JCTC Journal of Chemical Theory and Computation

Article

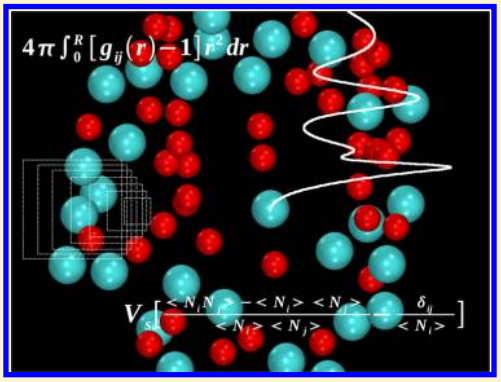
pubs.acs.org/JCTC

Convergence of Sampling Kirkwood–Buff Integrals of Aqueous Solutions with Molecular Dynamics Simulations

Pritam Ganguly and Nico F. A. van der Vegt*

Center of Smart Interfaces, Technische Universität Darmstadt, Petersenstrasse 17, 64287 Darmstadt, Germany

ABSTRACT: We discuss two methods for calculating Kirkwood–Buff integrals (KBIs) of aqueous cosolvent solutions from molecular simulations. The first method is based on computing running integrals over radial distribution functions obtained from NVT or NpT simulations. The second, more recent method, originally introduced by Schnell et al. (*J. Phys. Chem. B* **2011**, *115*, 10911), obtains the KBIs from direct analysis of particle number fluctuations in small, open subvolumes embedded in a larger reservoir as provided by the NVT (NpT) simulation cell. The thermodynamic limit is taken in the first method by using the plateau-values of the running KBIs for large distances, while in the second method an analytical finite-size scaling relation is applied to the KBIs of subvolumes of variable size. We find that direct analysis of particle number fluctuations at small scales provides more precise estimates of KBIs for methanol–water and urea–water solutions. Converged KBIs could, however, not be obtained from nanosecond time scale molecular dynamics simulations with either of the two methods. Based on 0.1 μs simulation trajectories of small and large system sizes time-converged KBIs were obtained with both methods. The running integral method suffers, however, from stronger finite-size artifacts than the sub-box method, also when empirical finite-size tail corrections are applied to the radial distribution functions.



1. INTRODUCTION

The Kirkwood–Buff (KB) theory[1] proposed in 1951 relates thermodynamic quantities of stable solution mixtures to the microscopic liquid structure. The theory is derived in the grand-canonical (μVT) ensemble with any number of components of any type and defines the so-called Kirkwood–Buff integrals (KBIs), which are the integrals of the radial distribution functions (RDFs) over volume. The KBI between mixture components i and j is defined as

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

where $g_{ij}^{\mu VT}(r)$ is the RDF between components i and j. These integrals can be related to thermodynamic quantities such as the isothermal compressibility, partial molar volumes, and derivatives of the chemical potentials or activity coefficients of solution components with solution composition.[2] Physically, KBIs provide a measure of the mutual affinity between solution components; that is, the quantity $\rho_j G_{ij}$, with ρ_j being the number density of molecules j, is the change in the average number of molecules j in a spherical region of radius R caused by placing a molecule i at the center of the region. Here, R is defined as $g_{ij}(r \geq R) = 1$ and typically corresponds to a distance between 1.0 and 2.0 nm, depending on the system under study. Hence, KB theory provides a link between local (<2.0 nm) properties (G_{ij}s) and global, thermodynamic properties. KB theory is exact, does not assume pairwise additivity of the potentials, and provides a powerful and computationally straightforward route to obtain thermodynamic properties from the RDFs obtained with molecular simulations. Using the inversion of the Kirkwood–Buff theory,[3] KBIs can, moreover, be obtained experimentally. This provides a means for parametrizing atomistic force-fields based on KB theory, early applications of which included aqueous cosolvent mixtures,[4−9] aqueous amide mixtures,[10] to name a few. Recently, KB theory has also been used to parametrize a coarse-grained urea–water force field.[11]

Although KBIs (eq 1) should in principle be obtained from calculations in open systems, they are usually obtained from closed-boundary NpT or NVT simulations.[4−12] This is justified, provided that the density fluctuations are local and the simulation-box size is significantly bigger than any of the correlation lengths which are typically <2.0 nm for aqueous solutions far away from critical points. If this condition is met, the KBI is a local quantity and the solution thermodynamics can be related to the solution structure in NpT and NVT simulations. For closed boundary systems, KBIs (G_{ij}) calculated (eq 1) by integrating over the overall box volume take the values $-1/\rho_j$ ($i = j$) or 0 ($i \neq j$). Clearly, these are not the required results. However, given the condition that the density fluctuations are local, we expect the integral to asymptotically approach the correct KBI for integration volumes smaller than the box size. The running-KBI (RKBI), defined as

Received: November 19, 2012
Published: February 18, 2013

$$G_{ij}(r) = 4\pi \int_0^r [g_{ij}^{\mathrm{NpT}}(r') - 1] r'^2 \, \mathrm{d}r' \tag{2}$$

thus approaches a plateau value that corresponds to the thermodynamic limit in eq 1 at a distance r, which is small compared to the box size. In practice, KBIs calculated this way are inaccurate due to the following reasons: (1) Poor convergence of the RDFs at longer distances as it takes long for the particles to move over large distances and, consequently, the tails of the RDFs suffer from poor statistics. This error is further weighted with r^2 while calculating the KBIs, (2) RKBIs calculated in closed systems do not show the correct asymptotic behavior because the RDFs go to a limit, which is not one.[2,13−15] The second issue has not been considered in several earlier works considering the analysis of RKBIs[4−11] where the authors obtained the KBIs by taking averages of the RKBIs in a finite range of the radial separations. Perera and Sokolić rescaled the pair correlation function in such a way that it approaches one asymptotically at half of the simulation box length[14] and, in a later work, used an alternative approach.[15] Another method of rescaling the RDFs was used by Hess and van der Vegt where the RDFs were renormalized with the actual particle densities in the bulk at large distance.[16] Christensen et al.[17] and Wedberg et al.[18] used parametric functions to fit RDFs and further extrapolated them to long-range.

