# Equilibrium Dialysis Data and the Relationships between Preferential Interaction Parameters for Biological Systems in Terms of Kirkwood−Buff Integrals

**Paul E. Smith***

*Department of Chemistry, 111 Willard Hall, Kansas State University, Manhattan, Kansas 66506-3701*

*Received: October 24, 2005; In Final Form: December 15, 2005*

Equilibrium dialysis data has provided valuable information concerning the preferential interaction of a cosolvent with a biomolecule in aqueous solutions. Here, we formulate the experimental data in terms of Kirkwood−Buff (KB) theory, resulting in equations that provide a simple physical picture of the dialysis experiment and thereby the interaction of a cosolvent with a biomolecule. These results are then used to establish exact relationships between preferential interaction coefficients, defined in different ensembles and/or using different concentration scales, in terms of KB integrals. It is then argued that the molality based equilibrium dialysis data represent the situation most relevant to computer simulations performed in either open or closed systems.

## Introduction

The concept of preferential interactions has been consistently used to quantify interactions between different species in solution. The general theory has been outlined in detail by Scatchard,[1] Wyman,[2] Eisenberg,[3,4] Tanford,[5,6] Timasheff,[7,8] Schellman,[9,10] and Record.[11,12] More recently, Kirkwood−Buff (KB) theory has also been used to analyze experimental data concerning preferential interactions in biomolecular systems.[13−21] KB theory provides a quantitative measure of the relative distribution of the different solution species in terms of KB integrals involving the corresponding radial distribution functions (rdfs).[22−24] The rdfs are easily obtained from computer simulations, and therefore this approach is particularly well suited for the analysis of simulations of biomolecules in mixed solvents.[13,14] However, some confusion has arisen in the literature due to the presence of several different thermodynamic definitions of the preferential interaction depending on the choice of ensemble and concentration units.[17,19−21]

Previously, we have provided relationships for the molality based preferential interaction parameters in terms of KB integrals using a somewhat indirect approach.[13,14] In this study we reanalyze the theory of preferential interactions in open systems and express the experimental equilibrium dialysis data in terms of KB integrals. The approach involves starting from an analysis of the fully open system to determine expressions for the preferential interactions in terms of KB integrals for the case of both finite and infinitely dilute biomolecule concentrations in solution. These are then transformed to expressions in other interesting semi open ensembles in a stepwise fashion. This is in contrast with the general matrix approach which is usually applied to fully closed systems.[24] The approach and resulting equations are quite simple in nature and provide a clear physical picture of the experiments. In doing so, we hope to clarify the use of KB theory for the description of preferential interactions and establish a rigorous basis for the analysis of computer simulation data.

## Background

The notation used here follows the usual definitions where the subscripts 1, 2, and 3 refer to the primary solvent (usually water), the biomolecule, and cosolvent, respectively. In addition, many equations refer to the limit of an infinitely dilute biomolecule and therefore we denote these with a zero superscript. For simplicity, the following equations do not include the Donnan term, which is required for systems where the biomolecule and cosolvent share a common ion. The application of KB theory in this situation has been described previously and requires the use of an indistinguishable ion approach for the evaluation of the KB integrals.[14] The present results should be applicable to the majority of protein systems.

The original source of preferential interaction data in biomolecular systems involved equilibrium dialysis experiments where the chemical potentials ($\mu$) of the solvent and cosolvent are constant at a given temperature ($T$).[4,25] The preferential interaction is then defined as a measure of the change in cosolvent concentration on introduction of a biomolecule in a system open to either cosolvent, water, or both. The earliest results involved changes in molal concentrations in systems open to both cosolvent and water. In this case the preferential interaction is defined as[4]

$$\left(\frac{\partial m_3}{\partial m_2}\right)_{T,\mu_1,\mu_3} = \frac{M_2}{M_3}\left(\frac{\partial w_3}{\partial w_2}\right)_{T,\mu_1,\mu_3} = \frac{M_2}{M_3}\xi_{23} \qquad (1)$$

where $m_i = N_i/N_1$, $M_i$, $N_i$, and $w_i$ are the molality, molecular mass, number of moles, and weight molality of species $i$, respectively. Other definitions of the preferential interaction can be defined based on either a different choice of constant thermodynamic variables and/or different concentration scales and will be discussed later. The above preferential interaction is obtained indirectly from a determination of the change in mass density ($\rho$) of an open system evaluated at a finite biomolecule concentration. This is related to the preferential interaction by[4]

$$\left(\frac{\partial \rho}{\partial c_2}\right)_{T,\mu_1,\mu_3} = \frac{(1-\bar{v}_2\rho) + \xi_{23}(1-\bar{v}_3\rho)}{1-(\bar{v}_2-\xi_{23}\bar{v}_3)c_2} + \kappa_T\rho\left(\frac{\partial \Pi}{\partial c_2}\right)_{T,\mu_1,\mu_3} \qquad (2)$$

where $c_i = \rho_i M_i = N_i M_i/V$ is the mass concentration of $i$, $\bar{v}_i$ is the partial specific molar volume of $i$ ($= \bar{V}_i/M_i$), $\kappa_T$ is the isothermal compressibility of the system, and $\Pi$ is the (osmotic)

* Tel: 785-532-5109. Fax: 785-532-6666. E-mail: pesmith@ksu.edu.

10.1021/jp056100e CCC: $33.50 
Published on Web 01/25/2006

pressure. The above expression is exact. Usually, the equilibrium dialysis data is determined as a function of biomolecule concentration, and one then evaluates the limiting expression as the initial value of $c_2$ tends to zero,

$$\left(\frac{\partial \rho}{\partial c_2}\right)^{\mathrm{O}}_{T,\mu_1,\mu_3} = (1 - \bar{v}_2\rho^{\mathrm{O}}) + \xi_{23}(1 - \bar{v}_3\rho^{\mathrm{O}}) + RT\kappa_T\rho^{\mathrm{O}}/M_2 \quad (3)$$

The last term is small and often ignored.

