Molecular Origin of Anticooperativity in Hydrophobic Association

Cezary Czaplewski,[†,‡] Adam Liwo,[†,‡] Daniel R. Ripoll,[§] and Harold A. Scheraga*[,†]

Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-1301, Faculty of Chemistry, University of Gdańsk, ul. Sobieskiego 18, 80-952 Gdańsk, Poland, and Cornell Theory Center, Ithaca, New York 14853-3801

Received: October 29, 2004; In Final Form: March 2, 2005

The structuring of water molecules in the vicinity of nonpolar solutes is responsible for hydrophobic hydration and association thermodynamics in aqueous solutions. Here, we studied the potential of mean force (PMF) for the formation of a dimer and trimers of methane molecules in three specific configurations in explicit water to explain multibody effects in hydrophobic association on a molecular level. We analyzed the packing and orientation of water molecules in the vicinity of the solute to explain the effect of ordering of the water around nonpolar solutes on many-body interactions. Consistent with previous theoretical studies, we observed cooperativity, manifested as a reduction of the height of the desolvation barrier for the trimer in an isosceles triangle geometry, but for linear trimers, we observed only anticooperativity. A simple mechanistic picture of hydrophobic association is drawn. The free energy of hydrophobic association depends primarily on the difference in the number of water molecules in the first solvation shell of a cluster and that in the monomers of a cluster; this can be approximated by the molecular surface area. However, there are unfavorable electrostatic interactions between the water molecules from different parts of the solvation shell of a trimer because of their increased orientation induced by the nonpolar solute. These electrostatic interactions make an anticooperative contribution to the PMF, which is clearly manifested for the linear trimer where the multibody contribution due to changes in the molecular surface area is equal to zero. The information theory model of hydrophobic interactions of Hummer et al. also explains the anticooperativity of hydrophobic association of the linear trimers; however, it predicts anticooperativity with a qualitatively identical distance dependence for nonlinear trimers, which disagrees with the results of simulations.

1. Introduction

Hydrophobic association of nonpolar solutes in water arises from solvent-induced interactions (i.e., *indirect* interactions among solutes) caused by the presence and the particular structure of the solvent. The structure of a liquid solvent is more difficult to capture than that of solids because of its dynamic and disordered nature.[1,2] Water, the most familiar of liquids, and aqueous solutions are studied today as extensively as ever. Water is a unique liquid with a 3D network of hydrogen bonds. Modification of the structure and dynamics of water due to the presence of a nonpolar solute molecule is termed hydrophobic hydration,[3] and the term hydrophobic interactions (or hydrophobic association) describes water-mediated interactions between nonpolar solutes.[3] The term hydrophobic effect frequently refers to both hydrophobic hydration and hydrophobic interactions. The hydrophobic effect is of fundamental importance in the molecular mechanism of formation and stability of various self-assembly aggregates and biological structures in aqueous environment. These include biophysical phenomena such as protein folding, the formation of micelles, biological membranes and macromolecular complexes, and protein−ligand binding as well as protein−lipid interactions.[4−11]

Well-defined thermodynamic properties of hydrophobic hydration are observed experimentally.[5,9−12] Nonpolar solutes are only slightly soluble in water. The observed hydration properties of nonpolar solutes are characterized by a negative enthalpy and a negative entropy of solution. Near room temperature, the hydration process is dominated by the entropy term, which results in a positive free energy of solution and accounts for the low solubility. The solubility of small nonpolar solute species in water decreases when the solution is heated and increases when the system is cooled near room temperature.[4,12] The partial molar heat capacity of nonpolar solutes in water is large and positive at room temperature,[5,13] which stands in sharp contrast to the observations made for hydrophilic solutes in water.[14,15] The tendency for hydrophobic association of nonpolar solutes in water can be regarded as a partial reversal of the thermodynamically unfavorable process of hydrophobic hydration.[3,5,11] Near room temperature, hydrophobic association is characterized by a positive enthalpy and a positive entropy of association. The largely favorable entropy of association dominates and provides a net attraction between nonpolar solutes in water.

All of these experimental observations have been attributed by Frank and Evans[16] to changes in the structure of water around nonpolar groups to a more ordered state compared to the structure of bulk water. Némethy and Scheraga proposed a statistical mechanical model of liquid water[1,17] and applied it to study hydrophobic hydration.[3,12] The model was based on a partial cage or clathrate picture,[18] with five energy states for the water molecules depending on the number of hydrogen bonds formed by each water molecule. The model led to the correct signs and magnitudes of the free energy, enthalpy, entropy, and heat capacity of hydrophobic hydration and

* Corresponding author. E-mail: has5@cornell.edu. Phone: (607) 255 4034. Fax: (607) 254 4700.
† Cornell University.
‡ University of Gdańsk.
§ Cornell Theory Center.

10.1021/jp040691b CCC: $30.25 © 2005 American Chemical Society
Published on Web 03/30/2005

hydrophobic interactions. The increased ordering of the water around nonpolar solutes was explained by the decrease in the energy of water molecules with four hydrogen bonds and the increase in the energy of water molecules with fewer hydrogen bonds formed.[3,11,12] This model of hydrophobic hydration was later verified by many Monte Carlo[19,20] and molecular dynamics[21−25] simulations of aqueous solutions of nonpolar molecules. Pratt and Chandler[26] developed an integral-equation theory to calculate the potential of mean force of hydrophobic association; this model was later extended by Hummer and coworkers[27−29] by incorporating concepts from information theory.

Despite the wide agreement about the importance of the ordering of water around hydrophobic solutes, there are remarkably few direct experimental measurements that reveal the specific changes in water structure in solutions of nonpolar molecules.[30,31] One difficulty is that techniques such as infrared, Raman, and magnetic resonance spectroscopy can provide evidence only for an increase in the strength of hydrogen bonds or a decrease in motion in a solution of nonpolar molecules and cannot show specific structural changes. X-ray scattering can provide a direct measure of the oxygen−oxygen distribution function, $g_{OO}(r)$, but this is a rather insensitive measure of structural changes and depends on the model used to extract $g_{OO}(r)$ from the X-ray data.[3,32] The most sensitive measure of water structure changes is neutron scattering using isotope substitution.[33] This technique provides a direct determination of the hydrogen−hydrogen distribution function, $g_{HH}(r)$, which is more sensitive to changes in water orientation than $g_{OO}(r)$. Another problem is that it is difficult to dissolve a sufficient quantity of a nonpolar substance in water to produce enough perturbation to be observed in the experiment. Most of the work on the structure of water around hydrophobic solutes has been done on polar and ionic compounds such as alcohols or tetraalkylammonium ions.[33]

Due in part to difficulties in direct experimental studies of water structure and other microscopic details of hydrophobic hydration, extensive use has been made of computer simulations in hydrophobic hydration studies.[19−22,34−36] Molecular simulations, indeed, provide valuable insight into the nature of hydrophobic interactions. For a detailed understanding of solute association, the potential of mean force (PMF) is very informative. The PMF represents the free energy of the solutes in a solvent as a function of solute−solute geometry. Most of the studies of hydrophobic interactions at the molecular level characterize only the pairwise PMF for two nonpolar particles in aqueous solution[19,22,37−41] and its dependence on environmental variables such as temperature,[5,42−45] pressure,[46−48] and ionic strength of the solution.[49]

Némethy and Scheraga[5] noticed that hydrophobic interactions between more than two solute particles are not additive and cannot in general be expressed as a sum of solute−solute pair PMFs. Hydrophobic interactions can be cooperative, with the actual free energy more favorable to association than the pairwise sum, or anticooperative, with the actual free energy less favorable to association than the pairwise sum. Simulations by Rank and Baker[50] of the PMF of three methane molecules in the specific configuration of an isosceles triangle, with the methane dimer at contact distance forming a fixed base, using the free-energy perturbation (FEP) method and the TIP4P water model, have shown that the three-body PMF is anticooperative for distances up to 6.5 Å. However, Rank and Baker reported that the errors associated with baseline estimation (i.e., the zeroing of the absolute free energies at long distances) were of the same order of magnitude (0.2−0.3 kcal/mol) as the three-body term itself, which did not provide a reliable conclusion about the magnitude and sign of the nonadditive contribution to the PMF.

Our molecular-dynamics study of the hydrophobic association of two model systems, three methane molecules in an isosceles triangle, as in the study of Rank and Baker, and an equilateral triangle geometry, using umbrella sampling and the weighted histogram analysis method (WHAM) and two water models (TIP3P and TIP4P), showed that hydrophobic interactions are cooperative and that the three-body term increases the strength of hydrophobic association by about 0.1 kcal per methane−methane pair, or about 10% of the depth of the contact minimum in the PMF.[23] In the case of an equilateral triangle geometry, the cooperativity extended up to a distance of 6.5−7 Å depending on the water model; in the case of the isosceles triangle, the three-body term was close to zero (showing additivity) for distances smaller than 4 Å and negative (showing cooperativity) for distances up to 5.5−6 Å with a minimum at 4.9 Å. To minimize errors associated with the zeroing of the absolute free energies at long distances, we estimated the baseline of the cooperative term by assuming that this term vanishes with distance faster than the PMF itself.

