us first consider the effects of repulsive hard-core exclusion forces between the water (species 1) and the macromolecule (species 2). The superscript ''ex'' is used to indicate a contribution arising from such forces. A substantial void region, where $g^{\mathrm{ex}}_{12}(r) \cong 0$, is expected around $r = 0$, as illustrated in Fig. 1. If both species 1 and 2 were perfectly spherical, then this void region would be followed at larger r by the region of the first coordination shell, where $g^{\mathrm{ex}}_{12}(r) > 1.0$ (11,17,28). This is true even in the case of hard spheres with no attractive interactions whatsoever. The first coordination shell would then be followed by a dip of $g_{12}(r)$ below 1.0, which in turn would be followed by a weaker second coordination shell, a second shallower dip, and so on, finally leveling off to $g_{12}(r) = 1.0$. In the case of a nonspherical macromolecule, the dips and peaks associated with the void volumes and coordination shells arising from different parts of the surface are superposed with a distribution of relative ''phases'', so that $g^{\mathrm{ex}}_{12}(r)$ likely exhibits simply a more or less smooth rise to a plateau at 1.0, as indicated in Fig. 1. Because typical neutral osmolytes (species 3) are larger than water, the void regions of $g^{\mathrm{ex}}_{32}(r)$ would extend outward somewhat farther than in the case of $g^{\mathrm{ex}}_{12}(r)$, as indicated also in Fig. 1. The volumes excluded to the centers of species 3 and 1 can be expressed as $V^{\mathrm{ex}}_3 = \int_0^R d\mathbf{r}\left(1 - g^{\mathrm{ex}}_{32}(r)\right)$ and $V^{\mathrm{ex}}_1 = \int_0^R d\mathbf{r}\left(1 - g^{\mathrm{ex}}_{12}(r)\right)$, respectively. The difference between the volumes accessible to the centers of species 1 and 3 within the macromolecular domain is defined by,

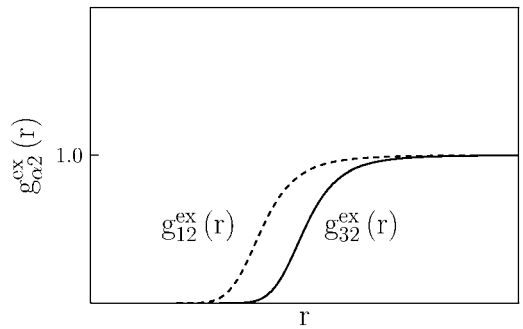

FIGURE 1 Schematic illustration of $g^{\mathrm{ex}}_{12}(r)$ and $g^{\mathrm{ex}}_{32}(r)$ versus the distance r between the central atoms of species 2 and either 1 or 3, respectively. The $g^{\mathrm{ex}}_{12}(r)$ and $g^{\mathrm{ex}}_{32}(r)$ are those parts of the pair correlation functions that arise solely from repulsive exclusion forces between species 2 and either 1 or 3, respectively.

$\Delta V^{ac} \equiv \int_0^R d\mathbf{r}\left(g_{12}^{ex}(r) - g_{32}^{ex}(r)\right) = V_3^{ex} - V_1^{ex}$, which is also the difference between the volumes excluded to species 3 and 1.

The $c_1\Delta V^{ac}$ contribution to $\Gamma_1(2)$ can be understood heuristically in terms of the osmotic pressure-volume work required to introduce a 2-molecule into the solution. The 2-molecule must effectively extrude the centers of the osmolytes (species 3) from a region occupied by the centers of the waters (species 1), which requires the input of work equal to $\pi\left(V_3^{ex} - V_1^{ex}\right)$, where π is the osmotic pressure of species 1 in the bulk solution. This work appears as a term in μ_2^0, which is the increase in solution free energy upon adding a 2-molecule to the solution. The variation of the osmotic pressure of species 1 with its activity is given by $\bar{V}_1 d\pi = -kTd\ \ln\ a_1$. Thus, the osmotic work contribution to $\Gamma_1(2)$ is $-\left(\partial\mu_2^0/\partial\mu_1\right)_{T,P,c_{2J}^\infty}^{ex} = -(1/kT)\left(\partial\mu_2^0/\partial \ln\ a_1\right)_{T,P,c_{2J}^\infty}^{ex} = (1/\bar{V}_1)\left(\partial\mu_2^0/\partial\pi\right)_{T,P,c_{2J}^\infty}^{ex} = (1/\bar{V}_1)\left(V_3^{ex} - V_1^{ex}\right) \cong c_1\Delta V^{ac}$, when c_3 is sufficiently dilute that $c_1 \cong 1/\bar{V}_1$. This simple analysis breaks down, when c_3 becomes comparable to c_1.

In the void regions, where $g_{12}^{ex}(r)$ and $g_{32}^{ex}(r)$ vanish, $g_{12}^{ex}(r)$ and $g_{32}^{ex}(r)$ are practically independent of either c_1 or c_3. The contribution of repulsive exclusion forces to $\Gamma_1(2)$ is obtained from Eq. 21 as

$$\Gamma_1^{ex}(2) = c_1\int_0^R d^3\mathbf{r}\left(g_{12}^{ex}(r) - g_{32}^{ex}(r)\right) \equiv c_1\Delta V^{ac} = c_1\left(V_3^{ex} - V_1^{ex}\right). \quad (24)$$

Any variation of ΔV^{ac} with c_1 or c_3 should be rather slight, due to the constancy of the void volumes, so $\Gamma_1^{ex}(2)$ should remain nearly constant, so long as c_1 doesn't change much from the value, $c_1 = 1/\bar{V}_1$, which will be the case, provided that $c_3 \leq 1.0$ M. Due to the generally larger void volume of $g_{32}(r)$ in comparison to $g_{12}(r)$, both $\Delta V^{ac} = \left(V_3^{ex} - V_1^{ex}\right)$ and $\Gamma_1^{ex}(2)$ should be generally positive. In view of Eqs. 22 and 24, it is also expected that

$$\Gamma_3^{ex}(2) = -c_3\Delta V^{ac}, \quad (25)$$

where $\Gamma_3^{ex}(2)$ denotes the contribution of purely repulsive exclusion forces. Hence, $\Gamma_3^{ex}(2)$ is expected to be proportional to c_3 and negative.

Generic long-range attractive forces

Let us now consider generic attractive forces, long-range van der Waals forces in particular, that may affect the densities of (centers of) species 1 and 3 in the region immediately beyond the void volume. Such mean densities are denoted by $c_1 g_{12}^{ga}(r)$ and $c_3 g_{32}^{ga}(r)$, where the superscript ''ga'' denotes generic attractions. For simplicity it will be assumed here that such generic attractions do not discriminate significantly between species 1 or 3, so that the ratio of their densities at any r beyond the void volume matches that of the bulk solution, that is $c_1 g_{12}^{ga}(r)/\left(c_3 g_{32}^{ga}(r)\right) \cong c_1/c_3$, which implies that $g_{12}^{ga}(r) \cong g_{32}^{ga}(r)$, even though both may differ significantly from 1.0. In that case, the net contributions to $\Gamma_1(2)$ and $\Gamma_3(2)$ reckoned from Eqs. 21 and 22, respectively, are $\Gamma_1^{ga}(2) \cong 0 \cong \Gamma_3^{ga}(2)$. Thus, generic, but nondiscriminating, attractions may alter the local densities of species 1 and 3, but make no net contribution to the preferential interaction coefficients. Nonvanishing contributions of attractive interactions presumably arise from discriminatory exchange reactions, as indicated in the following section.

Osmolyte-water exchange reactions

Schellman (21–26) introduced the notion that the relevant reactions in solution were exchange reactions at sites or regions near the surface of the macromolecule (species 2). The objective here is to incorporate such exchanges within this formulation of the preferential interaction coefficients in terms of integrals over particular pair correlation functions.

Let us consider first the jth site, which may contain either a single osmolyte (species 3) or ν_j water molecules (species 1). For osmolytes that do not bear charged groups, it is expected that $\nu_j \cong \bar{V}_3/\bar{V}_1$, but that assumption need not be invoked at this point. The exchange reaction for this site is written as

$$\mathrm{M}\cdot\left(\mathrm{H_2O}\right)_{\nu_j} + \mathrm{L} \rightleftharpoons \mathrm{M}\cdot\mathrm{L} + \nu_j\,\mathrm{H_2O}, \quad (26)$$

where $\mathrm{M}\cdot\left(\mathrm{H_2O}\right)_{\nu_j}$ denotes a complex with ν_j bound waters on average in the jth site and M·L denotes a complex with a single bound osmolyte at the jth site. It is not required that ν_j be an integer. When the macromolecule M (species 2) is sufficiently dilute, the equilibrium constant for Eq. 26 is

$$K_j = \frac{[\mathrm{M}\cdot\mathrm{L}](a_1)^{\nu_j}}{[\mathrm{M}\cdot\left(\mathrm{H_2O}\right)_{\nu_j}]a_3}, \quad (27)$$

where $a_1 = a_w$ is the water activity for the mol fraction 1.0 standard state and $a_3 = a_L$ is the osmolyte activity for its hypothetical Henry's Law mol fraction 1.0 standard state, wherein each osmolyte experiences only the environment of its infinitely dilute solution (in water). The fraction of occupied (by osmolyte) j-sites is

$$f_j = \frac{[\mathrm{M}\cdot\mathrm{L}]}{[\mathrm{M}\cdot\left(\mathrm{H_2O}\right)_{\nu_j}] + [\mathrm{M}\cdot\mathrm{L}]} = \frac{K_j a_3 (a_1)^{-\nu_j}}{1 + K_j a_3 (a_1)^{-\nu_j}}. \quad (28)$$

The instantaneous density of the central atom of a 3-molecule in the jth site for any fixed configuration of the 2-molecule is a three-dimensional δ-function, $\delta(\mathbf{r} - \mathbf{r}_j)$, where $\mathbf{r}_j$ is the variable position of the central atom of the 3-molecule in the jth site in a coordinate frame originating on the central atom of the 2-molecule. When this density is averaged (with appropriate statistical weights) over the $\mathbf{r}_j$ for all allowed positions and configurations of the 3-molecule in the site and over all configurations of the 2-molecule, and

that result is in turn rotationally averaged about the chosen central atom of the 2-molecule, there results a distributed or smeared density function, $P_j^{(3)}(r) = \langle \delta(\mathbf{r} - \mathbf{r}_j) \rangle$, which depends only on the distance r from that central atom and should be peaked near the average distance $r = \langle |\mathbf{r}| \rangle = \langle |\mathbf{r}_j| \rangle$. The preceding averages are taken only over those configurations, wherein $\mathbf{r}_j$ lies within the somewhat arbitrarily defined boundaries of the jth exchange site for each configuration of species 2. This density function is still normalized, so $\int_0^R d^3\mathbf{r}\, P_j^{(3)}(r) = 1.0$.

The density function for those 1-molecules that occupy the jth site, when the 3-molecule is absent, is defined in the following way. First the center of a 3-molecule with a fixed configuration is placed at $\mathbf{r}_j$ in the jth site of a 2-molecule with a fixed configuration. The surrounding solution is assumed to consist entirely of 1-molecules. The density of all the η_1 1-molecules in the solution, $\rho_j^1(\mathbf{r}) = \sum_{\ell=1}^{\eta_1} \delta(\mathbf{r} - \mathbf{r}_\ell^1)$, is then averaged over all positions and configurations of those same 1-molecules. The resulting mean density of 1-molecules will practically vanish over an excluded volume, $V(\mathbf{r}_j, \xi, \zeta)$, that depends upon the particular $\mathbf{r}_j$ and fixed configurations ξ and ζ of the 2- and 3-molecules, respectively. The quantities ξ and ζ should be regarded as generalized vectors, or lists, of the coordinates of all the atoms in the 2-molecule and 3-molecule, respectively. Now, the 3-molecule is removed, but the configuration of the 2-molecule is held fixed at ξ. The 1-molecules are allowed to equilibrate with the 2-molecule in that same configuration ξ. The mean density of those 1-molecules, whose centers lie within the particular excluded volume, $V(\mathbf{r}_j, \xi, \zeta)$, is defined by $P_j^1(\mathbf{r}, \mathbf{r}_j, \xi, \zeta) \equiv \langle \sum_{\ell=1}^{\omega_{1j}} \delta(\mathbf{r} - \mathbf{r}_\ell^1) \rangle$, where the sum runs only over the ω_{1j} (variable) 1-molecules in each configuration, whose centers at $\mathbf{r}_\ell^1$ lie within $V(\mathbf{r}_j,\xi,\zeta)$, and the average is taken over all configurations of 1-molecules. This mean density of 1-molecules in $V(\mathbf{r}_j,\xi,\zeta)$ is further averaged over the $\mathbf{r}_j$ (within the jth site), ξ, and ζ by repeating this initial averaging process for various $\mathbf{r}_j$, ξ, and ζ, and then averaging the results over $\mathbf{r}_j$, ξ, and ζ. One obtains $P_j^{(1)}(\mathbf{r}) \equiv \langle \sum_{\ell=1}^{\omega_{1j}} \delta(\mathbf{r} - \mathbf{r}_\ell^1) \rangle_{\mathbf{r}_j,\xi,\zeta}$, where the subscripts denote the final averages over $\mathbf{r}_j$, ξ, and ζ. By definition, the average value of ω_{1j} for the jth site is $\langle \omega_{1j} \rangle = \int_0^R d^3\mathbf{r} \cdot P_j^{(1)}(\mathbf{r}) \equiv \nu_j$. When species 3 has no charged groups, so electrostriction effects are negligible, it is expected that the average number of 1-molecules that occupy an empty exchange site is $\nu_j \cong \bar{V}_3/\bar{V}1$. Finally, rotational averaging of $P_j^{(1)}(\mathbf{r})$ around the central atom of the 2-molecule yields $P_j^{(1)}(r)$, which depends only upon the scalar distance r from the central atom of the 2-molecule. The normalization integral remains unchanged, so $\int_0^R d^3\mathbf{r} P_j^{(1)}(r) = \nu_j$. It is expected that the final smeared density, $P_j^{(1)}(r)$, will normally be peaked near $r = \langle |\mathbf{r}_j| \rangle$ and exhibit a slightly greater width than $P_j^{(3)}(r)$, because the centers of multiple 1-molecules are involved.

