A Theoretical Study of the Structure of Liquid Water Based on Quasi-Component Distribution Functions

S. Swaminathan and David L. Beveridge*

Contribution from the Chemistry Department, Hunter College of the City University of New York, New York, New York 10021. *Received February 7, 1977*

Abstract: Monte Carlo statistical thermodynamic computer simulations of liquid water at 25 °C based on potential functions representative of ab initio Hartree-Fock and configuration interaction calculations of the water-water pairwise interaction energy are reported. The results are analyzed in terms of quasi-component distribution functions for coordination number K and binding energy. The various ad hoc models for the structure of water are considered on this basis. The ratio of low density ($K \leqslant 4$) to high density species in the statistical state of the system is found to be 70/30. The quasi-component distribution functions imply a generalized continuum model for the system, with a composite of 3-, 4-, 5-, and 6-coordinate species giving rise to a unimodal distribution of binding energies. The effect of electron correlation on the structure of water is considered in terms of difference quasi-component distribution functions, and found to increase the mole fraction of four-coordinate species at the expense of both lower and higher coordinate quasi-components.

I. Introduction

The extensive use of liquid water as a solvent in chemical systems, the ubiquity of liquid water on the planet earth, and the unique role of water as a biological life support system combine to make the structure of liquid water a matter of central importance to chemistry, biochemistry, and molecular biology. With the advent of third generation digital computer hardware, research problems in molecular liquids previously approached by modelistic methods in statistical mechanics have become accessible to solution by computer simulation. The best calculated results on observable properties of liquid water are in quite good accord with experiment[1] and an analysis of the structure of water based on computer simulation is now in order. The simulation methods also afford the possibility of carrying out computer experiments on the system to elucidate the effect of various defined characteristics on the structure of the system

We present herein a theoretical analysis of the equilibrium, diffusionally averaged structure of liquid water at 25 °C based on a Monte Carlo statistical thermodynamic computer simulation of the system. The calculations are based on pairwise additive analytical potential functions representative of ab initio quantum mechanical calculations of the interaction energy of water dimer.[2,3] The analysis of the structure is carried out in terms of quasi-component distribution functions.[4] A comparison of results based on potential functions derived from calculations at the Hartree-Fock level and from calculations including configuration interaction permit a determination of the effect of electron correlation on the structure of water.

The background for this project is reviewed in section II, followed in section III by a description of the theory and methodology relevant to the calculations performed. The results are presented in section IV and discussed in detail in section V. A summary of the results and discussion is given in section VI.

II. Background

Early thoughts about the structure of liquid water developed as variations of ideas on the structure of ice. The conception of the structure of water as a quasi-solution of ice in equilibrium with a truly fluid polymorph can be traced at least as far back as Roentgen[5] in 1892. Modern quantitative treatments of the problem date from the classic paper of Bernal and Fowler[6] in 1933, featuring an icelike component consisting of a network of tetrahedrally coordinated water molecules. The theoretical development of the subject from that point can be developed in terms of essentially ad hoc and ab initio approaches to the problem.

The ad hoc approach involves a statistical thermodynamic treatment of a viable model for the entire system, with disposable parameters chosen empirically or for optimum agreement between calculated and experimentally observed values for selected properties. Representative examples of ad hoc models for liquid water are the interstitial models developed by Samoilov,[7] Mikhailov,[8] and Narten, Danford, and Levy,[9] the clathrate model due to Pauling[10] and Frank and Quist,[11] the two-state mixture model described by Davis and Litovitz,[12] the three-state model of Vand and Senior,[13] the flickering cluster model due to Frank and Wen and Hagler, Nemethy, and Scheraga,[14] and the significant structure approach by Jhon, Grosh, Ree, and Eyring.[15] Alternatively, the structure of water has been viewed as a continuous distribution of structures as proposed by Pople,[16] and elaborated into the continuum model for the system.[17]

The ab initio methods take as a point of departure the many-body problem in classical statistical mechanics or classical kinetic theory and treat the system as a molecular assembly interacting via a configurational potential. In statistical mechanics one proceeds to calculate properties of the system by configurational averaging based on the Monte Carlo method, whereas in the kinetic approach, molecular dynamics, the individual molecular trajectories are calculated by simultaneous solution of the Newton-Euler equations. Both Monte Carlo and molecular dynamics calculations are relatively large problems for a digital computer, and are collectively referred to as *computer simulations* of the system.[18] Other ab initio approaches to liquid structure besides computer simulation are the integral equation methods such as that of Percus and Yevick[19] and the perturbation theory of liquids due to Barker and Henderson.[20] There remain significant assumptions in the ab initio approach to molecular liquids since the configurational potential for the system is not known exactly to high order, but the assumptions enter at a lower point in the theoretical hierarchy than in ad hoc methods.

All of the early theoretical work on the structure of water was based on ad hoc methods, and much of our current thinking about the structural chemistry of molecular liquids developed from these studies. The fundamental idea that the anomalous behavior of liquid water as compared with normal liquids in the behavior of density and compressibility with respect to temperature was due to contributions from a low density four-coordinate icelike structure with anomalously low energy was well established in the early work. Detailed reviews of ad hoc models are available in the monographs of Eisenberg and Kauzmann,[21] Ben-Naim,[4] and the review articles by Davis and Jarzynski[22] and by Hagler, Nemethy, and Scheraga[14] on mixture models and Kell[17] on continuum models. Interesting additional perspective on the problem is due to Kamb,[22] Stevenson,[23] and Frank.[24] Extensive experimental data are available on the system, but the collected results have been unable to differentiate unequivocally between mixture models and continuum models or among the various mixture models. The experimental evidence for mixture models based on vibrational spectra has been given particularly by Walrafen,[25] and the arguments from diverse viewpoints are given in ref 17 and 22.

