Fluctuation theory of molecular association and conformational equilibria

Yuanfang Jiao and Paul E. Smith[a)]
Department of Chemistry, Kansas State University, Manhattan, Kansas 66506, USA

(Received 25 March 2011; accepted 27 May 2011; published online 5 July 2011)

General expressions relating the effects of pressure, temperature, and composition on solute association and conformational equilibria using the fluctuation theory of solutions are provided. The expressions are exact and can be used to interpret experimental or computer simulation data for any multicomponent mixture involving molecules of any size and character at any composition. The relationships involve particle-particle, particle-energy, and energy-energy correlations within local regions in the vicinity of each species involved in the equilibrium. In particular, it is demonstrated that the results can be used to study peptide and protein association or aggregation, protein denaturation, and protein-ligand binding. Exactly how the relevant fluctuating properties may be obtained from experimental or computer simulation data are also outlined. It is shown that the enthalpy, heat capacity, and compressibility differences associated with the equilibrium process can, in principle, be obtained from a single simulation. Fluctuation based expressions for partial molar heat capacities, thermal expansions, and isothermal compressibilities are also provided. © *2011 American Institute of Physics*. [doi:10.1063/1.3601342]

I. INTRODUCTION

Studies of protein denaturation play a central role in our efforts to understand the forces that stabilize protein structures and assemblies.[1] Proteins can be denatured by changes in temperature, pressure, and solution composition (cosolvents and pH) in closed systems[2–4] and by osmotic pressure or stress in open systems.[5] Experimentally, the thermodynamics of protein denaturation are well established and a large volume of data on protein denaturation is available. More recently, a growing amount of thermodynamic data concerning the factors that influence peptide and protein aggregation has also been determined.[6–9] Unfortunately, it is extremely difficult to relate this thermodynamic data to specific interactions with, or effects on, either the native or denatured forms. Consequently, the application of computer simulations for the study of protein denaturation has become increasingly more common. In principle, an atomic level picture of interactions and structural changes can be elucidated from these computer simulations. However, in practice this has proven difficult for two main reasons. First, one cannot typically follow the denaturation equilibrium with current computational resources, with the possible exception of a few extreme examples,[10,11] and thereby evaluate the equilibrium constant (K). Second, it is not clear exactly how to extract from a simulation the relevant properties of a protein that relate to thermal or pressure denaturation—unless one has already solved the first problem.

For example, simulations of a protein folding/unfolding equilibrium to a degree where a precise equilibrium constant can be determined are extremely rare. Hence, obtaining a reliable equilibrium constant for protein denaturation over a range of pressure, temperature, or cosolvent concentrations is essentially impossible at present. The temperature denaturation or folding of small proteins or peptides can be studied more easily. In particular, peptide simulations using replica exchange techniques,[10] essentially provide the enthalpy (first derivative of K) and heat capacity (second derivative of K) changes via an analysis of the equilibrium constant as a function of temperature. However, these simulations remain computationally expensive for larger proteins in explicit solvent and it is still not clear, for instance, exactly how one should decompose or interpret the resulting enthalpy changes.[12,13] In addition, the heat capacity changes associated with thermal denaturation are also typically difficult to quantify by simulation.[14–16]

Pressure denaturation simulations are also problematic. Thermodynamics relates the effect of pressure on the equilibrium constant to a difference in volume between the native and denatured forms.[3,17] The determination of protein volumes from a simulation are either numerically challenging (direct evaluation of the volume change on addition of the protein), or require somewhat subjective definitions of the protein volume which may or may not be correct.[18–20] Furthermore, the second derivative of K with respect to pressure is usually interpreted in terms of a difference in compressibility between the two protein forms.[6] This is often estimated using a protein volume fluctuation formula which is technically only valid for the total volume of a macroscopic closed system at constant pressure and temperature.[21,22] Clearly, a more rigorous and computationally efficient approach is desirable.

The simulation of protein denaturation by the addition of cosolvents such as urea has also received attention.[23–26] In fact, the effects of cosolvents and, in particular, how these effects may be determined from simulation in a way that can be compared with experimental data are essentially solved.[27] In our opinion, the most rigorous analysis of computer

a)Author to whom correspondence should be addressed. Electronic mail: pesmith@ksu.edu. Tel.: 785-532-5109. Fax: 785-532-6666.

simulation data involves the Kirkwood-Buff (KB) theory of solution mixtures to relate changes in the equilibrium constant to the relative distribution of cosolvent and solvent molecules in the vicinity of both protein forms.[28–33] A similar type of approach would be useful for the interpretation of pressure and thermal denaturation simulations. This is a major goal of the present study.

The KB theory is a general theory of solution mixtures which relates molecular distributions in solution to the thermodynamic properties of that solution.[34–36] We were initially drawn to this type of approach as the resulting expressions are exact and involved quantities that can be easily obtained from computer simulations. The KB theory quantifies the molecular distributions in terms of Kirkwood-Buff integrals (KBIs), involving integrals over the corresponding intermolecular radial distribution functions, or as particle number fluctuations corresponding to local regions of the solution. Hence, it is also referred to as the fluctuation theory (FT) of solutions. The application of KB or FT to understand solution mixtures has provided valuable insight into their behavior.[37] Recently, we have extended the analysis of solution mixtures, building on the work of Buff and Brout,[38] and also Debenedetti,[39–42] by determining particle-energy and energy-energy fluctuations obtained from experimental enthalpy of mixing, thermal expansion, and heat capacity data.[43]

The KB theory has also been applied to understand chemical equilibria. Several studies have used expressions derived for thermodynamically independent infinitely dilute solutes to study the equilibrium between two infinitely dilute forms.[28,29,44,45] More rigorous work by O'Connell and co-workers has provided general expressions for reactive systems, including sequential reactions, in terms of both total and direct correlation functions after explicitly including the material balance constraint resulting in a modified grand canonical distribution function.[46,47] Ben-Naim derived expressions for the effect of a cosolvent on association equilibria in a primary solvent using an alternative approach where the chemical equilibrium conditions were imposed on the usual multicomponent KB expressions.[48] More recently, we derived general expressions which could be applied to interpret real experimental data for complex systems in a variety of ensembles using a slightly different approach, from which the original Ben-Naim result for the effect of a cosolvent could be obtained.[49] The cosolvent effects were related to particle-particle fluctuations in the vicinity of each form present in the equilibrium. Here, we wish to extend this type of approach to provide general fluctuation based expressions which can be used to interpret the effects of temperature and pressure on association equilibria in solution, with a specific emphasis on protein denaturation. All the expressions provided are exact and can be used to interpret either experimental or simulation data concerning pressure, temperature, or cosolvent denaturation. In particular, first and second derivatives of the equilibrium constant with respect to pressure, temperature, and cosolvent concentration are developed which can (in principle) be determined from a single simulation, thereby eliminating the need for computationally intensive multiple simulations.

II. THEORY

Here, we develop relationships describing an equilibrium process in a system, which may be under a variety of different thermodynamic constraints, in terms of particle-particle, particle-energy, and energy-energy fluctuations. All ensemble averages—signified in this study by angular brackets—correspond to that of the Grand Canonical Ensemble. They can be used to describe properties of other ensembles which possess the same average thermodynamic quantities, chemical potential, pressure, etc., and then represent fluctuations observed for local regions within these systems.

In the Grand Canonical Ensemble the independent variables are the volume (V), the temperature ($\beta = 1/RT$), and the set of chemical potentials ($\{\beta\mu\}$). Hence, a differential for the number density of i particles ($\rho_i = \langle N_i\rangle/V$) in a solution mixture can be written as a function of the independent variables, such that

$$d\ln\rho_i = \left(\frac{\partial \ln\rho_i}{\partial\beta}\right)_{\{\beta\mu\}} d\beta + \sum_j \left(\frac{\partial \ln\rho_i}{\partial\,\beta\mu_j}\right)_{\beta,\{\beta\mu\}'} d\,\beta\mu_j, \tag{1}$$

where the summation is over all j components of the mixture, and the prime indicates that all chemical potentials except for the one of interest are held constant. Using the statistical mechanical equations associated with the grand canonical (μVT) ensemble one can show that the above derivatives are given by the following ensemble averages[38]:

$$\left(\frac{\partial \ln\rho_i}{\partial\,\beta\mu_j}\right)_{\beta,\{\beta\,\mu\}'} = \frac{\langle\delta N_i\delta N_j\rangle}{\langle N_i\rangle} = \delta_{ij} + N_{ij},$$
$$\left(\frac{\partial \ln\rho_i}{\partial\beta}\right)_{\{\beta\,\mu\}} = -\frac{\langle\delta N_i\delta E\rangle}{\langle N_i\rangle} = -F_{\mu,i}, \tag{2}$$

where δ_{ij} is the Kroenecker delta function, $\delta X = X - \langle X\rangle$ is the deviation in X from the ensemble average X for each member of the ensemble, and E is the total internal energy of each member of the ensemble.