Shimizu and Chan[51] presented results for a particle-insertion simulation of the hydrophobic interaction of three methane molecules in an isosceles triangle geometry using the TIP4P water model, which contradicted those of our molecular simulation study.[23] They suggested that there is anticooperativity in hydrophobic interactions that extends over a significant range of separations, from 3 to 13 Å. Shimizu and Chan estimated the baseline using the free energy of hydration of a single methane molecule from the particle-insertion method. In an exchange of comments,[52,53] we criticized Shimizu and Chan for not presenting any estimate of the errors of the pairwise PMF used for calculating the three-body term and for neglecting the minimum-image convention, and they criticized our method of baseline estimation. In their recent paper,[54] Shimizu and Chan presented a full free-energy landscape for a fixed methane dimer and an approaching methane molecule. They revised their two-methane PMF to a higher accuracy than the one used in their previous study[51] but have not addressed the issue of edge effect in which all their curves are terminated at 11 Å, close to half of the box distance. Their three-body PMF curve for the isosceles triangle arrangement is shifted up by about 0.1 kcal/mol compared to our results. Shimizu and Chan[53,54] tried to explain this by long-range effects in hydrophobic association. Nevertheless, the spatial range of hydrophobic interactions is finite, and the PMF must tend to zero with increasing solute−solute distance.[52] Distances larger than 10.5 Å, at which it is possible to observe the second solvent-separated minimum, are large enough to show that the PMF curves become flat at large distances.[52]

Recently, we carried out simulations of larger nonpolar solutes than methane in an equilateral triangle geometry using the TIP3P water model.[24] As in our study of the methane trimer,[23] we found that the three-body term is cooperative (i.e., it *increases* the strength of hydrophobic interactions). The simulations were carried out in a 28 Å periodic box with solute−solute distances up to 14 Å, at which all PMF curves become flat within the error inherent in simulations, which provided a reliable estimate of the baseline. Also, our most recent simulations of the association of clusters of four methane molecules, a trigonal pyramid with three methane molecules at contact distance forming a fixed base and a regular uniformly expanding tetrahedron, confirmed that multibody terms are cooperative for

most distances.[25] The multibody contribution was found to amount to about 0.2 kcal/mol per methane−methane pair, or about 25% of the depth of the contact minimum in the PMF. The main effect of the multibody contribution to the PMF was a reduction of the height of the barrier between the contact- and solvent-separated minima and a narrowing of the region of its maximum, whereas the region of the contact minimum was affected only weakly. In the case of the isosceles triangle (the side of a trigonal pyramid), a small anticooperativity was observed near the contact minimum and past the PMF maximum. It should be noted that our results showed that the multibody contribution to the PMF of a three-methane cluster in the isosceles triangle geometry can be slightly positive (anticooperative) at the contact minimum (3.9 Å) and in the region of the solvent-separated minimum (7.2 Å), in agreement with the data of Shimizu and Chan, but that there is a significant decrease of the barrier between the contact- and solvent-separated minima due to multibody interactions, which in turn means that there is cooperativity in the region of this barrier.[25] In all of our studies,[23−25] the distance dependence of the multibody contributions to the PMF was in a good agreement with that of the molecular surface area, which confirms the observation of Rank and Baker[50] that molecular surface area provides a good description of hydrophobic hydration. However, the distance dependence of the molecular surface area does not fully account for the anticooperativity observed for the methane trimer at the isosceles triangle geometry.

Ghosh et al.[55] used molecular dynamics simulations of aqueous solutions of 10 nonpolar solutes (one system containing 10 methane molecules, and one containing 10 larger nonpolar particles) in the TIP3P water model to compute two- and three-particle PMFs directly from pair and triplet density correlations. They also reported many-body effects for changes in hydrostatic pressure and in the ionic strength of a solution. Dominant features in the triplet correlations were observed only within the first solvation shell of nonpolar solutes, which suggests that the nonadditivity of hydrophobic interactions at the level of triplet correlations is limited to short distances, in agreement with our results and not those of Shimizu and Chan. Additivity/nonadditivity properties were reported[54] to depend on the geometric arrangement of the three solutes: overall, the contact- and solvent-separated minima configurations were anticooperative, whereas the barrier configuration was mostly cooperative, also in agreement with our results. Details of the distance dependence of the multibody term for the isosceles triangle geometry were not presented, thus direct comparison with our data is not possible. The solute−solute−solute three-particle correlations were calculated using a large bin width of 0.4 Å, which limits the resolution of the PMFs and multibody terms; for example, there is no second solvent-separated minimum visible in any PMF.[55] The conclusion of Ghosh et al. was that the sign of the nonadditivity effect (i.e., anticooperativity or cooperativity) appears to be correlated to the strength of the individual pair interactions involved: more favorable individual pair interactions at contact- and solvent-separated minima lead to anticooperativity, whereas extending one of the pair distances in the triplet to a barrier configuration with less favorable interactions leads to cooperativity.

In this study, we intend to provide a graphical picture of the molecular process of hydrophobic association and explain anticooperative effects on a molecular level. To investigate the PMF of many-body interactions, we simulated the formation of a dimer and trimers of methane molecules in three specific configurations in explicit water, as shown in Figure 1a−d. We carried out simulations in a large periodic box to address the baseline-estimation problem. We analyze the packing and orientation of water molecules in the vicinity of the solutes to explain the effect of ordering of the water around nonpolar solutes on many-body interactions.

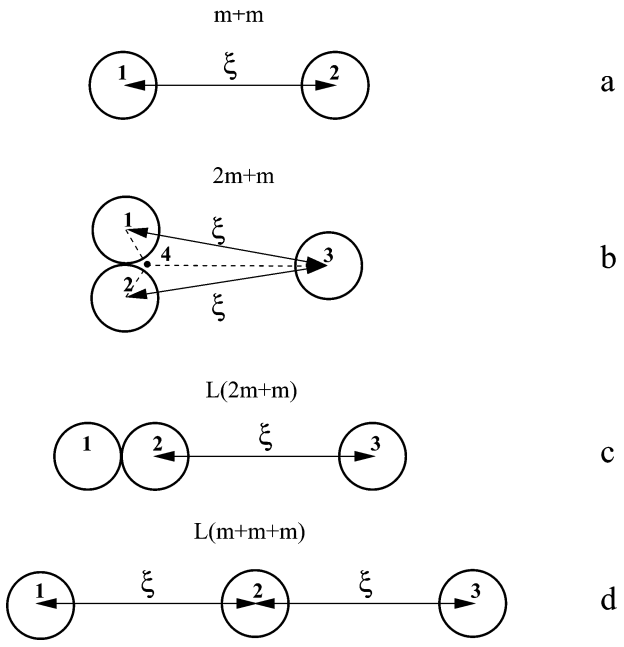

Figure 1. Scheme of the methane clusters studied: (a) dimer, (b) trimer with isosceles-triangle geometry, (c) linear trimer with molecules m_1 and m_2 at contact distance, and (d) symmetric linear trimer. ξ is the reaction coordinate (a dot marked 4 is a dummy atom positioned at the center of an equilateral triangle with side 3.9 Å and m_1 and m_2 molecules at two of its vertices). See section 2.1 of the text for an explanation.

2. Methods

2.1. Definition of Cooperativity in Hydrophobic Association. In general, the potential of mean force W of an n-particle interaction can be decomposed into single-body, pairwise, and multibody terms

$$W(1, 2, ..., n) = F(1, 2, ..., n) - \sum_{i}^{n} F^{(1)}(i)$$

$$= \sum_{i<j}^{n} \delta F^{(2)}(i, j) + \sum_{i<j<k}^{n} \delta F^{(3)}(i, j, k)$$

$$+ ... + \sum_{i<j<...}^{n} \delta F^{(l)}(i, j, ...) + ... + \delta F^{(n)}(i, j, ..., n) \quad (1)$$

where $F(1, 2, ..., n)$ is the shorthand for $F(\xi_1, \xi_2, ..., \xi_{3n-6})$, with ξ_i being the ith reaction coordinate. $F^{(1)}(i)$ corresponds to hydrophobic hydration (serving to define the reference state), $\delta F^{(2)}(i, j)$ is the two-body term of hydrophobic association, and $\delta F^{(l)}(i, j, ...)$ corresponds to subsequent higher-order terms. In this work, we investigated the lowest-order cooperative term, $\delta F^{(3)}$.

The three-body term $\delta F^{(3)}$ can be derived recursively from eq 1. For a three-body system, $\xi_1 = r_{12}$, $\xi_2 = r_{13}$, and $\xi_3 = r_{23}$, where r_{ij} represents the site−site distances identified with the respective reaction coordinate. Given the unnormalized total PMF, $F(r_{12}, r_{13}, r_{23})$, for the cluster of three solute molecules

in water, the three-body cooperativity term, $\delta F^{(3)}(r_{12}, r_{13}, r_{23})$, can be expressed as

$$\begin{aligned}\delta F^{(3)}(r_{12}, r_{13}, r_{23}) &= F(r_{12}, r_{13}, r_{23}) - \delta F^{(2)}(r_{12}) - \delta F^{(2)}(r_{13}) - \delta F^{(2)}(r_{23}) - 3F^{(1)} \\ &= F(r_{12}, r_{13}, r_{23}) - 3F^{(1)} - [F(r_{12}) - 2F^{(1)}] - [F(r_{13}) - 2F^{(1)}] - [F(r_{23}) - 2F^{(1)}] \\ &= W^{(3)}(r_{12}, r_{13}, r_{23}) - [W^{(2)}(r_{12}) + W^{(2)}(r_{13}) + W^{(2)}(r_{23})] \end{aligned} \quad (2)$$

where $F(r_{ij})$ is the unnormalized total PMF for an isolated solute pair in water and $F^{(1)}$ is the hydrophobic hydration free energy of a single solute molecule, which is an unknown constant. Computing the differences $F(r_{12}, r_{13}, r_{23}) - 3F^{(1)}$ and $F(r_{ij}) - 2F^{(1)}$ is equivalent to shifting the respective total PMFs to zero values at infinite (practically sufficiently large) distances between the solute molecules and provides the *normalized* PMF, $W^{(3)}(r_{12}, r_{13}, r_{23})$ and $W^{(2)}(r_{ij})$, respectively. Thus, $\delta F^{(3)}(r_{12}, r_{13}, r_{23})$ can be determined from the difference between the three-body and the sum of two-body normalized PMFs.

The determination of the PMF of two solute molecules in water is a 1D problem, and that of the PMF of three solute molecules is a 3D one. In this article, we study the three-particle PMF in selected 1D subspaces. We considered three geometries for the methane trimer presented in Figure 1b−d: (a) an isosceles triangle (2m + m) (Figure 1b); as in our previous work[23,25] (the distance between the first and second molecule is constrained to 3.9 Å, the contact-minimum distance between two methane molecules in water), the distance of the third methane molecule from the first and the second ones, respectively, is the reaction coordinate $\xi = \xi_{13} = \xi_{23}$; (b) a linear L(2m + m) system (Figure 1c) with a third methane molecule approaching a methane dimer along its axis, with the distance between the first and second molecules restrained to 3.9 Å as in 2m + m and the reaction coordinate $\xi = \xi_{23} = \xi_{13} - 3.9$; (c) a linear L(m + m + m) system (Figure 1d) with two methane molecules approaching the central one from both sides simultaneously, with the reaction coordinate $\xi = \xi_{12} = \xi_{23} = \xi_{13}/2$.