In light of the preceding remarks, the contribution of the jth site to the mean density of 1-molecules in the vicinity of the 2-molecule is

$$c_1 g_{12}(r) = P_j^{(1)}(r)\left(1/\left(1 + K_j\, a_3(a_1)^{-\nu_j}\right)\right), \tag{29}$$

and to the mean density of 3-molecules is

$$c_3 g_{32}(r) = P_j^{(3)}(r)\left(K_j\, a_3(a_1)^{-\nu_j}/\left(1 + K_j\, a_3(a_1)^{-\nu_j}\right)\right). \tag{30}$$

The fraction of occupied sites, f_j, from Eq. 28 appears in Eq. 30 and $1 - f_j$ appears in Eq. 29. The contributions of the exchange reaction at the jth site to the preferential interaction coefficients follow from Eqs. 21 and 22 and the respective normalizations of $P_j^{(3)}(r)$ and $P_j^{(1)}(r)$:

$$\Gamma_{1j}^{er}(2) = \frac{\nu_j - (c_1/c_3)K_j a_3(a_1)^{-\nu_j}}{\left(1 + K_j a_3(a_1)^{-\nu_j}\right)} \tag{31}$$

$$\Gamma_{3j}^{er}(2) = -(c_3/c_1)\frac{\left(\nu_j - (c_1/c_3)K_j\, a_3(a_1)^{-\nu_j}\right)}{\left(1 + K_j\, a_3(a_1)^{-\nu_j}\right)}. \tag{32}$$

The total contributions of exchange reactions at all such sites are $\Gamma_1^{er}(2) = \sum_j \Gamma_{1j}^{er}(2)$ and $\Gamma_3^{er}(2) = \sum_j \Gamma_{3j}^{er}(2)$, where the sums run over all sites (j), which lie beyond the macromolecular void volume.

A model for preferential interactions

Let us now consider a model system that exhibits simultaneously all of the aforementioned repulsive exclusion forces, generic attractions, and discriminatory interactions that are responsible for exchange. For simplicity, we shall assume that the contributions of the various interactions to the total mean densities, $c_1 g_{12}(r)$ and $c_3 g_{32}(r)$, are additive. This important assumption is not generally valid and merits some discussion. For any given fixed configuration of species 2, the repulsive hard-core exclusion forces between 2 and either 1 or 3 affect the densities of species 1 and 3 in one region of space, whereas attractions or repulsions of longer range act on 1 and 3 in a different region (outside the hard core, but still inside the macromolecular domain of radius R). Hence, the effects of the short-range and longer-range interactions are largely spatially complementary, and would be expected to be nearly additive, even after configurational and rotational averaging of species 2. Nondiscriminatory generic attractions make no net contribution to $\Gamma_1(1)$ or $\Gamma_3(2)$ and are not considered further here. In regard to exchange reactions, some interaction between exchanging sites is generally expected. The neglect of such interactions renders this discussion oversimplified in an important regard, whenever c_3 is not small compared to c_1. Nevertheless, useful insights may emerge, and quantitatively useful accuracy may be obtained whenever $c_3\bar{V}_3 \ll c_1\bar{V}_1$.

Under this additivity assumption

$$\Gamma_1(2) = \Gamma_1^{ex}(2) + \sum_j \Gamma_{1j}^{er}(2) \tag{33}$$

$$\Gamma_3(2) = \Gamma_3^{ex}(2) + \sum_j \Gamma_{3j}^{er}(2). \tag{34}$$

Equations 24 and 25 give the $\Gamma_k^{ex}(2)$ in terms of c_1, c_3 and ΔV^{ac}, and Eqs. 31 and 32 express the $\Gamma_{kj}^{er}(2)$ in terms of ν_j, K_j, a_3, and a_1.

To examine the regime of small c_3 in more detail, additional approximations are invoked. First, it is assumed that $\bar{V}_1$ and $\bar{V}_3$ are independent of c_3 (which has units of molecules/m^3) up to a molar concentration of 1.0. To lowest order in c_3, that gives $c_1 = (1 - c_3\bar{V}_3)/\bar{V}_1 \cong 1/\bar{V}_1$ and $a_3 = \gamma_3 X_3 = \gamma_3(c_3/(c_1 + c_3)) \cong \gamma_3 c_3\bar{V}_1$, where γ_3 is the activity coefficient of species 3. With these approximations, and the exact relation, $c_3/c_1 = m_3/m_1$, Eqs. 33 and 34 become,

$$\Gamma_1(2) = \frac{\Delta V^{ac}}{\bar{V}_1} + \sum_j \frac{\left(\nu_j - K_j\gamma_3(a_1)^{-\nu_j}\right)}{\left(1 + K_j\gamma_3 c_3\bar{V}_1(a_1)^{-\nu_j}\right)} \tag{35}$$

$$\Gamma_3(2) = -(m_3/m_1)\left[\Delta V^{ac}/\bar{V}_1 + \sum_j \frac{\left(\nu_j - K_j\gamma_3(a_1)^{-\nu_j}\right)}{\left(1 + K_j\gamma_3 c_3\bar{V}_1(a_1)^{-\nu_j}\right)}\right]. \tag{36}$$

Generalization of the exchange model

We imagine that a lattice of exchanging sites (or cells) with initial volume $\bar{V}_3$ fills the entire osmolyte-accessible region of the macromolecular domain of radius $R \geq R_{min}$. An osmolyte is regarded as bound to a particular site, when its central atom lies within that cell. The initial cell volume is taken as $V^0 = \bar{V}_3$, so the cell size matches the partial molecular volume of the osmolyte. Thus, if all of the initial sites were filled, species 3 would just fill the entire volume. The average number of 1-molecules that occupy a cell, when the osmolyte is absent, is assumed to be $\nu = \bar{V}_3/\bar{V}_1$, which is exact far from the macromolecular surface, and is almost certainly a fairly good approximation even near the macromolecular surface, except when electrostriction effects are large. Thus, the species 1 would just fill the lattice volume in the absence of species 3. While holding the overall lattice volume constant, one could now choose a smaller uniform cell size for the lattice of exchange sites, namely $\hat{V}_m = (1/m)\bar{V}_3$, where $m \geq 2$ is an integer, provided that the contributions of each site to $\Gamma_1(2)$ are reduced by the same factor, $1/m$, and that the j-sums in Eqs. 35 and 36 are extended from the original L sites of volume $\bar{V}_3$ to the mL smaller sites of volume $\hat{V}_m$. For some sufficiently large value of m, when $\hat{V}_m = \bar{V}_3/m \ll \bar{V}_1, \Gamma_1(2)$ should become entirely independent of m or $\hat{V}_m$. A lattice cell size in that range is adopted here. The center of the jth cell is taken at position $\mathbf{q}(j)$, and its exchange constant, $K_{\mathbf{q}(j)}$, may vary from one cell to the next in a limited way, so as to create a gradient of the $K_{\mathbf{q}(j)}$ along any reasonably smooth path in the discrete $\mathbf{q}(j)$ space. Both $\nu = \bar{V}_3/\bar{V}_1$ and the exchange constant, $K_{\mathbf{q}(j)}$, for each smaller cell of volume, $\hat{V}_m = \bar{V}_3/m$, are taken as the values typical of a site with the initial volume, $\hat{V}_0 = \bar{V}_3$, whose center lies within that smaller cell, with the understanding that $m - 1$ adjacent sites are closed, whenever an osmolyte binds to the smaller cell in question. In this way, any region of volume $\bar{V}_3$ will bind one and only one osmolyte in approximately the same way as a function of a_3 or a_1, regardless of the number of cells into which is it subdivided, and the maximum densities of species 1 and 3 will remain unchanged. The smaller lattice cell volumes are employed simply to represent the spatial variation of the exchange constants at higher resolution than is afforded by cells of volume $\bar{V}_3$. By suitable adjustment of the K_j associated with the various sites in the lattice, it is possible to create any conceivable mean densities $c_1 g_{12}(r)$ and $c_3 g_{32}(r)$ at a level of resolution set by the lattice cell size, subject to the implicit volume conservation rule invoked here (i.e., $\nu = \bar{V}_3/\bar{V}_1$). The approximate validity of this model is limited to the regime of small volume fraction of species 3, so that events in any one region of volume $\bar{V}_3$ do not affect events in neighboring regions of the same size. The large anticooperativity associated with the closure of $m - 1$ binding sites surrounding a given site, when it becomes occupied, generally has a very strong influence on the system, except when the volume fraction of species 3 is small. In that special case, for a cell volume $\hat{V}_m$, Eqs. 35 and 36 become

$$\Gamma_1(2) = \frac{\Delta V^{ac}}{\bar{V}_1} + (1/m)\sum_{j=1}^{mL} \frac{\left(\nu - K_j\gamma_3(a_1)^{-\nu}\right)}{\left(1 + K_j\gamma_3 c_3\bar{V}_1(a_1)^{-\nu}\right)} \tag{37}$$

$$\Gamma_3(2) = -(m_3/m_1)\left(\frac{\Delta V^{ac}}{\bar{V}_1} + (1/m)\sum_{j=1}^{mL} \frac{\left(\nu - K_j\gamma_3(a_1)^{-\nu}\right)}{\left(1 + K_j\gamma_3 c_3\bar{V}_1(a_1)^{-\nu}\right)}\right). \tag{38}$$

Neutral binding

When the standard free energy change for an exchange reaction vanishes, $K_j = 1.0$. First, let us consider the limit of small c_3, where, $\gamma_3 \to 1.0$ and $a_1 \to 1.0$, so the numerator of each term in the j-sums of Eqs. 37 and 38 becomes $\nu - 1.0 \cong (\bar{V}_3/\bar{V}_1) - 1.0$. For typical small neutral osmolytes, excluding molecules the size of trehalose and sucrose, one expects that $\nu \cong \bar{V}_3/\bar{V}_1 = 3 - 5$. Note that, if $\nu = 1.0$, as was assumed in early treatments of exchange by Schellman (21–26), then $\nu - 1.0 = 0$, and the entire exchange contribution of the jth site would vanish. Although the condition, $K_j = 1.0$, is the point of neutrality in terms of vanishing standard state free energy change, it is not generally the point of neutrality in regard to purely random binding in the neighborhood of a 2-molecule, because ν 1-molecules are released for every 3-molecule bound. The point of neutrality in regard to random binding of 1- and 3-molecules at the jth site, when $\gamma_3 = 1.0$ and $a_1 = 1.0$, is clearly $K_j = \nu$.

In general, sites that lie out in the bulk solution sufficiently far from the surface of the 2-molecule can make no net

contribution to $\Gamma_1(2)$ or $\Gamma_3(2)$, so for such sites it is absolutely required that $K_j = \nu/(\gamma_3(a_1)^{-\nu})$, which can be taken as the general condition for neutrality of any site in regard to random binding of 1's and 3's. Smaller values of K_j yield a positive contribution of the jth site to $\Gamma_1(2)$.