The N_{ij} values represent particle-particle number fluctuations within a local volume of the solution of interest and are the focus of the KB theory of solutions. They are related to the traditional KBIs between species i and j which can be expressed in terms of the corresponding radial probability distribution (g_{ij}),[34]

$$N_{ij} = \rho_j G_{ij} = 4\pi\rho_j \int_0^\infty [g_{ij}(r) - 1]r^2 dr \tag{3}$$

or as particle-particle number fluctuation densities

$$B_{ij} = \rho_i(\delta_{ij} + N_{ij}) = \frac{\langle\delta N_i\delta N_j\rangle}{V}. \tag{4}$$

The use of radial probability distributions imparts a useful physical picture to the N_{ij}'s. Namely, the change in the number of j particles resulting from the introducing an i particle to the reference volume, from the number of j particles observed in the same volume of the bulk solution.[50] The F_μ's correspond to particle number-energy correlations within the same local region of interest, and can be used to characterize solution mixtures in an analogous fashion to the KB theory.[38,43]

Previously, we used a more convenient property for the analysis of experimental data on solution mixtures that provided useful relationships for particle-energy and energy-energy fluctuations in terms of experimentally accessible excess thermodynamic properties.[43] This was achieved by defining an excess energy (ε), such that

$$\varepsilon = E - \sum_j N_j E_j^0, \qquad F_i = \frac{\langle \delta N_i \delta \varepsilon \rangle}{\langle N_i \rangle}, \tag{5}$$

where E_j^0 is the average internal energy per particle (molar energy) in the pure liquid j, or for any reference state, at the temperature and average pressure of interest. Unfortunately, when studying systems at constant pressure and temperature the above approach leads to rather cumbersome expressions. Much simpler, and totally equivalent, results can be obtained by defining an alternative fluctuating quantity, such that

$$\varepsilon_P = E - \sum_j N_j \bar{h}_j, \qquad F_{P,i} = \frac{\langle \delta N_i \delta \varepsilon_P \rangle}{\langle N_i \rangle}, \tag{6}$$

where $\bar{h}_j$ is the partial molar enthalpy of species j at the particular composition of interest. The value of ε_P measures the difference between the instantaneous energy E of a local region containing $\{N\}$ molecules and the average total enthalpy of the same set of molecules in the bulk solution. While it seems somewhat unusual to subtract an enthalpy term from an energy term, the following analysis is aided greatly by this substitution, especially for closed systems at constant pressure. Furthermore, while the use of partial molar enthalpies is unfortunate as they cannot be obtained from experiment, they can be easily extracted from computer simulation data, and the corresponding F_P's can be obtained from experimental data as we illustrate below. Combining Eqs. (1), (2), and (6) one finds

$$d \ln \rho_i = -F_{P,i} d\beta + \sum_j (\delta_{ij} + N_{ij})(d\beta \mu_j - \bar{h}_j d\beta) \tag{7}$$

for any species i in a mixture of j components at any composition. If T (β) is constant the above equations reduce to a series of differentials corresponding to the traditional KB theory of solution mixtures.[51,52] Using the above substitution, one finds that the isothermal compressibility (κ_T), isobaric thermal expansion coefficient (α_P), and constant pressure heat capacity (C_P) of any multicomponent solution mixture are then given by

$$\begin{aligned} RT\kappa_T &= -RT \left(\frac{\partial \ln V}{\partial P} \right)_{T,\{N\}} = \sum_j (\delta_{ij} + N_{ij}) \bar{V}_j, \\ RT^2 \alpha_p &= RT^2 \left(\frac{\partial \ln V}{\partial T} \right)_{P,\{N\}} = -F_{P,i} = -\sum_i x_i F_{P,i} \\ &= -\sum_i \phi_i F_{P,i}, \\ RT^2 C_P &= RT^2 \left(\frac{\partial \langle H \rangle}{\partial T} \right)_{P,\{N\}} = \langle \delta \varepsilon_P \, \delta \varepsilon_P \rangle, \end{aligned} \tag{8}$$

where $\bar{V}_i$ and $\phi_i = \rho_i \bar{V}_i$ are the partial molar volume and volume fraction of i, respectively. The above expressions are much simpler than the equivalent expressions provided previously for the thermal expansion and heat capacity.[38,43,53] It should be noted that, using this formulism, the same value of F_P is obtained for each thermodynamically independent species. However, different values will be obtained for the thermodynamically dependent solute forms (see later discussion).

In the following analysis, we shall also use the pseudo chemical potential (μ_i^*) concept, and its associated enthalpy (h_i^*) and volume (V_i^*), to indicate how one can extract relevant quantities from available experimental data. The pseudo chemical potential approach centers on the statistical mechanical definition of chemical potential and thereby eliminates the need for standard states.[36] The pseudo chemical potential is similar to the excess chemical potential used in computer simulations with the only difference being a term related to the internal partition function of the species. Using Eqs. (1) and (2) and rearranging one finds

$$\begin{aligned} d\,\beta \mu_i^* \equiv d\,\beta \mu_i - d \ln \Lambda_i^3 \rho_i &= -3d \ln \Lambda_i + F_{\mu,i} d\beta \\ &- \sum_j N_{ij} d\,\beta \mu_j, \end{aligned} \tag{9}$$

where Λ_i is the thermal de Broglie wavelength and is proportional to $T^{-1/2}$. From the above equation one obtains

$$\begin{aligned} V_i^* &= \bar{V}_i - RT\kappa_T = -\sum_j N_{ij} \bar{V}_j, \\ h_i^* &= \bar{h}_i - RT^2 \alpha_p - \frac{3}{2} RT = -\frac{3}{2} RT + F_{\mu,i} - \sum_j N_{ij} \bar{h}_j, \\ d\mu_i^* &= d\mu_i - RT d \ln \rho_i = -\sum_j N_{ij} d\mu_j \qquad T \text{ constant}, \end{aligned} \tag{10}$$

where the last equation can be used for changes in composition of the system. The volume term (V_i^*) can be determined experimentally, while only changes in the enthalpy and chemical potentials for various processes can be evaluated experimentally (see later discussion). In principle, all three properties can be obtained directly from computer simulations.

Equation (7) represents a series of source equations which can be used to obtain expressions for various properties of solution mixtures in terms of number-number, number-energy, and (later) energy-energy correlations characterizing local microscopic regions within the solution. Our primary focus here is that of chemical equilibria involving an associating solute, or a solute which can undergo a change in conformation. We examine a system with a solute (2) in a primary solvent (1) which may contain any number of additional co-solvents (3,4, ...). The solute is in equilibrium between two forms. One form being a monomer (M) and the other an aggregate (A) containing n monomers. This equilibrium is described by an equilibrium constant (K), such that

$$nM \rightarrow A, \qquad K = \frac{\rho_A}{\rho_M^n}, \qquad d \ln K = d \ln \rho_A - n d \ln \rho_M. \tag{11}$$

We note that the equilibrium constant involves the actual number densities (molar concentrations) present at the equilibrium

composition of interest, and not the activities approximated by concentrations—as is often the case in biological problems. The relationships between the number of solutes, number of monomers, and number of aggregates are given by

$$N_M + nN_A = N_2, \quad f_M = \frac{N_M}{N_2}, \quad f_A = \frac{nN_A}{N_2}. \tag{12}$$

Furthermore, the equilibrium conditions dictate that the following relationships:

$$\mu_A = n\mu_M = n\mu_2, \qquad d\mu_A = nd\mu_M = nd\mu_2, \tag{13}$$

must be obeyed. Using Eq. (7), the above relationships, and the approach outlined previously,[49] it is relatively easy to show that

$$d\ln K = -(F_{P,A} - nF_{P,M})d\beta + \sum_i (N_{Ai} - nN_{Mi})(d\,\beta\mu_i - \bar{h}_i d\beta), \tag{14}$$

where the summation is over all thermodynamically independent i components including component 2. The exact meaning of the N_{A2} and N_{M2} values has been discussed previously.[49] This equation, in combination with the Gibbs-Duhem (GD) expression,

$$SdT - VdP + \sum_i N_i d\mu_i = 0, \tag{15}$$

enables one to develop a complete picture of how pressure, temperature, and solution composition affect the above equilibrium in terms of particle and energy fluctuations within local regions of the solution under the conditions of interest.

III. RESULTS

In the following subsections we provide expressions (first and second derivatives) describing the effect of pressure, temperature, and composition on a general equilibrium process in solution. The equilibrium can involve molecules of any size and character at any composition in a variety of ensembles. In addition, we provide expressions for the simplest case—a two state conformational equilibrium in a single solvent at infinite dilution of the solute—such as often used to understand protein folding or denaturation.

A. The effect of hydrostatic pressure on chemical equilibria

Taking derivatives of Eq. (14) with respect to pressure while keeping temperature and composition constant immediately provides

$$\left(\frac{\partial \ln K}{\partial P}\right)_{\beta,\{N\}} = \beta \sum_i (N_{Ai} - nN_{Mi})\bar{V}_i. \tag{16}$$

The partial molar volumes can also be expressed in terms of KBIs if desired. Hence, in the absence of specific affinities between the two forms and any i species, then if A is "smaller" than n monomers there will be an excess of each i molecule in the vicinity of A compared to the vicinity of n monomers. Therefore, each term on the right hand side will be positive and an increase in P will increase K and thus favor the A form.

In many cases, one is also interested in the "compressibility" of the process as manifested in the second derivative of the equilibrium constant with respect to pressure. To develop expressions for this derivative we first note that

$$\left(\frac{\partial \ln K}{\partial P}\right)_{\beta,\{N\}} = \beta \sum_i \bar{V}_i \left(\frac{\partial \ln K}{\partial\, \beta\mu_i}\right)_{\beta,\{\beta\mu\}'}. \tag{17}$$

Consequently, taking pressure derivatives of the above expression, then interchanging the order of differentiation on the right hand side, one finds the second pressure derivative can be written as

$$\left(\frac{\partial^2 \ln K}{\partial P^2}\right)_{\beta,\{N\}} = -\beta \sum_i (N_{Ai} - nN_{Mi})\bar{V}_i\bar{\kappa}_{T,i} + \beta^2 \sum_{i,j} \bar{V}_i\bar{V}_j \times \left(\frac{\partial (N_{Aj} - nN_{Mj})}{\partial\, \beta\mu_i}\right)_{\beta,\{\beta\mu\}'}, \tag{18}$$

where $\bar{\kappa}_{T,i} = -(\partial \ln \bar{V}_i/\partial P)_{T,\{N\}}$ is the partial molar isothermal compressibility of species i, and $(\partial \bar{V}_j/\partial\, \beta\mu_i)_{\beta,\{\beta\,\mu\}'} = 0$. Fluctuation based expressions for the partial molar compressibilities are provided in Appendix B. The required derivative can be obtained from the equations of the grand canonical ensemble in the same manner as before to provide

$$\left(\frac{\partial N_{Aj}}{\partial\, \beta\mu_i}\right)_{\beta,\{\beta\,\mu\}'} = \frac{\langle \delta N_A \delta N_i \delta N_j\rangle}{\langle N_A\rangle} - N_{Ai}N_{Aj} = N_{Aij} - N_{Ai}N_{Aj}. \tag{19}$$

Hence, the above derivative provides information on triplet particle number fluctuations (N_{Aij}) in the region of interest. Similar expressions are obtained for the corresponding derivatives of N_{Mj}. The final expression for the second derivative of the equilibrium with respect to pressure for a solution containing any number of components is therefore

$$\left(\frac{\partial^2 \ln K}{\partial P^2}\right)_{\beta,\{N\}} = -\beta \sum_i (N_{Ai} - nN_{Mi})\bar{V}_i\bar{\kappa}_{T,i} + \beta^2 \sum_{i,j} \bar{V}_i\bar{V}_j[N_{Aij} - N_{Ai}N_{Aj} - n(N_{Mij} - N_{Mi}N_{Mj})]. \tag{20}$$

In principle, all the terms present in Eqs. (16) and (20) can be determined reasonably easily from a single computer simulation.

