

ger, D. L.; Mintz, E. *Organometallics* 1984, 3, 1756. (d)  
Reger, D. L.; Coleman, C. J. *Inorg. Chem.* 1979, 18, 3155.  
8. Womack, E. B.; McWhirter, J. In *Organic Syntheses*; Hor-

ning, E. C., Ed.; Wiley: New York, 1955; Coll. Vol. III,  
pp 714-715.

## Molecular Dynamics Simulation on thermodynamic and Structural Properties of Liquid Hydrocarbons : Normal Alkanes

Wonpil Im and Youngdo Won

*Department of Chemistry, Hanyang University, Seoul 133-792, Korea*

*Received June 23, 1994*

A series of aliphatic hydrocarbons, methane to hexane in the liquid state, are modeled with the molecular mechanical potential parameters treating all hydrogen degrees of freedom explicitly. Thermodynamic properties (heat capacities and heats of vaporization) are calculated from relatively short (20ps) molecular dynamics trajectories. The liquid state structures are also examined through various radial distribution functions. Molecular dynamics simulations reproduce experimentally measured properties within a few percent errors, thus indicate that the present set of all-hydrogen parameters is suitable for simulating macromolecular systems in bulk.

### Introduction

A detailed model for complex organic and biochemical systems includes not only reactants but also surrounding molecules acting as solvent. A complete description of such a system as bulk matter would contain sub-Avogadro's number of atoms,  $\sim 10^{22}$ . For each atom, we have to deal with 6 phase space coordinates (position and velocity) while carrying out molecular dynamics simulations.

One way to approach to the problem of  $10^{22}$  degrees of freedom is to utilize the model system with a stochastic boundary. Only a part of the universe within the boundary is simulated with the normal molecular dynamics and the rest is regarded as the stochastic thermal bath to the system.<sup>1</sup> Another way is to set up a periodic boundary surrounding the system of interests.<sup>2</sup> We need to include 1,000 to 10,000 atoms within the boundary depending on the problem at hand. The problem of a bulk system can be safely reduced to  $10^4$  degrees of freedom in size.

Commonly used theoretical methodologies in molecular simulations are Monte Carlo (MC) and molecular dynamics (MD). Both methods are based on a molecular mechanical energy function designed to handle thousands of degrees of freedom efficiently. The Born-Oppenheimer potential surface is parameterized with a standard form of empirical potential function. Through molecular simulations, we generate or sample phase space trajectories over the potential energy surface.

With limited computing resources, we have been using the united atom representation of molecules with polar hydrogen parameters. In the polar hydrogen model, nonpolar hydrogen is merged into the heavy atom that binds the hydrogen. For example, the methyl group is replaced with one atom that represents the force field of  $\text{CH}_3$  moiety. The united atom representation has certain advantages in reducing

computational demands with sacrificing less important degrees of freedom such as CH stretching or HCH bending motions. The previous MC simulation showed that a variety of thermodynamic and conformational properties of molecular systems could be calculated with a few percent errors using the united atom model.<sup>3</sup>

With much increased computing resources enough to fulfill the needs of extended molecular simulations, it is desired to include all degrees of freedom explicitly. All-hydrogen parameters are under development. Harvard group have made a preliminary release of all-hydrogen parameter sets for protein, nucleic acid and lipid simulations and is about to complete such the task.<sup>4</sup>

As a part of all-hydrogen parameter development efforts, it is appropriate and necessary to perform molecular simulations with the subset of parameters for model compounds. In this work, we have carried out molecular dynamics simulations on a series of saturated hydrocarbons. Thermodynamic and structural properties of the liquid state of methane to hexane series are evaluated from the dynamics trajectories. Comparison to known experimental results provides an indicator for goodness of the all-hydrogen parameter set.

Hydrocarbon is the backbone in most organic molecular structure and therefore the present work is considered to be the first step forward to developing molecular models for organic solvents.

