

# A molecular dynamics simulation of a water droplet by the implicit-Euler/Langevin scheme

Tamar Schlick and Samuel Figueroa

*Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012*

Mihaly Mezei

*Department of Chemistry and Center for Study in Gene Structure & Function, Hunter College and the Graduate Center of CUNY, New York, New York 10021*

(Received 4 June 1990; accepted 26 October 1990)

Results are presented from potential energy minimization of water clusters and from molecular dynamics and Monte Carlo simulations of a liquid water droplet model. A new method for molecular dynamics—the implicit-Euler/Langevin scheme—is used in combination with a truncated Newton minimizer for potential energy functions. Structural and thermodynamic properties are reported for the scheme (with time steps of 5 and 10 fs), compared to a standard explicit formulation (with  $\Delta t = 1$  fs), to a Monte Carlo simulation, and to available experimental data. Results demonstrate that the implicit scheme is computationally feasible for large-scale biomolecular simulations, and that the droplet model can reasonably reproduce general structural features of liquid water. Results also show that the desired behavior is obtained from the implicit formulation: stability over large time steps, and effective damping of the high-frequency vibrational modes. Thus, major “bulk” properties of the system of interest may be observed more rapidly.

## I. INTRODUCTION

Liquid water has been investigated extensively through Monte Carlo and molecular dynamics simulations.<sup>1–10</sup> Since a large body of structural and thermodynamic data is available from experiment, simulations have been used to test various solvent potentials that will later be incorporated into protein and nucleic-acid force fields.<sup>5–10</sup> In addition, the unique properties of water—such as the complex structure of the liquid state, the polymorphs of ice, the liquid–water/ice relationship, and the anomalies of supercooled water—have always provided intriguing stimuli for new hypotheses and subsequent simulations.<sup>11–17</sup>

In this work, our goal is to test a new *numerical* algorithm for molecular dynamics based on the implicit-Euler integration scheme and the Langevin equation.<sup>18–21</sup> We have already applied the method successfully to smaller systems for which structural and energetic data are available.<sup>18,19,21</sup> It is our current goal to develop the method for efficient use in macromolecular simulations. Since competitive performance of the scheme depends on formulation of the computational model itself (due to added complexity—see Sec. II), our goal in the present work is also to establish a feasible (water-droplet) model and verify that it is physically reasonable. We anticipate investigating more complex models at a later stage.

The numerical methods will be presented in Sec. II. Details of the force field and program setup will be summarized in Sec. III. Results from potential energy minimization of water clusters will be discussed in Sec. IV. In particular, we will analyze the connectivity network of an energy-minimized 125-molecular cluster, since it is used: (1) as the initial configuration for the molecular dynamics and Monte Carlo simulations, and (2) for establishing a procedure for radial-distribution-function analysis. In Sec. V, we will discuss results from the molecular dynamics and Monte Carlo

runs. Findings from both our implicit scheme and a common explicit scheme will be compared to each other and to available data. The Monte Carlo simulation will further help place these differences in appropriate perspective. We will summarize our results in Sec. VI.

## II. METHODS

### A. The implicit-Euler/Langevin scheme

Our scheme addresses two related issues that are of general interest in molecular dynamics simulations: (1) using larger time steps than those typically used in explicit schemes, and (2) damping effectively the high-frequency vibrational modes. We achieve stability over larger time steps by employing the implicit-Euler scheme, well known in numerical analysis for solving stiff differential equations.<sup>22</sup> This scheme discretizes the differential equation (written in vector form)  $dy/dt = f(y)$  by the formula  $(y^{n+1} - y^n)/\Delta t = f(y^{n+1})$ . The quantity  $\Delta t$  is the time step, and  $y^n$  is the approximation to  $y(n\Delta t)$ . We achieve a frequency-discriminating damping by introducing a cutoff frequency parameter,  $\omega_c$ , and then exploiting the two different types of damping that enter into our scheme—frictional and intrinsic. This establishes a regime where modes with frequencies  $\omega \gg \omega_c$  are effectively frozen by the method, while modes  $\omega \ll \omega_c$  are fully activated.<sup>18,19</sup> In this way, some quantum effects can be mimicked.

The Langevin equation couples the molecular system to a thermal reservoir by mimicking effects of molecular collisions. A simple analytic formulation that has been used in molecular dynamics simulations is given by the following differential equation:<sup>23,24</sup>

$$M \frac{d^2 \mathbf{x}(t)}{dt^2} = -\mathbf{g}_E(\mathbf{x}(t)) - \gamma M \frac{d\mathbf{x}(t)}{dt} + \mathbf{r}(t). \quad (1)$$

In this equation,  $M$  is the diagonal mass matrix,  $\mathbf{x}$  and  $\mathbf{v}$  are the coordinate and velocity vectors, respectively, of the molecular system,  $\mathbf{g}_E$  is the gradient vector of the potential energy,  $\gamma$  is a collision parameter, and  $\mathbf{r}$  is a random force vector. The random force is a stationary, Gaussian process with mean zero and covariance matrix given by  $\langle \mathbf{r}(t)\mathbf{r}(t')^T \rangle = 2\gamma k_B T M \delta(t-t')$ . The quantities  $k_B$  and  $T$  denote Boltzmann's constant and the temperature in Kelvin, respectively. The random force is chosen in this way to balance the frictional damping of the energy with fluctuations of the random force (that may increase the energy) and to establish the desired equilibrium.

Discretization by the implicit-Euler scheme produces the following pair of differential equations:

$$M[(\mathbf{v}^{n+1} - \mathbf{v}^n)/\Delta t] = -\mathbf{g}_E(\mathbf{x}^{n+1}) - \gamma M \mathbf{v}^{n+1} + \mathbf{r}^{n+1}, \quad (2a)$$

$$(\mathbf{x}^{n+1} - \mathbf{x}^n)/\Delta t = \mathbf{v}^{n+1}, \quad (2b)$$

where

$$\langle \mathbf{r}^n \rangle = 0, \quad (3a)$$

$$\langle \mathbf{r}^n(\mathbf{r}^m)^T \rangle = 2\gamma k_B T M (\delta_{nm}/\Delta t). \quad (3b)$$

Note that, since both  $\mathbf{x}$  and  $\mathbf{v}$  are evaluated at step  $n+1$  in the right-hand side of (2a), a procedure for calculating  $\mathbf{x}^{n+1}$  must be formulated. Typically in implicit schemes this is achieved by solving a nonlinear system of equations at every step. We have shown that we can obtain  $\mathbf{x}^{n+1}$  in our case by minimizing the "dynamics" function,  $\Phi(\mathbf{x})$ , where  $\Phi(\mathbf{x}) = \frac{1}{2}(1 + \gamma \Delta t)(\mathbf{x} - \mathbf{x}_0^n)^T M(\mathbf{x} - \mathbf{x}_0^n)$

$$+ (\Delta t)^2 E(\mathbf{x}), \quad (4)$$

$$\mathbf{x}_0^n = \mathbf{x}^n + [\Delta t/(1 + \gamma \Delta t)](\mathbf{v}^n + \Delta t M^{-1} \mathbf{r}^{n+1}). \quad (5)$$

Note that  $\Phi(\mathbf{x})$  contains a quadratic, "kinetic" term and a potential-energy term.<sup>20</sup> The vector  $\mathbf{x}_0$ , defined above, is available through quantities computed in the previous step (hence the superscript  $n$ ). Each step of our procedure thus consists of minimizing  $\Phi(\mathbf{x})$  to obtain  $\mathbf{x}^{n+1}$  and then calculating  $\mathbf{v}^{n+1}$  from (2b). A good initial guess for a minimum of  $\Phi$  is  $\mathbf{x}_0^n$ . Any minimum of  $\Phi(\mathbf{x})$  will satisfy (2a).

To minimize  $\Phi$  efficiently, we have developed a truncated Newton method for potential energy functions.<sup>22,25,26</sup> The truncated Newton approach is attractive for large-scale applications because: (1) computational effort is concentrated in conformational regions near the minima, (2) separability of the Hessian into bonded and nonbonded terms is exploited to accelerate convergence, (3) computational requirements are versatile and manageable for large problems, and (4) the method retains the quadratic convergence of full Newton methods.