Kirkwood−Buff theory provides relationships between particle number fluctuations and derivatives of the chemical potentials in the grand canonical ensemble where the volume, temperature, and chemical potential of all species are constant. The primary result used here is that[22,24]

$$\frac{RT}{V}\left(\frac{\partial N_i}{\partial \mu_j}\right)_{T,V,\mu_{k\neq j}} = \left(\frac{\partial \rho_i}{\partial \beta \mu_j}\right)_{T,\mu_{k\neq j}} = \rho_i\rho_j G_{ij} + \rho_i\delta_{ij} \quad (4)$$

where $G_{ij}$ is the Kirkwood−Buff integral between species $i$ and $j$, $\delta_{ij}$ is the Kroenecker delta function, and $\beta = 1/RT$. The KB integrals are defined in terms of the corresponding rdfs ($g_{ij}$) such that

$$G_{ij} = G_{ji} = 4\pi\int_0^\infty [g_{ij}^{\mu VT}(r) - 1]r^2\,\mathrm{d}r \quad (5)$$

An excess coordination number can be defined ($N_{ij} = \rho_j G_{ij} \neq N_{ji}$), which characterizes the excess number of $j$ molecules around a single $i$ molecule in a large volume of the open system above that observed in an equivalent volume in the reference solution.

KB theory provides expressions for thermodynamic properties in other ensembles by using suitable thermodynamic transformations. Several (four) expressions relating to the properties of binary solutions of 1 and 3 will be required for our analysis and are taken directly from Ben-Naim.[24] The partial molar volumes are given by

$$\bar{V}_i = \left(\frac{\partial \mu_i}{\partial P}\right)_{T,N_1,N_3} = \frac{1 + \rho_j(G_{jj} - G_{ij})}{\eta} \quad (6)$$

with $i$ and $j = 1$ or 3, and where $\eta = \rho_1 + \rho_3 + \rho_1\rho_3\,(G_{11} + G_{33} - 2G_{13})$. Derivatives of the chemical potentials of water and the cosolvent are provided by

$$\left(\frac{\partial \mu_3}{\partial m_3}\right)^{\mathrm{O}}_{T,P,m_2} = N_1\left(\frac{\partial \mu_3}{\partial N_3}\right)_{T,P,N_1} = -\frac{{\rho_1}^2 V}{\rho_3}\left(\frac{\partial \mu_1}{\partial N_3}\right)_{T,P,N_1} = \frac{{\rho_1}^2 RT}{\rho_3\eta} \quad (7)$$

in terms of molalities, where we have used the fact that $N_1$ is constant in closed systems. Molarity based derivatives are provided by

$$\left(\frac{\partial \mu_3}{\partial \rho_3}\right)_{T,P} = -\frac{\rho_1}{\rho_3}\left(\frac{\partial \mu_1}{\partial \rho_3}\right)_{T,P} = \frac{RT}{\rho_3(1 + \rho_3 G_{33} - \rho_3 G_{31})} \quad (8)$$

while mole fraction derivatives are given by

$$\left(\frac{\partial \mu_3}{\partial x_3}\right)_{T,P} = -\left(\frac{\partial \mu_3}{\partial x_1}\right)_{T,P} = -\frac{x_1}{x_3}\left(\frac{\partial \mu_1}{\partial x_3}\right)_{T,P} = \frac{RT(\rho_1 + \rho_3)^2}{\rho_3\eta} \quad (9)$$

All the above equations are exact and can be applied to the ternary system when the biomolecule concentration is very low.

**Results**

From the definition of the molality based preferential interaction in a system open to cosolvent and solvent one can write

$$\left(\frac{\partial m_3}{\partial m_2}\right)_{T,\mu_1,\mu_3} = \frac{\left(\frac{\partial m_3}{\partial \mu_2}\right)_{T,\mu_1,\mu_3}}{\left(\frac{\partial m_2}{\partial \mu_2}\right)_{T,\mu_1,\mu_3}} = \frac{\frac{1}{N_1}\left(\frac{\partial N_3}{\partial \mu_2}\right)_{T,\mu_1,\mu_3} - \frac{N_3}{{N_1}^2}\left(\frac{\partial N_1}{\partial \mu_2}\right)_{T,\mu_1,\mu_3}}{\frac{1}{N_1}\left(\frac{\partial N_2}{\partial \mu_2}\right)_{T,\mu_1,\mu_3} - \frac{N_2}{{N_1}^2}\left(\frac{\partial N_1}{\partial \mu_2}\right)_{T,\mu_1,\mu_3}} \quad (10)$$

and therefore, using the KB integrals defined in eq 4, it follows that

$$\left(\frac{\partial m_3}{\partial m_2}\right)_{T,\mu_1,\mu_3} = \frac{\rho_3(G_{23} - G_{21})}{1 + \rho_2(G_{22} - G_{21})} \quad (11)$$

for any biomolecule concentration. In the infinitely dilute biomolecule limit, this reduces to the numerator which has been derived previously:[14]

$$\left(\frac{\partial m_3}{\partial m_2}\right)^{\mathrm{O}}_{T,\mu_1,\mu_3} = \rho_3(G_{23} - G_{21}) = N_{23} - \frac{\rho_3}{\rho_1}N_{21} \quad (12)$$

Alternatively, if the cosolvent and biomolecule are measured using molar concentrations, one can write

$$\left(\frac{\partial \rho_3}{\partial \rho_2}\right)_{T,\mu_1,\mu_3} = \frac{\left(\frac{\partial \rho_3}{\partial \mu_2}\right)_{T,\mu_1,\mu_3}}{\left(\frac{\partial \rho_2}{\partial \mu_2}\right)_{T,\mu_1,\mu_3}} = \frac{\rho_3 G_{23}}{1 + \rho_2 G_{22}} \quad (13)$$

which again reduces to the numerator for infinitely dilute biomolecules,

$$\left(\frac{\partial \rho_3}{\partial \rho_2}\right)^{\mathrm{O}}_{T,\mu_1,\mu_3} = \rho_3 G_{23} = N_{23} \quad (14)$$