### Theory and Computational Methods

**Molecular Model.** The standard all-hydrogen potential function of CHARMM (Chemistry at HARvard Macromolecular Mechanics)<sup>5</sup> is employed to build molecular models for the series of hydrocarbons. The empirical energy function is partitioned as

**Table 1.** Potential Energy Functions

Energy term	Function	Parameters
$E_b$	$\Sigma k_b(b_i - b_{i0})^2$	$k_b, b_{i0}$
$E_\theta$	$\Sigma k_\theta(\theta_i - \theta_{i0})^2$	$k_\theta, \theta_{i0}$
$E_{UB}$	$\Sigma k_{UB}(s_i - s_{i0})^2 + \Sigma k_{UB}(s_i - s_{i0})$	$k_{UB}, s_{i0}$
$E_\phi$	$\Sigma k_\phi[1 + \cos(n_i\phi_i - \delta_i)]$	$k_\phi, n_i, \delta_i$
$E_{vdW}$	$\Sigma 4\epsilon_{ij}[(\sigma_{ij}/r_{ij})^{12} - (\sigma_{ij}/r_{ij})^6]$	$\epsilon_{ij}, \sigma_{ij}$
$E_{elec}$	$\Sigma(q_i q_j / r_{ij})$	$q_i$

**Table 2.** Atom Types and Non-Bonding energy Parameters

Atom Type	$\epsilon$ (kcal/mol)	$\sigma$ (Å)	Charge
CT4 carbon in methane	-0.080	4.120	-0.36
CT3 carbon in methyl group	-0.080	4.120	-0.27
CT2 carbon in methylene group	-0.055	4.350	-0.18
HA aliphatic hydrogen	-0.022	1.320	0.09

$$U = E_b + E_\theta + E_{UB} + E_\phi + E_{elec} + E_{vdW} \quad (1)$$

The energy terms and parameters are summarized in Table 1.

The series of normal alkane molecules are built with four atom types listed in Table 2; the aliphatic hydrogen (HA), the methylene carbon (CT2), the methyl carbon (CT3) and the carbon in methane (CT4). All bonding energy parameters are adopted from CHARMM version 22 parameter set for proteins; top\_all22\_prot.inp and par\_all22\_prot.inp. The bond and angle energy terms take the harmonic form and the Urey-Bradley term is for 1-3 interactions and has both quadratic and linear contributions. The improper dihedral energy is another harmonic term usually needed to keep four atoms in a plane, however, no improper term is needed for the present work. The dihedral energy is represented as a periodic cosine term, where  $k_\phi$  is the dihedral force constant,  $n$  is the periodicity and  $\delta$  is the phase shift. Parameters for bonding interactions are collected in Table 3.

Non-bonded interactions are considered for 1-4 or further separated atoms in a molecule and for atoms belonging to different molecules with each other. In the all22 parameter set, an aliphatic hydrogen has the atomic charge of 0.09.

Methylene and methyl groups are considered as a neutral unit in a molecule, therefore, carbons take the appropriate charge to keep the neutrality of the group. We adopt the all22 parameters for all the molecules except methane for which appropriate values are not found in the set. The atomic charge of HA is set to 0.09, then that for CT4 is -0.36. For the van der Waals energy term of methane, we take -0.080 for  $\epsilon$  and 4.120 for  $\sigma$ . These values are obtained by considering the trend of CT1 to CT3 values of the all22 parameter set. We might have performed the parameterization protocol based on fitting *ab initio* interaction energies, which would generate more accurate parameters. However, the scaled values are appropriate for the present work. The appropriateness is to be evaluated by comparing simulation results to experimental data.

**Model System Construction.** Our molecular system is built within a cubic periodic boundary and is consisted with the primary cube surrounded by 26 images of its own. The IMAGE facility of CHARMM is utilized in setting up the cubic images and the CUTIM value is set to the same as CUTNB value (13 Å). Nonbonding interactions are considered only for those atoms separated within the CUTNB distance. CUTIM is the cutoff distance for primary atom to image atom nonbonding interactions. Such construction warrants a realistic representation of homogeneous liquids in bulk. When a molecule moves into or out of the primary cube, the corresponding image molecule moves out of or into the primary box so that the whole system may maintain the structural integrity as liquid and density of the box unchanged.

We put 216 molecules in the cubic box of proper size determined by the density at the simulation temperature. The initial coordinates are generated in such a way that the 216 hydrocarbon molecules have the all-trans conformation. However, we find that the proper gauche/trans ratio is maintained during dynamics simulations. The box size and number of atoms for each system are listed in Table 4. Due to the periodic boundary in use, the atoms in the primary box and those image atoms within CUTIM (13 Å) from the primary atoms should be considered. The actual number of atoms considered in the energy calculation is about three times of that of the primary atoms.