2.2. Calculation of the Potential of Mean Force (PMF). To evaluate the PMF, we used umbrella-sampling molecular dynamics (MD)[56,57] with the weighted histogram analysis method (WHAM).[58,59] In the umbrella-sampling method, a series of restraints (usually harmonic) are imposed on the reaction coordinate to ensure that all regions are sampled sufficiently. Each restraint defines a sampling window.

The molecular dynamics umbrella-sampling simulations were carried out using 21 windows with a harmonic restraining potential

$$V(\xi) = k(\xi - d_o)^2 \quad (3)$$

with the force constant $k = 2$ kcal/mol/Å^2 and the "equilibrium" distance for a given window, d_o, equal to 3.5, 4.0, ..., 13.5 Å for windows 1 to 21. The reaction coordinate ξ was chosen as the solute−solute distance. In the case of the m + m dimer, the intermolecular distance was restrained by eq 3. In the case of the 2m + m system, umbrella-sampling simulations were carried out by imposing a 2 kcal/mol/Å^2 restraint on the m_1-m_3 and m_2-m_3 distances with d_o values varying as described above and with the m_1-m_2 distance fixed at 3.9 Å (the methane−methane contact-minimum distance in water) using the SHAKE algorithm.[60] Additionally, we restrained the angles $m_1-x_4-m_3$ and $m_2-x_4-m_3$ to 120° by a harmonic force constant of 100 kcal/mol/rad^2 (x_4 is a dummy atom positioned at the center of the equilateral triangle with side 3.9 Å and m_1 and m_2 molecules at two of its vertices) to obtain more efficient sampling of the isosceles triangle geometry. In the case of the linear L(2m + m) trimer simulations, we used a harmonic force constant of 2 kcal/mol/Å^2 on all distances (i.e., m_1-m_2, m_1-m_3, and m_2-m_3). The equilibrium distance $d_o^{m_2-m_3}$ varied from 3.5 to 13.5 Å as described above; $d_o^{m_1-m_2}$ was set at 3.9 Å, and $d_o^{m_1-m_3} = d_o^{m_1-m_2} + d_o^{m_2-m_3}$. Additionally, the angle $m_1-m_2-m_3$ was restrained to 180°, and the angles $m_2-m_1-m_3$ and $m_2-m_3-m_1$ were restrained to 0° with a harmonic force constant of 100 kcal/mol/rad^2 to maintain the linear geometry of the trimer. In the case of the linear L(m + m + m) trimer, the box size limited umbrella-sampling simulations to 15 windows with d_o of the harmonic restraining potential in eq 3 equal to 3.5, 4.0, ..., 10.5 Å for distances m_1-m_2 and m_2-m_3 in windows 1 to 15. The d_o value for the m_1-m_3 distance was doubled, the angle $m_1-m_2-m_3$ was restrained to 180°, and the angles $m_2-m_1-m_3$ and $m_2-m_3-m_1$ were restrained to 0° with a harmonic force constant of 100 kcal/mol/rad^2 to maintain the linear geometry of the L(m + m + m) trimer.

Simulations with the methane dimer (m + m) and an isosceles triangle trimer (2m + m) were carried out by using a cubic periodic box with a 28.7 Å side and a 9.5 Å cutoff for both the nonbonded and electrostatic interactions. The number of water molecules in the box was 788 for m + m and 787 for the 2m + m system. The size of a set of molecules connected by harmonic restraints or by constraints is limited to half of the box dimension, and to simulate linear trimers, the periodic box was increased to 41 Å and contained 2291 water molecules. The TIP4P model was used for the water molecules, with parameters (Table 1) taken from Jorgensen et al.[61] In our previous work,[23] we have shown that the use of the spherical cutoff instead of the expensive Ewald summation for electrostatic interactions does not affect the PMF of hydrophobic association. Moreover, the TIP4P model of water used in this study was parametrized[61] to work with a spherical cutoff and not the Ewald summation; its version designed for the Ewald summation (TIP4P-Ew) appeared only recently.[62] Each methane molecule was modeled by a single interaction site, with van der Waals (VDW) parameters[63] listed in Table 1. The VDW parameters of the oxygen−solute interaction were calculated by using the standard Lorentz−Berthelot mixing rules.[64] The simulations were carried out in the *NVT* ensemble with temperature 298 K using the Berendsen coupling algorithm.[65] The time step in the MD simulations was 2 fs, and a SHAKE algorithm[60] was applied to keep the water molecules rigid. For each umbrella-sampling window, 100 ps of equilibration was followed by 4 ns of data collection simulations. The total

TABLE 1: Force Field Parameters for the TIP4P Water Model[61] and the Methane Model[63] Used in This Study[a,b]

atom type	σ (Å)	ϵ (kcal/mol)	q (eu)
methane			
C4	3.7300	0.294	0.0
TIP4P			
HW	0.0	0.0	+0.52
OW	3.1535	0.155	0.0
M	0.0	0.0	−1.04

[a] C4 represents the methane molecule, HW and OW represent hydrogen and oxygen atoms of the water molecule, respectively, and M represents a dummy atom in the TIP4P water model. [b] The form of the pair potential is $U(r) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6] + q_i q_j (4\pi r)^{-1}$.

simulation time reached 84 ns for the m + m and 2m + m systems and 60 ns for the L(m + m + m) system. The Cartesian coordinates of the solute molecules were stored at every time step of the MD simulation. The simulations were carried out by using the GROMACS[66] program.

To calculate the PMF from umbrella-sampling simulation data, the data from all simulation windows must be combined, and the contribution due to the restraining potential must be eliminated. We used the weighted histogram analysis method (WHAM)[58,59] as the most advanced approach to this problem. Umbrella sampling combined with WHAM gives faster convergence and more stable results than other methods.[23] In WHAM, the *unbiased* probability density corresponding to a given value of the reaction coordinate ξ is computed as a weighted average over all windows. The formulas for the weights[58,59] were derived on the basis of the requirement that the sum of the squares of the errors over all ξ in the calculated distribution be minimized. For a more detailed description of the method, the reader is referred to our previous paper[23] and the original papers by Kumar et al.[58,59] The bin dimension applied in the WHAM calculations of the PMF was equal to 0.1 Å for all systems (a 1D bin for the dimer and 3D bins for the trimers). For the 2m + m and L(2m + m) systems, the bins corresponding to the distance of the fixed dimer (3.9 Å) were placed in the centers of the multidimensional bins corresponding to these variables. This assured that use was made of all data close to the subspace of variables considered in the calculation. All points collected from the MD simulations, including those not satisfying the target geometry (the isosceles triangle or linear trimers), were taken to calculate histograms using WHAM. Only the histogram values of the bins whose center positions satisfied the condition required for the target geometry of the systems described above were used for the calculation and visualization of the PMF. The PMF is evaluated from the unbiased histogram obtained from the WHAM method by using eq 4

$$W(\xi) = -RT \ln \frac{N(\xi)}{\Delta V(\xi)} \tag{4}$$

where $N(\xi)$ is the value of the histogram at ξ and $\Delta V(\xi)$ is the volume element corresponding to ξ. For the systems studied, $\Delta V(\xi)$ values are expressed by eqs 5−7, respectively,

$$\Delta V_{\mathrm{m+m}} = 4\pi\xi^2\Delta\xi \tag{5}$$

$$\Delta V_{\mathrm{2m+m}} = 8\pi^2 d\xi^2\Delta\xi^2 \tag{6}$$

$$\Delta V_{\mathrm{m+m+m}} = 16\pi^2\xi^3\Delta\xi^3 \tag{7}$$

where d is the length of the base of the equilateral triangle for the 2m + m system (equal to 3.9 Å) and $\Delta\xi$ is the bin size assumed to be equal in all dimensions.

The calculated PMFs can be identified with $W^{(2)}(r_{12})$ and $W^{(3)}(r_{12}, r_{13}, r_{23})$ of eq 2 for two- and three-solute molecules, respectively, and should therefore tend to zero with increasing distance (after subtracting the constant factor accounting for the hydrophobic-hydration free energy of an *isolated* solute molecule). The region 11.5−13.0 Å was used as a baseline to superimpose the PMFs of the trimer onto that of the dimer for all systems except the linear trimer L(m + m + m), for which the 8.5−10.0 Å region was used.

On the basis of eq 2 and the symmetry of the selected systems, we expressed the cooperative terms per pair of interactive methane molecules for systems 2m + m, L(2m + m), and L(m + m + m) as follows:

$$\delta F^{(3)}_{\mathrm{2m+m}}(\xi) = \frac{W^{(3)}_{\mathrm{2m+m}}(\xi)}{2} - W^{(2)}_{\mathrm{m+m}}(\xi) \tag{8}$$

$$\delta F^{(3)}_{\mathrm{L(2m+m)}}(\xi) = \frac{W^{(3)}_{\mathrm{L(2m+m)}}(\xi) - W^{(2)}_{\mathrm{m+m}}(\xi) - W^{(2)}_{\mathrm{m+m}}(\xi + 3.9)}{2} \tag{9}$$

$$\delta F^{(3)}_{\mathrm{L(m+m+m)}}(\xi) = \frac{W^{(3)}_{\mathrm{L(m+m+m)}}(\xi) - 2W^{(2)}_{\mathrm{m+m}}(\xi) - W^{(2)}_{\mathrm{m+m}}(2\xi)}{3} \tag{10}$$

2.3. Computing Molecular Surface and Cooperative Contributions to It. The molecular surfaces of the m + m methane dimer and the methane trimers studied in this work were calculated numerically using a standard triangulation method implemented in the MOLMOL program[67] with radius $r_o = 2.09$ Å for the methane molecule and a radius of 1.4 Å for the water probe.