We consider next the limit, wherein $K_j\gamma_3 c_3\bar{V}_1(a_1)^{-\nu} \ll 1.0$, so the second terms in the denominators of Eqs. 37 and 38 can be ignored. The product, $c_3\bar{V}_1$ is unitless, and has the same value in any units, so one can take c_3 in mol/L and $\bar{V}_1$ in L/mol. For small neutral osmolytes, one typically has $\gamma_3 \simeq 1.0$ and $a_1 \simeq 1.0$ up to $c_3 = 1.0$ M. Thus, for $c_3 \lesssim 1.0$ M, the inequality, $K_j\gamma_3 c_3\bar{V}_1 a_1^{-\nu} \ll 1.0$, will be satisfied, when $K_j \ll 1/\ (c_3\bar{V}_1) = (55.6)/c_3$. Hence, K_j could be as large as 10, and still satisfy this inequality for $c_3 = 1.0$ M. In other words, K could be up to 2–3 times greater than the neutral random binding value, $\nu = 3 - 5$, and still the second terms in the denominators of Eqs. 37 and 38 would be negligibly small for all c_3 up to 1.0 M. In this limit, Eqs. 37 and 38 can be written as

$$\Gamma_1(2) = X + (1/m)\sum_{j=1}^{mL}\left(\nu - K_j\gamma_3(a_1)^{-\nu}\right) \cong X + (1/m)\sum_{j=1}^{mL}(\nu - K_j) \quad (39)$$

$$\Gamma_3(2) = -\left(\frac{m_3}{m_1}\right)\left(X + \left(\frac{1}{m}\right)\sum_{j=1}^{mL}\left(\nu - K_j\gamma_3(a_1)^{-\nu}\right)\right) \cong -\left(\frac{m_3}{m_1}\right)\left(X + \left(\frac{1}{m}\right)\sum_{j=1}^{mL}(\nu - K_j)\right), \quad (40)$$

where $X \equiv \Delta V^{ac}/\bar{V}_1$ is the difference between the volumes accessible to 1 and 3 in units of $\bar{V}_1$. As noted above, $\Delta V^{ac} = V_3^{ex} - V_1^{ex}$ is generally positive, because the osmolyte generally exceeds the water in size, so X should also be generally positive. The j-sum (of binding terms) can in principle take either sign, depending upon the magnitude of K_j.

Variation of $\Gamma_1(2)$ and $\Gamma_3(2)$ with c_3

Equations 39 and 40 predict that $\Gamma_1(2)$ should be nearly constant independent of c_3, and that $\Gamma_3(2)$ should vary nearly in proportion to m_3 with constant slope, up to $c_3 = 1.0$ M. In fact, for seven different osmolytes interacting with BSA, it was found that $\Gamma^m_{\mu_1,\mu_3}$, hence also $\Gamma_3(2)$, varied in proportion to m_3 with a constant negative slope up to $m_3 = 1.0$ molal (5,29,30; J. G. Cannon, personal communication, 2005). The negative slope implies that the total j-sum is either positive or not so negative that it overwhelms the positive value of X. The constant slope indicates that the second terms in the denominators of Eqs. 37 and 38 are negligible up to $m_3 = 1.0$ m, which in turn implies that (in the case of BSA) most of the contribution from the j-sum must arise from sites with K_j-values that do not exceed by more than approximately threefold the random binding value, $K_j^{rand} = \nu/(\gamma_3(a_1)^{-\nu}) \simeq \nu \simeq \bar{V}_3/\bar{V}_1$.

Because X derives from a shell volume with a thickness equal to the difference in radius between the osmolyte and water, it is expected to vary nearly in proportion to the area in the case of macromolecules with homologous surfaces. Likewise, the j-sum concerns primarily just the contact layer and a few additional layers of osmolyte or water, so that it too is expected to vary nearly in proportion to the area in the case of macromolecules with homologous surfaces. Courtenay et al. (6) noted that numerous globular proteins exhibit similar values of the ratio, $-\Gamma_3(2)/(m_3A_s)$, where A_s is the water accessible area.

Analysis of $\Gamma_1(2)$-values for BSA

Experimental values of $\Gamma_1(2)$ for different osmolytes interacting with BSA are obtained from the corresponding $\Gamma^m_{\mu_1,\mu_3}$ determined by the Record group (5,29,30; and J. G. Cannon, personal communication, 2005) via the relations Eqs. 22, 23, and 40, which are combined to give

$$\Gamma_1(2) = X + S = -(55.6)\left(\Gamma^m_{\mu_1,\mu_3}/m_3\right), \quad (41)$$

where $X = \Delta V^{ac}/\bar{V}_1$, and $S = (1/m)\sum_{j=1}^{mL}(\nu - K_j)$. By combining the measured value of $\Gamma^m_{\mu_1,\mu_3}$ with an estimate of X obtained from the protein structure, it is possible to obtain an experimental estimate of S.

Although no crystal structure has been reported for BSA, it is assumed to be satisfactorily modeled by human serum albumin. BSA has 607 amino acid residues and HSA has 609, which are 76% homologous with the BSA sequence. Only 578 of the 609 residues of HSA are resolved in the crystal structure (31). Here the molecular volume reckoned for the crystal structure is simply scaled by 609/578 = 1.054 to estimate the corresponding volume for the full HSA (or BSA). However, the ΔV^{ac} reckoned for the crystal structure corresponds to the volume of a relatively thin shell of a given thickness about the macromolecule, so it is scaled by the factor $(1.054)^{2/3} = 1.035$. The osmolyte and water accessible areas are also scaled by 1.035 to estimate the corresponding areas for HSA (or BSA). The various volumes and areas are reckoned using the program MSROLL (32). The crystal structure contains an HSA dimer and seven water molecules. The nonhydrogen atoms of both the water molecules and the second dimer and their coordinates are deleted from the list of atomic coordinates, leaving just the atoms and coordinates of the 578 resolved residues of the first monomer. The program assigns a van der Waals radius to each atom or group of HSA. An effective radius, $R_i = (1/2)(\bar{V}_i)^{1/3}$, is assigned to water ($i = 1$) and to each osmolyte ($i = 3$).

The molecular displacement volume ($\bar{V}_{dis}$) of HSA in water is determined by rolling a water-size sphere of radius $R_1 = 1.48$ Å around its exterior van der Waals surface in

each of a series of closely spaced parallel planes. The program reckons the volume inside the continuous surface formed by the contact surface(s) of the sphere with the van der Waals surface of the protein plus the so-called reentrant surface(s) that bridge the gaps in the contact surface by following the interior surface of the bridging sphere. We obtain $\bar{V}_{\rm dis} = 76,762\text{Å}^3$ for HSA, which is then scaled by 1.054 to estimate $\bar{V}_{\rm dis} = 80,879\text{Å}^3$ for BSA. This molecular displacement volume cannot be occupied by any part of the 1.48 Å sphere, and for numerous globular proteins is found to lie within 1–2% of the partial molecular volume $\bar{V}_2$ (S. Aragon, unpublished data). Courtenay et al. (5) report $\bar{V}_2 = 81,651\text{Å}^3$ for BSA, which differs by ~1% from the $\bar{V}_{\rm dis}$ calculated above. This agreement provides an important check on the structure and computational protocols used, but does not pertain directly to the preferential interaction coefficients.

Next we obtain the volumes excluded by HSA to the centers of water-size or osmolyte-size spheres. This is the volume inside a surface that is traced out by the center of the osmolyte sphere or water sphere, as it rolls over the surface of the protein, and represents the void volume in $g_{12}^{\rm ex}(r)$ or $g_{23}^{\rm ex}(r)$. Although MSROLL does not calculate the volume inside the excluded-center surface directly, that volume can be reckoned by first inflating the atomic van der Waals radii by R_1 or R_3 and using a probe sphere of zero radius. The resulting contact plus (vanishing) reentrant surface is the same surface traced out by the center of a sphere of radius R_1 or R_3, as it rolls over the uninflated van der Waals surface, and its interior volume is calculated by the program. The difference between the volume excluded to an osmolyte and that excluded to a water center is just $\Delta V^{\rm ac} = V_3^{\rm ex} - V_1^{\rm ex}$ for that osmolyte, as illustrated in Fig. 2. After scaling $\Delta V^{\rm ac}$ by 1.035, it is divided by the molecular volume of water, $\bar{V}_1 = 29.9\text{Å}^3$, to obtain $X = \Delta V^{\rm ac}/\bar{V}_1$. Then the exchange contribution, $S = \Gamma_1(2) - X$, is finally evaluated. The results for X and S are presented for seven different osmolytes interacting with BSA in Table 1. In every case, X exceeds the magnitude of S. Thus, in the case of BSA, the largest contribution to $\Gamma_1(2)$ is simply a geometrical consequence of the fact that the osmolytes are substantially larger than water

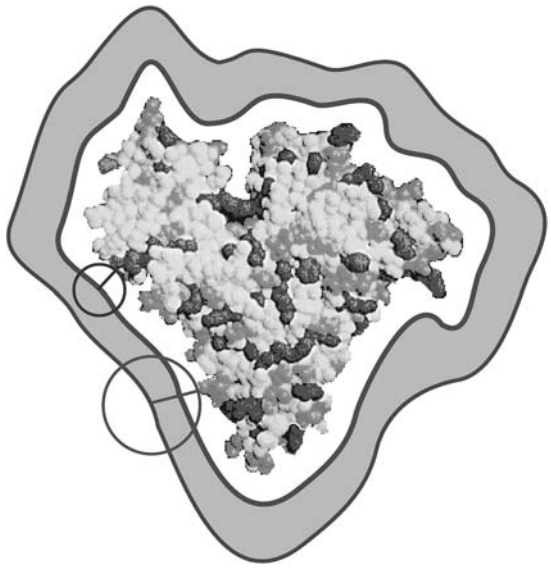

FIGURE 2 Schematic illustration (not to scale) of the difference between the volume excluded to osmolyte (*large sphere*) and to water (*small sphere*) by human serum albumin. The desired volume is that of the shaded shell between the surfaces traced out by the center of an osmolyte-size sphere and that traced out by a water-size sphere, as those spheres are rolled over the surface of the protein. The sizes of the water and osmolyte relative to that of BSA are exaggerated for illustrative purposes.

and therefore have a larger effective sphere radius. Four of the osmolytes, urea, glycerol, proline, and trehalose, exhibit a negative value of S, which indicates that $K_{\rm j}$ is on average greater than its neutral value, $\nu \cong \bar{V}_3/\bar{V}_1$, and implies a greater than random preference of the osmolyte for exchanging sites within the macromolecular domain. The remaining three osmolytes, trimethylamine N-oxide (TMAO), K^+glutamate, and betaine glycine, exhibit a positive value of S, which indicates that $K_{\rm j}$ is on average less than its neutral value, and implies a lower than random preference of these osmolytes for the exchanging sites.

In previous work in this field, it was commonly assumed that $\Gamma_1(2)$ (or $\Gamma_3(2)$) is determined primarily by exchange sites within the first surface-contact layer. We now also assume that the exchanging sites are confined to the first

TABLE 1 Excluded volume (*X*) and exchange reaction (*S*) contributions to $\Gamma_1(2)$ and $\Gamma_3(2)$ for osmolytes interacting with BSA

	$\bar{V}_3/\bar{V}_1$†	$-(55.6)(\Gamma_3(2)/m_3)$*	X*	S	$L_{\rm s}$	$\langle K\rangle/(\bar{V}_3/\bar{V}_1)$
Urea	2.45	217	560	−343	1364	1.10
Glycerol	3.92	250	777	−527	932	1.15
TMAO	4.01	1389	888	+501	911	0.86
Proline	4.60	778	981	−203	820	1.05
K^+glu^-	5.03	1111	1048	+63	755	0.95
Betaine	5.43	1283	1107	+176	712	0.96
Trehalose	11.61	1167	1762	−595	392	1.12

*The $\Gamma_3(2)/m_3 = \Gamma^{\rm m}_{\mu_1,\mu_2}/m_3 \cong \Gamma^{\rm m}_{\mu_3}/m_3$ and $\bar{V}_3$ values for the different osmolytes were reported by Courtenay et al. (5), Felitsky et al. (29), Hong et al. (30), and J. Cannon (2005, personal communication). The X-values were reckoned using the MSROLL program with effective osmolyte radii (R): urea (2.09 Å), glycerol (2.45 Å), TMAO (2.47 Å), proline (2.58 Å), K^+ glutamate (2.66 Å), betaine glycine (2.73 Å), and trehalose (3.52 Å), as described in the text.
†$\bar{V}_1 = 2.99 \times 10^{-23}\,\text{cm}^3/\text{molecule}$.

surface-contact layer. We take the number (L_s) of surface-contiguous sites for species 3 to be the accessible area traced out by the center of a sphere of radius R_3 rolled over the van der Waals surface of HSA, scaled by the factor 1.035, and divided by the area, $(\bar{V}_3)^{2/3}$, of a single site of volume $\bar{V}_3$. Values of L_s for the different osmolytes are also included in Table 1. For these sites, the cell volume is $\hat{V}_m = \bar{V}_3$, so $m = 1.0$, and $mL = L = L_s$. In this case, $S = L_s(\nu - \langle K_j \rangle)$, where $\langle K_j \rangle = (1/L_s)\sum_j K_j$ is the average exchange constant for the surface-contiguous sites of species 3. From the values of S, L_s, and $\nu = \bar{V}_3/\bar{V}_1$, we estimate $\langle K_j \rangle/(\bar{V}_3/\bar{V}_1)$, which is the ratio of the average exchange constant to its neutral (or random) value. These values are listed in the final column of Table 1. These $\langle K_j \rangle/(\bar{V}_3/\bar{V}_1)$ ratios are all remarkably close to 1.0, with a maximum deviation of $<\sim 0.15$. By this criterion the average interactions of these osmolytes with the BSA surface are all surprisingly similar.