**Molecular Dynamics Simulations.** A "normal" molecular dynamics simulation procedure is applied to each mo-

**Table 3.** parameters for bonding Interactions

Bond	$k_b$	$b_0$	Angle, and Urey-Bradley	$k_\theta$	$\theta_0$	$k_{UB}$	$s_0$	Dihedral	$k_\phi$	$n$	$\delta$
CT4-HA	322.0	1.111	CT3-CT2-CT3	53.35	114.0	8.00	2.561	CT3-CT2-CT2-CT3	0.15	1	0.00
CT3-HA	322.0	1.111	CT3-CT2-CT2	58.00	115.0	8.00	2.561	CT3-CT2-CT2-CT2	0.15	1	0.00
CT2-HA	309.0	1.111	CT2-CT2-CT2	58.35	113.6	11.16	2.561	CT2-CT2-CT2-CT2	0.15	1	0.00
CT3-CT3	222.5	1.530	HA-CT3-CT3	37.5	110.1	22.53	2.179	X-CT3-CT3-X	0.155	3	0.00
CT3-CT2	222.5	1.528	HA-CT3-CT2	34.6	110.1	22.53	2.179	X-CT3-CT2-X	0.160	3	0.00
CT2-CT2	222.5	1.530	HA-CT2-CT3	24.6	110.1	22.53	2.179	X-CT2-CT2-X	0.195	3	0.00
			HA-CT2-CT2	26.5	110.1	22.53	2.179				
			HA-CT4-HA	35.5	108.4	5.40	2.802				
			HA-CT3-HA	35.5	108.4	5.40	2.802				
			HA-CT2-HA	35.5	109.0	5.40	2.802				

Table 4. System Construction

System	$T_i$ (°C)	$T_r$ (°C)	Density (g/cm <sup>3</sup> )	Box Sizes (Å)	# of Primary Molecules	# of Primary Atoms	Total # of Atoms <sup>d</sup>
methane	-161.49	-161.49	0.424	23.88	216	1,080	2,700
ethane	-88.63	-88.63	0.546	27.03	216	1,728	4,300
propane	-42.07	-42.07	0.581	30.084	216	2,376	5,900
butane	-0.50	-0.50	0.602	34.842	216	3,024	7,500
pentane	36.07	25.00	0.621	34.668	216	3,672	9,000
hexane	68.74	25.00	0.655	36.138	216	4,320	10,000

<sup>a</sup>simulation temperature. <sup>b</sup>approximate number of primary and image atoms.

molecular system under the cubic boundary condition. First, the energy of the initial structure as a whole is minimized in order to remove any "hot" spots (close contacts within the van der Waals radius) generated during the coordinate building processes. The steepest descent (SD) and adopted basis Newton-Raphson (ABNR) method are used for minimization.

MD simulation is carried out at the boiling point for methane, ethane, propane and butane. For other systems, the simulation is carried out at the room temperature (25 °C). A minimized NVE ensemble is heated to the simulation temperature. Temperature is increased by 2K for every 50 steps of dynamics integration (each time integration step corresponds to one femtosecond) until it reaches the simulation temperature. Atomic velocities are assigned based on the Gaussian random deviates corresponding to the temperature. After the target temperature is established, 10ps of equilibration dynamics is performed. During the equilibration run, the temperature is monitored at every 50th step and the atomic velocities are reassigned if necessary. Verlet integrator is used for heating and equilibration dynamics with CH bonds "SHAKE".d. SHAKE is the way of time integration that keeps the "SHAKE".d bond lengths fixed during each integration step.

20 ps "production" runs for three different ensembles (NPT, NVT and NVE) are carried out with the equilibrated system. CPT (Constant Pressure and Temperature) dynamics program is used for simulation of NPT and NVT ensembles without using SHAKE. The Verlet integration method is used for NVE ensembles with SHAKE. Along each dynamics trajectory of NVT and NPT ensembles, the fluctuation in total energy is monitored and used for evaluating the heat capacities. The position and velocity trajectories are saved for further analyses every 25fs (25 steps) during the NVE dynamics.

**Analyses.** We have utilized various CHARMM analysis facilities and standard procedures for analyzing molecular dynamics trajectories. However, the procedure of evaluating the heat of vaporization needs to be described here.<sup>6</sup>

