

This is the author's final, peer-reviewed manuscript as accepted for publication. The publisher-formatted version may be available through the publisher's web site or your institution's library.

Kirkwood–Buff theory of molecular and protein association, aggregation, and cellular crowding

Moon Bae Gee, Paul E. Smith

How to cite this manuscript

If you make reference to this version of the manuscript, use the following information:

Gee, M.B. & Smith, P.E. (2009). Kirkwood–Buff theory of molecular and protein association, aggregation, and cellular crowding. Retrieved from <http://krex.ksu.edu>

Published Version Information

Citation: Gee, M.B. & Smith, P.E. (2009). Kirkwood–Buff theory of molecular and protein association, aggregation, and cellular crowding. *The journal of chemical physics*, 131(16), 11.

Copyright: © 2009 American Institute of Physics

Digital Object Identifier (DOI): doi:10.1063/1.3253299

Publisher's Link: http://jcp.aip.org/resource/1/jcpsa6/v131/i16/p165101_s1

This item was retrieved from the K-State Research Exchange (K-REx), the institutional repository of Kansas State University. K-REx is available at <http://krex.ksu.edu>

Kirkwood–Buff theory of molecular and protein association, aggregation, and cellular crowding

Moon Bae Gee and Paul E. Smith^{a)}

*Department of Chemistry, 213 CBC Building, Kansas State University,
Manhattan, Kansas 66506-0401, USA*

(Received 27 April 2009; accepted 1 October 2009; published online 26 October 2009)

An analysis of the effect of a cosolvent on the association of a solute in solution using the Kirkwood–Buff theory of solutions is presented. The approach builds on the previous results of Ben-Naim by extending the range of applicability to include any number of components at finite concentrations in both closed and semiopen systems. The derived expressions, which are exact, provide a foundation for the analysis and rationalization of cosolvent effects on molecular and biomolecular equilibria including protein association, aggregation, and cellular crowding. A slightly different view of cellular crowding is subsequently obtained. In particular, it is observed that the addition of large cosolvents still favors the associated form even when traditional excluded volume effects are absent. © 2009 American Institute of Physics. [doi:10.1063/1.3253299]

I. INTRODUCTION

Protein association, protein unfolding, protein aggregation, and cellular crowding are known to affect the normal function of cellular systems.^{1–7} In many cases, the resulting small changes in normal protein-protein intra- and intermolecular interactions are thought to lead to a variety of human diseases.^{8,9} Consequently, it is important to study these processes, at both the thermodynamic and atomic levels, in an effort to understand and eventually manipulate the behavior of such systems. One way to affect and help understand peptide and protein association is through the use of additives or cosolvents.¹⁰ A general theory describing these types of effects which relates their thermodynamic properties to the interactions between species in solution is therefore desired. This is a major aim of the present work.

The basic thermodynamic effects of additives on chemical equilibria were outlined some time ago using the concepts of binding and linkage.^{11–13} Binding polynomials are then typically used to help illustrate the resulting thermodynamic relationships. While binding polynomials can be used to provide an accurate representation of real experimental data, the physical interpretation of the binding constants requires some care for systems involving only weakly binding cosolvents.¹⁴ This is immediately apparent when attempting to define corresponding binding sites, occupation numbers, and equilibrium constants using coordinate data provided by computer simulations. The transient nature of the interactions between the cosolvent and biomolecule render such a comparison of the experimental and simulation data essentially impossible. Other thermodynamic approaches have been outlined,^{15–21} but it remains difficult to relate these to realistic distributions between the various species in solution.

The most common approach used to understand the effects of an additive on protein folding, association, and ag-

gregation involved scaled particle theory (SPT).^{6,19,22,23} Here, the change in free energy for association, denaturation, etc., is determined from the change in the free energy for insertion of the initial and final states into a system of hard spheres as a function of the additive volume fraction. These calculations generally suggest that the native and any associated states are favored on increasing the additive concentration. These are exactly the trends observed experimentally,⁶ and for simulations using simple excluded volume crowders.^{3,24} As SPT is dominated by repulsive excluded volume effects, the above trends are usually interpreted in terms of a decrease in the free volume, or increase in the excluded volume, within the solution. Unfortunately, it is difficult to extend SPT to include attractive interactions which undoubtedly occur between proteins in real systems.⁶ Hence, other models are still required which complement existing approaches.

The application of the Kirkwood–Buff (KB) theory of solutions to biological systems recently attracted some interest.¹⁴ KB theory provides a rigorous link between intermolecular distributions in solution and the thermodynamic characteristics of that solution.^{25,26} In particular, it has been used to understand the preferential interactions of cosolvents with proteins and small molecules,^{27–34} to study changes in the hydration of solutes,^{35–38} to investigate a variety of solution properties,^{39–41} and for hard sphere models of cellular crowding.⁴² Previously, Ben-Naim used KB theory to understand the effects of cosolvents on the association equilibrium of a solute in solution.⁴³ The current approach is based on the previous work of Ben-Naim and is similar to the approach of Hall for studying micelles.⁴⁴ Here, we extend the Ben-Naim approach to describe the effects of a cosolvent on the equilibrium constant for association to include semiopen and not just closed systems, and for systems with any number of components all of which can appear at finite concentrations. To achieve this we adopt a recent approach which avoids the usual matrices used for closed systems.⁴⁵ The resulting ex-

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

pressions are then used to illustrate some of the differences between open and closed systems, and between real and ideal solutions.

II. KIRKWOOD–BUFF THEORY

Kirkwood–Buff theory is an exact theory of solutions.⁴⁶ The principle quantities of interest are the Kirkwood–Buff integrals (KBIs) defined by

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

where g_{ij} is the radial distribution function (rdf) between the center of mass of species i and j . These integrals can be used to relate thermodynamic properties of a solution to the molecular distributions in that solution. This is achieved by starting with the grand canonical ensemble expression,²⁵

$$RT \left(\frac{\partial \ln \rho_i}{\partial \mu_j} \right)_{T, \mu_{k \neq j}} = \delta_{ij} + N_{ij}, \quad (2)$$

where R is the gas constant, T is the absolute temperature, $\rho_i = N_i/V$ is the number density of species i , μ is the chemical potential, and δ_{ij} is the Kronecker delta function. The $N_{ij} = \rho_j G_{ij}$ values describe the change in the number of j particles on introduction of a central i particle, from the number of j particles found in an equivalent volume of bulk solution at the same composition. More details concerning the application of KB theory to understand a variety of solution properties can be found elsewhere.^{14,26,39}

The major aim of this work is to develop expressions for cosolvent effects on biomolecular equilibria in both closed and semiopen ensembles. Traditionally, KB theory starts from Eq. (2) and uses a series of thermodynamic transformations to provide expressions for similar derivatives in closed ensembles.^{25,46} A general matrix formulation is available for chemical potential derivatives in closed systems—the most common system of interest.⁴⁶ However, we will avoid this approach as the corresponding expressions are difficult to manipulate for large numbers of components. Furthermore, it is also difficult to transform from the closed ensemble results back to expressions valid for semiopen systems.²⁹