Comparisons with prior work

An advantage of this formulation for $\Gamma_1(2)$ (or $\Gamma_3(2)$) in terms of the pair correlation functions (Eqs. 21 and 22) in comparison to the thermodynamic formulation in terms of numbers of molecules in the macromolecular domain (Eq. 3) is that the excluded volume contribution is unambiguously given by $\Delta V^{ac} = \int_0^R d\mathbf{r}\big(g_{12}^{ex}(r) - g_{32}^{ex}(r)\big) = V_3^{ex} - V_1^{ex}$, which is the volume of the shell in Fig. 2.

Shimizu (13) and Shimizu and Smith (14) employed a single ''excluded'' volume, V_E, that is independent of the water or osmolyte, and is essentially the macromolecular displacement volume, $\mathbf{V}_{dis}$, that is excluded to any part of a water or osmolyte molecule. Those authors approximated the excluded volume contribution to the numbers of water and osmolyte molecules in the macromolecular domain by $-c_1V_E$ and $-c_3V_E$, respectively. Their use of the same value, V_E, for both V_1^{ex} and V_3^{ex} leads to complete cancellation of the excluded volume contribution to $\Gamma_1(2)$, which is incorrect. However, the primary focus of their work was to determine G_{12} and G_{23} separately, and to interpret the excess quantities, c_1G_{12} and c_3G_{32}. For that purpose, the use of V_E in place of V_1^{ex} or V_3^{ex} may be a reasonable approximation.

Schellman (18) evaluated the cross-second virial coefficient, B_{23}, of the osmotic pressure for components 2 and 3 in terms of an integral over the potential of mean force between the osmolyte and macromolecule (averaged over all positions and numbers of the water molecules). The excluded volume contribution to that integral is the volume excluded by the macromolecule to the osmolyte centers, and is reckoned by a protocol identical to that employed here, except that the corresponding excluded volume for the water centers was not subtracted from that for the osmolyte. The exchange reaction of Schellman is the replacement of a single water molecule at a site by a single osmolyte at the same site, and its contribution to B_{23} is expressed in terms of the exchange constant, $\hat{K}_j = (f_3/f_1)/(\phi_3/\phi_1)$, where f_3 and f_1 are the fractional occupations of the jth site by osmolyte and water, respectively, and ϕ_1 and ϕ_3 are the volume fractions of osmolyte and water, respectively, in the bulk solution. The exchange reaction contribution was then summed over all sites. The primary objective was to express the change, ΔB_{23}, in B_{23} upon unfolding of the protein (species 2) in terms of the change in volume excluded to osmolyte centers and the change in the exchange reaction sum, and to assess their relative magnitudes. It was suggested that c_3B_{23} is ''the total excess of cosolvent molecules in the neighborhood of the protein. Its identity with the preferential interaction coefficient (in the absence of nonideality) is thus completely explained at the molecular level''. In fact, Schellman's c_3B_{23} is not identical to the $\Gamma_3(2)$ obtained here, or to $\Gamma^m_{\mu_1,\mu_3}$ or to $\Gamma^m_{\mu_3}$, because c_3B_{23} clearly lacks the contribution, $-c_3\int_0^R d\mathbf{r}\big(g_{12}^{ex}(r) - 1\big)$, that is explicit in Eqs. 20 and 22. Provided that $\bar{V}_1$ and $\bar{V}_3$ are independent of c_2 and c_3, this $\Gamma_3(2)$ is identical to $\Gamma^m_{\mu_1,\mu_3} = (\partial m_3/\partial m_2)_{T,\mu_1,\mu_3}$ and very close to $\Gamma^m_{\mu_3} = (\partial m_3/\partial m_2)_{T,P,\mu_3}$. However, Schellman's osmotic pressure calculation is carried out (implicitly) for constant μ_1, so that c_3B_{23} is conceivably identical to $\Gamma^m_{\mu_1} = (\partial m_3/\partial m_2)_{T,P,\mu_1}$, although a rigorous proof of that conjecture is lacking. As noted previously (5,19,20), $\Gamma^m_{\mu_1}$ differs considerably from $\Gamma^m_{\mu_1,\mu_3}$ and also from $\Gamma^m_{\mu_3}$ under the usual conditions of moderately low osmolyte concentration. In any case, Schellman's conclusion that urea occupation of the surface-contiguous layer exceeds the random value by modest amounts up to 15% for the five proteins analyzed (ribonuclease T, ribonuclease A, hen egg white lysozyme, staphylococcus nuclease, and T4 lysozyme) is consistent with the corresponding result of this analysis for urea interacting with BSA, where urea occupation exceeds the random value by 10% (c.f. Table 1).

Record and co-workers applied the thermodynamic two-domain model to analyze their preferential interaction coefficients (9). They demonstrated that $\Gamma_3(2)$ was proportional to m_3 and presented evidence that $\Gamma_3(2)$ is correlated with the osmolyte accessible surface area (5,29,30). They introduced a local domain-bulk domain partition coefficient, instead of an osmolyte-water exchange constant, and initially proposed that the volume of the local domain was that of the surface-contiguous water molecules (5). Because the centers of the larger osmolytes lie entirely outside such a domain, the osmolyte number within such a local domain is either vanishing or not well defined without additional assumptions, so the local-bulk partition coefficient in such cases is not well defined. In general, the local-bulk partition coefficient depends upon the total volume of the local domain, much as the average exchange constant depends upon the total number of exchange sites of volume $\sim\bar{V}_3$. In their studies to date, Record and co-workers have made no attempt to treat the separate contributions of excluded volume on one hand and the exchange reactions, or osmolyte partitioning into the accessible local domain volume, on the other. They have investigated the positive

correlation between $\Gamma_3(2)$ for betaine glycine and the fraction of the accessible macromolecular surface area that is associated with the anionic oxygen atoms of carboxylate or phosphate groups, and have proposed that the local domain includes two layers of more strongly bound water over those parts of the surface (29,30). That phenomenon can be analyzed in greater detail using this formulation and model, but that lies outside the scope of this article.

Strong water binding/weak osmolyte binding sites

It is conceivable that some macromolecules might exhibit a number of sites, wherein the water is rather more tightly bound and difficult to displace by common osmolytes. For such sites, K_j would lie far below the random binding value, $K_j^{rand} = \nu/(\gamma_3(a_1)^{-\nu})$, which declines with decreasing a_1, up to moderately high values of c_3 (or down to correspondingly low values of a_1). For such sites, the $K_j\gamma_3 c_3 \bar{V}_1(a_1)^{-\nu}$ terms in the denominators of the j-sum (in Eqs. 35 and 36) are negligibly small compared to 1.0 up to rather high values of c_3 (>1.0 M). Because the $K_j\gamma_3(a_1)^{-\nu}$ terms in the numerators are also negligibly small compared to their ν-terms, the ν-terms collectively provide the main contribution of such sites to the j-sum. Osmolytes of similar size should exhibit similar ν-values for any given site. Hence, for osmolytes of similar size, a dependence on osmolyte structure can enter the j-sum only through the K_j-terms. However, if those terms are negligibly small compared to ν, as is the case for strong water binding sites, then the contribution of such terms to the j-sum, to $\Gamma_1(2)$, and to $\Gamma_3(2)$, will be negligibly small. Thus, even though the K_j doubtless vary with osmolyte structure for osmolytes of comparable size, that variation will be negligible compared to the $\sum\nu$ and to the total contributions of the strong water-binding sites to the j-sum, or to $\Gamma_1(2)$, or to $\Gamma_3(2)$, provided that neither c_3 is too large, nor a_1 too small.

Upon increasing c_3, a_1 decreases, $(a_1)^{-\nu}$ increases, and the $K_j\gamma_3 c_3\bar{V}_1(a_1)^{-\nu}$ term in the denominator of each term in the j-sum increases toward (and eventually beyond) 1.0. In addition, the $K_j\gamma_3(a_1)^{-\nu}$ term in the numerator of each term in the j-sum increases toward its ν or beyond. Both effects act to decrease $\Gamma_1(2)$ and also to shift the negative slope of $\Gamma_3(2)$ with respect to c_3 toward less negative, or more positive, values. Such effects depend upon the K_j, and hence upon the chemical structure of the osmolyte. Thus, the contribution of strong water-binding sites to $\Gamma_1(2)$ and $\Gamma_3(2)$ is expected to be independent of osmolyte structure only for osmolytes of comparable size in the regime of strong exclusion of 3-molecules, when c_3 is not too large.

INTERPRETATION OF $\Delta\Gamma_1(2)$ FOR HYDRATION COUPLED REACTIONS

In experimental studies of the effects of osmolytes on equilibrium constants (K) for macromolecular reactions, it is typically found that ln K varies in proportion to ln a_1, when c_3 is not too large. Thus, $\Delta\Gamma_1(2)$ (c.f. Eq. 1) typically remains practically constant over the range of c_3 examined. This is in accord with Eqs. 19–21, 35, 37, 39, and 41, when c_3 is moderately small.

For simplicity we consider a conformational change of the macromolecule, $2_A \rightleftharpoons 2_B$, where the subscripts ''A'' and ''B'' denote different conformations of species 2. When $c_3 \lesssim 1.0$ M, Eqs. 39–41 apply to 2_A and 2_B separately. We adopt identical space-filling lattices of exchange sites centered at $\mathbf{q}(j)$, $j = 1,2,\ldots$, in the regions surrounding the hard-cores of species 2_A and 2_B. It is assumed that $\nu = \bar{V}_3/\bar{V}_1$ is the same for a given osmolyte in the lattices around both A and B. Then, Eq. 41 gives

$$\Delta\Gamma_1(2) \equiv \Gamma_1(2_B) - \Gamma_1(2_A) = X_B - X_A + (S_B - S_A), \quad (42)$$

where X_B and S_B are the excluded volume and exchange reaction contributions for species 2_B, and X_A and S_A are the corresponding quantities for species 2_A. Equation 42 applies only in the small c_3 limit, where $\gamma_3(a_1)^{-\nu} \cong 1.0$, so $\Delta\Gamma_1(2)$ remains nearly constant with increasing c_3, as found experimentally. We now divide the surfaces of both A and B into two regions, namely the ''passive'' regions that are the same in the B conformation as in the A conformation, and the ''active'' regions that differ between the two conformers. An active region may consist of surface that is either exposed or buried during the A → B transition, so that it is present in one species, but not in the other. It may also contain surface whose topography is reconfigured during the A→B transition, so as to alter the number and/or exchange constants of the exchanging sites associated with that part of the surface. A particular example would be the widening or narrowing of the angle of a cleft during the A → B transition. Under the assumption that the main contributions to X_A, X_B, S_A, and S_B involve only regions of the solution that are reasonably proximal to the surface of A or B, the contribution of the passive parts of the surface to $\Delta\Gamma_1(2)$ is expected to cancel. Furthermore, the sums over exchanging sites associated with the active parts of the surface can be divided into the terms arising from strong water-binding/weak osmolyte-binding sites, for which K_j can be neglected in favor of ν, and those terms arising from more neutral water-binding sites. Then,

$$\Delta\Gamma_1(2) = \left[(X_B - X_A) + \left(L_s^B - L_s^A\right)\nu + L_n^B\left(\nu - \langle K_n^B\rangle\right) - L_n^A\left(\nu - \langle K_n^A\rangle\right)\right]_a, \quad (43)$$

where the subscript ''a'' on the square bracket indicates that all quantities therein pertain only to the active parts of the surface, L_s^B and L_n^B are the numbers of strong (s) and more neutral (n) water-binding sites of volume $\hat{V}_m = \bar{V}_3$, respectively, associated with the active part of the surface of conformer 2_B, L_s^A, and L_n^A are the corresponding quantities for conformer 2_A, and $\langle K_n^B\rangle$ and $\langle K_n^A\rangle$ are the average equilibrium constants of the more neutral water-binding sites

associated with the active surfaces of 2_B and 2_A, respectively. Weak water-binding/strong osmolyte-binding sites are not considered here, because, if significant, they would cause $\Delta\Gamma_1(2)$ to vary with c_3, which is not observed. The L_n^B and L_n^A need count only those more neutral sites sufficiently near the surface that $\left(\nu - K_{nj}^B\right)$ and $\left(\nu - K_{nj}^A\right)$ differ significantly from zero.