The energy of vaporization for the liquid going to the ideal gas,  $\Delta E_v^0$  is defined by

$$\Delta E_v^0 = E(g) - \frac{E(l)}{N} \quad (2)$$

where  $N$  is number of molecules in the system. If we assume that the kinetic energies of the ideal gas and liquid are the same,

$$\Delta E_v^0 = E^{intra}(g) - \frac{E^{intra}(l) + E^{inter}(l)}{N} \quad (3)$$

where  $E^{intra}(g)$  is the intramolecular energy of the ideal gas (the total potential energy of an isolated molecule),  $E^{intra}(l)$  is the total intramolecular energy of the liquid and  $E^{inter}(l)$  is the intermolecular energy of the liquid. The intramolecular energy of an ideal gas molecule is obtained by running MD simulation of a molecule in vacuum at the simulation temperature.  $\Delta E_v^0$  is related to  $\Delta H_v^0$  as

$$\Delta H_v^0 = \Delta E_v^0 + P(V^0(g) - V(l)) \approx \Delta E_v^0 + RT \quad (4)$$

where the work term is approximated with  $RT$ . For the real gas, we need to subtract the nonideality correction (the enthalpy departure function)  $H^0 - H$  from  $\Delta H_v^0$

$$\Delta H_v = \Delta H_v^0 - (H^0 - H) \quad (5)$$

The enthalpy departure function can be calculated using the second virial coefficient for the gas according to

$$H^0 - H = \frac{RT}{V} \left( T \frac{dB}{dT} - B \right) \quad (6)$$

where  $B(T)$  is estimated from the Berthelot second virial coefficient,

$$B(T) = -\frac{9RT_c}{128P_c} \left( -\frac{18T_c^2}{T^2} - 1 \right) \quad (7)$$

The correction is small since the gases are relatively ideal under the simulation conditions. For example, the value for methane is 0.0337 kcal/mol.

## Results and Discussions

**Thermodynamics.** The heat capacities,  $C_p$  and  $C_v$ , are calculated from the total energy fluctuations in NPT and NVT ensembles respectively.<sup>7</sup> These quantities are well converged during the production dynamics of 20ps.

$$C_p = \left( \frac{\partial \langle H \rangle}{\partial T} \right)_p = \frac{\langle H^2 \rangle - \langle H \rangle^2}{NkT^2} \quad (8)$$

$$C_v = \left( \frac{\partial \langle E \rangle}{\partial T} \right)_v = \frac{\langle E^2 \rangle - \langle E \rangle^2}{NkT^2} \quad (9)$$

The results are listed in Table 5 and compared to the reported values in the literature.<sup>8-9</sup> The calculation out of energy fluctuations reproduces the experimental measurement very well. We obtain superior results to the previous

**Table 5.** Heat Capacities (in cal/mol K)

	$T_s$ (°C)	$C_p$ (calc.)	$C_p$ (expl.)	$C_v$ (calc.)	$C_v$ (expl.)
methane	-161.49	15.3	13.2	4.2	—
ethane	-88.63	16.8	17.6	5.5	—
propane	-42.07	19.5	23.5	9.5	—
butane	-0.50	29.9	31.8	25.6	23.1
pentane	25.00	35.2	39.9	27.7	—
hexane	25.00	38.8	46.8	36.0	—

**Table 6.** heats of vaporization (in kcal/mol)

	$T_s$ (°C)	NVE	NVT	NPT	Experimental
methane	-161.49	1.82	1.98	1.85	1.96
ethane	-88.63	3.43	3.20	3.20	3.52
propane	-42.07	4.52	4.62	4.35	4.49
butane	-0.59	4.55	4.87	5.19	5.04
pentane	25.00	6.27	6.64	6.32	6.32
hexane	25.00	7.53	7.81	7.48	7.54

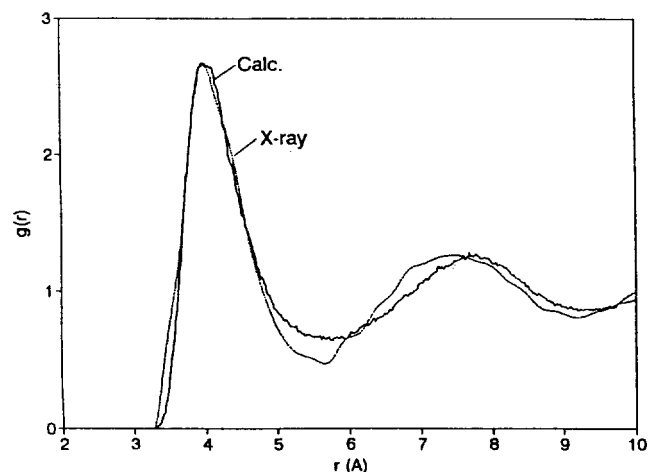
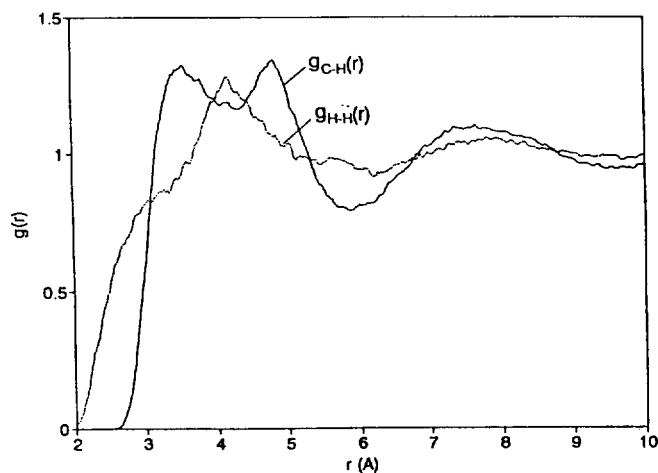
MC work for small hydrocarbons.<sup>3</sup> The MC simulation sampled 500,000 to 1,500,000 configurations of 128 united atom molecules, where methane was represented by a monatomic and ethane by a diatomic molecule. MD seems to be the better sampling strategy over well equilibrated representative ensembles.