Schnell et al. proposed an alternative method to calculate the thermodynamic limiting values of the KBIs from simulations of small boxes.[19] In their earlier work, they derived[20] expressions for thermodynamic quantities such as the thermodynamic correction factor (which appears in Fick diffusion coefficients) and molar enthalpy of small μVT systems using the formalism proposed by Hill.[21] Based on this formalism, they showed that the inverse of the thermodynamic correction factor and molar enthalpy of small μVT systems scale linearly with the inverse of the linear dimension of the system. In their later work,[19] they calculated KBIs of small nonperiodic systems embedded in a larger periodic system and showed that the KBIs also scale linearly with the inverse of the linear dimension of the small systems as

$$G_{ij}(L_s) = G_{ij} + \frac{A'}{L_s} \tag{3}$$

where $G_{ij}(L_s)$ is the KBI of the small embedded system (from this point on referred to as SKBI), G_{ij} is the SKBI in the thermodynamic limit for $L_s \to \infty$, and L_s is the linear size of the cubic small system. These small cubic boxes (referred to as sub-boxes from now on) can be thought to represent grand-canonical systems where the bigger simulation box acts as a large particle bath or reservoir with which the smaller sub-boxes can exchange energy and particles. This picture can be maintained provided that the size (L_s) of the sub-box is sufficiently smaller than the larger (NVT or NpT) simulation box with linear size L. The large periodic simulation boxes and small nonperiodic sub-boxes are shown schematically in Figure 1. The SKBIs were calculated from the particle number fluctuation as[19]

$$G_{ij}(L_s) = V_s \left[\frac{\langle N_i N_j \rangle - \langle N_i \rangle \langle N_j \rangle}{\langle N_i \rangle \langle N_j \rangle} - \frac{\delta_{ij}}{\langle N_i \rangle} \right] \tag{4}$$

where N_i is the number of particles of type i within the sub-box, V_s is the volume of the sub-box, and δ_{ij} is Kronecker delta. $\langle .. \rangle$ denotes the grand-canonical ensemble average. Calculations of $G_{ij}(L_s)$ for various sub-box sizes L_s allow to obtain the thermodynamic limiting value G_{ij} by means of eq 3.

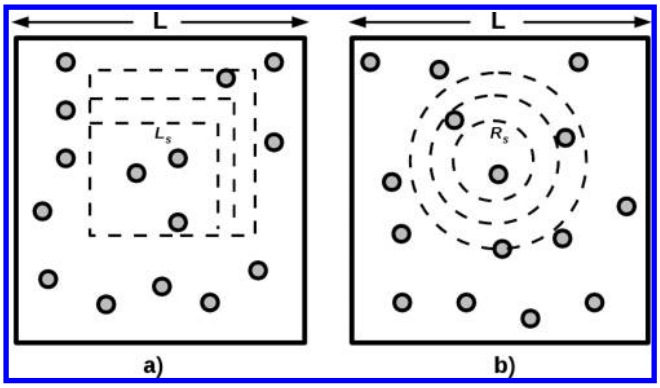


Figure 1. Cubic (a) and spherical (b) subsystems (dashed) embedded in a periodic simulation box of length L. L_s and R_s are the variable linear dimension and radius of the cubic and spherical subsystems, respectively. These open subsystems are nonperiodic, the periodic simulation box with the larger dimension L provides a particle-bath for the small subsystems.

In the present study we simulate urea−water and methanol−water solutions with different system sizes using Kirkwood−Buff-derived atomistic force fields.[5,7] We calculate KBIs using Schnell's sub-box method[19] with cubic and spherical sub-boxes and compare the results with KBIs obtained from the RKBIs (eq 2) in the limit of large r. As the force fields were originally developed using finite-sized systems, the analysis of finite-size effects provides additional information on the accuracy of the models. We show that KBIs obtained by taking the average values of the, usually oscillating, RKBIs in an arbitrary distance range contain relatively large fluctuations, since the KBIs depend on the range of distances used for averaging as well as on system size. We also study a finite-size correction to the RDFs and its effect on the KBIs calculated with different system sizes. Convergence issues of RKBIs and SKBIs, obtained from finite-time molecular dynamics simulations, are furthermore examined.

2. SIMULATION DETAILS

All-atomistic simulations were performed with the GROMACS molecular dynamics package (version 4.0).[22] The nonbonded parameters for methanol and urea were taken from Kirkwood−Buff-theory-derived force-fields[5,7] as mentioned earlier. The SPC/E water model was used for both the systems.[23] NpT simulations were carried out at a temperature of 300 K and a pressure of 1 bar. The temperature and pressure were kept constant using a Nose−Hoover thermostat[24,25] and a Parrinello−Rahman barostat,[26] respectively. Electrostatic interactions were evaluated using the particle mesh Ewald (PME) method.[27] The cutoff of nonbonded interactions was taken to be 1 nm.

3. RESULTS AND DISCUSSION

3.1. Convergence of KBIs with Simulation Time. Urea−water and methanol−water mixtures were simulated using large systems (approximately 10 000 to 12 000 molecules in total) as well as small systems (approximately 2000 to 2500 molecules in total). 100-ns-long trajectories were accumulated with an integration time-step of 2 fs. The RKBIs and SKBIs were obtained from 100 ns simulations. In particular, the long-range

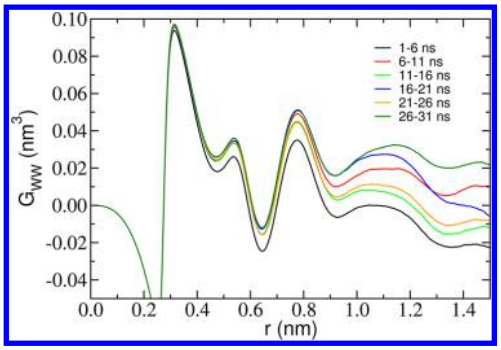


Figure 2. Water–water RKBIs for a methanol–water mixture (70 mol % methanol) calculated from simulations of a NpT system with 2000 molecules. RKBIs are presented for different sampling time intervals of 5 ns taken from a 100 ns trajectory.

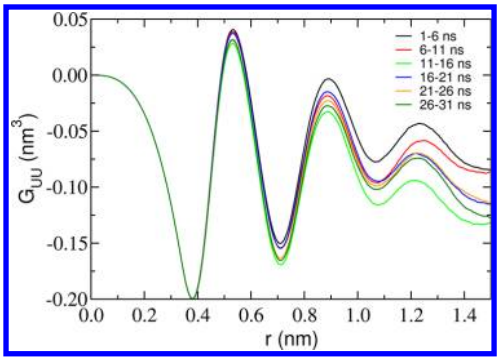


Figure 3. Urea–urea RKBIs for a urea–water mixture (6 m urea) calculated from simulations of a NpT system with 2000 water molecules. RKBIs are presented for different sampling time intervals of 5 ns taken from a 100 ns trajectory.

