JOURNAL
OF THE AMERICAN CHEMICAL SOCIETY

VOLUME 101, NUMBER 8 APRIL 11, 1979

Solvation. A Molecular Dynamics Study of a Dipeptide in Water

Peter J. Rossky† and Martin Karplus*

Contribution from the Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138. *Received May 15, 1978*

Abstract: A molecular dynamics study of a dilute aqueous solution of an alanine dipeptide in the C_7^{eq} configuration has been carried out. The intermolecular potentials are given by a sum of pairwise additive Lennard-Jones and electrostatic terms; the peptide internal degrees of freedom are governed by an empirical energy function; and the water molecules are described by a modification of the ST2 model of Stillinger and Rahman which includes internal flexibility. The average local structure of the dipeptide and the dynamics of its internal structural fluctuations are affected very little by the presence of the solvent; the only evidence for rapidly damped dynamic correlations is found for the lowest frequency motions (dihedral angle torsions; $\omega \sim 50$ cm^{-1}) and for those involving the the lightest masses (methyl group libration; $\omega \sim 185$ cm^{-1}). Analysis of the kinetic properties of the solvent shows that the influence of the solute is limited to the first solvation layer. Although the polar (C═O, NH) groups of the dipeptide have little influence on the mobility of the solvent, that near the nonpolar methyl groups is substantially hindered in both its translational and rotational motion. Consideration of the bonding and geometric disposition of the solvent molecules shows that the mobility loss near nonpolar groups is due to the maintenance of bulk-like hydrogen bonding within the constraint of a reduced number of possible bonding neighbors. For solvent molecules near polar dipeptide groups, this constraint is not present since water–dipeptide as well as water–water hydrogen bonds can form. It is pointed out that significant contributions to the enthalpy of solution and the solute partial molar heat capacity can result from small changes in the energetics of solvent–solvent hydrogen bonding. Further, although clathrate-like geometric characteristics are found near nonpolar groups, a description implying solid-like attributes is very misleading.

I. Introduction

Aqueous solutions of nonelectrolytes display a variety of thermodynamic properties which are qualitatively different from those of comparable nonaqueous systems.[1] For solutes composed primarily of polar groups (i.e., those which can participate in hydrogen bonds with the water molecules) the excess free energy of solution is usually dominated by a negative enthalpy contribution. Nonpolar solutes (typically containing alkyl groups) can also have a negative excess enthalpy. However, the excess free energy of solution is dominated at ambient temperatures by a substantial negative entropy that increases in magnitude systematically with the size of the nonpolar group.[1,2] It is this entropic term which is responsible for the observed small solubilities of nonpolar species. For the archetype alkane, methane, the standard free energy of solution at 25 °C is[1] about 6 kcal/mol; the entropy of solution contributes 9 kcal/mol ($\Delta S°$ = −31.2 cal/mol·deg) and the enthalpic term about −3 kcal/mol. It is generally believed that the negative entropy contribution is associated primarily with changes in the aqueous solvent, rather than with configurational constraints on the solute.[1,3] For rare gas atoms,[1] there can be no loss of solute internal configurational freedom, but the situation is less clear for more complicated nonpolar solutes.[4]

Behavior parallel to that of the entropy has been noted for solution heat capacities.[1,3] The partial molar heat capacities of nonpolar solutes in water are anomalously large and tend to increase systematically with solute size. For the example of methane, the excess heat capacity, ΔC_p, is 55 cal/mol·deg at 25 °C.[1] For mixed-functional solutes containing both polar and nonpolar groups (e.g., alcohols, amines, ketones), the polar group is found to contribute a small increment to the thermodynamic properties of solution; this increment is essentially constant, independent of the nature of the nonpolar substituent,[1,3,5] suggesting that the effect of the polar group is relatively local.

The observed entropy loss has been interpreted in terms of an increase in solvent "order" or "structure".[3] Relatively little is known, however, at the molecular level, about the nature of the increase in water "structure" due to nonpolar groups, and the microscopic interpretation of this "structure" is still highly speculative.[1,3,6] Few data bearing directly on the alterations of water structure in aqueous solutions are available. NMR studies on solutions of small molecules provide evidence that the translational and rotational mobility of water molecules is significantly decreased by the presence of nonpolar solutes.[3,7,9] In particular, the average translational diffusion rate of water molecules in such solutions is reduced considerably more than would be expected from the obstructive effect of the relatively large size of the solute molecule. The rotational reorientation time of the solvent molecules in contact with the nonpolar solute is increased by approximately a factor of 2–3. Direct measurements using X-ray scattering[10] and neutron scattering[9] methods have shown some evidence for the existence of solvent structural changes in aqueous solutions, as compared with the pure solvent. However, these studies do not

† Department of Chemistry, State University of New York at Stony Brook, Stony Brook, N.Y. 11794.

0002-7863/79/1501-1913$01.00/0

provide sufficiently detailed information to permit determination of the molecular arrangement of the solvent molecules in the neighborhood of the solute. Such information is required for a complete understanding of the origins of the phenomena associated with the interaction of water with nonpolar species.

Of interest also are the effects of the aqueous solvent on the solute properties. Owing to the strong infrared absorbance of water, solute structural studies present special problems. With other methods (e.g., NMR,[11,12] depolarized Rayleigh scattering[13]) certain aspects of molecular conformation in solution have been investigated. Some information on the rates of rotational reorientation and translational diffusion has been obtained by NMR.[7,8] However, little is known about solvent effects on the internal motions of the solute, although related studies have been made of nonaqueous systems using laser Raman methods.[14]

One reason for the intense interest in the nature of solute-solvent interactions in aqueous solutions is that they are believed to be of primary importance in the determination of protein structure and stability.[15] Hydrophobic bonding,[3] the tendency for nonpolar species to cluster in aqueous solution, has been suggested as the source of stability for the native structure of proteins. Such clustering derives its driving force from the increase in entropy associated with partial exclusion of the aqueous solvent from contact with the nonpolar species. Further, the competition between intramolecular hydrogen bonding and solute-solvent hydrogen bonding can have important consequences in determining the relative stability of various protein conformations.[15] Measurements of rotational mobility of water molecules near protein surfaces have been interpreted by dividing the solvent molecules into three groups.[16] The most rapidly reorienting group, which includes the bulk solvent, has a characteristic rotational reorientation time (τ_r) of not more than about 10^{-11} s. The next most rapid exhibits a rotational reorientation time of about 10^{-9} s and has been tentatively identified as involving the water molecules that are strongly associated with ionic groups. The third exhibits a τ_r of about 10^{-6} s, and these solvent molecules are considered to be essentially irrotationally bound to the macromolecule. The population exhibiting the fastest times is expected to include molecules which form hydrogen bonds to the peptide backbone and those which are influenced by the presence of nonpolar groups. In X-ray studies of protein structures, it is found that electron density peaks due to water molecules can be distinctly observed only at the sites which bind water strongly (e.g., through ionic or strongly polar hydrogen bonds).[17] Because of the difficulties involved in studies of protein solutions per se, it is of particular interest to investigate the properties of systems of small molecules that incorporate functional groups present in proteins.

In a previous paper,[18] a model was presented for the study of an aqueous solution of a dipeptide and the results of a preliminary molecular dynamics simulation were described. The ST2 model for water, due to Stillinger and Rahman,[19] was used, and corresponding interactions between the polar peptide links of the dipeptide solute and the water molecules were introduced. Both the water molecules and the dipeptide were treated as completely flexible; that is, no internal degrees of freedom were constrained. The alanine dipeptide was chosen for study as one of the simplest systems of the peptide type that includes both polar and nonpolar substituents, as well as several relatively low frequency ("soft") internal degrees of freedom.

In this paper, we describe the results of a more extensive molecular dynamics simulation of the alanine dipeptide solution. We examine structural and dynamic aspects of both the solute and solvent. For the dipeptide, primary emphasis is placed on the internal motions. The size and dynamical character of fluctuations relative to the average structure are investigated both under vacuum and in the presence of solvent. The dipeptide vibrational degrees of freedom have frequencies varying from approximately 50 (dihedral angle torsions) to 3500 cm^{-1} (bond stretching), corresponding to characteristic times in the range of 7×10^{-13} to 1×10^{-14} s. For such a range in characteristic times, a significant variation in solvent effects (e.g., damping of fluctuations) is expected.

The structural and dynamic properties of the aqueous solvent in the region immediately surrounding the dipeptide solute are of special interest. The principal questions which we address follow. First, how is the dynamic behavior of the solvent altered by the proximity of the solute? Second, what is the range of influence of the solute; that is, are the effective solvent-solute interactions of sufficiently short range that it is reasonable to regard the water molecules in contact with the polar (peptide) groups as qualitatively different from those in contact with the nonpolar (methyl) substituents? Finally, we investigate the structural origins of observed differences in the dynamic properties of the solvent and relate these to previous discussions of aqueous solutions.

In section II, we review the model and simulation procedure used in this study. Section III presents the results concerned with the structure and dynamics of the dipeptide in solution, as compared to those in the absence of solvent. Aspects of the solvent dynamics are described in section IV. We consider the translational and rotational mobility of solvent molecules in the immediate proximity of the solute polar and nonpolar groups, and compare the calculated results with those for the corresponding properties of water molecules which are further removed from the solute. The structural origins of the dynamics are investigated in section V by examining the differences in bonding energetics and association geometries of the solvent molecules in the various regions of the system. The implications of the analysis are discussed in section VI, where we evaluate models of aqueous solutions in light of our results. The conclusions are presented in section VII.

II. Model and Simulation Procedure

The details of the model used to simulate the dipeptide solution have been presented previously;[18] a brief review of the interactions present in the system and the methods used to carry out the simulation is given here. The alanine "dipeptide" solute ($CH_3C'ONHCHCH_3C'ONHCH_3$), shown in Figure 1, is a neutral molecule terminated by methyl groups, rather than by the carboxylic acid and amino groups of an amino acid. The structure shown in Figure 1 is known as the equatorial C_7 conformation[20] (C_7^{eq}) owing to the seven-atom central ring structure (OC'NCC'NH) and the equatorial orientation, with respect to this ring, of the alanine side chain ($-CH_3$); according to conventional definitions,[21] the conformation corresponds to $(\phi, \psi) \simeq (-60°, 60°)$. Under vacuum, the global minimum in the dipeptide potential surface occurs in this configuration, and depolarized Rayleigh scattering, NMR, and Raman experiments have been interpreted as suggesting that the C_7^{eq} structure is also the most favored conformation in both aqueous[13] and nonaqueous solutions;[20] because of difficulties with the interpretation of the experimental data, the results obtained in aqueous solution are open to question. The primary source of the relatively high vacuum stability of the C_7^{eq} conformation is the internal hydrogen bond (indicated in Figure 1 by a dashed line) between the C=O and N—H groups of the two peptide links. The internal degrees of freedom of the dipeptide are governed by a molecular mechanics force field,[18,22,23] that includes terms corresponding to harmonic bonds, anharmonic bond angles, dihedral angle torsions, and nonbonded Lennard-Jones and electrostatic interactions. No dipeptide degrees of freedom are constrained in the simulation.

The water molecules are modeled by a modification of the

ST2 model of Stillinger and Rahman.[19] The model consists of four point charges placed within a single Lennard-Jones sphere centered at the oxygen atom; two positive charges are located at the hydrogen atom positions, and two negative charges are located at positions representing the lone-pair orbitals. The only modification made in the ST2 model is to allow internal flexibility in the water molecules. The OH bond lengths and HOH bond angle are not rigid but can vibrate in the presence of an intramolecular potential of the form

$$V_W = V_{HH}(|\mathbf{r}_{H_1H_2}|) + V_{OH}(|\mathbf{r}_{OH_1}|) + V_{OH}(|\mathbf{r}_{OH_2}|)$$

where the potentials V_{HH} and V_{OH} are taken from the central force model for water.[24] The positions of the two virtual (lone-pair) charges are constructed from the positions of the three atoms, H_1, O, and H_2, as described previously.[18]

The intermolecular interactions among the water molecules are computed exactly as in the ST2 model.[19] For two molecules, W_1 and W_2, we have

$$V_{W_1W_2} = 4\epsilon_W\left\{\left(\frac{\sigma_W}{r_{O_1O_2}}\right)^{12} - \left(\frac{\sigma_W}{r_{O_1O_2}}\right)^{6}\right\} + \sum_{i,j=1}^{4}\frac{q_i^{W_1}q_j^{W_2}}{r_{ij}}S(r_{O_1O_2}) \quad (1)$$

where σ_W and ϵ_W are the parameters characterizing the single Lennard-Jones interaction, q_i^W is the ith charge in water molecule W, $r_{O_1O_2}$ is the intermolecular oxygen-oxygen distance, and the switching function, $S(r)$

$$S(r) = \begin{cases} 0 & r < R_L = 2.016\ \text{Å} \\ \dfrac{(r-R_L)^2(3R_U - R_L - 2r)}{(R_U - R_L)^3} & R_L \le r \le R_U = 3.1287\ \text{Å} \\ 1 & r > R_U \end{cases}$$

diminishes the size of the electrostatic term at close contact distances.[18] The interactions between each water molecule and the dipeptide are given by a sum of Lennard-Jones and electrostatic terms of the form[18]

$$V_{WD} = \sum_{\substack{\text{dipeptide}\\ \text{atoms},\lambda}}\left[4\sqrt{\epsilon_W\epsilon_\lambda}\left\{\left(\frac{\bar{\sigma}_\lambda}{r_{O\lambda}}\right)^{12} - \left(\frac{\bar{\sigma}_\lambda}{r_{O\lambda}}\right)^{6}\right\} + \sum_{j=1}^{4}\frac{q_j^W q_\lambda}{r_{j\lambda}}\right] \quad (2)$$

where $\bar{\sigma}_\lambda = (\sigma_W + \sigma_\lambda)/2$ and $r_{O\lambda}$ is the water oxygen-dipeptide atom distance.

For the chosen values[18] of dipeptide-atom Lennard-Jones parameters, σ and ϵ, and charges, q, the water molecules associated with the solute peptide groups have reasonable energies and geometries. In particular, the optimal association energies (kcal/mol) for the four types of hydrogen bonds in the system in order of increasing strength (the water HOH bond angle is fixed at the tetrahedral angle) follow: ∠NH···H_2O (−6.0), ∠H_2O···H_2O (−6.8), ∠C=O···H_2O (−7.4), ∠N-H···O=C (−8.1). As discussed in the previous paper,[18] the water-amide potentials are chosen to agree with quantum-mechanical and experimental results; further, the values of the association energies are adjusted to the water-water interaction energy given by the ST2 model[19] so that the hydrogen bond strengths are all similar, in accord with available data. We point out that the best amide-amide bond attainable in the dipeptide (see Figure 1) is not of the optimal, near linear, geometry, but differs in energy from the linear bond by only a few tenths of a kcal/mol.

The ST2 model has been found to reproduce at least qualitatively a variety of structural and dynamic properties of pure water,[19] such as the radial pair correlation function and the translational diffusion constant. Although the agreement with experiment is not quantitative,[19] the essentials of the behavior of liquid water appear to be described correctly. Correspondingly, one expects that the present model will reproduce the qualitative features of the dipeptide solution dynamics and structure which underlie the experimental results discussed in the Introduction.

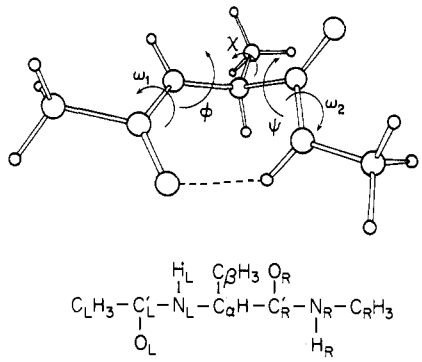


Figure 1. Alanine dipeptide in the equatorial C_7 conformation, $(\phi, \psi) \simeq (-60°, 60°)$. The structure is, left to right, $CH_3CONHCHCH_3CONHCH_3$; the dashed line represents the internal hydrogen bond.

The simulation is carried out on a sample consisting of one dipeptide and 195 water molecules in a cubic box with an edge length of 18.2194 Å; the density of 1.004 g/cm^3 is in accord with experiment.[25] In this system the dipeptide solute is surrounded by approximately two molecular layers of water at all points. To reduce boundary effects, we use periodic boundary conditions and the method of the minimum image,[26] with an intermolecular potential range cutoff of 8 Å. The numerical integration of the equations of motion is carried out using the Gear algorithm[27] with a time step of 3.67×10^{-16} s.

During the current simulation the dipeptide is in the neighborhood of a local minimum different from that which we studied previously.[18] An equilibrated system with the solute conformation of interest was obtained from that of our previous work by the temporary introduction of internal potentials to bring about the desired conformational change. Since the resulting disruption to the system's structure is rather small, the reequilibration in the new conformation is relatively rapid. We carried out approximately 2000 additional steps before generating the simulation which is analyzed here; monitoring various system properties during the equilibration, such as water-dipeptide interaction energy and association geometry, showed that the properties had stabilized by the end of the equilibration period. It was found necessary to artificially adjust the dipeptide and water temperatures during the equilibration period (by scaling of the corresponding velocities) in order to obtain a system in which the various temperatures were comparable. The simulation analyzed in the current work corresponds to 4000 steps, or 1.5 ps on a molecular time scale. The mean solvent kinetic temperature is 303 K (30 °C) and that of the dipeptide is 298 K (25 °C). The 4000-step simulation took approximately 5 h of computation time on an IBM Model 360/91 (about 4.5 s/step).

III. Solute Properties

In this section, we present results pertaining to the dipeptide mean structure and structural fluctuations. During the simulation, the solute remains in the vicinity of the C_7^{eq} minimum. This is not to be interpreted as implying that the C_7^{eq} is the most stable solution structure, because there is a very small probability of observing a large conformational change in such a short time. As a result of the limited sampling period, the average dynamic behavior of solute motions cannot be determined with high precision. We can, however, attempt to observe qualitative differences in the short-time behavior between

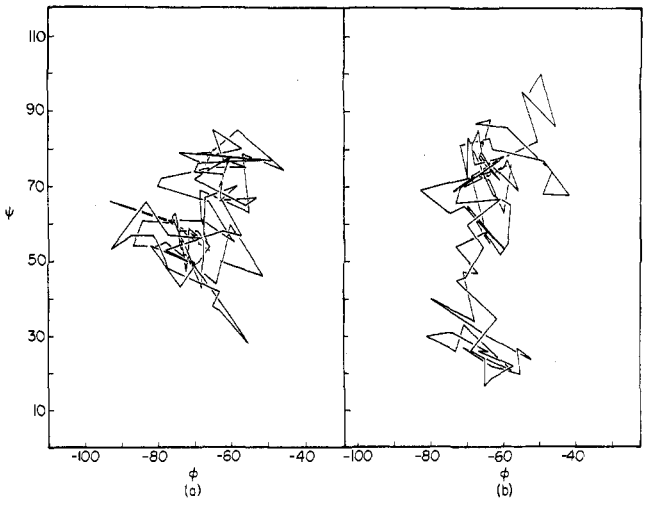


Figure 2. History of the dihedral angle pair, (ϕ, ψ) under vacuum (a) and solution (b). Consecutive points, separated by 50 steps in the simulation, are connected by straight lines.

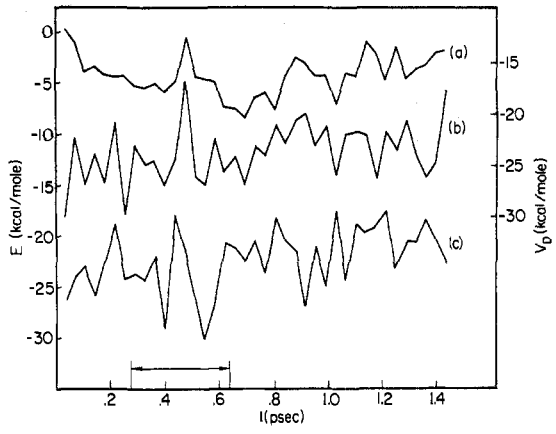


Figure 3. History of the total dipeptide energy (potential plus kinetic) (a); dipeptide potential energy (b); and dipeptide-solvent interaction energy (c). Curves (a) and (c) are referred to the scale on the left-hand side, (b) to the scale on the right-hand side. Values are given at intervals of 100 steps (0.036 ps). The time interval indicated on the t axis ($t \sim 0.3$–0.65 ps) corresponds to the interval during which (see Figure 2) $\psi < 40°$.

