Combined *Ab Initio*/Empirical Approach for Optimization of Lennard–Jones Parameters

DAXU YIN, ALEXANDER D. MACKERELL, JR.

Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, 20 North Pine Street, Baltimore, Maryland 21201

Received 15 April 1996; accepted 2 September 1997

ABSTRACT: Obtaining accurate Lennard–Jones (L-J) parameters is a vital part of the optimization of empirical force fields due to their significant contribution to condensed-phase properties. We present a novel approach to optimize L-J parameters via the use of *ab initio* data on interactions between rare gas atoms and model compounds combined with the reproduction of experimental pure solvent properties. Relative values of *ab initio* minimum interaction energies and geometries between helium or neon and model compounds were used to optimize the relative magnitude of the L-J parameters. Absolute values of the L-J parameters were determined by reproducing experimental heats of vaporization and molecular volumes for pure solvents. Application of the approach was performed on methane, ethane, and propane. Free energies of aqueous solvation and butane pure solvent and aqueous solvation calculations were used to test the developed L-J parameters. The new alkane parameters are similar or improved as compared with current empirical force field parameters with respect to experimental pure solvent properties and free energies of aqueous solvation. Also included is a description of the internal portion of the force field. © 1998 John Wiley & Sons, Inc. J Comput Chem 19: 334–348, 1998

Keywords: parameters; parametrization; neon; helium; CHARMM

Correspondence to: A. D. MacKerell Jr.

Contract/grant sponsor: Mildred Mindell Cancer Foundation

Contract/grant sponsor: National Institutes of Health; contract/grant number: GM-51501

This article includes Supplementary Material available from the authors upon request or via the Internet at ftp.wiley.com/public/journals/jcc/suppmat/19/334 or http://journals.wiley.com/jcc/

CCC 0192-8651/98/030334-15

Introduction

Empirical force fields have been used to study condensed-phase properties due to the simplicity of potential energy functions as required for calculations on systems that include 10,000 or more particles. In empirical force fields, dispersion–attraction interactions between atoms are often treated by the r^6 term in the Lennard Jones (L-J) portion of the potential energy function.[1] The L-J term has been shown to dominate the heats of vaporization of nonpolar organic molecules[2,3] and make significant contributions to the pure solvent properties of polar neutral organic compounds[4] and to the crystal structures and energies of nucleic acid base analogs.[5–8] Thus, obtaining accurate L-J parameters is a vital part in the optimization of empirical force fields.

L-J parameters for empirical force fields are often developed by reproducing model compound pure solvent properties, such as heats of vaporization and molecular volumes.[3,4,9–11] Crystal data have also been used to develop L-J parameters.[5–8] Previous parametrization work on the CHARMM force field has shown that different sets of L-J parameters for ethane satisfactorily reproduce the experimental heat of vaporization and molecular volume.[12] This illustrates the parameter correlation problem,[13] where significantly different parameters can yield virtually identical condensed-phase properties. This problem indicates that additional information beyond experimental macroscopic observables may be needed as goal data during L-J parameter optimization procedures.

Ab initio interaction geometries and energies have been considered as the goal data to develop L-J parameters for empirical force fields.[10,11,14,15] In some cases, a water molecule is used to probe hydrogen bonding sites, allowing both partial atomic charges and L-J parameters to be optimized via reproduction of the *ab initio* minimum interaction energies and distances. A water molecule has been used as a probe to determine the L-J parameters of aliphatic hydrogens in substituted aliphatic molecules based on *ab initio* interaction geometries and energies.[10,11] Due to the interaction between water and model compounds containing a significant electrostatic contribution[4,16] interpretation of the calculated minimum interaction energy and geometry may be limited. A force field for methane based only on methane-dimer *ab initio* results at the MP3/6-limited. A force field for methane based only on methane-dimer *ab initio* results at the MP3/6-311 + + G(3d,3p) level of theory with corrections for basis-set superposition error (BSSE) was developed by Tsuzuki et al.[14] Based on two sets of parameters developed using this approach, the experimental pure solvent properties were not well reproduced; that is, heats of vaporization were approximately 15% too low, densities were approximately 10% too low, and self-diffusion coefficients were about 10–20% too large. A recent study has used *ab initio*–determined interactions of neon and argon with methane to determine a force field for methane.[15] Application of one of the three derived sets of methane parameters for pure solvent calculations yielded a heat of vaporization and density within 5% and −2% of experiment. However, the self-diffusion coefficient was 48% too low. These results suggest that developing L-J parameters based solely on *ab initio* data is not sufficient at currently accessible levels of theory. To overcome these difficulties, an alternative approach to optimize L-J parameters is proposed in the present study: rare gas atoms will be used as probes of the van der Waals (vdW) surface of model compounds via *ab initio* calculations. The *ab initio* data will then be used to determine the relative values of L-J parameters, whereas the absolute values of the L-J parameters will be determined by reproducing experimental condensed-phase properties.