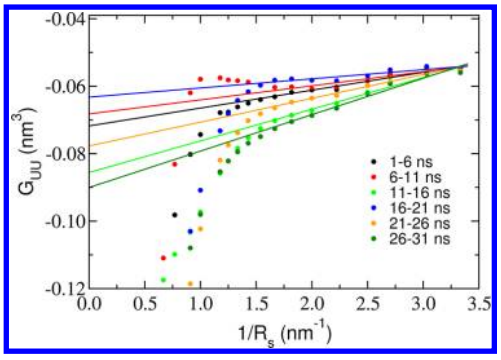


Figure 4. Urea–urea SKBIs obtained with the sub-box method for a urea–water mixture (6 m urea) calculated from simulations of a NpT system with 2000 water molecules. The particle number fluctuations have been calculated (eq 4) with different sampling time intervals of 5 ns taken from a 100 ns trajectory. The solid lines are fitted to the linear parts of the data. The linear extrapolations ($R_s \rightarrow \infty$) provide the thermodynamic limiting KBIs.

part of the RKBI equilibrates very slowly, as shown in Figure 2 and Figure 3 for the methanol–water and urea–water systems, respectively. Clearly, the limiting values of the RKBIs at distances larger than 1.0 nm cannot be obtained with reasonable accuracy from short (i.e., 5 ns) production runs. We note that equally long (100 ns) simulations for the same urea–water system (as in Figure 3) are required to obtain converged SKBIs as shown by the data presented in Figure 4.

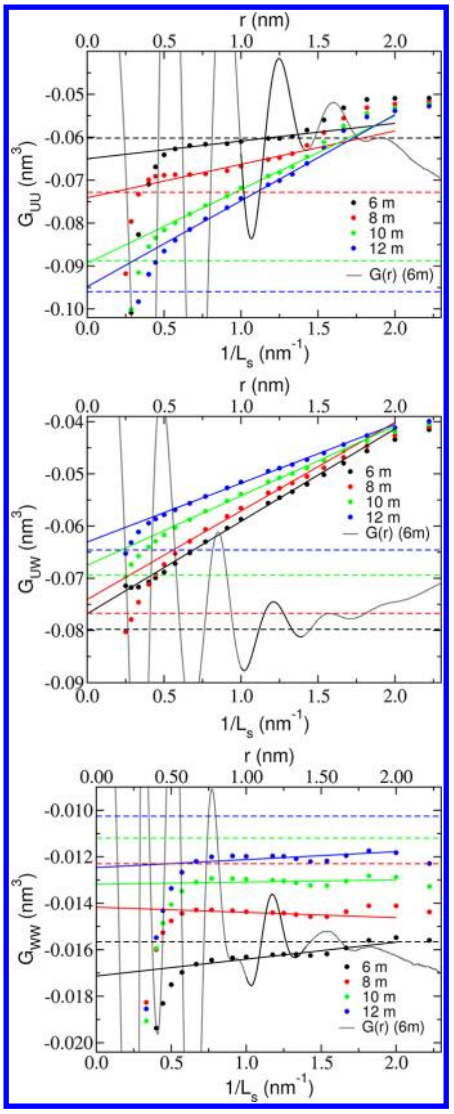


Figure 5. Urea–urea (upper panel), urea–water (middle panel), and water–water (lower panel) SKBIs calculated from the particle number fluctuations of cubic sub-boxes presented as functions of the inverse of the sub-box lengths for urea–water mixtures (big-box, 11111 water molecules) with different urea concentrations (molality). Solid straight lines are fitted to the linear regime of the plots. Horizontal dashed lines denote the limiting KBIs obtained from direct integration of the corresponding RDFs. Gray lines are the RKBIs at 6 m urea concentration. The averaging interval used to obtain the limiting KBIs is darkened on the curve.

The strong run length dependence shown in the data probably reflects the characteristic microheterogeneous nature of aqueous solutions[28−31] with a correspondingly slow domain like dynamics. Urea self-aggregation and correspondingly slow equilibration of concentrated aqueous urea solutions has previously been reported.[32,33] Neutron diffraction experiments and molecular simulations of mixtures of methanol and water showed that these systems exhibit extended structures, which, within a certain concentration region, appear to form separate, percolating networks.[34] In contrast to aqueous solutions of nonelectrolytes, oscillations of RKBIs of aqueous electrolyte solutions (≈ 1 m salt concentration) typically cease when r approaches 1 nm.[12,35]

3.2. Urea–Water KBIs from Large System Sizes. For urea–water large cubic boxes (7.5–8.0 nm) with periodic

dx.doi.org/10.1021/ct301017q | J. Chem. Theory Comput. 2013, 9, 1347−1355

Table 1. Limiting KBIs for Urea−Water Mixtures (in nm^3) Obtained from Direct Integration of RDFs (eq 2) and from Particle Number Fluctuations (eq 4) Using Cubic Sub-boxes and Spherical Sub-boxes by Extrapolating to the Thermodynamic Limit (eq 3)[a]

		integration of RDF						cubic sub-box	spherical sub-box	
		bigbox			smallbox					
urea molality	G_{ij}	0.8−1.2	1.0−1.4	1.1−1.5	0.8−1.2	1.0−1.4	1.1−1.5	bigbox	bigbox	smallbox
6	G_{UU}	−0.052	−0.060	−0.056	−0.064	−0.082	−0.085	−0.065	−0.064	−0.082
	G_{UW}	−0.076	−0.080	−0.078	−0.073	−0.074	−0.070	−0.077	−0.078	−0.073
	G_{WW}	−0.016	−0.016	−0.015	−0.016	−0.017	−0.017	−0.017	−0.017	−0.017
8	G_{UU}	−0.067	−0.073	−0.068	−0.081	−0.096	−0.098	−0.074	−0.075	−0.092
	G_{UW}	−0.072	−0.077	−0.075	−0.067	−0.068	−0.065	−0.074	−0.075	−0.069
	G_{WW}	−0.013	−0.013	−0.012	−0.014	−0.015	−0.015	−0.014	−0.014	−0.015
10	G_{UU}	−0.080	−0.089	−0.085	−0.088	−0.104	−0.106	−0.089	−0.090	−0.096
	G_{UW}	−0.072	−0.069	−0.075	−0.062	−0.063	−0.059	−0.068	−0.068	−0.065
	G_{WW}	−0.011	−0.011	−0.011	−0.013	−0.014	−0.014	−0.013	−0.013	−0.013
12	G_{UU}	−0.087	−0.096	−0.093	−0.089	−0.102	−0.100	−0.095	−0.096	−0.099
	G_{UW}	−0.061	−0.065	−0.062	−0.059	−0.062	−0.059	−0.063	−0.063	−0.062
	G_{WW}	−0.010	−0.010	−0.010	−0.011	−0.012	−0.012	−0.012	−0.012	−0.011