By analogy to eq 8 and taking advantage of the fact that, for nonclashing solute molecules, $\delta S^{(2)}(2\xi) = \delta S^{(2)}(\xi + 3.9\ \text{Å}) = 0$ (see eq 13), the cooperative contribution to the molecular surface area of the 2m + m and L(2m + m) trimers is expressed by eq 11, and that of the L(m + m + m) trimer is expressed by eq 12

$$\delta S^{(3)}_{\mathrm{2m+m/L(2m+m)}}(\xi) = \frac{\bar{S}^{(3)}_{\mathrm{2m+m}}(\xi)}{2} - \bar{S}^{(2)}_{\mathrm{m+m}}(\xi) \tag{11}$$

$$\delta S^{(3)}_{\mathrm{L(m+m+m)}}(\xi) = \frac{\bar{S}^{(3)}_{\mathrm{L(m+m+m)}}(\xi) - 2\bar{S}^{(2)}_{\mathrm{m+m}}(\xi)}{3} \tag{12}$$

where $\delta S^{(3)}$ denotes the total cooperative contribution to the molecular surface and $\bar{S}^{(n)}$ denotes the normalized molecular surface of the cluster of n solute molecules.

$$\bar{S}^{(n)} = S^{(n)} - nS^\circ \tag{13}$$

with $S^\circ = 4\pi r_o^2$ being the molecular surface of a single solute particle.

For linear trimers L(2m + m) and L(m + m + m), $\delta S^{(3)} = 0$ because there is no contact between the first and third methane molecule on the opposite sides of the trimer. Therefore, the surface can be decomposed into parts that arise from pairs of solute particles.

2.4. Calculating the PMFs by Using the Information Theory Model of Hummer and Co-workers. We also used the information theory (IT) model of hydrophobic interactions introduced by Hummer and co-workers[27,28] to compute the PMFs of the systems under study. Briefly, the method maximizes the cross entropy (η) computed from probabilities of finding 0, 1, ... solvent molecules in the cavity of the shape of the solute, as given by eq 14.

$$\eta(\{p_n\}) = -\sum_{n=0}^{\infty} p_n \ln\left(\frac{p_n}{\hat{p}_n}\right) \tag{14}$$

where p_n is the probability of finding n solvent molecules in the cavity and $\hat{p}_n$ is the probability corresponding to the default model of choice. The entropy is maximized subject to the conditions that the probabilities are normalized to unity, and the average number of solvent molecules ($\langle n \rangle$) and the average

square of the number of solute molecules ($\langle n^2 \rangle$) in the region of space to be occupied by the solute are constrained by available information regarding these quantities in the pure solvent, as given by eqs 15−17, respectively.

$$\sum_{n=0}^{\infty} p_n = 1 \tag{15}$$

$$\sum_{n=0}^{\infty} n p_n = \langle n \rangle \tag{16}$$

$$\sum_{n=0}^{\infty} n^2 p_n = \langle n^2 \rangle \tag{17}$$

The values of $\langle n \rangle$ and $\langle n^2 \rangle$ can be computed from the density and oxygen−oxygen pair correlation function of pure water, as given by eqs 18 and 19, respectively,[27]

$$\langle n \rangle = \rho V \tag{18}$$

$$\langle n(n-1) \rangle = \langle n^2 \rangle - \langle n \rangle = \rho^2 \int_\Omega \int_\Omega g(|\mathbf{r} - \mathbf{r}'|)\, d^3\mathbf{r}\, d^3\mathbf{r}' \tag{19}$$

where V is the volume of the cavity, Ω is the region of space occupied by the cavity, ρ is the density of the pure solvent, and $g(r)$ is the solvent pair correlation function, which we identify with the oxygen−oxygen correlation function.

Maximization of the expression for the cross entropy (eq 14) under constraints given by eqs 15−17 by using the method of Lagrange multipliers results in eq 20 for the probabilities[27]

$$p_n = \hat{p}_n \exp(\lambda_o + \lambda_1 n + \lambda_2 n^2) \tag{20}$$

where λ_o, λ_1, and λ_2 are the Lagrange multipliers determined by substituting eq 20 into eqs 15−17 and solving the resulting system of nonlinear equations.

The excess chemical potential, $\mu^{ex}(\xi)$, of a hard-sphere hydrophobic-solute cluster in water at reaction coordinate ξ is calculated from the probability $p_o(\xi)$ of finding the cavity of the solute shape (eq 21), and the potential of mean force of the cluster at the reaction coordinate ξ, $W(\xi)$, is calculated as the difference between the excess chemical potential at coordinate ξ and that at infinite separation of particles (eq 22).[27]

$$\mu^{ex}(\xi) = -RT \ln p_o(\xi) = -RT\lambda_\circ(\xi) - RT \ln \hat{p}_o \tag{21}$$

$$W(\xi) = \mu^{ex}(\xi) - \lim_{x \to \infty} \mu^{ex}(x) \tag{22}$$

The method outlined above correctly predicts the free energies of formation of spherical cavities of increasing size[27] and the shape of the PMF of two spherical hydrophobic solutes in water;[27] in the latter case, it performs better than the molecular surface area models because it reproduces the first and second solvent-separated minima.[27,28] This is probably caused by the incorporation of solvent granularity through the solvent pair correlation function (eqs 17 and 19).

As did Hummer and co-workers,[27,28] we assumed a flat default model in which $\hat{p}_n = 1$ for $n \le n_{max}$ and $\hat{p}_n = 0$ otherwise. We set $n_{max} = 20$. To calculate the oxygen−oxygen pair-correlation function, we carried out a 5 ns *NVT* simulation of 790 water molecules in a cubic water box with a 28.7 Å side length (corresponding to a density of 0.994 g/cm^3) using the TIP4P model of water. Snapshots spaced by 0.2 ps (a total of 25 000 snapshots) were used to calculate the oxygen−oxygen pair

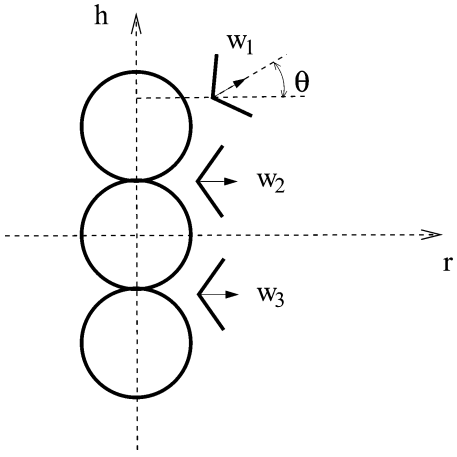

Figure 2. Scheme of orientations of water molecules w_2 and w_3 (with dipole-moment vectors shown) at the neighboring intersections of molecular surfaces of a linear trimer. An additional water molecule w_1 is presented (at an arbitrary orientation) to define the angle θ between a vector normal to the trimer axis and a dipole moment vector of the water molecule. The cylindrical axes h (the axis linking the molecules) and r (an axis perpendicular to it) are defined.

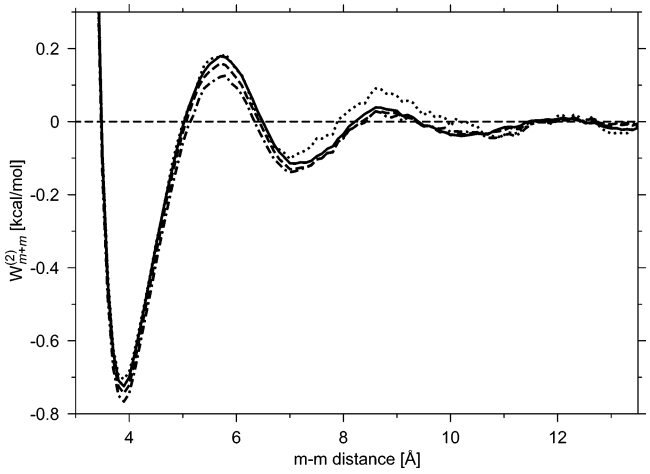

Figure 3. PMF curves for an increasing number of data points for the m + m system. The lines denote the duration of a run in each window: dotted line, 1 ns; dashed−dotted line, 2 ns; dashed line, 3 ns; solid line, 4 ns.

correlation function. A 3.3 Å effective solute radius was assumed as in the work of Hummer et al.[27] This parameter corresponds to the sum of the radius of the methane molecule and a water molecule (i.e., the minimum distance between the centers of a methane and a water molecule). Hummer et al.[27] justified the choice of such an effective solute radius by observation that this is the smallest distance at which methane−water pair correlations reach 1.0 in commonly used models.[68] We also find that the position of the solvent-separated minimum in the PMF of a pair of methane molecules (Figure 3) is about 6.6 Å, which gives exactly the 3.3 Å distance between a methane molecule and the intervening water molecule at this configuration. Thus, the radius assumed by Hummer et al.[27] is also consistent with the model used in our work. It should be noted that, by applying the combination rules to the van der Waals parameters of methane and oxygen collected in Table 1, we find that $\sigma_{C4\cdots OW} = 3.44$ Å, which is larger than the closest-approach radius of Hummer et al.; however, at this distance, the Lennard-Jones interaction energy of methane and water oxygen reaches zero, and these atoms can approach closer when the favorable interactions between water molecules (which tend to minimize the space not occupied by water) come into play.

The integrals defined by eq 19 were calculated by a Monte Carlo method. The system of nonlinear equations for λ_o, λ_1, and λ_2 (eqs 15−17) was solved by Newton's method. We determined the PMFs of the m + m, L(2m + m), 2m + m, and L(m + m + m) systems and also for the m + m + m system (a uniformly expanding equilateral triangle) studied in our previous work.[23] From the PMFs of the trimers and the PMF of the dimer, we computed the cooperativity terms, as described in section 2.2.

