Functional Group Contributions to Partial Molar Compressibilities of Alcohols in Water

Daren M. Lockwood and Peter J. Rossky*

Institute for Theoretical Chemistry, Department of Chemistry and Biochemistry, University of Texas at Austin, Austin, Texas 78712-1167

Ronald M. Levy

Department of Chemistry, Rutgers University, Piscataway, New Jersey 08855-0939

Received: November 30, 1999; In Final Form: February 15, 2000

A recently developed methodology for evaluating solute functional group contributions to the partial molar compressibility of solution [Lockwood, D. M.; Rossky, P. J. *J. Phys. Chem. B.* **1999**, *103*, 1982] is used to examine solvation of methanol and ethanol in water. Additive contributions to the compressibility are resolved for the methyl and hydroxyl groups of each solute, and the results are shown to be the same for both solutes within statistical error. Further, the effect of each functional group on the solvent is found to be localized in the vicinity of that functional group, explaining the apparent independence of functional group contributions observed experimentally by other workers. The difference in partial molar compressibility between methanol and ethanol can be attributed to localized solvent perturbation by the methylene group in ethanol. For the potential functions employed, compressibilities calculated via classical molecular dynamics simulations are in best agreement with experiments performed at temperatures higher than those at which the simulations are performed. Implications of calculated alcohol group contributions for studies of more complex solutes, including biomolecules, are discussed.

1. Introduction

Compressibilities of aqueous solutions of biomolecules have been the object of much attention in the recent literature.[1−4] One reason is the increasing precision and availability of experimental data.[1,3,4] A second reason is the high sensitivity of partial molar compressibilities to solvation effects.[1] Compressibility measurement has emerged as a useful way to probe protein interiors and conformational states, which is an important driving force for work in this area.[2,3] Investigators have also been intrigued by the fact that measurements of solution compressibility, while fundamentally thermodynamic or macroscopic in nature, suggest clear microscopic pictures of solutions,[5] creating the opportunity for fruitful interplay between experimental and molecular simulation investigations.[6]

One of the most intriguing aspects of solution compressibilities is the success of group-additive models in predicting experimental values. For example, in dilute aqueous solutions of solutes composed of repeating subunits—such as solutes containing alkyl components with differing numbers of methylene groups—the compressibility is a remarkably linear function of the number of subunits.[4,7−9] While exceptions to simple group additivity have been reported in the literature,[10] the applicability of group additive models is quite widespread.[4,7] Group additivity is observed for a number of thermodynamic properties other than partial molar compressibility,[11] but the sensitivity of compressibilities to solvation effects makes group additivity for this property particularly noteworthy. Indeed, group additivity would not be expected in the case of compressibilities unless functional group effects on the solvent were largely isolated from one another, suggesting that important solvent perturbation by a given functional group is rather localized to the region of that group. Solution compressibility data therefore provides important evidence for the so-called hydration shell model of solvation,[6] which attributes excess thermodynamic properties to a perturbed solvent layer near the solute surface and assumes that solvation effects fall off rapidly with distance from the solute surface.

Analysis of thermodynamic properties in terms of the hydration shell model is a long-standing practice.[12,13] It is only recently, however, that the range and extent of its validity has begun to be carefully explored.[6] Molecular dynamics calculations have been particularly useful in this regard. Such calculations provide detailed microscopic information about solvent molecular configurations, and so questions for which experimental studies do not provide definitive answers can be addressed. In the present context, we note that high-pressure simulations of a solvated protein showed significant changes in the structure of the hydration layer around certain functional groups;[14] the excess compressibilities of these groups has been implicated in a proposed mechanism for the pressure-induced unfolding of proteins.[15]

The present investigation was undertaken with the goals of evaluating functional group contributions to excess compressibilities for n-alcohols in water and of exploring the range of important solute effects on the solvent. We recently developed a methodology that may be used for this purpose.[16] n-Alcohols are ideal model compounds for gaining physical insight into the origin of the group additivity seen in partial molar compressibilities of more complex solutes. This is, in part, because n-alcohols contain both a polar and a hydrophobic group, and these groups play a prominent role in group-additive models for more complex biomolecules.[4] An additional advantage of studying n-alcohols is the fact that the partial molar

* To whom correspondence should be addressed.

10.1021/jp994197x CCC: $19.00
Published on Web 04/08/2000

compressibilities for these molecules have been determined experimentally and have been found to vary almost linearly with the number of methylene groups, providing a straightforward example of group additivity.[17]

The remainder of the paper is organized as follows. In section 2, we review the statistical mechanical formulation of excess compressibilities that makes it possible to isolate functional group contributions to solution compressibilities. Particular attention is paid to the role of so-called sphere distribution functions in extracting the net effects of solute hydration shells. Additionally, we discuss how the maximum entropy method can be used to analyze molecular dynamics distribution function data that is subject to significant statistical error. In section 3, we describe specific procedures for calculating the key quantities of interest. In section 4, we discuss results obtained by applying these procedures to the specific cases of methanol and ethanol solvation. Implications for solvation of large biomolecules are noted.

2. Theory

In previous work,[16] we outlined a methodology for evaluating functional group contributions to the partial molar compressibilities of solvated molecules. Here, we review only the important components of that methodology. In the first of two subsections, the statistical mechanical formulation that permits expression of the partial molar compressibility in terms of spatially localized contributions is reviewed. In the second section, we discuss how the maximum entropy method can be used to estimate these contributions when significant statistical errors exist in the data.