[a]Data are shown for big and small system sizes. For direct integration of RDFs three different regions for averaging were chosen, namely from 0.8 to 1.2 nm (0.8−1.2), from 1.0 to 1.4 nm (1.0−1.4), and from 1.1 to 1.5 nm (1.1−1.5).

boundary conditions were simulated. The urea concentration was varied from 6 m (4.7 M) to 12 m (7.7 M). The urea−urea, urea−water, and water−water KBIs were calculated using two methods: (1) from the direct integration of the respective RDFs (eq 2) and (2) from the particle number fluctuations (eq 4). In order to get the KBIs from particle number fluctuations, we inserted nonperiodic sub-boxes into the simulation box and counted the number of the urea and water molecules within the sub-boxes. Statistical averages were taken by inserting 5000 sub-boxes per frame at random positions within the simulation box along with a further averaging done using 15000 frames which span 90-ns-long trajectories. The shape of the embedded sub-boxes was chosen to be cubic, and the SKBIs were calculated for sub-boxes with varying linear size L_s. Figure 5 shows the SKBIs between urea−urea (G_{UU}), urea−water (G_{UW}), and water−water (G_{WW}) as functions of the inverse of the linear dimension (L_S) of the cubic sub-boxes for different urea concentrations. It is interesting to see that we find a regime corresponding to the sub-box length-scale of approximately 0.7 to 1.5 nm where the SKBIs vary linearly with the inverse of the sub-box lengths. This regime is used to obtain the limiting KBIs by linear extrapolation into the thermodynamic limit $1/L_S \to 0$ according to eq 3. The limiting KBIs derived by this procedure are reported in Table 1. The limiting KBIs obtained from the direct integration of the RDFs between urea−urea, urea−water, and water−water are also shown in Figure 5 (represented by dashed horizontal lines). These values were calculated by averaging the RKBIs (eq 2) in a distance range between r = 1.0 and 1.4 nm. We can see that the KBIs obtained with the two methods are in reasonable agreement although the values obtained from the RKBIs may fluctuate 10−15% if the spatial region used for averaging the RKBIs changes. Table 1 shows the variation of the KBIs obtained from RKBIs by carrying out the averaging in three different regions, namely from 0.8 to 1.2 nm, from 1.0 to 1.4 nm, and from 1.1 to 1.5 nm. For illustration, the gray curves in Figure 5 show the RKBIs at 6 m urea-concentration. We see that the RKBIs are still oscillating even after 1.0−1.5 nm, which leads to smaller precision (error 10−15%) of the KBIs obtained with the integration method compared to the limiting KBIs ($L_S \to \infty$) obtained from SKBI calculations. The observed, drifting asymptotes of the RKBIs shown in Figure 5 are due to finite system size effects and can be alleviated by applying corrections to the tails of the RDFs as discussed in the introduction. In particular, for smaller systems (2000 molecules, see below), tail corrections are required since direct integration of the uncorrected RDFs yields largely oscillating and drifting RKBIs. This aspect is discussed in greater detail later.

The SKBIs start to deviate from linearity with $1/L_S$ when the length of the sub-boxes is larger than ≈1.5 nm. This happens because, for the bigger sub-boxes, the simulation box (L = 7.5 − 8.0 nm) no longer acts as a sufficiently large particle reservoir; that is, the sub-boxes can no longer be considered as grand-canonical subsystems. This deviation was also seen in previous works.[19,20] For smaller sub-boxes ($L_S < \approx 0.7$ nm) where the size of the boxes can be compared to the molecular dimension, the KBIs cannot be calculated correctly from the particle number fluctuations because the particle-count within the sub-boxes would depend on the packing of the molecules at very small length scales and on the shape of the sub-boxes itself. We further examined the KBIs obtained using spherical sub-boxes with same averaging criteria as for cubic sub-boxes. The SKBIs calculated from particle number fluctuation within the spherical sub-boxes are shown in Figure 6 as functions of the inverse of the radii (R_s) of the sub-boxes. Again, we observe a linear regime in the plots between $R_s \approx 0.3$ to 1.4 nm which can readily be extrapolated ($1/R_S \to 0$) to obtain the KBIs in the thermodynamic limit. These limiting values are also listed in Table 1. For both shapes of the sub-boxes (i.e., cubical and spherical), we find a comparatively smaller linear regime for urea−urea SKBIs when plotted with the inverse of the linear dimension of the sub-boxes ($L_S \approx 0.8$ to 1.3 nm and $R_S \approx 0.45$ to 1.2 nm) as they suffer from poorer statistics than urea−water and water−water SKBIs. If we compare the SKBIs obtained from cubic and spherical sub-boxes, we see that the extension of the linear regime is larger with spherical sub-boxes than with cubic sub-boxes. Also, for urea−water and water−water SKBIs, spherical sub-boxes do a slightly better job than cubic sub-boxes at the smaller length scales, which can be expected from the spherical symmetry of the hydration shells around urea or water

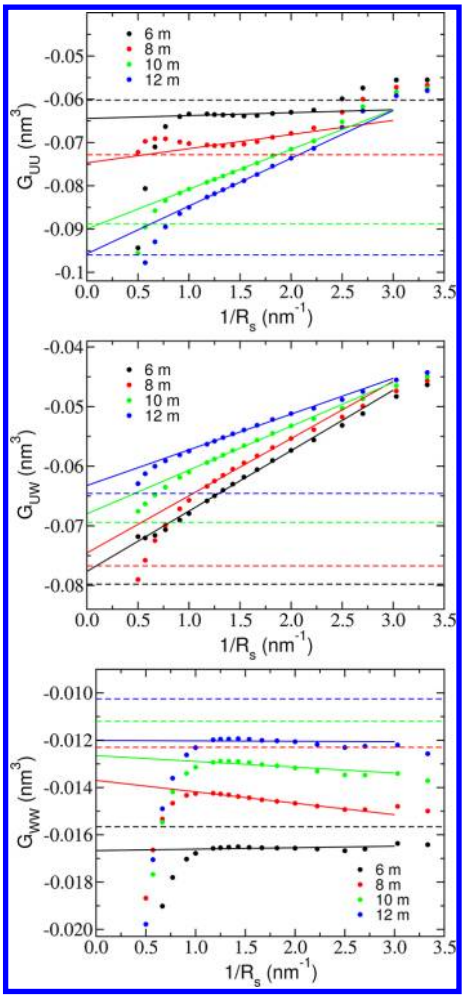


Figure 6. Urea–urea (upper panel), urea–water (middle panel), and water–water (lower panel) SKBIs calculated from the particle number fluctuations of spherical sub-boxes as functions of the inverse of the sub-box radii for urea–water mixtures (big-box, 11111 water molecules) with different urea concentrations (molality). Solid straight lines are fitted to the linear regime of the plots. Horizontal dashed lines denote the limiting KBIs obtained from direct integration of the corresponding RDFs.