In order to extract the specific terms associated with each form (A or M), one requires knowledge of the solute volume and how it varies with pressure. From Eq. (10) we have

$$V_2^* = -\sum_j N_{2j}\bar{V}_j = -\sum_j (f_A N_{Aj} + f_M N_{Mj})\bar{V}_j, \tag{21}$$

which provides the net sum over all j terms. To extract each N_{Aj} term would require knowledge of all the partial molar volumes in the mixture. The pressure derivative of the solute volume can then be expressed as

$$\left(\frac{\partial V_i^*}{\partial P}\right)_{\beta,\{N\}} = -\sum_j \left[\left(\frac{\partial N_{ij}}{\partial P}\right)_{\beta,\{N\}} \bar{V}_j - N_{ij}\bar{V}_j\bar{\kappa}_{T,j}\right], \tag{22}$$

where the required derivative is given by

$$\left(\frac{\partial N_{ij}}{\partial P}\right)_{\beta,\{N\}} = -N_{ij}\kappa_T + \beta \sum_k \frac{\langle \delta N_i \delta N_j \delta N_k \rangle}{\langle N_i \rangle} \bar{V}_k \quad (23)$$

and was evaluated by treating $\langle \delta N_i\, \delta N_j \rangle$ as a function of $\{\beta\mu\}$, β, and V, in a similar manner to Eq. (1) (see Appendix A). The final expression for the change in solute volume with pressure is then

$$\left(\frac{\partial V_2^*}{\partial P}\right)_{\beta,\{N\}} = RT(\kappa_T)^2 + \sum_i N_{2i}\bar{V}_i\bar{\kappa}_{T,i} - \beta \sum_{i,j} \bar{V}_i\bar{V}_j[f_A N_{Aij} + f_M N_{Mij}]. \quad (24)$$

We note that Eqs. (16), (20), and (21) could be simplified by using the notation $V_A^* = -\sum_j N_{Aj}\bar{V}_j$, as suggested by Eq. (10), although this requires some care (see later discussion).

A situation of common interest is that of an infinitely dilute solute in a single solvent. For instance, experimental data concerning the pressure denaturation ($N \rightarrow D$) of proteins is often interpreted in terms of a Taylor expansion of the standard free energy change for unfolding,[6]

$$-\beta\,\Delta\Delta G^{0,\infty} = \ln\left(\frac{K}{K_0}\right)^{\infty} \approx \left(\frac{\partial \ln K}{\partial P}\right)^{\infty}_{T,m_2} \Delta P + \frac{1}{2!}\left(\frac{\partial^2 \ln K}{\partial P^2}\right)^{\infty}_{T,m_2} (\Delta P)^2 + O[(\Delta P)^3], \quad (25)$$

where the subscript 0 refers to the reference pressure, $\Delta P = P - P_0$, and all derivatives are for an infinitely dilute protein evaluated at P_0 (usually 1 bar). In this case the first derivative is expressed in terms of fluctuating quantities by

$$\left(\frac{\partial \ln K}{\partial P}\right)^{\infty}_{T,m_2} = \beta(G^{\infty}_{D1} - G^{\infty}_{N1}) = \beta V_1^0(\langle N_1\rangle_D - \langle N_1\rangle_N) = -\beta\,\Delta V^{0,\infty}, \quad (26)$$

where $m_2 = \rho_2/\rho_1$, is a dimensionless molality, and the last equality arises from standard thermodynamics under conditions that activity and concentration are equivalent. The above expression has been presented before,[28] although it is seldom used for the analysis of simulation data on pressure denaturation.[54] The subscript D (or N) indicates an ensemble average within the same fixed volume of solution surrounding a single D (or N) molecule. Hence, the KBIs essentially quantify the volume of each solute form in terms of the number of solvent molecules that can be accommodated in the same fixed volume of solution.

The second derivative for pressure denaturation is given by

$$\left(\frac{\partial^2 \ln K}{\partial P^2}\right)^{\infty}_{T,m_2} = -\beta(G^{\infty}_{D1} - G^{\infty}_{N1})\kappa_T^0 + (\beta V_1^0)^2\big[\langle \delta N_1 \delta N_1\rangle_D - \langle \delta N_1 \delta N_1\rangle_N - \big[(N^{\infty}_{D1})^2 - (N^{\infty}_{N1})^2\big]\big]. \quad (27)$$

It is comforting to note the similarity of several terms in the above expression to the compressibility equation for a pure solvent, involving fluctuations in the number of solvent particles.[36] The above expressions indicate that the volume and "compressibility" associated with the equilibrium, and thereby the proteins themselves, are actually properties solely related to the water distribution in the vicinity of each form. In the present fluctuation based approach the protein volume does not enter directly into the expressions, and hence one does not have to be concerned as to exactly how this is defined or represented when analyzing computer simulation data. However, clearly the number of waters and their fluctuations will mimic the protein volume and fluctuations in the protein volume (see later discussion).

In order to extract specific fluctuations associated with each form, one requires information concerning the protein solute volume and how the volume changes with pressure. Namely,

$$V_2^{*,\infty} = -G^{\infty}_{21} = -f_D G^{\infty}_{D1} - f_N G^{\infty}_{N1} \quad (28)$$

and

$$\left(\frac{\partial V_2^*}{\partial P}\right)^{\infty}_{T,m_2} = RT\left(\kappa_T^0\right)^2 + G^{\infty}_{21}\kappa_T^0 - \beta\left(V_1^0\right)^2[f_D\langle \delta N_1 \delta N_1\rangle_D + f_N\langle \delta N_1 \delta N_1\rangle_N - \langle \delta N_1 \delta N_1\rangle_0]. \quad (29)$$

The zero subscript indicating an ensemble average obtained for the same volume of pure solvent. Hence, if one knows K and the compressibility of the pure solvent, together with $V_2^{*,\infty}$ for a series of pressures of interest, then the individual KBIs and fluctuations can be extracted from experimental data.

B. The effect of temperature on chemical equilibria at constant pressure

One of the most common ways to affect a chemical equilibrium involves changes in temperature at constant pressure. To our knowledge there are no fluctuation based expressions currently available for describing the associated enthalpy and heat capacity changes for chemical equilibria. Taking derivatives of Eq. (14) with respect to β while keeping pressure and composition constant one immediately obtains

$$\left(\frac{\partial \ln K}{\partial \beta}\right)_{P,\{N\}} = -(F_{P,A} - nF_{P,M}). \quad (30)$$

The above expression is valid for any number of components at any concentration. In order to develop second derivatives of the equilibrium constant with respect to β we first note that

$$\left(\frac{\partial \ln K}{\partial \beta}\right)_{P,\{N\}} = \left(\frac{\partial \ln K}{\partial \beta}\right)_{\{\beta\,\mu\}} + \sum_i \bar{h}_i \left(\frac{\partial \ln K}{\partial\, \beta\mu_i}\right)_{\beta,\{\beta\,\mu\}'}. \quad (31)$$

The second derivative is then obtained from the derivative of Eq. (31) after a change in the order of differentiation,

$$\left(\frac{\partial^2 \ln K}{\partial \beta^2}\right)_{P,\{N\}} = -RT^2 \sum_i (N_{Ai} - nN_{Mi})\bar{c}_{P,i} - \left(\frac{\partial(F_{P,A} - nF_{P,M})}{\partial \beta}\right)_{\{\beta\,\mu\}} - \sum_i \bar{h}_i \left(\frac{\partial(F_{P,A} - nF_{P,M})}{\partial\, \beta\mu_i}\right)_{\beta,\{\beta\,\mu\}'}, \quad (32)$$

where $\bar{c}_{P,i}$ is the partial molar constant pressure heat capacity of species i (see Appendix B for the relevant expressions). All the derivatives in the above expression can be evaluated in terms of local fluctuations using the equations of the grand canonical ensemble. Noting that $(\partial \bar{h}_j/\partial\, \beta\mu_i)_{\beta,\{\beta\,\mu\}'} = (\partial \bar{h}_j/\partial\beta)_{\{\beta\,\mu\}} = 0$, the required results are given by the following derivatives:

$$\left(\frac{\partial F_{P,A}}{\partial\, \beta}\right)_{\{\beta\,\mu\}} = -\frac{\langle \delta N_A \delta\, \varepsilon_P \delta E\rangle}{\langle N_A\rangle} + F_{P,A}F_{\mu,A} \tag{33}$$

and

$$\left(\frac{\partial F_{P,A}}{\partial\, \beta\mu_i}\right)_{\beta\{\beta\,\mu\}'} = \frac{\langle \delta N_A \delta N_i \delta\, \varepsilon_P\rangle}{\langle N_A\rangle} - F_{P,A}N_{Ai}, \tag{34}$$

which together provide

$$\left(\frac{\partial F_{P,A}}{\partial\beta}\right)_{\{\beta\,\mu\}} + \sum_i \bar{h}_i \left(\frac{\partial F_{P,A}}{\partial\, \beta\mu_i}\right)_{\beta,\{\beta\,\mu\}'} = -\frac{\langle \delta N_A \delta\, \varepsilon_P \delta\, \varepsilon_P\rangle}{\langle N_A\rangle} + F_{P,A}^2. \tag{35}$$

