Kirkwood–Buff integrals for ideal solutions

Elizabeth A. Ploetz,[1] Nikolaos Bentenitis,[2] and Paul E. Smith[1,a)]
[1]*Department of Chemistry, Kansas State University, Manhattan, Kansas 66506, USA*
[2]*Department of Chemistry and Biochemistry, Southwestern University, Georgetown, Texas 78626, USA*

(Received 20 May 2009; accepted 29 March 2010; published online 22 April 2010)

The Kirkwood–Buff (KB) theory of solutions is a rigorous theory of solution mixtures which relates the molecular distributions between the solution components to the thermodynamic properties of the mixture. Ideal solutions represent a useful reference for understanding the properties of real solutions. Here, we derive expressions for the KB integrals, the central components of KB theory, in ideal solutions of any number of components corresponding to the three main concentration scales. The results are illustrated by use of molecular dynamics simulations for two binary solutions mixtures, benzene with toluene, and methanethiol with dimethylsulfide, which closely approach ideal behavior, and a binary mixture of benzene and methanol which is nonideal. Simulations of a quaternary mixture containing benzene, toluene, methanethiol, and dimethylsulfide suggest this system displays ideal behavior and that ideal behavior is not limited to mixtures containing a small number of components. © *2010 American Institute of Physics*. [doi:10.1063/1.3398466]

I. INTRODUCTION

The Kirkwood–Buff (KB) theory of solutions has provided a wealth of data and insight into the properties of solution mixtures.[1–13] KB theory can be applied to any stable mixture containing any number of components. In particular, the theory provides exact relationships from the molecular distributions between the various species in solution to the corresponding thermodynamics of the mixture. The central quantities of interest in KB theory are the KB integrals (KBIs),[1]

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

where g_{ij} is the corresponding radial distribution function (rdf) between species i and j, and r is their intermolecular separation. In applying KB theory to understand solutions, it is often useful to compare the properties of a real solution with those of a corresponding ideal solution.[2,6,14,15] In particular, deviations from ideal solution behavior provide indications of specific associations or affinities between the solution components which affect the thermodynamics.[16,17] Furthermore, assuming ideal behavior for some components may be the only possible approach to describe solutions containing a large number of components, where experimental data are usually very rare.[18,19] Consequently, it is important to understand the nature of ideal solutions as formulated by KB theory.

Ideality in solution mixtures at constant pressure (P) and temperature (T) can be expressed using a variety of concentration scales. The most common is the mole fraction concentration scale where the corresponding activity coefficients are unity for all species at all solution compositions. These mixtures are also known as symmetric ideal (SI) or perfect solutions. In SI solutions the KBIs are neither zero nor independent of composition. The KBIs for these solutions can be expressed in terms of the isothermal compressibility (κ_T) and molar volumes (V_i) of the pure components at the same T and P. Expressions for binary, ternary, and quaternary solutions are available.[2,20,21] Ben-Naim has also shown that SI solutions including up to four components satisfy $\Delta G_{ij} = G_{ii} + G_{jj} - 2G_{ij} = 0$ for all i,j pairs.[22] Our recent analysis of the results obtained from the KB theory of solutions for $n=1$ to 4 components proposed the following general expression for the KBIs:[21]

$$G_{ij}^{\mathrm{SI}} = RT\kappa_T - V_i - V_j + S_n, \quad S_n = \sum_{k=1}^{n} \rho_k V_k^2, \tag{2}$$

where R is the gas constant and $\rho_i = N_i/V$ are the number densities of each species. The above expression was postulated to be valid for any i and j combination in any n component SI solution. This can be shown to be true for $n=1-4$ components.[21] Here we provide rigorous proof that this expression is valid for any number of components at constant T and P. Furthermore, we also determine the corresponding expressions for the KBIs which result in ideal behavior on the molality and molarity concentration scales. The results are then illustrated using data obtained from computer simulations.

II. THEORY

The chemical potential (μ) of a solute in solution can be expressed in a variety of ways using a series of different concentration scales. There are three major expressions related to whether the species concentrations are represented in terms of mole fractions (x), molalities (m), or molarities (c). It is then typical to write

$$\mu_i = \mu_i^{o,x} + RT \ln f_i x_i = \mu_i^{o,x} + \mu_i^{ex} + RT \ln x_i,$$

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

$$\mu_i = \mu_i^{o,m} + RT \ln \gamma_i m_i,$$

$$\mu_i = \mu_i^{o,c} + RT \ln y_i c_i,$$

$$\mu_i = \mu_i^{*} + RT \ln \Lambda_i^3 \rho_i, \tag{3}$$

where a set of activity coefficients (f, γ, y) are used to quantify deviations from ideal behavior on the mole fraction, molality, and molarity concentration scales, respectively. The final equation is the statistical mechanical expression for the chemical potential in terms of the pseudochemical potential (μ^*), thermal de Broglie wavelength (Λ), and number density.[2] The standard chemical potentials (μ^o) can refer to an infinitely dilute solute or the pure solute at the same T and P. They are constants and independent of the concentration of i, whereas the pseudochemical potential is composition dependent. The following discussion is focused on changes in the chemical potentials with composition, and therefore the choice of standard state is largely irrelevant. We also note that $d\mu^* = RT\ d \ln y$ when T is constant. It is generally known that the activity coefficients decrease when the solute displays significant solute self-association, whereas the activity coefficients increase when the solute is significantly solvated by the other species in solution.

The general approach used here to determine the required KBIs involves generating a series of expressions involving derivatives of the chemical potentials which are valid for real solutions, followed by application of the appropriate conditions corresponding to ideality for the various concentration scales. Let us consider the species number densities in the grand canonical ensemble to be functions of temperature and all the chemical potentials. One can then write[23]

$$d \ln \rho_i = \beta \sum_{j=1}^{n} (\delta_{ij} + N_{ij}) d\mu_j, \tag{4}$$

which is valid for changes in the number density of any component in any multicomponent system and any (thermodynamically reasonable) ensemble with T constant. Here, $N_{ij} = \rho_j G_{ij}$, $\beta = 1/RT$, and δ_{ij} is the Kroenecker delta function. Two further sets of equations can be derived from the above equation. If we take derivatives with respect to $\ln N_k$, while keeping T, P, and all other $N_{j \neq k}$ constant, we find an expression for the partial molar volumes $(\overline{V})$, written here in terms of volume fractions $(\phi_i = \rho_i \overline{V_i})$,

$$\delta_{ik} - \phi_k = \sum_{j=1}^{n} (\delta_{ij} + N_{ij}) \mu_{jk}, \tag{5}$$

for any i species and where we defined

$$\beta \left(\frac{\partial \mu_i}{\partial \ln N_j} \right)_{T,P,N_{k \neq j}} = \mu_{ij}. \tag{6}$$

The relationships to derivatives on other concentration scales are given by

$$\beta \left(\frac{\partial \mu_i}{\partial \ln x_j} \right)_{T,P} = \frac{\mu_{ij}}{\delta_{ij} - x_j},$$

$$\beta \left(\frac{\partial \mu_i}{\partial \ln m_j} \right)_{T,P,m_{k \neq j}} = \mu_{ij},$$

$$\beta \left(\frac{\partial \mu_i}{\partial \ln \rho_j} \right)_{T,P,N_{k \neq j}} = \frac{\mu_{ij}}{\delta_{ij} - \phi_j}, \tag{7}$$

which are general for any number of components. Alternatively, if one starts from Eq. (4) and then takes derivatives with respect to pressure with all N_j and T constant one finds that

$$RT\kappa_T = \sum_{j=1}^{n} (\delta_{ij} + N_{ij}) \overline{V_j} \tag{8}$$

for any i component.