2.1. Partial Molar Compressibility. The partial molar compressibility for a single solute in pure solvent is defined as the change in the solution compressibility that occurs due to the addition of a solute molecule under the conditions of constant pressure P, constant temperature T, and a constant number of solvent molecules, N:

$$\kappa_2^0 = -\frac{\partial}{\partial N_s}\left(\frac{\partial V}{\partial P}\right)_{NT}\Big|_{N_s\to 0} \tag{1}$$

where κ_2^0 is the partial molar compressibility, N_s is the number of solute molecules, and $N_s \to 0$ indicates the limit of infinite dilution. This partial molar compressibility may be expressed[16,18] as the sum of two components: an "ideal" contribution, which is equal to the partial molar compressibility for an ideal gas solute, and an "excess" contribution, which expresses the effects specific to the solution in question:

$$\kappa_2^0 = \kappa_2^{0,id} + \rho_0^{-1}\Delta\kappa \tag{2}$$

Here, $\kappa_2^{0,id}$ represents the ideal contribution, whereas $\Delta\kappa$ represents the excess contribution and ρ_0 is the average density of solvent in the solution. The excess contribution is amenable to evaluation via classical molecular dynamics simulations[6,16] using periodic boundary conditions, since it may be expressed as an integral over the volume of a unit cell that contains the solute in a fixed position and the solvent:

$$\Delta\kappa = \rho_0\int_V d\vec{r}\,\frac{\partial}{\partial P}(\rho(\vec{r})/\rho(\infty)) \tag{3}$$

where $\rho(\infty)$ is the average density of the solvent after removal of the solute at constant pressure, and $\rho(\vec{r})$ is the solvent density at the location $\vec{r}$ within the unit cell. This equation is correct if $\rho(\vec{r})$ is a nuclear particle density distribution function,[6] but other representations of the density are equally acceptable, as will be discussed below.

If the hydration shell model were strictly valid, then it would be possible to truncate this integral at a cutoff distance $|\vec{r}| = \lambda$ around the solute, beyond which the density equals $\rho(\infty)$. Furthermore, if the volume of space contained within this cutoff boundary is small compared to the unit cell volume, then $\rho(\infty)$ in eq 3 may be replaced within the boundary by ρ_0, the density obtained after constant V, rather than constant P, removal of the solute:[16]

$$\Delta\kappa = \rho_0\int_{|\vec{r}|\le\lambda} d\vec{r}\,\frac{\partial}{\partial P}g_{sx}(\vec{r}) \tag{4}$$

where $g_{sx}(\vec{r})$ is the normalized particle−particle distribution function $\rho_{sx}(\vec{r})/\rho_0$. The subscripts s and x have been introduced to emphasize the fact that $g_{sx}(\vec{r})$ and $\rho_{sx}(\vec{r})$ are particle−particle distribution functions and that they are uniquely defined functions if one specifies the solute molecular center s and solvent molecular center x on which they are based.[19] Since *any* selection of these molecular centers—including so-called "auxiliary sites" that do not contribute to the intermolecular potential[19]—will leave the value of the integral in eq 3 unchanged, the selection of defining molecular centers in eq 4 is important. As a simple example, the selection of s can affect the required size of the region enclosed by a cutoff boundary. This is because $\vec{r}$ represents displacement from the solute center s, and the condition $|\vec{r}| \le \lambda$ is only useful for gauging proximity to the solute insofar as $|\vec{r}|$ is itself a useful basis on which to judge proximity to the solute.

To resolve functional group contributions to the excess compressibility, the integral in eq 4 may be expressed as a sum of separate integrals over separate regions of space,[16] each containing different functional groups. Functional group contributions may be obtained if each integral is insensitive to the precise integral cutoff distance for distances reasonably proximate to a physical functional group center. We have demonstrated previously[16] that this condition is satisfied if one replaces the particle−particle distribution function in eq 4 by a more suitable distribution function; traditional particle−particle distribution functions[6] are not acceptable.[16] Here, we review the concept of "sphere distribution functions" and then explain why they have been found to be ideal for evaluating excess compressibilities via eq 4.[16]

Sphere distribution functions are generated by averaging over alternative solvent particle distribution functions, each of which is based on an alternative selection of the solvent molecular center x. As noted above, each of these alternative solvent particle distribution functions gives the same value for the integral in eq 3. Clearly, this implies that a distribution function $\rho(\vec{r})$, which is itself an average over solvent particle distribution functions, will also give the same result in eq 3. On the other hand, eq 4 is only valid for distribution functions that attain their asymptotic value within the integral cutoff distance λ of the solute, as discussed above. Sphere distribution functions are based upon a specific distribution of alternative solvent molecular centers x: molecular centers are distributed uniformly, and exclusively, within a certain radius R of a single solvent center, β.[16] As such, sphere distribution functions may be expressed in terms of the particle−particle distribution function $\rho_{s\beta}(\vec{r})$ as

$$\rho(\vec{r}) = \int d\vec{r}\,'\rho_{s\beta}(\vec{r}\,')P(\vec{r} - \vec{r}\,', R) \tag{5a}$$

where $P(\vec{r},R)$ is the probability distribution of centers x:

$$P(\vec{r},R) = \begin{cases} 3\pi^{-1}R^{-3}/4 \text{ for } |\vec{r}| < R \\ 0 \text{ for } |\vec{r}| \geq R \end{cases} \quad (5b)$$

Clearly, sphere distribution functions $\rho(\vec{r})$ depend on two parameters: a radius R for the spherical distribution of alternative solvent molecular centers x and the selection of a single solvent center β about which the alternative solvent molecular centers x are distributed. In our previous work on aqueous solutions,[16] we employed oxygen as our water molecule center β and determined that eq 4 is valid for an accessible cutoff distance λ if one uses a distance R commensurate with the minimum separation of solvent molecules in solution, which is about 2.3 Å for aqueous solutions. In this work, we will generally refer to water sphere distribution functions based on these parameters simply as 2.3 Å-sphere distribution functions. Calculation of these functions is facilitated if one uses the convolution theorem to reexpress eq 5 in the form[16,20]

$$g(\vec{r}) = (2\pi)^{-3} \int \mathrm{d}\vec{k} \mathrm{e}^{i\vec{k}\cdot\vec{r}} \hat{g}_{s\beta}(\vec{k})[3k^{-2}R^{-2}(k^{-1}R^{-1}\sin kR - \cos kR)] \quad (6a)$$

where $g(\vec{r})$ is expressed as an integral over momentum space and $\hat{g}_{s\beta}(\vec{k})$ represents the Fourier transform of the particle distribution function $g_{s\beta}(\vec{r})$ on which the sphere distribution function is based

$$\hat{g}_{s\beta}(\vec{k}) = \int \mathrm{d}\vec{r} \mathrm{e}^{-i\vec{k}\cdot\vec{r}} g_{s\beta}(\vec{r}) \quad (6b)$$

Again, β labels the solvent molecule center on which the sphere distribution function is based, and in this work the choice $\beta =$ oxygen is used in the case of water solvent molecules.