Similar expressions are found for the monomer form. Hence, the final result for a solution containing any number of components is given by

$$\left(\frac{\partial^2 \ln K}{\partial\beta^2}\right)_{P,\{N\}} = -RT^2 \sum_i (N_{Ai} - nN_{Mi})\bar{c}_{P,i} + \frac{\langle \delta N_A \delta\, \varepsilon_p \delta\, \varepsilon_p\rangle}{\langle N_A\rangle} - (F_{P,A})^2 - n\left[\frac{\langle \delta N_M \delta\, \varepsilon_P \delta\, \varepsilon_P\rangle}{\langle N_M\rangle} - (F_{P,M})^2\right]. \tag{36}$$

Again, it is satisfying that the above expression involves local energy fluctuations in the vicinity of both forms of the solute, which are typically characteristic of heat capacities.[36]

In order to extract the specific terms relating to each form, one requires knowledge of the thermal expansion and how the pseudo enthalpy varies with temperature. From Eq. (8) one has

$$RT^2\alpha_P = -F_{P,2} = -f_A F_{P,A} - f_M F_{P,M} \tag{37}$$

and from Eq. (10),

$$\left(\frac{\partial h_i^*}{\partial\beta}\right)_{P,\{N\}} = \frac{3}{2}(RT)^2 + \left(\frac{\partial F_{\mu,i}}{\partial\beta}\right)_{P,\{N\}} - \sum_j \left[\bar{h}_j \left(\frac{\partial N_{ij}}{\partial\beta}\right)_{P,\{N\}} - RT^2 N_{ij}\bar{c}_{P,j}\right]. \tag{38}$$

The two constant pressure derivatives can be evaluated using the same approach as used for Eq. (23) (see Appendix A) to provide

$$\left(\frac{\partial h_2^*}{\partial\beta}\right)_{p,\{N\}} = -RT^2 c_{P,2}^* = \frac{3}{2}(RT)^2 + (RT^2\alpha_P)^2 - f_A\frac{\langle \delta N_A \delta\, \varepsilon_P \delta\, \varepsilon_P\rangle}{\langle N_A\rangle} - f_M\frac{\langle \delta N_M \delta\, \varepsilon_P \delta\, \varepsilon_P\rangle}{\langle N_M\rangle} + RT^2\sum_j N_{2j}\bar{c}_{P,j} \tag{39}$$

Analogous expressions can be obtained for the other solution components if required.

Focusing again on an infinitely dilute protein solute in a single solvent, one finds a Taylor expansion of the equilibrium constant for a simple two state denaturation around a reference temperature provides the usual relationship,

$$\ln\left(\frac{K}{K_0}\right)^{\infty} \approx \left(\frac{\partial \ln K}{\partial T}\right)_{P,m_2}^{\infty}\Delta T + \frac{1}{2!}\left(\frac{\partial^2 \ln K}{\partial T^2}\right)_{P,m_2}^{\infty}(\Delta T)^2 + O[(\Delta T)^3], \tag{40}$$

where the first derivative is given by

$$\left(\frac{\partial \ln K}{\partial T}\right)_{P,m_2}^{\infty} = (RT^2)^{-1}\left(F_{P,D}^{\infty} - F_{P,N}^{\infty}\right) = (RT^2)^{-1}\Delta H^{0,\infty} \tag{41}$$

and provides an expression for the standard enthalpy change in terms of fluctuations in the local solution properties. This could be particularly useful for the analysis of computer simulation data. The value of $F_{P,D}$ at infinite dilution represents the relative enthalpy of the denatured form and for pairwise additive potentials is given by

$$F_{P,D}^{\infty} = \langle \varepsilon_P\rangle_D - \langle \varepsilon_P\rangle_0 = \left\langle E_{DD} + E_{D1} + E_{11} - N_1 E_1^0\right\rangle_D + PV_D^{*,\infty} - E_2^0, \tag{42}$$

which contains terms one would expect, namely, the intra and intermolecular energies E_{ij}, together with a less obvious additional term $N_1E_1^0$, which is non-negligible. The last term will cancel for differences between the F's even when $n \neq 1$. The above result can be written in a slightly more noticeable form using an alternative energy fluctuation term,

$$\left(\frac{\partial \ln K}{\partial\beta}\right)_{P,m_2}^{\infty} = -\left(F_D^{\infty} - F_N^{\infty}\right) + P\left(G_{D1}^{\infty} - G_{N1}^{\infty}\right) = -\Delta E^{0,\infty} - P\Delta V^{0,\infty}, \tag{43}$$

where we have used the definition of ε presented in Eq. (5). The volume term is the same as that derived for pressure denaturation (Eq. (26)), while the energy term provides a fluctuation expression for the energy contribution to the enthalpy change. The second derivative of the equilibrium constant is given by

$$\left(\frac{\partial^2 \ln K}{\partial T^2}\right)_{P,m_2}^{\infty} = (RT^2)^{-1}\Delta C_p^0 = -(RT^2)^{-1}C_{p,m}^0\left(N_{D1}^{\infty} - N_{N1}^{\infty}\right) + (RT^2)^{-2}\left[\langle \delta\, \varepsilon_P \delta\, \varepsilon_P\rangle_D - \langle \delta\, \varepsilon_P \delta\, \varepsilon_P\rangle_N - \left[\left(F_{P,D}^{\infty}\right)^2 - \left(F_{P,N}^{\infty}\right)^2\right]\right] \tag{44}$$

and provides a route to heat capacity changes associated with the chemical equilibrium in terms of local fluctuations. In the majority of cases, the volume changes associated with the equilibrium will be negligible in solution, compared to the corresponding energy changes, and it is often safe to ignore the PV_i terms (≈2 J/mol for water) and use $\varepsilon_P = \varepsilon$ in Eqs. (41)–(44).

Given an expression for the pseudo enthalpy of an infinitely dilute solute in a single solvent provided by Eq. (10) we have

$$h_2^{*,\infty} = -\frac{3}{2}RT + F_{\mu,2}^{\infty} - N_{21}^{\infty}H_1^0 = -\frac{3}{2}RT + F_{P,2}^{\infty} = -\frac{3}{2}RT + f_A F_{P,D}^{\infty} + f_M F_{P,N}^{\infty} \quad (45)$$

and from Eq. (39),

$$\left(\frac{\partial h_2^*}{\partial \beta}\right)_{P,m_2}^{\infty} = -RT^2 c_{p,2}^{*,\infty} = \frac{3}{2}(RT)^2 + \left(RT^2\alpha_p^0\right)^2 + RT^2 N_{21}^{\infty} C_{P,m}^0 - [f_D \langle \delta\varepsilon_P \delta\varepsilon_P\rangle_D + f_N \langle \delta\varepsilon_P \delta\varepsilon_P\rangle_N - \langle \delta\varepsilon_P \delta\varepsilon_P\rangle_0]. \quad (46)$$

Hence, if one knows K, the thermal expansion and heat capacity of the pure solvent, together with $h_2^{*,\infty}$ for a relevant process over a series of temperatures of interest, then the individual F_P's and fluctuations in ε_P can be extracted from experimental data.

Before leaving this section, we note that occasionally the pressure-temperature cross derivative may be useful and can be obtained from Eq. (16) to give

$$\left(\frac{\partial}{\partial\beta}\left(\frac{\partial \ln K}{P}\right)_{\beta,m_2}\right)_{P,m_2}^{\infty} = \left(G_{D1}^{\infty} - G_{N1}^{\infty}\right)\left(1 - T\alpha_P^0\right) - \beta V_1^0\left[\langle \delta N_1 \delta\varepsilon_P\rangle_D - \langle \delta N_1 \delta\varepsilon_P\rangle_N - \left[F_{P,D}^{\infty}N_{D1}^{\infty} - F_{P,N}^{\infty}N_{N1}^{\infty}\right]\right] \quad (47)$$

for protein denaturation at infinite dilution.

C. The effect of temperature on chemical equilibria at constant volume

Alternatively, the change in temperature could be performed for an equilibrium process under conditions of constant volume. While this is not a common situation, it is included here for completeness. In this case the general expression becomes

$$\left(\frac{\partial \ln K}{\partial\beta}\right)_{V,\{N\}} = -(F_{P,A} - nF_{P,M}) - \frac{T\alpha_P}{\kappa_T}\sum_i (N_{Ai} - nN_{Mi})\,\bar{V}_i, \quad (48)$$

where we have used the following standard thermodynamic relationship:

$$\left(\frac{\partial\,\beta\mu_i}{\partial\beta}\right)_{V,\{N\}} = \bar{h}_i - \bar{V}_i\frac{T\alpha_P}{\kappa_T}. \quad (49)$$

All the terms in the above equation represent average properties of the solution mixture and not fluctuating quantities. The change in equilibrium constant can be expressed in a far simpler form if one defines an alternative fluctuating property. We choose

$$\left(\frac{\partial \ln K}{\partial\beta}\right)_{V,\{N\}} = -(F_{V,A} - nF_{V,M}), \quad (50)$$

where

$$F_{V,i} = \frac{\langle \delta N_i \delta\varepsilon_V\rangle}{\langle N_i\rangle}, \qquad \varepsilon_V = E - \sum_j N_j\left(\bar{h}_j - \bar{V}_j\frac{T\alpha_P}{\kappa_T}\right). \quad (51)$$

If required, the thermal expansion and compressibility can be expressed in terms of fluctuating quantities. It should also be noted that using this definition, one can show that

$$RT^2 C_V = RT^2\left(\frac{\partial\langle E\rangle}{\partial T}\right)_{V,\{N\}} = \langle \delta\varepsilon_V \delta\varepsilon_V\rangle, \quad (52)$$

which is much simpler than previous expressions for C_V.[38,53]