The above equations relate the KBIs to properties of the solution. It is easier, albeit less general, if we select a particular species to continue. We choose $i=1$ and $k \neq 1$, and then write Eqs. (5) and (8) (after multiplication by ρ_1) in an $n \times n$ matrix form so that

$$\begin{bmatrix} \rho_1 \overline{V_1} & \rho_1 \overline{V_2} & \rho_1 \overline{V_3} & \cdots & \rho_1 \overline{V_n} \\ \mu_{12} & \mu_{22} & \mu_{32} & & \mu_{n2} \\ \mu_{13} & \mu_{23} & \mu_{33} & & \mu_{n3} \\ \vdots & & & \ddots & \vdots \\ \mu_{1n} & \mu_{2n} & \mu_{3n} & \cdots & \mu_{nn} \end{bmatrix} \begin{bmatrix} 1 + N_{11} \\ N_{12} \\ N_{13} \\ \vdots \\ N_{1n} \end{bmatrix} = \begin{bmatrix} \rho_1 RT\kappa_T \\ -\phi_2 \\ -\phi_3 \\ \vdots \\ -\phi_n \end{bmatrix}. \tag{9}$$

Hence, we have a set of simultaneous equations which can be solved quite easily for the required KBIs to give

$$\begin{bmatrix} 1 + N_{11} \\ N_{12} \\ N_{13} \\ \vdots \\ N_{1n} \end{bmatrix} = M_n^{-1} \begin{bmatrix} \rho_1 RT\kappa_T \\ -\phi_2 \\ -\phi_3 \\ \vdots \\ -\phi_n \end{bmatrix}, \tag{10}$$

where M_n is the matrix from Eq. (9). This is the approach we used recently for outlining a general KB inversion procedure.[21] After the appropriate chemical potential derivatives are used in the M_n matrix, and its inverse has been determined, the expressions for the KBIs representing our series of ideal solutions can be obtained from Eq. (10).

III. METHODS

All mixtures were simulated via classical molecular dynamics techniques using the GROMACS program (version 3.3).[24] The simulations were performed in the isothermal-isobaric NpT ensemble at the temperature and pressure of interest using Berendsen thermostats and barostats.[25] All solute and solvent bonds were constrained using the LINCS algorithm.[26] A 2 fs time step was used for integration of the equations of motion. Electrostatic interactions were evaluated using the particle-mesh-Ewald technique,[27] with cutoff distances of 1.2 nm and of 1.5 nm for the real space electrostatic and van der Waals interactions, respectively. Details of the force fields are provided elsewhere.[28–31] All mixtures were simulated with enough molecules to occupy a cubic

box of length of ≈6 nm, starting from random initial configurations, and were simulated for between 10 and 20 ns.

The simulated KBIs have to be truncated in closed systems. We discussed this issue in detail.[28,32,33] Typically, the integral is truncated after several solvation shells (1–1.5 nm) depending on the size of the solute and solvent. To eliminate statistical variations we averaged the integral values over a small range of integration distances corresponding to one solvation shell. The accuracy of this truncation procedure can be checked by determining the simulated partial molar volumes and compressibility using other approaches.[28] Error estimates were obtained from multiple 5 ns block averages.

IV. RESULTS

A. Ideal solutions on the mole fraction concentration scale

The most common and useful reference for solution mixtures is that of SI solutions. The following conditions must hold:

$$d\mu_i^{\mathrm{SI}} = RTd\ \ln x_i, \tag{11}$$

$$\mu_{ij}^{\mathrm{SI}} = \delta_{ij} - x_j, \tag{12}$$

and illustrate that the chemical potential of each component is affected by the addition of j through a change in their mole fractions. For SI solutions [defined by Eq. (11)] the required matrix reduces to

$$M_n^{\mathrm{SI}} = \begin{bmatrix} \rho_1 V_1 & \rho_1 V_2 & \rho_1 V_3 & \cdots & \rho_1 V_n \\ -x_2 & 1-x_2 & -x_2 & & -x_2 \\ -x_3 & -x_3 & 1-x_3 & & -x_3 \\ \vdots & & & \ddots & \vdots \\ -x_n & -x_n & -x_n & \cdots & 1-x_n \end{bmatrix}. \tag{13}$$

The determinant of this matrix equals x_1, and the inverse matrix is given by

$$(M_n^{\mathrm{SI}})^{-1} = \begin{bmatrix} 1 & -1+\rho_1(V_1-V_2) & -1+\rho_1(V_1-V_3) & \cdots & -1+\rho_1(V_1-V_n) \\ m_2 & 1+\rho_2(V_1-V_2) & \rho_2(V_1-V_3) & & \rho_2(V_1-V_n) \\ m_3 & \rho_3(V_1-V_2) & 1+\rho_3(V_1-V_3) & & \rho_3(V_1-V_n) \\ \vdots & & & \ddots & \vdots \\ m_n & \rho_n(V_1-V_2) & \rho_n(V_1-V_3) & \cdots & 1+\rho_n(V_1-V_n) \end{bmatrix}, \tag{14}$$

after evaluation of the required cofactors, or by the binomial inverse theorem. Here, $m_i=\rho_i/\rho_1$, are dimensionless molalities. Using this result in Eq. (10) the SI expressions for the KBIs can be obtained for any n component solution such that

$$1 + N_{11}^{\mathrm{SI}} = \rho_1 RT\kappa_T + \sum_{k=2}^{n} [1+\rho_1(V_k - V_1)]\phi_k, \tag{15}$$

$$N_{12}^{\mathrm{SI}} = \rho_2 RT\kappa_T - [1+\rho_2(V_1-V_2)]\phi_2 + \sum_{k=3}^{n} \rho_2(V_k - V_1)\phi_k. \tag{16}$$

The above equations can be simplified further to provide

$$G_{11}^{\mathrm{SI}} = RT\kappa_T - 2V_1 + \sum_{k=1}^{n} \rho_k V_k^2, \tag{17}$$

$$G_{12}^{\mathrm{SI}} = RT\kappa_T - V_1 - V_2 + \sum_{k=1}^{n} \rho_k V_k^2, \tag{18}$$

which are the expressions provided by the proposed Eq. (2). KBIs for the other components can be obtained by a simple index change. The expressions for the KBIs that result in ideal behavior on the molal and molar concentration scales can be found in Appendices A and B.