It is also possible to define the preferential interaction in terms of mole fractions ($x_i$). In this case the finite biomolecule concentration result is

$$\left(\frac{\partial x_3}{\partial x_2}\right)_{T,\mu_1,\mu_3} = \frac{\rho_3 G_{23} - x_3(1 + \rho_2 G_{22} + \rho_1 G_{21} + \rho_3 G_{23})}{1 + \rho_2 G_{22} - x_2(1 + \rho_2 G_{22} + \rho_1 G_{21} + \rho_3 G_{23})} \quad (15)$$

which reduces to the following expression for infinitely dilute biomolecule solutions:

$$\left(\frac{\partial x_3}{\partial x_2}\right)^{\mathrm{O}}_{T,\mu_1,\mu_3} = -x_3 + x_1\rho_3(G_{23} - G_{21}) \quad (16)$$

Equations 12, 14, and 16 represent expressions for the preferential interactions in terms of KB integrals relevant to equilibrium dialysis experiments obtained using the three different concentration scales.

Having established exact relationships for the different preferential interactions in the constant $T$, $V$, $\mu_1$, and $\mu_3$ ensemble in terms of KB integrals, we now make a connection to the

experimental density data. The solution mass density can be written

$$\rho = \frac{N_1M_1 + N_2M_2 + N_3M_3}{V} = \rho_1M_1 + \rho_2M_2 + \rho_3M_3 \quad (17)$$

which leads directly to an expression for the density derivative determined during an equilibrium dialysis experiment,

$$\left(\frac{\partial\rho}{\partial c_2}\right)_{T,\mu_1,\mu_3} = 1 + \left[\frac{M_1}{M_2}\left(\frac{\partial\rho_1}{\partial\mu_2}\right)_{T,\mu_1,\mu_3} + \frac{M_3}{M_2}\left(\frac{\partial\rho_3}{\partial\mu_2}\right)_{T,\mu_1,\mu_3}\right]\left(\frac{\partial\mu_2}{\partial\rho_2}\right)_{T,\mu_1,\mu_3} \quad (18)$$

which can be written in terms of KB integrals as

$$\left(\frac{\partial\rho}{\partial c_2}\right)_{T,\mu_1,\mu_3} = 1 + \frac{M_1}{M_2}\frac{\rho_1G_{21}}{1+\rho_2G_{22}} + \frac{M_3}{M_2}\frac{\rho_3G_{23}}{1+\rho_2G_{22}} \quad (19)$$

The above equation can be extended to include any number of diffusible components. Taking the infinitely dilute biomolecule limit one obtains

$$\left(\frac{\partial\rho}{\partial\rho_2}\right)^{O}_{T,\mu_1,\mu_3} = M_2 + c_1G_{21} + c_3G_{23} = M_2 + M_1N_{21} + M_3N_{23} \quad (20)$$

The above equation provides a very simple physical picture of the increase in density with biomolecule concentration in terms of the excess coordination numbers ($N_{ij}$), which we shall expand upon later. The individual values of $G_{21}$ and $G_{23}$ can be extracted from the density measurements using the KB results for the partial molar volume of the solute at infinite dilution in terms of properties of the reference solution,[15,24]

$$\overline{V_2}^{O} = RT\kappa_T - \rho_1\overline{V_1}G_{21} - \rho_3\overline{V_3}G_{23} = RT\kappa_T - N_{21}\overline{V_1} - N_{23}\overline{V_3} \quad (21)$$

A combination of eqs 6, 20, and 21 can be shown to be exactly equivalent to eq 3 as long as the osmotic pressure term is included. Equations 20 and 21 represent the primary results for the analysis of the density data in terms of KB integrals.

The determination of $G_{23}$ and $G_{21}$ is sufficient to quantify the preferential interactions in systems open to all diffusible species. However, preferential interactions in other ensembles open to just water or cosolvent require knowledge of the KB integrals characterizing the solution ($G_{11}$, $G_{33}$, and $G_{13}$), assuming the biomolecule concentration is low. Relationships between other preferential interaction coefficients can be determined by standard thermodynamic manipulation. For instance, it is possible to define a molality based preferential interaction with $P$ constant instead of $\mu_1$ constant by using the standard relationship between partial derivatives,

$$\left(\frac{\partial m_3}{\partial m_2}\right)_{T,\mu_1,\mu_3} = \left(\frac{\partial m_3}{\partial m_2}\right)_{T,P,\mu_3} + \left(\frac{\partial m_3}{\partial P}\right)_{T,m_2,\mu_3}\left(\frac{\partial\Pi}{\partial m_2}\right)_{T,\mu_1,\mu_3} \quad (22)$$

and therefore by using the Euler chain relationship we have

$$\left(\frac{\partial m_3}{\partial m_2}\right)_{T,P,\mu_3} = \left(\frac{\partial m_3}{\partial m_2}\right)_{T,\mu_1,\mu_3} + \left(\frac{\partial\mu_3}{\partial P}\right)_{T,m_2,m_3}\left(\frac{\partial m_3}{\partial\mu_3}\right)_{T,P,m_2}\left(\frac{\partial\Pi}{\partial m_2}\right)_{T,\mu_1,\mu_3} \quad (23)$$

The osmotic pressure derivative takes the following values for an infinitely dilute biomolecule,[4,24]

$$\left(\frac{\partial\Pi}{\partial m_2}\right)^{O}_{T,\mu_1,\mu_3} = \rho_1\left(\frac{\partial\Pi}{\partial\rho_2}\right)^{O}_{T,\mu_1,\mu_3} = x_1\left(\frac{\partial\Pi}{\partial x_2}\right)^{O}_{T,\mu_1,\mu_3} = \rho_1RT \quad (24)$$