Table I. Average Solute Structure

A[a]	$\langle A \rangle$ vacuum	$\langle A \rangle$ solution	$\langle \Delta A^2 \rangle^{1/2}$ vacuum	$\langle \Delta A^2 \rangle^{1/2}$ solution
		Bonds[b]		
C_L'-O_L	1.235	1.237	0.023	0.028
N_L-H_L	0.994	0.997	0.018	0.012
C_α-C_β	1.544	1.542	0.041	0.035
N_R-C_R	1.461	1.459	0.034	0.032
		Bond Angles[b]		
C_L-C_L'-O_L	122.31	121.89	3.30	3.23
N_L-$C_\alpha C_R'$	114.73	114.46	4.25	3.93
N_L-C_α-C_β	107.71	108.15	3.51	3.85
C_R'-N_R-H_R	120.97	120.68	3.97	4.29
		Dihedral Angles[b,c]		
ϕ	−67.21	−63.96	9.67	7.83
ψ	63.45	59.33	11.53	22.57
ω_1	−179.25	−179.67	8.74	9.67
ω_2	−179.68	178.08	14.72	12.39
χ	−59.10	−62.88	9.96	32.24

[a] All structural parameters, A, are defined in Figure 1; $\langle A \rangle$ = mean value, $\langle \Delta A^2 \rangle^{1/2} = \langle (A - \langle A \rangle)^2 \rangle^{1/2}$. [b] Bonds in ångstroms, bond angles and dihedral angles in degrees. [c] The vacuum minimum occurs at $\phi = -66.2°$, $\psi = 65.3°$, $\omega_1 = 179.2°$, $\omega_2 = -179.9°$.

the dipeptide in solution and a corresponding simulation of the dipeptide dynamics in the absence of solvent.

As found in the previous study,[18] neither the average structure nor the magnitude of the local fluctuations of the dipeptide is strongly affected by the solvent environment. The results are summarized in Table I, where we give results for typical bonds, bond angles, and dihedral angles (see Figure 1). With the possible exception of the fluctuation in the dihedral angle ψ, the observed differences are within the statistical error of the calculation.

We see from the average value entries in Table I for the dihedral angle, χ, associated with the alanine methyl side chain rotation, that the motion observed is oscillatory within a single minimum; i.e., no net reorientation of this methyl group occurs. During the simulation, only one of the three methyl groups in the dipeptide reorients in solution (it is C_LH_3 in Figure 1) and none of the three reorients during the comparable time simulated under vacuum. The relatively low frequency of reorientation is consistent with ^{13}C spin relaxation measurements which indicate that the methyl group rotational reorientation time is about 5 ps.[28] Thus, the averages obtained are representative of a particular short-time behavior and do not correspond to the equilibrium (infinite time) average. This consideration applies equally to the time dependence of fluctuations in χ considered below.

In Figure 2, we show the (ϕ,ψ) configuration space traversed by the dipeptide under vacuum, (a), and in solution, (b), in terms of a history of the dihedral angle pair; consecutive points, separated by 50 steps (0.018 ps), are connected by straight lines. The results confirm the data in Table I; that is, the behavior of ϕ in solution and under vacuum is similar, but that of ψ is significantly different. Although the increased fluctuation of ψ may be accidental, it appears reasonable to attribute it to competitive hydrogen bonding of the solvent with the peptide polar groups, and the *effectively* weaker internal dipeptide hydrogen bond which results. Based on the *vacuum* dipeptide potential surface,[18] roughly 2 kcal/mol of *additional* energy is required to obtain the larger observed fluctuations in ψ in solution, as compared to vacuum. This estimate corresponds to the energy change associated with rotation about the dihedral angles ϕ and ψ, keeping the remainder of the internal structure rigid.

To examine further the energetics associated with the fluctuation in ψ, we show in Figure 3a a history (in increments of 0.036 ps) of the total (internal potential + kinetic) dipeptide energy, in Figure 3b the dipeptide potential energy, and in Figure 3c the dipeptide-solvent interaction energy; the latter is the sum of the interactions of all 195 water molecules with the dipeptide. The interval during which ψ is less than 40° (i.e., the region outside the range covered by the dipeptide under vacuum; see Figure 2) is indicated on the horizontal axis. It is clear from the figure that all quantities fluctuate widely during the simulation. The fluctuations in the potential energies (Figures 3b and 3c) are due in part to rapid interconversion of potential and kinetic energy. We see that the hydration energy (Figure 3c) fluctuates by considerably more than 2 kcal/mol during the time interval associated with the largest fluctuations in ψ. Further, the hydration energy attains its most negative value during this time interval. The latter observation suggests that improved solvent-solute association is responsible for the conformational fluctuation. However, examination of particular configurations in the simulation does not reveal a significant rearrangement of the water molecule-dipeptide association, in that the same water molecules remain associated with the dipeptide during this time. Hence, the negative de-

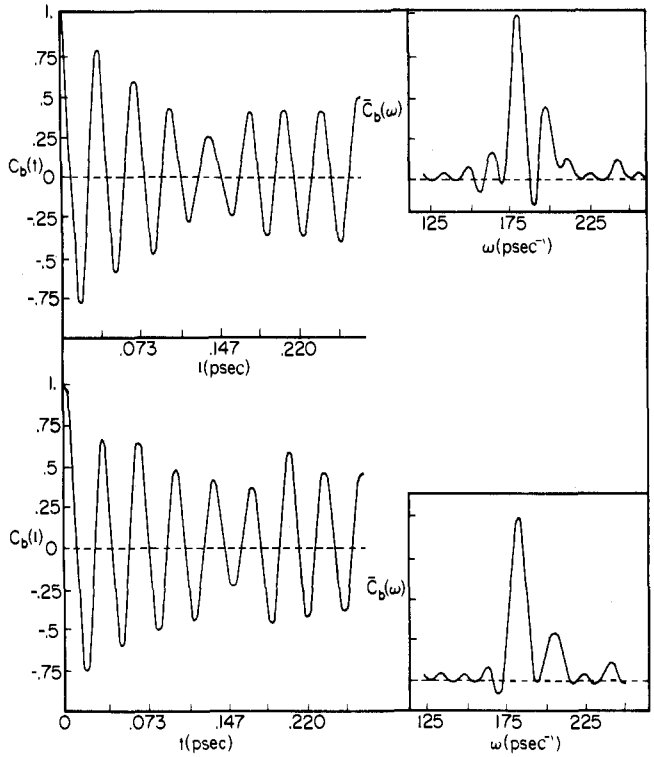


Figure 4. Time correlation function and spectral density for fluctuations in the bond length, C_α-C_β (Figure 1); top, in solution; bottom, under vacuum.

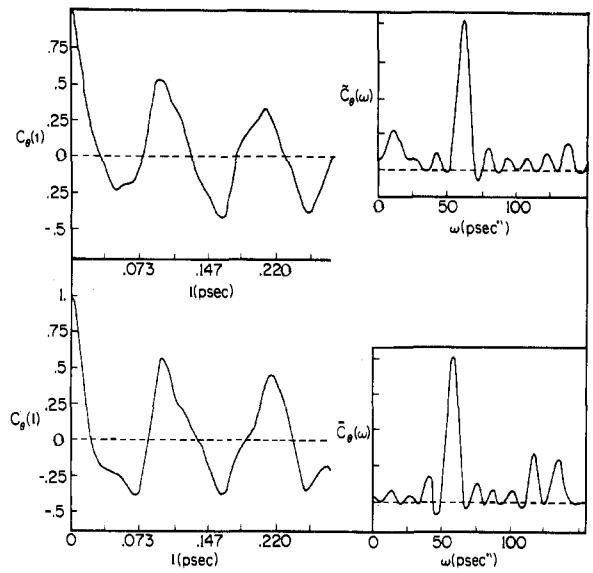


Figure 5. Time correlation function and spectral density for fluctuations in the bond angle, N_L-C_α-C_R' (Figure 1); top, in solution; bottom, under vacuum.

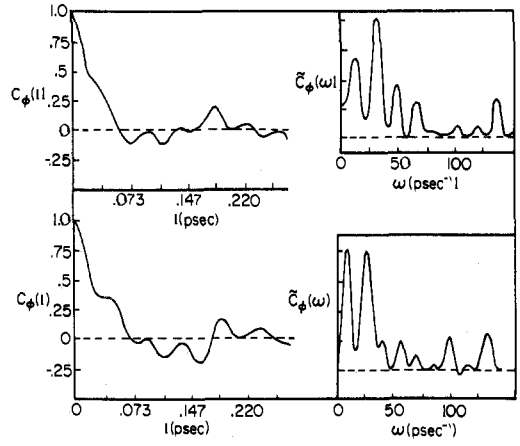


Figure 6. Time correlation function and spectral density for fluctuations in the dihedral angle, ϕ (Figure 1); top, in solution; bottom, under vacuum.

viations in the hydration energy that are associated with the structural fluctuation are most likely the result of changes in the orientation of the surrounding solvent. The variation in the hydration energy cannot be interpreted in terms of a simple description in which a solvent molecule is considered to be either bonded or not bonded to the solute.

Correlation Functions. We next consider dynamic correlations of the solute fluctuations. The time correlation function[29] for solute structural fluctuations is defined as

$$C_A(t) = \frac{\langle \Delta A(\tau)\Delta A\,(t+\tau)\rangle_\tau}{\langle \Delta A^2\rangle} \equiv \frac{\langle \Delta A(0)\Delta A(t)\rangle}{\langle \Delta A^2\rangle} \quad (3)$$

where A is a particular structural parameter (e.g., bond length, dihedral angle), $\Delta A(t) = A(t) - \langle A\rangle$, $\langle \Delta A^2\rangle = \langle (A - \langle A\rangle)^2\rangle$, and the brackets indicate an average over the simulation; in the numerator of eq 3, the average includes all values, τ, in the simulation.

The unperturbed harmonic vibration of an isolated bond of length b would lead to a correlation function, $C_b(t)$, which oscillates without decay for all times. However, shifts in either the frequency or phase of the oscillation results in an eventual decay of $C_b(t)$ to zero after a time when the phase of $\Delta b(t)$ is, on the average, completely random with respect to that of $\Delta b(0)$. Even in an isolated molecule, such a decay can occur due to coupling of the motion of various degrees of freedom with different frequencies, although over long times the correlation function cannot remain zero since there is no dissipation. The change in the correlation function in solution is determined by the effectiveness of solvent-solute collisions in dissipating the solute's dynamical information. In water, these "collisions" can involve the repulsive forces characteristic of hard sphere-like systems[30] and the strong hydrogen bonding forces; weak attractive van der Waals interactions are expected to have only a small effect. Collisions with solvent are more likely to affect the solute motion if the latter is associated with a small characteristic force constant or if the mass of the solute structural component involved is small; also, damping is generally more effective for motions involving structural components of increased spatial dimensions.[31,32] In the current study, a comparison of different dipeptide structural motions shows the expected qualitative differences in behavior.

We illustrate the results by presenting the correlation functions for (see Figure 1) a typical bond (C_α-C_β) in Figure 4, a typical bond angle (N_L-C_α-C_R') in Figure 5, and a number of dihedral angles, ϕ in Figure 6, ψ in Figure 7, ω_2 in Figure 8, and χ in Figure 9. The correlation functions are truncated at a time, t_{max}, equivalent to 750 simulation steps since statistical accuracy decreases with increasing t. Because of this statistical problem, only the qualitative behavior of the functions can be interpreted with confidence. In each figure, we show the result obtained in solution at the top and that obtained under vacuum at the bottom. The spectral density corresponding to $C_A(t)$ is

$$\tilde{C}_A(\omega) = \int_0^\infty dt \cos(\omega t) C_A(t)$$

Here, the limited knowledge of $C_A(t)$ forces us to truncate the time integral at t_{max}, rather than at infinite time; the calculated spectral density function is shown in an inset in each case (the

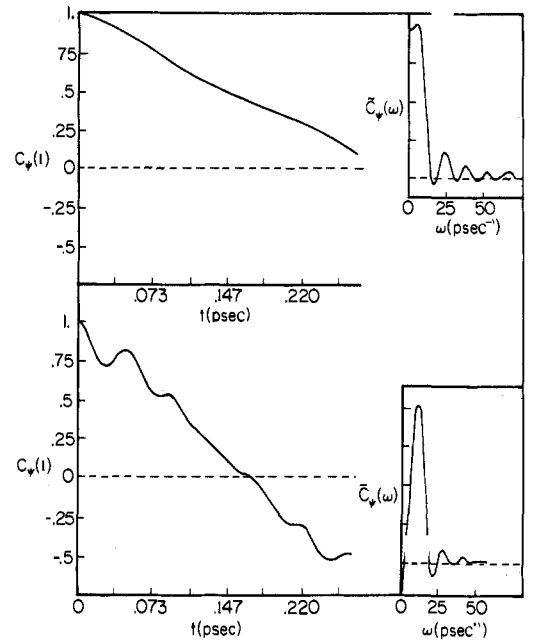


Figure 7. Time correlation function and spectral density for fluctuations in the dihedral angle, ψ (Figure 1); top, in solution; bottom, under vacuum.

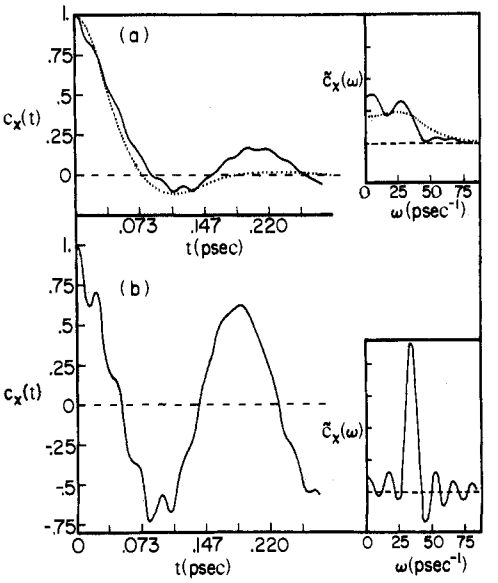


Figure 9. Time correlation function and spectral density for fluctuations in the dihedral angle, χ (Figure 1); top, in solution; bottom, under vacuum. The functions shown dotted at the top are obtained from a Langevin equation (see text).

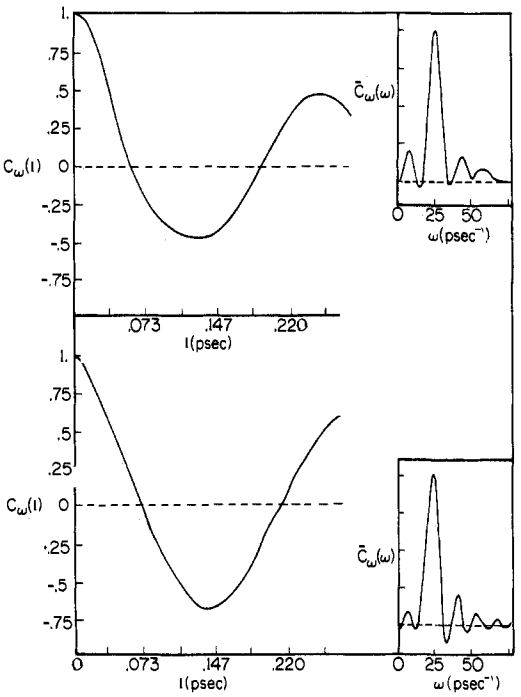


Figure 8. Time correlation function and spectral density for fluctuations in the dihedral angle, ω_2 (Figure 1); top, in solution; bottom, under vacuum.

amplitudes are in arbitrary units). Negative values of $\tilde{C}_A(\omega)$ result from the finite upper limit on the integration.

On the picosecond time scale considered, no significant damping is seen in the oscillatory correlation functions describing the high-frequency ($\omega \gtrsim 300$ cm^{-1}) bond-length stretching and bond-angle bending modes. A similar lack of damping is found for the relatively fast internal torsional oscillation of the peptide group defined by the angle ω_2 (Figure 1) (vacuum frequency $\simeq 120$ cm^{-1}). It is clear that, for these high-frequency motions, the behavior of the time correlation functions is very similar in solution as compared to vacuum, and that in both environments there is no evidence for a significant zero-frequency component in the spectral density. The latter is expected if the correlation function contains a decaying component. The only evidence of damping in the correlation functions is found for the motions associated with either low-frequency oscillation (ϕ and ψ, $\omega \simeq 50$ cm^{-1}) or small moments of inertia (methyl group libration, $\omega \simeq 185$ cm^{-1}). In each of these latter cases, a significant zero-frequency component is apparent in the spectral density for the dipeptide in solution. However, except at the lowest frequencies, the spectral densities of the dipeptide motions in the two environments are still rather similar.

Since the dipeptide structural component involved in the torsional angle χ is well defined (namely, the alanine methyl group) and the vacuum motion involves principally a single frequency (see Figure 9), we consider this motion in more detail. During the current simulation, the motion involves only libration and not overall reorientation of the methyl group. By comparison with the result in the absence of solvent, it can be seen that the solvent is effective in damping the oscillatory motion of the methyl group. The behavior is manifest by the appearance of a low-frequency component in the spectral density, $\tilde{C}_\chi(\omega)$. The short-time behavior of the solution correlation function ($t \lesssim 0.1$ ps) is roughly consistent with underdamped motion calculated from a Langevin equation:[31]

$$I\frac{d^2\chi}{dt^2} + I\omega_0^2\chi + f\frac{d\chi}{dt} = F_R(t)$$

where I is the moment of inertia of the methyl group (3 amu Å^2), ω_0 is the harmonic vacuum frequency (35 ps^{-1}; see Figure 9), f is a frictional coefficient, and $F_R(t)$ is a white noise random force. The time correlation function, $C_\chi^L(t)$, corresponding to motion governed by the Langevin equation is[31]

$$C_\chi^L(t) = e^{-\beta t/2}\left\{\cosh\left(\frac{\beta_1 t}{2}\right) + \frac{\beta}{\beta_1}\sinh\left(\frac{\beta_1 t}{2}\right)\right\}$$

where

$$\beta = f/I$$

and

$$\beta_1 = \{\beta^2 - 4\omega_0^2\}^{1/2}$$

Table II. Contributions to Solvent-Solute Hydrogen Bonding[a]

$\epsilon_{W\text{-}D}$, kcal/mol[b]	no.[c]	$\epsilon_{W\text{-}D}$, kcal/mol[b]	no.[c]
$-\infty < \leq -6.0$	0.125	$-\infty < \leq -6.0$	0.125
$-6.0 < \leq -5.0$	0.635	$-\infty < \leq -5.0$	0.760
$-5.0 < \leq -4.0$	1.058	$-\infty < \leq -4.0$	1.818
$-4.0 < \leq -3.0$	1.220	$-\infty < \leq -3.0$	3.037
$-3.0 < \leq -2.0$	2.280	$-\infty < \leq -2.0$	5.315

[a] Calculated from Figure 10. [b] Dipeptide-water molecule interaction energy. [c] Average number of water molecules interacting with the solute within the specified range of energy.

The function C_χ (*, ..nown as a dotted line at the top in Figure 9, is in reasonable agreement for short times with the correlation function $C_\chi(t)$ found in solution, if the characteristic time, $2/\beta = 2I/f$, is chosen as 0.05 ps; the corresponding spectral density is shown dotted in the inset. It is clear that for times greater than about 0.15 ps the correlation functions are not in agreement. This discrepancy is reflected in the spectral density; the Langevin equation predicts only a single peak in $\tilde{C}_\chi(\omega)$ ($\omega \sim 25$ ps^{-1}) and not two. It is not clear whether this disagreement results from insufficiency of the Langevin model or statistical inaccuracy in the evaluation of $C_\chi(t)$ from the simulation. We note that the initial decay of $C_\chi(t)$ corresponds to an apparent solvent drag that is much smaller than hydrodynamic estimates that assume stick boundary conditions. For example, treating the methyl group as a sphere of radius a = 2.5 Å (the van der Waals radius) to obtain the frictional coefficient f[33]

$$f = 8\pi\eta a^3$$

(η is the shear viscosity, 0.01 P), we find $2I/f$ equal to 0.0003 ps, i.e., about 150 times smaller than that observed. In this sense, the observed drag is nearer to the hydrodynamic slip boundary condition limit; the exact slip limit for a sphere corresponds to $f = 0$ and an infinite relaxation time. The relatively long relaxation time is consistent with the results of experimental studies of the rotational motion of small nonassociated molecules.[34]

IV. Structure and Dynamics of the Solvent

The structural and dynamic properties of the water molecules in the dipeptide solution are described in this section. We focus attention on the water molecules which are in the immediate vicinity of the solute, the so-called solvation "shell". We describe first some general aspects of solvation structure in terms of dipeptide-water hydrogen bond association and average solvent spatial distribution. This information is used to define solvation regions around the solute. It is found that a division of the solvent molecules into groups according to whether they are near solute polar groups, nonpolar groups, or outside of the first solvation layer provides an effective approach for evaluating the solute influence on solvent behavior.

A detailed analysis of the solvent dynamics is carried out in terms of these three regions. We are able to compare the behavior of the various water molecules in the first solvation layer with that of molecules farther from the solute. The observed differences in the dynamics are analyzed in section V. There the hydrogen bonding of molecules in each region is considered and the importance of energetics vs. structure in determining the solvent behavior is examined.