It is clear that applying the SI condition to a solution mixture will not lead to ideality for the alternative concentration scales. Consequently, it is interesting to note the consequences of SI behavior and their effect on the other activity coefficients. By taking appropriate derivatives of the expressions in Eq. (3), applying the SI condition for $d\mu_i$ [Eq. (11)], and then equating the results it can be shown that

$$\left(\frac{\partial \ln f_i}{\partial x_j}\right)_{T,P}^{\mathrm{SI}} = 0, \quad \left(\frac{\partial \ln \gamma_i}{\partial m_j}\right)_{T,P,m_{k\neq j}}^{\mathrm{SI}} = -x_1,$$

$$\left(\frac{\partial \ln y_i}{\partial \rho_j}\right)_{T,P,m_{k\neq j}}^{\mathrm{SI}} = \frac{V_j - V_m}{1-\phi_j}. \tag{19}$$

After integration one finds

$$f_i^{\mathrm{SI}} = 1, \quad \gamma_i^{\mathrm{SI}} = \frac{\Sigma_{k\neq j} m_k^{\circ}}{m_j + \Sigma_{k\neq j} m_k^{\circ}}, \quad y_i^{\mathrm{SI}} = \frac{V_m}{V_m^{\circ}} \tag{20}$$

for any i on addition of any j component. The zero superscript indicates a molality or molar volume of the solution (V_m) before addition of the j component. Therefore, γ_i always decreases, while y_i will increase if $V_j > V_m$ and vice versa.

B. Deviations from ideal solution behavior

KB theory provides a wealth of data on real solutions in terms of the molecular distributions. Thermodynamically, real solutions are characterized by activity coefficients. General expressions for changes in the activity coefficients can be derived and provide further information concerning the conditions for ideality and how real solutions deviate from these conditions. Using Eqs. (3) and (4), together with a similar manipulation that led to Eq. (C5), one can show that

$$RTd\ln f_i = -\sum_{j=1}^{n}\left(N_{ij} - \sum_{k=1}^{n} x_k N_{kj}\right)d\mu_j \tag{21}$$

for any real or ideal solution at constant T and P. Alternatively, Eq. (C2) leads directly to

$$RTd\ln\ \gamma_i = -\sum_{j=2}^{n} N_{ij}^{+} d\mu_j \tag{22}$$

for $i>1$ and solutions at constant T and P. Here, $N_{ij}^{+}=N_{ij}+m_j(1+N_{11}-N_{i1}-N_{j1})$. Finally, Eq. (4) provides

$$RTd\ln y_i = d\mu_i^{*} = -\sum_{j=1}^{n} N_{ij}d\mu_j = -\sum_{j=2}^{n}(N_{ij}-m_j N_{i1})d\mu_j \tag{23}$$

at constant T and P. Clearly, the simplest expression for real solutions uses the molarity, or pseudochemical potential, approach. Thermodynamically, stable solution mixtures require $\mu_{ii}\geq 0$ and $\mu_{ij}<0$ for all components.[34] Therefore, on increasing the concentration of i the molar activity coefficient of i decreases when $N_{ii}>0$ and $N_{ij}<0$, i.e., when self-association of i dominates, and vice versa. Equations (21)–(23) can also be used to obtain the expressions provided in Eq. (19).

Deviations from ideal behavior for an infinitely dilute solute (2) can be expressed in terms of the KBIs and chemical potential derivatives,

$$-\left(\frac{d\ln y_2}{d\rho_2}\right)_{T,P,m_{k\neq 2}}^{\infty} = (G_{22}^{\infty}-G_{21}^{\infty})+\rho_2^{-1} \times\sum_{j>2}^{n}\rho_j(G_{2j}^{\infty}-G_{21}^{\infty})\mu_{j2}, \tag{24}$$

where a general recursive relationship exists for the derivatives in Eq. (24).[23] Here, the molar activity coefficient will tend to decrease if the distribution of solutes around a central solute (G_{22}) exceeds the distribution of the primary solvent (1) around the solute (G_{21}), but will increase if the distribution of the other cosolvents (G_{2j}) exceeds the distribution of the primary solvent around the solute.

It is also possible to rewrite the usual KBI expressions for binary solutions to emphasize the similarities to the SI expressions. This is useful as the real partial molar volumes are often reasonably independent of composition. Changes in solute activity then dictate the variation of the real KBIs from the SI expressions. Here, one finds that for real solutions,

$$G_{12} = RT\kappa_T + \frac{S_2 - \overline{V_1} - \overline{V_2}}{1+f_{22}}, \tag{25}$$

$$G_{11} = RT\kappa_T + \frac{S_2 - \overline{V_1} - \overline{V_1}}{1+f_{22}} - \frac{f_{22}}{\rho_1(1+f_{22})}, \tag{26}$$

where $f_{22}=\mu_{22}-1=(\partial\ln f_2/\partial\ln x_2)_{P,T}$ and the expression for G_{22} can be obtained from a suitable index change. The second term on the right hand side (rhs) of Eq. (25) is always negative and typically dominates the first term. Both the second and third terms on the rhs of Eq. (26) can be negative or positive. Alternatively, one could express the above KBIs and the deviation from their ideal values using the KB expression for f_{22}.[22] Hence,

$$G_{ij} = RT\kappa_T + (S_2 - \overline{V_i} - \overline{V_j})(1+x_1\rho_2\Delta G_{12}) + \delta_{ij}(1-x_i)\Delta G_{12}. \tag{27}$$

The above equations illustrate that the expressions for the KBIs in real solutions are in fact strongly related to the SI expressions under most reasonable conditions. We also find that

$$\Delta G_{12} = \frac{G_{22} - (RT\kappa_T + S_2 - 2\overline{V_2})}{\phi_1^2} \approx \frac{G_{22}-G_{22}^{\mathrm{SI}}}{\phi_1^2}, \tag{28}$$

where the approximation should be valid when the partial molar volumes do not vary significantly with composition. The above relationships could be useful for modeling deviations from ideal solution behavior—as obtained using the Debye–Hueckel theory, for example.