2.5. Analysis of the Packing and Orientation of Water Molecules in the Vicinity of Solutes. In umbrella-sampling simulations, only the Cartesian coordinates of the solute molecules are stored for PMF calculations. To analyze the distribution and orientation of water molecules around the interacting solutes, we ran an additional 20 ns of MD simulations for both the methane dimer and linear methane trimers with harmonic restraints to keep them at contact distances while we stored the Cartesian coordinates of all atoms every 0.2 ps for further analysis. (It should be noted that this additional simulation pertains to both the L(m + m + m) and the L(2m + m) systems because the distances between the central methane molecule and its neighbors are both equal to 3.9 Å.) Because the m + m, L(2m + m), and L(m + m + m) systems have cylindrical symmetry, we expressed the distributions in cylindrical coordinates *h* (the axis passing through the line linking the centers of the solute molecules) and *r* (an axis perpendicular to *h*). We computed the distribution of the density of the water molecules and the average values of the angle θ between the dipole moment of a water molecule and the *r* axis (cf. Figure 2) at all points of the grid. Two-dimensional maps were prepared with a 0.2 Å grid. The definition of the angle θ is illustrated in Figure 2.

3. Results and Discussion

3.1. Two-Body PMF. We previously computed and discussed the two-body PMF of the methane−methane interaction,[23,25] and hence we will discuss it only briefly here. The PMF is shown in Figure 3 where it converges well as the number of data increases. The PMF has a characteristic shape with a deep contact minimum at 3.9 Å, a solvent-separated minimum at about 7 Å (the distance at which one water molecule enters the space between the two solutes[21]), and a second solvent-separated minimum at about 10 Å. Detection of a very shallow second solvent-separated minimum can be a measure of the quality of the PMF because it would be hidden within the ruggedness of the curve because of a simulation error if a less accurate method is used to determine the PMF. At larger distances, the plot is almost flat. The baseline can thus be established by averaging the PMF at distances greater than 11.5 Å, and its estimated error cannot exceed the depth of the second solvent-separated minimum (i.e., about 0.05 kcal/mol; this is the upper estimate of the error, and the actual error is very likely to be much smaller).

3.2. Three-Body PMFs and Multibody Terms. We repeated the simulations for an isosceles triangle (2m + m) using the TIP4P water model and a larger box because this system exhibits both cooperativity and anticooperativity depending on the range of distances when the TIP3P water model is used.[25] The anticooperativity with the TIP3P water model[25] could not be explained by the distance dependence of the difference between the molecular surface area of the trimer and those of the component dimers ($\delta S^{(3)}$ of eq 11). We did not consider an equilateral triangle (m + m + m) because previous studies showed that there is only cooperativity and the distance dependence of the multibody term was well explained[25] by changes in the molecular surface area.

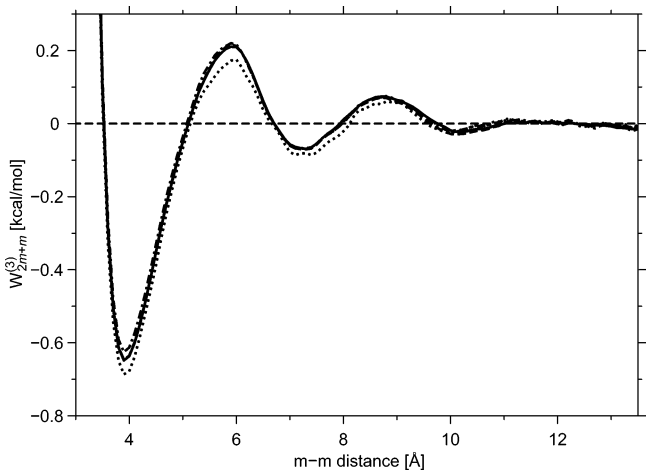

Figure 4. PMF curves for an increasing number of data points for the 2m + m system. The lines denote the duration of a run in each window: dotted line, 1 ns; dashed−dotted line, 2 ns; dashed line (which merges with the solid line at short distances), 3 ns; solid line, 4 ns.

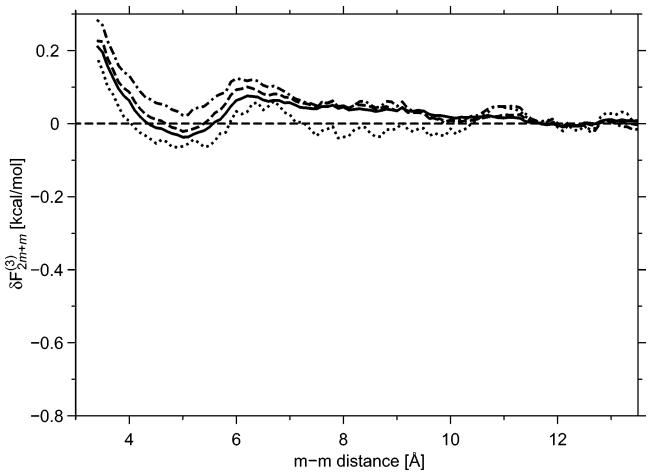

Figure 5. Curves for the cooperative term $\delta F^{(3)}$ for the 2m + m system. The lines denote the duration of a run in each window in the parent PMF simulations of Figure 4.

The three-body PMF for the 2m + m system is shown in Figure 4. It can be seen that the PMF curve has the same features as those corresponding to the two-body PMF (Figure 3) with one contact minimum and two solvent-separated minima present. With an increasing number of sampling points, the PMF curve converges quickly, and the plot becomes flat after 11.5 Å, thus enabling us to determine the baseline of the PMF.

The three-body contribution to the PMF ($\delta F^{(3)}_{2m+m}$ calculated from eq 8) of the 2m + m system is shown in Figure 5. It can be seen that the curves corresponding to increasing amounts of data converged to the final curve, although the convergence is slower than for the parent PMF curves of Figures 3 and 4. The $\delta F^{(3)}_{2m+m}$ curve exhibits a minimum with a slightly negative value at 5 Å. (This distance is located between the contact minimum and the desolvation maximum; the desolvation maximum is defined as the maximum in the PMF curve that is located between the contact and the first solvent-separated minimum.) There is anticooperativity (positive $\delta F^{(3)}_{2m+m}$) in the region of the contact minimum (3.9 Å) and also in the region of the solvent-separated minimum and beyond where $\delta F^{(3)}_{2m+m}$ decreases slowly to zero. An inspection of the $\delta S^{(3)}_{2m+m}$ curves in Figure 6 reveals that the anticooperativity in the contact-

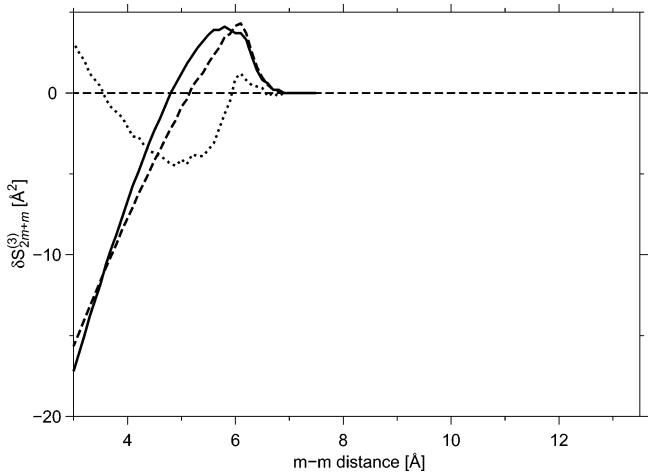

Figure 6. Distance dependence of the molecular surface area for the methane dimer ($\bar{S}^{(2)}_{m+m}$, solid line), the methane trimer in isosceles-triangle geometry ($\bar{S}^{(3)}_{2m+m}/2$, dashed line), and the contribution of the three-body term to the molecular surface area ($\delta S^{(3)}_{2m+m}$, dotted line).

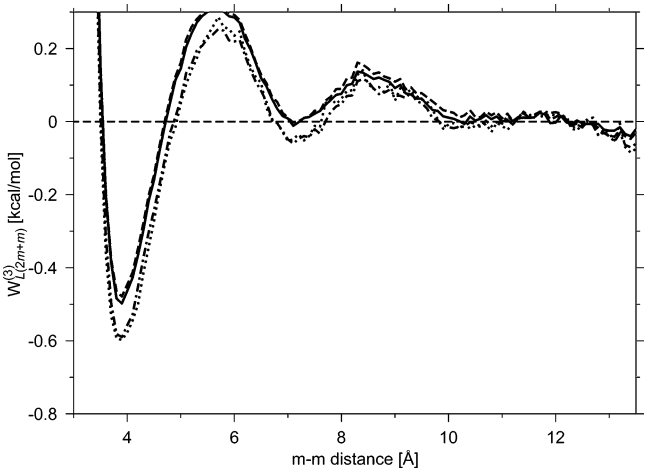

Figure 7. PMF curves for an increasing number of data points for the L(2m + m) system. The lines denote the duration of a run in each window: dotted line, 0.5 ns; dashed−dotted line, 1 ns; dashed line, 1.5 ns; solid line, 2 ns.

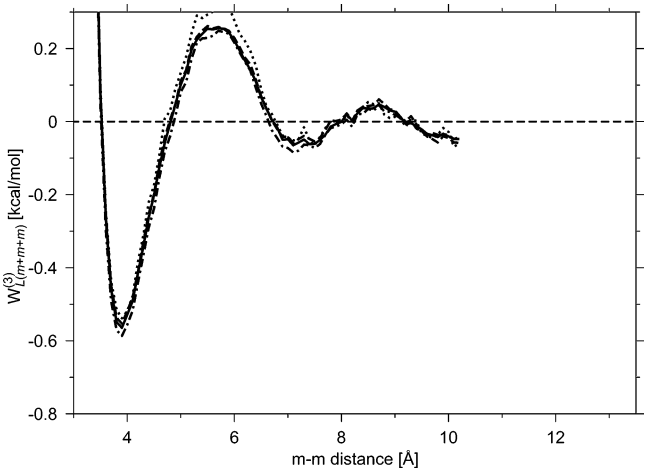

Figure 8. PMF curves for an increasing number of data points for the L(m + m + m) system. The lines denote the duration of a run in each window: dotted line, 0.5 ns; dashed−dotted line, 1 ns; dashed line, 1.5 ns; solid line, 2 ns.