A. Nature of the Solvation Shell. To define the solvation regions, we examine the average solute-solvent interaction energy and the spatial distribution of solvent molecules around the dipeptide.

Solvation Structure. In Figure 10, we show the calculated distribution function, $N(\epsilon)$, which gives the number of water-dipeptide pairs with interaction energy in the range ϵ to $\epsilon + d\epsilon$; the energy ϵ is computed by summing over all atoms in the solute for each water molecule. The area under the curve in Figure 10 is equal to 195, the total number of water molecules in the system. The contributions to $N(\epsilon)$ at the most negative energies correspond to water molecules that are hydrogen bonded to the solute, while those at the highest energies correspond to molecules in close, repulsive contact with the dipeptide. Such a distribution is expected since thermal motion in the system often leads to configurations with molecular interactions that are far from the minimum in the potential energy. In Table II, we give integrated contributions to the distribution in Figure 10. As can be seen from the values in the table and from the figure, the contributions to $N(\epsilon)$ begin to increase more rapidly as ϵ becomes less negative; this is a reflection of the fact that an increasing number of water molecules at larger distances from the dipeptide are included in the distribution as the specified interaction energy becomes less negative. A large number, which are typically not in direct contact with the solute, have nearly zero interaction energy; this corresponds to the main peak centered around $\epsilon = 0$.

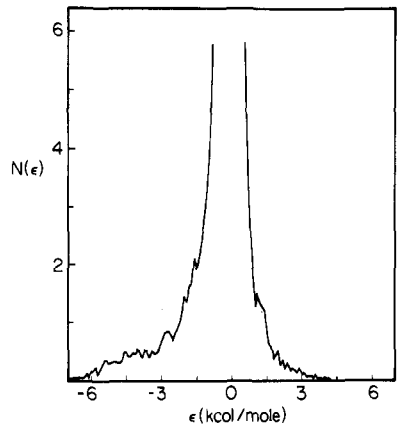


Figure 10. Relative distribution of water-dipeptide interaction energies, including all water molecules; the distribution is normalized so that the total integral of $N(\epsilon)$ is 195.

For the purposes of analysis, it is useful to define the difference between hydrogen-bonded molecular pairs and unbonded pairs according to an interaction energy criterion. Since the distribution of energies is continuous, this definition must be somewhat arbitrary. A primary consideration in the choice of criterion is to include only the molecular pairs which are geometrically disposed in a manner satisfactory for hydrogen bonding;[35-37] that is, we do not want to make the criterion so weak that, for example, second nearest neighbor molecules would be classified as hydrogen bonded. The distribution function shown in Figure 10 suggests a division near -3 kcal/mol, where the contributions to $N(\epsilon)$ begin to rise relatively more steeply. The criterion of -3 kcal/mol turns out to be equally appropriate for the analysis of water-water bonding, as will be seen below.

Examination of the water-dipeptide interaction potential[18] shows that an interaction energy of -3 kcal/mol corresponds to optimal association of a water molecule with a carbonyl group at an oxygen-oxygen separation of 4 Å. All water molecules with an interaction energy more negative than -3 kcal/mol are separated by a smaller distance; the energy minimum of -7.4 kcal/mol occurs at 2.60 Å. For comparison, the most frequent second neighbor distance in liquid water is[19] about 4.6 Å, for which the optimum water-water interaction energy is about -2 kcal/mol for the rigid ST2 geometry.[38] These considerations indicate that the energy criterion of -3 kcal/mol for a hydrogen bond is satisfactory on both energetic and geometric grounds.

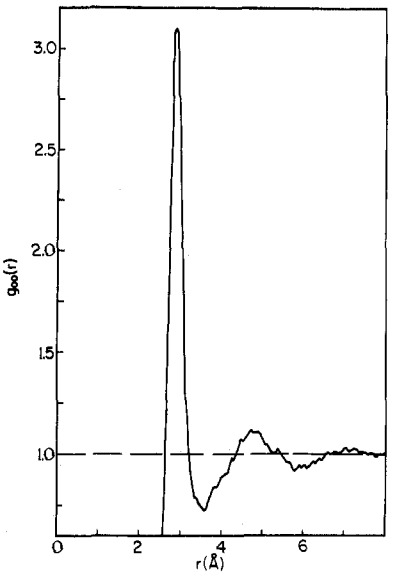


Figure 11. Water oxygen-oxygen pair correlation function computed including all water molecular pairs.

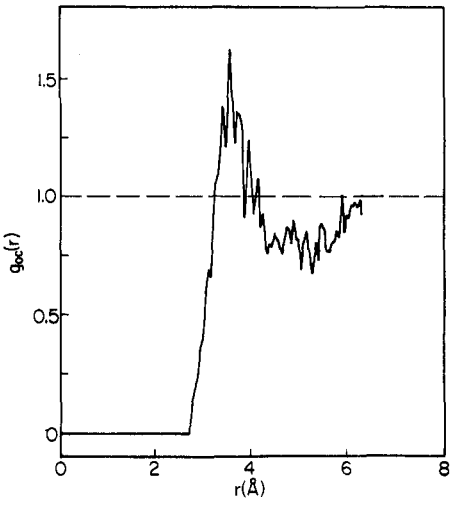


Figure 12. Water oxygen-methyl group carbon pair distribution (eq 4), averaged over the three solute methyl groups.

With this energy criterion, an average of 3.04 water molecules are hydrogen bonded to the dipeptide (see Table II). One of these molecules is associated with the amide NH external to the ring (left in Figure 1), and two are associated with the carbonyl group which is involved in the internal hydrogen bond. Two additional solvent molecules are more weakly associated with the other carbonyl group; their average energies are slightly above −3 kcal/mol. No water molecules are strongly associated with the NH group involved in the internal bond. The shielding of this amide proton from strong solvent bonding results from the requirement of a relatively linear NH···H_2O bond;[18] when the internal dipeptide hydrogen bond is formed, it is not possible for a water molecule to simultaneously form a satisfactory bond to the amide NH. NMR methods have been employed to probe the exposure of amide protons to solvent. A recent study makes use of line broadening by nitroxyl radicals to determine which amide protons are sequestered from the solvent.[11] Since the above result suggests that in the C_7^{eq} conformation one amide proton is not exposed to the solvent, such a measurement might be a useful test of the alanine dipeptide solution structure.

We next consider the average spatial distribution of solvent molecules around the solute. In general, the distribution functions which describe the disposition of solvent depend on the coordinates defining the dipeptide conformation, as well as the coordinates of the water molecules.[39] Functions of many coordinates, however, are inherently difficult to determine for statistical reasons,[40] and are, as well, often difficult to interpret. For the purpose of defining solvation regions, it is sufficient to examine atomic pair distributions. We consider the distribution of solvent atoms of type W around a particular solute (or solvent) atom of type A. Given $N_{WA}(r)$, the average number of W atoms within a sphere of radius r around atom A, we can define the probability distribution function, $g_{WA}(r)$ by

$$g_{WA}(r) = \frac{1}{4\pi\rho_W r^2}\frac{dN_{WA}(r)}{dr} \qquad (4)$$

where ρ_W is the density of W atoms in the bulk fluid. With the factor $(4\pi\rho_W r^2)^{-1}$, $g_{WA}(r)$ is normalized to one at positions in the bulk fluid far from the given atom, A. In a pure fluid, the distribution function of eq 4 is the conventional radial pair correlation function,[39] $g(r)$.

In Figure 11, we show the water molecule oxygen-oxygen pair distribution function, $g_{OO}(r)$, obtained from the current simulation including *all* solvent pairs. The result is the same as that obtained from simulation of bulk water,[18,19] within the statistical accuracy of the calculation. We note that the function $g_{OO}(r)$ is characterized by narrow peaks and troughs, a result of the hydrogen-bonded structure.[19,41] The first peak occurs at 2.85 Å corresponding to the energy minimum of the O—H···O hydrogen bond. The average distribution of water oxygen atoms around the methyl group carbon atoms, $g_{OC}(r)$, is shown in Figure 12; the result is the average over the three solute methyl groups. In contrast to Figure 11, the first peak is broad. The center of the peak occurs at about 3.7 Å, comparable to the average model water molecule-methyl group van der Waals contact distance of about 3.8 Å. Since water molecules can make contact with the methyl groups only within a restricted solid angle around the carbon atom (owing to the presence of the remainder of the solute attached to the methyl group), the height of the first peak in $g_{OC}(r)$ is reduced relative to the value that would be obtained if the group were completely exposed to solvent (e.g., as is a methane molecule in solution). The breadth and radial position of the first peak are comparable to that found in studies on argon-like systems;[42] that is, the first peak occurs at nearly the van der Waals contact distance and is relatively broader than that in $g_{OO}(r)$ for water, particularly on the large r side of the peak. in argon, the liquid structure is determined by the repulsive core of the Lennard-Jones spheres, rather than by the strong attractive hydrogen bond forces characterizing pure water.[30]

It has been suggested that the average density of the water surrounding nonpolar groups differs from that in the bulk fluid as a result of solvent "structuring".[40] To examine this aspect of solvent behavior requires accurate values for $g_{OC}(r)$, from which the average density profile at the methyl surface can be obtained. In the present simulation the first peak is significantly altered by the presence of the remainder of the solute (see above). One possible way of correcting for this effect might be the evaluation of $g_{OC}(r)$ by inclusion of only the solvent oxygen atoms that fall within the solid angle subtended by the tetrahedral face defined by the three protons of each methyl group (the appropriate solid angle is $(4\pi/4) = \pi$). The distribution function thus determined (and normalized to unity at $r \rightarrow \infty$) would be comparable to that for methane in water. Very recent Monte Carlo results for methane in water[43] indicate that the first peak in $g_{OC}(r)$ is broader than that found here, but this may well be a reflection of the very low density in that study (24% less than the experimental value).

The calculated average distribution of water molecules

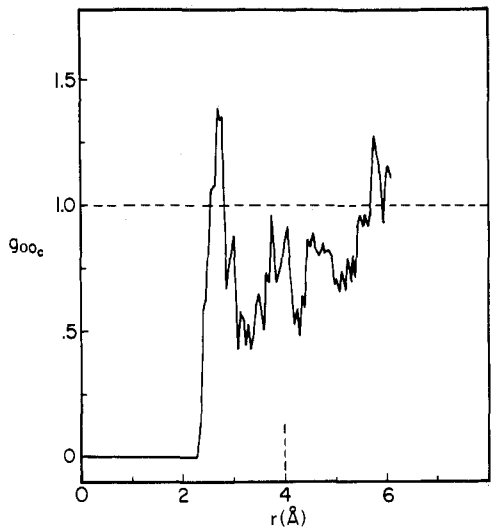


Figure 13. Water oxygen-carbonyl oxygen pair distribution (eq 4), averaged over the two solute oxygen atoms.

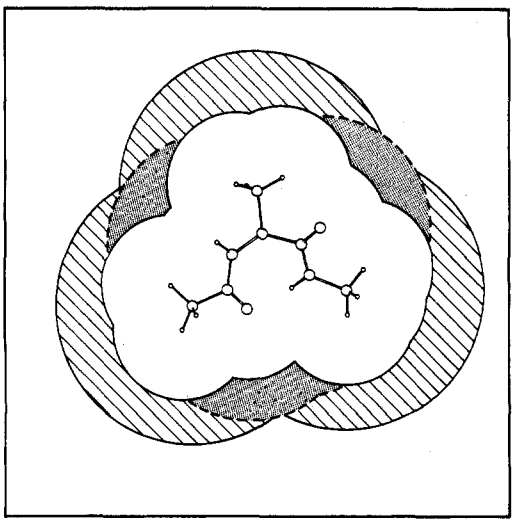


Figure 14. Schematic representation of the solvation regions defined in the text: "polar", dotted; "nonpolar", cross-hatched; "bulk", within outer square and outside shaded regions.

around a carbonyl oxygen (O_C), averaged over the two carbonyl groups in the dipeptide, is shown in Figure 13. The same geometric effects on the peak height (due to the remainder of the solute), mentioned above for $g_{OC}(r)$, apply here. We see that the first peak occurs at approximately the optimum hydrogen bond distance of 2.6 Å and is distinctly narrower than that found in the region surrounding the methyl groups; it corresponds more closely to $g_{OO}(r)$ for bulk water (Figure 11). The first peak is due primarily to the water molecules that are strongly associated with the carbonyl group. As in water,[41] the number of close neighbors is greater than the number of bonded neighbors; we find by integrating $g_{OO_C}(r)$ to 3.5 Å that 3.14 water molecules are on the average within 3.5 Å of each carbonyl oxygen, whereas (see above) only two molecules are hydrogen bonded (or nearly bonded) according to the energetic criterion.

The lack of smoothness apparent in $g_{OO_C}(r)$ for $r > 3$ Å is only in part due to statistical noise; some of this structure arises from the occurrence of the remaining dipeptide atoms at nearly fixed positions with respect to the carbonyl oxygen atoms and the solvent exclusion at these positions. The existence of this type of solvent structure can be attributed to molecular packing considerations in analogy to the behavior of a simple dense liquid.[30] A corresponding effect is not clearly manifest for the solvent in the vicinity of the methyl groups, since these groups protrude farther into the surrounding solvent and so are comparatively less shielded by the rest of the solute.

Decomposition of the Solvent. To progress in the analysis of solvation, it is necessary to consider explicitly the inequivalence of the surrounding solvent molecules. To this end, we divide the solvent molecules in accord with their average position with respect to the functional groups of the dipeptide solute; the approach is analogous to that in a pilot study of this system.[18] By this division, we obtain the advantage that we can investigate separately the dynamic and structural properties of solvent molecules which are more strongly influenced by one structural component of the solute than by another. At the same time, we retain the statistical advantages associated with the ability to average the observed behavior of a number of individual molecules. The division made in the current study is shown schematically in Figure 14. The central blank area immediately surrounding the solute corresponds to the region from which the centers (i.e., the oxygen atoms) of the solvent molecules are excluded by the solute. The outer square corresponds to the 18.2194-Å box which encloses the centers of all 195 water molecules. The intervening region, which contains all the solvent molecules, is divided into three subregions. The first is defined by spherical surfaces of radius 4 Å around each of the solute amide hydrogen and carbonyl oxygen atoms, as indicated by the dotted areas in Figure 14. The choice of a 4-Å radius is based on the shape of $g_{OO_C}(r)$ (Figure 13); 4 Å corresponds roughly to the center of the broad minimum in this distribution. Also, the choice made earlier for the minimum strength of a water-dipeptide hydrogen bond energy (−3 kcal/mol) corresponds to a maximum water-polar group separation of approximately 4 Å. This definition of radius is, therefore, the smallest which guarantees the enclosure of the centers of all molecules which are, by definition, hydrogen bonded to the dipeptide. The second region is defined by spherical surfaces of radius 5 Å around each methyl carbon atom, indicated in Figure 14 by the cross-hatched area. This distance encompasses the first solvation layer defined by the minimum in the radial distribution shown in Figure 12. We exclude from this second region those areas already assigned to the polar group. Thus, the first solvation layer consists of the sum of the water molecules allocated to the polar and nonpolar groups; that is, the spatial volume enclosed by the *defined* first shell corresponds to the region *observed* in the simulation to encompass all molecules contributing to the first peaks in the radial distribution functions. The remaining volume available to the solvent includes the second molecular layer of water, and, owing to the limited size of the sample, essentially nothing else.

Individual water molecules are classified into groups according to the *average* distance between their oxygen atom and each of the amide H and carbonyl O atoms of the two peptide links and three methyl carbon atoms; the average is taken over the time of the simulation. Each of the 195 water molecules is assigned to one of three solvent classes, denoted "polar", "nonpolar", or "bulk"; "polar" if the mean distance to any of the four polar atoms is less than 4 Å, "nonpolar" if the mean distance is greater than 4 Å to a polar atom but less than 5 Å to a methyl carbon, and "bulk" if neither preceding criterion is met.

The division into solvent groups results in the assignment of 14 molecules as "polar", 20 as "nonpolar", and the remaining 161 molecules as "bulk". As described earlier, only a fraction of the molecules in the polar group participate in strong bonds with the dipeptide. Their proximity to the polar atoms, however, suggests that they all have a significant interaction energy with the peptide groups.

We find that the average interaction energies with the dipeptide for the water molecules in each class are −0.02 ("bulk"), −0.18 ("nonpolar"), and −1.13 kcal/mol ("polar"). This corresponds roughly to $0.03k_BT$, $0.3k_BT$, and $2k_BT$, respectively.

The net displacement of a typical solvent molecule during the simulation is only about half the solvent molecular diameter (see part B), and the solute also moves relatively little. Consequently, the identification of particular molecules with well-defined spatial regions in the solution is a meaningful procedure. As a test of this, we can compare the solvent breakdown given above to the corresponding results obtained based on the average positions in the first half or last half of the simulation. In the latter case, the classification of 3 of the 14 polar molecules is changed, while in the former case 1 of the 14 is changed. Each of these four molecules has average positions close to the 4-Å dividing line defining the polar region (see above). As a corollary, it is clear that the structures seen in the present simulation represent not an overall average but a selected sample average. The length of time considered does not permit substantial rearrangement of the molecular centers of mass; such a time-averaged structure corresponds closely to that termed "V structure" by Eisenberg and Kauzmann.[44]

We note that the same consideration of relatively slow mobility is employed in the analysis of experimental results;[7] that is, an assignment of particular molecular behavior (e.g., rotational correlation time) to solvent molecules interacting with a solute is valid only if the residence time in the vicinity of the solute is longer than the characteristic time being examined. For times of the order of a few picoseconds (e.g., rotational correlation times), this so-called slow exchange assumption,[7] usually made in experimental analyses, is satisfactory. In the next subsection, we describe the results obtained from the analysis of kinetic properties using the solvent classification procedure we have developed.

B. Molecular Dynamics of the Solvent. Having formulated a procedure for dividing the solvent into regions, we employ it to examine the average dynamic properties of the water molecules in each class. We emphasize that, in the absence of additional information, the solvent groups are defined only to aid our analysis. If the results show that the water molecules in each class behave in a qualitatively different way, a physical basis for this division into groups will be obtained. Correspondingly, if the "bulk" water (in the present case, the second solvation layer) manifests the dynamics of pure water, we can conclude that the effect of the solute is primarily restricted to the first solvation layer. Further, the observation of qualitatively different behavior in the "polar" and "nonpolar" regions can support the view that the effect of individual solute groups is localized, as is suggested by the experiments on mixed-functional solutes[1,3,5] mentioned in the Introduction.

Since the number of molecules in each of the polar and nonpolar classes is not large, the statistical accuracy of results obtained for them is expected to be significantly worse than for the bulk water. The goal of our analysis is, therefore, limited to the observation of clear qualitative differences in molecular behavior rather than to the determination of precise values of, for example, correlation times.

During the simulation, the mean kinetic temperature of each of the three groups of water molecules was found to be identical and equal to 303 K. In Table III, the mean atomic temperatures in the solvent are summarized; the values given correspond to the average over all atoms of the species indicated (H or O) in each class. (The values given are mean values for the simulation; the instantaneous kinetic temperatures necessarily fluctuate substantially, corresponding to the exchange of potential and kinetic energy during molecular motion.) The variation among the atomic temperatures (291 to 319 K) shows that the exact agreement among the mean molecular temperatures (303 K) is partly fortuitous. Although the relative water/dipeptide temperatures were adjusted during the equilibration period (see section II), no attempt was made to equilibrate the different groups of water molecules. Thus, the agreement of the various temperatures is satisfactory.

Table III. Solvent Temperatures[a]

type[b]	T_H	T_O	T
total	297	315	303
bulk	295	319	303
nonpolar	309	291	303
polar	305	298	303

[a] Temperatures in K; H and O are the water molecule hydrogen and oxygen atoms; $T = \frac{1}{3}(2T_H + T_O)$. [b] Solvent class as defined in text; "total" includes all water molecules in the system.

Our principal interest is an examination of the translational and rotational mobility of the solvent molecules, since changes in the mobility have been associated experimentally with solute-solvent interaction.[3,7-9] These motions can be characterized by time correlation functions of both velocity and position (angular and translational). Since the velocity correlation functions are more difficult to interpret we describe first the results obtained from the position correlation functions; the velocity functions are considered at the end of the subsection. Correlation functions of rotational reorientation and translational position (see below) are generally monotonic,[19,41] since, unlike velocities, positions do not oscillate between positive and negative values. Furthermore, the values obtained at each time can be given an easily interpreted physical meaning.