A considerable literature exists on computer simulation of model fluids, but the application of computer simulation techniques to molecular liquids is at relatively early stage of development at present.[26] The predominant number of published studies deal with liquid water,[26-28] with simulations on other pure liquids and some solutions just now appearing.[29] The original molecular dynamics on liquid water study is due to Rahman and Stillinger[27] and incorporates the Ben-Naim–Stillinger (BNS) potential.[30] The first Monte Carlo simulations on liquid water were reported by Watts and Barker,[28a] and Sarkisov et al.[28b] based on empirical potentials.[30,31]

Concurrently an important series of papers on analytical potential functions representative of ab initio quantum mechanical calculation of pairwise interaction energies by Clementi and co-workers have appeared. Potentials for the water–water interaction were reported at the level of Hartree–Fock (HF) molecular orbital theory[2] and with a number of alternative corrections for electron correlation,[32] the best quality function being representative of moderately large configuration interaction (CI) calculations on the water dimer.[3] Each of these potentials was incorporated in Monte Carlo computer simulations and calculations on a wide variety of molecular properties of liquid water were reported. The Hartree–Fock potential produced an oxygen–oxygen radial distribution function with serious discrepancies between calculated and observed values beyond the first hydration shell.[32] The CI potential with only intermolecular correlation corrections included gave results for the oxygen–oxygen radial distribution reasonably close to the observed values for the entire range of interparticle separations,[1] and is currently considered the best available pairwise potential function for the water–water interaction.

A satisfactory quantitative agreement between experimentally observed radial distribution functions and corresponding quantities calculated for liquid water in computer simulation incorporating the CI potential having been demonstrated, it is of interest to analyze the results of the simulation with a view toward obtaining an ab initio theoretical idea of the nature of the structure of liquid water. Moreover, the availability of both the CI and the HF potential functions permits a comparison of results obtained with and without electron correlation and a study of the effect of electron correlation on the structure of liquid water.

III. Theory

Motional freedom of molecules in liquids mandates theoretical studies in this area to be problems in statistical mechanics. The theoretical description of the system follows from the partially classical partition function[4,26]

$$Q(T,V,N) = \frac{q^N}{(8\pi^2)^N \Lambda^{3N} N!} \int \dots \int \exp[-\beta E(\mathbf{X}^N)]\, d\mathbf{X}^N \quad (1)$$

appropriate for N particles in volume V at temperature T, the canonical (T,V,N) ensemble. Here q is the internal partition function and Λ is the one-dimensional translational partition function for a single particle. The integration ranges over all configurational coordinates $\mathbf{X}^N$ of the N particles of the system,

$$\mathbf{X}^N = \{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \dots, \mathbf{X}_N\} \quad (2)$$

where $\mathbf{X}_i$ represents the configurational coordinates of the ith particle consisting of a specification of position $\mathbf{R}_i$ and orientation $\mathbf{\Omega}_i$, $\mathbf{X}_i = \{\mathbf{R}_i, \mathbf{\Omega}_i\}$. The quantity $E(\mathbf{X}^N)$ is the configurational energy of the system, with β representing $(kT)^{-1}$. The multiple integration in eq 1 is referred to as configurational averaging.

Monte Carlo computer simulations focus on the determination of average properties of the system such as configurational internal energy U,

$$U = \int \dots \int E(\mathbf{X}^N) P(\mathbf{X}^N)\, d\mathbf{X}^N \quad (3)$$

where $P(\mathbf{X}^N)\, d\mathbf{X}^N$ is the probability of finding the system in configuration $\mathbf{X}^N$,

$$P(\mathbf{X}^N) = \frac{\exp[-\beta E(\mathbf{X}^N)]}{\int \dots \int \exp[-\beta E(\mathbf{X}^N)]\, d\mathbf{X}^N} \quad (4)$$

Expressions for other average properties of the system such as distribution functions can be formulated in a manner analogous to eq 3. Computer simulation in the Monte Carlo sense is the numerical integration of eq 3

$$U = M^{-1} \sum_{i=1}^{M} E(\mathbf{X}^N) P(\mathbf{X}_i{}^N) \quad (5)$$

with the M configurations $\mathbf{X}_i{}^N$ generated at random. In practice the numerical process represented by eq 5 is known to be very slowly convergent. An expedient suggested by Metropolis et al.[33] involves carrying out the calculation by means of a stochastic walk through configuration space, selecting points $\mathbf{X}_j{}^N$ which enter the averaging process with a frequency $P(\mathbf{X}_j{}^N)$. The determination of average quantities reduces to a simple summation, viz.

$$U = M^{-1} \sum_{j=1}^{M} E(\mathbf{X}_j{}^N) \quad (6)$$

where the configurations $\mathbf{X}_j{}^N$ are chosen by the Metropolis method. This process converges in 500K steps for most homogeneous systems of 0(100) particles. The N-particle system is given a condensed phase environment by the appropriate choice of T and V and the use of image cells, i.e., periodic boundary conditions.[34]

The principal numerical results of Monte Carlo computer simulation are the thermodynamic indices of the system such as internal energy and heat capacity. (The Metropolis method can be extended to produce quantities such as free energy (and thus entropy) by the use of auxiliary potential functions and conditional probabilities.) The spatial pair correlation function for the system follows from an integration over positional coordinates $\mathbf{R}^N$ of the N-particle system,

$$g(\mathbf{R}_1, \mathbf{R}_2) = \rho^{-2} \int \dots \int P(\mathbf{R}^N) \sum_{i \neq j}\sum \delta[L_i(\mathbf{R}^N) - \mathbf{R}_1]\, \delta[L_j(\mathbf{R}^N) - \mathbf{R}_2]\, d\mathbf{R}^N \quad (7)$$

where ρ is number density, $L_i(\mathbf{R}^N)$ gives the position vector of particle i in configuration $\mathbf{R}^N$, and the bracketed quantities under the double summation are Dirac δ counting functions[4] for the number of particles within $d\mathbf{R}_1$ of $\mathbf{R}_1$ jointly with particles within $d\mathbf{R}_2$ of $\mathbf{R}_2$. In numerical calculations the corresponding radial distribution function is obtained with $R = |\mathbf{R}_2 - \mathbf{R}_1|$ as

$$g(R) = \frac{N(R)}{\rho 4\pi R^2 \Delta R} \quad (8)$$

where $N(R)$ is the average number of particles in a spherical shell of width ΔR at a radial distance R from another particle; the function $g(R)$ gives the configurationally averaged deviation of the local microscopic environment of a particle from the value characteristic of bulk density. Experimental values of $g(R)$ for liquids are obtained from an analysis of the results of x-ray and neutron diffraction experiments.[9] The average coordination number can be determined from $g(R)$ as the average number of particles within a sphere of a given radius R_M,

$$N_C = \rho \int_0^{R_M} g(R) 4\pi R^2 \,\mathrm{d}R \qquad (9)$$

The parameter R_M can be chosen as the radial value of the first minimum in $g(R)$ to correspond to common chemical concepts of coordination number.