To achieve the damping of high-frequency modes, as discussed earlier, we choose  $\omega_c$  as

$$\omega_c = k_B T / \hbar \quad (6)$$

where  $\hbar$  is Planck's constant divided by  $2\pi$ . We then set  $\gamma$  to

$$\gamma = (\omega_c)^2 \Delta t. \quad (7)$$

We have shown theoretically and computationally that, for this choice, we can obtain for harmonic oscillators a parti-

tion of energies among the various modes that closely resembles the quantum-mechanical distribution. Low-frequency modes contribute nearly  $k_B T/2$  energy per mode, as predicted by classical statistical mechanics, while high-frequency modes contribute much less; a critical point for the distribution occurs near  $\omega = \omega_c$ .<sup>18,19</sup> Details of computational complexity and performance are described separately.<sup>20</sup>

## B. An explicit/Langevin scheme

For comparison, we investigate an explicit discretization of the Langevin equation (1) as well. Note that application of the explicit-Euler method [ $dy/dt = f(y) \Rightarrow (y^{n+1} - y^n)/\Delta t = f(y^n)$ ] would produce  $\mathbf{x}^n$  and  $\mathbf{v}^n$  in the right-hand sides of Eqs. (2a) and (2b), in place of  $\mathbf{x}^{n+1}$  and  $\mathbf{v}^{n+1}$ , respectively; however, once  $\mathbf{x}^{n+1}$  is solved by (2a), the existing form of (2b) can be used. The resulting scheme can be written as:

$$M[(\mathbf{v}^{n+1} - \mathbf{v}^n)/\Delta t] = -\mathbf{g}_E(\mathbf{x}^n) - \gamma M \mathbf{v}^n + \mathbf{r}^{n+1} \quad (8a)$$

$$(\mathbf{x}^{n+1} - \mathbf{x}^n)/\Delta t = \mathbf{v}^{n+1}. \quad (8b)$$

A convenient way of rewriting the explicit scheme (8a), (8b) so as to obtain  $\mathbf{x}^{n+1}$  and  $\mathbf{v}^{n+1}$  from  $\mathbf{x}^n$  and  $\mathbf{v}^n$  (we compute  $\mathbf{v}^n$  in order to calculate the kinetic energy at every step) is given by replacing (8a) with

$$\mathbf{x}^{n+1} = \mathbf{x}^n + \Delta t(1 - \gamma \Delta t)\mathbf{v}^n + \Delta t^2 M^{-1}[\mathbf{r}^{n+1} - \mathbf{g}_E(\mathbf{x}^n)]. \quad (9)$$

This scheme is related to the second-order accurate Verlet leap-frog scheme<sup>23</sup> used commonly for molecular dynamics without the friction and random forces. With our additional terms, however, the scheme is first-order accurate, as implicit Euler. The main difference between the implicit and explicit formulations is that, while implicit Euler is stable over a wide range of  $\Delta t$  choices, the explicit scheme is only stable when  $\Delta t$  is sufficiently small. Note that, throughout our discussion, superscripts of  $\mathbf{r}$  do not affect the scheme since  $\mathbf{r}$  is chosen independently at every step.

## III. THE MODEL AND PROGRAM

### A. The force field

We employ the SPC force field, an "effective pair potential" developed by Berendsen *et al.*,<sup>4</sup> known both for being computationally simple and physically reasonable.<sup>1</sup> Recently, the use of the SPC potential in nucleic acid simulations has demonstrated good numerical and physical behavior in relation to other potentials.<sup>8</sup>

The SPC potential consists of two intermolecular terms—van der Waals and electrostatic—as follows:

$$E_{\text{SPC}}(\mathbf{x}) = \sum_{\text{inter. oxygen pairs } (i,j), i < j} \left( \frac{-A}{r_{ij}^6} + \frac{B}{r_{ij}^{12}} \right) + \sum_{\text{inter. atom pairs } (k,l), k < l} \left( \frac{Q_k Q_l}{r_{kl}} \right). \quad (10)$$

The quantities  $r$  denote interatomic distances, and  $A$ ,  $B$ , and  $Q$  are energy parameters. Note that the first term is summed for all O—O atom pairs, while the second includes 4 H—H,

4 O—H, and 2 O—O interactions for each distinct molecular pair ( $i,j$ ). The vector  $\mathbf{x}$  represents the collective vector of Cartesian coordinates for the molecular system, from which the  $r_{ij}$ 's are determined [ $r = r(\mathbf{x})$ ]. It is crucial for this work to realize that parameters of the SPC potential were fitted to reproduce the experimentally known interaction energy and pressure for liquid water at 300 K as well as to provide reasonable agreement with the form of the radial distribution functions.<sup>4,27-31</sup> (Definition and details are discussed later.) SPC parameters were selected after performing 12 MD runs for 12 different points in the  $\{B, Q_H\}$  parameter space and then analyzing the thermodynamic and structural results. (The parameter  $A$  was fixed at a quantum-mechanical derived value, and  $Q_O$  was set to  $-2Q_H$ .) The complete resulting parameter set was:  $A = 625.5$  (kcal  $\text{\AA}^6$ )/mol,  $B = 629.4 \times 10^3$  (kcal  $\text{\AA}^{12}$ )/mol,  $Q_H = 0.41e$ ,  $Q_O = -0.82e$ . (The dielectric constant in the electrostatic potential is 1, and the numerical value of the factor necessary to produce energies in kcal/mol—with  $Q_O$ ,  $Q_H$  above and distances  $r$  in Angstroms—is 332.)

As for intramolecular geometries, the SPC potential assumes rigid O—H bond lengths,  $\bar{b}$ , of 1  $\text{\AA}$ , and tetrahedral H—O—H bond angles,  $\bar{\theta}$  ( $\cos \bar{\theta} = -1/3$ ,  $\bar{\theta} = 109.471^\circ$ ). However, actual enforcement of these geometries varies from formulation to formulation and may include soft constraints (restraints) on instantaneous values for  $\theta$ , the two  $b$ 's and/or on the (nonbonded) H—H distance in each molecule. In our MC simulation, intramolecular geometries were held rigid. In our MD implementation, we use soft constraints as a computational device in the following form:

$$E_{\text{intra}}(\mathbf{x}) = \sum_{\text{molecules } i} \left( S_1 \sum_{k=1,2} (b_{i,k}^2 - \bar{b}^2)^2 + S_2 (\cos \theta_i - \cos \bar{\theta}_i)^2 \right), \quad (11)$$

where  $b_{i,1}$ ,  $b_{i,2}$ , and  $\theta_i$  are the two O—H bond values and the H—O—H angle, respectively, in each molecule  $i$ . These forms of penalty potentials have been used in our nucleic acid program, MADPAC, because of noted computational and physical advantages.<sup>32,33</sup> Clearly, for small fluctuations about target values, the form is unimportant and parameters can be chosen to match standard harmonic potentials.<sup>33</sup> Here we set  $S_1$  and  $S_2$  to values corresponding to experimentally derived values:  $S_1 = 94$  kcal/(mol  $\text{\AA}^4$ ) and  $S_2 = 42$  kcal/mol, corresponding to the harmonic-potential parameters of 750 kcal/(mol  $\text{\AA}^2$ ) and 75 kcal/(mol  $\text{rad}^2$ ), respectively.<sup>34</sup>

## B. The program setup

Our program sets up water clusters of any given size, performs energy minimization or molecular dynamics with the potential described above, and analyzes resulting structural and energetic properties. For the cluster models, no periodic boundary conditions are used, but *all* pairwise interactions are computed. A model with boundary conditions will be considered later. Our goal here was to test the simplest possible model with the implicit-Euler/Langevin scheme so that long runs could be investigated; boundary

conditions would complicate and thus slow down the minimization phase (at each step) considerably. Thus, when we incorporate periodic boundary conditions, we anticipate implementing a fast multipole method for the electrostatic interactions.<sup>35</sup> Multivariate minimization of the potential energy and the dynamics function (in our implicit MD scheme) is performed with our large-scale truncated-Newton method.<sup>25</sup> The truncated Newton algorithm is especially efficient for these water models because the preconditioning matrix (roughly, a matrix chosen as a sparse approximation to the Hessian in order to accelerate convergence) is the  $9 \times 9$  diagonal-block matrix corresponding to the intramolecular (soft-constraint) forces. An efficient factorization of this matrix can then be performed without the need for variable reordering.<sup>25</sup>

Initial coordinates for an ensemble of water molecules are computed as follows. First, we specify the desired number of molecules in each coordinate direction (for example,  $n_x = n_y = n_z = 5$  sets up a cubic domain of 125 molecules). We also specify a desired O—O distance,  $d$  (typically  $d = 2.75$   $\text{\AA}$ ). Second, we position the oxygens in centers of cubes throughout the computational domain. Third, we position the two hydrogens of each molecule "pseudorandomly" in a unit sphere about their oxygen so that  $\theta = \bar{\theta}$  (the target tetrahedral value). More generally, arbitrary values for  $\bar{b}$  and  $\bar{\theta}$  can be specified. This procedure can be summarized in the following steps.

(1) Initialize  $n_x, n_y, n_z, d, \bar{\theta}, \bar{b}$ .