When is $\Delta\Gamma_1(2)$ independent of the chemical structure of the osmolyte?

What are the conditions under which $\Delta\Gamma_1(2)$ is or is not independent of the chemical structure of the osmolyte for osmolytes of comparable size? The $X_B - X_A$ term in Eq. 43 depends on the osmolyte's size, but not on its chemical structure. The $\left(L_s^B - L_s^A\right)\nu$ term also varies with the size, but not the chemical structure, of all those osmolytes that have a common set of strong water-binding sites around conformer 2_B and another common set around conformer 2_A. Nonionic, non-zwitterionic osmolytes of similar size, such as ethylene glycol and acetamide, should have the same set of strong water-binding/weak osmolyte-binding sites around a given conformer. In contrast, osmolytes of similar size, but different ionic character, such as glycerol and TMAO, may well have different sets of such sites around conformers 2_A and 2_B, and would likely exhibit different $\left(L_s^B - L_s^A\right)\nu$ terms. In general, the $\langle K_n^B\rangle$ and $\langle K_n^A\rangle$ vary with both the osmolyte's size and its chemical structure, especially for surface-contiguous sites, and the $L_n^B\left(\nu - \langle K_n^B\rangle\right) - L_n^A\ \left(\nu - \langle K_n^A\rangle\right)$ term is expected to vary from one osmolyte to another, except in the event of accidental cancellation, or in the event that active regions of the surface consist only of strong water-binding sites, in which case these terms vanish. Except in special cases, discussed below, one expects to find that $\Delta\Gamma_1(2)$ varies with both osmolyte chemical structure and size. Such variations of $\Delta\Gamma_1(2)$ with the osmolyte's chemical structure and size have often been reported (2,6,33–36). For many of these processes the magnitude of $\Delta\Gamma_1(2)$ was found to increase with osmolyte size, as would be expected, if the $2_A \rightarrow 2_B$ transition involved a significant change in macromolecular surface area, and if also X_B and X_A exceed $|S_B|$ and $|S_A|$, respectively, as is the case for BSA.

One scenario, wherein $\Delta\Gamma_1(2)$ is independent of the osmolyte's chemical structure, occurs when the active regions of the surface almost completely enclose pockets or channels that cannot be penetrated by any osmolyte exceeding a certain size. In such a case, there are no exchanging sites of any kind within the pocket or channel, and $\Delta\Gamma_1(2) = X_B - X_A$ arises entirely from the excluded volume contribution. Because all osmolytes exceeding a certain size are completely excluded from the pocket or channel in 2_B, one has $[X_B]_a = \left[V_3^{ex} - V_1^{ex}\right]_{Ba}/\bar{V}_1$, which is just the volume of the pocket or channel in 2_B that is accessible to water centers, but excluded to osmolyte centers, divided by $\bar{V}_1$, which is approximately the number of nondisplaceable water molecules within the pocket or channel of 2_B. A similar relation applies to $[X_A]_a$. In this scenario, $\Delta\Gamma_1(2)$ can be regarded as the change in the number of bound water molecules, or more precisely in the number of water molecules that cannot be displaced by any osmolyte above a certain size. This nondisplaceable water scenario is typically proposed to rationalize the observation that $\Delta\Gamma_1(2)$ (or $-\Delta\Gamma_3(2)/m_3$) is independent of the chemical structure and size of the osmolyte over a significant range of osmolyte kinds and sizes (6,7,38–39). However, this scenario seems unlikely to account for the observations of Spink and Chaires (33), who studied the effects of ethylene glycol, acetamide, glycerol, and sucrose on DNA melting. The inverse melting temperature, T_m^{-1}, varied linearly with ln a_1. The slope, $\partial\left(T_m^{-1}\right)/\partial \ln a_1$, was similar for ethylene glycol, acetamide, and glycerol, corresponding to $\Delta\Gamma_1(2) = -4$ water molecules per basepair, but was substantially greater for the much larger sucrose, as expected, because it is unable to penetrate the minor groove and perhaps other nooks, as well. However, crystal structures of duplex B DNAs appear to provide no spaces in which to sequester water molecules so that they cannot in principle be displaced by osmolytes as small as ethylene glycol or acetamide, or even by glycerol. Recent and ongoing work in our lab (40) suggests that ethylene glycol and acetamide at 37°C induce a transition to an alternative duplex conformation within the B-family, and may do so in a similar, though not completely identical, manner as a function of ln a_1. These findings suggest that another mechanism may exist by which osmolytes may exert effects that depend only weakly on their chemical structure or even size over a limited range.

Water in small confined spaces, whether accessible to osmolytes or not, is likely to be at least partially ordered. When two atomically smooth cylindrical mica surfaces in aqueous media are pressed together with perpendicular orientation, the force at first rises smoothly with decreasing distance, and then for distances $\leq 18\mathring{A}$ exhibits five to six oscillations, as successive water layers are squeezed out (41). This suggests that water molecules in clefts and grooves of width $\leq 18\mathring{A}$ are likely to be at least partially ordered. In fact, high resolution x-ray diffraction studies at -110°C revealed some four layers of partially ordered water in and above the minor groove of a B-DNA (42,43). The absence of crystallographic evidence for more widespread occurrence of more than two layers of ordered water molecules in clefts, grooves, and pits might arise from the coexistence of two or a few different partially ordered water structures within the same cavity, which would appear to be disordered with a concomitant loss of resolution of the translationally ordered water. In any case, it would be premature, in our view, to conclude that water molecules in clefts and grooves of DNA and protein surfaces are not partially ordered on the basis of the extant reported crystal structures. If partially ordered multilayers exist in certain clefts and grooves, as we suspect, then there also exist multiple exchanging sites within that

space, where the osmolyte does not contact the macromolecular surface. For such sites, the $S_B - S_A$ terms in $\Delta\Gamma_1(2)$ (Eq. 42) might vary much less with osmolyte chemical structure than is the case for surface-contiguous sites, as described below.

We speculate that the osmolyte is preferentially excluded from any partially ordered water multilayer, much as impurities are excluded from macroscopic crystalline ice. In such a case, the exchanging sites within the array are necessarily of the strong water-binding/weak osmolyte-binding variety. Hence, the exchange constant, $K_{\mathbf{q}(j)}$, along an approaching trajectory should decline from its random value, $\bar{V}_3/\bar{V}_1$, in bulk solution to a much smaller value typical of strong water-binding sites, as indicated in Fig. 3. Although the $K_{\mathbf{q}(j)}$-values in the interior and bottom of the array (i.e., to the left of $\mathbf{q}(M)$) may depend strongly on the osmolyte's chemical structure, that will not matter much, because for those sites, $K_{\mathbf{q}(j)} \ll \bar{V}_3/\bar{V}_1$ in any case. However, in the region immediately exterior to (i.e., to the right of) $\mathbf{q}(M)$, the $K_{\mathbf{q}(j)}$-values also depend on the osmolyte chemical structure and are large enough to make a significant contribution compared to the $\bar{V}_3/\bar{V}_1$ term for each site. Indeed, sufficiently far to the right of $\mathbf{q}(M)$, the $K_{\mathbf{q}(j)} = \bar{V}_3/\bar{V}_1$ everywhere in the bulk solution, and the corresponding $(\nu - K_{\mathbf{q}(j)})$ terms in S for those sites vanish altogether. It is conceivable that the deeper, strong water-binding sites in the cleft are much more numerous than the more neutral water-binding sites at the outer surface of the hydration multilayer, as illustrated in Fig. 4 *a*, so the X_B and $L_s^B\nu$ terms dominate the $L_n^B\left(\nu - \langle K_n^B\rangle\right)$ term in Eq. 43 for the B conformer, and a similar circumstance prevails for the A conformer. In such a case, the relative variation of $\Delta\Gamma_1(2)$ with osmolyte

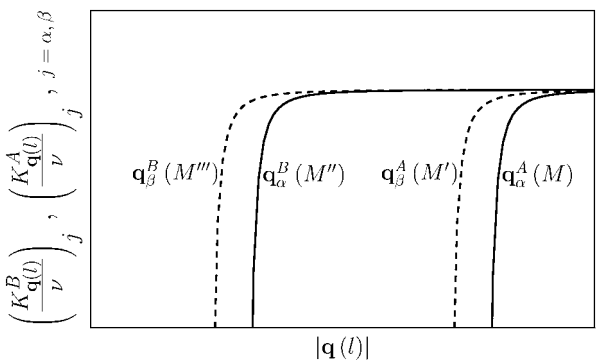

FIGURE 3 Variation of the equilibrium constants for osmolyte/solvent exchange of two different neutral nonzwitterionic osmolytes of different size, $\alpha$and β, along the same path passing from the bulk solution into a hydration multilayer from which the osmolytes are strongly excluded. The macromolecule (protein or DNA) exhibits two different conformations, A and B, and A has a more distant outer boundary of its hydration multilayer than does B. The larger osmolyte α is unable to exchange with sites as close to the outer edge of the hydration multilayer as those accessed by the smaller osmolyte β, and that is true for either conformation, A or B, of the macromolecule. Nevertheless, in this scenario, the difference in positions where the exchange constants turn down for the two conformations, namely $\left|q^B_\alpha(M'')\right| - \left|q^A_\alpha(M)\right|$ for osmolyte α and $\left|q^B_\beta(M''')\right| - \left|q^A_\beta(M')\right|$ for osmolyte β, is practically the same for both osmolytes, regardless of their chemical composition or size. The positions, $\mathbf{q}(M)$ where the exchange constants turn downward, correspond to the position where the osmolyte would first contact the outermost hydration layer, if that layer did not melt.

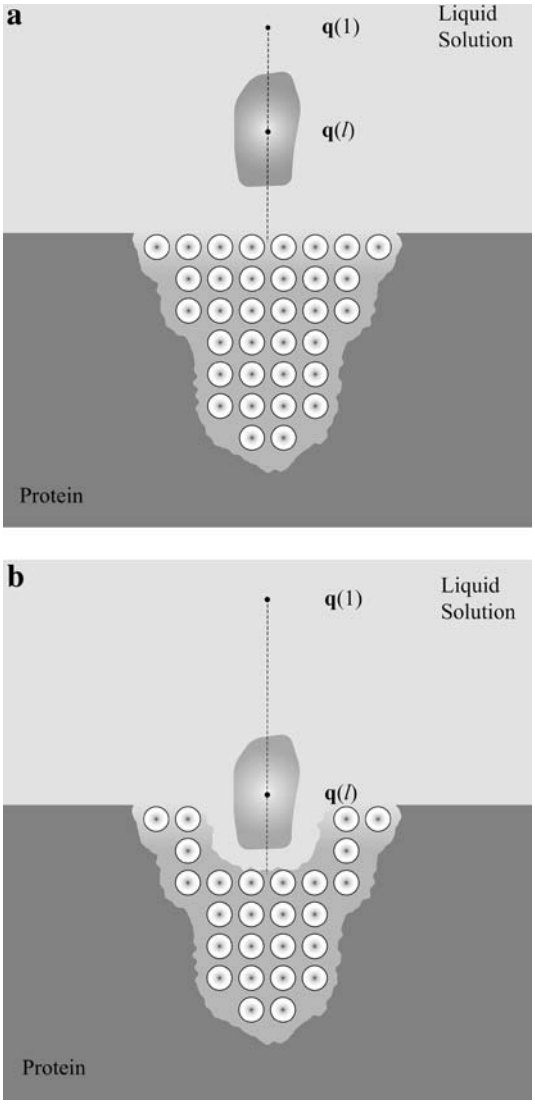

FIGURE 4 (*a*) Schematic illustration of an osmolyte molecule approaching a partially ordered hydration multilayer in a cleft of a protein (or DNA) molecule. The maximum occupation positions of the water molecules are denoted by the ordered array of spheres in this diagram. The path of the osmolyte through the lattice of osmolyte/solvent exchange sites surrounding the protein-water complex begins at position $\mathbf{q}(1)$ in the bulk solution and passes through a sequence of sites that are designated by $\mathbf{q}(\ell)$, $\ell = 1, 2, \ldots$ (*b*) As the osmolyte approaches the partially ordered hydration multilayer, the latter ''melts'' to form a liquid water film that wets the surface of the remaining partially ordered waters, and separates the osmolyte from the partially ordered water. Direct contact between the osmolyte and partially ordered water is thereby prevented, and the osmolyte remains surrounded by liquid everywhere along its advancing trajectory. The free energy cost to ''melt'' the ordered water resists the advance of the osmolyte toward the protein (or DNA), by decreasing the equilibrium constant for osmolyte/solvent exchange as the osmolyte begins to displace the partially ordered water.

chemical structure for osmolytes of a similar size might be rather modest, because the largest terms, $X_B - X_A$ and $(L_s^B - L_s^A)\nu$, are practically invariant to osmolyte structure.