He and Ne atoms have previously been used in *ab initio* studies to probe the vdW surface of water, NH_3, and CH_4.[15,17–21] These calculations used extensive basis sets and included treatment of electron correlation up to fourth-order Møller–Plesset perturbation theory. Such levels of theory, however, are computationally inaccessible when applied to larger molecules. *Ab initio* calculations at lower levels of theory in the absence of electron correlation, however, cannot accurately treat dispersion interactions[22–24] as required to obtain reasonable estimates of the values of the L-J parameters. To apply He or Ne atoms to study larger systems, certain compromises between the accuracy of the calculations and the accessibility of computational power must be made in selection of a proper level of theory. Previous work in our laboratory has systematically investigated the influence of different levels of theories on helium and neon dimers by comparing the calculated results with highly accurate empirical data.[25] *Ab initio* calculations were also performed on methane

dimers and interactions between helium or neon and methane to test the selected level of theory. Based on these computations, the MP3/6-311 + + G(3d,3p) level of theory was selected for use in the calculation of the interactions between rare gas atoms and model compounds.

The present study proposes and tests a novel methodology for the optimization of L-J parameters. This methodology focuses on overcoming limitations in previous parametrization efforts by the inclusion of both experimental and *ab initio* information as goal data. Alkanes were selected for validation of the methodology due to the extensive work previously performed on this class of compounds.[2,26–35] Computational approaches and theory used in the current study are presented in the ''Methods'' section. *Ab initio* and empirical computations for methane, ethane, and propane, including data for a number of parameter sets tested, are presented in the ''Results.'' Also included in the ''Results'' are condensed-phase calculations on butane performed as a test of the optimized parameters. Comparison with previous results and comments on the presented approach are included in the Discussion section. Via inclusion of *ab initio* interaction energies and geometries as goal data, in addition to experimental condensed phase properties, a new set of alkane L-J parameters were developed using the novel L-J parametrization methodology presented in this study. Also included are data on the internal portion of the force field optimized to be consistent with the new L-J parameters.

Methods

L-J parameters for alkanes were developed for use with the CHARMM all-hydrogen potential energy function.[12,33,36] The CHARMM potential energy function is shown in eq. (1), where b, θ, ϕ, r_{ij}, S, and w from, for example, an X-ray structure, are the bond length, bond angle, dihedral angle, nonbonded distance between atoms i and j, distance of 1,3 atom pairs, and improper dihedral angle, respectively. Intramolecular parameters K_b, b_0, K_θ, θ_0, K_ϕ, δ, n, K_{ub}, S_0, K_w, and w_0 are the bond-stretching force constant, equilibrium bond distance, angle bending force constant, equilibrium bond angle, force constant of the dihedral angle, phase of the dihedral angle, multiplicity of the dihedral angle, Urey–Bradley (1,3) force constant, Urey–Bradley equilibrium distance, improper dihedral force constant, and improper dihedral equilibrium angle, respectively. Intermolecular parameters ε_{ij}, $R_{\min,ij}$, q_i, and D are the Lennard–Jones well depth and radius between atoms i and j, the partial atomic charge of atom i, and the dielectric constant, respectively:

$$V_T = \sum_{\text{bonds}} K_b(b - b_0)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedrals}} K_\phi[1 + \cos(n\phi - \delta)] + \sum_{\text{1,3 pairs}} K_{ub}(S - S_0)^2 + \sum_{\text{improper}} K_w(w - w_0)^2 + \sum_{\text{nonbonded}} \varepsilon_{ij}\left[\left(\frac{R_{\min,ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{\min,ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{4\pi D r_{ij}} \quad (1)$$

Empirical interaction energy surfaces were calculated using versions 23 and 24 of CHARMM.[37] The starting parameters used in this study were obtained from the CHARMM parameter set.[33,36,38] Model compound structures used for empirical calculations were the *ab initio* MP2/6-31G(d)-optimized structures. For the interaction energy surfaces, the internal geometries of the model compounds were fixed. L-J parameters for helium ($R_{\min}$ = 1.48 Å, ε = 0.021 kcal/mol) and neon ($R_{\min}$ = 1.53 Å, ε = 0.085 kcal/mol) have been parameterized to reproduce experimental homodimer interaction energy surfaces.[39–41] Partial atomic charges for neon and helium were set to zero.