Set  $N_{\text{mol}} = n_x n_y n_z$ .

(2) Set positions of the oxygens:

$\mathbf{O}_i = (kd, ld, md)^T$ ,

$\left\{ \begin{array}{l} k = 0, \dots, n_x - 1 \\ l = 0, \dots, n_y - 1 \\ m = 0, \dots, n_z - 1 \end{array} \right\}$ , for  $i = 1, \dots, N_{\text{mol}}$ .

(3) Set positions of the hydrogens (see Fig. 1). For each molecule  $i$ , choose three pseudorandom numbers  $\{\alpha, \beta, \gamma\}$  uniformly between 0 and  $2\pi$ . Assume for now that our O (of the present molecule  $i$ ) is located at the origin. Set

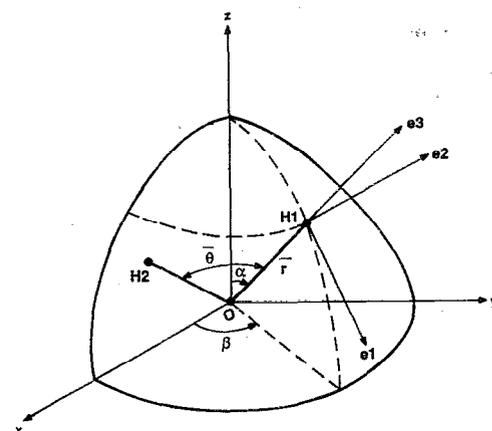


FIG. 1. Algorithm for positioning water molecules pseudorandomly in the computational domain (see text for details).

$$e_3 = (\cos \alpha \sin \beta, \sin \alpha \sin \beta, \cos \beta)^T,$$

and then set the first hydrogen vector,  $\mathbf{H1}$ , to  $\mathbf{H1} = \bar{r}e_3$ . Now construct a right-handed coordinate system  $\{\mathbf{e1}, \mathbf{e2}, \mathbf{e3}\}$  by choosing

$$\mathbf{e1} = (1/\sin \beta) (-e_3(2), e_3(1), 0)^T$$

and

$$\mathbf{e2} = \mathbf{e3} \times \mathbf{e1}.$$

Now position the second hydrogen vector,  $\mathbf{H2}$ , at

$$\mathbf{H2} = (\bar{r} \cos \gamma \sin \bar{\theta})\mathbf{e1} + (\bar{r} \sin \gamma \sin \bar{\theta})\mathbf{e2} + (\bar{r} \cos \bar{\theta})\mathbf{e3}.$$

By construction,  $\mathbf{H1} \cdot \mathbf{H2} = \|\mathbf{H1}\| \|\mathbf{H2}\| \cos \bar{\theta}$ . Finally (to remove the O-at-origin assumption above), translate the hydrogen vectors by their oxygens:  $\mathbf{H1} \leftarrow \mathbf{H1} + \mathbf{O}$ ,  $\mathbf{H2} \leftarrow \mathbf{H2} + \mathbf{O}$ .

This procedure is especially convenient for performing energy minimizations on water clusters where various starting configurations are desired. Not only can we set up different spatial arrangements for the same number of molecules (e.g., 16 molecules can be positioned in a linear array, a  $4 \times 4$  "sheet", or a  $4 \times 2 \times 2$  "box") but the O—O separation parameter,  $d$ , can be used to disperse the molecules at will.

#### IV. ENERGY MINIMIZATION RESULTS

Minimized configurations of water clusters are worth examining first to better understand the molecular dynamics results that follow. In Table I we summarize results from minimization of water clusters of various sizes. Figures showing some of these minimized structures can be obtained from an author (T. Schlick) upon request. For a given number of molecules, different minima were obtained from dif-

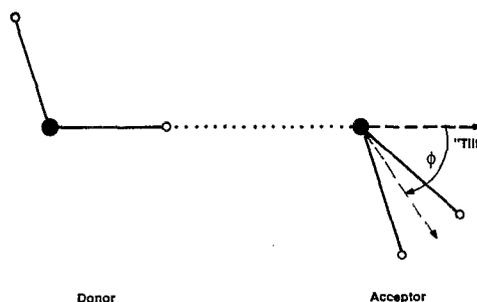


FIG. 2. Definition of the tilt angle in a linear hydrogen-bonded water dimer. The tilt is the angle between the hydrogen bond vector and the bisector of the acceptor molecule.

ferent starting configurations. For each calculated minimum, the energy, the number of hydrogen bonds (H bonds), and the mean values of  $b$  and  $\theta$  are listed. Additionally, geometry for the H bonds is described by the following quantities: (1) mean values of the O...O distance, (2) O—H...O angle, and (3) tilt angle  $\phi$  defined as the angle between the H-bond vector and the bisector of the acceptor molecule (see Fig. 2).

The following trends can be identified for the energy-minimized water clusters. Intramolecular O—H bonds tend to stretch to a value of 1.02 or 1.03 Å, and H—O—H angles attain values around 105° rather than the 109° (the target). Indeed, experimental and theoretical evidence for the ice polymorphs support the attainment of such O—H distances and H—O—H angles, as found in the "isolated" water molecule [Ref. 12, pp. 79,91]. Mean O...O distances for hydrogen-bonded molecular pairs range from 2.69 to 2.75 Å. Mean H-bond angles range from 163° to 176°, exhibiting a clear deviation from the ideal, linear H-bond geometry. Val-

TABLE I. Minimized water clusters.

# Mol	$E$ (kcal/mol)	#H Bonds	Intramolecular geometry		H-Bond Geometry		
			$b$ (Å)	$\theta$ (°)	$r_{O \cdots O}$ (Å)	$\theta_{O-H \cdots O}$ (°)	$\phi$ (°)
2	-6.95	1	1.01	108	2.73	176	22
4	-32.52	4	1.02	105	2.69	174	35
5	-38.87	5	1.02	106	2.69	174	29
6	-52.99	7	1.02	105	2.71	171	37
8	-75.00	12	1.02	105	2.75	163	56
	-72.49	10	1.02	105	2.72	169	39
16	-172.24	27	1.03	104	2.75	160	56
	-163.24	23	1.02	104	2.73	169	41
	-158.39	24	1.02	105	2.73	166	50
	-154.58	24	1.02	105	2.74	165	49
25	-266.41	39	1.02	104	2.73	167	46
	-264.79	39	1.02	104	2.73	166	48
	-259.74	40	1.02	104	2.74	163	52
	-255.39	36	1.02	105	2.73	167	43
27	-305.20	48	1.02	105	2.74	164	51
	-300.83	45	1.03	104	2.73	167	50
	-298.34	46	1.02	105	2.74	165	50
	-295.76	45	1.03	104	2.73	165	50
	-295.50	44	1.03	104	2.74	166	47
125	-1501.83	228	1.03	104	2.74	166	50

TABLE II. Number of hydrogen bonds per molecule in the energy minimized cluster of 125 molecules.

$N$	# molecules with $N$ H-bonds	%
0	0	0
1	0	0
2	6	5
3	39	31
4	73	58
5	7	6
6	0	0

ues of the tilt angles range from  $22^\circ$  (in the dimer) to  $56^\circ$ . All these values fall within expected ranges.<sup>3,4</sup> In particular, as the cluster size grows, we note increases in  $O \cdots O$  distances, larger deviations from linear H-bond angles, and attainment of tilt values near  $50^\circ$ . Indeed, the "bent" H-bond geometry ( $\sim 15^\circ$  deviation from linearity) characterizes the more random network of H-bonds in liquid water, where a complex mixture of ordered/disordered H-bonded clusters prevails.<sup>11-13,16</sup>

In comparison with the  $O \cdots O$ , tilt, and energy values obtained for the rigid SPC potential applied to a water dimer ( $2.75 \text{ \AA}$ ),  $26^\circ$ ,  $-6.6 \text{ kcal/mol}$ , respectively,<sup>1</sup> we note small differences, as expected, from our added intramolecular flexibility.