More complete invariance of $\Delta\Gamma_1(2)$ to the chemical structures of osmolytes of similar size requires that the identity of the osmolyte be concealed from the partially ordered water, as it approaches from bulk solution to the position, $\mathbf{q}(M)$, of the steep descent. We propose the following speculative mechanism. If the contact surface free energy between the hydration multilayer and the osmolyte sufficiently exceeds the sum of the surface free energy between the hydration multilayer and liquid water and that between liquid water and the osmolyte, then the hydration multilayer will ''melt'' in front of the advancing osmolyte to create an interposed liquid water film as indicated in Fig. 4 *b*. As an example, crystalline ice forms a stable liquid film at its interface with air, provided that the temperature is not too far below the freezing point (44,45). In this case, the chemical structure of the osmolyte is not directly sensed by the partially ordered water, from which it is separated by the liquid film. Of course, there may also be partially ordered water associated with the osmolyte. Under these conditions, for osmolytes of similar size, the $K_{\mathbf{q}(j)}$ are practically independent of the osmolyte's chemical structure from position $\mathbf{q}(1)$ in the bulk to $\mathbf{q}(M)$. The $K_{\mathbf{q}(j)}$ may vary with osmolyte structure at deeper sites to the left of $\mathbf{q}(M)$, not only for those sites that place the osmolyte in contact with the macromolecular surface, but also for those where the liquid film cannot form, perhaps due to the higher free energy cost of melting those layers. However, those $K_{\mathbf{q}(j)}$-values are very small compared to $\bar{V}_3/\bar{V}_1$ in any case, and contribute little to $\Gamma_1(2_B)$, $\Gamma_1(2_A)$, or $\Delta\Gamma_1(2)$. In this scenario, $\Delta\Gamma_1(2)$ is practically independent of chemical structure for osmolytes of the same hydrated size, but it cannot be identified simply with a change in the amount of bound water in the multilayer, because some of the $K_{\mathbf{q}(j)}$ terms make a significant contribution.

In an ideal case, $\Delta\Gamma_1(2)$ may also be largely independent of osmolyte size over a limited range of osmolyte sizes. To exhibit such behavior, both the excluded volume contribution, $[X_B - X_A]_a$, and the exchange reaction contribution, $[S_B - S_A]_a$, associated with the active part(s) of the surface must display practically the same values for two osmolytes, α and β, of different size. That is, one must have, $[X_B - X_A]_a^\beta = [X_B - X_A]_a^\alpha$, where $[X_B]_a^\beta = ([V_\beta^{ex}]_{Ba} - [V_1^{ex}]_{Ba})/\bar{V}_1$ is the difference between the volume excluded by species B to osmolyte β and that excluded to solvent 1 in units of $\mathbf{V}_1$, and similar definitions apply to $[X_A]_a^\beta$, $[X_B]_a^\alpha$, and $[X_A]_a^\alpha$. This condition is equivalent to $[V_\beta^{ex}]_{Ba} - [V_\beta^{ex}]_{Aa} = [V_\alpha^{ex}]_{Ba} - [V_\alpha^{ex}]_{Aa}$. The difference volume, $[V_\beta^{ex}]_{Ba} - [V_\beta^{ex}]_{Aa}$, may well be nearly independent of the size of the osmolyte β over a limited range of sizes, because it corresponds to the volume of a partial ''shell'', associated with the active parts of the surface, whose thickness is just the difference in the positions of the active surfaces of the B and A conformers relative to their own central atoms. In such a case, $[X_B - X_A]_a^\beta \cong ([X_B - X_A]_a^\alpha)$. In addition, an osmolyte of smaller size (β) begins to penetrate the hydration multilayer at a closer distance to the surface of conformer B than does an osmolyte of larger size (α), so that its $K_{\mathbf{q}(j)}$ turns downward at a correspondingly smaller distance ($\mathbf{q}_\beta^B(M')$ vs. $\mathbf{q}_\alpha^B(M')$), as indicated in Fig. 3. However, if the separation between the turndown positions for those two osmolytes (β and α) is the same for conformer A as for conformer B, as indicated in Fig. 3, then the contribution of the smaller osmolyte to $\Delta\Gamma_1(2)$ may be nearly the same as that of the larger osmolyte. This can be seen from the expression, $[S_B - S_A]_a^\alpha = [-(1/\bar{V}_1)\int d\mathbf{q}(j)((K_{\mathbf{q}(j)}^{B\alpha}/\nu_\alpha) - (K_{\mathbf{q}(j)}^{A\alpha}/\nu_\alpha))]_a$, wherein $\nu_\alpha = \bar{\mathbf{V}}_\alpha/\bar{V}_1, \bar{V}_\alpha$ is the volume of the exchange sites, which is here taken as the partial molecular volume of the osmolyte α, and $K_{\mathbf{q}(j)}^{B\alpha}$ and $K_{\mathbf{q}(j)}^{A\alpha}$ are the exchange constants for an osmolyte α with its center in the volume element $d\mathbf{q}(j)$ at position $\mathbf{q}(j)$ in the vicinity of the species B or A, respectively. The integral is taken over the volume of the osmolyte accessible region extending from the active part of the macromolecular surface out to the point, where both $K_{\mathbf{q}(j)}^{B\alpha}/\nu_\alpha = 1.0$ and $K_{\mathbf{q}(j)}^{A\alpha}/\nu_\alpha = 1.0$, so the integrand vanishes. Analogous considerations apply for $[S_B - S_A]_a^\beta$. For the case illustrated in Fig. 3, where the difference between the downturn positions of the osmolyte β in the vicinity of B and A is very similar to the corresponding difference for the osmolyte α, the integrals may well take rather similar values for two osmolytes of somewhat different size, hence their $[S_B - S_A]_a$-values may also be nearly identical. Under these conditions, then, $\Delta\Gamma_1(2)$ may become independent of osmolyte size, as well as osmolyte chemical structure, over a limited range of osmolyte sizes. However, $\Delta\Gamma_1(2)$ does not necessarily correspond to the change in the amount of ordered water, because some of the $K_{\mathbf{q}(j)}$ terms make a significant contribution.

It is not known whether a liquid film actually is formed between a hydration multilayer and an osmolyte, or at what, if any, depth it ceases to form, because the free energy required to ''melt'' an interior layer is simply too high. Nevertheless, this scenario may merit consideration in those cases, where $\Delta\Gamma_1(2)$ is found to be independent of osmolyte chemical structure for osmolytes of the same size, and perhaps also independent of osmolyte size over a limited range of sizes, but where also the crystal structure appears to provide no places to sequester water in such a way that it could not be displaced by those same osmolytes.

Possible biological relevance of highly excluded osmolytes

It is noteworthy that cells employ osmoprotectants that are zwitterionic (betaine glycine, glycine, proline, and trimethylamine N-oxide) or both ionic and zwitterionic (glutamate), and which are excluded from 2.3 to 5.3 times as much water-occupied volume as is glycerol, for which $\bar{V}_3 = 3.9\,\bar{V}_1$. Cells also employ neutral nonzwitterionic species, such as trehalose, which has a larger $\bar{V}_3$, and is also excluded from

three times as much water as glycerol. Such superexcluded osmolytes are unable to displace some of the innermost water molecules in clefts, grooves, and pits. It is conceivable that these osmolytes have been selected to lower a_1 without displacing important inner water molecules associated with the cell's macromolecules, which may be the most strongly coupled to changes in macromolecular structure and function. Felitsky et al. (29) proposed that betaine glycine was commonly selected as an osmoprotectant, because it has the least effect on protein-unfolding equilibria for a given osmolality. It is not unlikely that both functions, namely preserving inner waters and minimally altering unfolding equilibria, are important for an osmoprotectant.

APPENDIX A: DEFINITION OF $G_{\alpha\beta}(R)$ IN THE GRAND ENSEMBLE

The pair correlation function is defined in the following way. First, a particular atom at the same topological position in every molecule of a given kind is arbitrarily designated as its central atom. The coordinates of the central atom of the ith molecule of the αth kind are denoted by $\mathbf{r}_i^\alpha$, and the coordinates of its remaining $(\varsigma_\alpha - 1)$ atoms are denoted by a generalized $3(\varsigma_\alpha - 1)\times 1$ vector, $\mathbf{S}_i^\alpha$. The full set of coordinates of all the atoms in the ith molecule of kind α are denoted by a generalized $3\varsigma_\alpha \times 1$ vector, $\mathbf{R}_i^\alpha = (\mathbf{r}_i^\alpha, \mathbf{S}_i^\alpha)$. The full set of all of the coordinates of all the atoms in all the molecules in the volume V are denoted by a generalized vector $\mathbf{R}$ of dimension $M = \sum_{\alpha=1}^{N} N_\alpha \cdot 3\varsigma_\alpha$, which has associated volume element $d^M\mathbf{R}$. The potential energy function $U(\mathbf{R})$ depends upon all of the coordinates. The pair correlation function, or radial distribution function, for the central atom of a β-molecule at $\mathbf{r}_2$, given a central atom of an α-molecule at $\mathbf{r}_1$, is defined by

$$\frac{\langle N_\alpha\rangle\langle N_\beta\rangle}{V^2} g_{\alpha\beta}(\mathrm{r}_{12}) \equiv \sum_{N_\alpha=0}^{\infty}\sum_{N_\beta=0}^{\infty}\cdots\sum_{N_\eta=0}^{\infty}\left(\prod_{j=1}^{\nu}\left(\lambda_j^{N_j}(\Lambda_j)^{3N_j}/N_j!\right)\right)$$
$$\left(N_\alpha N_\beta - N_\alpha\delta_{\alpha\beta}\right)\int\cdots\int d^M\mathbf{R}\left((1-\delta_{\alpha\beta})\delta(\mathbf{r}_1^\alpha - \mathbf{r}_1)\delta(\mathbf{r}_1^\beta - \mathbf{r}_2)\right.$$
$$\left. + \delta_{\alpha\beta}\delta(\mathbf{r}_1^\alpha - \mathbf{r}_1)\delta(\mathbf{r}_2^\alpha - \mathbf{r}_2)\right)\exp[-U(\mathbf{R})/kT]/GPF, \qquad \text{(A1)}$$

wherein $\lambda_\ell = \exp[\mu_\ell/kT]$, $\Lambda_\alpha \equiv \Pi_{n=1}^{\varsigma_\alpha}((2\pi m_{\alpha n}kT)/h^2)^{1/2}$, $\delta_{\alpha\beta}$ is the Kronecker δ, k is Boltzmann's constant, T is absolute temperature, $m_{\alpha n}$ is the mass of the nth atom of an α-molecule, and GPF denotes the grand partition function,

$$GPF \equiv \sum_{N_\alpha=0}^{\infty}\cdots\sum_{N_\eta=0}^{\infty}\left(\prod_{j=1}^{\nu}\left(\lambda_j^{N_j}(\Lambda_j)^{3N_j}/N_j!\right)\right)$$
$$\int\cdots\int d^M\mathbf{R}\exp[-U(\mathbf{R})/kT]. \qquad \text{(A2)}$$

In the thermodynamic limit of extremely large systems, the terms in the sums of both the numerator and denominator on the right-hand side of Eq. A1 are strongly peaked near the mean values of the N_j, in which case the kinetic energy factors $(\Lambda_j)^{3N_j}$ cancel out of the $g_{\alpha\beta}(\mathrm{r}_{12})$, as expected in classical statistical mechanics.