Condensed phase simulations, including free energy perturbations, were performed with the program BOSS,[42] in which the combining rules were converted to the Lorentz–Berthelodt rules where the geometric mean is used for ε and the arithmetic mean is used for $R_{\min}$. In all cases, the simulations were performed in the NPT ensemble at 1 atm. Move sizes were adjusted to yield approximately 40% acceptance rates. Temperatures for pure solvent simulations were −162.00°, −88.63°, −42.07°, and −0.50°C for methane, ethane, propane, and butane, respectively. Aqueous solvation free energy perturbations were performed at 298 K. Internal geometries were fixed except for dihedral angles. The simulations were run for 1×10^6 configurations of equilibrium and 3×10^6 configurations of sampling. The pathways of the free energy perturbation calculations in

aqueous solution are represented by the arrows in Figure 1. The solute molecule was placed in the center of a periodic box consisting of 263 water molecules and any water molecules overlapping the solute were removed. Cutoffs used were 10 Å for solute–solvent interactions and 9.5 Å for solvent–solvent interactions. Instead of being perturbed to nothing directly, methane and ethane were perturbed to an extended-atom methane ($R_{\min} = 4.265$ Å, $\varepsilon = 0.20$ kcal/mol), while was previously perturbed to nothing. In the calculations, the charges and L-J parameters were linearly perturbed to 0.0, while simultaneously decreasing all bond lengths to 0.01 Å and growing in the extended-atom methane atom. In all of these calculations, only the L-J parameters of the selected model compounds were considered.

Ab initio calculations were performed with the GAUSSIAN-92 program.[43] Geometries of model compounds were optimized at the MP2/6-31G(d) level of theory. All *ab initio* calculations on interaction energy surfaces were performed at the MP3/6-311 + + G(3d,3p) level of theory with the MP2/6-31G(d) geometries. An *sp*-diffuse function for helium is not included in the standard GAUSSIAN package; however, one developed in this laboratory was used for the present computations.[25] Interaction energy surfaces were determined by systematically increasing the interaction distance by 0.1 Å and calculating the differences between the total energy and the sum of the isolated monomer energies.

Optimization of the L-J parameters was performed in an iterative fashion involving two steps: the relative and absolute optimization procedures. The relative optimization procedure was performed to yield parameters in which the relative values of the L-J parameters for different atom types (e.g., carbon and hydrogen in methane) were selected to reproduce relative *ab initio* minimum interaction energies and geometries for different rare gases to model compound interaction orientations. The absolute optimization procedure was included to insure that the absolute values of the L-J parameters yielded condensed-phase properties in good agreement with experiment. Use of the relative orientation procedure is designed to minimize problems with parameter correlation that can lead to different parameter sets yielding similar condensed phase properties. The absolute parameter optimization procedure overcomes limitations in the use of *ab initio* data as goal data due to their limited ability to accurately treat dispersion interactions, even when using large basis sets and various methods for the treatment of electron correlation (see Yin and MacKerell[25] and references therein), leading to poor agreement with experimental condensed-phase properties.

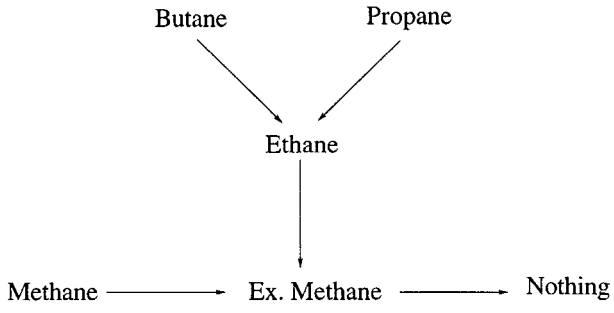

FIGURE 1. Scheme for free energy perturbation calculations in aqueous solution.

The relative L-J parameter optimization procedure was based on minimizing fluctuations of the *ab initio* versus empirical interaction energy and distance differences or ratios for selected interaction orientations between He and the model compound, or Ne and the model compound. For methane, data on the methane dimer were also included. Step 1 of the relative optimization procedure involved calculating the ratios, γ and ρ, and absolute differences, δ and ε, between *ab initio* and empirical minimum interaction energies and geometries, respectively, for a selected range of parameter sets as shown in eqs. (2a)–(2d).