C. Molecular dynamics simulations of symmetric ideal solutions

The KBIs represent integrals over the rdfs between the various species in solution. The rdfs are closely related to the potential of mean force ($W_{ij}=-RT\ln g_{ij}$) between species pairs obtained after averaging over all other molecules in the system, including other i and j molecules. The only way to gain reliable insights into these rdfs is from integral equation studies or molecular simulations performed with accurate potentials. Here, we present some of our recent molecular dynamics results using force fields specifically designed to reproduce the experimental KBIs. Two binary systems displaying close to SI behavior were studied. The first is the classic example of benzene (BEN) and toluene (TOL). The second involves mixtures of methanethiol (MSH) and dimethylsulfide (MSM). The resulting KBIs are displayed in Fig. 1.

Figure 1 clearly demonstrates that the simulations reproduce, within the statistical errors, the correct KBIs for both mixtures. In addition, the values of the integrals are close to those predicted for SI solutions using the pure molar volumes of 91, 109, 54, and 74 $\mathrm{cm}^3/\mathrm{mol}$ for benzene, toluene, methanethiol, and dimethylsulfide, respectively. We note that not all force fields accurately reproduce the experimental KBIs.[32,33,35,36] The simulated enthalpy of mixing (βH_m) for the BEN/TOL and MSH/MSM mixtures with $x_2=0.5$ was 0.02(1) and 0.01(1), respectively. The experimental value is 0.03 for the BEN/TOL mixture.[37] The corresponding simu-

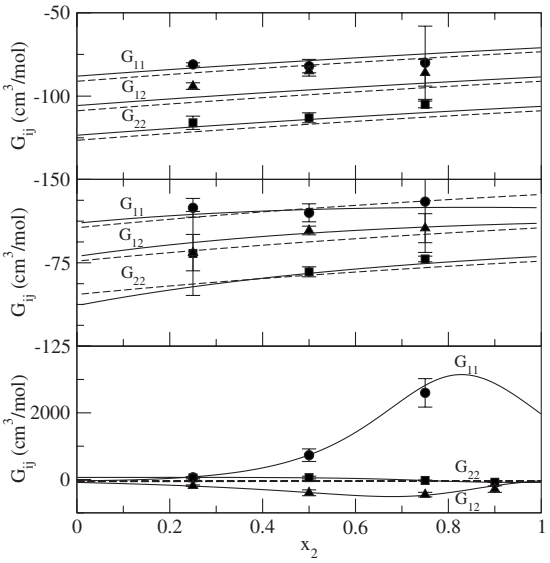

FIG. 1. KBIs for mixtures of benzene (1) and toluene (2) at 313 K (top), methanethiol (1) and dimethylsulfide (2) at 288 K (middle), and methanol (1) and benzene (2) at 308 K (bottom). The experimental KBIs are displayed as solid lines for G_{11} (circles), G_{22} (squares), and G_{12} (triangles), while dashed lines represent the SI solution results [Eq. (2)], and symbols correspond to the simulation data. The experimental KBIs were determined from literature data (Refs. 37, 39, and 48) as described previously (Ref. 29–31).

lated values of ΔG_{ij} were $-25(3)$ and $+10(7)$ $\mathrm{cm}^3/\mathrm{mol}$ for the two mixtures, compared to -1 and -10 $\mathrm{cm}^3/\mathrm{mol}$ observed experimentally.[38,39] All these values indicate close to perfect SI behavior and provide some confidence in the simulated rdfs.

The corresponding rdfs for mixtures with $x_2=0.5$ are displayed in Figs. 2 and 3. The expected similarity between the respective rdfs is apparent, although the two solution mixtures accomplish this in two different ways. The BEN/TOL mixture indicates solvation shells which essentially appear at

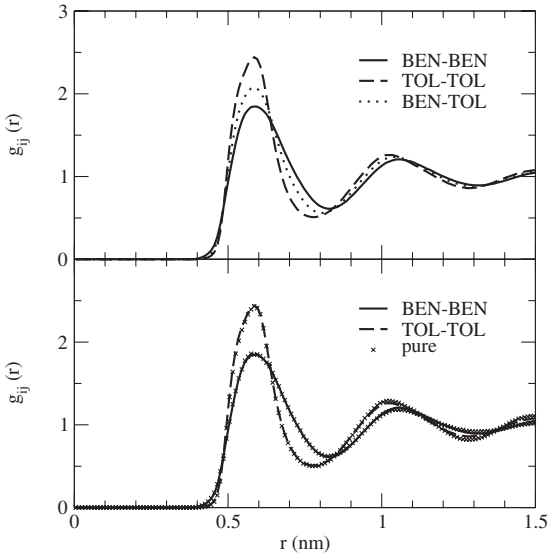

FIG. 2. Center of mass based rdfs obtained from a simulation of benzene (1) and toluene (2) at 313 K. The rdfs for a mole fraction of $x_2=0.5$ are displayed in the top panel. The rdfs for a mole fraction of $x_2=0.5$ are compared with the distributions obtained for the pure solvents at the same temperature in the bottom panel.

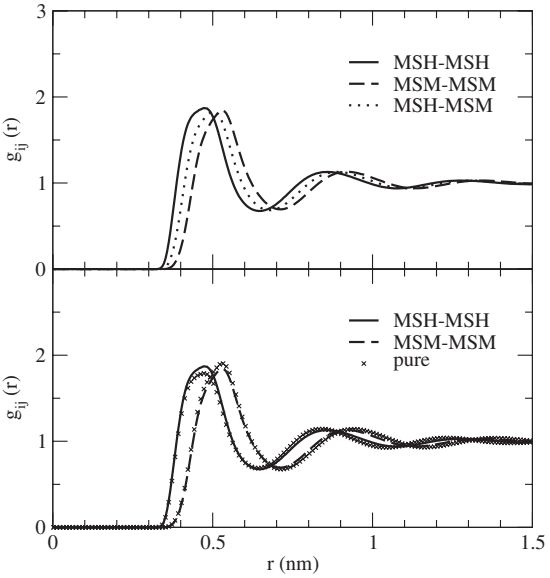

FIG. 3. Center of mass based rdfs obtained from a simulation of methanethiol (1) and dimethylsulfide (2) at 288 K. The rdfs for a mole fraction of $x_2=0.5$ are displayed in the top panel. The rdfs for a mole fraction of $x_2=0.5$ are compared with the distributions obtained for the pure solvents at the same temperature in the bottom panel.

the same distances but with small differences in their magnitude. In contrast, the MSH/MSM mixture displays similar magnitudes for the various solvation shells, but these are shifted to slightly different r values. In both cases the changes from the pure solution values are small. The different behavior observed for the rdfs in the two solutions can be explained by the fact that the BEN/TOL system will be dominated by ring-ring interactions and packing effects which will be essentially the same for both components. Hence, the variation of the rdfs with distance is also similar. However, in the case of MSH/MSM the different sizes of both molecules has a larger effect and the rdfs become shifted relative to each other. They display similar peak heights due to a presumed similarity of their interactions. In both systems g_{12} is intermediate between g_{11} and g_{22} and further ensures that $\Delta G_{ij} \approx 0$.