The difficulty that arises if one uses a simple water particle distribution function in eq 4 is that packing forces lead to large long-ranged oscillations in the calculated excess compressibility as a function of integral cutoff distance.[6] These oscillations make no net contribution to the compressibility,[16] but this is not obvious from results for simple particle distribution functions alone, nor is it convenient to determine the asymptotic converged value of the compressibility as a function of integral cutoff distance. On the other hand, 2.3 Å-sphere distribution functions effectively carry out a coarse-grained averaging over radial regions; the purely oscillatory contributions associated with packing effects are averaged out, and net contributions from solvent regions are obtained.[16] The asymptotic converged value is the same in both cases, but sphere distribution functions permit a significantly shorter integral cutoff distance, revealing the locality of functional group effects and helping one to isolate important functional group effects on the solvent from one another.

To employ these sphere distribution functions in eq 4, it is necessary to have their derivative with respect to pressure. This quantity may be estimated by finite difference.[16] Herein, we use the following expression:[6]

$$\partial g(\vec{r},\rho_0)/\partial P \approx (\partial\rho_0/\partial P)(g(\vec{r},\rho_0 + \Delta\rho_0/2) - g(\vec{r},\rho_0 - \Delta\rho_0/2))/\Delta\rho_0 \quad (7)$$

where $\Delta\rho_0$ represents a finite change in the average solvent density in solution. Errors that are due to the finite difference approximation scale as the square of this differential and may be neglected for a suitably small differential.[21]

In practice, the methodology outlined above is not sufficient to permit evaluation of functional group contributions to partial molar compressibilities via the molecular dynamics calculations performed in this work. This is because of significant statistical errors present in the distribution function data obtained from molecular dynamics calculations carried out for computationally reasonable simulation times. Therefore, we now discuss how the maximum entropy method can be used to estimate compressibility contributions in the face of the statistical errors.

2.2. Maximum Entropy Method. There are significant statistical errors associated with the estimation of functional group compressibility contributions from computationally convenient simulation times.[16] Here, we review and generalize our previous discussion of the way in which the maximum entropy method may be used to make improved estimates from noisy distribution function data.[16] A number of good references are available that discuss the maximum entropy method.[20,22,23]

The maximum entropy method concerns the calculation of a discretized probability density function ("pdf") $\{p_i\}$ from a set of data points $\{d_i\}$ with which there are associated significant statistical errors. In this case, a number of significantly different pdf's may be considered to be consistent with the data points. A Bayesian approach to data analysis[22,23] suggests that the probability associated with a given pdf $\{p_i\}$ is

$$P(\{p_i\}|\{d_i\}) = N \exp\left[\alpha S(\{p_i\}) - \frac{1}{2} L(\{p_i\},\{d_i\})\right] \quad (8)$$

Here, N is a normalization constant, and S is the so-called "entropy" of the pdf $\{p_i\}$ and a measure of the a priori probability of the pdf. L is a measure of the experimental likelihood of the pdf, based upon the calculated statistical error in the function $\{d_i\}$. The maximum entropy prescription for selecting $\{p_i\}$ is to select the pdf that maximizes the function $P(\{p_i\}|\{d_i\})$; in this case, α plays the role of a regularization parameter, balancing optimization of the a priori probability against optimization of the experimental likelihood. The entropy of a pdf $\{p_i\}$ is defined as

$$S = \sum_i p_i - \sum_i m_i - \sum_i p_i \ln(p_i/m_i) \quad (9)$$

where $\{m_i\}$ is the so-called "default model" for $\{p_i\}$ and is chosen on the basis of the "principle of indifference"[16] to express total ignorance of the dataset.

In the case where the errors in the data points $\{d_i\}$ are nearly independent and are well represented by Gaussian distributions, the function L may be equated to the reduced χ^2 function:[22,23]

$$L \approx \chi^2 \equiv N_D^{-1} \sum_i [(Kp_i - d_i)^2/\sigma_i^2] \quad (10)$$

where $\{\sigma_i^2\}$ represents the standard deviations in the data, N_D is the number of data points, and K is an operator which translates the function $\{p_i\}$ into the space of the data.[22] However, in the event that cross-correlations between data points have significant effects, a more general form of the function L must be employed.[22] In either case, the form of the function L remains valid even if $\{p_i\}$ is not strictly a pdf but may take on negative values. However, the entropy function S must be altered in such a case. In the present work, we are interested in the function $\Delta g(\vec{r}) = g(\vec{r},\rho_0 + \Delta\rho_0/2) - g(\vec{r},\rho_0 - \Delta\rho_0/2)$, which is a function of the difference between two pdf's with the same default model.[16] One reasonable measure of the information contained in such a function[16,23] is obtained by maximizing the sum of the entropies of the two pdf's, subject to a constraint on the difference between the two that enforces consistency with the data. As discussed previously,[16] for the case of a difference

between sphere distribution functions, the resulting entropy expression can be approximated by

$$S = -\sum_i V_i |\Delta g_i|^2 \quad (11)$$

where $\{V_i\}$ represents a set of finite, separate volumes that make up the total volume of interest and Δg_i denotes the average value of $\Delta g(\vec{r})$ within V_i. Therefore, in the absence of significant effects of cross-correlations in the data, optimizing $P(\{p_i\}|\{d_i\})$ corresponds to optimizing

$$Q = -\alpha \sum_i V_i |\Delta g_i|^2 - (1/2) \sum_i [(\Delta g_i - \Delta g_{\mathrm{exp},i})^2 / \sigma^2_{\Delta g_i}] \quad (12)$$

where $\Delta g_{\mathrm{exp},i}$ denotes the value of Δg_i derived from the data and $\sigma_{\Delta g_i}$ is the associated error due to the data. Since the pdf takes on both positive and negative values, the "historic" rather than the "classic" method is used to select α.[16,20] In the historic method, α is selected so that $\chi^2 \leq 1$.