A microscopic theoretical analysis of the system can be generated in terms of quasi-component distribution functions[4] (QCDF). QCDF with respect to coordination number and binding energy are the principal means of analysis of the computer simulation results in this investigation. The QCDF for coordination number $N_C^{(1)}(K)$ gives the average number of particles in the system with coordination number K. The mole fraction of such particles is given by

$$x_C(K) = \frac{N_C^{(1)}(K)}{N}$$

$$= \frac{1}{N} \int \ldots \int P(\mathbf{R}^N) \sum_i \delta[C_i(\mathbf{R}^N) - K] \,\mathrm{d}\mathbf{R}^N \qquad (10)$$

where $\Sigma_i \delta[C_i(\mathbf{R}^N) - K]$ is a Dirac δ counting function for the number of particles with coordination number K in configuration $\mathbf{R}^N$. The quantity $x_C(K)$ can be viewed as a component of the vector

$$\mathbf{x}_C(K) = \{x_C(0), x_C(1), x_C(2), \ldots\} \qquad (11)$$

the elements of which define the entire composition of the system in terms of coordination number. The average coordination number of eq 9 in terms of $x_C(K)$ is

$$N_C = \sum_{K=0}^{\infty} K x_C(K) \qquad (12)$$

The QCDF with respect to the binding energy B_i for particle i follows from the definition

$$B_i = E(\mathbf{X}_1, \ldots, \mathbf{X}_{i-1}, \mathbf{X}_i, \mathbf{X}_{i+1}, \ldots, \mathbf{X}_N)$$
$$- E(\mathbf{X}_1) \ldots, \mathbf{X}_{i-1}, \mathbf{X}_{i+1}, \ldots, \mathbf{X}_N) \qquad (13)$$

The mole fraction of particles with binding energy between ν and $\nu + \mathrm{d}\nu$ is given by

$$x_B(\nu) = \frac{N_B^{(1)}(\nu)}{N} = \frac{1}{N} \int \ldots \int P(\mathbf{X}^N)$$
$$\times \sum_i \delta[B_i(\mathbf{X}^N) - \nu] \,\mathrm{d}\mathbf{X}^N \qquad (14)$$

where $\delta[B_i(\mathbf{X}^N) - \nu]$ is a Dirac δ counting function for particles with binding energy between ν and $\nu + \mathrm{d}\nu$. The vector

$$\mathbf{x}_B = \{x_B(\nu_0), x_B(\nu_1), x_B(\nu_2), \ldots\} \qquad (15)$$

defines the composition of the entire system as a function of configurationally averaged binding energy for individual particles. The relationship between $x_B(\nu)$ and configurational internal energy is

$$U = \frac{N}{2} \int_{-\infty}^{+\infty} \nu x_B(\nu) \,\mathrm{d}\nu \qquad (16)$$

Note that $x_C(K)$ is a discrete function in K whereas $x_B(\nu)$ is a continuous function in ν. It is expected that the various quasi-components of a mixture model would be distinguished by separate maxima in a plot of $x_B(\nu)$ vs. ν.

Finally the analysis of the system in terms of QCDF can be extended to display the distribution of binding energies as a function of coordination number,

$$x_{B,C}(\nu,K) = (8\pi^2)^{-N} \int \ldots \int P(\mathbf{X}^N)$$
$$\times \sum_i \delta[B_i(\mathbf{X}^N) - \nu] \sum_j \delta[C_j(\mathbf{X}^N) - K] \,\mathrm{d}\mathbf{X}^N \qquad (17)$$

and define a compositional vector as

$$\mathbf{x}_{B,C} = \{x_{B,C}(\nu,K)\} \qquad (18)$$

Further characteristics of the system can be displayed in terms of QCDF formulated on the basis of considerations analogous to those defined above.

The principal variables in a Monte Carlo computer simulation are T, V, N, and the means of determining the configurational energy of the system $E(\mathbf{X}^N)$. This quantity can be expressed as

$$E(\mathbf{X}^N) = \sum_{i<j} E^{(2)}(\mathbf{X}_i, \mathbf{X}_j)$$
$$+ \sum_{i<j<k} E^{(3)}(\mathbf{X}_i, \mathbf{X}_j, \mathbf{X}_k) + \ldots \qquad (19)$$

where $E^{(n)}$ are n-body contributions to the interaction. The conventional way to proceed in computer simulation is to obtain a relatively simple algebraic expression for the $E^{(n)}$ and program this into the overall calculation. Most simulation work on molecular liquids has only considered the pairwise ($n = 2$) term with the others either neglected entirely or considered to be mapped onto an effective pairwise term in some averaged way. The analytical potential functions representative of quantum mechanical calculations are directly identified with $E^{(2)}$ by definition. The effect of higher order terms has been studied especially for multiple water interactions by Hankins, Moskowitz, and Stillinger[35] and Lenz and Scheraga;[36] higher order effects can contribute errors in the interaction energy of the order of 10–15%. Still little is presently known of reasonable analytical forms for higher order $E^{(n)}$, although this problem is presently under consideration. All calculations in the study described herein assume pairwise additivity and this should be considered as a possible source of error in the results.