Consequently, from eqs 6, 7, 12, and 24, one obtains that

$$\left(\frac{\partial m_3}{\partial m_2}\right)^{O}_{T,P,\mu_3} = \frac{\rho_3}{\rho_1} + \rho_3(G_{23} - G_{21} + G_{11} - G_{13}) \quad (25)$$

which is in agreement with our previous expression obtained from a different route,[14]

$$\left(\frac{\partial m_3}{\partial m_2}\right)^{O}_{T,P,\mu_3} = \rho_3(G_{23} - G_{21}) + \frac{\rho_3\overline{V_3}\eta}{\rho_1} = \rho_3(G_{23} - G_{21}) + \frac{\rho_3}{\phi_1}(RT\kappa_T - G_{13}) \quad (26)$$

where we have also used the fact that[24]

$$\overline{V_1V_3}\eta = RT\kappa_T - G_{13} \quad (27)$$

for a solution of components 1 and 3 only. One can also convert the molarity based preferential interactions into constant $P$ values using eqs 6, 8, 14, 24, and the molarity version of eq 23. In this case an additional thermodynamic manipulation of the derivative $(\partial\mu_i/\partial P)^{O}_{T,\rho_2,\rho_3} = (\partial\mu_i/\partial P)_{T,V,N_3}$ into $(\partial\mu_i/\partial P)_{T,N_1,N_3}$ has to be performed. The final result is the following expression:

$$\left(\frac{\partial\rho_3}{\partial\rho_2}\right)^{O}_{T,P,\mu_3} = \rho_3(G_{23} - G_{13}) \quad (28)$$

Converting the mole fraction based preferential interactions into constant $P$ values using eqs 6, 9, 16, 24, and the mole fraction version of eq 23 provides

$$\left(\frac{\partial x_3}{\partial x_2}\right)^{O}_{T,P,\mu_3} = x_1\rho_3(G_{23} - G_{21} + G_{11} - G_{13}) \quad (29)$$

Before leaving this section it should be noted that conditions corresponding to this ensemble are difficult to achieve experimentally and therefore direct evaluation of the above three preferential interactions is rare.

Preferential interactions can also be defined in the $T$, $P$, $\mu_1$ ensemble. Application of thermodynamic relationships converting $\mu_3$ to $P$ ($N_3$ is now constant) lead to

$$\left(\frac{\partial m_3}{\partial m_2}\right)^{O}_{T,P,\mu_1} = -1 + \rho_3(G_{23} - G_{21} + G_{13} - G_{33}) \quad (30)$$

which again is in agreement with our previous result[14]

$$\left(\frac{\partial m_3}{\partial m_2}\right)^{O}_{T,P,\mu_1} = \rho_3(G_{23} - G_{21}) - \overline{V_1}\eta = \rho_3(G_{23} - G_{21}) - \frac{\rho_3}{\phi_3}(RT\kappa_T - G_{13}) \quad (31)$$

Other preferential interactions in this ensemble are then given by

$$\left(\frac{\partial\rho_3}{\partial\rho_2}\right)^{O}_{T,P,\mu_1} = -1 + \rho_3(G_{23} - G_{33}) \quad (32)$$

and

$$\left(\frac{\partial x_3}{\partial x_2}\right)^{\mathrm{O}}_{T,P,\mu_1} = -1 + x_1\rho_3(G_{23} - G_{21} + G_{13} - G_{33}) \quad (33)$$

Contrary to the other two ensembles considered here, the preferential interaction is not zero at infinite dilution of the cosolvent and biomolecule. This arises from the fact that one can write $(\partial m_3/\partial m_2)_{T,P,\mu_1} = -(\partial\mu_1/\partial m_2)_{T,P,m_3}/(\partial\mu_1/\partial m_3)_{T,P,m_2}$, which indicates that the preferential interaction will change due to changes in the chemical potential of water on the introduction of the biomolecule, even in the absence of cosolvent.

Finally, we note the following relationship holds between the molality and molarity based preferential interactions in the constant $V$ ensemble,

$$\left(\frac{\partial \rho_3}{\partial \rho_2}\right)^{\mathrm{O}}_{T,\mu_1,\mu_3} = \phi_1\left(\frac{\partial m_3}{\partial m_2}\right)^{\mathrm{O}}_{T,\mu_1,\mu_3} - \rho_3(\overline{V_2}^{\mathrm{O}} - RT\kappa_T) \quad (34)$$

where $\varphi_i$ is the volume fraction of $i$. In the constant $P$ ensembles the relationship is given by

$$\left(\frac{\partial \rho_3}{\partial \rho_2}\right)^{\mathrm{O}}_{T,P,\mu_i} = \phi_1\left(\frac{\partial m_3}{\partial m_2}\right)^{\mathrm{O}}_{T,P,\mu_i} - \rho_3\overline{V_2}^{\mathrm{O}} \quad (35)$$

with $i$ equal to 1 or 3. The relationship between the molality and mole fraction definitions in any of the above ensembles is given by

$$\left(\frac{\partial x_3}{\partial x_2}\right)^{\mathrm{O}}_{T} = x_1\left(\frac{\partial m_3}{\partial m_2}\right)^{\mathrm{O}}_{T} - x_3 \quad (36)$$

and also between preferential interactions on the same concentration scale ($y$) but in different ensembles according to

$$\left(\frac{\partial y_3}{\partial y_2}\right)^{\mathrm{O}}_{T,\mu_1,\mu_3} = \phi_1\left(\frac{\partial y_3}{\partial y_2}\right)^{\mathrm{O}}_{T,P,\mu_3} + \phi_3\left(\frac{\partial y_3}{\partial y_2}\right)^{\mathrm{O}}_{T,P,\mu_1} \quad (37)$$

for $y$ equal to $m$ or $x$ and by

$$\left(\frac{\partial y_3}{\partial y_2}\right)^{\mathrm{O}}_{T,\mu_1,\mu_3} = \phi_1\left(\frac{\partial y_3}{\partial y_2}\right)^{\mathrm{O}}_{T,P,\mu_3} + \phi_3\left(\frac{\partial y_3}{\partial y_2}\right)^{\mathrm{O}}_{T,P,\mu_1} - \rho_3 RT\kappa_T \quad (38)$$

for $y$ equal to $\rho$. Equation 34 is a slightly different expression than that obtained by Eisenberg for open systems and corrects the small approximation made in the previous analysis.[4] Equations 35 and 36 have also been presented by Shulgin and Ruckenstein for semi open systems.[19] Equation 38 was originally presented by Anderson and more recently by Schurr et al.[11,20] It can also be determined rather easily from eq 7 by using eq 23 and the Gibbs−Duhem equation on any concentration scale, $y_1\partial\mu_1 + y_3\partial\mu_3 = 0$ at constant $T$ and $P$. Alternatively, one could use eqs 26 and 31 to arrive at the same result.