$$\delta_{ij}^{\kappa} = | D_{ij,emp}^{\kappa} - D_{ij,ai} | \tag{2a}$$

$$\gamma_{ij}^{\kappa} = \frac{D_{ij,emp}^{\kappa}}{D_{ij,ai}} \tag{2b}$$

$$\varepsilon_{ij}^{\kappa} = | E_{ij,emp}^{\kappa} - E_{ij,ai} | \tag{2c}$$

$$\rho_{ij}^{\kappa} = \frac{E_{ij,emp}^{\kappa}}{E_{ij,ai}} \tag{2d}$$

In eq. (2), $D_{ij,emp}^{\kappa}$ and $D_{ij,ai}$ represent the empirical and *ab initio* minimum interaction distances, respectively, and $E_{ij,emp}^{\kappa}$ and $E_{ij,ai}$ are the empirical and *ab initio* minimum interaction energies, respectively, where i represents the dimer interaction pairs (e.g., helium–methane, neon–methane and the methane dimer), j represents the interaction orientations shown in Figures 2–4, and κ represents a set of L-J parameters. The mean ratios or differences are then determined over the different interaction orientations, j, for each interacting

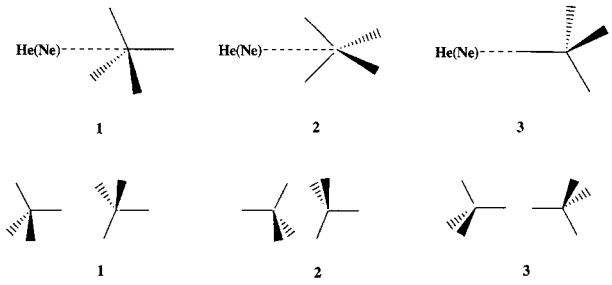

FIGURE 2. Interaction orientations between helium (neon) and methane (upper) and the methane dimer interactions (lower).

pair, i, as shown in eq. (3a)–(3d).

$$\delta_{i,\text{mean}}^{\kappa} = \frac{\sum_j \delta_{ij}^{\kappa}}{j} \tag{3a}$$

$$\gamma_{i,\text{mean}}^{\kappa} = \frac{\sum_j \gamma_{ij}^{\kappa}}{j} \tag{3b}$$

$$\varepsilon_{i,\text{mean}}^{\kappa} = \frac{\sum_j \varepsilon_{ij}^{\kappa}}{j} \tag{3c}$$

$$\rho_{i,\text{mean}}^{\kappa} = \frac{\sum_j \rho_{ij}^{\kappa}}{j} \tag{3d}$$

Root-mean-square (rms) fluctuations about the mean ratios or mean differences of the interaction orientations, j, are then obtained for each interacting pair, i, as shown in eqs. (4a)–(4d).

$$\delta_{i,\text{rms}}^{\kappa} = \sqrt{\frac{\sum_j (\delta_{ij}^{\kappa} - \delta_{i,\text{mean}}^{\kappa})^2}{j}} \tag{4a}$$

$$\gamma_{i,\text{rms}}^{\kappa} = \sqrt{\frac{\sum_j (\gamma_{ij}^{\kappa} - \gamma_{i,\text{mean}}^{\kappa})^2}{j}} \tag{4b}$$

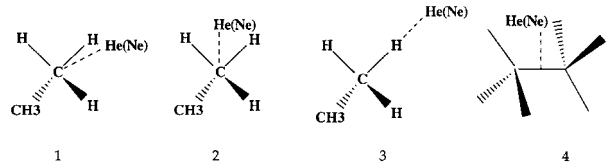

FIGURE 3. Interaction orientations between helium (neon) and ethane.

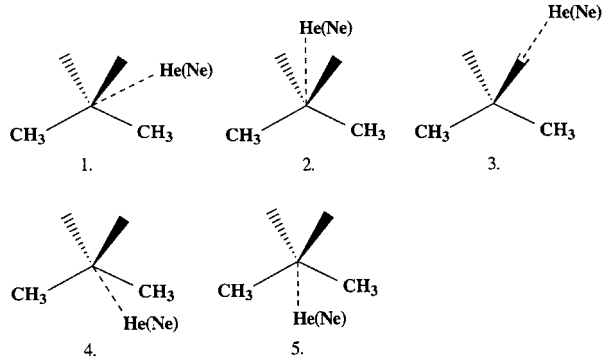

FIGURE 4. Interaction orientations between helium (neon) and propane.

$$\varepsilon_{i,\text{rms}}^{\kappa} = \sqrt{\frac{\sum_j (\varepsilon_{ij}^{\kappa} - \varepsilon_{i,\text{mean}}^{\kappa})^2}{j}} \tag{4c}$$

$$\rho_{i,\text{rms}}^{\kappa} = \sqrt{\frac{\sum_j (\rho_{ij}^{\kappa} - \rho_{i,\text{mean}}^{\kappa})^2}{j}} \tag{4d}$$