For comparison, we also included the simulated and experimental results for a nonideal solution involving a mixture of one of the above components (BEN) with methanol (MOH). The KBIs and corresponding rdfs are displayed in Figs. 1 and 4. Clearly, the experimental KBIs indicate significant nonideality, especially at high BEN mole fractions. This is reproduced by the simulations. The rdfs indicate only minor changes for the BEN-BEN distribution, but significant changes in the MOH-MOH distribution. Presumably, these differences arise from the need for MOH molecules to satisfy their hydrogen bonding requirements, which leads to a high degree of self-association at low MOH mole fractions. Similar trends are observed for mixtures of TOL/MOH, MSH/MOH, and MSM/MOH.[29,30]

Mixtures of BEN/TOL and, to some extent, MSH/MSM would be expected to display close to ideal behavior. In contrast, the behavior of BEN/MSH mixtures would be more difficult to predict using basic chemical principles. We performed additional simulations and found that a quaternary

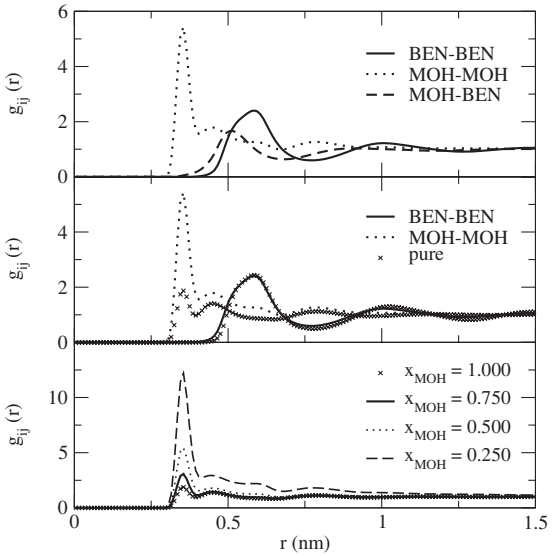

FIG. 4. Center of mass based rdfs obtained from a simulation of benzene (1) and methanol (2) at 308 K. The rdfs for a mole fraction of $x_2=0.5$ are displayed in the top panel. The rdfs for a mole fraction of $x_2=0.5$ are compared with the distributions obtained for the pure solvents at the same temperature in the middle panel. The MOH-MOH rdf is displayed as a function of methanol mole fraction in the bottom panel.

mixture of BEN/TOL/MSH/MSM at $x_i=0.25$ also displays near ideal behavior. The results from this simulation are displayed in Fig. 5 and Table I. Clearly, the rdfs are somewhat different and yet they all possess first and second solvation shells at essentially the same distances, and occur with similar probabilities. The enthalpy of mixing is negligible ($\beta H_m=0.00$). Unfortunately, there is no experimental data for this mixture, but the simulations suggest that this would be an interesting solution mixture for study and that ideal behavior is not limited to trivial solutions with a small number of components.

Finally, in Fig. 6 and Table I we present a comparison of the simulated and experimental ΔG_{ij} for all the mixtures considered here. The simulated, and presumably experimental, KBIs displayed in Fig. 6 clearly distinguish between ideal and nonideal behavior. While the simulated values of ΔG_{ij}

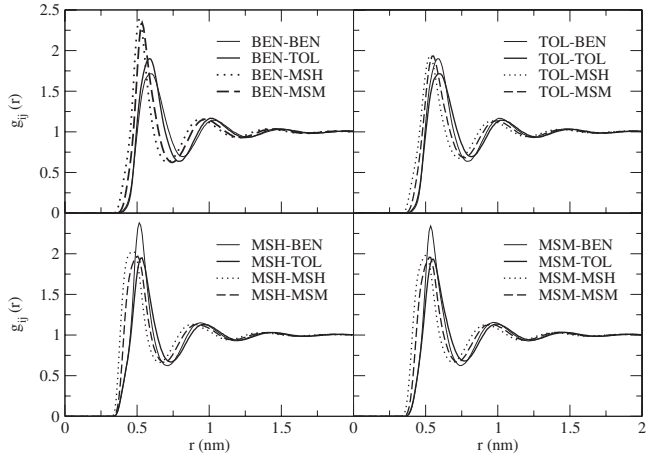

FIG. 5. Center of mass based rdfs obtained from a simulation of a quaternary mixture ($x_i=0.25$) of benzene, toluene, methanethiol, and dimethylsulfide at 300 K.

TABLE I. Simulated and experimental values of ΔG_{ij} for equimolar binary and quaternary solutions. (Estimated errors are displayed in parentheses.)

	ΔG_{ij} (cm^3/mol)	
System	Simulated	Experimental
$n=2$		
BEN-TOL	−25(3)	−1
MSH-MSM	10(7)	−10
BEN-MOH	1584(159)	1570
$n=4$		
BEN-TOL	−90(6)	
BEN-MSH	−5(8)	
BEN-MSM	−52(5)	
TOL-MSH	−14(13)	
TOL-MSM	−68(3)	
MSH-MSM	−43(10)	

for the different component pairs can differ from the SI value of zero, these differences are negligible in comparison with the changes observed for nonideal solutions—assuming one has a well parametrized force field.[5,30,36] It is possible that the KBIs or ΔG_{ij} for solution mixtures may be used to help define a variety of solution behavior. However, this requires a detailed analysis of a large number of solution mixtures and is beyond the scope of the current study.

V. DISCUSSION

In the previous sections we established expressions for the KBIs, in terms of the isothermal compressibility and molar volumes of the pure components, which correspond to ideal solutions using three different concentration scales. The different ideal solutions all result in the same form for the KBIs,

$$G_{ij}^{\text{ideal}} = RT\kappa_T - V_i - V_j + S_n, \tag{29}$$

except that when using the molality scale, one takes $V_1=0$ for all G_{ij} and includes a $-1/\rho_1$ term for the G_{11} integral, while for the molarity scale the same molar volume is used

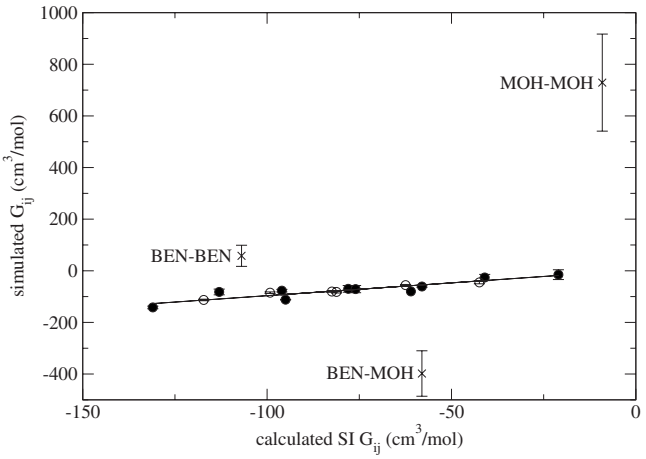

FIG. 6. A comparison between the simulated KBIs and the SI KBIs [Eq. (2)] for the two binary ideal solutions (open circles), one nonideal binary solution (crosses and labels), and one quaternary solution (filled circles) used in this study. All data refer to equimolar mixtures.