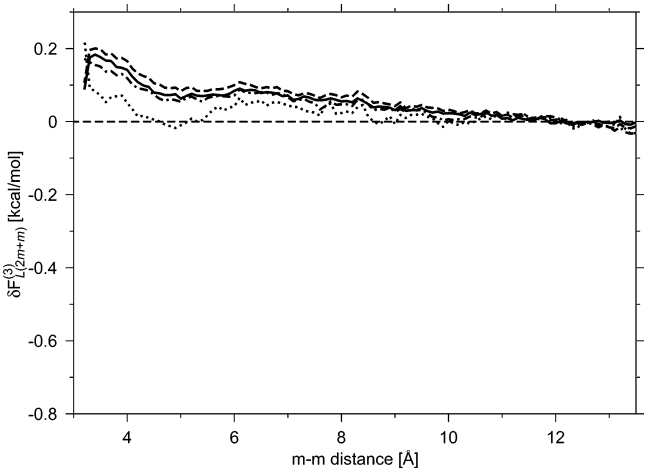

Figure 9. Curves for the cooperative term $\delta F^{(3)}$ for the L(2m + m) system. The lines denote the duration of a run in each window in the parent PMF simulations.

minimum region can, to some extent, be explained in terms of the molecular surface area because $\delta S^{(3)}_{2m+m}$ is positive to the left of the contact-minimum region. However, there is only a very small region with positive $\delta S^{(3)}_{2m+m}$ at about 6 Å, as opposed to a large region of PMF anticooperativity at distances beyond the solvent-separated minimum in Figure 5. Moreover, the $\delta F^{(3)}_{2m+m}$ curve appears to be shifted upward compared to that of $\delta S^{(3)}_{2m+m}$; this feature was not observed in our previous work for the PMF of the methane trimer in the equilateral-triangle geometry (m + m + m).[25] Therefore, the analysis of the molecular surface area of the 2m + m system in Figure 6 is not sufficient to explain the multibody contribution to the hydrophobic interactions for that system in Figure 5.

To find the origin of anticooperativity in the hydrophobic interactions in all methane trimers considered here and in refs 23 and 25 except for the equilateral triangle, we studied two systems in which the three-body contributions to $\delta S^{(3)}$ are exactly equal to zero, namely, a methane dimer at the contact minimum plus an additional methane molecule that moves along the m + m axis (the L(2m + m) system) and a linear methane trimer (the L(m + m + m) system), shown in Figure 1c and d, respectively. Because one dimension of these systems can take larger values than that of the 2m + m system, we had to use a larger periodic box, as pointed out in Methods. The PMFs of the trimers are shown in Figures 7 and 8, respectively, and the corresponding $\delta F^{(3)}$ graphs are shown in Figures 9 and 10, respectively. It can be seen that the curves for the total PMF exhibit the same features as that of the pairwise PMF and that of the 2m + m system, respectively, whereas the $\delta F^{(3)}$ graphs are different from those of the 2m + m system and show only anticooperativity with $\delta F^{(3)}$ decreasing to zero with distance. It can also be seen (Figures 9 and 10) that the shapes of the $\delta F^{(3)}$ curves of the L(2m + m) and the L(m + m + m) systems correspond to the part of the $\delta F^{(3)}$ graph of the 2m + m system beyond 6 Å (Figure 5); this suggests that the origin of anticooperativity in the PMF of the 2m + m system is similar to that in the linear trimers.

3.3. Origin of Anticooperativity of the PMF of Methane Trimers. The simplest explanation of the presence of anti-cooperative terms in the potentials of mean force of the linear trimers could be their identification with the cavity part of the PMF of the m_1-m_3 pair. The reason for this is that the presence of the m_2 molecule between m_1 and m_3 effectively prevents the water molecules from entering the space between them. Therefore, the dominant contribution to the PMF coming from the m_1-m_3 pair should be the direct Lennard-Jones interaction energy between these methane molecules, which cancels out

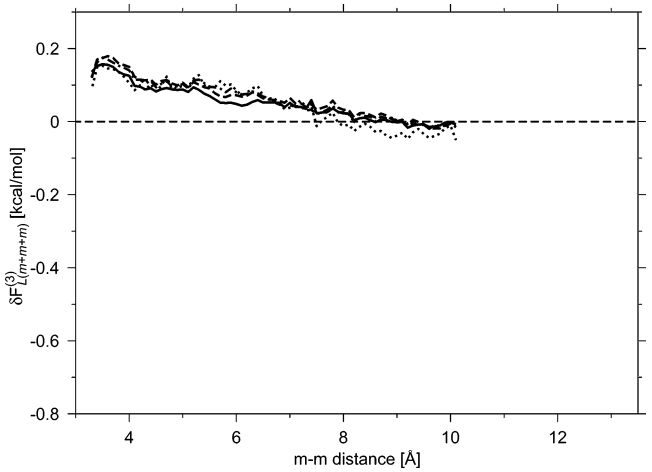

Figure 10. Curves for the cooperative term $\delta F^{(3)}$ for the L(m + m + m) system. The lines denote the duration of a run in each window in the parent PMF simulations.

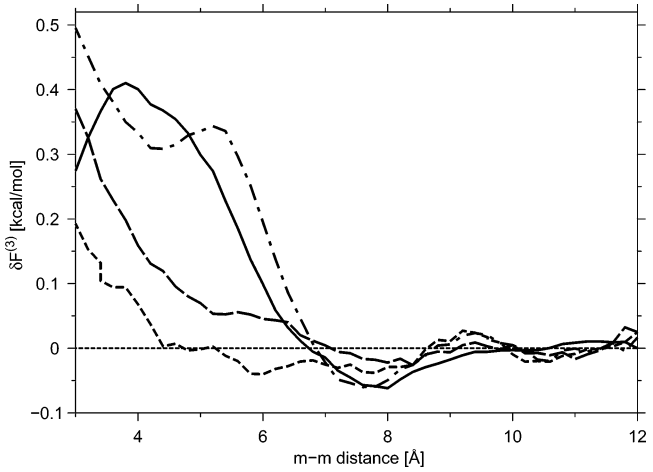

Figure 11. Curves for the cooperative term $\delta F^{(3)}$ for the L(m + m + m) (solid line), L(2m + m) (long-dashed line), 2m + m isosceles-triangle arrangement (short-dashed line), and m + m + m equilateral-triangle arrangement (dot−dashed line) systems computed from the information theory model of Hummer et al.[27]

when computing $\delta F^{(3)}$ from eqs 9 and 10, respectively. The putative cooperative contributions are therefore expressed by eqs 23 and 24, respectively,

$$\delta\tilde{F}^{(3)}_{\mathrm{L(2m+m)}}(\xi) \approx -\frac{W^{(2)}_{\mathrm{(m+m)}}(\xi + 3.9\ \text{Å}) - E^{\mathrm{LJ}}_{\mathrm{(m+m)}}(\xi + 3.9\ \text{Å})}{2} \tag{23}$$

$$\delta\tilde{F}^{(3)}_{\mathrm{L(m+m+m)}}(\xi) \approx -\frac{W^{(2)}_{\mathrm{m+m}}(2\xi) - E^{\mathrm{LJ}}_{\mathrm{m+m}}(2\xi)}{3} \tag{24}$$

where $E^{\mathrm{LJ}}_{\mathrm{m+m}}$ denotes the interaction energy between two methane molecules expressed by the Lennard-Jones potential function.

It can easily be demonstrated that $\delta\tilde{F}^{(3)}_{\mathrm{L(2m+m)}}$ and $\delta\tilde{F}^{(3)}_{\mathrm{L(m+m+m)}}$, in the range of ξ considered for the linear trimers, are proportional to the negative of the cavity part of the PMF (effectively the whole PMF) of two methane molecules from the solvent-separated minimum onward, which contrasts with the monotonic decrease of the cooperative term with distance found by simulations (Figures 9 and 10, respectively). Moreover, the values of $\delta\tilde{F}^{(3)}$ in the distance range considered are several times smaller than the *actual* values of $\delta F^{(3)}$ determined by simulations. Thus, $\delta F^{(3)}$ for linear trimers cannot be approximated by eq 23 or 24, respectively. The reason for this is that, while water molecules, in a large part, are eliminated from the space between m_1 and m_3 because of the presence of the m_2 molecule, there are still water molecules around the trimer; their rearrangement due to the presence of the nonpolar solute seems to constitute the dominant contribution to the PMF.

We also applied the information theory model of hydrophobic interactions of Hummer et al.[27,28] to compute the cooperative contribution to the PMF of the linear trimers (section 2.4). The results are shown in Figure 11. It can be seen that the distance dependence of $\delta F^{(3)}_{\mathrm{L(2m+m)}}$ calculated from the information theory of Hummer et al. is similar to that from simulations (Figure 9), and $\delta F^{(3)}_{\mathrm{L(m+m+m)}}$ is also quite similar except that there is a clear minimum at about 8 Å, which is not present in the curve obtained from simulations (Figure 10). This suggests that anticooperativity might in part be explained by the granularity of the solvent, which is present in the information theory of Hummer et al. through the pair correlation function of the solvent. It should be noted, however, that both curves (Figure 11) decrease with the distance faster than their counterparts obtained from simulations.