APPENDIX B: EVALUATION OF $(\partial c_1/\partial c_3)_{T,P,c_2}$

For the kth species, $c_k \equiv N_k/V$, so $dc_k = dN_k/V - N_k\,dV/V^2$, and $dN_k/V = dc_k + c_k d\ln V$. At constant T,

$$dV = \bar{V}_1\,dN_1 + \bar{V}_2\,dN_2 + \bar{V}_3\,dN_3 + (\partial V/\partial P)_{N_1,N_2,N_3}\,dP. \qquad \text{(B1)}$$

Dividing Eq. B1 by V and using the preceding relations for $\mathrm{d}N_k/V$ yields

$$d\ln V = \bar{V}_1(dc_1 + c_1\,d\ln V) + \bar{V}_2(dc_2 + c_2\,d\ln V) + \bar{V}_3(dc_3 + c_3\,d\ln V) - \kappa dP, \qquad \text{(B2)}$$

wherein the compressibility is defined by $\kappa \equiv -(\partial\ln V/\partial P)_{N_1,N_2,N_3}$. After collecting all of the $d\ln V$ terms on the left-hand side, one has $(d\ln V)(1 - c_1\bar{V}_1 - c_2\bar{V}_2 - c_3\bar{V}_3) = 0$ on that side. There remains then

$$0 = \bar{V}_1 dc_1 + \bar{V}_2 dc_2 + \bar{V}_3 dc_3 - \kappa dP, \qquad \text{(B3)}$$

so at constant c_2 and P one has finally

$$(\partial c_1/\partial c_3)_{P,c_2} = -\bar{V}_3/\bar{V}_1. \qquad \text{(B4)}$$

APPENDIX C: EVALUATION OF $\bar{V}_\alpha$

At constant T, $\mu_j(T,c_1,\ldots,c_\nu)$ depends upon all the concentrations. Hence,

$$\bar{V}_\alpha = (\partial\mu_\alpha/\partial P)_{N_\gamma} = \sum_{\beta=1}^{\nu}(\partial\mu_\alpha/\partial c_\beta)_{c_{\gamma\neq\beta}}(\partial c_\beta/\partial P)_{N_\gamma}. \qquad \text{(C1)}$$

Using $dc_\beta = (1/V)dN_\beta - c_\beta\,d\ln V$ from Appendix B, one finds

$$(\partial c_\beta/\partial P)_{N_\gamma} = -c_\beta(\partial\ln V/\partial P)_{N_\gamma} = c_\beta\kappa, \qquad \text{(C2)}$$

where κ is the compressibility defined in Appendix B. Use of Eq. C2 in Eq. C1 yields

$$\bar{V}_\alpha = \kappa\sum_{\beta=1}^{\nu} c_\beta(\partial\mu_\alpha/\partial c_\beta)_{c_{\gamma\neq\beta}}. \qquad \text{(C3)}$$

Equation C3 is multiplied by c_α on both sides and summed over α=1, . . . , ν to give

$$1.0 = \kappa\sum_{\alpha=1}^{\nu}\sum_{\beta=1}^{\nu} c_\alpha c_\beta(\partial\mu_\alpha/\partial c_\beta)_{c_{\gamma\neq\beta}}. \qquad \text{(C4)}$$

After inserting Eq. 9, there results

$$\kappa = \frac{|B|}{kT\sum_{\alpha=1}^{\nu}\sum_{\beta=1}^{\nu} c_\alpha c_\beta|B|_{\alpha\beta}}. \qquad \text{(C5)}$$

After inserting Eq. C5 into Eq. C3 and again using Eq. 9, there results

$$\bar{V}_\alpha = \frac{\sum_{\beta=1}^{\nu} c_\beta|B|_{\alpha\beta}}{\sum_{\gamma=1}^{\nu}\sum_{\delta=1}^{\nu} c_\gamma c_\delta|B|_{\gamma\delta}}. \qquad \text{(C6)}$$

Equations C5 and C6 are precisely the expressions of Kirkwood and Buff (10).

APPENDIX D: VERIFICATION OF BEN-NAIM'S EXPRESSION FOR $\bar{V}_2$

We adopt Eq. 6.17.22 of Ben-Naim (17) as a conjecture for the three-component system at constant T:

$$\bar{V}_2 = -\bar{V}_1 c_1 G_{12} - \bar{V}_3 c_3 G_{32} + kT\kappa. \tag{D1}$$

In the limit $c_2 \to 0, \bar{V}_1, \bar{V}_3$, and κ are properties of the two-component solution (1 + 3), whereas $\bar{V}_2$ is a property of the three-component solution. Equation C3 with $\alpha = 3$ can be written as

$$\kappa = \bar{V}_3 / \sum_\beta c_\beta (\partial \mu_3 / \partial c_\beta)_{T, c_{\gamma \neq \beta}} = |B'|/kTD', \tag{D2}$$

where the second equality was obtained from the first by using Eq. C6 for $\bar{V}_3$ and Eq. 9 for $(\partial \mu_3/\partial c_\beta)_{T, c_{\gamma \neq \beta}}$, the primes denote quantities pertaining to the two-component system, and

$$D' = \sum_\beta \sum_\gamma c_\beta c_\gamma |B|_{\beta\gamma} = c_1^2 B_{33} - 2c_1 c_3 B_{31} + c_3^2 B_{11}. \tag{D3}$$

Use was made of the symmetry of the B-matrix to obtain the final equality of Eq. D3. After substituting equations D2 and C6 into the right-hand side (*rhs*) of Eq. D1 and expanding out the two-component sums for $\bar{V}_1$, $\bar{V}_3$, and $|B'|$ (in κ) there results

$$\begin{aligned} rhs = \{&-c_1 G_{12}(c_1 B_{33} - c_3 B_{31}) - c_3 G_{32}(-c_1 B_{13} + c_3 B_{11}) \\ &+ B_{11} B_{33} - B_{13} B_{31}\}/D' \end{aligned} \tag{D4}$$

The left-hand side (*lhs*) of Eq. D1 is evaluated for the three-component system via Eq. C6, which after expansion of the various terms gives

$$\begin{aligned} lhs &= \{c_1 |B|_{21} + c_3 |B|_{23} + c_2 |B|_{22}\}/D \\ &= \{-c_1 (B_{12} B_{33} - B_{13} B_{32}) - c_3 (B_{11} B_{32} - B_{12} B_{31}) \\ &\quad + c_2 (B_{11} B_{33} - B_{13} B_{31})\}/D, \end{aligned} \tag{D5}$$

where

$$\begin{aligned} D = {} & c_1^2 (B_{22} B_{33} - B_{23} B_{32}) + c_2^2 (B_{11} B_{33} - B_{13} B_{31}) \\ & + c_3^2 (B_{11} B_{22} - B_{12} B_{21}) + 2c_1 c_3 (B_{21} B_{32} - B_{22} B_{31}) \\ & - 2c_1 c_2 (B_{21} B_{33} - B_{23} B_{31}) - 2c_3 c_2 (B_{11} B_{23} - B_{13} B_{21}). \end{aligned} \tag{D6}$$

Every term in the numerator of *lhs* in Eq. D5 contains a single factor of c_2. The terms in D in Eq. D6 all contain either one or two factors of c_2. Hence, both numerator and denominator of *lhs* can be divided by c_2. Any terms that still contain a factor of c_2 can be neglected against the constant terms in the limit $c_2 \to 0$. In this limit, D becomes identical to D' in Eq. D3, and finally

$$\begin{aligned} lhs = \{&-c_1 G_{12}(c_1 B_{33} - c_3 B_{31}) - c_3 G_{32}(-c_1 B_{13} + c_3 B_{11}) \\ &+ B_{11} B_{33} - B_{13} B_{31}\}/D' = rhs, \end{aligned} \tag{D7}$$

so the conjectured Eq. D1 is verified.

APPENDIX E: CONNECTION BETWEEN $\Gamma_2(3)$ AND OTHER PREFERENTIAL INTERACTION COEFFICIENTS

For a three-component system, μ_2 and μ_3 depend upon T, P, and the concentrations, c_2 and c_3. Thus, at constant T, P, and c_2,

$$d\mu_2 = \mu_{23}^{c} dc_3 \tag{E1}$$

$$d\mu_3 = \mu_{33}^{c} dc_3, \tag{E2}$$

where $\mu_{k3}^{c} \equiv (\partial \mu_k/\partial c_3) T, P, c_2, k = 2, 3$. Dividing Eq. E1 by Eq. E2 yields

$$\Gamma_3(2) \equiv -(\partial \mu_2/\partial \mu_3)_{T,P,c_2} = -\left(\mu_{23}^{c}/\mu_{33}^{c}\right). \tag{E3}$$

For the same system, μ_2 and μ_3 can also be expressed in terms of T, P, and the molalities, m_2 and m_3, so at constant T, P,

$$d\mu_2 = \mu_{22}^{m} dm_2 + \mu_{23}^{m} dm_3 \tag{E4}$$

$$d\mu_3 = \mu_{32}^{m} dm_2 + \mu_{33}^{m} dm_3, \tag{E5}$$

where $\mu_{kl}^{m} \equiv (\partial \mu_k/\partial m_l)_{T,P,m_{\gamma \neq l}}, k, l = 2, 3$. Hence, at constant T, P,

$$\mu_{23}^{c} = \mu_{22}^{m} (\partial m_2/\partial c_3)_{c_2} + \mu_{23}^{m} (\partial m_3/\partial c_3)_{c_2} \tag{E6}$$

$$\mu_{33}^{c} = \mu_{32}^{m} (\partial m_2/\partial c_3)_{c_2} + \mu_{33}^{m} (\partial m_3/\partial c_3)_{c_2}. \tag{E7}$$

The subscripts denoting constant T and P of the slopes, $\partial m_k/\partial c_1$, in Eqs. E6 and E7 are omitted for clarity and this convention applies in the sequel.

We assume for simplicity that the partial molar volumes, $\bar{V}_1$, $\bar{V}_2$, and $\bar{V}_3$, are constants independent of c_2 and c_3 at constant T, P over the range of c_3 up to ~1.0 M, which should be a good approximation, provided that c_2 is sufficiently dilute.

The molalities, m_2 and m_3, can be expressed in terms of the concentrations and solvent molar mass, M_1, by

$$m_2 = c_2/(c_1 M_1) = c_2 (\bar{V}_1/M_1)/(1 - c_2 \bar{V}_2 - c_3 \bar{V}_3) \tag{E8}$$

$$m_3 = c_3/(c_1 M_1) = c_3 (\bar{V}_1/M_1)/(1 - c_2 \bar{V}_2 - c_3 \bar{V}_3). \tag{E9}$$

After performing the derivatives and collecting terms, we obtain

$$(\partial m_2/\partial c_3)_{c_2} = m_2 \bar{V}_3/(1 - c_2 \bar{V}_2 - c_3 \bar{V}_3) = m_2 \bar{V}_3/c_1 \bar{V}_1 \tag{E10}$$

$$\begin{aligned} (\partial m_3/\partial c_3)_{c_2} &= m_3 (1 - c_2 \bar{V}_2)/(c_3 (1 - c_2 \bar{V}_2 - c_3 \bar{V}_3)) \\ &= m_3 (1 - c_2 \bar{V}_2)/(c_3 c_1 \bar{V}_1). \end{aligned} \tag{E11}$$

In the limit $c_2 \bar{V}_2 \ll 1.0$, the $c_2 \bar{V}_2$ term in Eq. E11 may be neglected, and we obtain

$$(\partial m_2/\partial c_3)_{c_2}/(\partial m_3/\partial c_3)_{c_2} = (m_2/m_3)(c_3 \bar{V}_3) = c_2 \bar{V}_3. \tag{E12}$$

After dividing Eq. E6 by Eq. E7 in accord with Eq. E3, dividing both numerator and denominator of the resulting quotient by $\mu_{33}^{m} (\partial m_3/\partial c_3)_{c_2}$, and rearranging somewhat, there results

$$\begin{aligned} \Gamma_3(2) &= (-)\left(\frac{\mu_{22}^{m}}{\mu_{33}^{m}} c_2 \bar{V}_3 - \Gamma_{\mu_3}^{m}\right) \Big/ \left((-)\Gamma_{\mu_3}^{m} c_2 \bar{V}_3 + 1\right) \\ &= \left(\Gamma_{\mu_3}^{m} - c_2 \bar{V}_3 \Gamma_{\mu_2}^{m} \Gamma_{\mu_3}^{m}\right) \Big/ \left(1 - c_2 \bar{V}_3 \Gamma_{\mu_3}^{m}\right). \end{aligned} \tag{E13}$$

The Maxwell relation, $\mu_{23}^{m} = \mu_{32}^{m}$, and the definition, $\Gamma_{\mu_3}^{m} \equiv -\mu_{32}^{m}/\mu_{33}^{m}$, were used to obtain the first equality, and the definition, $\Gamma_{\mu_2}^{m} \equiv -\mu_{22}^{m}/\mu_{23}^{m}$, was used to obtain the second. In the limit of small c_2, $\Gamma_{\mu_3}^{m}$ becomes independent of c_2, and the second term in the denominator may be neglected to obtain

$$\Gamma_3(2) = \Gamma_{\mu_3}^{m} \left(1 - c_2 \bar{V}_3 \Gamma_{\mu_2}^{m}\right). \tag{E14}$$

The second term in Eq. E14 cannot be neglected for small c_2, as seen in Eq. E15 below. The definitions of $\Gamma_{\mu_3}^{m}$, $\Gamma_{\mu_2}^{m}$, and $\Gamma_{\mu_1}^{m} \equiv (-)\mu_{12}^{m}/\mu_{13}^{m}$, and their alternative representations used here are given in Table 1 of Anderson et al. (19). An exact expression linking $\Gamma_{\mu_3}^{m}$, $\Gamma_{\mu_2}^{m}$, and $\Gamma_{\mu_1}^{m}$ was also derived in Eq. 12 of that same article, and can be rearranged without approximation to give

$$\Gamma_{\mu_2}^{m} = (m_3/m_2)\left(\left(\Gamma_{\mu_3}^{m} - \Gamma_{\mu_1}^{m}\right)/\Gamma_{\mu_3}^{m}\right) + \Gamma_{\mu_1}^{m}. \tag{E15}$$

Inserting Eq. E15 into Eq. E14 and taking the limit $c_2 \to 0$, where $\Gamma_{\mu_3}^{m}$ and $\Gamma_{\mu_1}^{m}$ remain constant, yields

$$\Gamma_3(2) = \Gamma^{\mathrm{m}}_{\mu_3}\left(1 - \phi_3\left(1 - \Gamma^{\mathrm{m}}_{\mu_1}/\Gamma^{\mathrm{m}}_{\mu_3}\right)\right), \tag{E16}$$

where $\phi_3 \equiv c_3\bar{V}_3$ is the volume fraction of species 3.