For isotropic translational diffusion, the mean square displacement of a particle is initially quadratic in time but becomes linear in the asymptotic limit;[29] that is

$$\langle \mathbf{R}^2(t) \rangle \rightarrow 6Dt; \quad t \rightarrow \infty \tag{5}$$

where $\mathbf{R}(t)$ is the vector connecting the particle's position at an initial time and its position at a time which is later by an amount t, and D is, by definition, the translational diffusion constant. Past experience with molecular dynamics calculations[19] shows that the asymptotic region is reached in a time much shorter than that required to diffuse a molecular diameter. In Figure 15, we show the calculated mean square displacements for center of mass translation of the water molecules, averaged separately within each defined solvent class. In this figure, the curves labeled (a), (b), and (c) correspond to the solvent classes "bulk", "nonpolar", and "polar", respectively; the same labeling is used for all figures appearing in this section. In addition, Figure 15 includes the mean square displacement for the dipeptide solute center of mass, curve (d).

The curve labeled (a), corresponding to the bulk class, is consistent with the result obtained for pure water in previous simulation studies.[19] The average polar class behavior, shown by curve (c), is very similar to that found for the bulk class, though there are differences between the curves for times greater than about 0.4 ps. Owing to statistical accuracy (which decreases at longer times), we conclude that the difference between curves (a) and (c) cannot be assigned any significance. Thus, the translational mobility of the water molecules in the polar group appears to be similar to the bulk. In striking contrast to this, the result for the nonpolar class (curve (b)) shows a reduced net displacement and reduced limiting slope as compared to the bulk class, corresponding to a significantly reduced diffusion rate. Furthermore, this difference occurs for all times longer than about 0.1 ps, so that the difference cannot be ascribed to statistical inaccuracy; the slight lack of linearity in curves (c) and (d) at long times ($t > 0.6$ ps) *is* almost certainly statistical error. The diffusion constants obtained from

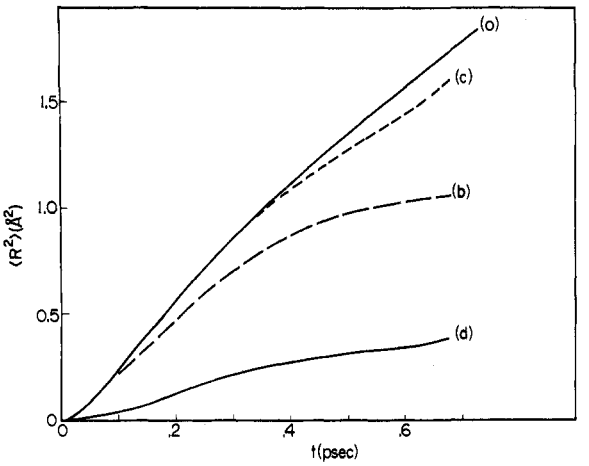


Figure 15. Mean square displacement for water molecule center of mass translation: (a) "bulk"; (b) "nonpolar"; (c) "polar", as defined in the text; (d) solute center of mass.

Table IV. Characteristics of Translational and Rotational Mobility

H_2O class[a]	D, 10^{-5} cm^2/s[b]	τ_1, ps[c]	τ_2, ps[c]	τ_1/τ_2
total[d]	3.24	3.4	1.4	2.43
bulk	3.45	2.7	1.1	2.45
polar	2.8	3.7	1.8	2.06
nonpolar	0.68	8.6	3.1	2.77
polar/bulk[e]	0.80	1.3	1.6	
nonpolar/bulk[e]	0.20	3.1	2.8	
total/bulk[e]	0.94	1.1	1.1	

[a] As defined in the text. [b] Estimated from the limiting slope of the curves shown in Figure 15, using eq 5. [c] Obtained by fitting a single exponential to the curves shown in Figures 16 and 18. [d] Weighted average of the polar, nonpolar, and bulk values, equivalent to slow exchange assumption (ref 7). [e] The value given is the ratio for the two classifications indicated of the characteristic (D, τ_1, τ_2) indicated at the top of the column.

these curves by use of eq 5 are given in Table IV. From curve (d), it is evident that, although the net motion of the solute is quite small, the asymptotic dipeptide displacement rate is indistinguishable from that of the nonpolar class of water molecules. The equivalence of these rates implies that, *on this short time scale*, the water molecules in the neighborhood of the methyl group behave as though they are bound to the dipeptide. The results shown in Figure 15 suggest that these "nonpolar" solvent molecules are not *tightly* bound to the solute, in that the magnitude of the solvent mean square displacement is much larger than that of the solute (the difference between the mean square displacements is about 0.7 Å^2 after 0.5 ps). In the case of *tight* binding, the displacements, as well as their slopes, are expected to be asymptotically equal. The observed behavior is analogous to that which would be expected if the water molecules were bound to the methyl groups by a weak spring, allowing relative motion, but no net diffusion. However, a larger mean square displacement of the nonpolar solvent molecules, compared to that of the solute center of mass, could arise even if the motion of these water molecules took place at a *fixed* distance from the solute center of mass; this would result if the water molecules translated principally parallel to the methyl group surface, or if these molecules "followed" the methyl groups during solute rotation. In the latter case, a difference in square displacement would be expected even in the case of tight binding.

The mean square displacement is expected to become linear in time for times, t, greater than the translational velocity correlation time, τ_v. For a spherical Brownian particle with a

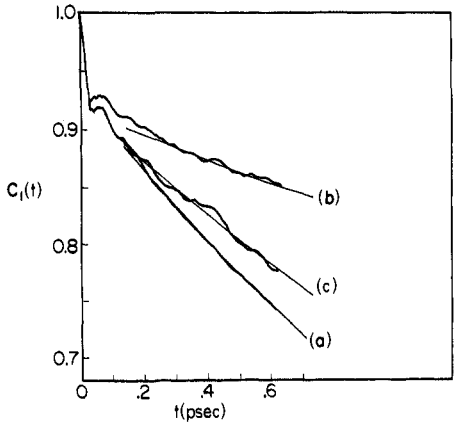


Figure 16. Rotational reorientation of water molecule dipole direction for $l = 1$ (eq 7). The curves are labeled as in Figure 15.

hydrodynamic radius, a, consistent with the size of the dipeptide ($a \sim 3.5$ Å), and with the dipeptide mass ($M = 144$ amu), this time is (stick boundary conditions)[31]

$$t \gtrsim \tau_v = \frac{M}{f} = \frac{M}{6\pi\eta a} \sim 0.04 \text{ ps}$$

where f is the translational friction coefficient. Although the time for the transition to the asymptotic regime is found to be larger than this rough estimate ($t < 0.3$ ps, Figure 15), the diffusion constant calculated from the limiting slope of the solute mean square displacement is consistent with a hydrodynamic estimate of the dipeptide diffusion constant using the Stokes-Einstein law (stick boundary conditions)

$$D = \frac{k_B T}{6\pi\eta a}$$

where η is the solvent shear viscosity (0.01 P). With an assumed hydrodynamic radius, a, of 3.5 Å, the diffusion constant is about 6.3×10^{-6} cm^2/s, as compared with the dynamics estimate of 6.5×10^{-6} cm^2/s.

The time correlation functions for rotational motion[45]

$$C_l(t) = \lim_{t' \to \infty} \frac{1}{t'} \int_0^{t'} dt\, P_l(\hat{\mu}(\tau) \cdot \hat{\mu}(t+\tau)) = \langle P_l(\hat{\mu}(0) \cdot \hat{\mu}(t)) \rangle \equiv \langle P_l(\cos\theta(t)) \rangle \quad (6)$$

where $P_l(x)$ is the Legendre polynomial of order l, measure the average reorientation rate of the molecular dipole direction, given at time t by the unit vector, $\hat{\mu}(t)$. As in eq 3, the averages in eq 6 are evaluated using all pairs of configurations separated by a time t during the simulation of *finite* length, t'. In particular, we have

$$C_1(t) = \langle \cos\theta \rangle \quad (7)$$

and

$$C_2(t) = \langle (3\cos^2\theta - 1)/2 \rangle \quad (8)$$

These correlation functions decay to zero as the molecular orientation becomes randomized with respect to its initial value; $C_2(t)$ typically decays more quickly than $C_1(t)$.[45] The calculated results for $C_1(t)$ (eq 7) are shown in Figure 16 for the different groups of water molecules. The initial rapid decay of the correlation function during the first 0.05 ps (approximately the decay time of the angular velocity correlations; see below) corresponds to overall molecular oscillation (libration) with a loss of phase memory, but without significant net reorientation. In this short time period, the three groups appear to behave similarly. However, for longer times, the differences in the decay rates of $C_1(t)$ correspond to those found for translational diffusion. The molecules in the polar class reorient

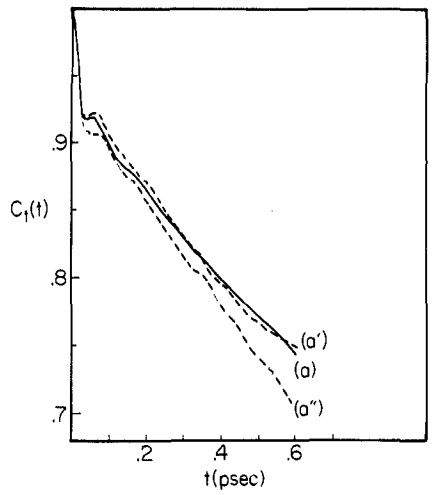


Figure 17. Comparison of $C_1(t)$ (eq 7) evaluated for the "bulk" from 161 molecule average (solid curve, a) and two randomly chosen groups of 20 molecules of the 161 (dashed curves, a', a").

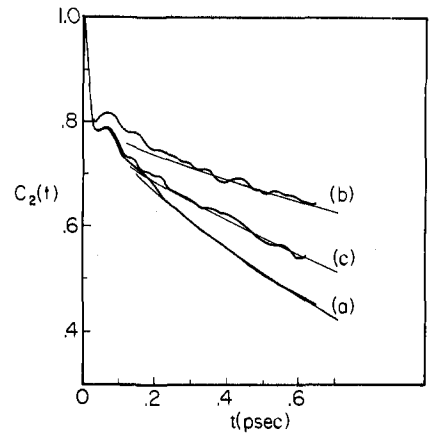


Figure 18. Rotational reorientation of water molecule dipole direction for $l = 2$ (eq 8). The curves are labeled as in Figure 15.

at a rate similar to that exhibited by the bulk class, while those in the nonpolar class reorient more slowly. As for translational motion, the nonpolar curve differs from the others at all times after about 0.1 ps.

It is of interest to have a quantitative measure of the relaxation rates for reorientation. Relaxation times determined experimentally (e.g., by NMR methods) correspond to[8]

$$\tau_l = \int_0^\infty dt\, C_l(t)$$

where the correlation function, $C_l(t)$, is defined in eq 6. To evaluate the time integral requires a knowledge of the correlation function for all times; this is not available here. We can, however, obtain an easily comparable measure of the decay rates for each solvent class from the computed functions by carrying out a least-squares fit to a single exponential; for this fit, we consider the period from 0.25 to 0.6 ps. The narrow solid line drawn through each curve in Figure 16 corresponds to such an exponential fit. The relaxation times obtained in this way (Table IV) are denoted τ_l. From the values of τ_l, there is some indication that the polar group reorients slightly more slowly than the bulk, though, as for translation, the statistical errors are too large to be certain.

It is important to estimate the significance of statistical errors in the above comparison; that is, to determine whether the differences between classes are real or whether they are due to the errors introduced by using a small number of molecules and studying them for a relatively short time. Since the "bulk" values are based on 161 molecules they are likely to provide valid averages. In contrast, the polar and nonpolar classes include only 14 and 20 molecules, respectively, and so are expected to have larger statistical errors due to the limited sample size. As a test of these errors, we have selected groups of 20 molecules at random from the bulk class and evaluated the corresponding functions for the chosen subset. A comparison of the 161 molecule average and various 20 molecule averages provides a guide to the significance of the results obtained. In Figure 17, we show such a comparison for $C_1(t)$; curve (a) corresponds to the 161-molecule average, curves (a') and (a") to representative 20-molecule averages. All three curves yield similar reorientation rates; the τ_1 values are 2.7, 2.4, and 2.1 ps for curves (a), (a'), and (a"), respectively. This suggests that uncertainties in τ_1 of 10–12% may be expected. Comparing Figure 16 and Figure 17, we see that the difference between the polar (c) and bulk (a) curves in Figure 16 could be a reflection of statistical error, but that this is not true of the much larger difference between (b) and either (a) or (c).

The calculated functions for $l = 2$ reorientation (eq 8) are shown in Figure 18. We note that the decay rate of $C_2(t)$ is significantly higher than that of $C_1(t)$ (Figure 16), as expected. Again, the "nonpolar" water molecules reorient significantly more slowly than those in either of the other two classes. The exponential fits to the curves in Figure 18 yield the values, τ_2, given in Table IV.

From Table IV, it is clear that, both for translational and rotational motion, the water molecules in the neighborhood of the methyl groups (nonpolar) are slower (small D, larger τ_1 and τ_2) than the water molecules associated with the C=O or N—H groups (polar) or the "bulk" water. It is of considerable interest to compare these results with the available experimental data. A variety of techniques have been used to examine molecular motion in the pure solvent and in solutions. These include neutron scattering,[9,46] nuclear magnetic resonance,[7,8,47] dielectric relaxation,[48,49] Raman and infrared spectroscopy,[50,51] depolarized Rayleigh scattering,[52] and ultrasonic absorption.[53,54] Each of these techniques gives somewhat different information concerning molecular motion and the possibilities and limitations of the various approaches are discussed in the cited references. For translational diffusion, the results are relatively unequivocal, but for molecular rotation and the associated relaxation times, the values obtained from a particular experiment are dependent on the choice of motional model involved.

Considering first studies of pure water, the most recent results for the translational diffusion constant,[55] obtained using isotopic tracer methods, yield a value of 2.7×10^{-5} cm^2/s at 30 °C; the results of earlier work lead to values about 10% higher.[56] The diffusion constant is rather sensitive to temperature; at 25 and 35 °C, it equals 2.3×10^{-5} and 2.9×10^{-5} cm^2/s, respectively.[55] This temperature variation can be described by an activation energy of about 4.7 kcal/mol.[55] Simulation studies of pure water lead to rather higher values of the diffusion constant. With the BNS model,[41] a value of 4.2×10^{-5} cm^2/s is obtained at 34.3 °C; the ST2 model[19] leads to 1.9×10^{-5} cm^2/s at 10 °C and 4.3×10^{-5} cm^2/s at 41 °C, suggesting a value of about 3.5×10^{-5} cm^2/s for the translational diffusion constant at 30 °C.

A direct experimental measurement of the rotational correlation time, τ_1, cannot be carried out with known techniques. The dielectric relaxation time, τ_d, is a closely related quantity; however, since dielectric relaxation is a collective (many-particle) effect, the values of τ_1 and τ_d are different: and their relationship is not established.[57] Theoretical considerations for polar fluids[58,59] suggest that τ_d is about $3/2$ to 2 times as large as τ_1. Although the development leading to these esti-

mates is based on models much simpler than that appropriate for water, it is nevertheless of some interest to consider results obtained with these approximate ratios. From the BNS model[41] a value of 5.6×10^{-12} s is obtained for τ_1, leading to approximate values of τ_d in the range $8.4\text{-}11.2 \times 10^{-12}$ s at the temperature considered, 34.3 °C. The experimental value,[60] 6.7×10^{-12} s, is somewhat smaller. For the temperature used in the present work (30 °C), the experimental value is[60] $\tau_d = 7.4 \times 10^{-12}$ s. No corresponding results obtained from the ST2 model have been published.

The relaxation time corresponding to $l = 2$ reorientation (eq 8), τ_2, can be obtained directly from NMR spin-lattice relaxation measurements, and the values characterizing pure water are reasonably well established.[61] At 30 °C, the experimental τ_2 is about 2×10^{-12} s. The temperature dependence of τ_2 is comparable to that of the diffusion constant.[52,61] The BNS model[41] yields a value of τ_2 of 2.1×10^{-12} s at 34.3 °C, in rather good agreement with the experimental value[61] of 1.9×10^{-12} s. As for τ_1, values of τ_2 from the rigid ST2 model are not available.

Because of the experimental uncertainties in the determination of τ_1 and τ_2,[62] and the model dependence of the derived value of τ_1,[57-59] the ratio, τ_1/τ_2, cannot be obtained with high accuracy. However, this ratio is of considerable interest, since it appears as one parameter that is in principle both measurable and derivable from inertial and stochastic theories of molecular reorientation in fluids.[45,63] For small-step Brownian rotational diffusion, $\tau_1/\tau_2 = 3$; for models involving a series of finite steps or "jumps", $\tau_1/\tau_2 < 3$. It must be stressed that τ_1/τ_2 is not a rapidly varying function of the jump size, as has been pointed out previously.[64] For a diffusional model assuming a Gaussian distribution of step sizes, θ, a τ ratio of 2.7 is consistent with a rather large root mean square step, $\langle\theta^2\rangle^{1/2}$, of 20°. Hence the interpretation of the τ ratio in terms of average step size is very sensitive to the precision of the determination of τ. For the BNS model at 34.3 °C,[41] $\tau_1/\tau_2 = 5.6/2.1 \sim 2.7$. Allowing a 10% variation in each value (ps, $5.0 \lesssim \tau_1 \lesssim 6.2$; $2.4 \lesssim \tau_2 \lesssim 3.0$), we obtain

$$2.2 \lesssim (\tau_1/\tau_2) \lesssim 3.3$$

demonstrating the difficulty of comparing the simulation with theory.

Considering the present results, the translational diffusion constant D, obtained for the bulk class, 3.45×10^{-5} cm^2/s (see Table IV), agrees well with that obtained using the ST2 model[19] (3.5×10^{-5} cm^2/s; see above), although not as well with the experimental value[55] of 2.7×10^{-5} cm^2/s. The rotational reorientation time, τ_1, obtained for the bulk class, 2.7×10^{-12} (see Table IV), is smaller than that obtained from the BNS model[41] (5.6×10^{-12} s). As for pure water, above, the rough estimates for the dielectric relaxation time corresponding to the present "bulk" value of τ_1 are

$$4.0 \times 10^{-12}\ \text{s} \lesssim \tau_d \lesssim 5.4 \times 10^{-12}\ \text{s}$$

compared to the experimental value of 7.4×10^{-12} s. For τ_2, we obtain a "bulk" value of 1.1×10^{-12} s, compared to 2.0×10^{-12} s, experimentally.[61] These comparisons suggest that the reorientation rates in the present model are rather more rapid than in the real fluid. However, our primary consideration here is a comparison among the various classes (polar, nonpolar, and bulk). The decrease in D by a factor of 5 and the increase by a factor of 3 in τ_1 and τ_2 for the nonpolar class relative to the bulk (see Table IV) are large effects. For the reorientation times, it appears to be consistent with experimental estimates for nonpolar solutes,[3,4,8] namely, an increase of roughly a factor of 2-3 depending on the system studied. (The τ_l values in Table IV denoted "total" correspond to the slow exchange assumption[7], i.e., to the weighted average of the τ_l from each of the three classes). Considering the likely precision in τ_1 and τ_2 discussed above, it is unreasonable to interpret differences in the ratios, τ_1/τ_2, in terms of microscopic models for rotational motion (e.g., jump diffusion, small diffusion, etc).[45,62] The ratios are, however, included in Table IV for completeness.

Since a direct measurement of the translational diffusion constants of different solvent groups is not possible, it is of interest to compare the "total" system value with the changes observed in solution as a function of concentration. Here, the value for the "total" system (see Table IV) is the average value for all solvent molecules in the system in accord with the slow-exchange assumption.[7] In a study of Franks et al.[9] the decrease in the average water translational diffusion constant was measured over a range of *tert*-butyl alcohol concentrations. At their lowest concentration (2 mol %; 1 molecule of *t*-BuOH to 49 molecules of water) they observed a decrease of 14% in the diffusion constant To compare our value to theirs, we require an estimate of the number of water molecules surrounding each *tert*-butyl group. This can be obtained by assuming that the three CH_3 groups in *tert*-butyl are comparable to the three in the dipeptide. Thus we infer that the 14% observed reduction is due to 20 of each of the 49 water molecules per alcohol molecule, if only the first solvation layer is influenced by the solute. Scaling to our concentration, 0.51 mol %, and assuming a linear concentration dependence, we obtain an experimental estimate of a 3.5% difference in the *tert*-butyl alcohol case as compared to the calculated value of 6%. Considering the crude assumptions made in this comparison, the agreement is satisfactory. We note that the relatively small change in the effective diffusion constant results from a large reduction (by a factor of 5) in the diffusion rate of the 10% of the water molecules near the nonpolar solute groups and unchanged values for the rest of the system.