To determine whether the information theory of Hummer et al. can reproduce the cooperativity in the PMF of the 2m + m isosceles-triangle trimer and the m + m + m equilateral-triangle trimer studied in our earlier work,[23] we used it to compute the cooperative terms for these systems. It can be seen (Figure 11) that, for these systems, the information theory of Hummer et al. predicts largely the same distance dependence as for the linear trimers, which is in clear contradiction with the results of simulations, which show cooperativity at distances beyond the contact minimum (Figures 5 and 7B of ref 23, respectively). Therefore, the results obtained for linear trimers might be an artifact of the information theory of Hummer et al., which can, in turn, be connected with the fact that this theory has problems with the additivity of the calculated excess chemical potentials of hydrophobic interactions at large distances between hydrophobic solutes.[28]

We found that the most convincing explanation for the origin of anticooperativity in the PMF of the methane trimers studied in this work in water, which does not contradict the cooperativity observed for the 2m + m and m + m + m trimers,[23,24] comes from an analysis of the structure of water in the vicinity of the solute particles. Water−water interactions in hydrophobic association were studied for a pair of methane molecules by molecular-dynamics simulation.[69] The orientation distribution functions of the water dipole directions with respect to the methane−oxygen vector showed a high-order structure of water only in the first shell around the methane molecules. The dependence of the surface topography of hydrophobic hydration was reported for simulations of the melittin dimer.[70] It was observed that clathratelike structures dominate near convex surface patches, whereas the hydration shell near flat surfaces fluctuates between clathratelike and less-ordered or inverted structures.[70]

Paschek[45] studied the cylindrical distribution function of the center of mass of the SPCE water model[71] and the distribution function of the configurational contribution to the partial molar heat capacity of the SPCE water molecules around a pair of xenon particles at selected Xe−Xe distances.[45] He obtained the contribution to the partial molar heat capacity by a linear fit to the corresponding potential energies of the water molecules at five temperatures. The water molecules close to the bisector plane between the two nonpolar particles contributed to an increase of the partial molar heat capacity both at the contact

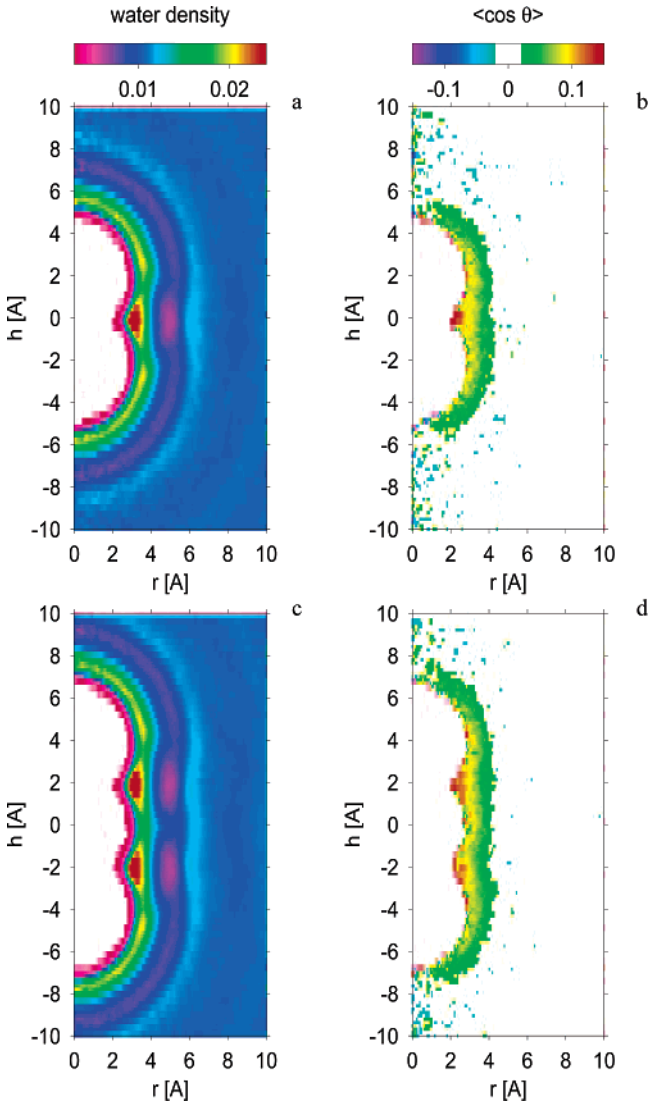

Figure 12. Distribution of the water density in the vicinity of (a) the methane dimer and (c) a linear trimer at the contact distance ($\xi = 3.9$ Å). Distribution of the cosine of the angle θ between the vector normal to the axis of (b) the methane dimer or (d) a linear trimer and the water dipole vector defined with the aid of the arbitrarily oriented water molecule w_1 of Figure 2. The red portions of panel d represent the values of $\langle\cos\theta\rangle$ for water molecules w_2 and w_3 of Figure 2. The color scale is shown above panel a (for density) and panel b (for $\langle\cos\theta\rangle$). Cylindrical coordinates h and r are defined in Figure 2.

distance and at the desolvation barrier distance, which explained the presence of a maximum in the partial molar heat capacity at the desolvation barrier.

We analyzed the water density and the average value of the cosine of the angle θ of the orientation of water molecules with respect to the axis perpendicular to the axis of the cluster (Figure 2) from our molecular dynamics simulations for these systems. The 2D maps of the water density around the m + m and the linear trimer systems at the contact distance are shown in Figure 12a and c, respectively. It can be seen that water accumulates close to the bisector plane ($h = 0$ in a and ± 2 in c) between the neighboring methane molecules (Figure 12) as in Paschek's study of the Xe−Xe pair. Therefore, for the linear trimers, a substantial contribution to the free energy of the system can be expected from the potential energy of interaction of the water molecules at one bisector plane with those at the second bisector plane, as outlined in Figure 2. If these water molecules were to exhibit similar orientations, the interaction energy would be positive (parallel dipole moments as shown in Figure 2). To check this, we computed the map for the average value of the cosine of the angle θ between the dipole-moment vector of a water molecule and the normal to the axis of the methane dimer or a linear trimer. The 2D plots of cos θ for the methane dimer and trimers at the contact distance are shown in Figure 12b and d, respectively. It can be seen that the water molecules are not oriented at random as in bulk water (this would correspond to cos $\theta = 0$) but the maximum value of cos θ is ~0.15 (red portion of Figure 12b and d); this leaves a small net component of the dipole moment pointing outward from the molecular surface in the region of the bisector planes.

Let us consider two water molecules w_2 and w_3 located as in Figure 2. The average electrostatic-interaction energy of these two molecules can be approximated as the average energy of interaction of their permanent dipole moments, as expressed by eq 24[72]

$$\langle E_{\mathrm{dip}}\rangle(\mathrm{kcal/mol}) = 332\left\langle \frac{\mathbf{p_1}\cdot\mathbf{p_2}}{r^3} - \frac{(\mathbf{p_1}\cdot\hat{\mathbf{r}})(\mathbf{p_2}\cdot\hat{\mathbf{r}})}{r^3}\right\rangle \approx \frac{332\langle\mathbf{p_1}\cdot\mathbf{p_2}\rangle}{r^3} = \frac{332(\langle\cos\theta\rangle p)^2}{r^3} \quad (25)$$

where $\mathbf{p_1}$ and $\mathbf{p_2}$ are the dipole-moment vectors of water molecules w_2 and w_3, respectively, p is the value of the dipole moment of a water molecule, $\hat{\mathbf{r}}$ is the unit vector pointing from molecule w_2 to molecule w_3, and r is the distance between the two interacting water molecules. In eq 25, we set the dielectric constant at 1 rather than using the dielectric constant of water because water molecules w_2 and w_3 are not likely to be separated by any other water molecule.

The average value of the $(\mathbf{p_1}\cdot\hat{\mathbf{r}})(\mathbf{p_2}\cdot\hat{\mathbf{r}})$ term is zero because the average cosine of the angle between the dipole moment of a water molecule in the first solvation shell of a methane dimer/trimer and the dimer/trimer axis in the vicinity of the intersection of the molecular surfaces of the methane molecules is only about ± 0.01, as opposed to the value of the projection on the perpendicular axis (cos $\theta \approx 0.15$), as found by analysis of the MD snapshots. Substituting the dipole moment of water of 1.8 D ≈ 0.21 electron charge unit × Å, $\langle\cos\theta\rangle = 0.15$, and $r = 4$ Å into eq 24, we obtain $\langle E_{\mathrm{dip}}\rangle \approx +0.05$ kcal/mol. This interaction energy makes a direct contribution to $\delta F^{(3)}$ because such interactions are not present for the isolated dimers. This explains the anticooperativity in the hydrophobic interactions of the L(2m + m) and L(m + m + m) trimers. Because up to four water molecules can be accommodated at the intersection of the molecular surfaces of the methane molecules, the total energy of interaction will approximately amount to about 0.2 kcal/mol (i.e., it is of the order of the magnitude of $\delta F^{(3)}$ determined from our simulations). This explanation for the anticooperativity in the hydrophobic interaction of the 2m + m system may be assumed to be the same.

4. Conclusions

We have shown that multibody contributions to the free energy of hydrophobic associations cannot be fully explained in terms of the change of the molecular surface area with solute geometry. In particular, anticooperativity in the hydrophobic interactions of linear methane trimers cannot be explained in terms of the change of the molecular surface area. Application of the information-theory model of hydrophobic interactions of Hummer et al.,[27] which incorporates the granularity of the solvent through inclusion of the solvent−solvent pair correlation function, correctly predicts anticooperativity for linear trimers

qualitatively; however, it also predicts anticooperativity for nonlinear trimers, in contrast to the results of simulations and molecular surface area models. Therefore, the theory of Hummer et al. does not provide a full explanation of multibody effects in hydrophobic interactions.

It appears that the change of the pattern of interactions between the water molecules that belong to the solvation shell makes a remarkable contribution to multibody effects in hydrophobic association because of additional partial local ordering of water in the solvation shell. This ordering results in a net positive energy of electrostatic interaction between those water molecules that are close to the intersections of the molecular surface of a methane molecule with its neighbors. We demonstrated this clearly for the linear methane trimers, where the multibody contribution to the molecular surface area is exactly equal to zero (cf. section 2.3), yet the three-body contributions to the free energy of hydrophobic associations amount to about 10% of the depth of the contact minimum and are anticooperative.

In conclusion, a simple mechanistic picture of hydrophobic association can be drawn. The free energy of hydrophobic association first of all depends on the difference in the number of water molecules in the solvation shell of a cluster and that in the solvation shells of its monomers; this can be approximated by the molecular surface area. However, when the matter is considered in more detail, the change of the water−water interaction pattern must be taken into account. Obviously, this pattern depends on the local partial orientation of water in the first solvation shell and can be expected to be anticooperative because of the same orientation of the dipole moments of the water molecules involved in the interaction.