## Discussion

All the above expressions are exact. Although many are restricted to biomolecules at very low concentrations, this is the usual situation assumed in most experimental and theoretical studies. In principle, one could start with the finite biomolecule concentration expressions given in eqs 11, 13, and 15 and derive equations for the other ensembles by using KB expressions for the chemical potential derivatives and partial molar volumes in ternary systems. However, we have primarily restricted ourselves to the infinitely dilute biomolecule, as this provides the simplest physical picture of the nature of the preferential interactions through the KB integrals. Some aspects of the above results will now be discussed.

The equilibrium dialysis experiment corresponds to a system at constant $T$, $V$, and chemical potentials of all diffusible species. This situation corresponds to the grand canonical ensemble in statistical mechanics, which is the foundation of KB theory. In practice, the equilibrium dialysis experiments may actually involve a flexible dialysis bag, rather than a rigid dialysis chamber. Hence, the volume of the dialysis bag may change slightly. However, this does not correspond to a constant $T$, $P$, and chemical potential ensemble. In fact, this ensemble is meaningless as there are no criteria for equilibrium in this situation.[24] Hence, even if the dialysis bag initially swells slightly, the thermodynamics must correspond to that of a constant volume ensemble. Consequently, the relevant ensemble involves a constant volume, even though this subscript is usually absent in the common notation used for preferential interactions.

The preferential interaction corresponding to the equilibrium dialysis experiments is defined by eq 12 on the molality scale and by eq 14 on the molarity scale. It is clear that the molality scale definition probes changes in $m_3$ due to changes in the number of cosolvent molecules ($N_{23}$) and the number of water molecules ($N_{21}$) in the dialysis chamber, both of which can lead to a change in cosolvent molality. On the other hand, the molarity based definition probes only changes in the cosolvent concentration ($N_{23}$). Preferential exclusion of the cosolvent will result in a negative value of $N_{23}$, while preferential binding will produce a more positive value of $N_{23}$. This effect is enhanced for the molality based preferential interaction, as preferential binding of the cosolvent (positive $N_{23}$) must be accompanied by preferential exclusion of water (negative $N_{21}$) in order to keep the volume constant. Hence, molality based preferential interactions are always larger in magnitude than the corresponding molarity based definitions (see below).

The increase in density during the dialysis experiment observed on introduction of the biomolecule is given by eq 20. This is a much simpler and informative equation than that derived from purely thermodynamic arguments. The excess coordination numbers are, by definition, a measure of the increase in water or cosolvent in the open system above that in the reference solution. Obviously, the classical preferential interaction cannot be obtained directly from the density increment alone. However, a combination of the density increment and partial molar volume of the biomolecule can be used to extract values of $G_{23}$ and $G_{21}$, which then lead to the required preferential interaction on any concentration scale and in any ensemble, provided one knows the concentration-dependent values of $G_{11}$, $G_{33}$, and $G_{13}$ for the reference solution.

The preferential interactions at finite biomolecule concentrations involve the KB integral between biomolecules ($G_{22}$), which is also a measure of the change in osmotic pressure due to the presence of the biomolecule,[4,24]

$$\left(\frac{\partial \Pi}{\partial \rho_2}\right)_{T,\mu_1,\mu_3} = \rho_2\left(\frac{\partial \mu_2}{\partial \rho_2}\right)_{T,\mu_1,\mu_3} = \frac{RT}{1 + \rho_2 G_{22}} \quad (39)$$

In principle, a study of the density derivative as a function of biomolecule concentration could provide information on cosolvent modified biomolecule−biomolecule interactions, just as variations in osmotic pressure provide information concerning the interaction between biomolecules in pure water. However, it appears that no studies of this kind exist.

Record and co-workers have argued that the most useful definition of the preferential interaction is given at constant $T$, $P$, and $\mu_3$ using the molality scale.[26] This is primarily because one can use the Euler reciprocity relation (noting the fact that $N_1$ is constant) to write

$$\left(\frac{\partial m_3}{\partial m_2}\right)_{T,P,\mu_3} = -\frac{\left(\frac{\partial \mu_3}{\partial m_2}\right)_{T,P,m_3}}{\left(\frac{\partial \mu_3}{\partial m_3}\right)_{T,P,m_2}} = -\frac{\left(\frac{\partial \mu_2}{\partial m_3}\right)_{T,P,m_2}}{\left(\frac{\partial \mu_3}{\partial m_3}\right)_{T,P,m_2}} = -\left(\frac{\partial \mu_2}{\partial \mu_3}\right)_{T,P,m_2} \tag{40}$$

which closely represents the experimental conditions used for most biomolecule denaturation studies, as also noted previously by Timasheff.[8] The above relationship is not true of the other preferential interactions, or for the corresponding molarity based definitions, contrary to what has appeared in the literature.[17] Alternatively, we have argued that the most appropriate definition for analyzing computer simulation data from a statistical mechanics approach corresponds to the equilibrium dialysis conditions.[13,14] In this situation one can relate the denaturation equilibrium to changes in the pseudo chemical potential ($\mu^*$) as defined by Ben-Naim,[24]

$$\beta\mu_i = \beta\mu_i^* + \ln(\Lambda_i\rho_i) \tag{41}$$

where $\Lambda_i$ is the thermal de Broglie wavelength. The pseudo chemical potential isolates interactions between the solute and any solvent species and is equivalent to the change in Gibbs energy on transferring the solute from a fixed position in the gas phase to a fixed position in the solution, including changes to the internal partition function. Hence, the molarity (or number density) is the most natural definition for concentration in this case. One can apply this approach to protein denaturation by equating the chemical potentials of the native and denatured forms and then expressing the equilibrium constant ($K = \rho_D/\rho_N$) for protein unfolding as[14]