Ideally, if the ratios or differences for the different interacting orientations were identical, the rms fluctuations calculated in eq. (4) would equal zero. For example, such a case occurs when the empirical parameters produce minimum interaction energies and distances that are all offset by the same amount from the *ab initio* minimum interaction energies and distances. To allow for the simultaneous use of the both the difference and ratio rms fluctuations as criteria for selection of optimal L-J parameters, these terms must be normalized. Normalization, as shown in eq. (5), is performed based on the set of L-J parameters yielding the smallest rms fluctuations, represented by κ_{min} in eq. (5), obtained from all parameters tested in each iteration:

$$\Delta_i^{\kappa} = \frac{\delta_{i,\text{rms}}^{\kappa}}{\delta_{i,\text{rms}}^{\kappa_{\text{min}}}} \tag{5a}$$

$$\Gamma_i^{\kappa} = \frac{\gamma_{i,\text{rms}}^{\kappa}}{\gamma_{i,\text{rms}}^{\kappa_{\text{min}}}} \tag{5b}$$

$$E_i^{\kappa} = \frac{\varepsilon_{i,\text{rms}}^{\kappa}}{\varepsilon_{i,\text{rms}}^{\kappa_{\text{min}}}} \tag{5c}$$

$$P_i^{\kappa} = \frac{\rho_{i,\text{rms}}^{\kappa}}{\rho_{i,\text{rms}}^{\kappa_{\text{min}}}} \tag{5d}$$

Four selection criteria were tested to choose parameter sets yielding minimum values when sum-

ming over all the interacting pairs, i, as shown in eqs. (6a)–(6d). These criteria include: (i) rms fluctuations of energy differences plus geometry differences [eq. (6a)]; (ii) rms fluctuations of energy ratios plus geometry ratios [eq. (6b)]; (iii) rms fluctuations of energy ratios plus geometry differences [eq. (6c)], and (iv) rms fluctuations of energy differences plus geometry ratios [eq. (6d)]:

$$I^{\kappa} = \sum_{i=1}^{m} (\Delta_i^{\kappa} + E_i^{\kappa}) \tag{6a}$$

$$II^{\kappa} = \sum_{i=1}^{m} (\Gamma_i^{\kappa} + P_i^{\kappa}) \tag{6b}$$

$$III^{\kappa} = \sum_{i=1}^{m} (\Delta_i^{\kappa} + P_i^{\kappa}) \tag{6c}$$

$$IV^{\kappa} = \sum_{i=1}^{m} (\Gamma_i^{\kappa} + E_i^{\kappa}) \tag{6d}$$

Based on these criteria, the L-J parameters, κ, which produced the smallest values, were selected for the absolute optimization procedure.

The absolute optimization procedure involved using the parameter set, κ, selected from the relative optimization procedure in Monte Carlo simulations to compute the pure solvent properties of the model compound. Comparison of the calculated and experimental condensed-phase properties was then performed to select a new range of parameters to be tested in the relative optimization procedure. For example, iteration n of the relative optimization procedure for ethane selected C and H L-J parameter set 8 in Table III. These were then used in pure solvent simulations from which a heat of vaporization of 3.13 kcal/mol, which is too small compared with the experimental value of the heat of vaporization, and a molecular volume of 98.2 Å^3, which is too large, were determined. Due to the calculated heat of vaporization being not favorable enough, a new range of ε values that were larger than the values of 0.071 and 0.023 for C and H, respectively, were selected for iteration $n + 1$ of the relative optimization procedure. For the $R_{\min}$ terms, a range of values smaller than those listed for parameter set 8 (see Table III) were selected for the relative optimization procedure due to the molecular volume being too large.

Iterations between the relative and absolute L-J parametrization steps were continued until satisfactory pure solvent properties were obtained. Methane, ethane, and propane were selected as the model compounds in this study. Calculations on butane were performed to test the developed L-J parameter set.

Optimized internal geometries of the selected alkanes were obtained by minimization in the gas phase with infinite cutoffs to an rms gradient $< 10^{-6}$ kcal/mol · Å. Vibrational frequencies were calculated via the VIBRAN module in CHARMM and potential energy distributions were calculated with the MOLVIB module. Adiabatic potential energy surfaces were determined by constraining the specified dihedral angle and minimizing all remaining degrees of freedom. In the case of propane, rigid energy surfaces were obtained by fixing the internal geometry at the minimum structure and rotating a single methyl group.

Calculations were carried out on Silicon Graphics IRIS and Indigo workstations, and Cray YMP and C90 supercomputers. The relative optimization and normalization procedures, as outlined previously in eqs. (2)–(6), were carried out using a program written by our group.