Acknowledgment. This work was supported by grants from the National Science Foundation (MCB00-03722), the National Institutes of Health (GM-14312), the Fogarty Foundation (TW1064), and grant DS 8372-4-0138-4 from the Polish State Committee for Scientific Research (KBN). This research was conducted using the resources of (a) the National Science Foundation Terascale Computing System at the Pittsburgh Supercomputer Center, (b) the Informatics Center of the Metropolitan Academic Network (IC MAN) in Gdańsk, and (c) our own 40-processor Linux cluster at the Faculty of Chemistry, University of Gdańsk.

References and Notes

(1) Némethy, G.; Scheraga, H. A. *J. Chem. Phys.* **1962**, *36*, 3382−3400.
(2) Errington, J. R.; Debenedetti, P. G. *Nature* **2001**, *409*, 318−321.
(3) Griffith, J. H.; Scheraga, H. A. *J. Mol. Struct.: THEOCHEM* **2004**, *682*, 97−113.
(4) Kauzmann, W. *Adv. Protein Chem.* **1959**, *14*, 1−63.
(5) Némethy, G.; Scheraga, H. A. *J. Phys. Chem.* **1962**, *66*, 1773−1789. Erratum: *J. Phys. Chem.* **1963**, *67*, 2888.
(6) Scheraga, H. A.; Némethy, G.; Steinberg, I. Z. *J. Biol. Chem.* **1962**, *237*, 2506−2508.
(7) Poland, D. C.; Scheraga, H. A. *J. Phys. Chem.* **1965**, *69*, 2431−2442.
(8) Tanford, C. *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*; Wiley: New York, 1973.
(9) Ben-Naim, A. *Hydrophobic Interactions*; Plenum Press: New York, 1980.
(10) Ravishanker, G.; Beveridge, D. L. Theoretical Studies of the Hydrophobic Effect. In *Theoretical Chemistry of Biological Systems*; Náray-Szabó, G. Ed.; Elsevier: Amsterdam, 1986; Chapter 7, pp 423−494.
(11) Scheraga, H. A. *J. Biomol. Struct. Dyn.* **1998**, *16*, 447−460.
(12) Némethy, G.; Scheraga, H. A. *J. Chem. Phys.* **1962**, *36*, 3401−3417.
(13) Shimizu, S.; Chan, H. S. *J. Am. Chem. Soc.* **2001**, *123*, 2083−2084.
(14) Edsall, J. T. *J. Am. Chem. Soc.* **1935**, *57*, 1506−1507.
(15) Wicke, E. *Angew. Chem., Int. Ed. Engl.* **1966**, 5106−5122.
(16) Frank, H. S.; Evans, M. W. *J. Chem. Phys.* **1945**, *13*, 507−532.
(17) Némethy, G.; Scheraga, H. A. *J. Chem. Phys.* **1964**, *41*, 680−689.
(18) Frank, H. S.; Wen, W. Y. *Discuss. Faraday Soc.* **1957**, *24*, 133−140.
(19) Owicki, J. C.; Scheraga, H. A. *J. Am. Chem. Soc.* **1977**, *99*, 7413−7418.
(20) Kincaid, R. H.; Scheraga, H. A. *J. Comput. Chem.* **1982**, *3*, 525−547.
(21) Geiger, A.; Rahman, A.; Stillinger, F. H. *J. Chem. Phys.* **1979**, *70*, 263−276.
(22) Rapaport, D. C.; Scheraga, H. A. *J. Phys. Chem.* **1982**, *86*, 873−880.
(23) Czaplewski, C.; Rodziewicz-Motowidło, S.; Liwo, A.; Ripoll, D. R.; Wawak, R. J.; Scheraga, H. A. *Protein Sci.* **2000**, *9*, 1235−1245.
(24) Czaplewski, C.; Ripoll, D. R.; Rodziewicz-Motowidło, S.; Liwo, A.; Wawak, R. J.; Scheraga, H. A. *Int. J. Quantum Chem.* **2002**, *88*, 41−55.
(25) Czaplewski, C.; Rodziewicz-Motowidło, S.; Dabal, M.; Liwo, A.; Ripoll, D. R.; Scheraga, H. A. *Biophys. Chem.* **2003**, *105*, 339−359. Erratum: *Biophys. Chem.* **2004**, *111*, 267−271.
(26) Pratt, L. R.; Chandler, D. *J. Chem. Phys.* **1977**, *67*, 3683−3704.
(27) Hummer, G.; Garde, S.; Garcıa, A. E.; Pohorille, A.; Pratt, L. R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8951−8955.
(28) Hummer, G.; Garde, S.; Garcıa, A. E.; Paulaitis, M. E.; Pratt, L. R. *J. Phys. Chem. B* **1998**, *102*, 10469−10482.
(29) Hummer, G. *J. Am. Chem. Soc.* **1999**, *121*, 6299−6305.
(30) Blandamer, M. J.; Fox, M. F. Spectroscopic Properties. In *Water: A Comprehensive Treatise*; Franks, F. Ed.; Plenum: New York, 1973; Vol. 2, Chapter 8.
(31) Enderby, J. E.; Neilson, G. W. X-ray and Neutron Scattering by Aqueous Solutions of Electrolytes; In *Water: A Comprehensive Treatise*; Franks, F. Ed.; Plenum: New York, 1979; Vol. 6, Chapter 1.
(32) Hura, G.; Russo, D.; Glaeser, R. M.; Head-Gordon, T.; Krack, M.; Parrinello, M. *Phys. Chem. Chem. Phys.* **2003**, *5*, 1981−1991.
(33) Turner, J.; Soper, A. K. *J. Chem. Phys.* **1994**, *101*, 6116−6125.
(34) Swaminathan, S.; Harrison, S. W.; Beveridge, D. L. *J. Am. Chem. Soc.* **1978**, *100*, 5705−5712.
(35) Ravishanker, G.; Mehrotra, P. K.; Mezei, M.; Beveridge, D. L. *J. Am. Chem. Soc.* **1984**, *106*, 4102−4108.
(36) Southall, N. T.; Dill, K. A.; Haymet, A. D. J. *J. Phys. Chem. B* **2002**, *106*, 521−533.
(37) Swaminathan, S.; Beveridge, D. L. *J. Am. Chem. Soc.* **1979**, *101*, 5832−5833.
(38) Ravishanker, G.; Mezei, M.; Beveridge, D. L. *Faraday Symp. Chem. Soc.* **1982**, *17*, 79−91.
(39) Smith, D. E.; Haymet, A. D. J. *J. Chem. Phys*. **1993**, *98*, 6445−6454.
(40) van Belle, D.; Wodak, S. J. *J. Am. Chem. Soc*. **1993**, *115*, 647−652.
(41) Young, W. S.; Brooks, C. L., III. *J. Chem. Phys*. **1997**, *106*, 9265−9269.
(42) Lüdemann, S.; Schreiber, H.; Abseher, R.; Steinhauser, O. *J. Chem. Phys*. **1996**, *104*, 286−295.
(43) Shimizu, S.; Chan, H. S. *J. Chem. Phys.* **2000**, *113*, 4683−4700.
(44) Paschek, D. *J. Chem. Phys*. **2004**, *120*, 6674−6690.
(45) Paschek, D. *J. Chem. Phys*. **2004**, *120*, 10605−10617.
(46) Chau, P.-L.; Mancera, R. L. *Mol. Phys*. **1999**, *96*, 109−122.
(47) Ghosh, T.; Garcia, A. E.; Garde, S. *J. Am. Chem. Soc*. **2001**, *123*, 10997−11003.
(48) Ghosh, T.; Garcia, A. E.; Garde, S. *J. Chem. Phys*. **2002**, *116*, 2480−2486.
(49) Mancera, R. L. *Chem. Phys. Lett*. **1998**, *296*, 459−465.
(50) Rank, J. A.; Baker, D. *Protein Sci.* **1997**, *6*, 347−354.
(51) Shimizu, S.; Chan, H. S. *J. Chem. Phys.* **2001**, *115*, 1414−1421.
(52) Czaplewski, C.; Rodziewicz-Motowidło, S.; Liwo, A.; Ripoll, D. R.; Wawak, R. J.; Scheraga, H. A. *J. Chem. Phys*. **2002**, *116*, 2665−2667.
(53) Shimizu, S.; Chan, H. S. *J. Chem. Phys.* **2002**, *116*, 2668−2669.
(54) Shimizu, S.; Chan, H. S. *Proteins: Struct., Funct., Genet*. **2002**, *48*, 15−30.
(55) Ghosh, T.; Garcia, A. E.; Garde, S. *J. Phys. Chem. B* **2003**, *107*, 612−617.
(56) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys*. **1977**, *23*, 187−199.
(57) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic Press: San Diego, CA, 1996; Chapter 7, pp 176−181.
(58) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem*. **1992**, *13*, 1011−1021.
(59) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem*. **1995**, *16*, 1339−1350.
(60) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem*. **1992**, *13*, 952−962.

(61) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys*. **1983**, *79*, 926−935.

(62) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. *J. Chem. Phys.* **2004**, *120*, 9665−9678.

(63) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc*. **1984**, *106*, 6638−6646.

(64) Allen, M. P.; Tildesley, D. G. *Computer Simulation of Liquids*; Oxford University Press: New York, 1987, p 21.

(65) Berendsen, H. J. C.; Postma, J. P. M.; DiNola, A.; Haak, J. R. *J. Chem. Phys*. **1984**, *81*, 3684−3690.

(66) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun*. **1995**, *91*, 43−56.

(67) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51−55.

(68) Hummer, G.; Pratt, L. R.; Garcıa, A. E. *J. Phys. Chem*. **1996**, *100*, 1206−1215.

(69) Tsunekawa, N.; Miyagawa, H.; Kitamura, K.; Hiwatari, Y. *J. Chem. Phys*. **2002**, *116*, 6725−6730.

(70) Cheng, Y. K.; Rossky, P. J. *Nature* **1998**, *392*, 696−699.

(71) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem*. **1987**, *91*, 6269−6271.

(72) Hirschfelder, J. O.; Curtiss, C. F.; Bird, R. B. *Molecular Theory of Gases and Liquids*; Wiley & Sons: New York, 1954, p 851.