We now consider briefly velocity correlation functions. The interpretation of such functions is more difficult than those of position. This results from the fact that, for water, both the angular and translational velocity correlation functions have considerable oscillatory structure.[19,41] Such behavior arises since, in general, each successive "collision" of a molecule with its neighbors tends to reverse the direction of the velocity; these "rebounds" are manifest in the velocity correlation function as successive oscillations. For water, the oscillations in the translational velocity autocorrelation function are much more pronounced than in the corresponding function for simple liquids (e.g., argon).[43]

For the center of mass velocity, the normalized correlation function $C_v(t)$

$$C_v(t) = \frac{\langle v(0) \cdot v(t) \rangle}{\langle v^2 \rangle}$$

is shown in Figure 19. The plotted curves are, as above, (a) "bulk", (b) "nonpolar", and (c), "polar". Curves (b) and (c) are truncated since we believe that the calculated functions are not meaningful beyond the time plotted. The corresponding correlation functions for the total molecular angular momentum, $C_J(t)$,

$$C_J(t) = \frac{\langle \mathbf{J}(0) \cdot \mathbf{J}(t) \rangle}{\langle J^2 \rangle}$$

are given in Figure 20. It is clear from Figures 19 and 20 that, for times less than about 0.15 ps, the correlation functions, $C(t)$, are qualitatively similar; at longer times they are different, but it is likely that such differences are, in part, the result of statistical inaccuracy. Since the actual values of the quantities $C(t)$ are close to zero in this region ($t > 0.15$ ps), statistical inaccuracy in their evaluation can have large effects on the values of the functions and hence make comparison among the curves nearly impossible. Further, a qualitative

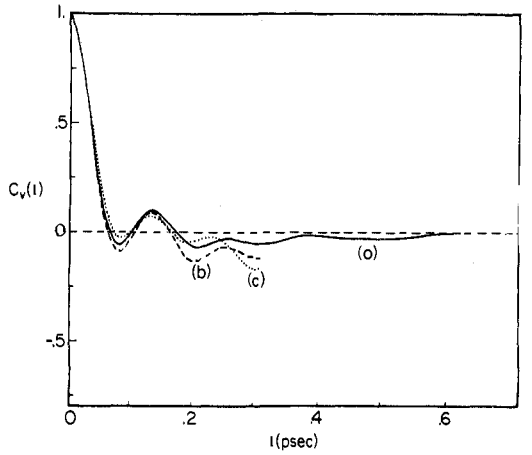


Figure 19. Water molecule center of mass velocity autocorrelation function: (a) "bulk"; (b) "nonpolar"; (c) "polar".

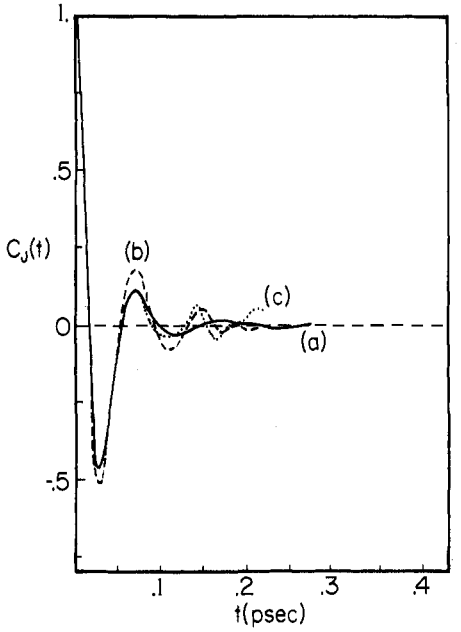


Figure 20. Water molecule total angular momentum autocorrelation function; (a) "bulk"; (b) "nonpolar"; (c) "polar".

similarity at short times does not suffice to form useful conclusions as to the relatively mobility of molecules in the several classes.

The precise values of the characteristic diffusion constants, which are obtained from the zero frequency component of the Fourier transform of the correlation function[29]

$$\tilde{C}(\omega) = \int_0^\infty dt \cos(\omega t) C(t)$$

result from substantial cancellation between the positive and negative contributions to the Fourier integral which occur in various time intervals. Statistical errors at longer times can therefore lead to significant errors in the values obtained. These errors can also confuse the interpretation of relatively small differences in the calculated spectral densities at nonzero frequencies.

As has been seen by a number of spectroscopic methods,[65] as well as in simulations,[19,41] water exhibits broad spectral bands corresponding to hindered translational motion, centered near 60 (11 ps^{-1}) and 185 cm^{-1} (35 ps^{-1}), and a very broad librational band, from about 300 to 900 cm^{-1} (55-170 ps^{-1}).

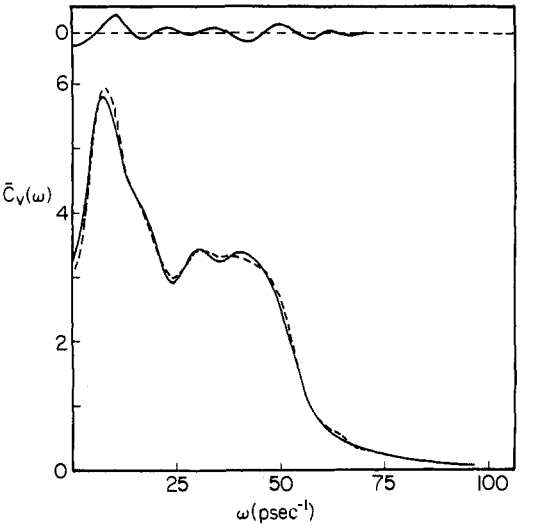


Figure 21. Spectral density from water center of mass velocity autocorrelation function: "bulk", solid; total system, dashed. The difference is shown at the top of the figure.

Variation in these bands might present one possible experimental probe of solvation. In one such study on a protein solution, it has been suggested that the translational band is shifted slightly to higher frequency owing to solvent "structuring".[66] In order to estimate the precision in band shape measurements which would be required to observe such changes, we compare the calculated spectral densities resulting from consideration of, first, all water molecules in our system and, second, only those in the bulk class.

The hindered translational bands, calculated from the center of mass velocity autocorrelation function

$$\tilde{C}_v(\omega) = \int_0^\infty dt \cos(\omega t) C_v(t)$$

are shown in Figure 21. The solid curve corresponds to the bulk class and the dashed curve to the total system. The spectral bands are located in the expected regions (see above). At the top of the figure the difference spectrum is shown (relative to zero as indicated). The spectra differ by 6% (relative to $\tilde{C}(\omega)$ for the bulk) at zero frequency, corresponding to the observed change in diffusion constant (3.45 vs. 3.24×10^{-5} cm^2/s; see Table IV). At nonzero frequencies the difference is ±4% or less. Furthermore, these differences are not uniform over any significant frequency range, suggesting that they result, at least in part, from statistical error.

The librational bands, calculated from the angular momentum autocorrelation function

$$\tilde{C}_J(\omega) = \int_0^\infty dt \cos(\omega t) C_J(t)$$

are shown in Figure 22. As for translation, the position of this broad band is in accord with experiment. The two spectra differ even less than for the translational case. The calculated difference spectrum (top) oscillates in sign and has a maximum at the center of the band corresponding to only 3% of the total spectral amplitude. These results suggest that changes in the spectra, if any, would be extremely difficult to observe.

V. Structural Origins of the Solvent Dynamics

We have seen that the dynamics of the solvent molecules near the nonpolar groups differs considerably from that of the bulk solvent, a result in agreement with experimental inferences.[3,7-9] The essential conclusion is that water molecules near the nonpolar groups of the solute have longer translational and rotational relaxation times than either the bulk water or that

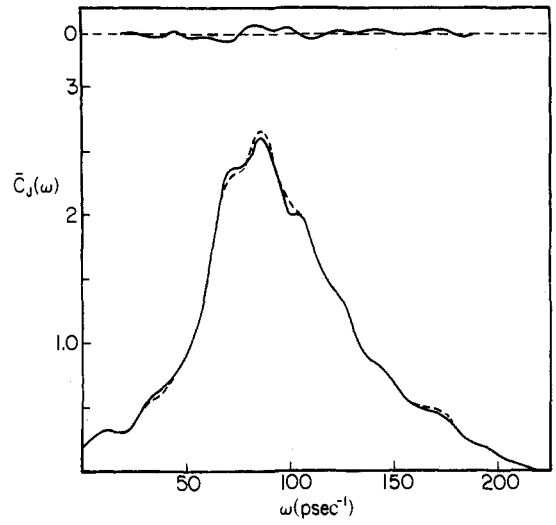


Figure 22. Spectral density from water angular momentum autocorrelation function: "bulk", solid; total system, dashed. The difference is shown at the top of the figure.

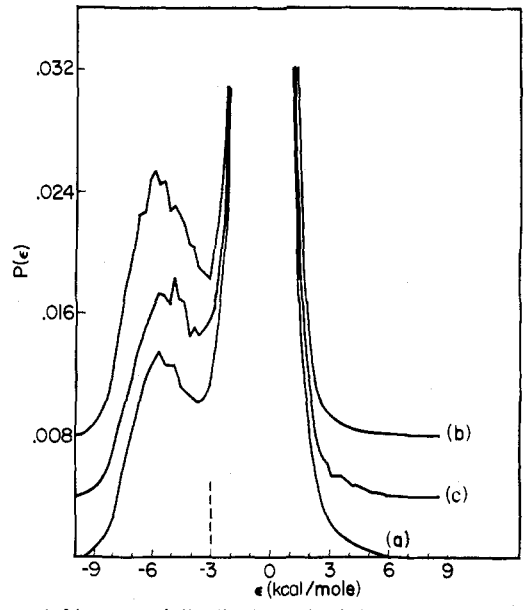


Figure 23. Normalized distributions of pair interaction energies among water molecules; a hydrogen bond is defined by $\epsilon \leq -3$ kcal/mol, indicated by the vertical mark on the abscissa. (a), all pairs in the system; (b), one of the pair in the "nonpolar" class; (c), one of the pair in the "polar" class. Each curve is integrally normalized to unity.

near the polar groups of the solute. To determine the origin of this difference, we examine certain time-averaged structural and energetic properties of the various solvent species.

A. Energetic Factors. We consider first some average properties related to the bonding energetics. Of interest are the strengths of the hydrogen bonds involving the different solvent species, the total interaction energies in the presence and absence of the solute, and the number of hydrogen bonds formed by each molecule.

Figure 23 shows the calculated distributions of water-water pair interaction energies, analogous to the water-dipeptide pair energy distribution (Figure 10) presented earlier; in the figure, the relative origins of the curves labeled (b) and (c) are shifted upward by 0.008 and 0.004, respectively; the particular values given on the ordinate refer to the curve labeled (a). Each curve gives the probability $P(\epsilon)$ of observing a pair of molecules with interaction energy, ϵ, when all pairs within the potential range (8 Å) of each other are included; the curves are individually normalized such that their integral is unity. The three separate distributions correspond to (a) all distinct molecular pairs in the system; (b) all pairs which include at least one molecule in the nonpolar class; and (c) all pairs which include at least one molecule in the polar class. It is not possible to completely separate the distributions for molecules in each class, since the pair energy is a function of the class of two particles; for example, a pair involving one "polar" solvent molecule and one "nonpolar" molecule contributes to all three distributions. The peak at $\epsilon = 0$ includes the relatively large number of molecular pairs which are well separated in space and therefore have very small average interaction energies; the number of such pairs is finite owing to the finite potential range. The general shape of these curves, including the appearance of a local maximum in the negative energy region, is that expected from studies of pure water.[19] The precise value of the positions of this local peak and of the local minimum differ somewhat from previous studies owing to the difference in potential functions used. Here, each occurs at a value about 1 kcal/mol more negative than for the ST2 model.[19]

To define hydrogen-bonded water pairs, we use the same energy criterion (−3 kcal/mol) as we employed earlier for the water-dipeptide pairs (section IV). The choice of −3 kcal/mol is in accord with the shape of the distribution in Figure 23; this value corresponds approximately to the position of the local minimum that separates the central peak from the bonding region. However, the precise value is not essential to the interpretation of the calculated results (as will be seen below).

Table V. Water-Water Hydrogen Bond Energies[a]

class[b]	mean bond energy, $\bar{\epsilon}$
total	−5.25
nonpolar	−5.35
polar	−5.21

[a] Calculated from results shown in Figure 23; a hydrogen bond is defined as $\epsilon \leq -3$ kcal/mol. [b] As defined in the text.

The strength of intermolecular bonding for each type of solvent is reflected by characteristics of the peak occurring near $\epsilon = -5$ kcal/mol in the curves of Figure 23. It is clear that there are only small differences, if any, in the positions of the peak; that of the polar group is at a slightly more positive energy, and that of the nonpolar group at a slightly more negative energy than that of the total system. A corresponding negative shift in pair interaction energies for water molecules in the first solvation layer of methane has been noted in a recent simulation study.[43] An alternative comparison is obtained from the calculated mean pair energy, $\bar{\epsilon}$, for bonded pairs ($\epsilon \leq -3$ kcal/mol). The relative values of $\bar{\epsilon}$, which are given in Table V, are in accord with the positions of the peaks in the figure; nevertheless, considering the noise level in the figure, the observed differences (~0.1 kcal/mol) may not be statistically significant. In any case, the shift is much smaller than k_BT (~0.6 kcal/mol), indicating that changes in hydrogen bond energies, per se, cannot account for the observed differences in dynamic behavior.

It is important to note that very small changes in the mean bond energy can have significant thermodynamic effects. The bonding region for the nonpolar group (curve (b)) includes contributions from 40 distinct molecular pairs in a typical configuration. Consequently, a shift of only −0.05 kcal/mol in the mean bond energy (Table V) can contribute an enthalpy change of −2 kcal/mol. This result suggests that enthalpies of solution are sensitive to small changes in the average bond energies; they are analyzed further below.

It is clear from a comparison of Figure 23 and the corresponding interaction energy distribution for solvent-solute association (Figure 11) that the solvent-solute hydrogen bonds

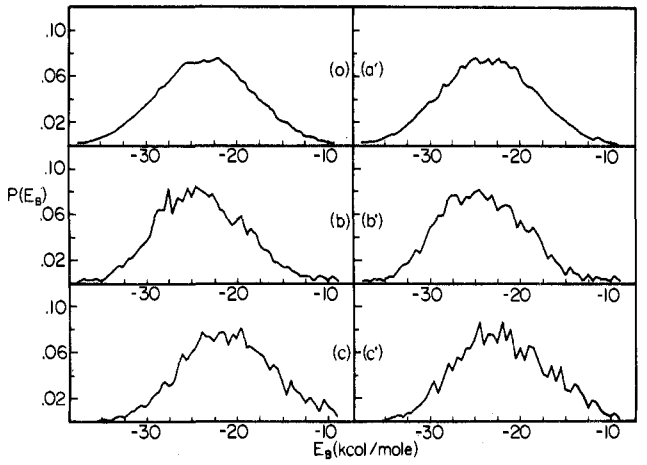


Figure 24. Binding energies for solvent molecules. "Bulk" (a, a'), "nonpolar" (b, b'), "polar" (c, c'). Curves a, b, c exclude interactions with the solute; a', b', c' include them.

are typically weaker than the solvent-solvent bonds. From the results shown in Figure 11, we can determine a mean solvent-solute bond energy corresponding to that obtained above for water-water association. We find that the average water-dipeptide interaction energy for the pairs which interact with an energy more negative than −3 kcal/mol is −4.21 kcal/mol. This can be compared with the value appropriate for water-water (see Table V) of −5.25 kcal/mol. Some caution must be exercised in regarding this difference in bonding strengths, since it depends on the hydrogen-bond potential functions used. In light of the comparable relative vacuum energies of the hydrogen bonds involved (see section II), the observed energy difference appears to be a result of the average structure in the solution. The relative weakness of water-dipeptide bonds could arise from limitations placed on the disposition of the intermolecular geometry for water molecules involved in such multiple bonding. However, it seems more likely that the difference results from the flexibility of the water molecules in the model solution. The latter allows closer approach of proton and lone-pair charges (which are not individually surrounded by repulsive van der Waals spheres) than does the rigid ST2 structure[19] and leads to a shift to stronger hydrogen bonds. This shift is apparent in the water-water bond energies in Figure 23, as noted earlier; the strongest bonds attained in solution approach 10 kcal/mol, while the rigid (ST2) structure (corresponding to the vacuum calculations[18]) leads to a maximum strength of 6.8 kcal/mol. However, unlike the water molecules, the dipeptide charges are each surrounded by a repulsive core potential. Hence, the solute charges cannot approach the center of a water molecule more closely than in the rigid structure, and the influence of the flexibility of the water charges is expected to be less for the solute-solvent bond energies. This result is clear in Figure 10; the strongest water-solute bonds approach 8 kcal/mol, comparable to the optimum vacuum value of 7.4 kcal/mol for a C=O···water bond,[18] obtained with the rigid structure.

An alternative description of the intermolecular bonding is provided by the total binding energy of each molecule in the solution; that is, by the sum of interaction energies of a given molecule with all other molecules. The potential range cutoff of 8 Å is employed in this calculation, as it is in the simulation; the error in the energies resulting from the cutoff is expected to be about 2%.[19] Letting i and j denote water molecules, the total binding energy for molecule i in the solution, denoted $E_B^T(i)$, is

$$E_B^T(i) = \epsilon_{iD} + \Sigma\, \epsilon_{ij}$$

Table VI. Average Binding Energies of Solvent Molecules[a]

	bulk[c]	nonpolar[c]	polar[c]
$\langle E_B' \rangle^b$	−23.53	−23.87	−20.77
$\langle E_B^T \rangle^b$	−23.55	−24.05	−21.90
$\langle E_B^T - E_B' \rangle^b$	−0.02	−0.18	−1.13
$\langle E_B' \rangle - \langle E_B^{bulk} \rangle^b$	0.0	−0.34	2.76
$\langle E_B^T \rangle - \langle E_B^{bulk} \rangle^b$	−0.02	−0.52	1.63
$\langle E_B(D) \rangle^b$	−22.64		
$\Delta E_c{}^b$	−6.72		

[a] All values in kcal/mol. [b] Energy quantities as defined in text. [c] Water classes as defined in text.

where ϵ_{ij} is the pair interaction energy between molecules i and j and ϵ_{iD} is the energy of interaction with the dipeptide. The water binding energy excluding the dipeptide, $E_B'(i)$, is

$$E_B'(i) = \sum_{\substack{j=1 \\ (j \neq i)}}^{195} \epsilon_{ij}$$

and the dipeptide binding energy is

$$E_B(D) = \sum_{i=1}^{195} \epsilon'_{iD}$$

The total intermolecular potential energy, E_c, is

$$E_c = \tfrac{1}{2}\left(\left(\sum_{i=1}^{195} E_B^T(i)\right) + E_B(D)\right) \equiv \tfrac{1}{2}\left(\sum_{i=1}^{195} E_B'(i)\right) + E_B(D)$$

(The factor of $\frac{1}{2}$ accounts for the fact that all pair energies are included twice in the summation.) In Figure 24, we show normalized distributions of the quantities $E_B'(i)$ (Figures 24a-c) and $E_B^T(i)$ (Figures 24a'-c') subdivided according to solvent class; the distributions correspond to "bulk" (a, a'), "nonpolar" (b, b'), and "polar" (c, c'). In Table VI, the corresponding average values are listed; the brackets ($\langle\rangle$) indicate averages over the simulation.

It is of interest, first, to compare the bulk value with the results of previous calculations. To do this, we assume that $\langle E_B' \rangle$ for the bulk class (denoted $\langle E_B^{bulk} \rangle$) is comparable to the binding energy in the pure solvent. The value of $\langle E_B^{bulk} \rangle$, −23.53 kcal/mol (see Table VI), includes the pair interaction energy of each molecule with all other water molecules (within range). To obtain an estimate of the total potential energy in the bulk fluid, we must divide by two, yielding −11.77 kcal/mol. This is somewhat more negative than the corresponding value for the ST2 model[19] at 30 °C (interpolated) of −10.20 kcal/mol, and the experimental internal potential energy of −9.86 kcal/mol. The difference between the two calculated results is due to a shift in hydrogen bond energies associated with the flexibility of the water molecules in the present simulation, as noted earlier.