Second derivatives of the equilibrium constant will eventually lead to an expression for the constant volume heat capacity change associated with the equilibrium. First, we note that

$$\left(\frac{\partial \ln K}{\partial\beta}\right)_{V,\{N\}} = \left(\frac{\partial \ln K}{\partial\beta}\right)_{\{\beta\,\mu\}} + \sum_i\left(\bar{h}_i - \bar{V}_i\frac{T\alpha_P}{\kappa_T}\right) \times \left(\frac{\partial \ln K}{\partial\,\beta\mu_i}\right)_{\beta,\{\beta\,\mu\}'} \quad (53)$$

and hence,

$$\left(\frac{\partial^2 \ln K}{\partial\beta^2}\right)_{V,\{N\}} = -RT^2\sum_i (N_{Ai} - nN_{Mi})\,\bar{c}_{V,i} - \left(\frac{\partial(F_{V,A} - nF_{V,M})}{\partial\beta}\right)_{\{\beta\,\mu\}} - \sum_i\left(\bar{h}_i - \bar{V}_i\frac{T\alpha_P}{\kappa_T}\right) \times \left(\frac{\partial(F_{V,A} - nF_{V,M})}{\partial\,\beta\mu_i}\right)_{\beta,\{\beta\,\mu\}'} \quad (54)$$

using the same approach as before. Fluctuation based expressions for $\bar{c}_{V,i}$ can be found in Appendix B. The derivatives in the above expression are analogous to Eqs. (33)–(35) and lead to the final result

$$\left(\frac{\partial^2 \ln K}{\partial\beta^2}\right)_{V,\{N\}} = -RT^2\sum_i (N_{Ai} - nN_{Mi})\,\bar{c}_{V,i} + \frac{\langle \delta N_A \delta\varepsilon_V \delta\varepsilon_V\rangle}{\langle N_A\rangle} - (F_{V,A})^2 - n\left[\frac{\langle \delta N_M \delta\varepsilon_V \delta\varepsilon_V\rangle}{\langle N_M\rangle} - (F_{V,M})^2\right], \quad (55)$$

which takes the same form as the constant P expression, although the fluctuating quantities are clearly different.

D. The effect of cosolvents on chemical equilibria

The effect of cosolvents on chemical equilibria has been the subject of many KB related studies.[27] Recently, we provided a general multicomponent expression in terms of chemical potential derivatives.[49] A similar expression, but using molarities instead of molalities, can be obtained from Eq. (14) after taking derivatives with respect to the (molar) concentration of species j keeping pressure, temperature, and the number of all other species constant,

$$\left(\frac{\partial \ln K}{\partial \ln \rho_j}\right)_{P,T,\{m\}'} = \sum_i (N_{Ai} - nN_{Mi})\mu_{ij} = \sum_{i\neq 1} (\Gamma_{Ai} - n\Gamma_{Mi})\mu_{ij}, \quad (56)$$

where $\mu_{ij} = (\partial \mu_i / \partial \ln \rho_j)_{P,T,\{m\}'}$ and the last summation excludes the primary solvent after elimination of $d\mu_1$ using the GD expression at constant T and P. The Γs are defined by

$$\Gamma_{Aj} = \rho_j (G_{Aj} - G_{A1}) = N_{Aj} - m_j N_{A1} \quad (57)$$

and can be considered as preferential binding parameters which quantify the excess binding of species j relative to that of the primary solvent in the vicinity of each solute form. The chemical potential derivatives can also be expressed in terms of KBIs if desired.[34,51] For an infinitely dilute solute in a primary solvent the addition of a single cosolvent (3) results in a change to the equilibrium provided by[29,45]

$$\left(\frac{\partial \ln K}{\partial \ln \rho_3}\right)^{\infty}_{P,T,m_2} = \left(\Gamma^{\infty}_{A3} - n\Gamma^{\infty}_{M3}\right)\mu_{33} \quad (58)$$

with

$$\mu_{33} = \beta \left(\frac{\partial \mu_3}{\partial \ln \rho_3}\right)^{\infty}_{P,T,m_2} = \frac{1}{1 + N_{33} - N_{13}}. \quad (59)$$

Note that the above derivative is different (molarity versus molality) than used in the traditional notation. Derivatives using other concentration scales can be found using the relationships provided by standard thermodynamics and KB theory

$$\left(\frac{\partial \ln \rho_3}{\partial \ln m_3}\right)^{\infty}_{P,T,m_2} = \phi_1 = \frac{1 + N_{33} - N_{13}}{1 + N^{+}_{33}},$$
$$\left(\frac{\partial \ln x_3}{\partial \ln m_3}\right)^{\infty}_{P,T,m_2} = x_1 \quad (60)$$

with $N^{+}_{ij} = N_{ij} + m_j(1 + N_{11} - N_{i1} - N_{j1})$. The above result (Eq. (58)) was also obtained by Ben-Naim using a different approach and alternative definitions of the equilibrium constant and cosolvent concentration.[48]

In order to isolate the binding to either form, one can take one of two approaches. First, rearranging Eq. (10), then eliminating $d\mu_1$ using the GD equation at constant T and P, provides an expression valid only for systems at constant temperature and pressure,

$$-d\mu^{*}_i = -RT d\ln y_i = \sum_{j\neq 1} \Gamma_{ij} d\mu_j, \quad (61)$$

where y_i is the molar activity coefficient and μ_i* is the pseudo chemical potential of species i ($\neq 1$). Derivatives of this expression then provide the following:

$$-\left(\frac{\partial \mu^{*}_2}{\partial \mu_3}\right)^{\infty}_{P,T,m_2} = -\left(\frac{\partial \ln y_2}{\partial \ln a_3}\right)^{\infty}_{P,T,m_2} = \Gamma^{\infty}_{23} = f_A \Gamma^{\infty}_{A3} + f_M \Gamma^{\infty}_{M3}. \quad (62)$$

Therefore, experimental data concerning the behavior of the solute activity coefficient at low concentrations can be used to extract values for Γ_{23}, etc. Alternatively, equilibrium dialysis studies that measure density changes in osmotic systems on the introduction of a non-diffusible solute (such as a protein) provide[55]

$$\left(\frac{\partial m_3}{\partial m_2}\right)^{\infty}_{T,\mu_1,\mu_3} = \Gamma^{\infty}_{23} = f_A \Gamma^{\infty}_{A3} + f_M \Gamma^{\infty}_{M3} \quad (63)$$

and also enable the isolation of the various Γ values. Finally, before leaving this section we note that typical cosolvent denaturation studies monitor the change in equilibrium constant as a function of cosolvent molarity,

$$\left(\frac{\partial \ln K}{\partial \rho_3}\right)^{\infty}_{P,T,m_2} = \frac{1}{\rho_3} \frac{\Gamma^{\infty}_{D3} - \Gamma^{\infty}_{N3}}{1 + N_{33} - N_{13}} = m \quad (64)$$

from which a fluctuation based expression for the protein m-value is obtained.[56]

It is possible to determine second derivatives of the equilibrium constant with respect to cosolvent concentration. Indeed, nonlinear cosolvent effects are observed.[57] The general expression is

$$\rho_j^2 \left(\frac{\partial^2 \ln K}{\partial \rho_j^2}\right)_{P,T,\{m\}'} = \sum_i (N_{Ai} - nN_{Mi})\mu_{ijj} + \sum_{i,k} \mu_{ij}\mu_{kj}[N_{Aik} - N_{Ai}N_{Ak} - n(N_{Mik} - N_{Mi}N_{Mk})], \quad (65)$$

where $\mu_{ijj} = (\partial \mu_{ij} / \partial \ln \rho_j)_{P,T,\{m\}'}$. Some of the above terms can be eliminated using the GD equation and Eq. (10). The most useful relationship is obtained for the denaturation of an infinitely dilute protein solute in a mixture of solvent (1) and a single cosolvent (3). Here, one finds

$$\rho_3^2 \left(\frac{\partial^2 \ln K}{\partial \rho_3^2}\right)^{\infty}_{P,T,m_2} = \sum_{i=1,3} (N_{Ai} - nN_{Mi})\mu_{i33} + \mu^2_{33}\left[\left\langle \delta\Gamma^{\infty}_{D3}\delta\Gamma^{\infty}_{D3}\right\rangle_D - \left\langle \delta\Gamma^{\infty}_{N3}\delta\Gamma^{\infty}_{N3}\right\rangle_N - \left[\left(\Gamma^{\infty}_{D3}\right)^2 - \left(\Gamma^{\infty}_{N3}\right)^2\right]\right], \quad (66)$$

where $\Gamma^{\infty}_{D3} = \langle N_3 \rangle_D - m_3 \langle N_1 \rangle_D$ and $\delta\Gamma^{\infty}_{D3} = \delta N_3 - m_3 \delta N_1$, which corresponds to a fluctuation in the binding parameter, i.e., cosolvent and water distributions, in the vicinity of a single denatured form. To our knowledge, general fluctuation expressions for the μ_{ijj} derivatives are not available. However, based on calculations described in Sec. IV, we suspect the contribution from the first term on the right hand side will be negligible under ambient conditions.

E. Osmotic systems

Many equilibria of biological importance occur under osmotic (or cellular) conditions. In addition, the study of osmotic pressures in protein solutions is interesting in that it provides information concerning protein-protein interactions.[58] An expression for the change in osmotic pressure (Π) on the addition of a biomolecule can be obtained from Eqs. (7) and (15),

$$\left(\frac{\partial \Pi}{\partial \rho_2}\right)_{\beta,\{\beta\mu\}'} = \frac{RT}{1+N_{22}} \tag{67}$$

and was noted in the original KB study.[34] This expression is valid for systems containing any number of additional cosolvents as long as their chemical potentials are held constant. Higher derivatives of the osmotic pressure provide information on higher order correlations between solute molecules. For instance, the second derivative of the osmotic pressure is given by

$$\left(\frac{\partial^2 \Pi}{\partial \rho_2^2}\right)_{\beta,\{\beta\mu\}'} = \frac{RT}{\rho_2(1+N_{22})}\left[1-\frac{\langle \delta N_2 \delta N_2 \delta N_2\rangle}{\langle N_2\rangle (1+N_{22})^2}\right] \tag{68}$$

and includes information concerning triplet solute correlations. Kirkwood and Buff showed that these higher order terms can form part of a series expansion which reduces to the McMillan-Mayer theory of solutions when the derivatives are obtained at infinite dilution of the solute.[34] However, the above expressions are valid for any solution composition. For chemical equilibria in an osmotic system one finds from Eq. (14) that

$$\begin{aligned}\left(\frac{\partial \ln K}{\partial \beta}\right)^{\infty}_{\rho_2,\{\beta\mu\}'} &= -\left(F^{\infty}_{P,A} - nF^{\infty}_{P,M}\right) - \left(N^{\infty}_{A1} - nN^{\infty}_{M1}\right) H^0_1 \\ &= -\left(F^{\infty}_{\mu,A} - nF^{\infty}_{\mu,M}\right)\end{aligned} \tag{69}$$

for an infinitely dilute solute in an osmotic solution. This result is to be expected as the conditions are essentially those of the grand canonical ensemble. Second derivatives can be obtained and provide the same expression as found in Eq. (44), but where ε_P is replaced by E and F_P is replaced by F_μ, and there is no term involving the heat capacity.