III. CHEMICAL EQUILIBRIA

Let us consider a formal n_c component system containing a primary solvent (1), a biomolecule (2), and a series of cosolvents (3, 4, ...). The cosolvents can be small molecules such as urea, a salt, a proton, or larger molecules such as different proteins. We will refer to species 2 as a biomolecule, but it can easily refer to any associating species. Furthermore, the biomolecule can exist in two forms, one being a monomer (M) and the other an aggregate (A) of any shape formed from n monomers. Any counterion associated with the biomolecule will be assumed to have no effect on the equilibrium and to be different from any of the cosolvent if the latter are salts (no Donnan effect). Hence, we have the equilibrium process,



which can be described by an equilibrium constant K . Formally, the equilibrium constant involves the activities of the various species. In the majority of thermodynamic approaches it is then assumed that the activities of the biomolecule species can be replaced by their concentrations as the biomolecules usually appear at low concentrations, and concentrations are relatively easy to determine experimentally. Here, we do not assume ideal behavior of any species, but simply define our equilibrium constant in terms of concentrations to match the usual experimental representation of the data. Consequently, the equilibrium constant can be defined in several ways using a variety of concentration scales. We chose

$$K = \frac{\rho_A}{\rho_M^n} \quad (4)$$

in terms of number densities or molarities. Other choices will be discussed later. The number of each form of the biomolecule present in solution are related by

$$N_M + nN_A = N_2, \quad dN_M + ndN_A = dN_2, \quad (5)$$

and,

$$f_M = \frac{N_M}{N_2}, \quad f_A = \frac{nN_A}{N_2}, \quad f_M + f_A = 1, \quad (6)$$

where f_i is the fraction of either A or M at equilibrium. The material equilibrium condition indicates that

$$\mu_A = n\mu_M, \quad d\mu_A = nd\mu_M = nd\mu_2, \quad (7)$$

as long as we remain at equilibrium. From Eq. (4) a general change in the equilibrium constant can be written as

$$d \ln K = d \ln \rho_A - nd \ln \rho_M. \quad (8)$$

We note that all the above expressions can be applied to any chemical equilibrium that follows Eq. (3) in any thermodynamically reasonable ensemble. Our main aim is to express the changes in the number densities of A and M in terms of the KB integrals and thereby provide a simple physical picture of the effect of a cosolvent on the equilibrium constant.

IV. GENERAL KIRKWOOD–BUFF THEORY OF CHEMICAL EQUILIBRIA

Our system corresponds to a pseudo- n_c+1 component system with thermodynamic constraints between two of the components (M and A). Before proceeding we note that the application of KB theory to this type of problem involves some subtle issues. First, N_2 is a true independent thermodynamic variable. However, N_M and N_A are not as they are related through Eq. (5). Hence, we will avoid taking derivatives with respect to the chemical potentials or concentrations of M and A , although derivatives involving either one (M or A) can be used in place of the chemical potential or concentration of 2. The application of KB theory implies that the KB integrals used here for semiopen or closed systems correspond to an equivalent system at the same composition but open to all species (including M and A). This issue has

been discussed by Ben-Naim.²⁶ One can apply KB theory to understand such a system as long as one does not treat M and A as independent thermodynamic variables.⁴⁷

The traditional approach to this type of problem starts with the matrix formulation of KB theory for closed systems. However, the evaluation of the matrix determinants for a large number of components where all components are present at finite concentrations is rather cumbersome. In addition, transforming back to an open or semi-open system from the closed system results is also difficult. Hence, we use a different approach which avoids the direct evaluation of any matrices for small n_c values; thereby greatly simplifying the problem.⁴⁵ Let us consider the number density of each species to be functions of T and all the chemical potentials. The differential of the number densities at constant T is then provided by

$$d \ln \rho_i = \sum_j \left(\frac{\partial \ln \rho_i}{\partial \mu_j} \right)_{T, \mu_{k \neq j}} d\mu_j, \quad (9)$$

for any i , and where the summation is over all j components. The partial derivatives can be expressed in terms of KBIs through Eq. (2) to give

$$RTd \ln \rho_i = \sum_j (\delta_{ij} + N_{ij})d\mu_j \quad (10)$$

for any constant T ensemble. To our knowledge the above equation was first derived by Hall,⁴⁸ but using a different route. The general Gibbs–Duhem (GD) relation at constant temperature can be written as

$$dP = \sum_j \rho_j d\mu_j, \quad (11)$$

where P is the pressure. Equations (10) and (11) can be applied to any number of components in any ensemble with T constant.

Hence, for our n_c+1 component system of species 1, M , A , 3, 4, \dots , one can use Eq. (7) to eliminate $d\mu_A$ from the relationships provided in Eq. (10) for $i=A$ and M to give

$$RTd \ln \rho_A = (n + nN_{AA} + N_{AM})d\mu_M + \sum_{j \neq A, M}^{n_c+1} N_{Aj}d\mu_j, \quad (12)$$

$$RTd \ln \rho_M = (1 + N_{MM} + nN_{MA})d\mu_M + \sum_{j \neq A, M}^{n_c+1} N_{Mj}d\mu_j.$$

One could have focused on species A instead of M , but the results will be the same. However, we cannot retain *both* $d\mu_M$ and $d\mu_A$ terms for the remainder of the analysis as this implies they are thermodynamically independent. To generate a relationship for changes in the equilibrium constant we use Eq. (8) and the relationships in Eq. (12) to provide

$$RTd \ln K = (nN_{AA} + N_{AM} - nN_{MM} - n^2N_{MA})d\mu_M + \sum_{j \neq A, M}^{n_c+1} (N_{Aj} - nN_{Mj})d\mu_j. \quad (13)$$

However, the above expressions can be simplified further.

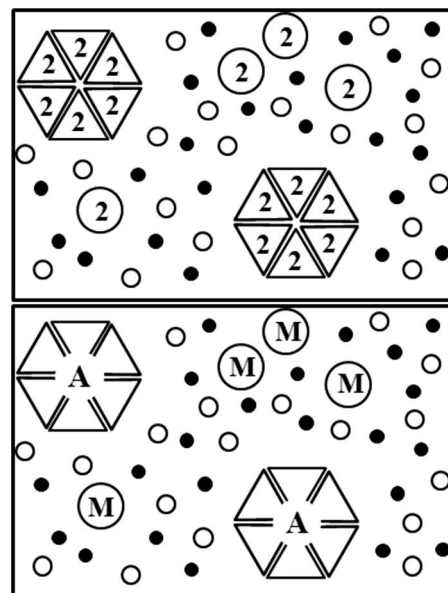


FIG. 1. The two representations of the same system used in this study. The system contains a solvent (1, shaded spheres), a solute (2), and a cosolvent (3, open spheres). In this case the solute can exist in two forms: One being the monomer (M) and the other being an aggregate (A) of $n=6$ monomers. The monomer can adopt different shapes in the associated and free forms.

A series of relationships between KBIs involving the biomolecule can be established. These can be written as

$$N_{i2} = N_{iM} + nN_{iA}, \quad (14)$$

$$N_{2i} = f_M N_{Mi} + f_A N_{Ai},$$

for $i \neq A, M$, or 2, and

$$N_{M2} = 1 + N_{MM} + nN_{MA},$$

$$N_{A2} = n + nN_{AA} + N_{AM}, \quad (15)$$

$$1 + N_{22} = f_M(1 + N_{MM} + nN_{MA}) + f_A(n + nN_{AA} + N_{AM}) = f_M N_{M2} + f_A N_{A2}.$$

The above expressions were obtained from the general fluctuation formula in the grand canonical (μVT) ensemble,

$$\delta_{ij} + N_{ij} = \frac{\langle N_i N_j \rangle - \langle N_i \rangle \langle N_j \rangle}{\langle N_i \rangle}, \quad (16)$$

by suitable substitutions of $N_2 = N_M + nN_A$. We note that the above relationships do not assume that N_2 is constant. They merely reflect a change in index for the M , A , and 2 species as illustrated in Fig. 1. Their meaning will be discussed later. Application of the above relationships to Eqs. (10)–(13) provides

$$RTd \ln K = \sum_j^{n_c} (N_{Aj} - nN_{Mj})d\mu_j, \quad (17)$$

$$RTd \ln \rho_i = \sum_j^{n_c} (\delta_{ij} + N_{ij})d\mu_j, \quad (18)$$

$$RTd \ln m_i = \sum_j^{n_c} (\delta_{ij} + N_{ij} - N_{1j} - \delta_{1j}) d\mu_j, \quad (19)$$

$$dP = \sum_j^{n_c} \rho_j d\mu_j, \quad (20)$$

where $d \ln m_i = d \ln \rho_i - d \ln \rho_1$, and m_i is the dimensionless molality (ρ_i/ρ_1). The above equations represent the basic expressions required for this study. The summations over the formal n_c components only involve indices 1, 2, 3, etc., and not M or A . They can be applied to any constant T ensemble and are valid for any concentration of solvent, biomolecule, and cosolvents. We note that Eqs. (18)–(20) could have been written directly for n_c component systems. However, Eq. (17) is not so obvious.