In our previous work,[16] we used a weighted χ^2 constraint to obtain a good fit to data in the vicinity of the solute. Also, information obtained using this constraint was incorporated into the entropy optimizations via modification of the default model introduced in eq 9. The need for these two additions to the method arose because of the large response of the distribution functions to pressure in the vicinity of the solute. This created an impetus for the maximum entropy method to remove features in this region beyond what was called for by the uncertainties in the data. Indeed, it was found that data in the vicinity of the solute was subject to the least statistical error. Applying this method, we determined that the dominant contributions to the compressibility—calculated with 2.3 Å-sphere distribution functions—come from the subspace of points whose distances from a solute Lennard-Jones center is less than the sum of the Lennard-Jones radii of the solute and solvent molecules. We will refer herein to this region as the "vdw separation" region of the solute and solvent.

On the basis of this determination, a simpler method was employed in this work: data outside the vdw separation region was entropy-optimized, whereas data inside that region was not. Results obtained in this way are very similar to those obtained by the method used in our previous work, as just described.[16] But the need for either a weighted χ^2 constraint or further modification of the default model is obviated. It should be noted that the justification for this simpler method rests upon results obtained with the previous method, and so the simpler method should not be regarded as independent of the previous one.

For the systems investigated in this work, we did not find that use of the more general form of L, which allows for cross-correlations between data points, had any significant effect on the optimized pdf, and so the less general expression $L = \chi^2$ was used in this work. Errors $\sigma_{\Delta g_i}$ used in the χ^2 function were estimated by dividing the simulated trajectory into five equal parts and calculating a standard deviation over the values corresponding to each segment. It seems reasonable that cross-correlations are relatively less important as the error level in the data increases. In our case, the maximum entropy method mainly mitigates the impact of the high level of statistical error present in solvent regions far from the solute, so that contributions from data in this region may be largely neglected.[16]

3. Simulation Procedures

Molecular dynamics simulations[24] were carried out in the canonical ensemble, at temperatures of 10 °C and 65 °C, using the periodic boundary conditions convention with cubic unit cells. Cell edge lengths of 27.31 and 26.41 Å were used to define the pair of densities, and the number of solvent water molecules in every case was 647, yielding average solvent densities of 0.95 and 1.05 g/cm^3, respectively. In each simulation performed, the cell contained a fixed solute molecule in addition to the solvent molecules. Solutes considered were methanol (at 10 °C) and ethanol (at both temperatures), where ethanol was fixed in the trans conformation. Methanol and ethanol molecules were described by the OPLS[25] parameter set, whereas the SPC[26,27,28] potential was used for the water molecules. Velocities were sampled according to the Boltzmann distribution every 500 time steps throughout the entire simulation;[24] the time step was 2 fs. After 50 ps of equilibration, coordinates were recorded every 25 time steps in a 4-ns trajectory and were used to determine solute−oxygen particle distribution functions.[24]

Equation 6 was used to convert these solute−oxygen particle distribution functions into 2.3 Å-sphere distribution functions, and the derivatives with respect to pressure were calculated by means of eq 7 (the required experimental value of the derivative of the solvent density with respect to pressure can be found in ref 29). These derivatives were in turn used in eq 4 to calculate excess compressibilities: they replace the derivatives of the particle−particle distribution functions in eq 4, as discussed in the text accompanying this equation. The calculated excess compressibilities can be converted to experimentally accessible partial molar compressibilites of dilute aqueous solutions by means of the expression:[16]

$$\kappa_2^0 = \rho_0^{-1} \Delta\kappa - kT(\partial\kappa_0/\partial P) \quad (13)$$

where $\Delta\kappa$ is the value obtained by simulation, and $(\partial\kappa_0/\partial P)$ is taken from experiment.[30]

4. Results and Discussion

We begin by examining the effect of the methyl group on the total excess compressibility of solvation for methanol in water. To do so, we consider a plane which bisects the carbon−oxygen bond and is composed of those points equidistant from the carbon and the oxygen atom centers in the methanol molecule. This plane serves to divide solvent molecules into two groups: those closest to the carbon atom and those closest to the oxygen atom. This is not the only possible selection of subspaces, but it will be seen to lead to isolation of important functional group effects. Subsequently, the plane is used to define "partial" carbon−oxygen (solute−solvent) distribution functions[16] at each density, centered on the carbon atom of the methyl group. This partial distribution function differs from a traditional distribution function in that all values of the distribution function at points on the other side of the plane from the carbon atom are zero. Related partial distribution functions based on proximity criteria have been described previously by Mehrotra and Beveridge[31] and have been employed by other authors to investigate thermodynamic properties other than compressibility.[32] The partial distribution functions obtained here are used to calculate partial 2.3 Å-sphere distribution functions via eq 6, which are in turn used in conjunction with eqs 7 and 4 (see section 3) to calculate an additive contribution to the compressibility. We note that a second additive contribution may be obtained on the other side of the plane from the carbon atom and that the sum of the two contributions will give the total excess compressibility of solvation for methanol in water.[16]

The calculated magnitude of the methyl group contribution to the excess compressibility, as a function of integral cutoff