Results

Sets of L-J parameters for carbon and hydrogen atoms in methane, selected based on the four relative parameter optimization selection criteria, are listed in Table I. The selections were based on interactions between methane and helium or neon and methane dimer interactions (see Fig. 2). The pure solvent properties of liquid methane calculated for the selected parameters are also listed in Table I. During this portion of the study it was observed that criteria I and IV and criteria II and III always select the same sets of parameters. Parameters selected by criteria II and III generally gave better pure solvent properties than those selected by criteria II and III. This suggests that the ratios of the *ab initio* and empirical interaction energies are most critical in determining the relative L-J parameters rather than the interaction distances. The first set of parameters in Table I best reproduced the pure solvent properties and were selected for further calculations.

Empirical minimum interaction energies and distances between helium or neon and methane and methane dimers calculated using the selected parameter set are compared with the *ab initio* results in Table II. From these data it can be seen that both the differences and the ratios of the minimum interaction distances are being satisfactorily reproduced ($> 93\%$; < 0.3 Å, respectively). The reproduction of the relative minimum interac-

TABLE I.
Comparison of Methane Experimental and Calculated Heats of Vaporization and Molecular Volumes for Parameter Sets Chosen Based on Relative Selection Criteria.

	$C(\varepsilon/R_{min}/2)$	$H(\varepsilon/R_{min}/2)$	ΔH_{vap}	MV	Diff. %
Exp.			**1.96**	**62.91**	
Results based on selection criteria II and III					
1	**0.095 / 2.11**	**0.017 / 1.34**	**1.94**	**62.90**	**−1.0 / 0.0**
2	0.080 / 2.06	0.020 / 1.30	1.83	59.74	−6.6 / −5.0
3	0.077 / 2.08	0.018 / 1.32	1.68	63.65	−14.3 / −1.2
4	0.075 / 2.06	0.022 / 1.28	1.81	59.73	−7.6 / −5.1
5	0.080 / 2.07	0.020 / 1.31	1.83	60.99	−6.6 / −3.1
6	0.082 / 2.06	0.020 / 1.30	1.86	59.44	−5.1 / −5.5
7	0.082 / 2.07	0.018 / 1.32	1.76	62.26	−10.2 / −1.0
8	0.082 / 2.05	0.020 / 1.29	1.85	58.83	−5.6 / −6.5
9	0.078 / 2.06	0.022 / 1.29	1.89	59.00	−3.6 / −6.2
Results based on selection criteria I and IV					
1	0.077 / 2.04	0.024 / 1.34	2.06	58.20	5.1 / −7.5
2	0.075 / 2.02	0.024 / 1.34	2.01	58.01	2.6 / −7.2
3	0.080 / 2.04	0.026 / 1.36	2.26	57.44	15.3 / −8.7
4	0.075 / 2.02	0.026 / 1.36	2.17	57.22	10.7 / −9.0
5	0.080 / 2.05	0.024 / 1.33	2.10	58.09	7.1 / −7.7
6	0.078 / 2.05	0.024 / 1.34	2.08	58.38	6.1 / −7.2
7	0.078 / 2.05	0.026 / 1.35	2.21	57.87	12.8 / −8.0
8	0.078 / 2.05	0.024 / 1.35	2.08	59.17	6.1 / −5.9
9	0.078 / 2.03	0.024 / 1.35	2.09	58.03	6.6 / −7.8

ε: kilocalories per mole; $R_{min}/2$: angstroms; heat of vaporization: kilocalories per mole, molecular volume: cubic angstroms. Diff. % = (calculated value-experimental value) / experimental value * 100 for ΔH_{vap}/MV. Typical calculated errors for ΔH_{vap} are ±0.005 and for MV are ±0.16. Parameter set 1 (**boldfaced**) is selected as the CHARMM96 L-J parameters for methane.

tion energies is complicated. The empirical interaction energies between helium and methane are generally more favorable than the *ab initio* results (99–138%), whereas the empirical interaction energies between neon and methane and the methane dimer are less favorable (35–53% and 52–56%, respectively). Differences between the *ab initio* and empirical results for the relative interaction energies upon going from helium to neon are consistent with *ab initio* calculations at the MP3/6-311 +

TABLE II.
Comparison of Minimum Interaction Energies and Geometries Calculated for MP3 / 6-311 + + G(3d,3p) Level and Empirical Force Field for Final Parameters.