V. RESULTS

We will apply Eqs. (17)–(20) to a variety of thermodynamic ensembles to develop expressions for the effect of a cosolvent on the chemical equilibrium in terms of KB integrals. During the following analysis several combinations of KBIs appear repeatedly. Hence, in an effort to simplify the results we will define the following:

$$P_{jk}^i = N_{ij} - \frac{\rho_j}{\rho_k} N_{ik} = \rho_j (G_{ij} - G_{ik}) = -\frac{\rho_j}{\rho_k} P_{kj}^i, \quad (21)$$

which should be read as the preference or affinity of j over k for species i . For example, P_{31}^A quantifies the preference of 3 over 1 for the biomolecule in form A . This notation will only be used when it applies to specific differences between affinities of the various components for the two biomolecule forms. A simple physical interpretation of the sign associated with the above expression is that when $P_{jk}^i > 0$ the local ratio of j to k molecules around a central i molecule is larger than the bulk ratio of j to k molecules, and vice versa. We will also make use of the notation,

$$N_{ij}^+ = N_{ij} + m_j(1 + N_{11} - N_{i1} - N_{j1}), \quad (22)$$

originally introduced by Hall,⁴⁸ to help simplify many of the results.

VI. GENERAL EXPRESSIONS FOR ANY NUMBER OF COMPONENTS IN A CLOSED SYSTEM

Our initial focus will be on fully closed systems at constant pressure. Eliminating $d\mu_1$ from Eq. (17) using Eq. (20) and then taking derivatives with respect to one of the cosolvent molalities, one obtains the expression

$$\left(\frac{\partial \ln K}{\partial \ln m_j} \right)_{T,P,m_{k \neq j}} = \sum_{i>1} (P_{i1}^A - nP_{i1}^M) \mu_{ij}, \quad (23)$$

or alternatively,

$$\begin{aligned} RT \left(\frac{\partial \ln K}{\partial \mu_j} \right)_{T,P,m_{k \neq j}} &= \left(\frac{\partial \ln K}{\partial \ln a_j} \right)_{T,P,m_{k \neq j}} \\ &= \mu_{jj}^{-1} \sum_{i>1} (P_{i1}^A - nP_{i1}^M) \mu_{ij}, \end{aligned} \quad (24)$$

where the sum involves species 2 (not M and A), a_j is the activity of j , and we defined

$$\mu_{ij} = \frac{1}{RT} \left(\frac{\partial \mu_i}{\partial \ln m_j} \right)_{T,P,m_{k \neq j}} = \frac{1}{RT} \left(\frac{\partial \mu_i}{\partial \ln N_j} \right)_{T,P,N_{k \neq j}}. \quad (25)$$

The molality derivative has been chosen here as there is a general recursion relationship for these derivatives in any closed multicomponent solution.⁴⁵ Reference to the stability requirements for solutions indicates that $\mu_{ii} \geq 0$ and $\mu_{ij} < 0$.⁴⁹ Consequently, on increasing the concentration of j the association process is favored when $P_{i1}^A - nP_{i1}^M$ is positive for $i = j$ and negative for $i \neq j$. A change in concentration scales can be performed using the following thermodynamic relationships:

$$\left(\frac{\partial \ln \rho_i}{\partial \ln m_i} \right)_{T,P,m_{j \neq i}} = 1 - \phi_i \quad \left(\frac{\partial \ln x_i}{\partial \ln m_i} \right)_{T,P,m_{j \neq i}} = 1 - x_i, \quad (26)$$

where ϕ_i is the volume fraction and x_i the mole fraction. It should be noticed that these derivatives are positive and so a change in cosolvent concentration variable does not affect the sign of the previous effects.

In a closed system the affinity of all species for the M and A forms contribute to the overall effect. These contributions are expressed relative to species 1 (usually taken as the primary solvent). This is a direct consequence of our choice to eliminate $d\mu_1$ from Eq. (17) using the GD equation, and not due to the use of molality based derivatives. Clearly, one could eliminate any $d\mu_i$ to obtain a set of equivalent expressions. Finally, one could obtain an expression for a closed system at constant volume ($T, \rho_{j \neq i}$) from Eqs. (17)–(20). However, we have not pursued this further as the resulting expressions are rather complicated and this ensemble is not relevant for most biological systems, although it is the ensemble adopted in SPT.

VII. CONSTANT T, P , and m_2 Ensemble

The most common situation involves a closed system with $n_c = 3$ at constant T and P , where one is interested in the effect of a single cosolvent (3) on the biomolecular equilibrium when the biomolecule (2) and primary solvent (1) concentrations are constant. To develop expressions for the effect of a cosolvent on the equilibrium in this situation, we eliminate $d\mu_1$ from Eqs. (17)–(19) using Eq. (20). Then, noting that m_2 is constant we have from Eq. (19) with $i=2$ that

$$0 = (1 + N_{22}^+) d\mu_2 + N_{23}^+ d\mu_3. \quad (27)$$

This can then be used to eliminate $d\mu_2$ from Eq. (17) to give

$$RT \left(\frac{\partial \ln K}{\partial \mu_3} \right)_{T,P,m_2} = P_{31}^A - nP_{31}^M - \frac{(P_{21}^A - nP_{21}^M)N_{23}^+}{1 + N_{22}^+}. \quad (28)$$

To relate this to the cosolvent concentration one can eliminate $d\mu_1$ from Eq. (18) with $i=3$ using Eq. (20), and then combine with Eq. (27) to provide

$$RT \left(\frac{\partial \ln \rho_3}{\partial \mu_3} \right)_{T,P,m_2} = 1 + N_{33} - N_{13} - \frac{(N_{32} - m_2 N_{31})N_{23}^+}{1 + N_{22}^+}. \quad (29)$$

Alternatively, if one is interested in the effect of cosolvent molality one can eliminate $d\mu_1$ from Eq. (19) using Eq. (20) with $i=3$, and then combine with Eq. (27) to generate

$$RT \left(\frac{\partial \ln m_3}{\partial \mu_3} \right)_{T,P,m_2} = 1 + N_{33}^+ - \frac{N_{32}^+ N_{23}^+}{1 + N_{22}^+}, \quad (30)$$

which can then be combined with Eq. (28) if desired. Equation (30) is in agreement with previous results.⁴⁵ We note that the final terms in Eqs. (28)–(30) all disappear as the biomolecule concentration tends to zero, a common situation, and are then consistent with previous published expressions.^{27,29,47}

In some cases the cosolvent concentration might be constant and the biomolecule concentration may vary. The correct expressions for this situation can be obtained from Eqs. (28)–(30) via a simple $2 \leftrightarrow 3$ index change, or by the same approach that led to Eqs. (28)–(30) but where m_3 is constant. In principle, Eq. (18) can be used to obtain expressions for $d \ln x_i$ and thereby chemical potential derivatives with respect to mole fractions. These expressions are rather cumbersome and, as mole fractions are rarely the concentration scale of choice for biological systems, we have not pursued this further here.