$$\ln K = -(\beta\mu_D^* - \beta\mu_N^*) \tag{42}$$

To investigate how the equilibrium constant changes with concentration one can take the derivative of the above equation and use the expression for the change in pseudo chemical potential given by[13,14,24]

$$-\left(\frac{\partial \beta\mu_2^*}{\partial \rho_3}\right)_{T,P}^{O} = \left(\frac{G_{23} - G_{21}}{1 + \rho_3(G_{33} - G_{13})}\right) \tag{43}$$

where we have used the fact that $(\partial\rho_3/\partial x_3)_{T,P}^{O} = (\rho_1 + \rho_3)^2\bar{V}_1$ to give

$$-\left(\frac{\partial \beta\Delta G^\circ}{\partial \rho_3}\right)_{T,P}^{O} = \left(\frac{\partial \ln K}{\partial \rho_3}\right)_{T,P}^{O} = \left(\frac{\Delta G_{23} - \Delta G_{21}}{1 + \rho_3(G_{33} - G_{13})}\right) \tag{44}$$

and the value of $\Delta G_{2j}$ ($= G_{Dj} - G_{Nj}$) denotes the difference in KB integrals for the denatured and native states of the biomolecule. The above expression completes the full analysis of the relationships between preferential interactions, KB integrals, and the effect of a cosolvent on the denaturation equilibrium. The last term in parentheses in eq 44 is constant within the range of validity of the $m$ value characterization of protein denaturation.[27]

Using eqs 8, 12, and 44, it is quite easy to show that molality based preferential interactions are related to the change in equilibrium according to

$$\left(\frac{\partial \ln K}{\partial \ln a_3}\right)_{T,P}^{O} = \left(\frac{\partial m_3}{\partial m_2}\right)_T^D - \left(\frac{\partial m_3}{\partial m_2}\right)_T^N = \Delta N_{23} - \frac{\rho_3}{\rho_1}\Delta N_{21} \tag{45}$$

The above equation has been derived from many approaches.[2,3,6,14,24,28] The expression is valid for the molality based preferential interactions in any ensemble as the differences between them involve KB integrals for the solution only. Alternatively, using eq 34 or 35 one can write for the molarity based preferential interactions that

$$\phi_1\left(\frac{\partial \ln K}{\partial \ln a_3}\right)_{T,P}^{O} = \left(\frac{\partial \rho_3}{\partial \rho_2}\right)_T^D - \left(\frac{\partial \rho_3}{\partial \rho_2}\right)_T^N + \rho_3\Delta\overline{V}_2^O \tag{46}$$

Previously, we have argued that the contribution from the change in biomolecule volume on denaturation is negligible compared to the change in preferential interaction,[14,15,29] even though the change in excluded volume may be significant.[30] In this case, the above equation can be written

$$\left(\frac{\partial \ln K}{\partial \ln a_3}\right)_{T,P}^{O} = \frac{\Delta N_{23}}{\phi_1} \tag{47}$$

and arises due to the conservation of volume on denaturation, which implies that any additional cosolvent molecules that interact with the denatured state must result in an equal volume of water molecules, which are excluded; i.e., $\Delta N_{23}$ and $\Delta N_{21}$ must have opposite signs. More precisely, from eq 21 we have[14]

$$\Delta N_{23}\overline{V}_3 + \Delta N_{21}\overline{V}_1 = 0 \tag{48}$$

as long as the change in biomolecule volume is small. Stabilization of the native state, also known as preferential hydration, occurs when $\Delta N_{23} < 0$ and therefore $\Delta N_{21} > 0$, which is a property displayed by most osmolytes.[28,31] The opposite is true for destabilization. Finally, the mole fraction based preferential interactions can be related to the derivative of the equilibrium constant by

$$x_1\left(\frac{\partial \ln K}{\partial \ln a_3}\right)_{T,P}^{O} = \left(\frac{\partial x_3}{\partial x_2}\right)_T^D - \left(\frac{\partial x_3}{\partial x_2}\right)_T^N \tag{49}$$

The above equations illustrate that the thermodynamic definition of cosolvent binding involves changes in both the cosolvent and water distribution as emphasized by Timasheff.[8]

The differences between the preferential interactions defined in the ensembles examined here involve terms corresponding to properties of the solution mixture of 1 and 3 alone. The significance of these will now be discussed. The clearest way to illustrate these differences is to consider the introduction of a particle, in place of the biomolecule, which occupies no volume and has no interactions with either the solvent or cosolvent. This implies that $G_{23}$ and $G_{21}$ are zero. In this case, the preferential interactions in the $T$, $\mu_1$, $\mu_3$ ensemble are all zero. However, the preferential interactions in the constant $P$ ensembles are non zero and are dependent on the properties of the solution mixture. The reason for this is that addition of the particle, which has kinetic energy according to the value of T, contributes to the pressure, which results in a small volume increase. This volume increase will change the chemical potential (activity) of the cosolvent and water molecules, giving rise to a contribution to the preferential interactions. This is not true in the constant volume ensemble. Clearly, if one is interested in interpreting preferential interactions in terms of

**TABLE 1: KB Expression for the Preferential Interaction Parameter Defined in Different Ensembles and Using Various Concentration Scales**

| | $T, \mu_1, \mu_3$ | $T, P, \mu_1$ | $T, P, \mu_3$ |
|---|---|---|---|
| $m$ | $\rho_3(G_{23} - G_{21})$ | $-1 + \rho_3(G_{23} - G_{21} + G_{13} - G_{33})$ | $(\rho_3/\rho_1) + \rho_3(G_{23} - G_{21} + G_{11} - G_{13})$ |
| $\rho$ | $\rho_3 G_{23}$ | $-1 + \rho_3(G_{23} - G_{33})$ | $\rho_3(G_{23} - G_{13})$ |
| $x$ | $-x_3 + x_1\rho_3(G_{23} - G_{21})$ | $-1 + x_1\rho_3(G_{23} - G_{21} + G_{13} - G_{33})$ | $x_1\rho_3(G_{23} - G_{21} + G_{11} - G_{13})$ |

favorable or unfavorable interactions with the solute, it is best to remove contributions that do not involve these interactions. This is precisely the reason the pseudo chemical potential is chosen: it removes contributions to the chemical potential of the biomolecule arising from the volume properties of the solution, i.e., the liberational free energy or entropy.[24] Hence, in our opinion the equilibrium dialysis based preferential interactions provide the most useful data for developing atomic level descriptions of the relevant molecular interactions.