IV. Calculations and Results

Programs based on the theoretical development described in the preceding section were coded and implemented on the CUNY IBM 370/168 computer. Calculations were carried out on liquid water by means of a 125 particle simulation at 25 °C and a volume commensurate with a density of 1 g/cm^3 as in the Rahman-Stillinger dynamical simulation.[27] Step size in the stochastic walk in the Metropolis procedure was selected to produce an approximately 50% acceptance rate. All calculations were found to be well converged after 500K steps of the Monte Carlo process, 200K points having been initially discarded.

Two separate computer simulations are described, one based on the CI potential of Matsuoka, Clementi, and Yoshimine[3] and the other based on the HF potential of Popkie, Kistenmacher, and Clementi.[2] The difference between these functions is displayed for one representative slice of the potential energy hypersurface in Figure 1. Here the interaction energy for the linear water dimer is plotted as a function of the oxygen-oxygen separation for the linear water dimer; this is the optimum geometry for intermolecular hydrogen bonding.[35] The HF potential has a minimum at 3.01 Å and an intermolecular binding energy of −4.55 kcal/mol. The CI potential function

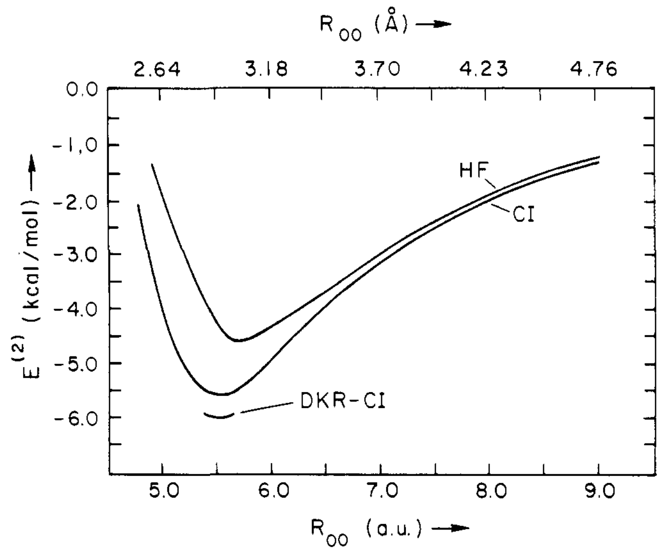

Figure 1. Calculated pairwise interaction energy $E^{(2)}$ as a function of interoxygen separation R_{OO} for the linear configuration of the water dimer; adapted from Figure 6 of ref 3 and Figure 1 of ref 32.

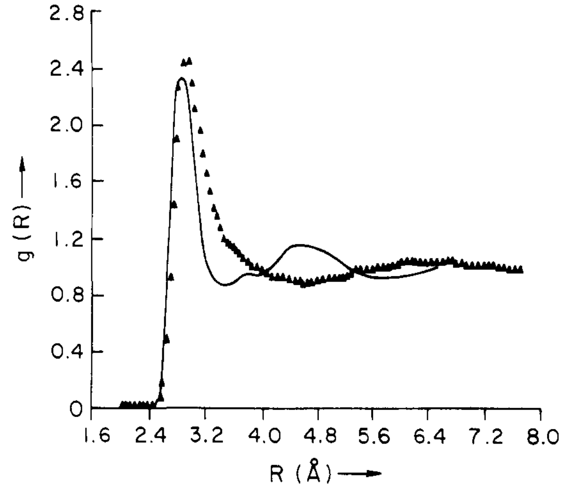

Figure 2. A comparison of the experimentally observed radial distribution function for liquid water (solid line) with points on the radial distribution function calculated from a Monte Carlo computer simulation based on the HF potential. The experimental curve is redrawn from Figure 2 of ref 1; experimental data are due to Narten, Danford, and Levy.[9]

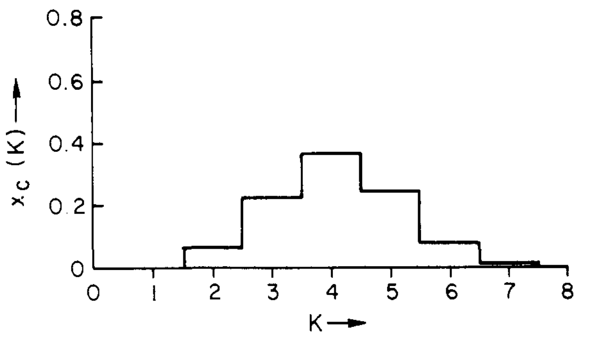

Figure 3. Calculated quasi-component distribution function $x_C(K)$ vs. coordination number K based on the HF potential function.

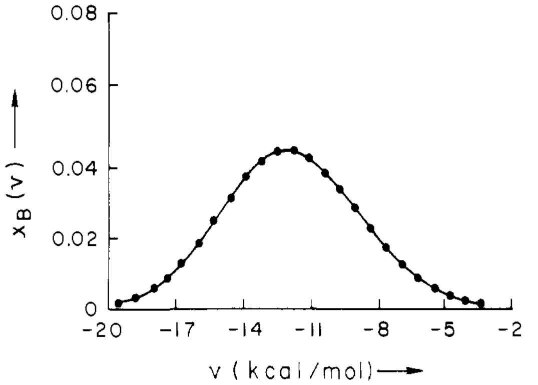

Figure 4. Calculated quasi-component distribution function $x_B(\nu)$ vs. binding energy ν for liquid water based on the HF potential function.

has a minimum at 2.92 Å and an intermolecular binding energy of −5.62 kcal/mol. The CI potential function gives results parallel to and within 0.43 kcal/mol of the most extensive configuration calculation to date on the water dimer reported by Diercksen, Kraemer, and Roos,[37] labeled "DKR-CI" in Figure 1. The effect of electron correlation on the pairwise interaction is to stabilize the energy by approximately 1 kcal/mol (~19%) in the binding region and reduce the minimum energy intermolecular separation by 0.09 Å.