VIII. GENERAL EXPRESSIONS FOR ANY NUMBER OF COMPONENTS IN SEMIOPEN SYSTEMS

If the system corresponds to that of a real cell, or some approximation to a real cell, then it may be open to one or more components. Furthermore, the addition of the cosolvent can occur with volume or pressure held constant. Let us consider a system which contains a set of species at fixed concentrations (ρ), together with a set of species with fixed chemical potentials (μ). Taking derivatives of Eq. (17) with respect to one of the cosolvent molarities with volume fixed one obtains the expression

$$\left(\frac{\partial \ln K}{\partial \ln \rho_j} \right)_{T,\rho_{k \neq j},\mu} = \sum_i^{n'_c} (N_{Ai} - nN_{Mi}) \mu'_{ij}, \quad (31)$$

where the sum is over the restricted set of components for which the system is closed (n'_c). Alternatively, with pressure constant one finds

$$\left(\frac{\partial \ln K}{\partial \ln \rho_j} \right)_{T,P,N_{k \neq j},\mu} = \sum_i^{n'_c} (N_{Ai} - nN_{Mi}) \mu''_{ij}. \quad (32)$$

The corresponding chemical potential derivatives are defined

$$\mu'_{ij} = \frac{1}{RT} \left(\frac{\partial \mu_i}{\partial \ln \rho_j} \right)_{T,\rho_{k \neq j},\mu} = \frac{1}{RT} \left(\frac{\partial \mu_i}{\partial \ln N_j} \right)_{T,V,N_{k \neq j},\mu}, \quad (33)$$

$$\mu''_{ij} = \frac{1}{RT} \left(\frac{\partial \mu_i}{\partial \ln \rho_j} \right)_{T,P,N_{k \neq j},\mu}.$$

Expressions for the required derivatives can be obtained after solving the set of simultaneous equations provided by Eqs. (18)–(20). This is illustrated in the following two sections.

IX. CONSTANT T , μ_1 , and ρ_2 Ensemble

A common osmotic system involves a semiopen system with $n_c=3$ at constant T , V , and μ_1 , where one is interested in the effect of a single cosolvent (3) on a biomolecular equilibrium when the biomolecule concentration is constant and the system is closed to the cosolvent. To develop expressions for the effect of a cosolvent on the equilibrium in this situation, we note that ρ_2 is constant and hence from Eq. (18) with $i=2$ one has

$$0 = (1 + N_{22})d\mu_2 + N_{23}d\mu_3. \quad (34)$$

This can then be used to eliminate $d\mu_2$ from Eq. (17) to give

$$RT \left(\frac{\partial \ln K}{\partial \mu_3} \right)_{T,\mu_1,\rho_2} = N_{A3} - nN_{M3} - \frac{(N_{A2} - nN_{M2})N_{23}}{1 + N_{22}}. \quad (35)$$

To relate this to the cosolvent concentration one can eliminate $d\mu_2$ from Eq. (18) with $i=3$, and then use Eq. (34) to provide

$$RT \left(\frac{\partial \ln \rho_3}{\partial \mu_3} \right)_{T,\mu_1,\rho_2} = 1 + N_{33} - \frac{N_{32}N_{23}}{1 + N_{22}}. \quad (36)$$

Alternatively, if one is interested in the effect of cosolvent molality one can eliminate $d\mu_2$ from Eq. (19) with $i=3$, and then use Eq. (34) to generate

$$RT \left(\frac{\partial \ln m_3}{\partial \mu_3} \right)_{T,\mu_1,\rho_2} = 1 + N_{33} - N_{13} - \frac{(N_{32} - N_{12})N_{23}}{1 + N_{22}}, \quad (37)$$

which can then be combined with Eq. (35) if desired. We note that the final terms in Eqs. (35)–(37) all disappear as the biomolecule concentration tends to zero.

X. CONSTANT T , P , μ_1 , and N_2 Ensemble

A second common semiopen ensemble involves a system with $n_c=3$ at constant T , P , and μ_1 , where one is interested in the effect of a single cosolvent (3) on a biomolecular equilibrium when the number of biomolecules is constant, and the system is also closed with respect to the cosolvent. To develop expressions for the effect of a cosolvent on the equilibrium in this situation we note that P is constant and hence from Eq. (20) one has

$$0 = \rho_2 d\mu_2 + \rho_3 d\mu_3. \quad (38)$$

This can then be used to eliminate $d\mu_2$ from Eq. (17) to give

$$RT \left(\frac{\partial \ln K}{\partial \mu_3} \right)_{T,P,\mu_1,N_2} = N_{A3} - nN_{M3} - \frac{\rho_3}{\rho_2} (N_{A2} - nN_{M2}). \quad (39)$$

To relate this to the cosolvent concentration one can eliminate $d\mu_2$ from Eq. (18) with $i=3$, and then use Eq. (38) to provide

$$RT \left(\frac{\partial \ln \rho_3}{\partial \mu_3} \right)_{T,P,\mu_1,N_2} = 1 + N_{33} - N_{23}. \quad (40)$$

Alternatively, if one is interested in the effect of cosolvent molality one can eliminate $d\mu_2$ from Eq. (19) with $i=3$, and then use Eq. (38) to generate

$$RT \left(\frac{\partial \ln m_3}{\partial \mu_3} \right)_{T,P,\mu_1,N_2} = 1 + N_{33} - N_{13} - N_{23} + m_3 N_{21}, \quad (41)$$

which can then be combined with Eq. (39) if desired.

XI. SYMMETRIC IDEAL SOLUTIONS

Ideal behavior occurs when the right hand side of Eqs. (29), (30), (36), and (37), etc., is unity. Clearly, the corresponding relationships between the KBIs depend on the concentration scale and the ensemble. Symmetric ideal (SI) solutions provide a useful reference point for understanding the properties of real solution mixtures in closed ensembles at constant temperature and pressure.²⁶ SI solutions are defined by the fact that the mole fraction scale activity coefficients are unity for all components over all compositions, i.e., $d\mu_i = RTd \ln x_i$. In terms of the KBIs this requirement is fulfilled when $\Delta G_{ij} = G_{ii} + G_{jj} - 2G_{ij} = 0$ for all i, j pairs.²⁶ Recently, we provided a general expression for the KB integrals in SI solutions of any number of components,⁵⁰

$$G_{ij} = RT\kappa_T - V_i - V_j + S_{n_c} \quad S_{n_c} = \sum_{k=1}^{n_c} \rho_k V_k^2, \quad (42)$$

where the sum is over all n_c components of the mixture, κ_T is the isothermal compressibility, and V_i is the molar volume of pure component i at the same T and P . For our pseudo- $n_c + 1$ component system, one can write

$$S_{n_c+1} = \phi_2 \langle V_2 \rangle + \sum_{k \neq M,A}^{n_c+1} \rho_k V_k^2, \quad (43)$$

$$\langle V_2 \rangle = f_M V_M + f_A V_A = V_2 (f_M + n f_A),$$

$$\phi_2 = \rho_M V_M + \rho_A V_A = (\rho_M + n\rho_A) V_2 = \rho_2 V_2.$$

In developing the above expressions we made the very reasonable assumption that

$$V_A = nV_M = nV_2. \quad (44)$$

For instance, estimated differences in volume between native and denatured proteins are typically small compared to their total volumes.^{27,51,52} The value of $\langle V_2 \rangle$ is the average molar volume of the biomolecule at that particular composition and varies between $V_M = V_2$ for $f_M = 1$ and $V_A = nV_2$ for $f_A = 1$.