For the 125-molecular cluster, one minimization run was performed from a starting configuration corresponding to a  $5 \times 5 \times 5$  cube. The corresponding minimum structure was subsequently used for the molecular dynamics and Monte Carlo simulations. In addition to the structural and energetic features of this minimum listed in Table I, we describe in Table II its H-bond distribution. Even with the inevitable surface effects in our droplet model, we can observe an overall tetrahedral coordination structure, similar to the coordination in ice. To analyze the neighbor distribution in more detail, we show in Fig. 3 a histogram of the number of neighbors present in an inner subcluster (this removes surface effects). Good agreement can be noted with this description and the one determined experimentally for water at  $4^\circ \text{C}$ .<sup>12,13,27,28</sup> It is also interesting to note the com-

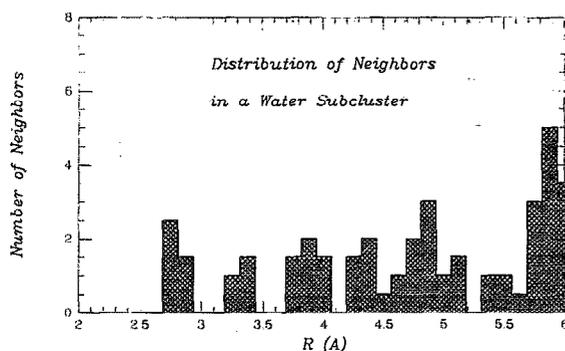


FIG. 3. Distribution of neighbors in an inner water subcluster of the energy-minimized water cluster (of 125 molecules). Data were collected by counting the number of oxygen neighbors in a volume element (bin)  $R \text{ \AA}$  away from a central oxygen.

mon features between our histogram and the  $O-O$  distribution in the common polymorphs of ice, namely I, II, and III (see Fig. 6 of Ref. 13). For these ices,  $O-O$  peaks occur near  $2.75 \text{ \AA}$  (for ice I, II, III), near  $3.6 \text{ \AA}$  (ice II),  $4.3 \text{ \AA}$  (ice III),  $4.5 \text{ \AA}$  (ice I), and  $5.25 \text{ \AA}$  (ice I). The primary features of  $O-O$  distance distributions in low-temperature water structures are: (1) a first-neighbor peak around  $2.75\text{--}2.9 \text{ \AA}$ ; (2) a broad, next-neighbor peak around  $4.5 \text{ \AA}$  (to  $5.0 \text{ \AA}$ ); and (3) a significant group of non-nearest neighbors in the interval  $3.0\text{--}4.0 \text{ \AA}$ , with a peak near  $3.6 \text{ \AA}$ . Thus, the distorted H-bond structure along with the increase in H-bond lengths permit these more dense arrangements, as found in the denser polymorphs of ice.<sup>13</sup> Clearly, results for the MD simulation at room temperature will tend to smooth out these distributions and lower the peak heights.

In order to determine an appropriate procedure for calculating radial distribution functions from our simulations of the droplet model, the following analysis was performed. In Fig. 4, we show a distribution of the number of molecular neighbors each molecule has within a  $6 \text{ \AA}$  sphere. For example, there are two (innermost) molecules with 38 molecular neighbors in a  $6 \text{ \AA}$  sphere centered around each; there are 9 molecules that have 20 neighbors in  $6 \text{ \AA}$  spheres surrounding each of them. For obtaining reliable statistics over the run, it is desirable to average properties for an inner subcluster that is not subject to surface effects on one hand, but is sufficiently large, on the other, to avoid measuring mere random fluctuations of one subgroup. In particular, to be able to resolve well first neighbors and next-neighbor peaks, the subcluster should include molecules that have at least 16 neighbors within a  $6 \text{ \AA}$  sphere, corresponding to an approximate tetrahedral-coordination structure. We can thus choose a number  $m$  (corresponding to the abscissa of Fig. 4) and include in the simulation averages over all molecules that have  $m$  neighbors or more within a  $6 \text{ \AA}$  sphere. For the energy minimized structure, the sample size corresponding to Fig. 4 is shown in Fig. 5. For example, a choice of  $m = 30$  will select 20 inner molecules for our sample, while  $m = 20$  will pro-

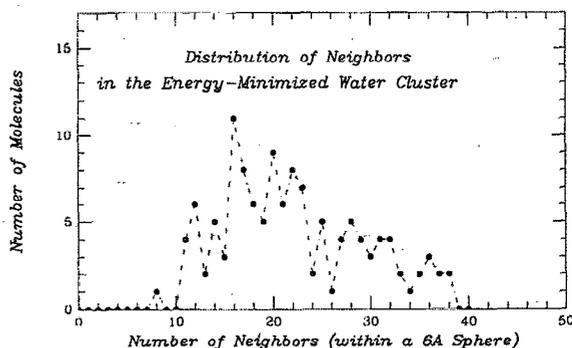


FIG. 4. Distribution of neighbors throughout the energy-minimized water cluster of 125 molecules. Data were collected by counting the number of oxygen neighbors each molecule had within a  $6 \text{ \AA}$ -sphere about its oxygen. Thus, there are few molecules that have many neighbors (only the innermost subgroup) and, correspondingly, few molecules that have few neighbors (only the surface, outermost subgroup). We use molecules in the right "tail" of this distribution for our set of accumulating interatomic distances for the MD runs—see Fig. 5.

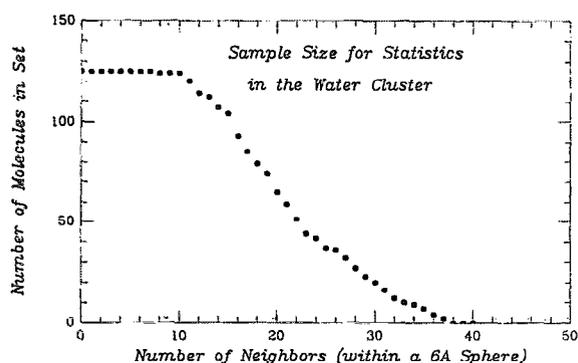


FIG. 5. Size of the sample set, corresponding to Fig. 4, when we require that only molecules with  $m$  neighbors or more—within a 6 Å sphere—will be considered. For example, there would be only 10 molecules in our set if we require  $m = 34$  but 64 molecules when  $m = 20$ .

duce 64. All 125 molecules will be considered when  $m = 7$ .

This analysis shows that approximately 2/3 of the molecules in the minimized cluster should provide useful information on distribution of first and second neighbors. This suggests that qualitative agreement with experimental radial distribution functions may be obtained from the simulation with this procedure. Since the distributions shown in Figs. 4 and 5 are likely to change somewhat during the MD run (the distribution, analogous to Fig. 4, tends to a sharper, more clustered Gaussian-like curve—see Table III), we choose the value  $m = 20$ . In addition, we impose the restriction that our sample size for the statistics contain at least 40 mole-

cules. (If it is less,  $m$  is decreased until the sample size is 40 or more.) This procedure resulted in an average sample size of 45–55 molecules during all our runs.

## V. RESULTS AND DISCUSSION

One Monte Carlo (MC) simulation and five MD simulations were performed from the starting point described in the previous section. We refer to the explicit/Langevin scheme described in Sec. II as “explicit” and to the implicit-Euler/Langevin scheme as “implicit.” For the explicit scheme, we performed two runs—with a time step of 1 fs ( $10^{-15}$  s)—for  $T = 300$  K and  $T = 325$  K. For the implicit scheme, we performed three runs:  $T = 300$  K,  $\Delta t = 5$  fs;  $T = 300$  K,  $\Delta t = 10$  fs; and  $T = 325$  K,  $\Delta t = 5$  fs. All these runs covered a period of 20 ps.

The main purpose of the MC simulation was to provide further perspective on the explicit versus implicit results. The MC run was performed with the force-biased MC technique<sup>36,37</sup> and covered one million configurations, after 200 000 steps of equilibration. The force field was also SPC, as in the MD runs, but intramolecular geometry was held rigid. As for the MD runs, all pairwise interactions were computed, and no boundary conditions were imposed. The same starting point and procedure for accumulating bin data for the  $g(R)$ 's were also used.

Simulation results are summarized in Table III and Figs. 6–14. Below, we discuss the following issues in turn: thermodynamics, overall structure, and radial distribution functions. Structural and energetic details, in addition to ra-

TABLE III. Average quantities from the 5 MD runs and 1 MC run.<sup>a</sup>

	Starting Point	$T = 300$ K				$T = 325$ K	
		Exp. 1 fs	Imp. 5 fs	Imp. 10 fs	MC	Exp. 1 fs	Imp. 5 fs
$E_k$	0	342	88	84		369	101
$E_p$	-1502	-1094	-1442	-1427	-1072	-1029	-1427
$E_{NB}$	-1641	-1309	-1577	-1558		-1247	-1561
$E$	-1502	-752	-1354	-1343		-659	-1325
$\bar{E}^b$		(-760)	(-1369)	(-1342)		(-658)	(-1331)
$N_{HB}$	228	213	226	224	207	206	226
$N_{HB/atom}$	3.7	3.4	3.6	3.6	3.3	3.3	3.6
H-bond bins							
0	0	0	0	0	0	0	0
1	0	1	0	0	3	3	0
2	5	18	9	9	19	20	10
3	31	33	31	30	33	32	27
4	58	40	54	52	35	35	56
5	6	9	7	7	9	9	7
6	0	1	0	1	1	1	0
O...O	2.74	2.85	2.78	2.79	2.90	2.87	2.79
O-H...	166	157	165	165	155	156	164
$\phi$	49	49	48	48	49	49	48
H-O-H	104	105	104	104		105	104
O-H	1.03	1.02	1.03	1.03		1.02	1.03

<sup>a</sup> All energies are given in kcal/mol, distances in Å, and angles in degrees. Error estimates for the energies were calculated by the method of batch means and are  $\pm 8$  kcal/mol for all MD runs and  $\pm 7$  kcal/mol for the MC run.