The calculated radial distribution function for the computer simulation incorporating the HF potential is shown in comparison with the corresponding experimentally observed result in Figure 2. This reproduced Figure 10 of Kistenmacher, Popkie, Clementi, and Watts,[32] and serves as a confirmation of our computer programs. Comparing the calculated and observed results, the HF potential as noted previously by Clementi and co-workers shows significant deviations from the experimentally determined function, with the first peak (first hydration shell) biased toward larger distances and the position of the second hydration shell well displaced from the observed values; in fact the HF potential produces a minimum in $g(R)$ where the experimental values show a maximum. The calculated configurational internal energy is −6.9 kcal/mol with a standard deviation estimated at 0.03 kcal/mol.[38] For comparison with experimentally observed internal energy this requires addition of $3RT$, the kinetic energy of free molecular and translation. This results in −5.12 kcal/mol as compared with the experimentally observed value of −8.1 kcal/mol. The discrepancy is attributed mainly to electron correlation, the assumption of pairwise additivity, and basis set truncation errors. The calculated heat capacity is 18 cal/deg mol, in (fortuitously) exact agreement with the experimental value at 25 °C.

An analysis of the computer simulation based on the HF potential was carried out in terms of the QCDF for coordination number and binding energy. The structure of "HF water" in terms of coordination number is shown in Figure 3. Here the mole fraction $x_C(K)$ based on $R_M = 3.3$ is plotted vs. K in histogram form. The predominant coordination number is $K = 4$, coming in at 37%. The distribution is markedly symmetric about $K = 4$ with $K = 3$ and $K = 5$ represented to the extent of 23% and $K = 2$ and $K = 6$ at about 7%. Approximately 2% of the molecules have a coordination of 7 in the equilibrium statistical state of the system. The average coordination number is 4.05 compared with an experimental value of 4.4.

The structure of HF water in terms of binding energy is shown in Figure 4. Here the mole fraction $x_B(\nu)$ is displayed as a function of ν. We find a continuous distribution of binding energies in the system essentially symmetric about −12 kcal/mol. In Figure 5 the distribution of particles as a function of both coordination number and binding energy, $x_{B,C}(\nu,K)$ vs. ν and K, is presented. The curves for each value of K in Figure 5 are individually normalized and must be multiplied by $x_C(K)$ to assess their respective contributions to the statistical state of the system. The $x_C(K)$ histogram is displayed along the K axis of the figure to facilitate comprehension of

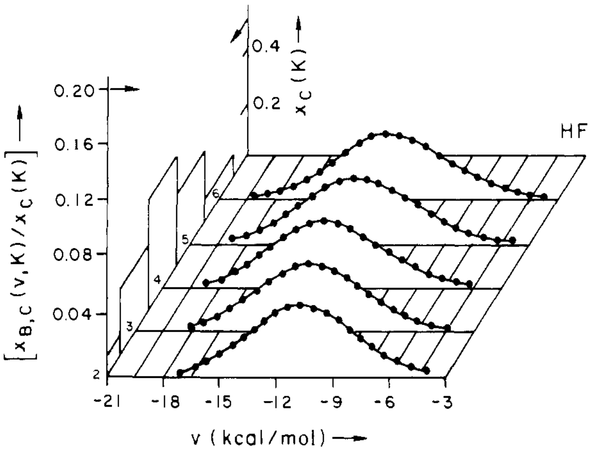

Figure 5. Calculated quasi-component distribution function $x_{B,C}(\nu,K)$ vs. binding energy ν and coordination number K for liquid water based on the HF potential function. Each curve is normalized by $x_C(K)$ and must be weighted by this quantity to assess its relative contribution to the statistical state of the system. To aid in the comprehension of this plot, $x_C(K)$ from Figure 3 is plotted along the K axis.

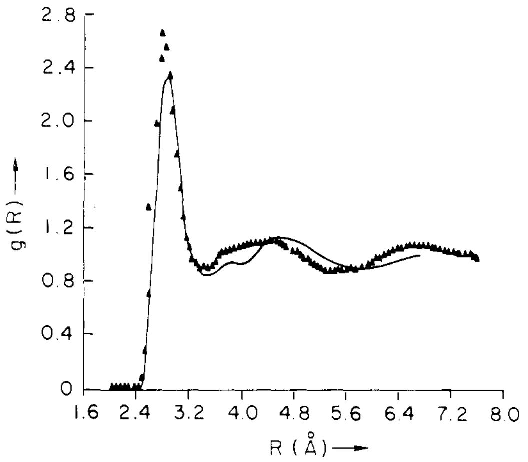

Figure 6. A comparison of the experimentally observed radial distribution function for the liquid water (solid line) with points on the radial distribution function calculated from a Monte Carlo computer simulation based on the CI potential. See Figure 2 for notation on the experimental data.

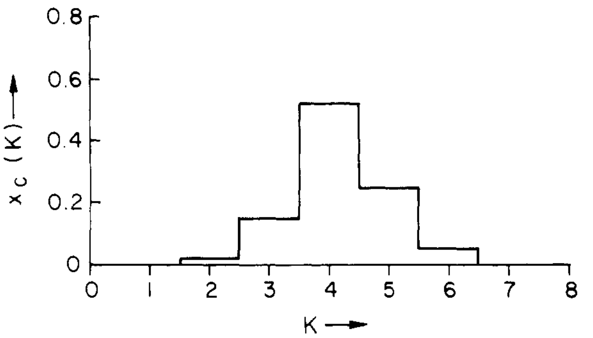

Figure 7. Calculated quasi-component distribution function $x_C(K)$ vs. coordination number K for liquid water based on the CI potential function.

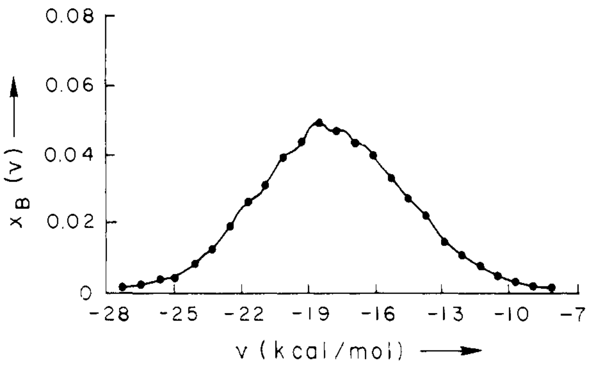

Figure 8. Calculated quasi-component distribution function $x_{B,C}(\nu,K)$ binding energy ν for liquid water based on the CI potential function.

the plots. Considering first the behavior of $x_{B,C}(\nu,K)$ vs. ν, each of the distributions are individually continuous and essentially symmetric about a single maximum value. The behavior of $x_{B,C}(\nu,K)$ vs. K shows a single minimum at $K = 4$, the most predominant coordination number in the system. The binding energies for $K > 4$ are lower in energy than those for $K < 4$.