Using Eq. (42) one can develop expressions for the various combinations of KB integrals that appeared in the previous sections. Hence, one finds

$$N_{ij} - N_{kj} = \rho_j (V_k - V_i),$$

$$N_{ij} - m_j N_{i1} = \rho_j (V_1 - V_j), \quad (45)$$

$$N_{ij}^+ = m_j,$$

for any i and $j=1, 2$, etc. For KB integrals involving specific forms of the biomolecule we have

$$N_{Ai} - nN_{Mi} = (n-1)\rho_i (V_i - RT\kappa_T - S_{n_c+1}), \quad (46)$$

$$P_{ij}^A - nP_{ij}^M = (n-1)\rho_i (V_i - V_j),$$

for any i and $j=1, 2$, etc., after using Eq. (44). In both cases the mean value of the molar volume $\langle V_2 \rangle$ should be used when i or $j=2$.

For SI solutions one finds that $\mu_{ij} = \delta_{ij} - x_j$ and therefore insertion of the SI expressions into Eq. (23) and performing the summation leads to a general result for SI solutions,

$$\left(\frac{\partial \ln K}{\partial \ln m_j} \right)_{T,P,m_{k \neq j}}^{\text{SI}} = (n-1)\rho_j (V_j - V_m) = (n-1)(\phi_j - x_j), \quad (47)$$

valid for any number of components. Here, $V_m = V/(N_1 + N_2 + \dots)$ is the molar volume, or the average volume per molecule of the solution, and $\phi_j = \rho_j V_j$ is the volume fraction. Hence, an increase in the concentration of the biomolecule or an additive has no effect on the equilibrium in SI solutions when $n=1$, while they increase the equilibrium constant for $n>1$ when their molar volume is larger than the average volume of the solution components. This appears to follow the result expected for a simple excluded volume effect where the excluded volume is smaller for the aggregate compared to an equivalent number of monomers. However, this is incorrect as we have not assumed anything concerning the character (size or shape) of either the M or A species. The real origin of the result is discussed later. SI solution results for other concentration scales are given by

$$\left(\frac{\partial \ln K}{\partial \ln \rho_j} \right)_{T,P,m_{k \neq j}}^{\text{SI}} = (n-1)\rho_j \frac{V_j - V_m}{1 - \phi_j}, \quad (48)$$

$$RT \left(\frac{\partial \ln K}{\partial \mu_j} \right)_{T,P,m_{k \neq j}}^{\text{SI}} = \left(\frac{\partial \ln K}{\partial \ln x_j} \right)_{T,P,m_{k \neq j}}^{\text{SI}} = (n-1)\rho_j \frac{V_j - V_m}{1 - x_j},$$

where we used the previous relationships between the concentration derivatives.

XII. ALTERNATIVE DEFINITIONS OF THE EQUILIBRIUM CONSTANT

Our choice of an equilibrium constant utilizing number densities was motivated by the pseudochemical potential (pcp) approach pioneered by Ben-Naim.⁴⁶ The pcp (μ^*) is related to the total chemical potential by the equation

$$\mu_i = \mu_i^* + RT \ln(\Lambda_i^3 \rho_i), \quad (49)$$

where Λ is the thermal de Broglie wavelength. The equilibrium condition then provides

$$RTd \ln K = -(d\mu_A^* - nd\mu_M^*). \quad (50)$$

The pcg approach helps to simplify the resulting expressions for the same reasons as outlined previously.²⁶ Other definitions of the equilibrium constant are possible, but typically generate extra terms which complicate the analysis. This can be seen from the following expressions which relate various definitions of the equilibrium constant to the one used here. For instance,

$$d \ln K_m = d \ln K + (n-1)d \ln \rho_1, \quad K_m = \left(\frac{m_A}{m_M^n}\right), \quad (51a)$$

$$d \ln K_x = d \ln K + (n-1)d \ln \rho, \quad K_x = \left(\frac{x_A}{x_M^n}\right), \quad (51b)$$

$$d \ln K_f = d \ln K + (n-1)d \ln \rho_2, \quad K_f = \left(\frac{f_A}{f_M^n}\right), \quad (51c)$$

$$d \ln K_\rho = d \ln K + (n-1)d \ln \rho_M, \quad K_\rho = \left(\frac{\rho_A}{\rho_M^n}\right), \quad (51d)$$

where ρ is the total number density. The final terms in Eqs. (51a)–(51d) represent additional contributions which depend on the ensemble. They are properties of the solution mixture itself, and not directly related to the affinity of any of the species to the different forms of the biomolecule (see later). Also, we note that defining an equilibrium constant by $K' = n\rho_A/\rho_M^n$ does not affect any of the results presented here as $d \ln K' = d \ln K$. The final expression [Eq. (51d)] does not include the factor of n for the equilibrium constant. Hence, this does not correctly reflect the stoichiometry of the association reaction. However, it may relate more meaningfully to the experimental data which are typically indirect measures of biomolecular concentrations obtained through spectroscopic or activity measurements, or for solutions where a unique value of n might not be known. Clearly, different results are obtained with different definitions of the equilibrium constant, and therefore comparisons with experimental data should be performed with care.

XIII. APPROXIMATE FREE ENERGY CURVES

KB theory provides expressions for derivatives of the equilibrium constant (or free energy) for a particular process in terms of the intermolecular distributions observed in solution. To obtain changes in the equilibrium constant one has to then integrate. Unfortunately, the general dependence of the N_{ij} on composition is unknown. However, one can make some reasonable approximations and investigate the results. First, most studies involve the solute (2) at low concentrations. Second, if the G_{ij} are assumed to be relatively constant then one can obtain approximate free energy curves. Therefore, for two of the main ternary systems [Eqs. (28) and (29), and Eqs. (35) and (36)] one finds

$$-\beta\Delta\Delta G = \ln\left(\frac{K}{K_0}\right)_{T,P,m_2 \rightarrow 0} = \frac{P_{31}^A - nP_{31}^M}{N_{33} - N_{13}} \ln(1 + N_{33} - N_{13}), \quad (52)$$

$$-\beta\Delta\Delta G' = \ln\left(\frac{K}{K_0}\right)_{T,\mu_1,\rho_2 \rightarrow 0} = \frac{N_{A3} - nN_{M3}}{N_{33}} \ln(1 + N_{33}), \quad (53)$$

which may be compared to the corresponding exact SI result,

$$-\beta\Delta\Delta G^{\text{SI}} = \ln\left(\frac{K}{K_0}\right)_{T,P,m_2 \rightarrow 0}^{\text{SI}} = -(n-1) \ln[1 + \rho_3(V_1 - V_3)], \quad (54)$$

where $\Delta\Delta G = \Delta G(\rho_3) - \Delta G(0)$, $\beta = 1/RT$, and K_0 is the equilibrium constant in the absence of cosolvent. A similarity to denaturation binding models is clearly apparent.^{53–55} For small ρ_3 the right hand side reduces to $P_{31}^A - nP_{31}^M$ and $N_{A3} - nN_{M3}$ for the closed and open systems, respectively. Hence, using these approximations the cosolvent effect is predicted to be linear in cosolvent molarity for low cosolvent concentrations—a result observed experimentally.⁵⁶ We note that the infinitely dilute biomolecule limit applied to Eq. (38) results in $d\mu_3 = 0$ for the T , P , and μ_1 ensemble.