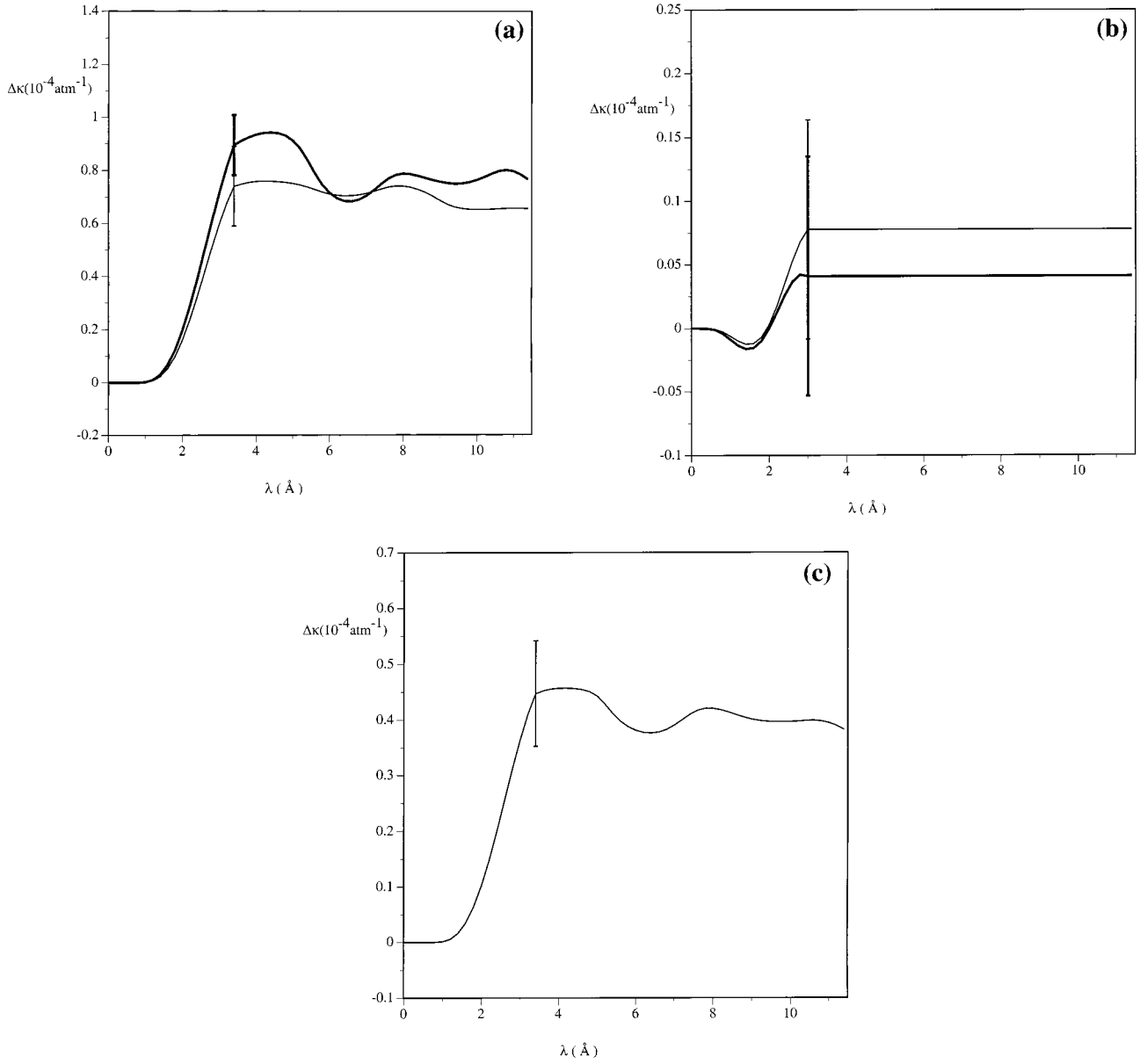

Figure 1. Estimated excess compressibility contributions, as a function of integral cutoff distance, at a simulated temperature of 10 °C. Error bars represent standard deviations over five equal subdivisions of the simulation time: (a) Contributions for the methyl groups in methanol (bold solid line) and ethanol (thin solid line), based on carbon−oxygen partial 2.3 Å-sphere distribution functions. Data outside the vdw separation region (λ < 3.5 Å) is entropy-optimized (see section 2.2); (b) Contributions for the hydroxyl groups in methanol (bold solid line) and ethanol (thin solid line), based on oxygen−oxygen partial 2.3 Å-sphere distribution functions. Data outside the vdw separation region (λ < 3.1 Å) is entropy-optimized; and (c) Contribution for the methylene group in ethanol, based on a carbon−oxygen partial 2.3 Å-sphere distribution function. Data outside the vdw separation region (λ < 3.5 Å) is entropy-optimized.

distance, is displayed in Figure 1a. The error bar indicates the standard deviation in the magnitude of this contribution for an integral cutoff distance equal to the vdw separation of the solute and solvent molecules. Points on the plot that correspond to greater separation were calculated using the maximum entropy method, as described in section 2.2.

For comparison, the calculated contribution of the methyl group in ethanol to the excess compressibility of ethanol in water is also shown in Figure 1a. In the ethanol case, a plane perpendicular to the carbon−carbon bond in ethanol was employed, and the distance from the methyl carbon was selected to match the displacement of the analogous plane in methanol from the methyl carbon atom. Within statistical error, the contributions of the methyl group in methanol and the methyl group in ethanol are the same. This is consistent with the success of group additive models in reproducing the experimental excess compressibilities of solvated alcohols in water.[17] Furthermore, important effects of the methyl groups on the solvent appear to be localized to the vicinity of those methyl groups, providing an explanation for additivity, namely, that group contributions arise largely in separate solvent regions and are thus substantially independent of one another. Before discussing the comparison between experimental and simulated values of the methyl group contribution, we discuss calculated contributions of the hydroxyl and methylene groups.

The calculated hydroxyl group contributions for methanol and ethanol are shown in Figure 1b. Apparently, the hydroxyl group contributions are subject to greater relative statistical error than are the methyl group contributions for the simulation times employed in this study. Furthermore, the maximum entropy method indicates that data beyond the vdw separation is too noisy to provide any information regarding the compressibility

(hence, the plot is flat). However, we can draw some significant and useful conclusions about the magnitude of the hydroxyl group contribution to the compressibility. First, we note that the experimental value for the excess compressibility of a water molecule in water is 0.45×10^{-4} /atm.[33] The plot in Figure 1b thus indicates that the magnitude of the hydroxyl contribution to the solution compressibility is less than half that of a water molecule and is significantly smaller than the compressibility contribution of the methyl group indicated in Figure 1a. Also, the hydroxyl contributions are the same for methanol and ethanol within statistical error.