($V_i=V_0$ and $S_n=V_0$) for all species. All the expressions reduce to the compressibility equation as $\varphi_1 \to 1$. The use of $V_1=0$ for all KBIs, and the addition of the $-1/\rho_1$ term for G_{11}, in Im solutions leads to a description of the solvent as a continuum characterized by a number density, but no structural or volume information.

Only the S_n term in the KBI expressions for SI and ideal molal scale (Im) solutions depends strongly on composition (κ_T is also composition dependent but only slightly so when away from critical points). This dependence is the same for all G_{ij} pairs. The S_n term plays no role in the expressions for either the partial molar volumes or chemical potential derivatives as these involve differences in pairs of KBIs.[2,15,23] It is only necessary to obtain the correct expression for the compressibility, as illustrated by Eq. (8). We note that one could also write $S_n=\langle V^2\rangle/\langle V\rangle$, where the angular brackets denote an average (using mole fractions) over the different components in the solution mixture, and are not to be considered an ensemble average. Alternatively, $S_n=\langle V\rangle$, if one uses volume fractions in the averaging.

There are several implications that arise from the KBI expressions derived above. The Ben-Naim result,[22]

$$\Delta G_{ij}^{\text{ideal}} = G_{ij} + G_{jj} - 2G_{ij} = 0, \tag{30}$$

can be shown to be valid for any number of components and any ideal solution as long as $i,j \neq 1$ in the Im case. The difference between any two KBIs in ideal solutions is given by

$$G_{ij}^{\text{ideal}} - G_{kl}^{\text{ideal}} = V_k + V_l - V_i - V_j = \text{constant} \quad i,j \neq 1 \quad \text{for Im} \tag{31}$$

and is zero for ideal molar scale (Ic) solutions. Finally, the difference between the KBIs for SI and Im solutions is simply related to properties of the primary solvent,

$$G_{ij}^{\text{SI}} - G_{ij}^{\text{Im}} = \rho_1 V_1^2 \quad i,j \neq 1, \tag{32}$$

for any number of components. Another situation of interest arises when all the components except for the primary solvent (1) are present at low concentrations, i.e., when they would be expected to display some ideal behavior. Here, one finds that as $\varphi_1 \to 1$,

$$G_{ij}^{\text{ideal}} = RT\kappa_T - V_i - V_j + V_1 \tag{33}$$

for all i and j (except for $i=j=1$ in the Im case). This may be particularly relevant for biological systems where the concentrations of the peptides and/or proteins of interest are usually small. However, it should be noted that ideal behavior for a component appearing at very low concentrations does not imply that the KBIs adopt the corresponding ideal values. For example, salts display large positive values for G_{22} even at very low salt concentrations.

Experimental and theoretical data on solutions containing a large number of components are rare. This is especially true for biological systems. In these situations it is tempting to assume that many of the components behave ideally and that one can then use the expressions provided by Eq. (2) for the ideal components. It is immediately apparent, however, on examining Eqs. (9) and (10) that even if only a couple of the components exhibit nonideal behavior, this affects the KBIs between all of the ideal components as the solutions to Eq. (10) involve the determinant of the matrix in Eq. (9). Hence, this type of assumption should be used with care.

The properties of SI solutions and their relationships to the underlying rdfs and originating potential functions have been discussed in detail by Ben-Naim.[2] It should be noted that all the KBI expressions corresponding to the three ideal solution types represent thermodynamic models of solution behavior. This is clear from Eqs. (11), (A1), and (B1), which are thermodynamic statements of ideality, and are not clearly related to any physical models of the solution. It is tempting to present simple pictures for the underlying potential or rdfs which might lead to ideal behavior, but this may not be possible, especially for Im solutions where V_1 is zero. This is in contrast, for example, to the physical model of solution mixtures provided by scaled particle theory (SPT),[40,41] or simple hard sphere potentials. On the other hand, a more physical model is unlikely to obey all the thermodynamic relationships required by Eqs. (4), (5), and (8) for real solutions.

In Figs. 1–3 we provided data for two systems which approach SI behavior. SI behavior over the whole composition range is rare. In particular, it is unlikely that any salt will display ideal behavior since G_{22} approaches infinity at low concentrations due to local electroneutrality constraints between the ions.[42] Examples of systems exhibiting Im and Ic behavior appear unlikely, but ideality in urea and water mixtures on the molarity scale is observed up to 5M–6M urea ($m_2 \approx 6.5m$ or $x_2 \approx 0.10$).[16,43] Here, the activity derivative presented in Eq. (C6) deviates from unity by less than 4% over this range. However, the molar volumes of urea and water are 46 and 18 cm^3/mol, respectively, certainly not identical. Ideal behavior is observed because $G_{22}-G_{12}$ varies from 50 cm^3/mol at very low concentrations to -1 cm^3/mol at 5M. Consequently, when the value of $G_{22}-G_{12}$ is non-negligible the value of ρ_2 is small, whereas when the urea concentration is much larger the above difference is small. Furthermore, computer simulations of urea and water mixtures indicate that the g_{22} and g_{12} rdfs are quite different and suggest that neither the size of the molecules, nor the effective interactions between them, could be considered similar.[32] Hence, in our opinion, this is an interesting system but not a clear example of an Ic mixture.

Finally, we note that there has been some controversy concerning the interpretation of the KBIs in real and ideal solutions.[44,45] This has primarily revolved around the exact meaning of N_{ij} and the comparison of real and SI KBIs. The results presented here have no bearing on the above issue, they merely provide the appropriate expressions which can be used in either case.