There are additional relationships relating the various KBIs which can be related to the properties of solution mixtures. For instance, taking derivatives of Eq. (17) with respect to P and keeping T and all N_i constant provides,

$$(1-n)RT\kappa_T = \sum_i^{n_c} (N_{Ai} - nN_{Mi})\bar{V}_i, \quad (55)$$

where \bar{V}_i is the partial molar volume (pmv). The above expression assumes that N_A and N_M are reasonably independent of pressure. Equation (55) can be developed further for the case of $n_c = 3$ and an infinitely dilute biomolecule, by assuming that the left hand side of Eq. (55) is small and may be neglected. This is true for $n = 1$ and should be reasonable for $n < 10$ for all but small cosolvent concentrations. In this case one can relate the preference of 3 over 1 for the biomolecule to just the affinity of 3 for both forms,

$$P_{31}^A - nP_{31}^M \approx \frac{N_{A3} - nN_{M3}}{\phi_1}. \quad (56)$$

This is a generalization of a previous result.^{27,57} The relationship exists because, under these approximations, an increase in the local density of one species around the biomolecule should be accompanied by some degree of depletion of another species in the same region. Using Eq. (56) in Eq. (52) and comparing with Eq. (53) one can predict the relative effects of the addition of a cosolvent in both open and closed systems. The result is

$$\phi_1 \Delta\Delta G \approx \Delta\Delta G' \quad (57)$$

for low biomolecule and cosolvent concentrations. This also assumes that the KBIs are reasonably independent of the osmotic pressure, i.e., $G_{ij}(T, N, P) \approx G_{ij}(T, N, P + \Pi)$.

Therefore, the cosolvent effect should be larger in magnitude in a closed system compared to a similar corresponding semiopen system.

XIV. RELATIONSHIP TO THE BEN-NAIM RESULT FOR CLOSED SYSTEMS

Previously, Ben-Naim studied the effects of cosolvents on molecular association using KB theory.⁴³ In particular, the effect of a cosolvent on an equilibrium involving an infinitely dilute solute expressed in terms of the quantity $(\partial N_A / \partial N_3)_{T,P,m_2 \rightarrow 0}$ in closed systems. The results presented here are different from the Ben-Naim expression for several reasons. First, we used an equilibrium constant in terms of numbers densities instead of just N_A . Second, the cosolvent concentration has been expressed in terms of molarity or molality rather than N_3 . This does not change the general conclusions obtained from the respective expressions, although it does change the expressions themselves. To illustrate further we will transform our result [Eqs. (28) and (30)] into the corresponding Ben-Naim expression. To do this we note that one can define an equilibrium constant using the molecule numbers (K_N). This is related to the equilibrium constant used here by

$$d \ln K = d \ln \left(\frac{N_A}{N_M^n} \right) - d \ln \left(\frac{V}{V^n} \right) = d \ln K_N + (n-1)d \ln V. \quad (58)$$

Taking the appropriate derivative in the required ensemble one obtains

$$\left(\frac{\partial \ln K}{\partial \ln N_3} \right)_{T,P,m_2} = \left(\frac{\partial \ln K_N}{\partial \ln N_3} \right)_{T,P,m_2} + (n-1)\rho_3 \bar{V}_3. \quad (59)$$

Finally, the relationship between changes in the equilibrium constant and changes in N_A can be obtained from Eq. (5) with N_2 constant,

$$d \ln K_N = \frac{N_M + n^2 N_A}{N_A N_M} dN_A. \quad (60)$$

Combining these relationships with Eqs. (28) and (30) one obtains the Ben-Naim result in our notation,

$$\begin{aligned} & \left(\frac{\partial N_A}{\partial N_3} \right)_{T,P,m_2 \rightarrow 0} \\ &= \frac{N_A N_M}{N_3 (N_M + n^2 N_A)} \frac{P_{31}^A - n P_{31}^M - (n-1)m_3(1 + N_{11} - N_{31})}{1 + N_{33}^+}, \end{aligned} \quad (61)$$

where we used the KB expression for the pmv of 3 in a binary solution of 1 and 3.⁴⁶ Clearly, the expression presented here is simpler and easier to interpret. This is a direct consequence of investigating the changes in an equilibrium constant defined using number densities, i.e., the pcp approach.

XV. DISCUSSION

In the above sections we provided an analysis, using KB theory, of the effects of increasing biomolecule and cosolvent concentrations on molecular association in solution. KBIs involving species 2, M , and A appear in these expressions, even though they are different representations of the same species. This was done deliberately. The use of N_{22} in the above expressions serves to indicate that this corresponds to a KBI between all forms of the biomolecule where one does not distinguish between monomer and aggregate. Alternatively, the use of N_{M3} (or N_{A3}) indicates a correlation between M and 3 (or A and 3) which is specific for one particular form of the biomolecule. In addition, the use of N_{M2} (or N_{A2}) signifies a correlation between the monomer (or aggregate) form and any other form of the biomolecule, both monomer and aggregate. Hence, we have nonspecific effects involving the KBIs for species 2, together with specific effects involving the KBIs for species M or A .

The resulting expressions for closed systems involve terms of the form $P_{i1}^A - n P_{i1}^M$. These terms quantify the excess (or depletion) of i over 1 in the vicinity of an aggregate over the same excess (or depletion) of i in the vicinity of n monomers. Therefore, if a species i displays a higher affinity for the aggregate than n individual monomers, then $P_{i1}^A - n P_{i1}^M > 0$ and an increase in the concentration of i leads to an increase in equilibrium constant and a shift in the equilibrium toward the associated form. The cosolvent urea serves as a reasonable example. Urea is well known to denature proteins.⁵⁶ Hence, when $n=1$ the preference of urea (over water) for the denatured state must be larger, on average, than that for the native state. Assuming this is related to the corresponding increase in exposed surface area upon denaturation,⁵⁸ this suggests that urea should stabilize the monomer over the aggregate as an aggregate typically possesses less surface area than n monomers. Hence, urea is a denaturant of proteins but should help prevent protein aggregation. This behavior has been observed experimentally.^{59,60} The opposite effects will be observed for classic osmolytes.

Some general trends are observed in the results. From Eq. (17) it is clear that the effect on the equilibrium does not involve any KBIs between the biomolecule and any species that is kept at a constant chemical potential. Any species at constant N will contribute to changes in the equilibrium as quantified through the KBIs. Hence, all the species contribute in closed systems [see Eq. (23)]. The main reason for this is quite simple. When a cosolvent displaces a solvent molecule in the vicinity of the biomolecule both the addition of cosolvent and the removal of a solvent (or other) molecule from the biomolecule affect the Gibbs free energy of the system. On the other hand, when a cosolvent displaces a solvent molecule in a system open to the solvent, only the addition of the cosolvent affects the free energy of the system. A change in concentration scale ($\mu \rightarrow m \rightarrow \rho$) typically results in a scaling effect [Eq. (26)], but this does not alter the sign of the effects contributing to the change in equilibrium constant.

The expression provided in Eq. (47) for closed SI solutions indicates that the association equilibrium is increased

by the addition of additives with a larger molar volume than the average molecular volume of the solution at that composition. The same expression can be obtained from Eq. (51b) noting that $d \ln K_x = 0$ for SI solutions [see Eq. (7)]. Hence, we have two contributions to changes in the equilibrium constant for SI solutions. First, on addition of a cosolvent the volume of the solution increases. This affects the number densities of M and A , and directly increases the equilibrium constant when $n > 1$. Second, the addition of another particle to the system changes the total number of species present and therefore the mole fraction of each species. This causes an increase in the monomer form to maintain the equilibrium condition. Finally, we note that dilution, i.e., the addition of solvent, will favor dissociation as long as the solvent has a low molar volume. These effects dominate in the absence of a particular affinity between any of the pairs of species in solution. There is no effect on a simple denaturation equilibrium in SI solutions.