IV. DISCUSSION

The expressions provided here can be used to interpret data concerning chemical equilibria. In particular, the individual terms can be extracted under favorable conditions where the relevant experimental data are available. Alternatively, they can be used to analyze computer simulation data. In the latter case fluctuations of the form $\langle \delta X\, \delta Y\, \delta Z\rangle$ need to be evaluated from the trajectory. In principle, these can be expressed in terms of correlation functions in a similar way to the usual KBIs.[34] However, in practice, this is actually more cumbersome than evaluating the fluctuations within the local regions directly. The expressions are all exact. We are currently using computer simulations and applying Eqs. (26), (41), and (58) to study the effects of temperature, pressure, and composition on a simple conformational equilibrium in pure water.[59] The numerical results support the validity of the expressions presented here. We also note that, while primarily developed for the description of solution mixtures, the expressions are also valid for gaseous systems.

One of the more interesting results obtained in this study is the expression given in Eq. (41) which provides a route to the standard enthalpy change for association or denaturation of a solute at infinite dilution. Simulated enthalpy changes are difficult to determine and generally require multiple simulations of the equilibrium constant at different temperatures. The ability to calculate enthalpy changes for these processes from a single computer simulation is therefore particularly attractive. It is also clear from Eq. (7) that this is a general result for any equilibrium and can also be applied to study protein (P) + ligand (L) to protein-ligand (PL) complex equilibria. In this case one finds that

$$\begin{aligned} d\ln K = &-(F_{P,PL} - F_{P,P} - F_{P,L})d\beta \\ &+ \sum_i (N_{PLi} - N_{Pi} - N_{Li})(d\beta\mu_i - \bar{h}_i d\beta), \end{aligned} \tag{70}$$

where $K = \rho_{\mathrm{PL}}/\rho_{\mathrm{P}}\rho_{\mathrm{L}}$, and we have used the fact that $d\mu_{PL} = d\mu_P + d\mu_L$ to eliminate $d\mu_{PL}$. The summations in the ε_P term and above expression only involve thermodynamically independent species (solvent, protein, ligand, etc.).

The expressions provided in Eqs. (41) and (42) also have significant consequences for pairwise additive potentials. First, the enthalpy change can be decomposed into a series of terms related to the average intra and intermolecular energies and solvent distribution around each solute form. Further decomposition into van der Waals and electrostatic contributions is also possible without additional approximation. Second, the decomposition is exact and different from typical *ad hoc* approaches.[12,13] Third, one can rewrite Eq. (42) to give

$$\begin{aligned} F^{\infty}_{P,D} = &\langle E_{DD}\rangle + \langle E_{D1}\rangle + \langle N_1\rangle_D \left(\frac{\langle E_{11}\rangle_D}{\langle N_1\rangle_D} - E^0_1\right) \\ &+ PV^{*,\infty}_D - E^0_2, \end{aligned} \tag{71}$$

which indicates that the true measure of the local solvent contribution involves both the number of solvent molecules in the local region, and how their average energy differs from the molar enthalpy of the pure solvent. The neglect of the $E_1{}^0$ contribution in the calculation of simulated enthalpy changes would lead to a significant error, even when the difference in the number of solvent molecules between both forms is small (typically 2–5 for most proteins),[17] as the average potential energy for common water models is large; −46.5 kJ/mol for SPC/E water,[60] for example. Hence, one cannot just determine the change in the solvent-solvent energy when attempting to determine simulated enthalpy changes. Furthermore, it is unclear to what degree implicit solvent or coarse grained models include or approximate terms involving the number of solvent molecules.

Many of the expressions provided here involve differences between extensive quantities that are then intensive in nature. A prime example is Eq. (71) as used in Eq. (41). The last two terms are both extensive (dependent on the reference volume), but their difference is independent of this volume. Furthermore, manipulation of the terms in Eq. (71)

has to be performed specifically recognizing that these are grand canonical ensemble averages. For instance, as the local reference volume increases one might be tempted to write $\langle E_{11}\rangle_D/\langle N_1\rangle_D$ as ${E_1}^0$, implying that the last two terms cancel for large volumes. This is incorrect and it is the additional change in the last two terms with increasing volume that will cancel leaving the same constant value independent of the reference volume (as long as it is large enough to include all the perturbing effects of the solute). In addition, when decomposing the terms found in Eq. (44), for example, one can only isolate the intensive term $\langle \delta\varepsilon_P \delta\varepsilon_P\rangle_D - \langle \delta\varepsilon_P \delta\varepsilon_P\rangle_0$, and not the extensive term $\langle \delta\varepsilon_P \delta\varepsilon_P\rangle_D$ itself, even though the latter (pure solvent) term cancels in Eq. (44).

The energy and enthalpy terms used in the definition of the various ε's involve the total internal energy. The kinetic energy contributions can be removed or ignored in some cases. For example, the kinetic energy contributions to the expression for α_P in Eq. (8), and the expressions provided in Eqs. (30), (37), (41), (42), (43), and (50) all cancel and therefore one could replace E with just the potential energy and ignore the kinetic energy contribution to the partial molar enthalpies. However, this is not the case for the C_P expression provided in Eq. (8) or the expressions provided in Eqs. (36), (44), and (55), where the ideal terms do not cancel, although the contribution (1/2 R per classical degree of freedom) is often small compared to the heat capacity change associated with the equilibrium itself, or will cancel in the case of protein denaturation ($n = 1$).

When studying protein denaturation it is clear that one is dealing with a transition where the protein changes from a set of relatively few nativelike conformations to a (potentially) very large number of denatured or unfolded conformations. However, this does not significantly affect the results presented here. If we consider a collection of denatured forms it can be shown that

$$d\ln\rho_D = \rho_D^{-1}\sum_i \rho_{D,i} d\ln\rho_{D,i} = \sum_i x_{D,i} d\ln\rho_{D,i}, \qquad (72)$$

where $x_{D,i}$ is the number fraction of denatured form i. Therefore, the F's and N_{ij}'s used in Eqs. (26) and (41) simply become averages over the individual denatured forms weighted by their fractional populations, or a simple time average in a simulation.

In Sections II and III, we attempted to outline how one could obtain specific contributions to, or correlations with, each form present in the equilibrium. This required additional experimental information. The experimental data can come in a variety of forms and therefore in our previous discussion we just provided expressions for changes in h_2^* with temperature—without invoking a specific process. One process relevant to the present discussion is the process of solvation. In this case the enthalpy change (Δh_2^*) corresponds to the process of transferring a molecule of i from a fixed position in an ideal gas phase to a fixed position in the solution at the T, P, and composition of interest. In principle, this information should be amenable to experiment. However, this might not always be true in practice. For instance, the (gas to solution) solvation enthalpy may be available for small volatile solutes, but not for proteins. In this case, a more practical application would involve the study of changes in enthalpy between pure water and mixed solvent systems, where the change in enthalpy would then involve the enthalpies of transfer between solvent systems. Derivatives of the protein pseudo enthalpy could be replaced by experimentally available protein heat capacities noting that $c_{p,2}^* = \bar{c}_{p,2} - (\partial RT^2\alpha_P/\partial T)_{P,\{N\}} - 3R/2$, as indicated by Eq. (10).

In an effort to establish the relative importance of each term in Eqs. (26), (27), (41), and (44) one can examine existing data concerning protein denaturation. Experimental data for the pressure and thermal denaturation of Ribonuclease A at 295 K and pH 2 in D_2O is available.[61,62] The observed difference in volume of -21 cm^3/mol obtained from pressure denaturation is small, especially compared to the native state volume of 9500 cm^3/mol, and corresponds to slightly more than one water molecule (${V_1}^0 = 18$ cm^3/mol). The second derivative of the equilibrium constant with respect to pressure is observed to be 6.1×10^{-6} bar^{-2}. Using the above data, one finds the ${\kappa_T}^0$ term in Eq. (27) to contribute a negligible -2×10^{-9} bar^{-2}, while the last contribution (difference in the squared terms) equals 2×10^{-6} bar^{-2}. Therefore, the last contribution and the difference between the $\langle \delta N_1 \delta N_1\rangle$ terms are similar in magnitude. Thermal denaturation data provide a standard enthalpy change and heat capacity of 200 kJ/mol and 4800 J/mol/K, respectively. Hence, the ${C_{P,m}}^0$ term in Eq. (44) is negligible, while the other two terms would appear to be significant. While this data only represent one system, we expect similar results for other proteins.