The calculated and observed oxygen-oxygen radial distribution function from the computer simulation of liquid water incorporating the CI potential is shown in Figure 6; cf. Figure 2a of ref 1. There is quite good accord between calculated and observed values for the position of all three main peaks. In the region of the first hydration shell the shape of the calculated peak agrees well with experiment, although the calculated maximum is slightly too high. For the second hydration shell the calculated position of the maximum in $g(R)$ is in close accord with experiment, but the shape is biased toward short distances. The small peak or shoulder at ca. 3.5 Å, much discussed by previous investigators, is clearly evident in the calculated $g(R)$. The third shell appears generally well accommodated. As noted by Lie et al., overall the results are in satisfactory accord with experiment and as good as or better than corresponding results based on any of the available empirical potential functions. The calculated internal energy, corrected for kinetic energy as discussed above, is −6.8 kcal/mol, some 16% above the experimentally observed value. The calculated standard deviation on internal energy is 0.03 kcal/mol.[38] This discrepancy is attributed primarily to the assumption of pairwise additivity in the potential function, and appears to be of the order of magnitude expected for this effect. The calculated heat capacity is 17.9 cal/mol deg.

The analyses of the computer simulation of liquid water based on the CI potential in terms of QCDFs are displayed in Figures 7, 8, and 9. The calculated structure of water in terms of coordination number for $R_m = 3.3$ is shown in Figure 7. The predominant coordination number is $K = 4$ at 47%. The distribution of coordination numbers in the system is biased towards values of $K = 4$, with $K = 5$ at 24% and $K = 6$ at 6% compared with $K = 3$ at 19% and $K = 2$ at 4%. The calculated average coordination number is 4.15.

The calculated structure of water in terms of binding energy is shown in Figure 8. There is a major maximum in the plot at −17.7 kcal/mol, a relatively symmetric distribution of binding energies about this value, and a slight but clearly discernible bias in the distribution toward lower binding energies. There is evidence of some structure in the curve, but overall the calculated distribution is continuous.

The calculated distribution of water molecules as a function of both binding energy and coordination number is shown in Figure 9. Considering $x_{B,C}(\nu,K)$ vs. ν, each of the curves is individually continuous about a single maximum value. The behavior of $x_{B,C}(\nu,K)$ vs. K shows a single minimum at $K = 4$. The binding energy curves for $K > 4$ are displaced toward lower energies relative to those for $K < 4$. The $x_{B,C}(\nu,K)$ distribution for $K = 4$ is biased toward low binding energies whereas the distributions of $K \neq 4$ are (slightly) biased toward higher binding energies. The anomalously low binding energy of the four-coordinate quasi-component is clearly evident in Figure 9.

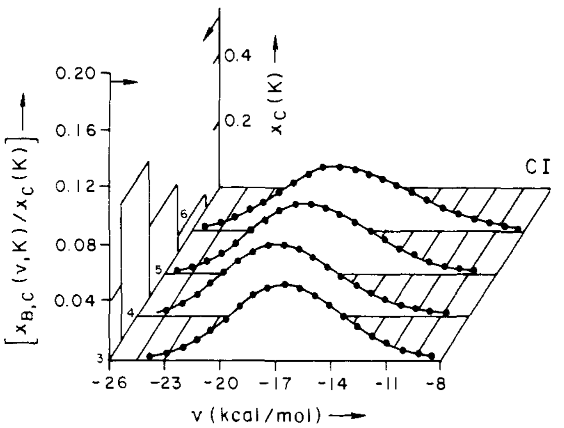

Figure 9. Calculated quasi-component distribution function $x_{B,C(\nu,K)}$ vs. binding energy ν and coordination number K for liquid water based on the CI potential function. Each curve is normalized by $x_C(K)$ and must be weighted by this quantity to assess its relative contribution to the statistical state of the system. To aid in the comprehension of this plot, $x_C(K)$ from Figure 7 is plotted along the K axis.

The analysis described above must be considered in perspective of the calculated error bounds on each quantity. A detailed error analysis based on the procedure described by Wood[38] was carried out on the $g(R)$ and $x_C(K)$ determined for the CI function. The standard deviation on $g(R)$ was found to be a maximum of 0.037 in the region of the first hydration shell and typically of the order of 0.01-0.02 elsewhere. The standard deviation in $x_C(K)$ was at its maximum for $K = 4$, a value of 0.007. Similar statistics are implied for the HF results.

V. Discussion

Since the Monte Carlo simulation of liquid water incorporating the CI potential produces a radial distribution function in close accord with experiment, we can look to the analysis of this calculation (Figures 7, 8, and 9) for a theoretical view of the structure of water. Considering first the idea of low density and high density contributions to water structure, the analysis of the results in terms of $x_C(K)$ in Figure 7 allows these contributions to be quantified. The mole fraction of low density component x_L is

$$x_L = \sum_{K=0}^{4} x_C(K) = 0.70 \tag{20}$$

and the mole fraction of high density component x_H is

$$x_H = \sum_{K=5}^{\infty} x_C(K) = 0.30 \tag{21}$$

for the statistical state of the system at 25 °C.

The predominant quasi-component with respect to coordination number at $K = 4$ in Figure 7 can be identified as part of the low density component considered in the ad hoc models. For an icelike structure we expect four hydrogen bonds per molecule at -5.62 kcal/mol each; a binding energy of -23 kcal/mol is near the lower limit of the $x_{B,C}(\nu,4)$ curve in Figure 9. Note that binding energy as defined in eq 13 includes contributions not just from the first hydration shell but from the entire assembly; thus the minimum energy in this plot is <-23 kcal/mol. The fact that the maximum in $x_{B,C}(\nu,4)$ falls at >-23 kcal/mol indicates that slightly distorted hydrogen bonds or other bound structures play a significant role in the system.