The heat of vaporization is determined as described in the previous section over NVE, NVT and NPT ensembles. While MD on NVE ensemble is carried out by the Verlet integrator with SHAKE, CPT dynamics of CHARMM is performed on NVT and NPT ensembles. The CPT dynamics is based on the Leap-Frog (LF) integrator. The temperature is deviated from the target temperature during the CPT dynamics when SHAKE is in effect. For some reason, the LF integrator is not able to maintain the constant temperature with using SHAKE. Therefore SHAKE has to be turned off during the production run under CPT dynamics. A separate dynamics on the NVE ensemble without using SHAKE indicates that the present integration step size of 1fs produces stable trajectories even without using SHAKE.

The heats of vaporization obtained from these ensembles are listed in Table 6. All are in excellent agreement to the experimental data.<sup>8-9</sup> In the present work, the largest deviation from experiments is 0.3 kcal/mol. The previous MC simulation well reproduced the experimental results too.<sup>3</sup> Because the heat of vaporization is obtained as the energy difference, one might expect that both methods, MD and MC sampling, produce the heat of vaporization in similar quality.

**Liquid Structure.** Radial distribution functions (rdf's) represent deviations in the distribution of atoms in a liquid from that expected from the bulk density. The radial distribution function  $g(r)$  gives the local density  $\rho(r) = \rho g(r)$  about a fixed atom, where  $\rho$  is the bulk density.

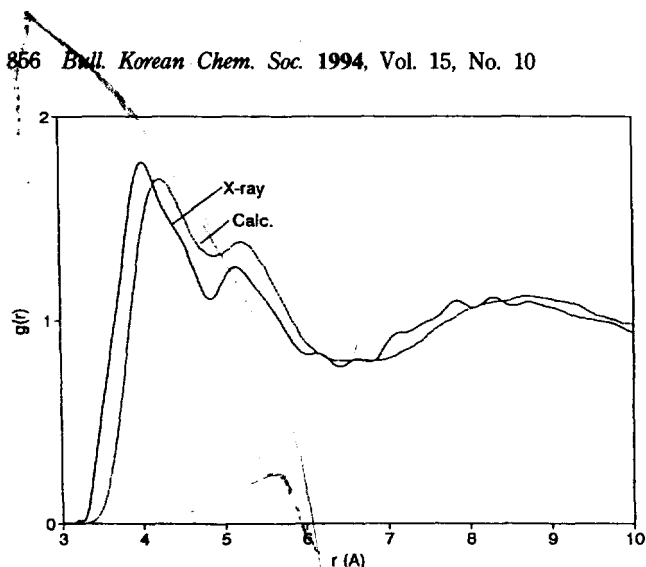
A variety of rdf's for each system are determined from the MD trajectories over the NVE ensemble. The CC rdf for liquid methane is compared with that determined by the X-ray crystallography<sup>10</sup> in Figure 1. The CH and HH rdf's

**Figure 1.** The C-C radial distribution function of methane calculated from the NVE ensemble trajectory compared with that determined by the X-ray scattering experiment.**Figure 2.** The C-H and H-H radial distribution functions of methane calculated from the NVE ensemble trajectory.