It is well known that cellular crowding, as described by the addition of rather large cosolvents, will tend to favor association.⁶ The usual explanation for this observation is that aggregates display less surface area, and therefore less excluded volume, compared to a collection of their monomers. Excluded volume effects also favor the native state over the denatured state for the $n=1$ case. In contrast, the SI result discussed above indicates that any excluded volume contributions are balanced by other (favorable) contributions so that $\Delta G_{ij}=0$ for all i, j pairs. In a hard sphere fluid model, as adopted in SPT, the SI condition does not hold and excluded volume contributions are dominant. These also favor the associated or native forms. However, we emphasize that even for ideal (mole fraction scale) solutions the addition of a large cosolvent favors association when the equilibrium constant is defined using molarities, and this effect is not related to any change in excluded volume.

The results predicted by Eq. (54) for closed ternary SI solutions are illustrated in Fig. 2. A variety of cosolvents are considered with different molar volumes relative to the primary solvent. The effects are linear in volume fraction at low cosolvent concentrations but display deviation from linear behavior at higher concentrations. All the results scale with $n-1$ and are independent of the size or shape of the monomer or aggregate, although Eq. (44) was used during the derivation. The magnitude of the effects obtained from Eq. (54) for closed SI solutions is similar to that predicted by SPT depending on the biomolecule shapes used (see Fig. 3 from Ref. 6, for instance). Clearly, we have two solution models that display similar results in qualitative agreement with experiment. Both are approximations to real solution conditions of biological interest. The SI approach represents a thermodynamic model, while SPT presents a physical model for solutions. It is unclear at present which, if any, is more reasonable.

The analysis of cosolvent effects on a chemical equilibrium using KB theory is exact. The KBIs quantify the local composition of the solution surrounding the various species corresponding to a four component system for the case of a single cosolvent. There is no problem extracting all the required KBIs from computer simulations as long as one can

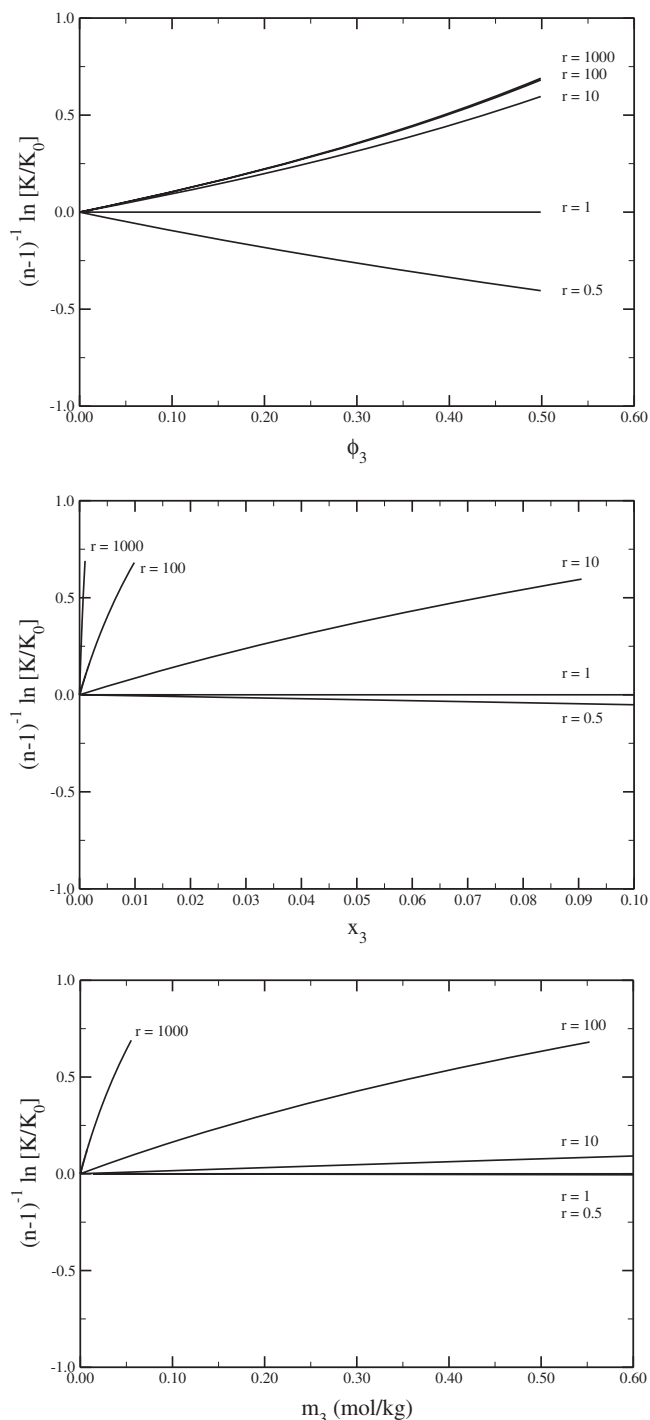


FIG. 2. SI results for the effects of a single crowder (3) on the association equilibrium ($nM \rightarrow A$) of an infinitely dilute solute (2) in a primary solvent (1) in a closed system. The data were obtained using Eq. (54) for different molar volume ratios (r) of the crowder and solvent such that $V_3 = rV_1$. The results are plotted as a function of cosolvent volume fraction ($\phi_3 = \rho_3 V_3$), cosolvent mole fraction (x_3), and cosolvent molality (m_3) with water as the solvent. In this case $x_3 = \phi_3 / [\phi_3 + r(1 - \phi_3)]$ and $m_3 = 1000\phi_3 / r(1 - \phi_3) / 18.015$. For reference, the molar volume of pure water is $V_1 = 18 \text{ cm}^3/\text{mol}$ and so $r=1000$ would correspond to a 25 kDa protein of approximately 225 residues. All curves are truncated at a cosolvent volume fraction of 0.5.

define the aggregate in a consistent manner. From a thermodynamic point of view, however, this is formally a three component system. Consequently, it is difficult to extract the individual KBIs from an analysis of the thermodynamic data

alone. For instance, a general n_c component system has $n_c(n_c+1)/2$ unique G_{ij} integrals. To determine the integrals from experimental data using the KB inversion approach requires 1 compressibility, n_c-1 independent pmvs, and therefore $n_c(n_c-1)/2$ independent chemical potential derivatives (μ_{ij}) as a function of composition. However, for our pseudo-four component system one can use the relationships in Eqs. (28)–(30) as long as one knows K and n for a particular system as a function of composition. Hence, one still has only six ($n_c=3$) unique KBIs as many of the ten ($n_c=4$) KBIs are related through Eqs. (14) and (15) and therefore are not unique.

XVI. CONCLUSIONS

In this study we applied KB theory to investigate the effects of cosolvents on molecular equilibria in solution. This is an extension of the previous approach of Ben-Naim to include any number of components at finite concentrations in both closed and semiopen systems. No assumption concerning the character, specifically the size and/or shape, of the molecules has been made. Hence, the results are totally general and can be applied to a variety of processes such as protein denaturation, protein aggregation, and cellular crowding, where the cosolvent can be either small (H^+ , urea) or large (another protein). In addition, the extension to include semiopen systems also allows one to consider systems under typical biological (cellular) conditions. Here, the effect of a cosolvent appears to be reduced compared to closed systems.

The results obtained here depend on the definition of the equilibrium constant. We advocate the use of molarities in accord with the pcp approach of Ben-Naim. In this case, even SI solutions, where significant excluded volume effects are absent, indicate an increase in association on the addition of a cosolvent crowder. This does not mean that excluded volume effects are not important in these systems. Merely that at least some of the typical effects can be explained without invoking excluded volume. Significant deviations from SI behavior will be observed when the cosolvent displays a preference for either form of the biomolecule.