Due to the fact that the simulated methyl and hydroxyl contributions for methanol and ethanol agree within statistical error and to the evident localization of the solute group effects on the solvent, we expect the difference in the excess compressibility of ethanol and methanol to be due to localized effects of the methylene group on nearby solvent molecules. In Figure 1c, we show the calculated contribution of the methylene group to the compressibility of solution for ethanol, based on the region of space containing the center of the methylene carbon and bound by the planes used in generating hydroxyl and methyl contributions. Localization to the region of the methylene carbon atom is indeed indicated by the rapid convergence. It should be noted that, rigorously, the sum of the three additive contributions evaluated for ethanol (corresponding to the methyl, methylene, and hydroxyl groups) is not equal to the total excess compressibility of ethanol in water, due to the fact that the subspaces containing the methyl and hydroxyl groups on ethanol overlap. However, this has no practical effect. If one subtracts from the "methylene" partial distribution function a second partial distribution function, centered on the methylene carbon and nonzero only in the "overlap region," then the resulting effective partial distribution function can be used to obtain a methylene contribution that does represent rigorous subdivision of the excess compressibility into additive components. The magnitude of the methylene contribution so calculated is not found to differ significantly from that calculated without taking the overlap region into account.

To test that the calculated excess compressibilities are not sensitive to either the selection of subspaces nor the selection of molecular centers for calculating compressibility contributions, we calculated the total excess compressibilities for methanol and ethanol based on full 3-dimensional distribution functions. The results are shown in Figure 2 as a function of integral cutoff distance, where the distance in question is based on the distance of a point from the "closest" Lennard-Jones center on the solutes. The word "closest" is placed in quotes here because cutoff distances from carbon atoms were incremented by the difference in Lennard-Jones radii of the carbon and oxygen atoms in order for the integral cutoff distance to reflect distance from the solute's Lennard-Jones surface. The results for the total excess compressibilities of the alcohols in water are within statistical error of the values obtained by adding up individual group contributions.

Because molecular dynamics calculations suggest that methyl, hydroxyl, and methylene contributions are nearly constant for n-alcohols, we expect the sum of the methyl and hydroxyl contributions to ethanol derived from simulation to compare well with the experimental excess compressibility for methanol, and we expect the difference between the experimental compressibilities for ethanol and methanol to compare well with the compressibility contribution of the methylene group derived from simulation. We begin with the experimental data for the methylene group, which provides a severe challenge for the molecular dynamics calculations. This is because the methylene contribution to the partial molar compressibility of ethanol in water is reported experimentally to depend very strongly on the temperature. Specifically, the experimental compressibility contribution of methylene rises from[34,35] -0.39×10^{-4} bar^{-1} at 10 °C to 0.40×10^{-4} bar^{-1} at 65 °C.

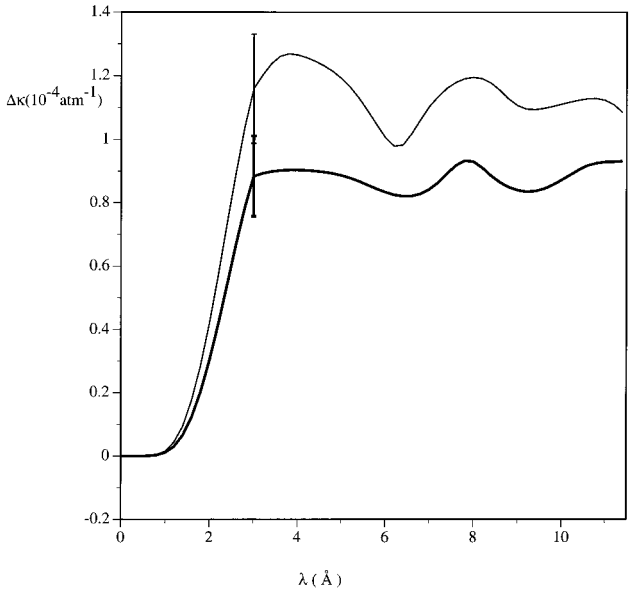

Figure 2. Total calculated excess compressibilities for methanol (bold solid line) and ethanol (plain solid line), as a function of integral cutoff distance from the vdw surface of the solutes, at a simulated temperature of 10 °C. Calculations are based on oxygen 2.3 Å-sphere distribution functions. Error bars represent standard deviations over five equal subdivisions of the simulation time. Data outside the vdw separation region is entropy-optimized (section 2.2).

The value of the methylene contribution estimated from simulation at 10 °C is 0.4×10^{-4} bar^{-1}. Comparison of the simulated methylene contribution at 10 °C, shown in Figure 1c, with that at 65 °C, shown in Figure 3a, indicates a derivative of the methylene contribution with temperature of roughly 0.003×10^{-4} bar^{-1}/°C (with large uncertainty), as compared to the experimental value of about 0.014×10^{-4} bar^{-1}/°C. Whereas the simulation-derived compressibility contribution is considerably less sensitive to temperature than the experimental value, the observed decrease of the compressibility with temperature is in accord with experiment. This fact, and the agreement of the simulated methylene contribution at 10 °C with the experimental value at 65 °C, indicate that it may be most consistent to compare the present simulation results with experiments performed at higher temperatures.

As far as the hydroxyl contribution is concerned, we show the calculated contribution to ethanol at 65 °C in Figure 3c and note that no statistically significant change with temperature is observed. Correspondingly, it is interesting to note that the excess compressibility of a water molecule in water is also rather insensitive to temperature.[33] This fact makes our previous comparison of the hydroxyl compressibility contribution with half of the excess compressibility of water more compelling. Assuming, as our simulation supports, that hydroxyl contributions to n-alcohol partial molar compressibilities in water are less than half the excess compressibility of a water molecule in water suggests a simple way to estimate methyl group contributions from experimental data. We subtract 25% of the excess compressibility of a water molecule (i.e., about 0.11×10^{-4} bar^{-1}) from the experimental methanol excess compressibility and expect the result to be within about 0.11×10^{-4} bar^{-1} of