The situation corresponding to a computer simulation is illustrated in Figure 1.[14,32] Analysis of simulation data involves the determination of the rdfs or coordination numbers for the cosolvent and water surrounding the central biomolecule to a distance ($R$) greater than that for which the rdf deviates from unity. One can consider a virtual dialysis membrane located at this distance such that within this fixed volume the cosolvent and solvent molecules are free to enter or leave and their concentrations deviate from that in the bulk solution in order to equalize the chemical potentials that are perturbed by the presence of the biomolecule. In principle, the KB integrals can be calculated by using the center of mass of the biomolecule as the origin. Indeed, this must be performed if absolute values of the KB integrals are to be determined for comparison with experiment. However, in many cases the distribution of solvent and cosolvent has been determined from the protein surface not the center of mass.[15,33,34] This is acceptable for determining $\rho_3(G_{23} - G_{21})$, as the volume excluded by the biomolecule cancels in this case. However, the molarity based preferential interactions, which only require $G_{23}$, cannot be calculated using the surface atoms as the origin, as this omits the space occupied by the biomolecule and requires a rather difficult normalization procedure to determine the corresponding rdf, $g_{23}(r)$. Consequently, the molarity based definition of the preferential interaction is not well suited for the analysis of simulation data.

Some of the equations presented above have appeared before in the literature, although not always in terms of KB integrals.

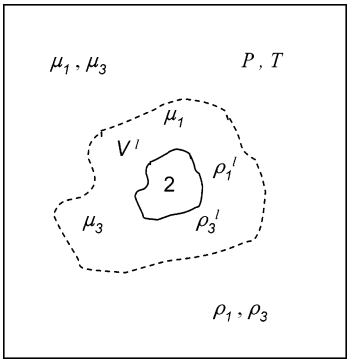


**Figure 1.** Schematic illustration of the analysis of computer simulation data. A fixed biomolecule solute (2) perturbs the chemical potential of the water (1) and cosolvent (3) molecules in the local region defined by $V^l$. Molecules then move to and from the NPT bath to equalize the chemical potential between the local region and the particle bath. The local and bulk solvent can be considered to be separated by a virtual dialysis membrane (dashed line) permeable to both cosolvent and water located at a distance $R$ from the surface. The excess coordination numbers $N_{2j}$ measure the change in local concentration over several solvation shells.

To our knowledge, an analysis of the density increment has not appeared, neither have expressions for the preferential interaction in fully open systems at finite biomolecule concentrations in terms of KB integrals, nor has the present approach been used to determine the expressions in different semi open ensembles. Equations 34 and 38 are similar to other literature expressions except for the compressibility term.[4,11,19] While this is small and therefore negligible for most situations, it is required to provide the exact relationships between the different preferential interactions that are possible due to the KB approach. Equation 28 is different from that quoted by Shimizu.[17]

The preferential interactions defined using different concentration scales and in different ensembles are summarized in Table 1. To illustrate the similarities and differences between the various preferential interactions a simple model can be used. The derivative in eq 44 should be constant within the range of validity of the $m$-value approach typically used to characterize protein denaturation. We will assume this is true for each form of the protein. The molality based preferential interaction in the open system is then given by

$$\left(\frac{\partial m_3}{\partial m_2}\right)^{\mathrm{O}}_{T,\mu_1,\mu_3} = \rho_3(G_{23} - G_{21}) = m\rho_3[1 + \rho_3(G_{33} - G_{13})] \tag{50}$$

where $m$ is positive for preferential binding of the cosolvent. Using approximate values for the KB integrals and molecular volumes, one finds the results provided in Figure 2. Cleary the molality based preferential interactions provide very similar results, especially for systems open to the cosolvent. In contrast, changing the concentration scale has a much larger effect. In particular, the molarity based definition provides large negative values. This is a direct consequence of the fact that $G_{23}$ in eq 14 includes the excluded volume effect of the protein which

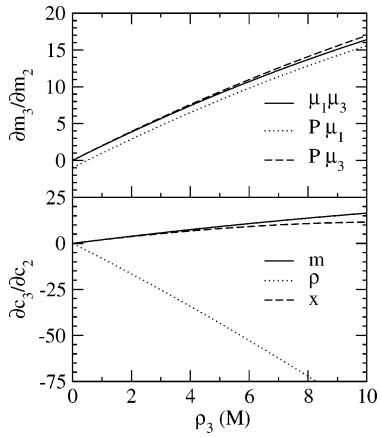


**Figure 2.** Preferential interaction parameters defined using the different concentration scales and in different ensembles all at constant $T$. The molality based parameters as a function of ensemble (top). The preferential interaction parameters obtained using the different concentration scales (c) in the open ($T$, $\mu_1$, $\mu_3$) ensemble (bottom). The calculations were based on eqs 21 and 50 and the expressions in Table 1. Values used for the above calculations were $G_{11} = -V_1 = -18$ cm$^3$/mol, $G_{33} = -V_3 = -54$ cm$^3$/mol, $G_{13} = 1/2\ (G_{11} + G_{33})$, $V_2 = 10000$ cm$^3$/mol, and $m = 2.0$ M$^{-1}$.

results in a large negative contribution. This dominates over the contribution from the cosolvent interaction with the protein. Obviously, it is more sensible to use a definition of the preferential interaction which is positive for preferential binding and negative for preferential exclusion of the cosolvent.