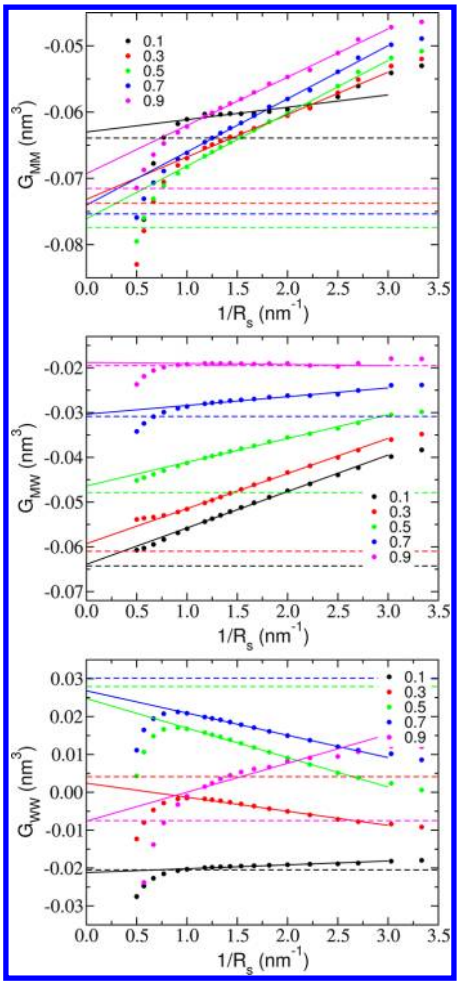


Figure 7. Methanol–methanol (upper panel), methanol–water (middle panel), and water–water (lower panel) SKBIs calculated from the particle number fluctuations of spherical sub-boxes presented as functions of the inverse of the sub-box radii for methanol–water mixtures (big-box, 10 000 molecules) with different methanol concentrations (mole-fraction). Solid straight lines are fitted to the linear regimes of the plots. Horizontal dashed lines denote the limiting KBIs obtained from direct integration of the corresponding RDFs.

molecules, so spherical sub-boxes provide better packing of the particles.

3.3. Methanol–Water KBIs from Large System Sizes. A series of methanol–water systems were also studied with methanol mole fractions ranging between 0.1 to 0.9. Each system consisted of 10 000 molecules (methanol+water) with cubic simulation boxes of length 7–9 nm having periodic boundaries. KBIs between methanol–methanol (G_{MM}), methanol–water (G_{MW}), and water–water (G_{WW}) were calculated from the particle number fluctuations within the sub-boxes of varying linear dimension embedded in the simulation box. Figure 7 shows the resulting SKBIs presented versus the inverse spherical sub-box radius. The SKBIs again follow eq 3 for a regime $R_s \approx 0.3$ to 1.5 nm, and their values in the thermodynamic limit are reported in Table 2. The limiting KBIs calculated using cubic sub-boxes are also listed in Table 2. For cubic sub-boxes, we found the linear regime to be $L_s \approx 0.5$ to 1.75 nm (data not shown). The SKBIs that suffer from poorer statistics, that is, methanol–methanol at lower methanol concentrations (10%, 30%) and water–water at higher methanol concentrations (70%, 90%), show a comparatively narrower linear regime ($R_s \approx 0.45$ to 1.2 nm for spherical sub-boxes and $L_s \approx 0.6$ to 1.3 nm for cubic sub-boxes). KBIs obtained from the direct integration of the corresponding RDFs are also presented in Table 2 (with three averaging regions, from 0.8 to 1.2 nm, from 1.0 to 1.4 nm, and from 1.1 to 1.5 nm) and in Figure 7 (see horizontal dashed lines) where the average values were taken between 1.0 and 1.4 nm.

3.4. KBIs with Varying Simulation Box Size. *3.4.1. Effect of RDF Tail Corrections on RKBIs of Small and Large Methanol–Water Systems.* KBIs evaluated based on simulations of finite-sized boxes suffer from finite-size effects as discussed in the introduction. To illustrate this problem we simulated methanol–water mixtures at 50% methanol concentration with varying box sizes, namely with 10 000, 5000, and 2000 molecules in total with simulation box-lengths of 7.9, 6.3, and 4.6 nm, respectively. The RKBIs between methanol–methanol, methanol–water, and water–water are shown in Figure 8 as a function of the radial distance (see the solid curves). It is clear that we hardly find a distance range with a well-defined plateau needed to estimate the KBIs in the

dx.doi.org/10.1021/ct301017q | J. Chem. Theory Comput. 2013, 9, 1347–1355

Table 2. Limiting KBIs for Methanol−Water Mixtures (in nm^3) Obtained from Direct Integration of RDFs (eq 2) and from Particle Number Fluctuations (eq 4) Using Cubic Sub-boxes and Spherical Sub-boxes by Extrapolating to the Thermodynamic Limit (eq 3)[a]

		integration of RDF						cubic sub-box	spherical sub-box	
		bigbox			smallbox					
methanol %	G_{ij}	0.8−1.2	1.0−1.4	1.1−1.5	0.8−1.2	1.0−1.4	1.1−1.5	bigbox	bigbox	smallbox
10	G_{MM}	−0.057	−0.064	−0.062	−0.108	−0.100	−0.078	−0.066	−0.063	−0.078
	G_{MW}	−0.063	−0.064	−0.063	−0.051	−0.055	−0.057	−0.062	−0.064	−0.060
	G_{WW}	−0.020	−0.020	−0.020	−0.023	−0.022	−0.021	−0.021	−0.021	−0.021
30	G_{MM}	−0.068	−0.074	−0.072	−0.076	−0.076	−0.069	−0.073	−0.073	−0.074
	G_{MW}	−0.059	−0.061	−0.059	−0.054	−0.057	−0.057	−0.058	−0.059	−0.057
	G_{WW}	0.005	0.004	0.004	0.000	0.001	0.003	0.001	0.002	0.001
50	G_{MM}	−0.072	−0.077	−0.076	−0.078	−0.079	−0.073	−0.075	−0.076	−0.077
	G_{MW}	−0.045	−0.048	−0.046	−0.043	−0.045	−0.043	−0.045	−0.046	−0.045
	G_{WW}	0.029	0.028	0.028	0.019	0.021	0.024	0.022	0.025	0.023
70	G_{MM}	−0.070	−0.075	−0.074	−0.074	−0.076	−0.070	−0.073	−0.074	−0.073
	G_{MW}	−0.027	−0.031	−0.029	−0.026	−0.028	−0.025	−0.031	−0.030	−0.029
	G_{WW}	0.030	0.030	0.030	0.007	0.012	0.019	0.024	0.027	0.021
90	G_{MM}	−0.066	−0.072	−0.070	−0.069	−0.071	−0.065	−0.069	−0.069	−0.068
	G_{MW}	−0.014	−0.019	−0.017	−0.016	−0.019	−0.014	−0.020	−0.019	−0.020
	G_{WW}	−0.003	−0.007	−0.010	−0.016	−0.010	−0.000	−0.002	−0.008	−0.002