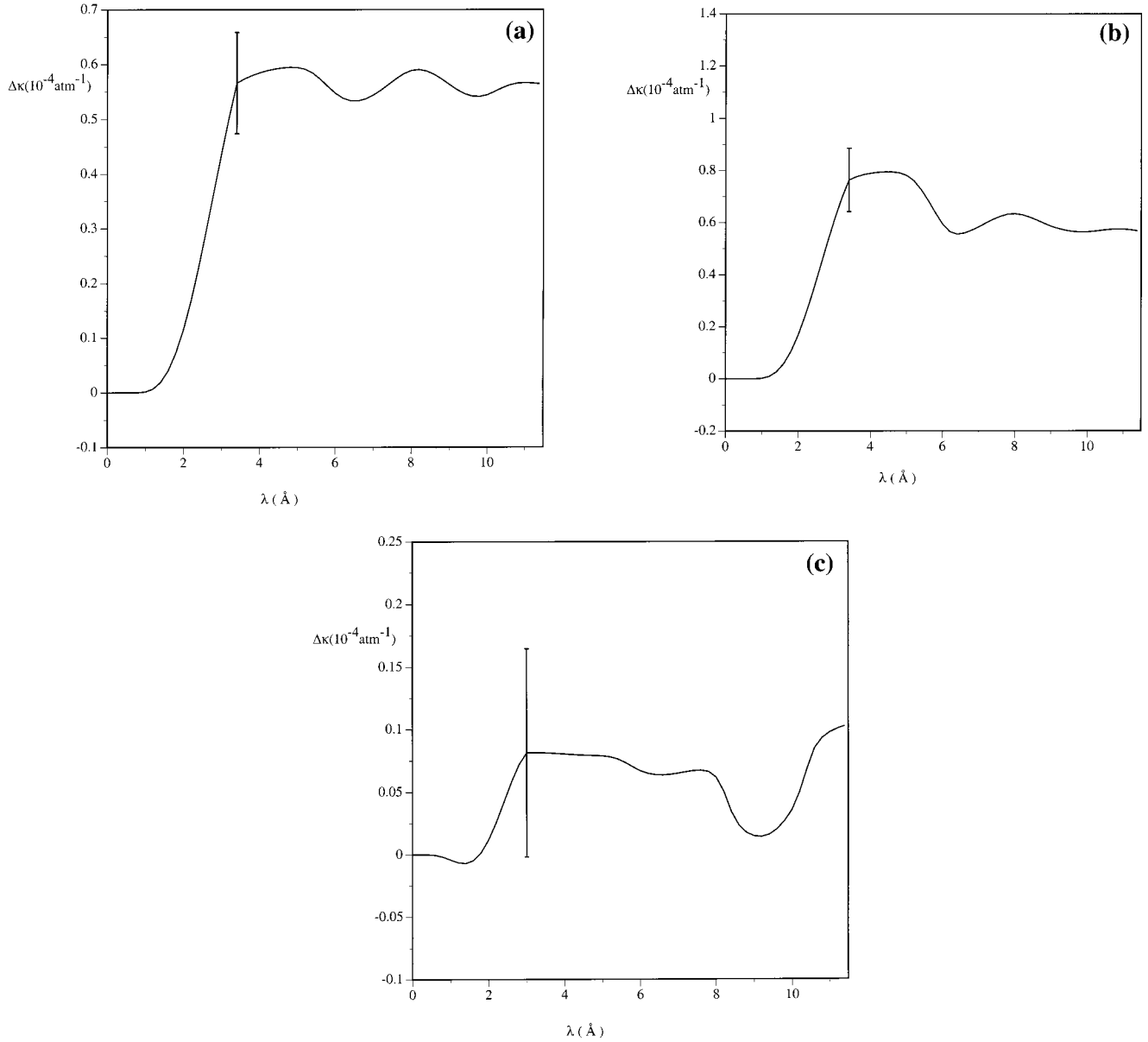

Figure 3. Estimated excess compressibility contribution, as a function of integral cutoff distance, for the groups in ethanol at a simulated temperature of 65 °C: (a) Contributions for the methylene group, based on a carbon−oxygen partial 2.3 Å-sphere distribution function. Data outside the vdw separation region (λ < 3.5 Å) is entropy-optimized (see section 2.2); (b) Contribution for the methyl group, based on a carbon−oxygen partial 2.3 Å-sphere distribution function. Data outside the vdw separation region (λ < 3.5 Å) is entropy-optimized; and (c) Contribution for the hydroxyl group. The calculation is based on an oxygen−oxygen partial 2.3 Å-sphere distribution function, and data outside the vdw separation region (λ < 3.1 Å) is entropy-optimized.

the "correct" value for the methyl contribution. Doing so, we find that the methyl contribution is indicated experimentally to rise from about 0.23×10^{-4} bar^{-1} at 10 °C to 0.87×10^{-4} bar^{-1} at 65 °C.

It is satisfying to observe that the experimentally indicated value of the methyl group at 65 °C compares well with the simulated value at 10 °C indicated in Figure 1a. This indicates that simulation-derived contributions from the methyl and methylene groups are each in agreement with the experimental result obtained at the *same* elevated temperature, a consistency that supports our analysis. The experimentally indicated temperature derivative for the methyl group is less than that for the methylene group, in agreement with our result at 65 °C shown in Figure 3b. As one might expect from the small methylene contribution derivative we observe, the methyl group derivative we observe is actually not different from zero within statistical uncertainty. One is then faced with the remarkable implication that under ambient experimental conditions, the methyl and methylene groups actually contribute to the compressibility with opposite sign. It is thus not at all clear that solvent surrounding hydrophobic surfaces on proteins should contribute in a uniform way to compressibility. A resolution of this question is needed for further analysis of the proposal that positive excess compressibilities of hydrophobic surfaces provide a mechanism for pressure-induced protein denaturation.[15]

Whether such uniformity in solvent compressibility contributions does indeed arise about more extended hydrophobic surfaces, or whether current additive models benefit from a sort of averaging over compressibility variations, is a very interesting question. Unfortunately, the potential functions used in this study limit our ability to answer this question at present, since at the higher temperatures that they apparently imitate, the large and intriguing disparity in methyl and methylene contributions per unit of solvent-accessible surface area inferred here at ambient

conditions does not arise. What the results from our methodology do supply is a satisfying explanation for group-additivity in model compounds, which apparently arises from highly localized and independent effects of different functional groups on the solvent.