Finally, it should be noted that while many workers have used the term preferential interaction to describe the change in cosolvent concentration with biomolecule concentration, this is not a universal definition. Timasheff refers to the preferential interaction parameter as that defined by $(\partial\mu_3/\partial m_2)_{T,P,m_3} = (\partial\mu_2/\partial m_3)_{T,P,m_2}$, and the term preferential binding as that defined by $(\partial m_3/\partial m_2)_{T,P,\mu_3}$, although this is assumed to be equal to $(\partial m_3/\partial m_2)_{T,\mu_1,\mu_3}$ within a negligible error.[8] More recently, Schurr et al. have defined a preferential interaction $-(\partial\mu_2^o/\partial\mu_3)_{T,P,\rho_2\to 0}$ in a closed system which involves the use of the standard chemical potential and not the total chemical potential as used in most applications.[20] This is equivalent to the pseudo chemical potential if the molar concentration scale is used.

## Conclusions

Some very simple equations have been developed that relate the change in solution density observed in an equilibrium dialysis experiment to changes in the distribution of cosolvent and water molecules around a biomolecule solute in terms of KB integrals. Exact relationships are then presented for the preferential interaction defined in three different ensembles and using three different concentration scales (summarized in Table 1). It is argued that the analysis of computer simulation data, obtained from either closed or open systems, is directly analogous to the situation observed in equilibrium dialysis experiments and is best interpreted using the molality based definition of the preferential interaction given by $(\partial m_3/\partial m_2)^{O}_{T,\mu_1,\mu_3} = \rho_3(G_{23} - G_{21})$. In this situation the preferential interaction quantifies changes in the relative distribution of the cosolvent and solvent which result from specific interactions with the biomolecule only ($G_{23}$ and $G_{21}$), rather than including additional non-specific changes related to the properties of the solution ($G_{11}$, $G_{33}$, and $G_{13}$).

**Acknowledgment.** Acknowledgment is made to the donors of the Petroleum Research Fund, administered by the ACS, for partial support of this research. I would like to thank one of the reviewers for correcting a mistake in an earlier version of the manuscript.

## References and Notes

(1) Scatchard, G. *J. Am. Chem. Soc.* **1946**, *68*, 2315−2319.
(2) Wyman, J., Jr. *Adv. Protein Chem.* **1964**, *19*, 223−286.
(3) Casassa, E. F.; Eisenberg, H. *Adv. Protein Chem.* **1964**, *19*, 287−395.
(4) Eisenberg, H. *Biological Macromolecules and Polyelectrolytes in Solution*, Clarendon Press: Oxford, 1976.
(5) Tanford, C. *Adv. Protein Chem.* **1968**, *23*, 121−282.
(6) Tanford, C. *J. Mol. Biol.* **1969**, *39*, 539−544.
(7) Timasheff, S. N. *Biochemistry* **1992**, *31*, 9857−9864.
(8) Timasheff, S. N. *Adv. Protein Chem.* **1998**, *51*, 355−432.
(9) Schellman, J. A. *Annu. Rev. Biophys. Biophys. Chem.* **1987**, *16*, 115−137.
(10) Schellman, J. A. *Biophys. Chem.* **1990**, *37*, 121−140.
(11) Anderson, C. F.; Felitsky, D. J.; Hong, J.; Record, M. T., Jr. *Biophys. Chem.* **2002**, *101−102*, 497−511.
(12) Anderson, C. F.; Record, M. T., Jr. *Biophys. Chem.* **2004**, *112*, 165−175.
(13) Chitra, R.; Smith, P. E. *J. Phys. Chem. B* **2001**, *105*, 11513−11522.
(14) Smith, P. E. *J. Phys. Chem. B* **2004**, *108*, 18716−18724.
(15) Aburi, M.; Smith, P. E. *J. Phys. Chem. B* **2004**, *108*, 7382−7388.
(16) Shimizu, S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 1195−1199.
(17) Shimizu, S. *J. Chem. Phys.* **2004**, *120*, 4989−4990.
(18) Shimizu, S.; Boon, C. L. *J. Chem. Phys.* **2004**, *121*, 9147−9155.
(19) Shulgin, I. L.; Ruckenstein, E. *J. Chem. Phys.* **2005**, *123*, 054909.
(20) Schurr, J. M.; Rangel, D. P.; Aragon, S. R. *Biophys. J.* **2005**, *89*, 2258−2276.
(21) Shulgin, I. L.; Ruckenstein, E. *Biophys. J.* **2006**, *90*, 704−707.
(22) Kirkwood, J. G.; Buff, F. P. *J. Chem. Phys.* **1951**, *19*, 774−777.
(23) Ben-Naim, A. *J. Chem. Phys.* **1977**, *67*, 4884−4890.
(24) Ben-Naim, A. *Statistical Thermodynamics for Chemists and Biochemists*; Plenum Press: New York, 1992.
(25) Reisler, E.; Haik, Y.; Eisenberg, H. *Biochemistry* **1977**, *16*, 197−203.
(26) Anderson, C. F.; Courtenay, E. S.; Record, M. T., Jr. *J. Phys. Chem. B* **2002**, *106*, 418−433.
(27) Greene, R. F., Jr.; Pace, C. N. *J. Biol. Chem.* **1974**, *249*, 5388−5393.
(28) Parsegian, V. A.; Rand, R. P.; Rau, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3987−3992.
(29) Smith, P. E. *Biophys. Chem.* **2005**, *113*, 299−302.
(30) Schellman, J. A. *Biophys. J.* **2003**, *85*, 108−125.
(31) Bolen, D. W.; Baskakov, I. V. *J. Mol. Biol.* **2001**, *310*, 955−963.
(32) Smith, P. E. *J. Phys. Chem. B* **2004**, *108*, 16271−16278.
(33) Baynes, B. M.; Trout, B. L. *J. Phys. Chem. B* **2003**, *107*, 14058−14067.
(34) Tang, K. E. S.; Bloomfield, V. A. *Biophys. J.* **2002**, *82*, 2876−2891.