		R_{min}				E_{min}			
		Ab initio	Emp.	δ	γ	*Ab initio*	Emp.	δ	γ
Meth_helium	1	3.4	3.5	0.1	1.03	−132	−131	−1	0.99
	2	3.7	3.5	−0.2	0.95	−98	−135	−37	1.38
	3	2.9	2.7	−0.2	0.93	−102	−108	−6	1.06
Meth_neon	1	3.3	3.5	0.2	1.06	−705	−264	441	0.37
	2	3.5	3.5	0.0	1.00	−507	−271	236	0.53
	3	2.8	2.7	−0.1	0.96	−625	−219	406	0.35
Meth_meth	1	4.0	4.2	0.2	1.05	−930	−490	440	0.53
	2	3.6	3.9	0.3	1.08	−1050	−549	501	0.52
	3	4.6	4.6	0.0	1.00	−618	−344	274	0.56

R_{min}; angstroms; E_{min}: kilocalories per mole. δ = empirical − *ab initio;* γ = empirical / *ab initio*. See Figure 2 for interaction orientations.

+ G(3d,3p) level of theory. Those calculations showed the helium dimer interaction energy to be less favorable than the experimentally based empirical data, whereas results for the neon dimer were more favorable.[25]

Interaction orientations between helium or neon and ethane are shown in Figure 3 and the selected parameter set is shown in Table III. As occurred with methane, criteria II and III and criteria I and IV selected the same sets of parameters. Several tests of the condensed phase calculations for the parameters selected by criteria I and VI were performed (not shown), and the results were compared with the corresponding results based on the parameters selected by criteria II and III. Consistent with methane, the criteria II- and III-based parameters yielded better condensed-phase properties than I and IV. Thus, only parameter sets selected by criteria II and III are listed in Table III where the calculated pure solvent properties for the selected parameter sets are compared with the experimental data. The first set of parameters in Table III best reproduced both the pure solvent properties and solvation free energy. This set of parameters was selected as the new L-J parameters for the methyl group.

Empirical minimum interaction energies and distances between helium or neon and ethane calculated using the selected parameter set are compared with the *ab initio* results in Table IV. Consistent with methane, the agreement of the interaction distances is good, whereas that of the interaction energies is poorer. The magnitudes of the empirical interaction energies between helium and ethane were again overestimated and those between neon and ethane were underestimated.

In the calculations on propane, L-J parameters for the methyl group developed for ethane were used directly and only the methylene L-J parameters were optimized. The interaction orientations between helium or neon and propane are shown in Figure 4. Again, the parameter sets selected by criteria II and III gave better calculated pure solvent properties than those selected by criteria I and IV with respect to the experimental data (not shown). Only the parameter sets selected by criteria II and III are listed in Table V. Pure solvent properties calculated based on these parameters are also compared with the experimental data in the table. Parameter set 1 best reproduced the experimental pure solvent properties. The empirical minimum interaction energies and distances calculated based on this set of parameters are compared with the *ab initio* results in Table VI. As with methane and ethane, the *ab initio* interaction distances were well reproduced by the empirical force field. For the interactions of helium and propane, the ratio of the empirical and *ab initio* distances was about 92–93%, whereas for the interaction between neon and propane, the ratio was 96–100%. The agreement of empirical and *ab initio* interaction energies was again not as good. The ratio of the empirical and *ab initio* minimum interaction energies fell in the range of 122–163%

TABLE III.
Comparison of Ethane Experimental and Calculated Heats of Vaporization and Molecular Volumes for Parameter Sets Selected by Criteria II and III.

	$C(\varepsilon/R_{min}/2)$	$H(\varepsilon/R_{min}/2)$	ΔH_{vap}	MV	Diff. %
Exp.			**3.52**	**91.5**	
1	**0.078 / 2.04**	**0.024 / 1.34**	**3.57**	**91.5**	**1.4 / 0.0**
2	0.074 / 2.00	0.026 / 1.38	3.50	93.5	−0.6 / 2.2
3	0.073 / 2.05	0.027 / 1.37	3.68	93.4	4.5 / 2.1
4	0.074 / 2.00	0.028 / 1.38	3.71	91.3	5.4 / − 0.2
5	0.077 / 2.03	0.025 / 1.35	3.60	91.5	2.3 / 0.0
6	0.073 / 2.01	0.027 / 1.39	3.63	92.6	3.1 / 1.2
7	0.073 / 1.99	0.021 / 1.37	3.02	95.8	−14.2 / 4.7
8	0.071 / 1.99	0.023 / 1.41	3.13	98.2	−11.1 / 7.3
9	0.071 / 1.97	0.028 / 1.41	3.49	94.5	−0.9 / 3.3
10	0.070 / 1.99	0.028 / 1.405	3.51	94.8	−0.3 / 3.6

ε: kilocalories per mole, $R_{min}/2$: angstroms; heat of vaporization: kilocalories per mole; molecular volume: cubic angstroms. Diff. % = (calculated-experimental value) / experimental value * 100%. Typical errors for ΔH_{vap} are ±0.01 and for MV are ±0.25. Parameter set 1 (**boldfaced**) is selected as the CHARMM96 L-J parameters for the methyl group.