The distributions in Figure 24 are seen to be continuous, unimodal curves, all having qualitatively the same shape; as expected, the polar and nonpolar groups have larger statistical fluctuations than the bulk water. The curves are all approximately symmetric and have nearly the same widths. There is no evidence in the figures for a multimodal distribution characteristic of a mixture of species with different binding energies.[40] However, there is a change in the curve position, with the nonpolar group shifted to larger (more negative) binding energies and the polar group to smaller (more positive) binding energies, relative to the bulk water. A small shift in the binding energy for the "nonpolar" molecules is in accord with the results of a simulation study of methane in water.[43] This is consistent with the values given in Table VI. The table shows that molecules in the nonpolar group are on the average energetically slightly stabilized with respect to the bulk by about 0.5 kcal/mol. Of the 0.5-kcal stabilization, only about 0.2 kcal/

Table VII. Average Solvent Hydrogen Bonding[a]

ϵ_{HB}, kcal/mol[b]	bulk[c]	nonpolar[c]	polar[c]
	A. Number Including Solute		
−2	4.72	4.08	4.20
−3	3.45	3.35	3.28
−4	2.73	2.75	2.57
	B. Number Excluding Solute		
−2	4.72	4.08	3.88
−3	3.45	3.35	3.08
−4	2.73	2.75	2.45
	C. Percent of Bonds to Solute		
−2	0.0	0.0	7.6
−3	0.0	0.0	6.1
−4	0.0	0.0	4.7

[a] Calculated from Figure 25. [b] Energy defining a hydrogen bond. [c] Classes as defined in text.

mol arises from the dipeptide-water interaction, and the rest (0.3 kcal/mol) from an increase in the water-water interaction. The "polar" molecules are energetically destabilized with respect to the bulk by about 1.65 kcal/mol. If the peptide interaction is excluded (e.g., $\langle E_B \rangle$), the relative destabilation of the polar molecules is considerably larger, 2.76 kcal/mol. The origin of these shifts will be considered subsequently.

The overall energy of solution includes contributions from both solvent-solute bonding ($\langle E_B{}^T \rangle$ vs. $\langle E_B' \rangle$) and from the changes in solvent-solute bonding ($\langle E_B' \rangle$ for each class). Using the data in Table VI, we can estimate the energy difference between the dipeptide under vacuum plus pure solvent, and the dipeptide solution (ΔE_c). In making this estimate, we assume that $\langle E_B{}^{bulk} \rangle$ can be equated with the binding energy in the pure solvent and, further, that the binding energies obtained from the current simulation, which involve an average for a typical solution structure ("V structure"),[44] correspond to those that would be obtained in a complete time average. Contributions to the energy of solution from changes in the internal solute potential energy are neglected. Unless the conformations in the gas phase and in solution differ significantly (in the simulation they do not; see Table I), such contributions are not expected to play a major role. Hence, we have

$$\Delta E_c = \tfrac{1}{2}\left[\sum_{i=1}^{195} (\langle E_B'(i) \rangle - \langle E_B{}^{bulk} \rangle)\right] + \langle E_B(D) \rangle$$

We obtain a value (see Table VI) of −6.72 kcal/mol. This can be thought of as resulting from a cancellation between the change in "nonpolar" water-water bonding (−3.4 kcal/mol), "polar" water-water bonding (+19.32 kcal/mol), and water-solute interaction (−22.64 kcal/mol). Of the latter contribution, −15.82 kcal/mol arises from the "polar" water molecules, −3.60 kcal/mol from the "nonpolar" molecules, and −3.22 kcal/mol from the "bulk". Hence, although the long-range interactions are individually small (−0.02 kcal/mol for each "bulk" molecule), they are not insignificant in the total. We emphasize that the exact values obtained depend on the hydrogen-bond potentials employed in this study.

The value of ΔE_c obtained above corresponds to the energy of solution from the gas phase, rather than the energy of solution with respect to the pure solute. Experimental work has focused principally on the latter[67] or on the heat of transfer of solutes from a solution in an organic solvent to that in water.[68] Although heats of solvation from the gas phase are available for a number of alkanes,[1] alcohols,[1] amines,[67] and acids,[67] there are no corresponding results for peptides. Hence, a comparison of ΔE_c with experimental data is not possible at the present time.

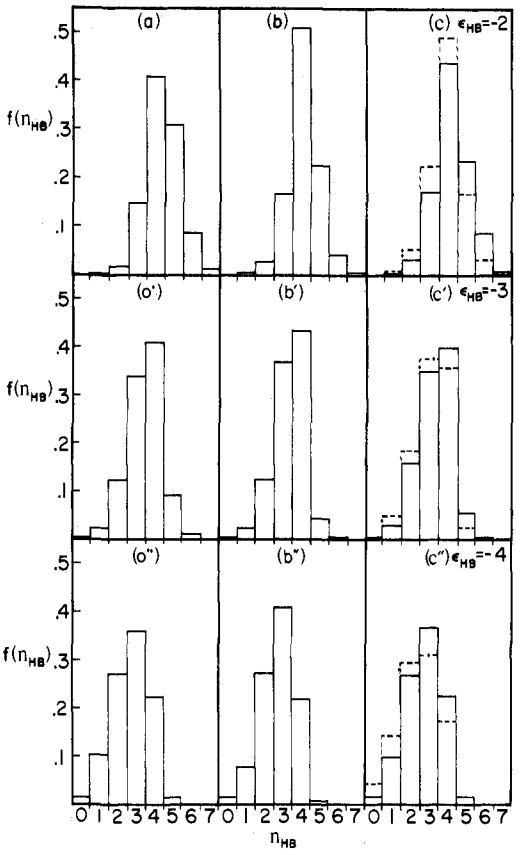


Figure 25. Fraction of water molecules participating in n_{HB} hydrogen bonds, according to the various bond definitions, ϵ_{HB}. The distributions are "bulk" (a, a', a"), "nonpolar" (b, b', b"), "polar" (c, c', c"). The dashed lines include only water-water bonds; the solid lines include also water-dipeptide bonds.

We now examine the number of hydrogen bonds formed by a typical water molecule in each of the three groups. Histograms presenting the average fraction, $f(n_{HB})$, of water molecules which participate in n_{HB} hydrogen bonds are shown in Figure 25. The middle set (top to bottom) corresponds to the energy criterion $\epsilon_{HB} = -3$ kcal/mol, which is the one suggested by the distribution functions for both water-water (Figure 23) and water-dipeptide (Figure 11) interactions, as discussed above (sections IVA and VA). To show that the analysis is not highly sensitive to the exact value of ϵ_{HB}, we include in Figure 25 (upper and lower sets) the corresponding results for $\epsilon_{HB} = -2$ and −4 kcal/mol. In each case, we label the "bulk" as a (a, a', a"), the nonpolar class as b (b, b', b"), and the polar class as c (c, c', c"). The average numbers of hydrogen bonds to each water molecule obtained from these distributions are tabulated in Table VII. For molecules in the polar class, the values indicated by dashed lines in Figure 25 are obtained by including only water-water bonds; the solid lines include, in addition, water-dipeptide bonds.

Focusing on the middle set, $\epsilon_{HB} = -3$ kcal/mol (a', b', c'), it is clear that the distributions indicated by solid lines are very similar. Thus, a typical water molecule participates in roughly the same number of hydrogen bonds in any of the three environments. For the polar group, we see that the bonds to the dipeptide contribute significantly; they tend to shift the peak in the distribution to higher values of n_{HB} and make the result more similar to that in the bulk than is that excluding these bonds. However, the shift is not large, since, as noted earlier (section IVA), only about 25% of the "polar" water molecules participate in strong bonds ($\epsilon < -3$ kcal/mol) to the dipeptide.

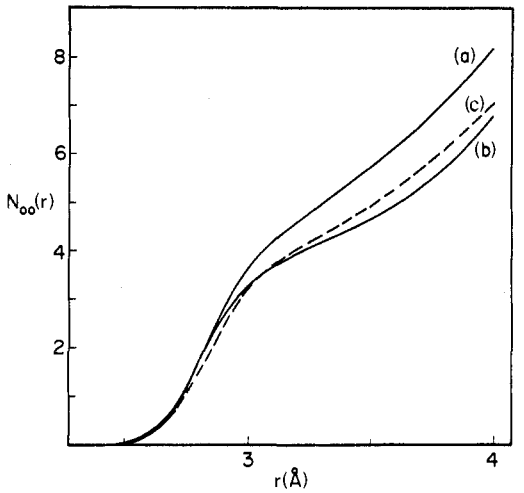


Figure 26. Number of solvent neighbors within a radius r of solvent molecules in each defined class (based on oxygen-oxygen distance). The curves are labeled as in Figure 15.

Comparing the nonpolar and bulk class, we note that the number of molecules contributing to the largest fraction ($f(4)$ for $\epsilon_{HB} = -3$ kcal/mol) is somewhat larger for the nonpolar class (b′) than for the bulk class (a′). It is also apparent from the figure that this increase occurs at the expense of the fraction associated with *higher* values of n_{HB}; i.e., the number of five bonded molecules ($f(5)$) is reduced in the nonpolar group relative to the bulk system. Corresponding results apply to the alternative energy criteria. Comparing the average number of bonds given in part A of Table VII for successively higher values of ϵ_{HB}, it is clear that this shift occurs due to loss of the weakest bonds; that is, the average number of hydrogen bonds in the bulk and nonpolar classes are nearly identical, except for the bond criterion $\epsilon_{HB} = -2$ kcal/mol. The latter value leads to a larger number of bonds in the bulk than in the nonpolar group. For the "nonpolar" solvent molecules, only the number of relatively weak interactions is reduced and these are, not unexpectedly, associated with five and higher bonded species. Hence, it is clear that the shift toward more negative binding energy for the "nonpolar" water molecules noted above (0.5 kcal/mol; see Table VI and Figure 24) results principally from an average strengthening of the water-water hydrogen bonds for this species (by ~0.1 kcal/mol, see Table V) relative to the bulk.

For the polar species, we find (see Table VII, part B) that a significant loss in solvent-solvent bonding occurs. This is expected since some molecules in the polar group participate in bonding with the solute. However, the loss in solvent-solvent bonding is not completely recovered by solvent-solute bonding as is evident from part A of the table. By considering the dependence of the number of bonds on ϵ_{HB}, we find that the fraction of the bonds between the "polar" species and the solute decreases as the criterion defining bonding, ϵ_{HB}, increases; that is, the solvent-solute interaction is generally weaker than the solvent-solvent interaction, as noted earlier in the comparison of Figures 11 and 23. From Table VII, we find for $\epsilon_{HB} = -3$ kcal/mol that molecules in the polar class lose 0.37 solvent-solvent bonds and gain 0.20 solvent-solute bonds. Taking the mean bond energies calculated above (based on the same bond energy criterion), namely, −5.2 kcal/mol for solvent-solvent interaction and −4.2 kcal/mol for solvent-solute interaction, we estimate the net loss in binding energy for the polar class to be 1.08 kcal/mol. The result is in reasonable agreement with the calculated value of 1.65 kcal/mol (see Table VI), if one considers the approximations involved in treating the bonding in the discrete manner. This analysis shows that the loss in "polar" binding energy with respect to the bulk is due to the partial sacrifice by the "polar" molecules of solvent-solvent bonds in favor of weaker solvent-solute interactions. The importance of this effect on the dynamics is discussed in section VI.

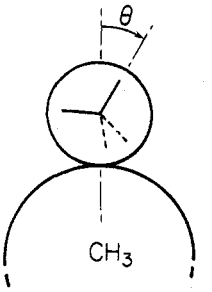


Figure 27. Schematic representation of water molecule orientation near a nonpolar ($-CH_3$) group.

The analysis presented here has shown that the three groups of water molecules (bulk, polar, and nonpolar) differ relatively little in their bonding energetics. In particular, the "nonpolar" water molecules have nearly the same number of strong hydrogen bonds as do those in the "bulk"; further, the hydrogen bonds for the nonpolar group are only very slightly stronger ($\sim 0.2 k_B T$) than are those in the bulk class. The "polar" solvent molecules are somewhat more weakly bonded than are those in "bulk"; they have nearly the same number of hydrogen bonds, but those to the solute are somewhat weaker than are those to other water molecules. These results show that the bonding energetics of the "average" configuration cannot explain the fact that the "polar" and "bulk" water have similar mobilities, while the "nonpolar" water molecules have decreased mobilities.

B. Geometric Factors. The presence of the dipeptide is expected to reduce the number of nearest-neighbor *solvent* molecules around any water molecule in the first solvation layer. In Figure 26, we show the calculated average number of *solvent* neighbors, $N_{OO}(r)$ (section IVA), for water molecules in each class. The number in the first shell around each solvent molecule is given by the value of $N_{OO}(r)$ at the first minimum in $g_{OO}(r)$ at 3.5 Å (see Figure 11). We find at this distance the values are 5.75 (bulk), 4.95 (polar), and 4.70 (nonpolar); that is, the water molecules in the bulk have roughly one more nearest-neighbor water molecule than those in the first solvation layer of the dipeptide. Since the average number of hydrogen bonds formed by molecules in the nonpolar group is essentially equal to that in the bulk (Table VII), there must be bonding among a significantly higher *fraction* of nearest-neighbor pairs; e.g., for the nonpolar group water molecules having four hydrogen bonds (see Figure 25), these bonds are distributed in an average sense among only 4.7 nearest neighbors rather than the 5.75 available in the bulk. Thus, relative to the bulk, a more developed bonding network exists among the water molecules in the immediate vicinity of a nonpolar group. For the solvent molecules near a polar group, the total number of neighbors capable of bonding is *not* reduced, since a hydrogen bond to the solute can take the place of that to a water molecule. Consequently the restrictions on the molecular arrangement of hydrogen bonds in the vicinity of a polar group are not qualitatively different from those present in the already relatively structured bulk fluid. As discussed in section VI, this difference in bonding capability is essential to the interpretation of the differences in dynamic behavior.

The reduction in the number of solvent neighbors for molecules in the nonpolar class does reduce the fraction of molecules with higher numbers of hydrogen bonds (n_{HB}). As seen

in Figure 25, the frequency of configurations with five or more hydrogen bonds is decreased, in accord with the lower average solvent coordination number. This shift is a result of geometric, rather than energetic, factors.

The formation of the same number of hydrogen bonds by the water molecules at the surface of a nonpolar solute as by those in bulk water entails significant restrictions on the orientation of the former.[3,69] From the schematic representation shown in Figure 27, we see that the maximum number of favorable water-water interactions can occur if none of the hydrogen atoms or lone-pair orbitals of the solvent molecule is directed toward the nonpolar group. If θ is the angle between an O-H bond direction (or O-lone pair direction) and the axis defined by the methyl carbon and solvent oxygen, the ideal orientation corresponds to $\theta = 0$. This molecular orientation is typical of crystalline clathrate hydrate compounds; the value $\theta = 0$ characterizes all of the water molecules in an ideal structure I clathrate geometry, that found in crystals containing the smaller nonpolar guest molecules[70] (e.g., CH_4). For other clathrate crystals, a number of water molecules can be orientated so that $\theta \neq 0$, but in each such case, the orientation is such that θ is far from 180°.[69,70] For example, the two OH bonds can be oriented so that the bisector of the H-O-H bond angle is directed at the carbon atom, and the OH bonds bridge the methyl group (the OH bond vectors correspond to $\theta \simeq 125°$ and the O-lone pair vectors to $\theta \simeq 55°$). In the simplest picture, optimum solvent hydrogen bonding can be achieved as long as one of the four bonding directions in the water molecule points away from the nonpolar surface. Consequently, we consider the four charges in our model water molecules as equivalent and compute a distribution for their orientations. If, for example, we included only the protons the accuracy of the distribution would be substantially reduced, because, over the time span of the simulation, molecular reorientation is not complete (see Figure 16). To examine the calculated distribution of orientations, the angle θ is redefined as the angle formed by *any* one of the four charges of the solvent molecule, the solvent-center of mass, and the methyl carbon atom. The lower curve in Figure 28 shows the calculated distribution of charge directions, averaged over the "nonpolar" solvent neighbors of the three methyl groups. (The distribution is integrally normalized to unity). The expected orientational bias of charges away from the nonpolar group is seen; that is, the distribution peaks around $\theta = 0$ ($\cos\theta = 1$) and has its minimum at $\theta = 180°$ ($\cos\theta = -1$). The probability of orientations with one charge directed away from the methyl group (i.e., at $\theta = 0$) is approximately three times that found with one toward it (i.e., at $\theta = 180°$). There is a broad secondary peak in the region $-0.1 > \cos\theta > -0.8$. This corresponds to the maximum at $\cos\theta = 0$ and is expected from the three other charges at $\cos\theta = -1/3$ for four tetrahedrally arranged charges; there is a corresponding minimum in the neighborhood of $\cos\theta = 1/3$. It is evident from the width of these maxima and minima that there is a significant dispersion in the orientations found in solution; contributions from nonoptimal orientations (e.g., that involving the two OH bonds bridging the methyl group) are clearly important. In particular, the distribution shown in Figure 28 is much less sharp than that for solvent molecules in an aqueous ionic solution.[71] Also, it is relatively diffuse compared to that found by examining the orientation of water molecules around any particular water molecule in the bulk fluid (ref 40, Figure 13).

To test the statistical significance of these results, we show, in the upper curve of Figure 28, the distribution obtained by choosing water molecules at random from the bulk; the same number (20) as in the nonpolar class is included in the sample. If the orientations of these molecules were completely uncorrelated with the positions of the methyl groups, and the configurational sampling in the simulation were sufficiently

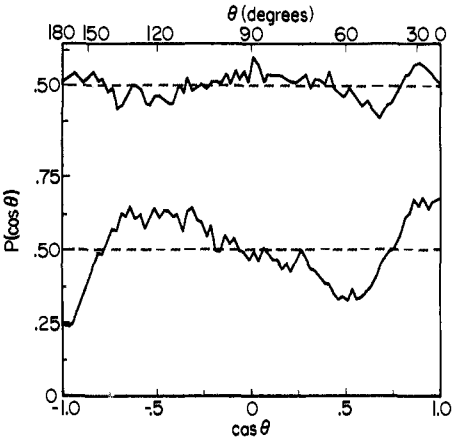


Figure 28. Distribution of orientations of water molecules near methyl groups (lower curve); θ as shown in Figure 27. All four solvent charges are included in the distribution, as described in the text. The upper curve shows a corresponding distribution obtained by choosing an equal number of water molecules randomly from the bulk.

complete, a uniform value of 0.5 would be expected. Although statistical fluctuations are present, the orientational effects in the lower curve are clear.

A careful study of dipolar relaxation of the NMR of methyl group protons by isotopic species of solvent water (D_2O, HDO, H_2O) makes possible an estimate of the orientation of the water molecules.[72] Although the quantitative conclusions depend on the details of the model, it is clear that there is an orientational preference of the type found here. The results also suggest that water protons are more likely to be pointing away from the methyl group (i.e., toward $\theta = 0$) than are the oxygen lone pairs. This point could be investigated in a more extensive simulation.

In Figure 29, we show a stereoscopic view of a typical configuration of water molecules surrounding the alanine dipeptide; the orientation of the dipeptide is that shown in Figure 1. The figure includes all 34 first-shell water molecules. The water molecules in the nonpolar class are shaded with dots; those in the polar class are not. Hydrogen bonds are indicated by dashed lines. The polar solvent molecule at the left center of the figure is bonded to the dipeptide NH group (left side of Figure 1) as indicated. We emphasized in section III that the division of the solvent into classes is necessarily an imprecise procedure. This imprecision is illustrated in Figure 29. Since the solvent molecules are not small on the scale of the solvation regions, molecules found near a boundary between regions may not behave in the manner attributed to the class as a group. For example, the "polar" molecule at the upper left of the alanine methyl side chain is clearly oriented with respect to the methyl group in the way ascribed above to "nonpolar" solvent molecules, although it is within the defined "polar" spatial region. We also see that significant bonding occurs between the "polar" and "nonpolar" classes (as it must also be between these and the bulk).