One can look to the plot of $x_B(\nu)$ in Figure 8 for information on the relative viability of mixture models as opposed to continuum models. According to the calculations on model systems,[4] the presence of distinct quasi-components of a mixture model should be indicated by structure in the $x_B(\nu)$ curve, with a separate distinct peak for each component, whereas a continuum model would be represented by a continuous distribution of individual binding energies. The results in Figure 8 favor the latter alternative; similar conclusions emerged from molecular dynamics. [27]This result should be considered in the perspective of the assumption of pairwise additivity in the configurational energy, since the formation of certain types of quasi-components may depend on cooperative effects.

The effect of electron correlation on the calculated structure of liquid water can be determined by comparing the results based on the HF potential function (Figures 2-5) with those obtained from the CI potential function (Figures 6-9). Comparing the results on radial distribution function obtained from the HF potential function (Figure 2) and the CI potential function (Figure 6), the inclusion of electron correlation is found to bring the calculated $g(R)$ into close accord with experiment. As noted at the opening of the preceding section, electron correlation increases the binding energy and decreases the intermolecular separation in the water dimer, and it is of interest to inquire as to how these effects operate in bringing about improved agreement with experiment. Matsouka et al. in considering potential functions representative of CI calculations on the water dimer reported another analytical potential function for the system which displayed essentially the correlation effect on binding energy but not on geometry. We have carried out an additional simulation based on this potential and find the results on $g(R)$ to be very close to that obtained using the HF potential function given in Figure 2. Therefore the correlation effect on geometry as well as energy is essential for obtaining agreement between calculated and observed radial distribution functions.

The effect of electron correlation on the calculated structure

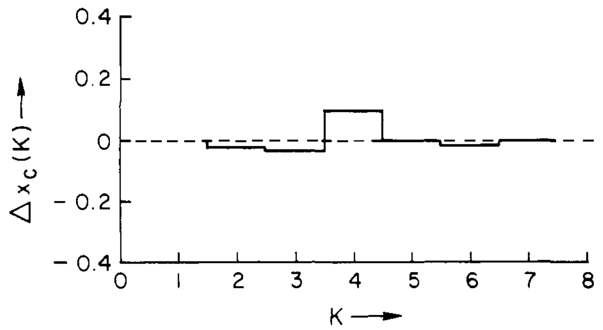

Figure 10. Calculated difference quasi-component distribution function $x_C(K)$ for liquid water; $x_C(K)$ for the CI potential function minus $x_C(K)$ for the HF potential function.

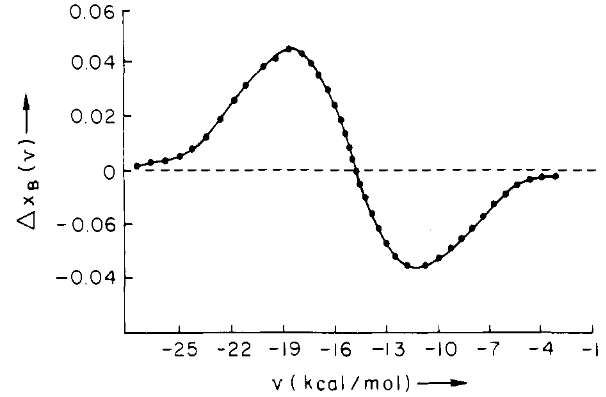

Figure 11. Calculated difference quasi-component distribution function $x_B(\nu)$ for liquid water; $x_B(\nu)$ for the CI potential function minus $x_B(\nu)$ for the HF potential function.

of liquid water can be quantified in terms of both coordination number and binding energy. This is most directly displayed in terms of difference (Δ) QCDFs,

$$\Delta x_C(K) = x_C(K)^{CI} - x_C(K)^{HF} \tag{22}$$

$$\Delta x_C(\nu) = x_C(\nu)^{CI} - x_C(\nu)^{HF} \tag{23}$$

The Δ QCDF for coordination number is shown in Figure 10. Correlation effects are seen to increase the percentage of four-coordinate species at the expense of low density $K = 2$ and $K = 3$ and the higher density $K = 6$ quasi-component. The Δ QCDF for binding energy is shown in Figure 11 and displays essentially a simple displacement of the binding energy curve toward lower values. Thus the effect of electron correlation on the calculated structure of liquid water is to increase the four-coordinate structure of the system by 10% and decrease the average binding of a water molecule in the liquid by about 6 kcal/mol.

VI. Summary and Conclusions

The analysis of results of Monte Carlo computer simulations on liquid water at 25 °C which are in satisfactory accord with experimentally observed thermodynamic properties and radial distribution functions quantifies the distribution of coordination number and binding energies in the system. The predominance of four-coordinate quasi-components in the diffusionally averaged structure of the system is confirmed, and a 70/30 ratio of low density to high density components is indicated. Analysis of the results in forms of the distribution of binding energies favors the idea of a continuous distribution of structures rather than a mixture model of liquid water, and included in the continuum of structures are contributions from 3-, 5-, and 6-coordinate species as well as the 4-coordinate component of "Pople water". The use of quasi-component distribution functions is seen to be a powerful means for the analysis of the structural chemistry of the statistical state of a molecular liquid.

The effect of electron correlation on the structure of water, obtained by comparing results based on a configuration interaction intermolecular potential function with those obtained from a Hartree-Fock intermolecular potential function, is to increase the four-coordinate structure of the system. The successful use of difference quasi-component distribution functions to quantify this effect suggests analogous use of this approach to quantify the effect of solutes on solvent structure. Such work is presently in progress in this laboratory.