It is tempting to simplify some of the expressions provided previously. However, this should be performed with care. For instance, the pressure effect can be written by reference to Eq. (10) as

$$\left(\frac{\partial \ln K}{\partial P}\right)_{\beta,\{N\}} = \beta \sum_i (N_{Ai} - nN_{Mi})\,\bar{V}_i = -\beta(V_A^* - nV_M^*), \qquad (73)$$

which seems logical. It might then be tempting to write

$$\left(\frac{\partial^2 \ln K}{\partial P^2}\right)_{\beta,\{N\}} = -\beta\left[\left(\frac{\partial V_A^*}{\partial P}\right)_{\beta,\{N\}} - n\left(\frac{\partial V_M^*}{\partial P}\right)_{\beta,\{N\}}\right] \qquad (74)$$

and to use the expression provided by Eq. (24) for the two derivatives. This is incorrect and differs from the correct result provided in Eq. (20). The reason is that Eq. (24) was developed for a system of thermodynamically independent composition variables and therefore requires that all N (including N_A and N_M) are held constant. This is clearly not the case according to Eq. (16). In contrast, the development of Eq. (20) correctly captures the inherent dependence of N_A (and N_M) on pressure and only assumes that their sum (N_2) is constant. Similar arguments also explain why one cannot simply replace the F_P's in Eq. (30), for instance, with their values suggested by Eq. (8). Furthermore, computer simulations or experiments which determine how the volume of a single protein form varies with pressure will provide information concerning the compressibility of that form. However, the difference in compressibility between two independent forms

is not simply the compressibility associated with the chemical equilibrium itself.

The fluctuation based results for the effect of pressure on protein denaturation involve differences between the water distributions surrounding each protein form. This is clearly related to the protein volume. For an infinitely dilute solute (2) in a single solvent (1) the pseudo volume is given by $-N_{21}^{\infty}V_1^0$ and can be considered as the ensemble average of a series of protein volumes provided by $V - N_1 V_1^0$, where N_1 is the number of solvent molecules surrounding a single protein in the same fixed volume V for each member of the ensemble. In this case the fluctuations in the protein volume are then given by $\langle \delta N_1 \delta N_1 \rangle_2 (V_1^0)^2$, which is one of the terms that appears in Eqs. (27) and (29). Therefore, fluctuations in the water distribution and fluctuations in the protein volume are intimately related. However, this relationship is only exact when one measures the protein volume using the local number of water molecules in the solution and, unfortunately, this is not the only term that appears in Eqs. (27) and (29).

Experimental information concerning the compressibility of proteins can also be obtained from sound velocity studies.[63–65] This approach has the advantage of probing the compressibility of proteins under normal pressures and temperatures. The resulting isentropic protein compressibilities (κ_S) are, however, much more difficult to interpret both experimentally and theoretically. For instance, the isentropic compressibility of a solution mixture is given by the thermodynamic relationship

$$RT\kappa_S = RT\kappa_T - RT\alpha_P \left(\frac{\partial T}{\partial P}\right)_{S,\{N\}} = RT\kappa_T - \frac{(RT^2\alpha_P)^2 V_m}{RT^2 C_{P,m}}, \tag{75}$$

which can be expressed in terms of fluctuating quantities using Eq. (8). Experiments provide partial molar or apparent molar isentropic protein compressibilities. These involve derivatives of the above expression and thereby contain a variety of fluctuating quantities. The analysis is greatly simplified by transforming to partial molar isothermal compressibilities,[65] which can then be analyzed in a more traditional manner (see Appendix B). Hence, we have not provided the expressions for partial molar isentropic compressibilities here.

Finally, it is important to recognize that changes to the equilibrium constant also involve contributions from the internal partition function—specifically changes to the vibrational modes— which are especially important when the temperature is varied. To illustrate this one can write (from Eqs. (9) and (14)) for protein denaturation

$$\left(\frac{\partial \ln K}{\partial \beta}\right)_{P,\{N\}}^{\infty} = -\left(\frac{\partial\, \beta\left(\mu_D^{*,\infty} - \mu_N^{*,\infty}\right)}{\partial \beta}\right)_{P,\{N\}} = -\left(h_D^{*,\infty} - h_N^{*,\infty}\right) = -\left(F_{P,D}^{\infty} - F_{P,N}^{\infty}\right). \tag{76}$$

The pseudo chemical potential terms contain the internal partition function whose dependence on temperature can be significant. Hence, analysis of experimental data on protein denaturation will explicitly include the vibration contributions in the extracted F_P's through the energy terms. However, simulations of classical systems performed with bond constraints will not include such contributions, or they will only be included to the degree that the force field has implicitly accounted for such effects during the parameter development.

V. CONCLUSIONS

Expressions describing how a chemical equilibrium responds to changes in pressure, temperature, and composition have been provided in terms of local fluctuations around the relevant chemical forms in solution. The expressions can be used to analyze experimental data regarding any chemical equilibrium which follows Eq. (11) in any multicomponent mixture at any composition, or they can be used to analyze or predict such effects using computer simulation. In particular, we provide exact expressions for determining enthalpy, heat capacity, and compressibility changes associated with a chemical equilibrium from a single simulation. The resulting expressions contain terms, which involve particle-energy, energy-energy, and particle-particle correlations for the enthalpy, heat capacity, and compressibility, respectively, for processes at infinite dilution in a single solvent. However, additional terms are also present which render the expressions non-trivial and different from other intuitive, but more approximate, approaches.

ACKNOWLEDGMENTS

The Authors would like to thank Samantha Weerasinghe and Elizabeth Ploetz for valuable discussion. The project described was supported by Grant No. R01GM079277 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

APPENDIX A: DERIVATIVES OF FLUCTUATING QUANTITIES IN THE GRAND CANONICAL ENSEMBLE

Here, we outline the general approach for obtaining fluctuating quantities, which can be used to help develop derivatives of KBIs with respect to pressure, temperature, and composition. The differential for a general fluctuating quantity in the grand canonical ensemble can be written

$$d\langle \delta X \delta Y \rangle = -\langle \delta X \delta Y \delta E \rangle d\beta + \langle \delta X \delta Y \rangle d\ln V + \sum_k \langle \delta X \delta Y \delta N_k \rangle d\beta \mu_k, \tag{A1}$$

where we have used the fact that $d\langle \delta X\ \delta Y \rangle = \langle \delta X \delta Y \rangle / V\ dV$ when $\langle \delta X\ \delta Y \rangle / V$ is intensive and therefore independent of V. Taking derivatives of this equation with the appropriate values of X and Y provides the expressions required for Eqs. (23), (38), and (65). This approach can also be used for other partial molar quantities as illustrated below (Appendix B). When X and Y are both particle numbers, this provides a route to derivatives of the KBIs or other fluctuating quantities (thermal expansion, compressibility, heat

capacity) with respect to pressure, temperature, and composition, which could be used to further analyze the properties of solution mixtures. The simplest results are those provided for the B_{ij}'s,

$$\left(\frac{\partial B_{ij}}{\partial P}\right)_{T,\{N\}} = \beta V^{-1}\sum_k \langle \delta N_i \delta N_j \delta N_k\rangle \bar{V}_k,$$

$$\left(\frac{\partial B_{ij}}{\partial \beta}\right)_{P,\{N\}} = -V^{-1}\langle \delta N_i \delta N_j \delta \varepsilon_P\rangle,$$

$$\left(\frac{\partial B_{ij}}{\partial \langle N_l\rangle}\right)_{P,T,\{N\}'} = V^{-1}\sum_k \langle \delta N_i \delta N_j \delta N_k\rangle \mu_{kl}, \quad \text{(A2)}$$

where $\mu'_{kl} = \beta(\partial\mu_k/\partial\langle N_l\rangle)_{P,T,\{N\}'}$.

APPENDIX B: PARTIAL MOLAR HEAT CAPACITIES, THERMAL EXPANSIONS, AND COMPRESSIBILITIES

In this section, we derive expressions for several partial molar quantities used in the previous discussion, but not available in the literature. The partial molar constant pressure heat capacity can be obtained starting from the definition

$$RT^2\bar{c}_{p,i} = RT^2\left(\frac{\partial \bar{h}_i}{\partial T}\right)_{P,\{N\}} = \left(\frac{\partial RT^2 C_P}{\partial \langle N_i\rangle}\right)_{T,P,\{N\}'} \quad \text{(B1)}$$

Therefore, using the expression in Eq. (8) for C_P and then Eq. (A1) with $X = Y = \varepsilon_P$ one finds

$$RT^2\bar{c}_{P,i} = \rho\bar{V}_i RT^2 C_{P.m} + \sum_j \langle \delta\varepsilon_P \delta\varepsilon_P \delta N_j\rangle \mu'_{ji}, \quad \text{(B2)}$$

where ρ is the total number density. As the chemical potential derivatives and partial molar volume can also be expressed in terms of fluctuations (KBIs) this is the desired result, although including these additional fluctuations here is not particularly informative. We note that $C_{P,m} = \sum_i x_i \bar{c}_{P,i}$ and hence the double summation over the last term in Eq. (B2) must be zero. The same approach can be used for the corresponding constant volume quantities. First, we note that the quantities used in Eq. (54) are given by

$$RT^2\bar{c}_{V,i} \equiv RT^2\left(\frac{\partial \bar{u}_i}{\partial T}\right)_{V,\{N\}} = \left(\frac{\partial RT^2 C_V}{\partial \langle N_i\rangle}\right)_{T,P,\{N\}'} \quad \text{(B3)}$$

in terms of the partial molar energies ($\bar{u}_i$). Using the expression for C_V provided in Eq. (51) leads to

$$RT^2\bar{c}_{V,i} = \rho\bar{V}_i RT^2 C_{V.m} + \sum_j \langle \delta\varepsilon_V \delta\varepsilon_V \delta N_j\rangle \mu'_{ji} \quad \text{(B4)}$$

as a final result.