[a]Data are shown for big and small system sizes. For direct integration of RDFs three different regions for averaging were chosen, namely from 0.8 to 1.2 nm (0.8−1.2), from 1.0 to 1.4 nm (1.0−1.4), and from 1.1 to 1.5 nm (1.1−1.5).

thermodynamic limit. For the smaller systems the RKBIs deviate significantly from those of the bigger systems after ≈1.0 nm, particularly the methanol−water and water−water RKBIs.

Given this problem, we wanted to test how the KBIs obtained from particle number fluctuations within the sub-boxes depend on the overall simulation box-size. The respective SKBIs calculated with spherical sub-boxes are also shown in Figure 8 presented versus the inverse of the sub-box radius (see solid dots) and the fits of the linear regime are shown as dashed straight lines. The SKBIs and extrapolated KBIs do not vary much with system size (see green, red, and black dashed straight lines) while the RKBIs do not converge to any plateau, rendering the integration approach more erroneous in predicting the limiting KBIs. Hence, thermodynamic quantities obtained by analyzing particle number fluctuations in small subvolumes exhibit reasonable accuracy, even if small overall (NVT or NpT) systems are simulated where direct integration of RDFs seems not to obtain a steady limiting KBI.

The asymptotic behavior of the RDF in a finite-sized system goes to a limit, which is not one.[15] If not corrected, this error leads to RKBIs with drifting asymptotes. Empirical corrections can, however, be made, and here, we follow the idea that the observed drift happens because of the fact that excess (depletion) of particle type j around particle type i at a local-scale is compensated by depletion (excess) of particle type j at long distances as the total number of particles of type j is fixed. We, therefore, made a correction to the RDFs, which accounts for the correct bulk density of particle type j at a distance r from particle type i depending on the excess or depletion of particle type j within a sphere of radius r around particle type i. The corrected RDFs are given by

$$g_{ij}^{\text{correct}}(r) = g_{ij}(r)\frac{N_j\left(1 - \frac{(4/3)\pi r^3}{V}\right)}{N_j\left(1 - \frac{(4/3)\pi r^3}{V}\right) - \Delta N_{ij}(r) - \delta_{ij}} \quad (5)$$

where N_j is the number of particles of type j in the system, V is the volume of the system, and $\Delta N_{ij}(r)$ is the excess number of particles of type j within a sphere of radius r around particle type i. The RKBIs calculated from the corrected RDFs do reasonably well in terms of finding a plateau, and the corresponding results for a 50% methanol mixture are shown in Figure 8 (see dashed black (10 000 molecules), red (5000 molecules), and green (2000 molecules) curves). The differences in the KBIs obtained from big and small simulation boxes using the RKBI method are significantly reduced after applying a finite-size correction but are still larger than the differences in KBIs between box sizes obtained from the sub-box method. The data in Figure 8 show that the limiting behavior of the methanol−water and water−water RKBIs remains system size depend despite the tail correction. The water−water RKBIs for large r are shifted up (more positive) relative to the limiting SKBI ($L_s \rightarrow \infty$). This may reflect poor converge of the water−water RDF in this system. Since tail corrections of the RDF are empirical and may lead to different limiting RKBIs, we further tested the correction proposed by Perera and Sokolić.[14] Figure 9 shows the water−water RKBIs (for the system with 50% methanol) corrected using the method of Perera and Sokolić (dotted lines) and eq 5 (dashed lines). Clearly, the two correction methods produce different RKBIs at distances larger than 1 nm.

The entire concentration range of methanol−water solutions has been studied using small boxes with 2000 molecules. Limiting KBIs were calculated using the sub-box method (with spherical sub-boxes) and the running-integral method. The resulting KBIs are summarized in Table 2. KBIs obtained from the RKBIs were not corrected for finite-size and were obtained by averaging the RKBI in three distance ranges, between 0.8 and 1.2 nm, between 1.0 and 1.4 nm, and between 1.1 and 1.5 nm. It should be pointed out that one obtains quite different KBIs (deviation of 10−25%) in case the averaging is performed using a different distance range of the same span which is a consequence of the oscillating nature and inaccuracies in the long-range parts of the RDFs for small system sizes. These

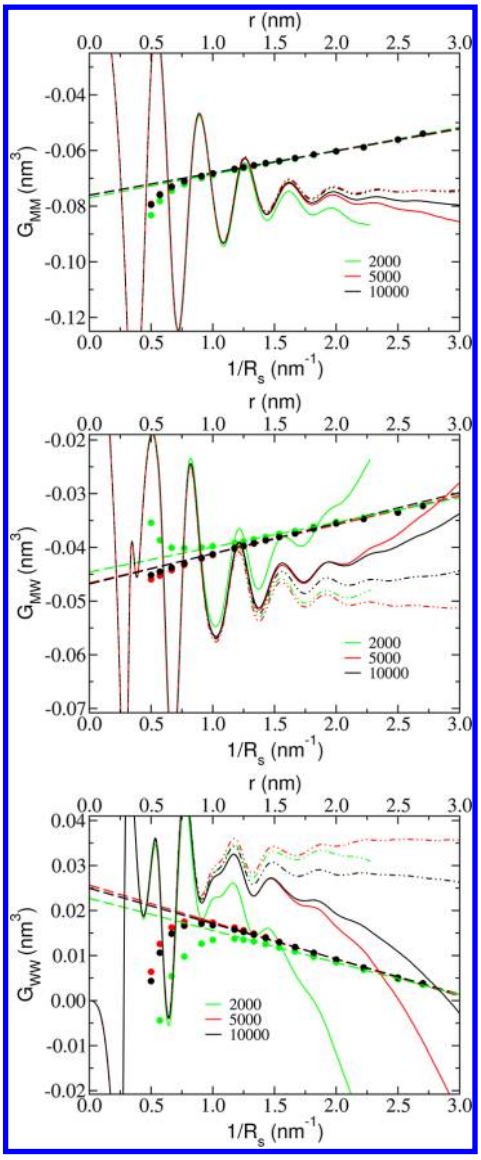


Figure 8. Methanol−methanol (upper panel), methanol−water (middle panel), and water−water (lower panel) SKBIs calculated from the particle number fluctuations of spherical sub-boxes presented as functions of the inverse of the sub-box radii for 50% methanol−water mixtures with different system size (solid dots), namely, with 2000, 5000, and 10 000 molecules. Dashed straight lines are fitted to the linear regimes of the plots. The RKBIs are presented as solid curves (without finite-size corrections to the RDFs) and dashed curves (with finite-size corrections using eq 5). The distance scale for the RKBIs is presented on the alternative, upper horizontal axis.