VI. CONCLUSIONS

Expressions for the KBIs that lead to ideal solution behavior using a variety of concentration scales have been presented. The relationships are valid for any number of components covering the full range of composition. The expressions for the KBIs are very similar for the different concentration scales. These expressions correspond to ther-

modynamic models for solutions which can be used as a reference for understanding real solution behavior. In addition, they provide a basis for studying complicated multicomponent systems for which experimental data might not be available. Computer simulations of BEN/TOL and MSH/MSM mixtures, both displaying close to SI behavior, indicate that different patterns for the rdfs are obtained in these systems and suggest that even these solutions exhibit some variability in their molecular distributions. However, these differences are small in comparison with the behavior of nonideal solutions.

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.

APPENDIX A: IDEAL SOLUTIONS ON THE MOLALITY CONCENTRATION SCALE

The conditions for ideality depend on the concentration scale adopted. If one desires ideal behavior on the molality scale (Im) then the following relationships must hold:

$$d\mu_i^{\mathrm{Im}} = RTd \ln m_i, \tag{A1}$$

$$\mu_{ij}^{\mathrm{Im}} = \delta_{ij}, \quad i,j \neq 1, \quad \mu_{1j}^{\mathrm{Im}} = -m_j, \tag{A2}$$

where the latter is obtained from the Gibbs–Duhem equation at constant T and P.[23] Clearly, the addition of j has no affect on the molality of i unless $i=j$. The above expressions can be used in the matrix shown in Eq. (9) to obtain the expressions for the KBIs which are required for ideality on the molality scale. The only additional problem lies in the fact that solving Eq. (10) provides only the KBIs to the primary solvent 1, which occupies a unique position for the molality scale. However, choosing $i=2$ and $k \neq 1$ in Eqs. (5) and (8) provides a new set of equations which can be solved in the same way as before to generate expressions for $i,j \neq 1$. We will not present all the matrices in the interest of brevity. The final results are

$$G_{11}^{\mathrm{Im}} = RT\kappa_T + S_n' - 1/\rho_1, \quad S_n' = \sum_{k=2}^{n} \rho_k V_k^2,$$

$$G_{1j}^{\mathrm{Im}} = G_{j1}^{\mathrm{Im}} = RT\kappa_T - V_j + S_n', \quad j \neq 1,$$

$$G_{ij}^{\mathrm{Im}} = RT\kappa_T - V_i - V_j + S_n', \quad i,j \neq 1. \tag{A3}$$

The above expressions can be generalized by setting $V_1=0$ to give

$$G_{ij}^{\mathrm{Im}} = RT\kappa_T - V_i - V_j + S_n - \delta_{i1}\delta_{j1}\rho_1^{-1}, \tag{A4}$$

which is valid for any number of components.

APPENDIX B: IDEAL SOLUTIONS ON THE MOLAR CONCENTRATION SCALE

Finally, ideality on the molar concentration scale (Ic) requires the following relationships to hold:

$$d\mu_i^{\mathrm{Ic}} = RTd \ln \rho_i, \tag{B1}$$

$$\mu_{ij}^{\mathrm{Ic}} = \delta_{ij} - \phi_j. \tag{B2}$$

Here, the addition of j affects all the chemical potentials as a result of a change in solution volume. These can be used in the matrix shown in Eq. (9) to obtain expressions for the KBIs which are required for ideality on the molarity scale. The initial results are

$$G_{ij}^{\mathrm{Ic}} = V_j(RT\kappa_T - V_i)/S_n, \tag{B3}$$

which are valid for any number of components. These expressions obey Eqs. (5) and (8), but give the appearance that $G_{ij} \neq G_{ji}$, an impossible outcome according to the well known fluctuation formula for the KBIs,[1]

$$G_{ij} = V\left[\frac{\langle N_iN_j\rangle - \langle N_i\rangle\langle N_j\rangle}{\langle N_i\rangle\langle N_j\rangle} - \frac{\delta_{ij}}{\langle N_i\rangle}\right]. \tag{B4}$$

Consequently, the expression provided in Eq. (B3) can only be true if $V_i=V_j$, and therefore it is more appropriate to write Eq. (B3) as

$$G_{ij}^{\mathrm{Ic}} = V_0(RT\kappa_T - V_0)/S_n = RT\kappa_T - V_0, \tag{B5}$$

where V_0 is the same for all components. Hence, all G_{ij} have to be identical and are independent of composition for ideal behavior on the molarity scale, a much more restrictive condition than for SI or Im solutions. This extra restriction arises as the addition of any component j changes the concentration of all the species through the solution volume. The only way this change can be the same for all j is if they have the same molar volumes.

APPENDIX C: PROOF OF THE KBI EXPRESSIONS FOR IDEAL SOLUTIONS

To demonstrate that Eqs. (2), (A4), and (B5) are indeed correct, we will derive the appropriate expressions for $d\mu_i$ using the different ideal KBIs in solutions containing any number of components. This ensures that the symbolic matrix inversion was performed correctly. The easiest to prove is the Ic case. Combining Eqs. (4) and (B5), one finds

$$RTd \ln \rho_i = d\mu_i^{\mathrm{Ic}} + \sum_{j=1}^{n} \rho_j(RT\kappa_T - V_0)d\mu_j^{\mathrm{Ic}}. \tag{C1}$$

Using the Gibbs–Duhem equation at constant T and P, $\Sigma_j \rho_j d\mu_j=0$, the summation term disappears and the required relationship [Eq. (B1)] is obtained.

To prove ideality on the molality scale one can start with Eq. (4), eliminate $d\mu_1$ using the Gibbs–Duhem relationship, and then convert to molalities using $d \ln m_i = d \ln \rho_i - d \ln \rho_1$ to provide the following expression:[1,23,46]

$$RTd \ln m_i = \sum_{j=2}^{n} (\delta_{ij} + N_{ij}^{+}) d\mu_j, \tag{C2}$$

for $i=2,n$. Using the KBIs for Im solutions [Eq. (A4)] one finds that $N_{ij}^{+}=0$ for all $i,j \neq 1$. Therefore, Eq. (A1) is obeyed.

Finally, for SI solutions, we start with Eq. (4) and convert to molalities to provide the expression[23]

$$RTd \ln m_i = \sum_{j=1}^{n} (\delta_{ij} + N_{ij} - \delta_{1j} - N_{1j}) d\mu_j. \tag{C3}$$

Using the SI KBI expressions presented in Eq. (2) and the Gibbs–Duhem relation, this reduces to

$$RTd \ln m_i = d\mu_i^{\mathrm{SI}} - d\mu_1^{\mathrm{SI}}. \tag{C4}$$

Noting that changes in the mole fractions can be written as $d \ln x_i = d \ln m_i - \sum_j x_j d \ln m_j$,[47] and using Eq. (C4) leads gives

$$RTd \ln x_i = d\mu_i^{\mathrm{SI}} - d\mu_1^{\mathrm{SI}} - \sum_{j=1}^{n} x_j (d\mu_j^{\mathrm{SI}} - d\mu_1^{\mathrm{SI}}). \tag{C5}$$

After a further application of the Gibbs–Duhem expression the above equation reduces to Eq. (11) as required.