Expressions for the partial molar thermal expansions are slightly more complicated. First, we note that from our definition one has

$$RT^2\bar{\alpha}_{p,i} \equiv RT^2\left(\frac{\partial \ln \bar{V}_i}{\partial T}\right)_{P,\{N\}} = RT^2\alpha_P + \frac{V}{\bar{V}_i}\left(\frac{\partial RT^2\alpha_P}{\partial \langle N_i\rangle}\right)_{T,P,\{N\}'}. \quad \text{(B5)}$$

From Eq. (8) the general expression for the thermal expansion can be written $RT^2\alpha_P = -V^{-1}\sum_j \langle \delta N_j \delta\varepsilon_P\rangle \bar{V}_j$ and therefore,

$$\left(\frac{\partial RT^2\alpha_P}{\partial \langle N_i\rangle}\right)_{T,P,\{N\}'} = -\frac{\bar{V}_i}{V}RT^2\alpha_P - \frac{1}{V}\sum_j \left(\frac{\partial(\langle \delta N_j \delta\varepsilon_P\rangle \bar{V}_j)}{\partial \langle N_i\rangle}\right)_{T,P,\{N\}'}, \quad \text{(B6)}$$

which, after using Eq. (A1) with $X = N_j\bar{V}_j$ and $Y = \varepsilon_P$, provides the final result

$$RT^2\bar{\alpha}_{p,i} = RT^2\alpha_P - \frac{1}{\bar{V}_i}\sum_{j,k}\bar{V}_j\langle \delta N_j \delta N_k \delta\varepsilon_p\rangle \mu'_{ki}. \quad \text{(B7)}$$

We note that $\alpha_P = \sum_i \phi_i \bar{\alpha}_{P,i}$ and hence the triple summation over the last term in Eq. (B7) must be zero.

Finally, the partial molar isothermal compressibilities can be obtained from our initial definition

$$RT\bar{\kappa}_{T,i} \equiv RT\left(\frac{\partial \ln \bar{V}_i}{\partial P}\right)_{T,\{N\}} = RT\kappa_T + \frac{V}{\bar{V}_i}\left(\frac{\partial RT\kappa_T}{\partial \langle N_i\rangle}\right)_{T,P,\{N\}'}. \quad \text{(B8)}$$

From Eq. (8) the general expression for the compressibility can be written $RT\kappa_T = V^{-1}\sum_{j,k}\langle \delta N_j \delta N_k\rangle \bar{V}_j\bar{V}_k$ and therefore,

$$\left(\frac{\partial RT\kappa_T}{\partial \langle N_i\rangle}\right)_{T,P,\{N\}'} = -\frac{\bar{V}_i}{V}RT\kappa_T + \frac{1}{V}\sum_{j,k}\left(\frac{\partial(\langle \delta N_j \delta N_k\rangle \bar{V}_j\bar{V}_k)}{\partial \langle N_i\rangle}\right)_{T,P,\{N\}'}, \quad \text{(B9)}$$

which, after using Eq. (A1) with $X = N_j\bar{V}_j$ and $Y = N_k\bar{V}_k$, provides the final expression

$$RT\bar{\kappa}_{T,i} = RT\kappa_T + \frac{1}{\bar{V}_i}\sum_{j,k,l}\bar{V}_j\bar{V}_k\langle \delta N_j \delta N_k \delta N_l\rangle \mu'_{li}. \quad \text{(B10)}$$

We note that $\kappa_T = \sum_i \phi_i \bar{\kappa}_{T,i}$ and hence the quadruple summation over the last term in Eq. (B10) must be zero.

[1] K. A. Dill, Biochemistry **29**, 7133 (1990).
[2] J. F. Brandts, J. Am. Chem. Soc. **86**, 4291 (1964).
[3] A. Zipp and W. Kauzmann, Biochemistry **12**, 4217 (1973).
[4] C. Tanford, J. Am. Chem. Soc. **86**, 2050 (1964).
[5] V. A. Parsegian, R. P. Rand, N. L. Fuller, and D. C. Rau, Methods Enzymol. **127**, 400 (1986).
[6] F. Meersman, C. M. Dobson, and K. Heremans, Chem. Soc. Rev. **35**, 908 (2006).
[7] R. Liu, H. Barkhordarian, S. Emadi, C. B. Park, and M. R. Sierks, Neurobiol. Dis. **20**, 74 (2005).
[8] S. Narayanan and B. Reif, Biochemistry **44**, 1444 (2005).
[9] D. K. Klimov, J. E. Straub, and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **101**, 14760 (2004).
[10] D. R. Canchi, D. Paschek, and A. E. Garcia, J. Am. Chem. Soc. **132**, 2338 (2010).
[11] D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. B. Shan, and W. Wriggers, Science **330**, 341 (2010).

[12]P. Setny, R. Baron, and J. A. McCammon, J. Chem. Theory Comput. **6**, 2866 (2010).
[13]L. V. Schafer, D. H. de Jong, A. Holt, A. J. Rzepiela, A. H. de Vries, B. Poolman, J. A. Killian, and S. J. Marrink, Proc. Natl. Acad. Sci. U.S.A. **108**, 1343 (2011).
[14]T. Lazaridis and M. Karplus, Biophys. Chem. **78**, 207 (1999).
[15]N. V. Prabhu and K. A. Sharp, Annu. Rev. Phys. Chem. **56**, 521 (2005).
[16]S. W. Rick, J. Phys. Chem. B **104**, 6884 (2000).
[17]C. A. Royer, Biochim. Biophys. Acta-Protein Struct. Mol. Enzymol. **1595**, 201 (2002).
[18]V. M. Dadarlat and C. B. Post, J. Phys. Chem. B **105**, 715 (2001).
[19]E. Paci, Biochim. Biophys. Acta-Protein Struct. Mol. Enzymol. **1595**, 185 (2002).
[20]V. M. Dadarlat and C. B. Post, Biophys. J. **91**, 4544 (2006).
[21]A. Cooper, Proc. Natl. Acad. Sci. U.S.A. **73**, 2740 (1976).
[22]C. Scharnagl, M. Reif, and J. Friedrich, Biochim. Biophys. Acta-Proteins Proteomics **1749**, 187 (2005).
[23]A. Wallqvist, D. G. Covell, and D. Thirumalai, J. Am. Chem. Soc. **120**, 427 (1998).
[24]B. J. Bennion and V. Daggett, Proc. Natl. Acad. Sci. U.S.A. **100**, 5142 (2003).
[25]P. J. Rossky, Proc. Natl. Acad. Sci. U.S.A. **105**, 16825 (2008).
[26]L. Hua, R. H. Zhou, D. Thirumalai, and B. J. Berne, Proc. Natl. Acad. Sci. U.S.A. **105**, 16928 (2008).
[27]V. Pierce, M. Kang, M. Aburi, S. Weerasinghe, and P. E. Smith, Cell Biochem. Biophys. **50**, 1 (2008).
[28]S. Shimizu, Proc. Natl. Acad. Sci. U.S.A. **101**, 1195 (2004).
[29]P. E. Smith, J. Phys. Chem. B **108**, 18716 (2004).
[30]I. L. Shulgin and E. Ruckenstein, J. Chem. Phys. **123**, 054909 (2005).
[31]E. Ruckenstein and I. L. Shulgin, Adv. Colloid Interface Sci. **123**, 97 (2006).
[32]I. L. Shulgin and E. Ruckenstein, Biophys. J. **90**, 704 (2006).
[33]J. Rosgen, B. M. Pettitt, and D. W. Bolen, Protein Sci. **16**, 733 (2007).
[34]J. G. Kirkwood and F. P. Buff, J. Chem. Phys. **19**, 774 (1951).
[35]J. P. O'Connell, Mol. Phys. **20**, 27 (1971).
[36]A. Ben-Naim, *Molecular Theory of Solutions* (Oxford University Press, New York, 2006).
[37]E. Matteoli and G. A. Mansoori, *Fluctuation Theory of Mixtures* (Taylor & Francis, New York, 1990).
[38]F. P. Buff and R. Brout, J. Chem. Phys. **23**, 458 (1955).
[39]P. G. Debenedetti, J. Chem. Phys. **87**, 1256 (1987).
[40]P. G. Debenedetti, J. Chem. Phys. **86**, 7126 (1987).
[41]P. G. Debenedetti, J. Chem. Phys. **88**, 2681 (1988).
[42]P. G. Debenedetti, Mol. Simul. **2**, 33 (1989).
[43]E. A. Ploetz and P. E. Smith, "Local fluctuations in solution mixtures," J. Chem. Phys. (submitted).
[44]S. Shimizu and C. L. Boon, J. Chem. Phys. **121**, 9147 (2004).
[45]M. Aburi and P. E. Smith, J. Chem. Phys. B **108**, 7382 (2004).
[46]R. L. Perry, J. C. Telotte, and J. P. O'Connell, Fluid Phase Equilib. **5**, 245 (1981).
[47]R. L. Perry and J. P. O'Connell, Mol. Phys. **52**, 137 (1984).
[48]A. Ben-Naim, J. Chem. Phys. **63**, 2064 (1975).
[49]M. B. Gee and P. E. Smith, J. Chem. Phys. **131**, 165101 (2009).
[50]K. E. Newman, Chem. Soc. Rev. **23**, 31 (1994).
[51]M. Kang and P. E. Smith, J. Chem. Phys. **128**, 244511 (2008).
[52]P. E. Smith, J. Chem. Phys. **129**, 124509 (2008).
[53]D. J. Adams, Mol. Phys. **29**, 307 (1975).
[54]T. Imai and Y. Sugita, J. Phys. Chem. B **114**, 2281 (2010).
[55]P. E. Smith, J. Phys. Chem. B **110**, 2862 (2006).
[56]R. F. Greene Jr and C. N. Pace, J. Biol. Chem. **249**, 5388 (1974).
[57]G. I. Makhatadze, J. Phys. Chem. B **103**, 4781 (1999).
[58]P. E. Pjura, M. E. Paulaitis, and A. M. Lenhoff, AIChE J. **41**, 1005 (1995).
[59]S. Weerasinghe and P. E. Smith (unpublished).
[60]H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).
[61]G. I. Makhatadze, G. M. Clore, and A. M. Gronenborn, Nat. Struct. Biol. **2**, 852 (1995).
[62]K. E. Prehoda, E. S. Mooberry, and J. L. Markley, Biochemistry **37**, 5785 (1998).
[63]H. Shiio, T. Ogawa, and H. Yoshihashi, J. Am. Chem. Soc. **77**, 4980 (1955).
[64]K. Gekko and H. Noguchi, J. Phys. Chem. **83**, 2706 (1979).
[65]K. Gekko and Y. Hasegawa, Biochemistry **25**, 6563 (1986).