deviations in the KBIs give rise to large scatter in thermodynamic quantities derived from them. On the other hand, Schnell's sub-box method does not exhibit this problem as small variation of the linear-extrapolation regime within the region $R_s \approx 0.3$ to 1.3 nm (for bigger systems) and ≈0.3−0.9 nm (for smaller boxes) show very small or no deviation at all (results not shown). KBIs obtained from RKBIs using a finite-size correction (eq 5) show much smaller fluctuation (5−15%) both for bigger and smaller systems (data not shown).

3.4.2. Effect of RDF Tail Corrections on RKBIs of Small and Large Urea−Water Systems. Urea−water mixtures with smaller boxes (2000 water molecules) were also studied. Again KBIs were calculated using the sub-box method and the

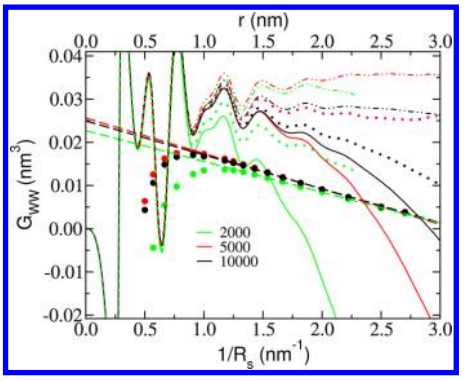


Figure 9. Tail corrected water−water RKBIs for 50% methanol−water mixtures with different system sizes (2000, 5000, and 10 000 molecules). The uncorrected RKBIs are presented as solid curves, RKBIs corrected with eq 5 as dashed curves, and RKBIs corrected with the method of ref 14 as dotted curves. The distance scale for the RKBIs is presented on the upper horizontal axis. The SKBIs (thick dots) and their linear fits are also shown.

running-integral method with different averaging regions. Data obtained with the sub-box method and the running-integral method are summarized in Table 1, values obtained from RKBIs were not corrected for finite-size. With both methods, we observe significant variations in the urea−urea KBIs for different system sizes. It is interesting to point out that the values obtained with the running-integral method using an averaging region being 1.0 to 1.4 nm are very close to the values obtained with sub-box method for the same system size. However, similar to the case of the methanol−water systems, if the region for averaging the RKBIs is changed, the KBIs deviate largely for the smaller systems without finite-size correction. Table 3 shows the KBIs obtained with tail-corrected RDFs. The differences between the resulting KBIs for different system sizes are significantly smaller in this case even compared to those obtained from the sub-box method for low urea concentrations. The fluctuations in the data obtained by varying the averaging regions of the RKBIs are 5−10% after the finite-size correction.

4. CONCLUSIONS

Kirkwood−Buff integrals (KBIs) of aqueous solution mixtures provide a link between local solution structure and global thermodynamic solution properties. In recent years, KBIs have been used in force field parametrization[4−12] as well as in simulation studies aiming to relate ion-specific thermodynamic changes and atomic scale correlations.[16,35,36] Notwithstanding the powerful scope of Kirkwood−Buff theory, sampling of the particle number fluctuations that determine the KBIs requires to address convergence issues and finite-size effects. In this paper, we have performed molecular dynamics simulations of methanol−water and urea−water mixtures with earlier-published Kirkwood−Buff-derived force fields. KBIs have been calculated by integrating the radial distribution functions over the volume (running-KBI method) as well as by direct analysis of particle number fluctuations in small sub-boxes embedded in the overall NpT simulation box (sub-box method). While in the former approach the thermodynamic limiting value of the KBI is obtained from the running integral in the limit of large distances, it is obtained in the latter approach by examining different sub-box sizes and applying an analytical finite-size scaling relation.[19] We find that with either of these two approaches converged KBIs can only be obtained

dx.doi.org/10.1021/ct301017q | J. Chem. Theory Comput. 2013, 9, 1347−1355

Table 3. Limiting KBIs for Urea−Water Mixtures (nm^3) Obtained from Direct Integration of RDFs (eq 2) after the Correction for the Finite-Size (eq 5) Applied to the Corresponding RDFs and from Particle Number Fluctuations (eq 4) Using Cubic Sub-boxes and Spherical Sub-boxes by Extrapolating to the Thermodynamic Limit (eq 3)[a]

		integration of RDF (corrected for finite-size)						cubic sub-box	spherical sub-box	
		bigbox			smallbox					
urea molality	G_{ij}	0.8−1.2	1.0−1.4	1.1−1.5	0.8−1.2	1.0−1.4	1.1−1.5	bigbox	bigbox	smallbox
6	G_{UU}	−0.049	−0.055	−0.049	−0.049	−0.055	−0.049	−0.065	−0.064	−0.082
	G_{UW}	−0.077	−0.081	−0.080	−0.071	−0.076	−0.074	−0.077	−0.078	−0.073
	G_{WW}	−0.015	−0.015	−0.015	−0.015	−0.015	−0.015	−0.017	−0.017	−0.017
8	G_{UU}	−0.065	−0.069	−0.064	−0.071	−0.079	−0.075	−0.074	−0.075	−0.092
	G_{UW}	−0.077	−0.081	−0.080	−0.071	−0.076	−0.074	−0.074	−0.075	−0.069
	G_{WW}	−0.012	−0.012	−0.011	−0.013	−0.013	−0.012	−0.014	−0.014	−0.015
10	G_{UU}	−0.079	−0.086	−0.082	−0.082	−0.093	−0.091	−0.089	−0.090	−0.096
	G_{UW}	−0.067	−0.070	−0.069	−0.066	−0.070	−0.068	−0.068	−0.068	−0.065
	G_{WW}	−0.011	−0.011	−0.010	−0.011	−0.012	−0.011	−0.013	−0.013	−0.013
12	G_{UU}	−0.086	−0.094	−0.091	−0.087	−0.094	−0.090	−0.095	−0.096	−0.099
	G_{UW}	−0.062	−0.065	−0.063	−0.063	−0.068	−0.067	−0.063	−0.063	−0.062
	G_{WW}	−0.010	−0.010	−0.009	−0.010	−0.009	−0.008	−0.012	−0.012	−0.011