Acknowledgments. This research is submitted by S.S. in partial fulfillment of the requirements for the Ph.D. in Chemistry at the City University of New York. Support for this project comes in part from NIH Grant 1-RO1-NS12149 from the National Institute of Neurological and Communication Diseases and Stroke, and a CUNY Faculty Research Award. D.L.B. acknowledges U.S. Public Health Service Research Career Development Award 6TK04-GM21281 from the National Institute of General Medical Studies. Helpful discussions with Dr. M. Mezei and Dr. S. W. Harrison are gratefully acknowledged.

References and Notes

(1) G. C. Lie, E. Clementi, and M. Yoshimine, *J. Chem. Phys.*, **64**, 2314 (1976).
(2) H. Popkie, H. Kistenmacher, and E. Clementi, *J. Chem. Phys.*, **59**, 1235 (1973).
(3) O. Matsuoka, E. Clementi, and M. Yoshimine, *J. Chem. Phys.*, **64**, 1351 (1976).
(4) A. Ben-Naim, "Water and Aqueous Solutions", Plenum Press, New York, N.Y., 1974.
(5) W. C. Roentgen, *Ann. Phys.*, **45**, 91 (1892).
(6) J. D. Bernal and R. H. Fowler, *J. Chem. Phys.*, **1**, 515 (1933).
(7) O. Ya. Samoilov, *Zh. Fiz. Khim.*, **20**, 1411 (1946); *Zh. Strukt. Khim.*, **4**, 499 (1963).
(8) V. A. Mikhailov, *Zh. Strukt. Khim.*, **8**, 189 (1967); **9**, 397 (1968).
(9) A. H. Narten, M. D. Danford, and H. A. Levy, *Discuss. Faraday Soc.*, **43**, 97 (1967); A. H. Narten and H. A. Levy, *Science*, **165**, 3892 (1969).
(10) L. Pauling, *J. Am. Chem. Soc.*, **57**, 2680 (1935); *Science*, **134**, 15 (1961).
(11) H. S. Frank and A S. Quist, *J. Chem. Phys.*, **34**, 604 (1961).
(12) C. M. Davis and T. A. Litovitz, *J. Chem. Phys.*, **42**, 2563 (1965). See also the interesting alternative proposed by C. A. Angell, *J. Phys. Chem.*, **75**, 3698 (1971).
(13) V. Vand and W. A. Senior, *J. Chem. Phys.*, **43**, 1878 (1965).
(14) A. T. Hagler, H. A. Scherega, and G. Nemethy, *J. Phys. Chem.*, **76**, 3229 (1972); *Ann. N.Y. Acad. Sci.*, in press.
(15) M. S. Jhon, J. Grosh, T. Ree, and H. Eyring, *J. Chem. Phys.*, **44**, 1465 (1966).
(16) J. A. Pople, *Proc. R. Soc. London, Ser. A*, **205**, 163 (1951).
(17) G. S. Kell in "Water and Aqueous Solutions", R, A. Horne, Ed., Wiley, New York, N,Y., 1972.
(18) I. R. McDonald and K. Singer, *Q. Rev., Chem. Soc.*, **24**, 238 (1970).
(19) J. K. Percus and G. L. Yevick, *Phys. Rev.*, **110**, 1 (1958).
(20) J. A. Barker and D. H. Henderson, *Annu. Rev. Phys.*, **23**, 439 (1972); *Acc. Chem. Res.*, **4**, 303 (1971).
(21) D. Eisenberg and W. Kauzmann, "The Structure and Properties of Water", Oxford University Press, London, 1969.
(22) B. Kamb in "Structural Chemistry and Molecular Biology", A. Rich and E. Davidson, Ed., W. H. Freeman, San Francisco, Calif., 1968.
(23) D. P. Stevenson in "Structural Chemistry and Molecular Biology", A. Rich and E. Davidson, Ed., W. H. Freeman, San Francisco, Calif., 1968.
(24) H. S. Frank in "Water—A Comprehensive Treatise", Vol. I, F. Franks, Ed., Plenum Press, New York, N.Y., 1972, p 515.
(25) G. E. Walrafen in ref 24, p 151.
(26) J. A. Barker and D. H. Henderson *Rev. Mod. Phys.*, **48**, 587 (1976).
(27) A. Rahman and F. H. Stillinger, *J. Chem. Phys.*, **55**, 336 (1971); *J. Am. Chem. Soc.*, **95**, 7943 (1973).
(28) (a) J. A. Barker and R. O. Watts, *Chem. Phys. Lett.*, **3**, 144 (1969); R. O. Watts, *Mol. Phys.*, **28**, 1069 (1974); (b) G. N. Sarkisov, V. G. Dashevsky, and G. G. Malenkov, *Mol. Phys.*, **27**, 1249 (1974).
(29) I. R. McDonald and M. L. Klein, *J. Chem. Phys.*, **64**, 4791 (1976); D. J. Evans and R. O. Watts, *Mol. Phys.*, **32**, 93 (1976); R. O. Watts, *ibid.*, **32**, 659 (1976).
(30) A. Ben-Naim and F. H. Stillinger in "Water and Aqueous Solutions", R. A. Horne, Ed., Wiley, New York, N.Y., 1974.
(31) J. S. Rowlinson, *Trans. Faraday Soc.*, **47**, 120 (1951).
(32) H. Kistenmacher, H. Popkie, E. Clementi, and R. O. Watts, *J. Chem. Phys.*, **60**, 4455 (1974).
(33) N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.*, **21**, 2087 (1953).
(34) F. Kohler, "The Liquid State", Crane Russack Co., 1972, Chapter 3.
(35) D. H. Hankins, J. W. Moskowitz, and F. H. Stillinger, *J. Chem. Phys.*, **53**, 4544 (1970).
(36) B. R. Lentz and H. A. Scheraga, *J. Chem. Phys.*, **58**, 5296 (1973).
(37) G. H. F. Diercksen, W. P. Kraemer, and B. O. Roos, *Theor, Chim Acta*, **36**, 249 (1975).
(38) Error bounds in all calculations were computed based on control functions as described by W. W. Wood in "Physics of Simple Liquids", H. N. V. Temperly, J. S. Rowlinson, and G. S. Rushbrook, Ed., North-Holland Publishing Co., Amsterdam, 1968.