One can also demonstrate that the KBI expressions for SI, Im, and Ic solutions are correct by applying them to the established relationships for binary solutions. The solute chemical potential derivatives for a binary system of solvent (1) and solute (2) are given by[2]

$$\beta \left(\frac{\partial \mu_2}{\partial \ln x_2} \right)_{T,P} = \frac{\rho_1 + \rho_2}{\rho_1 + \rho_2 + \rho_1 \rho_2 \Delta G_{12}} = \frac{1}{1 + x_2 \rho_1 \Delta G_{12}},$$

$$\beta \left(\frac{\partial \mu_2}{\partial \ln m_2} \right)_{T,P} = \frac{1}{1 + N_{22}^{+}} = \frac{1}{1 + m_2 (1 + \rho_1 \Delta G_{12})},$$

$$\beta \left(\frac{\partial \mu_2}{\partial \ln \rho_2} \right)_{T,P,N_1} = \frac{1}{1 + \rho_2 (G_{22} - G_{12})}. \tag{C6}$$

The above derivatives all reduce to unity after inserting the appropriate expressions for the KBIs. This is also true for the corresponding expressions describing ternary and quaternary solutions.[23] Furthermore, the KBI expressions for ideal solutions provided above all lead to the correct expressions for the partial molar volumes and compressibility for solutions with any number of components. This can be shown by inserting the ideal KBI expressions into Eq. (5) for the partial molar volumes and into Eq. (8) for the compressibility.

[1] J. G. Kirkwood and F. P. Buff, J. Chem. Phys. **19**, 774 (1951).
[2] A. Ben-Naim, *Statistical Thermodynamics for Chemists and Biochemists* (Plenum, New York, 1992).
[3] E. Matteoli and G. A. Mansoori, *Fluctuation Theory of Mixtures* (Taylor & Francis, New York, 1990).
[4] E. Matteoli and L. Lepori, J. Chem. Phys. **80**, 2856 (1984).
[5] V. Pierce, M. Kang, M. Aburi, S. Weerasinghe, and P. E. Smith, Cell Biochem. Biophys. **50**, 1 (2008).
[6] I. L. Shulgin and E. Ruckenstein, J. Chem. Phys. **123**, 054909 (2005).
[7] S. Shimizu, Proc. Natl. Acad. Sci. U.S.A. **101**, 1195 (2004).
[8] J. Rosgen, B. M. Pettitt, and D. W. Bolen, Protein Sci. **16**, 733 (2007).
[9] J. P. O'Connell, Mol. Phys. **20**, 27 (1971).
[10] A. A. Chialvo, J. Phys. Chem. **97**, 2740 (1993).
[11] Y. Marcus, Monatsch. Chem. **132**, 1387 (2001).
[12] P. E. Smith and R. M. Mazo, J. Phys. Chem. B **112**, 7875 (2008).
[13] F. Chen and P. E. Smith, J. Phys. Chem. B **112**, 8975 (2008).
[14] E. Ruckenstein and I. Shulgin, Fluid Phase Equilib. **180**, 281 (2001).
[15] P. E. Smith, Biophys. J. **91**, 849 (2006).
[16] H. Kokubo, J. Rosgen, D. W. Bolen, and B. M. Pettitt, Biophys. J. **93**, 3392 (2007).
[17] A. Ben-Naim, A. M. Navarro, and J. M. Leal, Phys. Chem. Chem. Phys. **10**, 2451 (2008).
[18] A. Ben-Naim, J. Chem. Phys. **63**, 2064 (1975).
[19] M. B. Gee and P. E. Smith, J. Chem. Phys. **131**, 165101 (2009).
[20] E. Ruckenstein and I. Shulgin, Fluid Phase Equilib. **180**, 345 (2001).
[21] P. E. Smith, J. Chem. Phys. **129**, 124509 (2008).
[22] A. Ben-Naim, *Molecular Theory of Solutions* (Oxford University Press, New York, 2006).
[23] M. Kang and P. E. Smith, J. Chem. Phys. **128**, 244511 (2008).
[24] E. Lindahl, B. Hess, and D. van der Spoel, J. Mol. Model. **7**, 306 (2001).
[25] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola, and J. R. Haak, J. Chem. Phys. **81**, 3684 (1984).
[26] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, J. Comput. Chem. **18**, 1463 (1997).
[27] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).
[28] S. Weerasinghe and P. E. Smith, J. Phys. Chem. B **109**, 15080 (2005).
[29] N. Bentenitis, N. R. Cox, and P. E. Smith, J. Phys. Chem. B **113**, 12306 (2009).
[30] E. A. Ploetz, N. Bentenitis, and P. E. Smith, Fluid Phase Equilib. **290**, 43 (2010).
[31] E. A. Ploetz and P. E. Smith (unpublished).
[32] S. Weerasinghe and P. E. Smith, J. Phys. Chem. B **107**, 3891 (2003).
[33] S. Weerasinghe and P. E. Smith, J. Phys. Chem. B **118**, 10663 (2003).
[34] I. Prigogine and R. Defay, *Chemical Thermodynamics* (Longmans, London, 1954).
[35] R. Chitra and P. E. Smith, J. Chem. Phys. **115**, 5521 (2001).
[36] A. Perera and F. Sokolic, J. Chem. Phys. **121**, 11272 (2004).
[37] S. Murakami, V. T. Lam, and G. C. Benson, J. Chem. Thermodyn. **1**, 397 (1969).
[38] K. D. Kassmann and H. Knapp, Ber. Bunsenges. Phys. Chem. **90**, 452 (1986).
[39] A. W. Jackowkski, Pol. J. Chem. **54**, 1765 (1980).
[40] H. Reiss, H. L. Frisch, and J. L. Lebowitz, J. Chem. Phys. **31**, 369 (1959).
[41] R. A. Pierotti, Chem. Rev. (Washington, D.C.) **76**, 717 (1976).
[42] P. G. Kusalik and G. N. Patey, J. Chem. Phys. **86**, 5110 (1987).
[43] R. H. Stokes, Aust. J. Chem. **20**, 2087 (1967).
[44] A. Ben-Naim, J. Phys. Chem. B **112**, 5874 (2008).
[45] I. L. Shulgin and E. Ruckenstein, J. Phys. Chem. B **112**, 5876 (2008).
[46] D. G. Hall, Trans. Faraday Soc. **67**, 2516 (1971).
[47] P. E. Smith, J. Phys. Chem. B **110**, 2862 (2006).
[48] G. M. Wilson, J. Am. Chem. Soc. **86**, 127 (1964).