[a]Data are shown for big and small system sizes. For direct integration of RDFs three different regions for averaging were chosen, namely from 0.8 to 1.2 nm (0.8−1.2), from 1.0 to 1.4 nm (1.0−1.4), and from 1.1 to 1.5 nm (1.1−1.5).

from molecular dynamics simulations on sufficient long time scales, typically of the order of 100 ns. Previous simulations of aqueous solutions have all used significantly shorter simulations (several nanoseconds) to obtain the KBIs. A possible explanation of the slow time convergence may be sought in the microheterogeneous nature of aqueous solutions, which leads to slow domain-like dynamics. We, furthermore, find that the running-KBI method suffers from stronger finite-size artifacts than the sub-box method. Simulations with different systems sizes (2000, 5000, 10000 molecules) indicate that the sub-box method provides limiting KBIs, which are in good agreement for all three system sizes. The running-KBI method, however, provides estimates of the limiting KBIs, which, depending on system size and the region for averaging the running-KBIs, shows larger variations, which are more severe for smaller systems. If, however, a finite-size correction (eq 5) for the RDFs is carried out, the running-integral method provides improved KBIs with smaller variations upon varying the region for averaging the running-KBIs. For larger systems and for longer simulation runs both the methods seem to produce very similar results with the sub-box method being more precise in most cases.

We finally point out that the running-KBI method provides a very powerful tool to analyze contributions of local correlations to the thermodynamic quantities. Examples of these include microscopic explanations for ion-specific osmotic properties of aqueous electrolyte solutions[16,35] and salting-in and salting-out mechanisms of water-soluble polymers and small peptides,[36,37] which cannot be provided by the sub-box method.

AUTHOR INFORMATION

Corresponding Author

*E-mail: vandervegt@csi.tu-darmstadt.de.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors wish to thank Debashish Mukherji for useful discussions during the course of this work. This research was supported by the German Research Foundation (DFG) within the Cluster of Excellence 259 "Smart Interfaces—Understanding and Designing Fluid Boundaries".

REFERENCES

(1) Kirkwood, J. G.; Buff, F. P. *J. Chem. Phys.* **1951**, *19*, 774.
(2) Ben-Naim, A. *Molecular Theory of Solutions*; Oxford University Press: New York, 2006.
(3) Ben-Naim, A. *J. Chem. Phys.* **1977**, *67*, 4884.
(4) Weerasinghe, S.; Smith, P. E. *J. Chem. Phys.* **2003**, *118*, 10663.
(5) Weerasinghe, S.; Smith, P. E. *J. Phys. Chem. B* **2003**, *107*, 3891.
(6) Weerasinghe, S.; Smith, P. E. *J. Chem. Phys.* **2004**, *121*, 2180.
(7) Weerasinghe, S.; Smith, P. E. *J. Phys. Chem. B* **2005**, *109*, 15080.
(8) Lee, M. E.; van der Vegt, N. F. A. *J. Chem. Phys.* **2005**, *122*, 114509.
(9) Gee, M. B.; Cox, N. R.; Jiao, Y. F.; Bentenitis, N.; Weerasinghe, S.; Smith, P. E. *J. Chem. Theory Comput.* **2011**, *7*, 1369.
(10) Kang, M.; Smith, P. E. *J. Comput. Chem.* **2006**, *27*, 1477.
(11) Ganguly, P.; Mukherji, D.; Junghans, C.; van der Vegt, N. F. A. *J. Chem. Theory Comput.* **2012**, *8*, 1802.
(12) Fyta, M.; Netz, R. R. *J. Chem. Phys.* **2012**, *136*, 124103.
(13) Lebowitz, J. L.; Percus, J. K. *Phys. Rev.* **1961**, *124*, 1673.
(14) Perera, A.; Sokolić, F. *J. Chem. Phys.* **2004**, *121*, 11272.
(15) Perera, A.; Zoranić, L.; Sokolić, F.; Mazighi, R. *J. Mol. Liq.* **2011**, *159*, 52.
(16) Hess, B.; van der Vegt, N. F. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 13296.
(17) Christensen, S.; Peters, G. H.; Hansen, F. Y.; O'Connell, J. P.; Abildskov, J. *Mol. Simul.* **2007**, *33*, 449.
(18) Wedberg, R.; Peters, G. H.; Abildskov, J. *Fluid Phase Equilib.* **2008**, *273*, 1.
(19) Schnell, S. K.; Liu, X.; Simon, J.-M.; Bardow, A.; Bedeaux, D.; Vlugt, T. J. H.; Kjelstrup, S. *J. Phys. Chem. B* **2011**, *115*, 10911.
(20) Schnell, S. K.; Vlugt, T. J. H.; Simon, J.-M.; Bedeaux, D.; Kjelstrup, S. *Chem. Phys. Lett.* **2011**, *504*, 199.
(21) Hill, T. L. *Thermodynamics of Small Systems, Part 1*; W. A. Benjamin: New York, 1963.
(22) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306.
(23) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.
(24) Nose, S. *Mol. Phys.* **1984**, *52*, 255.
(25) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.
(26) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182.

(27) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(28) Bowron, D. T.; Finney, J. L.; Soper, A. K. *J. Phys. Chem.* **1998**, *102*, 3551.
(29) Dixit, S.; Crain, J.; Poon, W. C.; Finney, J. L.; Soper, A. K. *Nature* **2002**, *416*, 829.
(30) Allison, S. K.; Fox, J. P.; Hargreaves, R.; Bates, S. P. *Phys. Rev. B* **2005**, *71*, 024201.
(31) Perera, A.; Sokolić, F.; Almasy, L.; Koga, Y. *J. Chem. Phys.* **2006**, *124*, 124515.
(32) Sokolić, F.; Idrissi, A.; Perera, A. *J. Chem. Phys.* **2002**, *116*, 1636.
(33) Stumpe, M. C.; Grubmüller, H. *J. Phys. Chem. B* **2007**, *111*, 6220.
(34) Dougan, L.; Bates, S. P.; Hargreaves, R.; Fox, J. P.; Crain, J.; Finney, J. L.; Réat, V.; Soper, A. K. *J. Chem. Phys.* **2004**, *121*, 6456.
(35) Ganguly, P.; Schravendijk, P.; Hess, B.; van der Vegt, N. F. A. *J. Phys. Chem. B* **2011**, *115*, 3734.
(36) Algaer, E.; van der Vegt, N. F. A. *J. Phys. Chem. B* **2011**, *115*, 13781.
(37) Mukherji, D.; van der Vegt, N. F. A.; Kremer, K. *J. Chem. Theory Comput.* **2012**, *8*, 3536.