An expression relating the ''dialysis'' preferential interaction coefficient, $\Gamma^{\mathrm{m}}_{\mu_1,\mu_3} \equiv (\partial m_3/\partial m_2)_{\mathrm{T},\mu_1,\mu_3}$, to $\Gamma^{\mathrm{m}}_{\mu_1}$ and $\Gamma^{\mathrm{m}}_{\mu_3}$ was derived by Anderson et al. (20) (their Eq. 20). In the limit $m_2 \to 0$, that relation becomes

$$\begin{aligned}\Gamma^{\mathrm{m}}_{\mu_1,\mu_3} &= \Gamma^{\mathrm{m}}_{\mu_3} + (m_3\bar{V}_3)(\Gamma^{\mathrm{m}}_{\mu_1} - \Gamma^{\mathrm{m}}_{\mu_3})/(m_1\bar{V}_1 + m_3\bar{V}_3)\\ &= \Gamma^{\mathrm{m}}_{\mu_3}(1-\phi_3) + \phi_3\Gamma^{\mathrm{m}}_{\mu_1}.\end{aligned} \tag{E17}$$

The relation between $\Gamma_3(2)$ and $\Gamma^{\mathrm{m}}_{\mu_1,\mu_3}$ is obtained by solving Eq. E17 for $\Gamma^{\mathrm{m}}_{\mu_3}$,

$$\Gamma^{\mathrm{m}}_{\mu_3} = \left(\Gamma^{\mathrm{m}}_{\mu_1,\mu_3} - \phi_3\Gamma^{\mathrm{m}}_{\mu_1}\right)/(1-\phi_3), \tag{E18}$$

and inserting that into Eq. E16 in both places where $\Gamma^{\mathrm{m}}_{\mu_3}$ occurs. There results finally

$$\begin{aligned}\Gamma_3(2) &= \frac{\left(\Gamma^{\mathrm{m}}_{\mu_1,\mu_3} - \phi_3\Gamma^{\mathrm{m}}_{\mu_1}\right)}{(1-\phi_3)}\left(1 - \phi_3\left(1 - \frac{(1-\phi_3)\Gamma^{\mathrm{m}}_{\mu_1}}{\Gamma^{\mathrm{m}}_{\mu_1,\mu_3} - \phi_3\Gamma^{\mathrm{m}}_{\mu_1}}\right)\right)\\ &= \Gamma^{\mathrm{m}}_{\mu_1,\mu_3}.\end{aligned} \tag{E19}$$

Thus, in the limit $m_2 \to 0$, this concentration-based $\Gamma_3(2)$ is equal to the molality-based $\Gamma^{\mathrm{m}}_{\mu_1,\mu_3}$, provided that $\bar{V}_1$, $\bar{V}_2$, and $\bar{V}_3$, are constants independent of c_3 at constant T, P over the range considered.

The authors thank Profs. Bob Mazo and John Schellman for pointing out an error in an earlier version of this manuscript, and Dr. Jonathan Cannon and Prof. Tom Record for providing recent experimental results before publication, and also for many helpful comments and suggestions.

This work was supported in part by grant RO1 GM61685 from the National Institutes of Health.

REFERENCES

1. Timasheff, S. N. 1993. The control of protein stability and association by weak interactions with water: how do solvents affect these processes? *Annu. Rev. Biophys. Biomol. Struct.* 22:67–97.
2. Parsegian, A., R. P. Rand, and D. C. Rau. 1995. Macromolecules and water: probing with osmotic stress. *Methods Enzymol.* 259:43–94.
3. Record, M. T., Jr., W. Zhang, and C. F. Anderson. 1998. 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 of salts. *Adv. Protein Chem.* 51:281–353.
4. Timasheft, S. N. 1998. Control of protein stability and reactions by weakly interacting cosolvents: the simplicity of the complicated. *Adv. Protein Chem.* 51:355–432.
5. Courtenay, E. S., M. W. Capp, C. F. Anderson, and M. T. Record Jr. 2000. 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.* 39:4455–4471.
6. Parsegian, V. A., R. P. Rand, and D. C. Rau. 2000. Osmotic stress, preferential hydration, and binding: a comparison of perspectives. *Proc. Natl. Acad. Sci. USA.* 97:3987–3992.
7. Timasheff, S. N. 2002. Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components. *Proc. Natl. Acad. Sci. USA.* 99:9721–9726.
8. Inoue, M., and S. N. Timasheff. 1972. Preferential and absolute interactions of solvent components in mixed solvent systems. *Biopolymers.* 11:737–743.
9. Record, M. T., Jr., and C. F. Anderson. 1995. Interpretation of preferential interaction coefficients of nonelectrolytes and of electrolyte ions in terms of a two-domain model. *Biophys. J.* 68:786–794.
10. Kirkwood, J. G., and F. P. Buff. 1951. The statistical mechanical theory of solutions. *J. Chem. Phys.* 19:774–777.
11. Chitra, R., and P. E. Smith. 2001. Preferential interaction coefficients with hydrophobic solutes. *J. Phys. Chem. B.* 105:11513–11532.
12. Shimizu, S. 2004. Estimating hydration changes upon biomolecular reactions from osmotic stress, high pressure, and preferential interaction coefficients. *Proc. Natl. Acad. Sci. USA.* 101:1195–1199.
13. Shimizu, S. 2004. Estimation of excess solvation numbers of water and cosolvents from preferential interaction and volumetric experiments. *J. Chem. Phys.* 120:4989–4990.
14. Shimizu, S., and D. J. Smith. 2004. Preferential hydration and the exclusion of cosolvents from protein surfaces. *J. Chem. Phys.* 121: 1148–1154.
15. Ben-Naim, A. 1975. Solvent effects on chemical equilibria. *J. Chem. Phys.* 63:2064–2073.
16. Ben-Naim, A. 1988. Theory of preferential solvation of nonelectrolytes. *Cell Biophys.* 12:255–269.
17. Ben-Naim, A. 1992. Statistical Thermodynamics for Chemists and Biochemists. Plenum Press, New York, NY.
18. Schellman, J. A. 2003. Protein stability in mixed solvents: A balance of contact interaction and excluded volume. *Biophys. J.* 85:108–125.
19. Anderson, C. F., E. S. Courtenay, and M. T. Record Jr. 2002. Thermodynamic expressions relating different types of preferential interaction coefficients in solutions containing two solute components. *J. Phys. Chem. B.* 106:418–433.
20. Anderson, C. F., D. J. Felitsky, J. Hong, and M. T. Record Jr. 2002. Generalized derivation of exact relationship linking different coefficients that characterize the thermodynamic effects of preferential interactions. *Biophys. Chem.* 101–102:497–511.
21. Schellman, J. A. 1987a. The thermodynamic stability of proteins. *Annu. Rev. Biophys. Biophys. Chem.* 16:115–137.
22. Schellman, J. A. 1987b. Selective binding and solvent denaturation. *Biopolymers.* 26:549–559.
23. Schellman, J. A. 1990. A simple model for solvation in mixed solvents. Applications to stabilization and destabilization of macromolecular structures. *Biophys. Chem.* 37:121–140.
24. Schellman, J. A. 1993. The relation between the free energy of interaction and binding. *Biophys. Chem.* 45:273–279.
25. Schellman, J. A. 1994. The thermodynamics of solvent exchange. *Biopolymers.* 34:1015–1026.
26. Schellman, J. A., and N. C. Gassner. 1996. The enthalpy of transfer of unfolded proteins into solutions of urea and quanidinium hydrochloride. *Biophys. Chem.* 59:259–275.
27. Hill, T. L. (1960). Introduction to Statistical Thermodynamics. Addison-Wesley Publishing, Reading, MA.
28. McQuarrie, D. A. 1976. Statistical Mechanics. Harper and Row, New York, NY.
29. Felitsky, D. J., J. G. Cannon, M. W. Capp, J. Hong, A. W. Van Wynsberghe, C. F. Anderson, and M. T. Record Jr. 2004. The exclusion of glycine betaine from anionic biopolymer surface: why glycine betaine is an effective osmoprotectant but also a compatible solute. *Biochemistry.* 43:14732–14743.
30. Hong, J., M. W. Capp, C. F. Anderson, R. M. Saecker, D. J. Felitsky, M. W. Anderson, and M. T. Record, Jr. 2004. Preferential interactions of glycine betaine and urea with DNA: implications for DNA hydration and for effects of these solutes on DNA stability. *Biochemistry.* 43: 14744–14758.

31. Sugio, S., A. Kashima, S. Mochizuki, M. Noda, and K. Kobayashi. 1999. Crystal structure of human serum albumin at 2.5 Å resolution. *Protein Eng.* 12:439–446.

32. Connolly, J. L. 1993. The molecular surface package. *J. Mol. Graph.* 4: 139–141.

33. Spink, C. H., and J. B. Chaires. 1999. Effects of hydration, ion release, and excluded volume on the melting of triplex and duplex DNA. *Biochemistry.* 38:496–508.

34. Preisler, R. S., H. M. Chen, M. F. Colombo, Y. Choe, B. J. Short, and D. C. Rau. 1995. The B-form to Z-form transition of poly (dG – m(5)dC) is sensitive to neutral solutes through an osmotic stress. *Biochemistry.* 34:14400–14407.

35. Ruggiero-Neto, J., and M. F. Colombo. 2000. Water regulation of actinomycin-D binding to DNA: the interplay among drug affinity, DNA long-range conformation and hydration. *Biopolymers.* 53:46–59.

36. Garner, M. M., and D. C. Rau. 1995. Water release associated with specific binding of gal repressor. *EMBO J.* 14:1257–1263.

37. Siderova, N. Y., and D. C. Rau. 1996. Differences in water release for the binding of EcoRI to specific and non-specific DNA sequences. *Proc. Natl. Acad. Sci. USA.* 93:12272–12277.

38. Colombo, M. F., D. C. Rau, and V. A. Parsegian. 1992. Protein solvation in allosteric regulation: a water effect on hemoglobin. *Science.* 256:655–659.

39. Vodyanoy, I., S. M. Bezrukov, and V. A. Parsegian. 1993. Probing alamecithin channels with water-soluble polymers. Size modulated osmotic action. *Biophys. J.* 65:2097–2105.

40. Rangel, D. P., C. A. Sucato, C. H. Spink, B. S. Fujimoto, and J. M. Schurr. 2004. Effects of small neutral osmolytes on the supercoiling free energy and intrinsic twist of p30δ DNA. *Biopolymers.* 75:291–313.

41. Israelachvili, J. 1987. Solvation forces and liquid structure, as probed by direct force measurements. *Acc. Chem. Res.* 20:415–421.

42. Shui, X., C. C. Sines, L. McFail-Isom, D. Van Derveer, and L. D. Williams. 1998. Structure of the potassium form of CGCGAATTCG-CG:DNA. Deformation by electrostatic collapse around inorganic ions. *Biochemistry.* 37:16877–16887.

43. Shui, X., L. McFail-Isom, G. G. Hu, and L. D. Williams. 1998. The B-DNA dodecamer at high resolution reveals a spine of water on sodium. *Biochemistry.* 37:8341–8355.

44. Elbaum, M., S. B. Lipson, and J. G. Dash. 1993. Optical study of surface melting on ice. *J. Cryst. Growth.* 129:491–505.

45. Wilen, L. A., J. S. Wettlaufer, M. Elbaum, and M. Schick. 1995. Dispersion force effects in interfacial premelting of ice. *Phys. Rev. B.* 52:12426–12433.