<sup>b</sup> Averages from the last part of the simulation (over 5 ps).

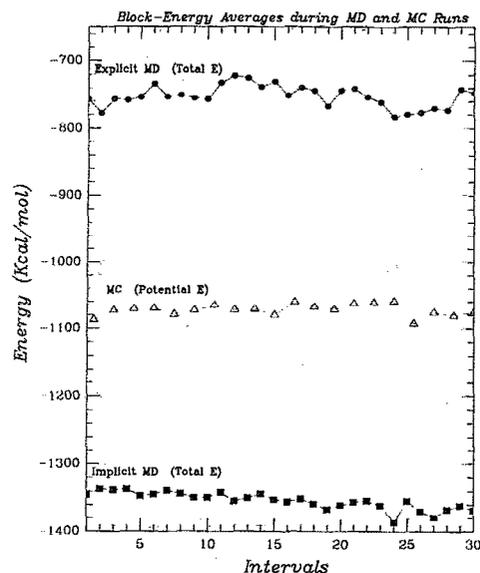


FIG. 6. Energy averages for subintervals during two MD runs and one MC run. For the MD runs at 300 K (explicit, and implicit with  $\Delta t = 5$  fs), points correspond to averages of the total energy during each 0.5 ps interval from 5–20 ps. For the MC run, points correspond to averages of the potential energy for each 500 000-configuration segment.

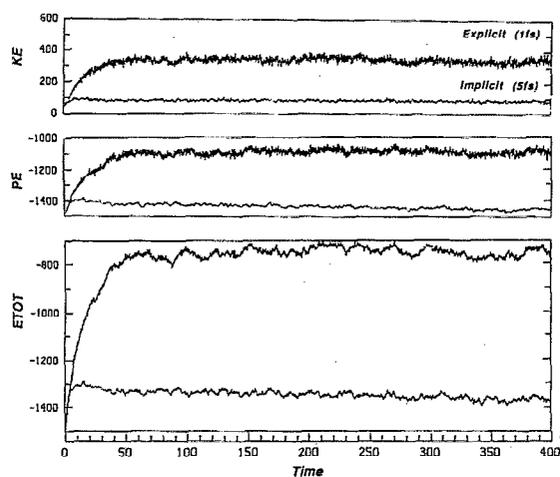


FIG. 7. Kinetic, potential, and total energy accumulated over two 20 ps MD runs at  $T = 300$  K: explicit scheme,  $\Delta t = 1$  fs; implicit scheme,  $\Delta t = 5$  fs. Our unit of time, shown here and in all subsequent plots, is 0.05 ps.

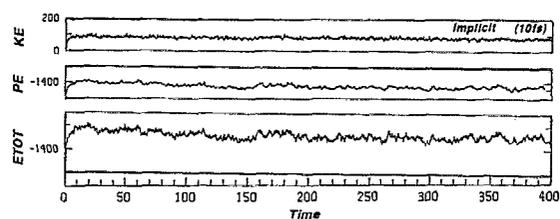


FIG. 8. Kinetic, potential, and total energy, accumulated over the 20 ps MD run at  $T = 300$  K with the implicit scheme,  $\Delta t = 10$  fs.

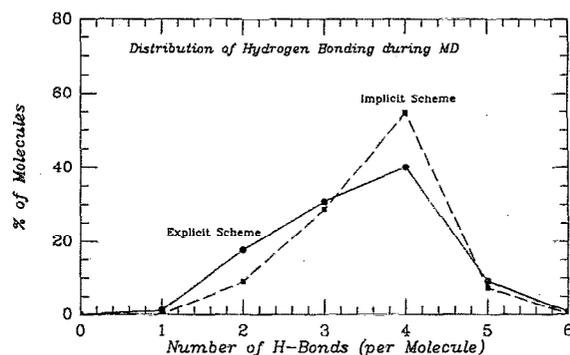


FIG. 9. Distribution of hydrogen bonding patterns during molecular dynamics—implicit versus explicit scheme. The explicit run corresponds to  $T = 300$  K,  $\Delta t = 1$  fs, and the implicit run to  $T = 300$  K,  $\Delta t = 5$  fs. Data were averaged over the 20 ps runs from the number of hydrogen-bonded neighbors each molecule had in every configuration. The basic tetrahedral coordination structure can be seen from these distributions, with a sharper distribution for the implicit run.

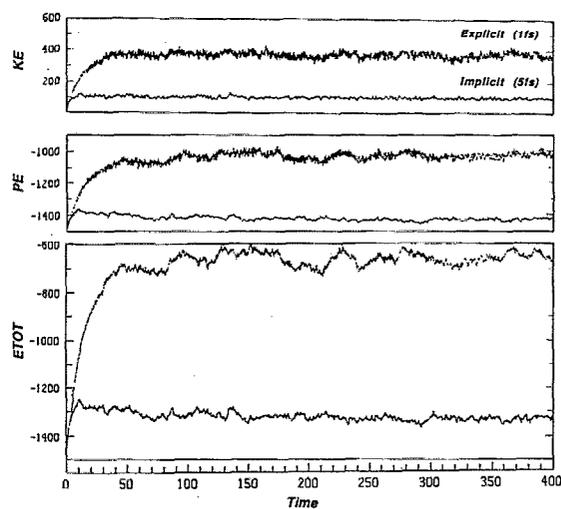


FIG. 10. Kinetic, potential, and total energy, accumulated over two 20 ps MD runs at  $T = 325$  K: explicit scheme,  $\Delta t = 1$  fs; implicit scheme,  $\Delta t = 5$  fs.

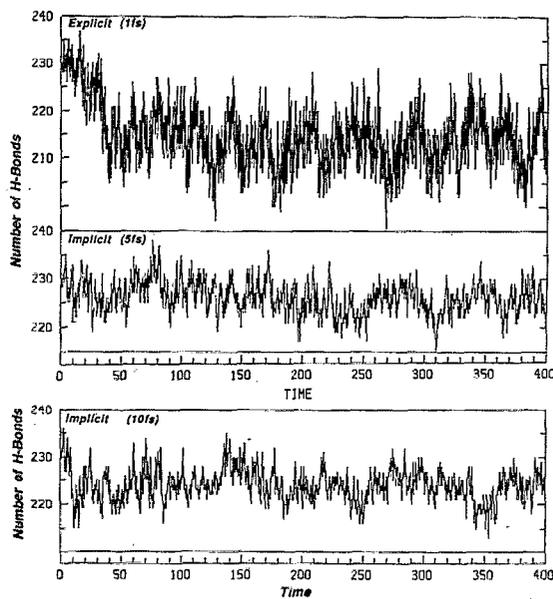


FIG. 11. Number of hydrogen bonds as a function of time for three 20 ps MD runs at  $T = 300$  K: explicit scheme,  $\Delta t = 1$  fs; implicit scheme,  $\Delta t = 5$  fs; and implicit scheme,  $\Delta t = 10$  fs. The hydrogen-bond classification is described in the text.

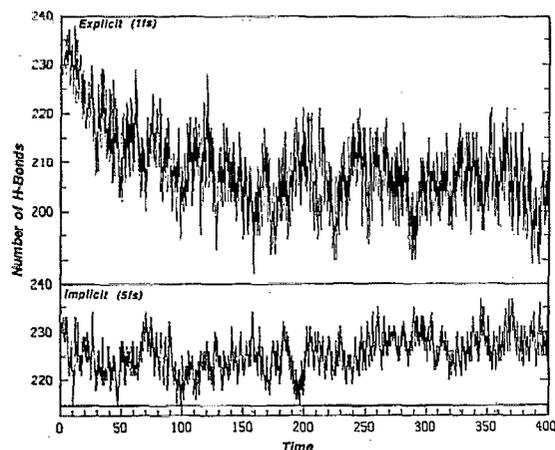


FIG. 12. Number of hydrogen bonds as a function of time for two 20 ps MD runs at  $T = 325$  K: explicit scheme,  $\Delta t = 1$  fs; implicit scheme,  $\Delta t = 5$  fs.

dial distribution functions, are important for understanding the behavior of the implicit scheme in relation to the explicit and MC runs. Computational differences will be discussed along with the summary, in the last section.