Conclusions

We have evaluated additive functional group contributions for excess compressibilities of n-alcohols in water. Methyl and hydroxyl group contributions are each found to be the same for different n-alcohols, within statistical error. It is possible to regard the effects of different solute functional groups on the solvent as localized to separate regions of space and thus highly independent of one another. This provides an explanation for the success of group-additive models in predicting experimental values for partial molar compressibilities. Our calculations suggest that hydrophobic and hydrophilic groups on solutes differ markedly in their contributions to the partial molar compressibility. Most notably, hydrophobic group contributions are indicated to vary significantly with temperature, whereas hydrophilic group contributions do not. Furthermore, our analysis of both simulated and experimental data would imply that methylene and methyl groups contribute to the compressibility with different sign under ambient conditions. If this is indeed the case, in some cases it may be necessary to differentiate between different hydrophobic groups when developing group-additive models for the compressibility. The extent to which hydrophobic group contributions vary on complex biomolecules is thus an important question for future investigations, as is the sensitivity of simulated compressibility results to the details of the molecular models used.

Acknowledgment. The authors are grateful for the support of this research through a grant from the Robert A. Welch Foundation (F-0761). Additional support was provided by the NIH (GM49204). The authors also thank Nobuyuki Matubayasi and Emilio Gallicchio for stimulating discussions.

References and Notes

(1) Chalikian, T. V.; Totrov, M.; Abagyan, R.; Breslauer, K. J. *J. Mol. Biol.* **1996**, *260*, 588.

(2) Chalikian, T. V.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 1012.

(3) Kharakoz, D. P.; Sarvazyan, A. P. *Bioploymers* **1993**, *33*, 11.

(4) Chalikian, T. V.; Sarvazyan, A. P.; Breslauer, K. J. *Biophys. Chem.* **1994**, *51*, 89.

(5) Iqbal, M.; Verral, R. E. *Can. J. Chem.* **1989**, *67*, 729.

(6) Matubayasi, N.; Levy, R. M. *J. Phys. Chem.* **1996**, *100*, 2681.

(7) Iqbal, M.; Verral, R. E. *J. Phys. Chem.* **1987**, *91*, 967.

(8) Kharakoz, D. P. *J. Phys. Chem.* **1991**, *95* (5), 5634.

(9) Chalikian, T. V.; Sarvazyan, A. P.; Funck, T.; Breslauer, K. J. *Biopolymers* **1994**, *34*, 541.

(10) Chalikian, T. V. *J. Phys. Chem. B* **1998**, *102*, 6921.

(11) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Solution Chem.* **1981**, *10*, 563.

(12) Frank, H. S.; Evans, M. W. *J. Chem. Phys.* **1945**, *13* , 507.

(13) Rossky, P. J.; Karplus, M. *J. Am. Chem. Soc.* **1979**, *101* (1), 1913 and references therein.

(14) Kitchen, D. B.; Reed, L. H.; Levy, R. M. *Biochemistry* **1992**, *31*, 10083.

(15) Payne, V. A.; Matubayasi, N.; Murphy, L. R.; Levy, R. M. *J. Phys. Chem. B* **1997**, *101*, 2054.

(16) Lockwood, D. M.; Rossky, P. J. *J. Phys. Chem. B* **1999**, *103*, 1982.

(17) *Water Science Reviews 1*; Franks, F., Ed.; Cambridge University Press: New York, 1985 and references therein.

(18) Ben-Naim, A. *Solvation Thermodynamics*; Plenum Press: New York, 1987.

(19) Hansen, J. P.; McDonald, I. R. *Theory of Simple Liquids*; Academic Press: New York, 1986.

(20) *Numerical Recipes*; Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannerey, B. P., Eds.; Cambridge University Press: New York, 1992.

(21) Kincaid, D.; Cheney, W. *Numerical Analysis*; Brooks/Cole Publishing Company: Pacific Grove, CA, 1991.

(22) Gallicchio, E.; Berne, B. J. *J. Chem. Phys.* **1996**, *105*, 7064.

(23) Sivia, D. S. *Data Analysis: A Bayesian Tutorial*; Oxford University Press: New York, 1996.

(24) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: Oxford, 1987.

(25) Jorgensen, W. L. *J. Phys. Chem.* **1986**, *90*, 1276.

(26) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(27) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, G., Ed.; Reidel: Dordrecht, Holland, 1981.

(28) Andersen, H. C. *J. Comput. Phys.* **1983**, *52*, 24.

(29) Fine, R. A.; Millero, F. J. *J. Chem. Phys.* **1973**, *59*, 5529.

(30) Gibson, R. E.; Loeffler, O. H. *J. Am. Chem. Soc.* **1941**, *63* (3), 898.

(31) Mehrotra, P. K.; Beveridge, D. L. *J. Am. Chem. Soc.* **1980**, *102* (2), 4287. Methanol−water oxygen partial distribution functions used in this work are equivalent to "1°" distribution functions described by Mehrotra and Beveridge, but not all of the partial distribution functions in this work are equivalent to such "1°" distribution functions.

(32) Ashbaugh, H. S.; Paulaitis. M. E. *J. Phys. Chem.* **1996**, *100*, 1900.

(33) Ref 30 contains experimental compressibilities for water. An equation for conversion into excess compressibilities is given in ref 16.

(34) Cabani, S.; Conti G.; Matteoli, E. *J. Solution Chem.* **1979**, *8*, 11. Equation 13 can be used to convert partial molar compressibilities into excess compressibilities. Whereas the experimental data only spans the range of temperatures from 10 °C to 40 °C, data for methylene contributions to amino acids (see ref 8) strongly suggests that a linear extrapolation based on the data at 25° and 40° to higher temperatures will give good estimates of the compressibilities at higher temperatures.

(35) We note that the wording in ref 16 may cause confusion. The experimental value for the excess compressibility of methanol cited there (0.67×10^{-4}/atm) was based on the density and compressibility of water at the elevated pressure that was simulated. The excess compressibility of methanol under ambient conditions is about 0.55×10^{-4}/atm.