A particular advantage of the KB approach is that one has exact expressions for the cosolvent effects in terms of KBIs which are directly related to molecular distributions. Hence, one can develop a series of models, each providing different approximations to the real KBIs, and thereby determine the resulting thermodynamic effects of various approaches. We are currently using this type of approach to compare in more detail the KB based results for hard sphere crowders to those obtained from the traditional SPT approach.

ACKNOWLEDGMENTS

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.

- ¹P. van den Berg, E. W. Chung, C. V. Robinson, P. L. Mateo, and C. M. Dobson, *EMBO J.* **18**, 4794 (1999).
- ²H. X. Zhou and K. A. Dill, *Biochemistry* **40**, 11289 (2001).
- ³M. S. Cheung, D. Klimov, and D. Thirumalai, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 4753 (2005).
- ⁴R. J. Ellis and A. P. Minton, *Biol. Chem.* **387**, 485 (2006).
- ⁵K. Richter, M. Nessling, and P. Lichter, *Biochim. Biophys. Acta* **1783**, 2100 (2008).
- ⁶H. X. Zhou, G. N. Rivas, and A. P. Minton, *Annual Review of Biophysics* **37**, 375 (2008).
- ⁷L. A. Munishkina, A. Ahmad, A. L. Fink, and V. N. Uversky, *Biochemistry* **47**, 8993 (2008).
- ⁸C. M. Dobson, *Nature (London)* **426**, 884 (2003).
- ⁹C. A. Ross and M. A. Poirier, *Nat. Med.* **10**, S10 (2004).
- ¹⁰S. N. Timasheff, *Adv. Protein Chem.* **51**, 355 (1998).
- ¹¹J. Wyman and S. J. Gill, *Binding and Linkage* (University Science Books, Mill Valley, California, 1990).
- ¹²C. Tanford, *Adv. Protein Chem.* **24**, 1 (1970).
- ¹³J. A. Schellman, *Biopolymers* **17**, 1305 (1978).
- ¹⁴V. Pierce, M. Kang, M. Aburi, S. Weerasinghe, and P. E. Smith, *Cell Biochem. Biophys.* **50**, 1 (2008).
- ¹⁵T. L. Hill, *J. Chem. Phys.* **34**, 1974 (1961).
- ¹⁶C. Tanford, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 1811 (1974).
- ¹⁷F. H. Stillinger and A. Ben-Naim, *J. Chem. Phys.* **74**, 2510 (1981).
- ¹⁸A. P. Minton, *Mol. Cell. Biochem.* **55**, 119 (1983).
- ¹⁹A. P. Minton, G. C. Colclasure, and J. C. Parker, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 10504 (1992).
- ²⁰A. P. Minton, *Biophys. J.* **63**, 1090 (1992).
- ²¹V. A. Parsegian, R. P. Rand, and D. C. Rau, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 3987 (2000).
- ²²S. B. Zimmerman and A. P. Minton, *Annu. Rev. Biophys. Biomol. Struct.* **22**, 27 (1993).
- ²³H. X. Zhou, *Proteins* **72**, 1109 (2008).
- ²⁴M. Wojciechowski and M. Cieplak, *Biosystems* **94**, 248 (2008).
- ²⁵J. G. Kirkwood and F. P. Buff, *J. Chem. Phys.* **19**, 774 (1951).
- ²⁶A. Ben-Naim, *Molecular Theory of Solutions* (Oxford University Press, New York, 2006).
- ²⁷M. Aburi and P. E. Smith, *J. Phys. Chem. B* **108**, 7382 (2004).
- ²⁸M. Kang and P. E. Smith, *Fluid Phase Equilib.* **256**, 14 (2007).
- ²⁹P. E. Smith, *Biophys. J.* **91**, 849 (2006).
- ³⁰S. Shimizu and C. L. Boon, *J. Chem. Phys.* **121**, 9147 (2004).
- ³¹S. Shimizu, W. M. McLaren, and N. Matubayasi, *J. Chem. Phys.* **124**, 234905 (2006).
- ³²I. L. Shulgin and E. Ruckenstein, *J. Chem. Phys.* **123**, 054909 (2005).
- ³³D. Trzesniak, N. F. A. Van Der Veegt, and W. F. van Gunsteren, *Mol. Phys.* **105**, 33 (2007).
- ³⁴J. Rosgen, B. M. Pettitt, and D. W. Bolen, *Protein Sci.* **16**, 733 (2007).
- ³⁵P. E. Smith, *J. Phys. Chem. B* **103**, 525 (1999).
- ³⁶R. Chitra and P. E. Smith, *J. Phys. Chem. B* **105**, 11513 (2001).
- ³⁷S. Shimizu, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1195 (2004).
- ³⁸S. Shimizu and N. Matubayasi, *Chem. Phys. Lett.* **420**, 518 (2006).
- ³⁹E. Matteoli and G. A. Mansoori, *Fluctuation Theory of Mixtures* (Taylor & Francis, New York, 1990).
- ⁴⁰Y. Marcus, *Monatsch. Chem.* **132**, 1387 (2001).
- ⁴¹A. Ben-Naim, A. M. Navarro, and J. M. Leal, *Phys. Chem. Chem. Phys.* **10**, 2451 (2008).
- ⁴²J. Roesgen, in *Methods in Molecular Biology*, edited by J. W. Shriver (Humana, New York, 2009), Vol. 490, Chap. 9, pp. 195–225.
- ⁴³A. Ben-Naim, *J. Chem. Phys.* **63**, 2064 (1975).
- ⁴⁴D. G. Hall, in *Thermodynamics of Micellar Solutions*, edited by E. Wyn-Jones and J. Gormally (Elsevier, Amsterdam, 1983), Vol. 26, Chap. 2, pp. 7–69.
- ⁴⁵M. Kang and P. E. Smith, *J. Chem. Phys.* **128**, 244511 (2008).
- ⁴⁶A. Ben-Naim, *Statistical Thermodynamics for Chemists and Biochemists* (Plenum, New York, 1992).
- ⁴⁷P. E. Smith, *J. Phys. Chem. B* **108**, 18716 (2004).
- ⁴⁸D. G. Hall, *Trans. Faraday Soc.* **67**, 2516 (1971).
- ⁴⁹I. Prigogine and R. Defay, *Chemical Thermodynamics* (Longmans, London, 1954).
- ⁵⁰P. E. Smith, *J. Chem. Phys.* **129**, 124509 (2008).
- ⁵¹C. A. Royer, *Biochim. Biophys. Acta* **1595**, 201 (2002).
- ⁵²P. E. Smith, *Biophys. Chem.* **113**, 299 (2005).
- ⁵³J. A. Schellman, *Biophys. Chem.* **96**, 91 (2002).

- ⁵⁴ K. C. Aune and C. Tanford, *Biochemistry* **8**, 4586 (1969).
⁵⁵ G. I. Makhatadze, *J. Phys. Chem. B* **103**, 4781 (1999).
⁵⁶ R. F. Greene, Jr., and C. N. Pace, *J. Biol. Chem.* **249**, 5388 (1974).
⁵⁷ S. Shimizu, *J. Chem. Phys.* **120**, 4989 (2004).

- ⁵⁸ J. K. Myers, C. N. Pace, and J. M. Scholtz, *Protein Sci.* **4**, 2138 (1995).
⁵⁹ D. K. Klimov, J. E. Straub, and D. Thirumalai, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14760 (2004).
⁶⁰ N. J. Nosworthy and A. Ginsburg, *Protein Sci.* **6**, 2617 (1997).