### A. Thermodynamics

In Table III, we provide mean values of the energy for each run: kinetic, potential, nonbonded, and total. These means were averaged over the run period 5–20 ps. Error

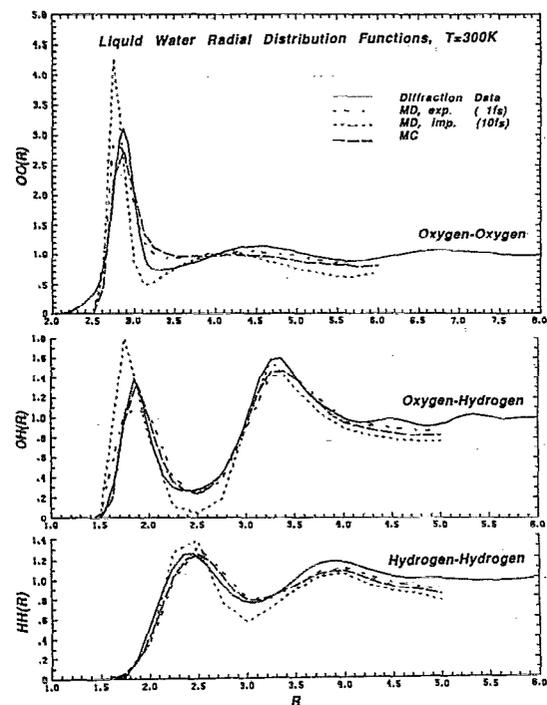


FIG. 13. Computed radial distribution functions for O—O, O—H, and H—H from two MD runs at 300 K (explicit,  $\Delta t = 1$  fs; implicit,  $\Delta t = 10$  fs) and MC, shown against the latest x-ray diffraction data of Soper and Phillips [Chem. Phys. 107, 47 (1986)]. Distances are given in Angstroms. Details of the computations are provided in the text. Results from the implicit MD run with  $\Delta t = 5$  fs are almost identical to the 10 fs run, so corresponding curves are omitted for clarity.

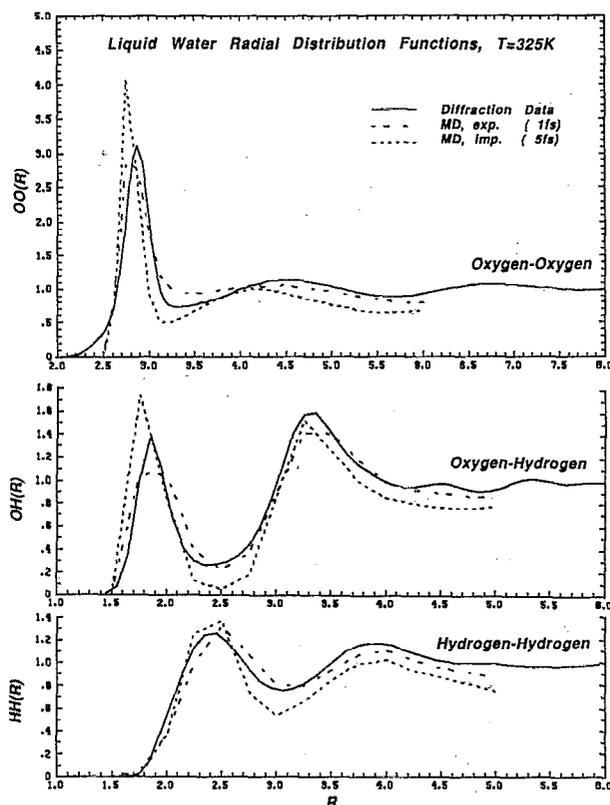


FIG. 14. Computed radial distribution functions for O—O, O—H, and H—H from two MD runs at 325 K (explicit,  $\Delta t = 1$  fs; implicit,  $\Delta t = 5$  fs), shown against the x-ray results of Soper and Phillips. Distances are given in Angstroms.

estimates for the energies were calculated by the method of batch means<sup>38,39</sup> and represent two standard deviations. These estimates are  $\pm 8$  kcal/mol for all the MD runs and  $\pm 7$  kcal/mol for the MC run. The resulting block averages for the MC run and two representative MD runs are shown in Fig. 6. For comparison, corresponding energy values for the starting configuration are also listed in Table III.

The following trends can be noted. First, the energy differences between the explicit and implicit runs are large. All values for the implicit scheme are lower (see Figs. 6–8). This behavior is expected, as the high-frequency modes are effectively damped out by the method. Second, larger relative differences occur for the kinetic energy rather than the potential energy components. In other words, more kinetic energy than potential energy is dissipated by the implicit scheme. Overall, the configurations resulting from the implicit scheme stay closer to the initial structure, while greater deviations can be observed for the explicit scheme. This can also be seen from Fig. 9. The average potential energy from the MC run is close to the potential energy of the explicit run at 300 K.

Another important observation is that energies from the implicit simulation with a time step of  $10^{-14}$  s (one order of magnitude higher than typical time steps) are very close to results obtained with 1/2 that value (see Figs. 7 and 8). This demonstrates the high stability of the implicit-Euler scheme. Somewhat higher energies are predicted from the implicit run with the larger time step, and this can be explained by the

fact that greater "jumps" in the configuration/time space bring about larger perturbations from the initial structure. In the minimization phase, a larger  $\Delta t$  can produce a solution  $\mathbf{x}^{n+1}$  more distant than  $\mathbf{x}^n$  in configuration space.

When comparing the two corresponding sets of MD runs at different temperatures, we note the expected increases in energies for the higher-temperature simulations (see Figs. 7 and 10). In particular, the implicit scheme produces a larger percent of increase in the kinetic component.

The total energy values appearing in parentheses for each run in Table III are averages obtained in the last part of the simulation. From these values, we can roughly estimate heat capacities at constant pressure for the two schemes. We obtain the values  $C_p = 30$  cal/(mol K) from the explicit scheme, and  $C_p = 12$  cal/(mol K) from the implicit scheme. The experimental value in this temperature range of 300–325 K is nearly constant and has a value of 18 cal/(mol K).<sup>12</sup> This experimental value serves as an upper bound, because the heat capacity for a droplet is expected to be lower than for the bulk liquid. This suggests that classical Newtonian molecular dynamics tends to overestimate the accessible energy, while quantum effects, some of which can be generally mimicked by our implicit scheme,<sup>18</sup> produce lower expected energies. In the case of these measured heat capacities, the value of  $C_p$  estimated from the implicit scheme comes to greater agreement to the experimental value than  $C_p$  from the explicit scheme. This provides a good qualitative test for the amount of damping introduced by our implicit formulation.

## B. Overall structure

The structural features summarized in Table III and Figs. 9 and 11–14 explain the differences in results among the various simulations, as described above. Basically, the damping in the explicit scheme produces configurations that have closer structural similarities to the initial structure. The explicit MD and MC runs produce similar configurations.

The number of H-bonds, as defined by our criteria in the previous section, is 228 for the initial structure and averages 207, 213, and 225 for the MC, explicit, and implicit runs, respectively, at room temperature. Thus, a greater number of H bonds is retained in the implicit MD (see Figs. 11 and 12). In terms of the distribution of the number of H bonds per molecule (see Table III), we see a greater percentage of molecules involved in 4 H-bonds for the implicit than for the explicit (54% vs 40%). These distributions can be seen more clearly in Fig. 9. The MC H-bond distribution is very similar to that of the explicit MD. The mean H-bond O...O lengths are 2.85 Å for the explicit scheme and 2.9 Å for the MC, compared to 2.79 Å for the implicit scheme and 2.74 Å for the minimum structure. Recall that our intramolecular flexibility permits closer approaches through accommodation of nearby geometry. For the higher-temperature runs, these O...O values increase slightly, as expected. As for the typical H-bond angles, results from the explicit scheme and the MC predict over 20° bending from the ideal linear arrangement, while the implicit scheme predicts a 15° bending. Both these values fall well within ranges observed for liquid

water. All H-bond tilt angles favor values of 48°–49°. Intramolecular geometry in the MD runs is uniformly characterized by H—O—H bond angles of 104°–105° and O—H bond lengths of 1.02–1.03 Å.

## C. Radial distribution functions

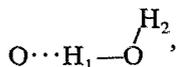
In Figs. 13 and 14 we show the computed radial distribution functions from the explicit, implicit, and MC runs at 300 and 325 K. As discussed earlier, bin data were accumulated for a subcluster at each time-step configuration. Typically, more than 1/3 of the entire cluster of 125 molecules was considered. Bin widths for all the MD runs were 0.125 Å for O—O, and 0.25 Å for O—H and H—H. For the MC run, since only cumulative data were obtained, all bin widths were 0.125 Å.