The nonpolar class (shaded) of solvent molecules is, in a general way, oriented so that no charge points at a CH_3 group. However, the orientations are not all characterized by the value $\theta = 0$ (see Figure 27); for example, the "nonpolar" molecules to the right of the right-hand terminal methyl group in Figure 29 are oriented in one alternative way mentioned above; two of the OH (O_q) bonds form a "fork" which bridges the CH_3 group. Such orientations have been suggested previously.[69]

A qualitative description of the observed geometry and intermolecular bonding pattern as solid-like and clathrate-like would be very misleading. The orientational *preference* apparent in Figures 28 and 29 *is* clathrate-like. The nonpolar

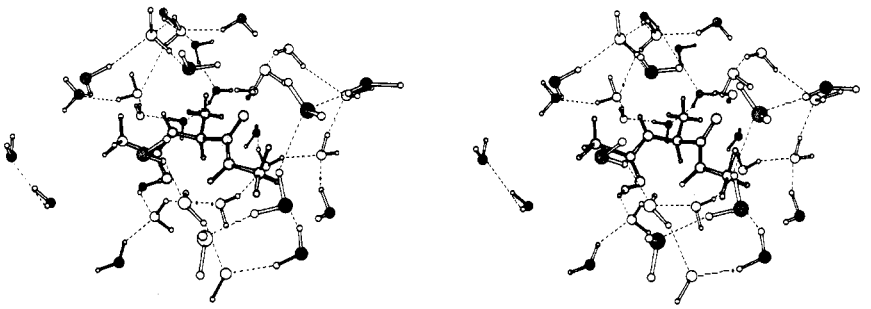


Figure 29. Typical configuration of water molecules around the alanine dipeptide. "Nonpolar" solvent molecules are shaded with dots; hydrogen bonds are indicated by dashed lines.

solute characteristic (namely, the inability to form hydrogen bonds) determining the special solvent behavior near nonpolar groups in solution is apparently the same as that determining the structure of crystalline clathrate hydrates. However, the intermolecular bonding in solution is very different from that expected in a solid; in solution, it is transient and highly disordered. As we have seen earlier, the bond strengths and number of bonds are bulk-like rather than solid-like; that is, in ice each molecule participates in four hydrogen bonds[73] (i.e., there is a sharp peak at $n_{HB} = 4$, Figure 25) and the decreased hydrogen bond distortion,[73] relative to the liquid,[40] is expected to lead to bonds that are distinctly stronger than in the bulk fluid. By contrast, we found only very small changes in bond energies near nonpolar groups, as compared to the bulk and no significant change in the bond number distribution. A description of liquid water structure in accord with these results is given by Stillinger[69] and by Franks.[3]

VI. Discussion

In this section, we discuss in some detail certain thermodynamic properties of the solution (excess enthalpy, entropy, and the heat capacity), and the relation between the results obtained for the solvent dynamics and its structure. We have seen that the water molecules in the first solvation layer of the dipeptide can be divided into two classes—those associated with the nonpolar groups and those associated with the polar groups—and that the members of the two classes have significantly different structural and dynamic properties. The rates of translational diffusion and rotational reorientation are decreased near the nonpolar groups, while near polar dipeptide groups they are similar to the "bulk" fluid. Further, the dynamics of the "bulk" are consistent with those found in previous simulations of the pure solvent.[19] Since the "bulk" fluid in the current simulation corresponds to the second molecular layer of water molecules around the solute, we conclude that, as monitored by dynamic properties, the effective solvent-solute forces are short ranged, being limited in their influence to the molecules in the first solvation layer around each functional group. This conclusion is in agreement with the substantial experimental evidence for locality obtained in thermodynamic studies on mixed functional solutes.[1,3,5]

A concomitant analysis of the average properties of the solvent has shown that the solvent molecules near nonpolar solute groups participate in nearly the same number of strong hydrogen bonds ($\epsilon \leq -3$ kcal/mol) as do those in the bulk solvent and that these bonds are only slightly stronger (~0.1 kcal/mol) than are those in the bulk. Maintenance of the same number of hydrogen bonds, in the presence of fewer water neighbors, is achieved by significant orientational restrictions on the "nonpolar" water molecules. The solvent molecules near the polar groups are less strongly bonded than are those in the bulk; these molecules partially replace water-water hydrogen bonds by water-peptide hydrogen bonds, which are somewhat weaker in the solution than are the solvent-solvent bonds. For the "polar" groups, the average number of hydrogen bonds is also found to be decreased slightly with respect to the bulk.

The observed mobility of molecules near polar groups does not conflict with the observation of electron density peaks in X-ray crystallographic studies of proteins at preferred positions of polar group-water association.[17] Such studies show a substantially increased probability of occupation of certain sites, but do not determine the rate of molecular exchange. Further, the range of calculated mobilities for the various classes of solvent molecules is small compared to the differences that have been suggested for protein solutions.[16] Hence, the inference that all solvent molecules in the current simulation fall into the so-called type I category ($\tau_r \lesssim 10^{-11}$ s),[16] mentioned in the Introduction, is in accord with the results.

It must be emphasized that the diffusive motion of water molecules is slow on the time scale considered in this work (i.e., 1.5 ps). Thus, the results obtained here pertain to a typical solution structure rather than to the true (infinite time) average. The simulation corresponds to a vibrationally averaged "V structure", as discussed by Eisenberg and Kauzmann,[44] in that averages obtained include contributions from relatively small fluctuations in the solvent and solute structure, and not from configurations differing from the initial conditions by large molecular displacements (i.e., of the order of a molecular diameter). It cannot be excluded, therefore, that other "V structures" would show different time-independent and dynamic behavior.

A. Thermodynamic Properties. We consider three thermodynamic quantities that characterize the solution. They are the excess enthalpy of solution, the excess entropy of solution, and the change in heat capacity accompanying solution. (The excess quantities are the differences between the observed quantity and that expected for the corresponding ideal system.)

We have described the energetics of the solution in section VA and shown that the energy change accompanying the process of dissolving the (gaseous) dipeptide in water involves a balance between solvent-solvent bonding and solvent-solute bonding. In the simulation, the net estimated energy of solution (−6.7 kcal/mol) results from a balance between a large attractive dipeptide-water interaction (−22.6 kcal/mol) and the large decrease in bonding (+19.3 kcal/mol) among the water molecules near the polar groups (as compared to the bulk). The small increase in the strength of water-water bonds near the nonpolar groups is also important; its contribution of −3.4 kcal/mol accounts for one-half of the net energy of solvation.

The experimentally observed negative excess entropy of solution for nonpolar species has been associated with an increase in "order" or "structure" in the solvent. In section V, we showed that this "structuring" arises from the maintenance of bulk-like hydrogen bonding for each of the water molecules near nonpolar groups. Since nonpolar groups cannot participate in hydrogen bonding, the number of neighboring mole-

cules capable of participating in such solvent-solvent bonds is reduced; an increased fraction of neighboring molecular pairs must be bonded to each other. Consequently, there are important configurational restrictions on these water molecules. Since bulk water is itself a highly structured liquid, the difference between the structure in bulk water and that near nonpolar surfaces is one of degree; that is, molecules near a nonpolar solute have their positions and orientations even more highly correlated with one another than does an average pair of water molecules in the bulk. Such an increased correlation near nonpolar solutes has been postulated previously as a primary characteristic of "structuring" due to nonpolar solutes.[3] It is manifested in the present simulation by a significant orientational preference for the water molecules near the nonpolar surface. This might appear to suggest that the entropy loss could be calculated by simply counting the reduction in permitted single molecule orientations.[74] However, since the orientational bias results from requirements of intermolecular bonding, the concomitant loss in configurational freedom, and the resulting negative entropy change, can be treated correctly only by considering the water molecules as a group; that is, it is necessary to include the requirements that such specific orientations of a given water molecule place on its molecular neighbors.

Although, in principle, entropies can be calculated using simulation methods,[75] the current simulation does not permit a direct determination of a numerical value for the entropy loss associated with the increased water structure. The interpretation we have given suggests that an estimate might be obtained from appropriate extensions of lattice or cell theories for water,[76] since these approaches attempt to include the configurational degrees of freedom that we find to be essential to the structuring phenomena. Such methods are discussed further below (section VIC).

Configurational correlations are relatively long ranged in water (about 8 Å).[41] We note, therefore, that, although the changes in dynamic properties are apparently limited to the first solvation shell, we cannot conclude from this study that the changes in entropy can be accounted for by consideration of the first shell alone. Exploratory Monte Carlo studies of nonpolar species in water[77] have suggested that the range of the potential of mean force between two nonpolar spheres is limited to a distance which permits less than two solvent layers between solute particles. This result is consistent with an entropic effect which is restricted to the first solvation layer around each solute particle. Additional support for this view is provided by the high degree of correlation found between the value of the excess entropies of solution and the molecular surface area of nonpolar solutes.[2]

In section V it was shown that water molecules near polar groups (C=O, N—H) need not be configurationally restricted relative to the bulk, since the polar group is capable of participating in hydrogen bonding. This suggests that the polar groups do not lead to *large* changes in solvent entropy in the solution. It is clear, however, that such a conclusion is a qualitative one, in that quantitative results for the entropy must depend on the hydrogen-bonding ability of the polar group and its effect on the configurational freedom of the solvent molecules in its vicinity.

For any fluid, the heat capacity arises from the ability of the various degrees of freedom to absorb potential and kinetic energy. In water, the intermolecular interactions (primarily the hydrogen bonds) are strong and provide a very important "sink" for potential energy. The hindered translational and rotational (librational) contributions associated with the interactions are not sufficient to account for the large heat capacity of liquid water.[78] There is an additional important contribution (called "configurational" heat capacity) associated with structural changes in the fluid, and usually assigned to the distortion or breaking of hydrogen bonds with increasing temperature.

As mentioned in the Introduction, an increase in the heat capacity, ΔC_p, is associated with the process of solution of species containing nonpolar groups.[1,5] A corresponding increment from polar groups is absent for solutions of species such as peptides.[79] For methane, $\Delta C_p°$ is[1] 55 cal/mol·deg; for the lower *n*-alkyl alcohols, the increment in $\Delta C_p°$ for addition of a $-CH_2-$ group is about[1] 15 cal/mol·deg. The origin of this heat capacity is uncertain, but it has been assumed to be associated with the hydrogen bonds among the water molecules at the nonpolar surface.[1] Since the bonding of these molecules takes place at the expense of reduced entropy, it is reasonable that these bonds should be more effectively disrupted by an increase in temperature than are those in the bulk; that is, the positive entropic contribution to the free energy of solution is increasingly important as the temperature is raised.

Although no attempt has been made to study this contribution to the heat capacity by molecular dynamics, the difference between the present results for 303 K and the preliminary solution simulation[18] at ~370 K is suggestive. The latter showed that, in contrast to the 303 K behavior, there is apparently little, if any, difference in the mobility of the water molecules in the first solvation layer as compared to those in the bulk. This indicates that at the higher temperature, the molecules in this layer are not configurationally restricted in comparison to those in the bulk. Consequently, the reduction in configurational restriction must be larger for the water molecules near the solute than for those in the bulk, since the mobilities at the higher temperature are comparable, while at the lower temperature they are not.

The concept that the anomalously large value of ΔC_p associated with the process of solution of nonpolar species corresponds to the "melting" of solvent "structure" originated with the discussion of Frank and Evans.[80] Although these authors were careful to point out that the term "ice-like" as applied to this "structure" should be interpreted very loosely, it is worthwhile to emphasize that the observed contributions to the solution heat capacity can be accounted for by a temperature-dependent shift in the hydrogen bond energies that is quite small compared to the total bond energy. To illustrate this, we consider the increment in the heat capacity upon dissolving ethane in water,[1] +66 cal/deg·mol at 25 °C. If we assume that the number of "nonpolar" water molecules surrounding ethane is 14 (i.e., seven for each CH_3 as found in the current work; the specific value is not crucial), the total heat capacity increment is accounted for by an increment of 0.005 kcal/deg·mol for each solvent molecule. Assuming that this resides principally in the bond strengths (rather than the number of bonds), a correspondingly small relative change in individual hydrogen bond energies is required. Such changes are in addition to those required to account for the large total heat capacity of water; the configurational contribution is about 0.010 kcal/mol·deg.[78] Thus, the estimated temperature dependence of the bond energies for "nonpolar" molecules is 50% greater than that in the bulk at 25 °C. Correspondingly, the net change in bond energy associated with a 5 °C increment in the temperature is comparable to that in the bulk associated with an increment of 7.5 °C; since such a change in the bulk would not normally be referred to as "melting", the application of this term to the nonpolar interactions does not seem appropriate. We point out that the positive increment in the heat capacity need not be associated with stronger hydrogen bonds near the nonpolar species; the heat capacity reflects only the temperature derivative of the internal energy and not its absolute value.

The relation of the small temperature variation in bond energies to changes in solvent structure and dynamics is not clear. To determine what is happening would require analyses

of dynamical simulations and experimental data for the thermodynamic parameters at a series of temperatures.

B. Microscopic Interpretation of the Dynamic Solvent Properties. A complete interpretation of the origin of *differences* in dynamic behavior between water molecules near the solute and far from it requires a satisfactory description of the microscopic motion in the bulk fluid. For water, such a description is not available, although a number of models have been presented. Based on dielectric relaxation measurements, it has been suggested that water molecules rotate as individual entities by breaking one hydrogen bond,[81] as clusters by breaking several hydrogen bonds, or both.[82] Studies employing NMR[61] and depolarized Rayleigh scattering[52] have been interpreted as suggesting that water molecules reorient through a sequence of finite jumps, each involving the breaking of a hydrogen bond. Based on molecular dynamics simulation it has been suggested that such motion occurs through "jumps", involving rapid changes in orientation which are of short duration,[83] or by relatively continuous variation in position involving no such jumps.[41,84] Further, based on the very small dispersion in dielectric relaxation times for water (suggesting a lack of dependence of this time on details of the local molecular environment), it has been proposed that water molecules reorient in a cooperative manner.[85]

A priori, the differences in dynamic behavior between the solvent classes could have been due to a change in bonding energetics or to a change in the configurational distribution of water molecules near nonpolar groups. In the first case, the reduced mobility results from the increased confinement produced by deeper local minima; that is, there is an increased activation energy for the molecular motion. In the second alternative, the average local environment of a water molecule near a nonpolar group is not energetically significantly different from typical local configurations which occur in the bulk liquid. However, there are a significant number of additional molecular configurations which are of comparable energy in the bulk fluid, but which are of prohibitively high energy for the water molecules which neighbor the nonpolar group; for example, such configurations could require the breaking of hydrogen bonds. The reduced mobility is then a manifestation of the reduced number of low-energy paths available for molecular rearrangements. This would be expected to result in an increased entropy of activation. From the analysis of section V, we have seen that it is the second of the two extreme possibilities that dominates the molecular dynamics results. In what follows, we interpret our observations in terms of a simplified picture that is in accord with molecular dynamics studies of the bulk solvent.[41,84]

The results obtained in the current work are consistent with the following schematic mechanism. We suppose that the translation and rotation of a molecule occur by more or less continuous motion; that is, hydrogen bonds between particular neighboring molecules are gradually strained and simultaneously hydrogen bonds are gradually formed to new molecular neighbors. Such a motion is expected to correspond to a low-energy path relative to one in which a bond is completely broken. The postulated bond exchange, which is consistent with molecular dynamics data,[41] is most easily described in the context of molecular reorientation. The resulting motion has certain "gear-like aspects, in that the reorientation of the solvent molecule involves a significant degree of correlation between the rotation of a given molecule and its hydrogen-bonded neighbors. For translational motion, the repulsive interactions, dominant in ordinary fluids,[30] must play an important role; however, a corresponding low-energy path involving rearrangement of hydrogen bonds is expected to contribute significantly to the ease of translational motion.

In the present framework, the variation in mobility among the solvent classes is attributed to differences in the capability of the solvent molecules to carry out such a coordinated rearrangement of bonding. For the "polar" molecules, the peptide polar groups can participate in bonding. Consequently, the replacement of a water-water hydrogen bond by one involving the peptide group should not inhibit the required rearrangements and, hence, have a small effect on the dynamics. For example, a solvent molecule near a carbonyl oxygen atom could rotate or translate by exchanging a solvent-solvent bond for a solvent-solute bond.

For the molecules near nonpolar groups, such bonding rearrangements are clearly restricted. Referring to Figure 27 we see that a molecule with four hydrogen bonds and an orientation near $\theta = 0$ cannot rotate far away from $\theta = 0$ without the disruption of a hydrogen bond; the nonpolar methyl group is distinguished from the polar groups precisely by its inability to participate in a hydrogen bond. Possible low-energy paths for rotational reorientation include motions in which θ remains near zero, and those in which the molecule reorients so as to maintain two (rather than three) bonding groups bridging the methyl group (such orientations are discussed above; see section VB). The latter orientation does permit the maintenance of four hydrogen bonds,[69] although, for the two "bridging" charges, it is likely that the bonds formed would typically be somewhat more strained than for the former ideal reorientation. For translational motion, the allowed rearrangement of bonding is similarly hindered. The facile translation of a "nonpolar" water molecule is assumed to require the replacement of a solvent-solvent bond by another solvent-solvent bond; the ways in which such bond exchange can occur are limited since the "nonpolar" water is not surrounded on all sides by possible bonding partners, but must move so that no bonding group on a water molecule is directed toward the nonpolar (nonbonding) group of the solute.

For a more detailed description of the modes and rates of translational and rotational motion, a more extensive study, including evaluation of the energetics and probabilities of various motional paths, would be required.

C. Models for Aqueous Solutions. One goal of a molecular dynamics simulation of the kind considered here is to ascertain the essential elements determining the characteristic behavior of the studied system. Such elements can, in principle, be incorporated into models that can then be used to study related systems. Further, models formulated to include only particular aspects of the complete molecular description are expected to be valuable objects for study, in that the relative importance of these elements for a satisfactory account of system properties can be investigated in detail. For aqueous solutions of nonelectrolytes, properties of particular interest include thermodynamic quantities, solvent structure (e.g., as characterized by molecular position distribution functions), and the molecular dynamics of solvent molecules.

In this subsection, we consider models for aqueous solutions. We discuss several alternative approaches in order to assess their validity and utility in light of the present analysis.

Mixture and Continuum Models. The most widely considered models of water can be classified as mixture models.[44,86,87] In such models, water is pictured as a mixture of two or more species that are differentiated by a particular property. Among the properties considered are the local solvent density[87] and the number of hydrogen bonds to each water molecule.[88,89] Continuum models[90] are distinguished from mixture models in that the solvent is characterized by a continuously varying parameter; such models can be considered as mixture models involving an infinite number of components. For example, solvent molecules could be characterized by binding energy (see section VA); in a continuum picture, the energy is a continuous variable, while in a mixture model one could require the binding energy to take on one of, say, five discrete values, depending on the number of hydrogen bonds (i.e., 0, 1, . . . , 4).

Although, in this example, the continuous picture is the most accurate *physical* description (as is clear from the energy distributions discussed in section V), it is possible that a mixture description can be useful as a model.

It is a prerequisite for any model of a solution that it be based on a satisfactory model of the pure solvent. The results of molecular dynamics simulation,[19,41] as well as experiment,[91] for pure water have provided substantial evidence that pure liquid water is not a mixture of "interstitial", essentially unbonded, molecules and highly bonded molecules. Since the solution models of Frank and Quist,[92] Frank and Franks,[93] and Mikhailov[94] assume such a model of the bulk fluid, it is difficult to justify their application to a solution and we shall not consider them here.

More generally, a description in terms of the relative populations of molecules with different numbers of hydrogen bonds (0, 1, 2 . . .) is of some interest. The distributions of the number of hydrogen bonds obtained from pure water simulations[41] (corresponding to Figure 25) suggest that a theoretical scheme based on the classification of water molecules into such separate species is not, a priori, an unreasonable approach to a model of the bulk fluid. It may well be useful, in particular, for evaluating equilibrium properties. However, the current simulation shows that a model of nonpolar hydration which postulates that the relative populations of bonded species are the basic variables is not able to elucidate the origin of the structuring phenomenon near nonpolar species; we find that the distributions are not substantially changed by the proximity of the solute. In particular, the shifts (relative to bulk water) in the populations of zero through four bonded species in solutions of nonpolar solutes that are an essential element of the Nemethy-Scheraga model[89] are not in accord with our results.

We note that the exclusion of five and higher bonded species in such mixture models of water is satisfactory for the qualitative description of the liquid structure since the fraction of such species is small for a reasonable choice of the bond strength criterion (<10% for $\epsilon_{HB} = -3$ kcal/mol; see Figure 25). In describing the distribution of these populations as very similar near the solute and in the bulk, we have discounted the difference in this minor fraction (see section VA). Since the contributions from zero- and one-bonded species are comparable to those from five-bonded species, the present results suggest that a mixture model that excludes zero-, one-, and five-bonded molecules is a reasonable one. The species included in such a model correspond to those of the Weres-Rice cell model.[76]

A related approach, involving a less detailed characterization of the solvent, attempts to describe it in terms of two species: a "bulky", low-density phase and a "dense" phase.[3] This division arises from the concept that the "structured" water near nonpolar solutes is "ice-like"[80] or "clathrate-like",[92] and hence less dense than the less ordered bulk water. However, the present analysis (i.e., the examination of the bond energies and numbers of hydrogen bonds; section VA) does not suggest solid-like trends that would be associated with the formation of such a "bulky" phase near nonpolar groups. In this context, we stress that the observed decrease in solvent coordination number ($N_{OO}(r)$, Figure 27) near the solute is a geometric effect (i.e., it is due to the volume excluded by the solute) and does not reflect an increase in "bulky", lower coordinate species.

Although, as mentioned above, the continuum picture is apparently a more accurate view of liquid water, such an approach has been considered in detail in only one case. Pople[95] regarded each water molecule as participating in four hydrogen bonds with each bond potential independent of all others; the potential was determined independently by the deviation of the proton *or* lone-pair direction from the O-O line between two solvent molecules. The bond potential, and hence the binding energy, varied continuously. Such a picture is consistent with bulk water simulation results,[41] in that one observes a generally random, deformed hydrogen bond network among molecules. Further, the model was found to be reasonably successful in the description of bulk water, despite the crude assumption that all bonds are independent of one another.