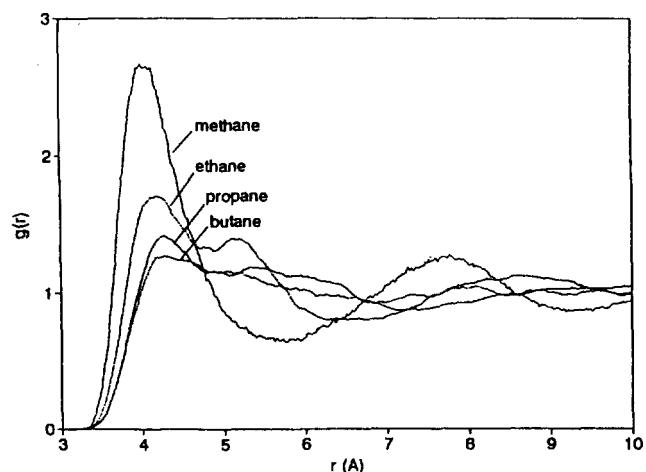
are also determined from the trajectory as shown in Figure 2. The structural features of methane molecule are well represented in those rdf's. The first two peaks of CH rdf are located at each side of the first peak of CC rdf. The HH rdf shows a rather dispersed feature over 2 Å.

Figure 3 is the comparison of the CC rdf's of ethane, one determined from MD simulation and the other is from X-ray data.<sup>11</sup> Both peaks are in good agreement with the experiment; less than 0.1 Å in position. Note that the second peak is distinctive and well reproduced. The previous MC simulation with the united atom model was not able to reproduce the feature.<sup>3</sup> The superior result is obviously due to more accurate model with explicit hydrogens.

The effect of increasing chain length on CH<sub>3</sub>-CH<sub>3</sub> rdf's is illustrated in Figure 4. The shielding by the rest of molecule containing the reference methyl group, causes the height of the first peak to be decreased with increasing chain length. The shielding effect, however, is weakened over the size of butane. The second peak of the ethane rdf indicates the other methyl group of the neighboring molecule. As



**Figure 3.** The C-C radial distribution function of ethane calculated from the NVE ensemble trajectory compared with that determined by the X-ray scattering experiment.



**Figure 4.** The C-C radial distribution functions of liquid methane, ethane, propane and butane calculated from the NVE ensemble trajectories.

shown in the figure, the rdf's are becoming featureless with increasing chain length. The short range order distinctive in methane and ethane is disrupted in larger hydrocarbons because of the rapid departure from spherical symmetry.

### Conclusion

The previous MC works on united atom liquid hydrocar-

bon systems<sup>3</sup> are re-examined by MD simulations with all-hydrogen model. For equilibrated systems, MD simulation results converge rapidly than those of similar MC simulations. Heats of vaporization are calculated with errors less than 0.3 kcal/mol for all the ensembles simulated. All hydrogen representation yields better results in heat capacities in general than the united atom model. The explicit hydrogen degrees of freedom are particularly important in methane and ethane systems. Structural features of experimental rdf's of methane and ethane are well reproduced, which indicates assurance of nonbonding interaction parameters used in this work. The aliphatic part of the pre-release version of all22 parameter set is confirmed to be suitable for simulating organic and biochemical systems.

**Acknowledgment.** This paper was supported by NON DIRECTED RESEARCH FUND, Korea Research Foundation, 1993.

### References

1. Brooks, C. L. III; Karplus, M.; Pettitt, B. M. *Proteins-A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*; Wiley: New York, 1988; p. 38.
2. Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, 1987; p. 24.
3. Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 6638.
4. CHARMM Developer's Meeting; CHARMM c22 Release Note 1993.
5. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comp. Chem.* **1983**, *4*, 187.
6. Jorgensen, W. L.; Binning, R. C. Jr.; Bigot, B. *J. Am. Chem. Soc.* **1981**, *103*, 4393.
7. Jorgensen, W. L. *J. Am. Chem. Soc.* **1981**, *103*, 4721.
8. American Petroleum Institute Research Project 44 *Selected Values of Physical and Thermodynamic Properties of Hydrocarbons and Related Compounds*; Carnegie Press: Pittsburgh, 1953.
9. ASTM Technical Publication No. 109A *Physical Constants of Hydrocarbon*; American Society for Testing and Materials: Philadelphia, 1963.
10. Habenschuss, A.; Johnson, E.; Narten, A. H. *J. Chem. Phys.* **1981**, *74*, 5234.
11. Sandler, S. I.; Lombardo, M. G.; Wong, D. S. -H; Habenschuss, A.; Narten, A. H. *J. Chem. Phys.* **1982**, *77*, 2144.