The radial distribution function  $g(R)$  measures how the local density of neighbors around a given molecule differs from the "bulk" density. It thus provides a good measure of positions and degree of occupancy of molecular neighbors. Various x-ray and neutron diffraction experiments have provided data to estimate interatomic distances in liquid water for O—O, O—H, and H—H pairs.<sup>27–31</sup> The experimentally derived value,  $\rho_{AB}(R)$ , represents the average number of atomic neighbors B at a distance  $R$  from atom A of a central molecule. The function known as  $g_{AB}(R)$  is obtained by dividing  $\rho(R)$  by  $\rho_0$ , the bulk density (expressed as number of molecules per unit volume). Thus, a plot of  $g(R)$  vs  $R$  is greater than one where the local density is greater than the bulk density, and it should tend to one as  $R$  increases.

From the MD and MC simulations, we accumulate bin data for a function  $\beta_{AB}(R)$ , defined as the average number of atomic neighbors B in a  $\Delta R$  volume element where  $\Delta R$  is our chosen bin width. For comparison with  $\rho(R)$ ,  $\beta(R)$  is divided by  $V(R)$ , the bin volume ( $V(R) = \frac{4}{3}\pi[R^3 - (R - \Delta R)^3]$ ). Furthermore, for comparison with the experimental function  $g(R) = \rho(R)/\rho_0$ , division by  $\tilde{\rho}_0$ , a normalizing factor, is needed:  $\tilde{g}(R) = \beta(R)/[V(R)\tilde{\rho}_0]$ . We estimate  $\tilde{\rho}_0$  for the droplet by calculating the average number of molecules per unit volume, from a central molecule, at increasing values of  $R$ . These numbers tend to some value just before surface effects enter (and the local density goes down monotonically). The values used in the present study are 0.032 molecules/(Å)<sup>3</sup> for the explicit scheme and the MC, and 0.038 molecules/(Å)<sup>3</sup> for the implicit scheme. These numbers are different because the density distributions are different (see Fig. 9).

The experimental curves for  $g(R)$  from the most recent works of Soper and Phillips<sup>31</sup> for O—O, O—H, and H—H are shown as solid curves in the radial-distribution-function plots in Figs. 13 and 14. These functions support the view that liquid water molecules form a connected network of H bonds with a local tetrahedral coordination that undergoes continuous architectural rearrangements. Changes result from the competition between compact, local H-bond arrangements (as in the ices) and the many ways of forming longer-range interactions between various H-bonded subclusters.

The main features of the O—O distribution is a first peak centered at  $\sim 2.8$  Å, corresponding to the nearest neighbors in a tetrahedrally coordinated arrangement. A second, broader peak is centered at about 4.5 Å, corresponding to second-neighbor separations in the same tetrahedral arrangement. The two highest peaks in the O—H curves occur at values near 1.85 and 3.35 Å. They correspond to the two O—H distances between the O of the acceptor molecule and the two H's of the donor molecule. Thus for the H-bond sequence



the first peak corresponds to O—H<sub>1</sub>, and the second peak corresponds to O—H<sub>2</sub>. Lower peaks occur for O—H distances between second molecular neighbors. The sharp, first O—H peak is consistent with a near-linear H-bond arrangement. In correspondence to the geometry described above, H—H peaks occur near 2.4 and 3.8 Å and are associated with the two distances originating from one H, of the acceptor molecule of an H-bonded pair: one to the donor H (H<sub>1</sub> above), and one to the nondonor H (H<sub>2</sub> above), of the donor molecule.

Our computed  $\bar{g}(R)$ 's for the different runs are compared with the experimental curves in Figs. 13 and 14. The functions obtained from the two implicit MD runs with time steps of 5 fs and 10 fs were almost identical, so one representative set of curves is shown. No smoothing was applied to these curves; points were plotted at the bin points and then connected by dashed lines.

Overall, we can note a good agreement between the explicit and experimental results. The O—O curve is particularly well reproduced since bin widths were small. (Coarser resolution of the O—H and H—H bins was used to reduce the volume of data generated from the MD runs). Surface effects of the droplet model begin to enter at about 6 Å, but up to that point locations and peak heights are reproduced well.

The computed  $g(R)$ 's from the MC run also exhibit good agreement with the experimental curves. These MC results are very similar to results of the explicit MD run. Since rigid intramolecular geometry was used in the MC run, the first O—O peak occurs at a slightly higher value of  $R$  than for the MD runs. The smaller bin widths used for O—H and H—H in the MC run produced smoother curves. The second O—O peak from MC is, however, somewhat flatter than that from our MD. Nonetheless, these basically similar results from two different computational techniques suggest that the droplet model, with the associated statistical procedure for accumulating bin data, is the major factor influencing the results. The MD technique and intramolecular flexibility may produce "finer tuning" and finer resolution of the data.

Results from the implicit runs exhibit radial distribution functions with different patterns than the MC and explicit MD: higher first peaks and lower populations of neighbors between the first two peaks. What is *crucial* for correct interpretation of these results is realizing that parameters of the SPC model were optimized to reproduce agreeable radial distribution functions.<sup>4</sup> Results with SPC by the implicit

scheme demonstrate that we indeed obtain the expected behavior: damping high-frequency modes. This is expressed in a more compact and ordered H-bond network, where the effects of the long-range competition between various subclusters—that lead to looser H bonds and more flexible geometries for H-bond networks—are reduced. Clearly, in order to get a better agreement between the implicit simulation results and the experimental  $g(R)$ 's, reparametrization of the water potential is required, as was done in the explicit case. Incorporation of periodic boundary conditions may further improve results. Nonetheless, the better agreement we obtain in the estimates for the heat capacity  $C_p$  (from  $\Delta E/\Delta T$ ) with the implicit scheme suggests that the scheme may provide better estimate of relative energy differences for different simulation temperatures.

As for differences in the radial distribution functions among the two different temperatures and among the two different time steps in the implicit scheme, only small differences can be observed. At higher temperatures (Fig. 14), peak heights are lowered, and the distribution is smoothed out further, indicating that the tetrahedral coordination in water is weaker at higher temperatures. The  $g(R)$ 's obtained from the two implicit runs produce almost identical distributions. Again, this illustrates the high stability of the implicit Euler discretization.

## VI. SUMMARY

Our liquid water simulations with the SPC pair potential by the implicit-Euler/Langevin versus the Verlet/Langevin scheme and MC have provided interesting answers to the following questions.

(1) Can a simple model be formulated that will be computationally feasible for the implicit scheme and still provide good agreement with available physical data? The water droplet model has here been tested as a first approach.

(2) Will the implicit scheme be computationally stable over time steps one order-of-magnitude greater than the explicit scheme?

(3) What will the energetic and structural differences between the two schemes be? In particular, will the implicit scheme mimic some general aspects of quantum mechanics by effectively damping the high frequency modes?

(4) What will the computational differences between the two schemes be?

The water droplet model as formulated, along with the procedure for choosing an adequate subcluster from which to calculate interatomic distances, has proven successful at resolving well the first two peaks of the radial distribution functions. Surface effects begin to enter only after a 6 Å radius for the 125-molecular cluster. A larger cluster will undoubtedly increase the range of "bulk"-property resolution, and incorporation of periodic boundary conditions should further improve results.

Numerical stability and computational feasibility of the implicit-Euler/Langevin scheme has been demonstrated for time steps of 5 and 10 fs. This behavior is vastly different from that of explicit schemes, for which a time step greater than 1 fs generally produces numerical instability. This suggests that the method has potential for biomolecular applica-

tions where the key motions of interest are typically slower and less frequent than the high-frequency vibrational modes. (The latter modes dictate very fine time resolutions in numerical schemes.) Of course, the effects of high-frequency damping will have to be monitored closely, since in some cases the high-frequency modes may facilitate collective motions of the system. Thus, application of the implicit-Euler/Langevin scheme to biomolecules should be targeted to answer specific conformational questions, for which it may be more suitable than explicit schemes. Finer resolutions of the motion can then be explored with small time-step simulations and no (or little) damping.

In comparing thermodynamic and structural properties obtained from the implicit and explicit schemes, we observed the expected damping of the high-frequency vibrational modes by the implicit scheme. Average energies are much smaller for the implicit scheme, and larger percentages of energy differences occur in the kinetic, rather than potential, energy components. The water structures simulated by this damping tend to be more compact, involving more rigid hydrogen-bonding networks. There is less competition between the local, more ordered structures of tetrahedrally coordinated hydrogen bonds and the various less-ordered subclusters of hydrogen-bonded networks. This is exhibited by sharper and higher first-neighbor peaks in the radial distribution functions and smaller occupancy of the region between the first and second peaks. Since the SPC parameters were optimized—by an explicit molecular dynamics simulation—to reproduce accurately experimentally derived radial-distribution-function curves, the resulting effective pair potential already incorporates quantum-mechanical effects. Indeed, we obtained very good agreement with these curves by the explicit integration scheme and by the MC. While results from the implicit runs produced the differences noted above, they provided better estimates for the heat capacities  $C_p$  than the explicit scheme. This suggests that relative energy differences may be better computed by the implicit formulation, where the thermally accessible energy per mode is not equally partitioned, as in classical statistical mechanics, but rather partitioned approximately in a quantum-like fashion.