More recently, a generalization of the Pople model has been applied to aqueous solutions.[96] In the generalization, the correlation between lone-pair and proton directions in a hydrogen bond, lacking in the Pople model, is incorporated by including the average orientation of neighboring molecules around a central molecule. By assuming that the average solvent molecular orientation is a continuous function of the spatial coordinate in the fluid (i.e., a field), the model is cast in the form of a Landau theory.[97] Of most interest here is that the effect of a nonpolar solute is introduced as a boundary condition on the solvent orientation at the surface of the solute particle. This assumption is in accord with the present analysis; that is, we have concluded that the restriction on orientations at the solute surface is of primary importance. Further, in the model, the correlation length describing the propagation of the effect of the boundary condition into the fluid is found to be of the order of one solvent molecular diameter; i.e., it is a local effect. Hence, the concepts involved in the Pople model and this generalization appear to be worthy of further study.

Other Equilibrium Theories. Some effort has been applied to the implementation of lattice and cell theories of liquids to pure water.[76,83] Such methods involve representation of the liquid by molecules with centers distributed among a set of fixed sites (lattice theories) or within spatial cells (cell theories). In either approach, molecular configurations are specified by the assignment of the location of molecular centers and an assignment of molecular orientations among a restricted set of fixed orientations. Further, the energy of a given configuration is typically obtained by including only nearest-neighbor interactions. The details of such models and their applications to pure water are reviewed elsewhere.[76]

Although some objections[41,84] have been raised to a lattice-like view of liquid water on the grounds that the liquid forms a highly deformed bonding network, such an approach includes elements of the liquid structure of particular importance for aqueous solutions. We have emphasized the importance of the configurational restrictions on solvent molecules near the nonpolar (nonbonding) solute groups and the cooperativity involved in these restrictions. By an appropriate extension of lattice or cell theories (e.g., by the introduction of fixed vacancies) it seems that such effects could be investigated. Care must be exercised, however, in the formulation of such a model, so that the underlying lattice structure does not unduly influence the preferred solvent structure near the "solute"

VII. Conclusion

By analysis of a molecular dynamics simulation of an aqueous solution of an alanine dipeptide, we have been able to elucidate several aspects of the characteristic dynamic behavior and the structural origins of solute and solvent motion.

As in our preliminary study of the dipeptide solution,[18] we find that the local solute structure and the dynamics of its high-frequency modes are, within the accuracy of the calculation, essentially unaffected by the presence of the solvent on the short time scale examined (~1 ps). We see somewhat more flexibility of the dihedral angle ψ in solution than under vacuum, but the mean value differs relatively little. Only the motions associated with the lowest frequency modes and the lightest masses show damped kinetic behavior. These dynamic properties are in qualitative agreement with those expected, based on the simplest descriptions of viscous damping.

Study of kinetic and structural properties of the solvent has resulted in a consistent, qualitative picture of solvation. The significant influence of the solute on the dynamic properties of the water molecules is limited to a first solvation layer. Further, the influence of individual functional groups is localized. The solvent "structure" which is induced in the vicinity of nonpolar groups, and the concomitant configurational confinement of water molecules, is a result of the maintenance of bulk-like intermolecular hydrogen bonding within the constraint of a reduced number of neighbors capable of participating in bonds. The nonpolar groups are incapable of participating in bonding and it is this property which distinguishes them from the polar groups. We interpret the decreased mobility of the solvent near nonpolar groups as arising primarily from configurational (entropic) barriers rather than energetic barriers. Although certain geometrical aspects of the system are "clathrate-like", the term is misleading in its implications with respect to the number and strength of intermolecular bonds. The bulk-like dynamics observed for solvent near the solute polar groups is consistent with the interpretation that these polar groups interact with neighboring water molecules in the same way as do other water molecules.

Of particular interest for future work is a detailed examination of the temperature dependence of the quantities considered in this work, such as bond energies and mobilities. In addition, a detailed study of intermolecular spatial correlations by use of a simulation of higher statistical accuracy would be very desirable. From such information, a quantitative description of the solvation structure and its implications for solvent dynamics can be obtained.

Note Added in Proof. After the submission of this article, there was published a report of a molecular dynamics simulation of two neon-like spheres in water [A. Geiger, A. Rahman, and F. H. Stillinger, *J. Chem. Phys.*, **70**, 263 (1979)], in which corresponding values for many of the quantities studied in the present work are given. Although there are certain quantitative differences, a comparison of the two sets of results supports the conclusion that the solvent near the dipeptide nonpolar (methyl) groups behaves in the same way as that near a simple nonpolar sphere.

Acknowledgments. We wish to thank Aneesur Rahman for his invaluable help in the initial stages of this study. We thank Carl Moser and the Centre Europeen de Calcul Atomique et Moleculaire, Orsay, France, for hospitality and support at the 1976 summer workshop where the solution study was begun We would also like to thank J. Andrew McCammon and Peter G. Wolynes for many valuable discussions and Robert Bryant, John Edsall, and Lawrence Pratt for comments on the manuscript. The stereopicture was prepared by James Patrician with a program written by John Ramsdell. This work was supported in part by grants from the National Science Foundation and the National Institutes of Health.

References and Notes

(1) F. Franks and D. S. Reid in "Water, a Comprehensive Treatise," Vol. 2, F. Franks, Ed., Plenum Press, New York, 1975, Chapter 5. This excellent review contains numerous references.
(2) C. Chothia, *Nature (London)*, **248**, 338 (1974).
(3) F. Franks in ref 1, Vol. 4, Chapter 1.
(4) O. W. Howarth, *J. Chem. Soc., Faraday Trans. 1*, **71**, 2303 (1975).
(5) K. P. Prasad and J. C. Ahluwalia, *J. Solution Chem.*, **5**, 491 (1976).
(6) A. Holtzer and M. F. Emerson, *J. Phys. Chem.*, **73**, 26 (1969).
(7) M. D. Zeidler in ref 1, Chapter 10.
(8) H. G. Hertz, *Prog. Nucl. Magn. Reson. Spectrosc.*, **3**, 159 (1967).
(9) F. Franks, J. Ravenhill, P. A. Egelstaff, and D. I. Page, *Proc. R. Soc. London, Ser. A*, **319**, 189 (1970).
(10) A. H. Narten and S. Lindenbaum, *J. Chem. Phys.*, **51**, 1108 (1969).
(11) K. D. Kopple and A. Go, *Biopolymers*, **15**, 1701 (1976).
(12) See, for example, J. D. Glickson, W. D. Cunningham, and G. R. Marshall, *Biochemistry*, **12**, 3684 (1973).
(13) M. Avignon, C. Garrigou, and P. Botherel, *Biopolymers*, **12**, 1651 (1973).
(14) A. Laubereau and W. Kaiser, *Annu. Rev. Phys. Chem.*, **26**, 83 (1976).
(15) W. Kauzmann, *Adv. Phys. Chem.*, **14**, 1 (1959).
(16) R. Cooke and I. D. Kuntz, *Annu. Rev. Biophys. Bioeng.*, **3**, 95 (1974).
(17) H. J. C. Berendsen in ref 1, Vol. 5, Chapter 6.
(18) P. J. Rossky, M. Karplus, and A. Rahman, *Biopolymers*, to be published.
(19) F. H. Stillinger and A. Rahman, *J. Chem. Phys.*, **60**, 1545 (1974).
(20) M. Avignon, P. V. Huong, and J. Lascombe, *Biopolymers*, **8**, 69 (1969).
(21) IUPAC–IUB Commission on Biochemical Nomenclature, *Biochemistry*, **9**, 3471 (1970).
(22) E. L. Ellel, N. C. Allinger, S. J. Angyal, and G. A. Morrison, "Conformational Analysis", Wiley, New York, 1965.
(23) B. R. Gelin, Thesis, Harvard University, 1976.
(24) H. L. Lemberg and F. H. Stillinger, *J. Chem. Phys.*, **62**, 1677 (1975); A. Rahman, F. H. Stillinger, and H. L. Lemberg, *ibid.*, **63**, 5223 (1975). The precise values of potential parameters used in the current study appear in ref 18.
(25) L. Bøje and A. Hvidt, *J. Chem. Thermodyn.*, **3**, 663 (1971).
(26) W. W. Wood in "Physics of Simple Liquids", H. N. V. Temperley, J. S. Rowlinson, and G. S. Rushbrooke, Eds., American Elsevier, New York, 1968.
(27) C. W. Gear, Argonne National Laboratory Report No. ANL-7126, 1966.
(28) A. Allerhand, D. Dodrell, and R. Komoroski, *J. Chem. Phys.*, **55**, 189 (1971).
(29) R. Zwanzig, *Annu. Rev. Phys. Chem.*, **16**, 67 (1965); B. J. Berne in "Physical Chemistry, an Advanced Treatise", Vol. VIII B, D. Henderson, Ed., Academic Press, New York, 1971, Chapter 9.
(30) D. Chandler, *Acc. Chem. Res.*, **7**, 246 (1974).
(31) S. Chandrasekhar, *Rev. Mod. Phys.*, **15**, 1 (1943).
(32) J. A. McCammon and P. G. Wolynes, *J. Chem. Phys.*, **66**, 1452 (1977).
(33) B. J. Berne and R. Pecora, "Dynamic Light Scattering", Wiley, New York, 1976.
(34) Reference 33, p 144 ff.
(35) W. A. P. Luck in ref 1, Chapter 4.
(36) L. C. Allen, *J. Am. Chem. Soc.*, **97**, 6921 (1975).
(37) A. Johansson, P. Kollman, S. Rothenberg, and J. McKelvey, *J. Am. Chem. Soc.*, **96**, 3794 (1974).
(38) A. Rahman and F. H. Stillinger, *J. Am. Chem. Soc.*, **95**, 7943 (1973).
(39) R. Balescu, "Equilibrium and Non-Equilibrium Statistical Mechanics", Wiley, New York, 1975, Chapter 7.
(40) A. Ben-Naim, *Trans. Faraday. Soc.*, **66**, 2749 (1970).
(41) A. Rahman and F. H. Stillinger, *J. Chem. Phys.*, **55**, 3336 (1971).
(42) L. Verlet, *Phys. Rev.*, **159**, 98 (1967); **165**, 201 (1968); A. Rahman, *ibid.*, **136**, A405 (1964).
(43) J. C. Owicki and H. A. Scheraga, *J. Am. Chem. Soc.*, **99**, 7403, 7413 (1977); see also S. Swaminathan, S. W. Harrison, and D. L. Beveridge, to be published.
(44) D. Eisenberg and W. Kauzmann, "The Structure and Properties of Water" Oxford University Press, New York, 1969.
(45) W. A. Steele, *Adv. Chem. Phys.*, **34**, 1 (1976).
(46) D. I. Page in ref 1, Vol. 1, Chapter 9.
(47) J. A. Glasel in ref 1, Vol. 1, Chapter 6.
(48) J. B. Hasted in ref 1, Vol. 1, Chapter 7.
(49) J. L. Salefran, G. Delbos, C. Marzat, and A. M. Bottreau, *Adv. Mol. Relaxation Interaction Processes*, **10**, 35 (1977).
(50) G. E. Walrafen in ref 1, Vol. 1, Chapter 5.
(51) M. J. Blandamer and M. F. Fox in ref 1, Chapter 8.
(52) C. J. Montrose, J. A. Bucaro, J. Marshall-Coakley, and T. A. Litovitz, *J. Chem. Phys.*, **60**, 5025 (1974).
(53) C. M. Davis and J. Jarzynski in ref 1, Vol. 1, Chapter 12.
(54) M. J. Blandamer in ref 1, Chapter 9.
(55) R. Mills, *J. Phys. Chem.*, **77**, 685 (1973).
(56) Reference 44, p 218.
(57) J. M. Deutch, *Annu. Rev. Phys. Chem.*, **24**, 319 (1973).
(58) J. G. Powles, *J. Chem. Phys.*, **21**, 633 (1953).
(59) T. W. Nee and R. Zwanzig, *J. Chem. Phys.*, **52**, 6353 (1970).
(60) Reference 44, p 207.
(61) K. Krynicki, *Physica (Utrecht)*, **32**, 167 (1966).
(62) J. G. Powles, *Ber. Bunsenges. Phys. Chem.*, **80**, 259 (1976).
(63) J. G. Powles and G. Rickayza, *Mol. Phys.*, **33**, 207 (1977).
(64) Reference 45, p 46.
(65) Reference 44, Table 4.10.
(66) F. Cavatorta, M. P. Fontana, and A. Vecli, *J. Chem. Phys.*, **65**, 3635 (1976).
(67) J. Konicek and I. Wadsö, *Acta Chem. Scand.*, **25**, 1541 (1971).
(68) See, for example, P. K. Nandi, *Int. J. Pept. Protein Res.*, **8**, 253 (1976).
(69) F. H. Stillinger, *J. Solution Chem.*, **2**, 141 (1973); *Philos. Trans. R. Soc. London, Ser. B*, **278**, 97 (1977).
(70) D. W. Davidson in ref 1, Chapter 3.
(71) K. Heinzinger, *Z. Naturforsch. A*, **31**, 1073 (1976).
(72) H. G. Hertz and C. Rädle, *Ber. Bunsenges. Phys. Chem.*, **77**, 521 (1973).
(73) F. Franks in ref 1, Vol. 1, Chapter 4.
(74) S. J. Gill and I. Wadsö, *Proc. Natl. Acad. Sci. U.S.A.*, **73**, 2955 (1976).
(75) See, for example, C. H. Bennett, *J. Comput. Phys.*, **22**, 245 (1976).
(76) See F. H. Stillinger, *Adv. Chem. Phys.*, **31**, 1 (1975).
(77) G. Dashevsky and G. N. Sarkisov, *Mol. Phys.*, **27**, 1271 (1974).
(78) Reference 44, Chapter 4.
(79) S. Cabani, G. Conti, and E. Matteoli, *Biopolymers*, **16**, 465 (1977).
(80) H. S. Frank and M. W. Evans, *J. Chem. Phys.*, **13**, 507 (1945).
(81) J. C. Hindman, *J. Chem. Phys.*, **60**, 4488 (1974), and references cited therein.
(82) B. M. Fung and T. W. McGaughy, *J. Chem. Phys.*, **65**, 2970 (1976).
(83) D. E. O'Reilly, *J. Chem. Phys.*, **60**, 1607 (1974).
(84) F. H. Stillinger and A. Rahman in "Molecular Motions in Liquids", J. Lascombe, Ed., Reidel, Dordrecht-Holland, 1974.
(85) Reference 44, p 209.
(86) C. M. Davis and J. Jarzynski in "Water and Aqueous Solutions", R. A. Horne,

Ed., Wiley-Interscience, New York, 1972.
(87) See A. Ben-Naim in ref 1, Chapter 11.
(88) For example, G. H. Haggis, J. B. Hasted, and T. J. Buchanan, *J. Chem. Phys.*, **20**, 1425 (1952); V. Vand and W. A. Senior, *ibid.*, **43**, 1878 (1965).
(89) G. Némethy and H. A. Scheraga, *J. Chem. Phys.*, **36**, 3401 (1962).
(90) G. S. Kell in ref 82.
(91) Reference 44, Section 47.
(92) H. S. Frank and A. S. Quist, *J. Chem. Phys.*, **34**, 604 (1961).
(93) H. S. Frank and F. Franks, *J. Chem. Phys.*, **48**, 4746 (1968).
(94) V. A. Mikhailov, *J. Struct. Chem.*, **9**, 332 (1968); V. A. Mikhailov and L. I. Ponomarova, *ibid.*, **9**, 8 (1968).
(95) J. A. Pople, *Proc. R. Soc. London, Ser. A*, **205**, 163 (1951).
(96) S. Marcelja, D. J. Mitchell, B. W. Ninham, and M. J. Sculley, *J. Chem. Soc., Faraday Trans. 2*, **5**, 630 (1977).
(97) L. D. Landau and E. M. Lifshitz, "Statistical Physics", Pergamon Press, Oxford, 1969, Chapter 14.

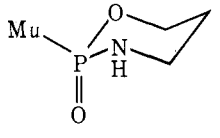
Solution Stereochemistry of Cyclophosphamide and Rigid Model Analogues

D. W. White, D. E. Gibbs, and J. G. Verkade*

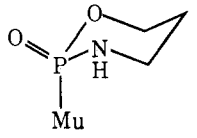
Contribution from Gilman Hall, Iowa State University, Ames, Iowa 50011.
Received September 11, 1978

Abstract: The structural properties of the anticancer drug cyclophosphamide (**1**, Mu = $(ClCH_2CH_2)_2N$) in solution are clarified by comparison of its spectroscopic characteristics with those of the rigid *cis*-4,6-dimethyl analogues **2a** and **2b**. The latter compounds, along with the *trans*-4,6-dimethyl isomers **3a** and **3b** (in which Mu is either axial or equatorial), were synthesized by ammonolysis of cyclic sulfate esters of *dl*- and *meso*-HOCHMeCHMeOH (**5a** and **5b**) followed by hydrolysis to the threo and erythro amino alcohols, respectively, and ring closure with $MuP(O)Cl_2$. In this synthetic scheme, isomerization of **5b** to **5a** was observed in acid and inversion at carbon apparently takes place in the ammonolysis step. ^1H NMR evidence is presented for the stereochemistries shown for the methyl groups in the principal conformers of **2a,b**, **3a,b**, and **5a,b**. Isomerism at phosphorus in **2a** and **2b** is indicated by examination of the N–H stretching region in the IR as well as by comparison of ^{31}P and ^1H NMR parameters. Detailed LIS investigations yield low *R* factors at high confidence levels for the structures shown. Similar experiments with **1** reveal a tendency toward conformation **1a** which is stronger in the presence of an LIS reagent but weaker in the presence of a hydrogen-bonding solvent such as water or chloroform. Although the stereochemistry of the methyl groups in **3a** and **3b** can be deduced from ^1H NMR data, assignment of the phosphorus stereochemistries in their dominant conformers is somewhat ambiguous. Preliminary antitumor cell screening indicates that **2a** is more active against KB cell cultures than **2b**.

As one of the most widely used chemotherapeutic agents in the treatment of many types of cancer, much effort has been made to understand the mode of action of cyclophosphamide (**1**) and to develop analogues with improved action.[1] Reports



1a



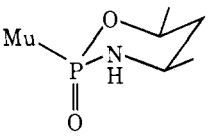
1b

Mu = $(ClCH_2CH_2)_2N$

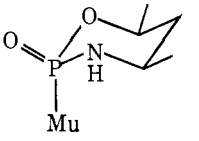
of structural investigations of cyclophosphamide and its ring carbon substituted analogues have been almost totally confined to solid-state X-ray investigations which reveal the presence of conformation **1a**.[2] Recently cyclophosphamide has been resolved[3] and the (+) enantiomer, which is more readily metabolized in human patients,[4] has been shown to have the *R* absolute configuration[5] while the (−) enantiomer, which is more active against PC6 mouse tumors,[6] has been demonstrated to possess the *S* configuration.[7]

In this paper we address ourselves to the stereochemical nature of **1** in solution, the state in which it displays its biological action. To obtain spectroscopic information characteristic of conformers **1a** and **1b**, we have synthesized the isomeric *cis*-4,6-dimethyl analogues **2a** and **2b**. Also reported are the isomeric *trans*-4,6-dimethyl compounds **3a** and **3b**.

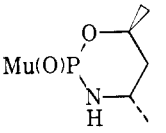
Previous efforts to elucidate the solution behavior of **1** have been few. Data from variable-temperature ^{13}C and ^1H NMR studies of **1** are consistent with a low barrier to ring reversal in the equilibrium **1a** ⇄ **1b**.[8] Attempts to elucidate the solution configurations of the isomeric 4-methyl derivatives of cyclophosphamide **4a** and **4b** by NMR and IR spectroscopic means have led to conflicting structural assignments.[9,10]



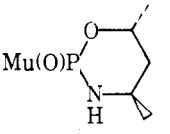
2a



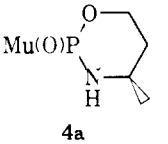
2b



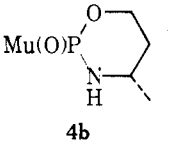
3a



3b



4a



4b

Experimental Section

Materials. Solvents and reactants, unless specifically noted otherwise, were reagent grade or better. Aromatic solvents were dried with NaK alloy, 4A molecular sieves, or KOH pellets, THF with K_2CO_3 or $LiAlH_4$, triethylamine with KOH pellets, ether with Na or Na/K alloy, and 1,2-dichloroethane, acetone, hexanes, carbon tetrachloride, and ethyl acetate with 4A molecular sieves. Ethanol was removed from preserved chloroform by washing several times with half a volume of water per volume of the solvent followed by drying for at least 1 day over anhydrous calcium chloride and distilling onto magnesium sulfate or 4A molecular sieves. Phosphoryl trichloride was distilled be-

0002-7863/79/1501-1937$01.00/0