Finally, in regard to computational performance, we mention that our main focus in this work was to investigate differences in thermodynamic and structural properties resulting from the explicit/Langevin and implicit/Langevin formulations; thus, no particular effort was devoted to improving the competitiveness, in terms of computer time, of the implicit scheme. Recall that work per iteration is greater in our implicit scheme because minimization of  $\Phi$  is required at every step (see Sec. II). However, only a few iterations were required at every step because we have a good initial guess for  $\mathbf{x}^{n+1}$ . For  $\Delta t = 5$  fs, two or three iterations of minimization were typically required in every dynamic iteration, while  $\Delta t = 10$  fs resulted in about eight minimization iterations per step. For the same total simulation time of 20 ps, 20 000 iterations were required with the explicit scheme, while 4 000 and 2 000 were required for the two implicit schemes, respectively. Overall, this resulted in a factor of 5 of additional computation time for the implicit scheme. We

expect to make this factor far smaller with future improvements and variations in our minimization algorithm.

Nonetheless, our results demonstrate that it is not necessary to run the implicit simulation for as long as the explicit one; the various thermodynamic and geometric averages settle down much more rapidly in the implicit than in the explicit run. In particular, we obtain the same radial distribution functions in the implicit scheme after 5 ps, 10 ps, and so on, whereas the curves from the explicit scheme produce greater variations for such cumulative subintervals. This finding supports our original motivation for searching for a numerical formulation that permits larger time steps and damps out high-frequency modes: certain “bulk” properties of the system can be observed more rapidly, without a fine resolution of all motions involved.

These physical and computational issues will be investigated further through applications to other physical systems for which experimental data are available. One such application, simple structural transitions and reproduction of transition rates in butane, is currently underway.

## ACKNOWLEDGMENTS

We thank the following individuals for their interest and contributions: David Beveridge, Jan Hermans, Shneior Lifson, Jerry Percus, Charles Peskin, and Paula Whitlock. Generous support for the research was provided by the National Science Foundation, the National Institutes of Health, the New York State Science and Technology Foundation, the American Association for University Women Educational Foundation, the San Diego Supercomputer Center, the DOE Supercomputer Facilities at Livermore, the Academic Computing Facility at Courant Institute, NYU, the University Computing Center of CUNY, and through a CUNY/PSC grant.

<sup>1</sup>W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).

<sup>2</sup>D. L. Beveridge, M. Mezei, P. K. Mehrotra, F. T. Marchese, G. R.-Shanker, T. Vasu, and S. Swaminathan, in *ACS Advances in Chemistry, No. 204, Molecular-Based Study of Fluids*, edited by J. M. Haile and G. A. Mansoori (American Chemical Society, Washington, DC, 1983), pp. 297–351.

<sup>3</sup>J. Hermans, A. Pathiaseril, and A. Anderson, *J. Am. Chem. Soc.* **110**, 5982 (1988).

<sup>4</sup>H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, in *Intermolecular Forces*, edited by B. Pullman (Reidel, Dordrecht, The Netherlands, 1981), pp. 331–342.

<sup>5</sup>B. M. Pettitt and M. Karplus, *Chem. Phys. Lett.* **136**, 383 (1987).

<sup>6</sup>M. Mezei, P. K. Mehrotra, and D. L. Beveridge, *J. Am. Chem. Soc.* **107**, 2239 (1985).

<sup>7</sup>E. Clementi, *J. Phys. Chem.* **89**, 4426 (1985).

<sup>8</sup>D. L. Beveridge, P. Subramanian, B. Jayaram, S. Swaminathan, and G. Ravishanker, in *Structure and Methods, Volume III*, edited by R. H. Sarma and M. H. Sarma (Adenine, New York, 1990), pp. 79–111.

<sup>9</sup>V. Madison, D. J. Osguthorpe, P. Dauber, and A. Hagler, *Biopolymers* **22**, 27 (1983).

<sup>10</sup>D. L. Beveridge, G. Ravishanker, M. Mezei, and B. Gedulin, in *Biomolecular Stereodynamics III*, edited by R. H. Sarma and M. H. Sarma (Adenine, New York, 1986), pp. 237–252.

<sup>11</sup>F. Franks (Ed.), *Water: A Comprehensive Treatise* (Plenum, New York, 1972–1982), Vols. 1–7.

- <sup>12</sup>D. Eisenberg and W. Kauzmann, *The Structure and Properties of Water* (Oxford U.P., New York, 1969).
- <sup>13</sup>B. Kamb, in *Structural Chemistry and Molecular Biology*, edited by A. Rich and N. Davison (Freeman, San Francisco, 1968), pp. 507-542.
- <sup>14</sup>J. D. Bernal and R. H. Fowler, *J. Chem. Phys.* **1**, 515 (1933).
- <sup>15</sup>A. Geiger, F. H. Stillinger, and A. Rahman, *J. Chem. Phys.* **70**, 4185 (1979).
- <sup>16</sup>F. H. Stillinger, *Science* **209**, 451 (1980).
- <sup>17</sup>R. J. Speedy and M. Mezei, *J. Phys. Chem.* **89**, 171 (1985).
- <sup>18</sup>C. S. Peskin and T. Schlick, *Comm. Pure Appl. Math.* **42**, 1001 (1989).
- <sup>19</sup>T. Schlick and C. S. Peskin, *Comm. Pure Appl. Math.* **42**, 1141 (1989).
- <sup>20</sup>T. Schlick, in *Proceedings of the Molecular Mechanics and Molecular Dynamics Workshop*, Florida State University, 26-30 March 1990 (in press).
- <sup>21</sup>T. Schlick, B. E. Hingerty, C. S. Peskin, M. L. Overton, and S. Broyde, in *Theoretical Biochemistry and Molecules Biophysics, Volume I: Nucleic Acids*, edited by D. L. Beveridge and R. Lavery (Adenine, New York, in press), pp. 39-58.
- <sup>22</sup>G. Dahlquist and Å. Björk, *Numerical Methods*, (Prentice-Hall, Englewood Cliffs, NJ, 1974).
- <sup>23</sup>J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge U.P., Cambridge, 1987).
- <sup>24</sup>W. F. van Gunsteren, H. J. C. Berendsen, and J. A. C. Rullmann, *Mol. Phys.* **44**, 69 (1981).
- <sup>25</sup>T. Schlick and A. Fogelson, *ACM Trans. Math. Software* (in press).
- <sup>26</sup>T. Schlick and M. Overton, *J. Comp. Chem.* **8**, 1025 (1987).
- <sup>27</sup>A. H. Narten, M. D. Danford, and H. A. Levy, *J. Chem. Soc. Discuss. Faraday* **43**, 97 (1967).
- <sup>28</sup>A. H. Narten and H. A. Levy, *J. Chem. Phys.* **55**, 2263 (1971).
- <sup>29</sup>A. H. Narten, W. E. Thiessen, and L. Blum, *Science* **217**, 1033 (1982).
- <sup>30</sup>A. K. Soper and R. N. Silver, *Phys. Rev. Lett.* **49**, 471 (1982).
- <sup>31</sup>A. K. Soper and M. G. Phillips, *Chem. Phys.* **107**, 47 (1986).
- <sup>32</sup>T. Schlick, C. Peskin, S. Broyde, and M. Overton, *J. Comp. Chem.* **8**, 1199 (1987).
- <sup>33</sup>T. Schlick, *J. Comp. Chem.* **10**, 951 (1989).
- <sup>34</sup>J. Hermans (private communication).
- <sup>35</sup>J. Carrier, L. Greengard, and V. Rokhlin, *SIAM J. Sci. Stat. Comp.* **9**, 669 (1987).
- <sup>36</sup>N. A. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, H. A. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- <sup>37</sup>M. Rao, S. Pangali, and B. J. Berne, *Mol. Phys.* **37**, 1779 (1979).
- <sup>38</sup>R. B. Blackman and J. W. Tuckey, *The Measurement of Power Spectra* (Dover, New York, 1958).
- <sup>39</sup>W. W. Wood, in *Physics of Simple Liquids*, edited by H. N. V. Temperley, J. S. Rowlinson, and G. S. Rushbrooke, (North-Holland, Amsterdam, 1968).