TABLE IV.
Comparison of Minimum Interaction Energies and Geometries Between Ethane and Helium or Neon Calculated at MP3 / 6-311 + + G(3d,3p) Levels of Theory and Empirical Force Field for Final Parameters.

		R_{min}				E_{min}			
		Ab initio	Emp.	δ	γ	*Ab initio*	Emp.	δ	γ
Eth_helium	1	3.4	3.3	0.1	0.99	−153	−164	−11	1.07
	2	3.7	3.4	−0.3	0.92	−113	−175	−62	1.55
	3	2.9	2.7	−0.2	0.92	−112	−135	−23	1.20
	4	3.6	3.4	−0.2	0.95	−166	−225	−59	1.35
Eth_neon	1	3.3	3.4	0.1	1.02	−863	−332	531	0.38
	2	3.5	3.4	−0.1	0.98	−687	−353	334	0.51
	3	2.8	2.7	−0.1	0.98	−721	−274	447	0.38
	4	3.5	3.5	0.0	1.00	−900	−453	448	0.50

R_{min}: angstroms; E_{min}: kilocalorie per mole. δ = empirical − *ab initio*; g = empirical / *ab initio*. See Figure 3 for interaction orientation.

for helium−propane interactions and 39−55% for neon−propane interactions.

Based on the optimized L-J parameters for the methyl and methylene groups, pure solvent properties of butane were calculated. The calculated heat of vaporization and molecular volume were 5.26 ± 0.02 kcal/mol and 162.9 ± 0.6 Å^3, which are in good agreement with the experimental values of 5.35 kcal/mol and 160.3 Å^3, respectively. These values are within 2% of the experimental results.[27, 44]

Additional tests of the optimized parameters were obtained by calculating free energies of solvation in aqueous solution. These calculations were performed following the perturbation pathways shown in Figure 1 and the results from the calculations are included in Table VII. In all cases, the present parameters are within the estimated errors of the experimental data. It should be noted that the pathways for ethane, propane, and butane did not go through all-atom methane due to the poor agreement of that model with experiment.

TABLE V.
Comparison of Propane Experimental and Calculated Heats of Vaporization and Molecular Volumes for Parameter Sets Selected by Criteria II and III.

	$C(\varepsilon/R_{min}/2)$	$H(\varepsilon/R_{min}/2)$	ΔH_{vap}	MV	Diff. %
Exp.			**4.49**	**126.0**	
1	**0.056 / 2.01**	**0.028 / 1.34**	**4.51**	**126.0**	**0.4 / 0.0**
2	0.046 / 2.06	0.026 / 1.30	4.36	126.4	2.9 / − 0.3
3	0.052 / 2.04	0.024 / 1.34	4.39	126.6	2.2 / − 0.5
4	0.052 / 2.04	0.020 / 1.31	4.23	128.7	5.8 / − 2.1
5	0.055 / 2.11	0.022 / 1.31	4.54	125.3	−1.1 / 0.6
6	0.050 / 2.06	0.026 / 1.335	4.36	128.6	2.9 / − 2.1
7	0.050 / 2.06	0.026 / 1.33	4.42	126.6	1.6 / − 0.5
8	0.0515 / 2.08	0.027 / 1.35	4.56	125.3	−1.6 / 0.6
9	0.054 / 2.11	0.023 / 1.31	4.52	126.0	−0.7 / 0.0
10	0.054 / 2.11	0.023 / 1.305	4.55	125.1	−1.3 / 0.7
11	0.051 / 2.045	0.026 / 1.32	4.47	125.5	0.5 / 0.4
12	0.0515 / 2.01	0.0278 / 1.36	4.39	127.7	2.2 / − 1.4
13	0.0515 / 2.01	0.0278 / 1.33	4.40	126.6	2.0 / − 0.5
14	0.053 / 2.01	0.028 / 1.335	4.44	126.4	1.1 / − 0.3
15	0.052 / 2.08	0.028 / 1.36	4.57	126.0	−1.8 / 0.0

ε: kilocalories per mole; R_{min} / 2: angstroms heat of vaporization: kilocalories per mole; molecular volume: cubic angstroms. Diff. % = [(calculated-experimental value) / experimental value] ∗ 100 for ΔH_{vap}/MV. Typical errors for ΔH_{vap} are ±0.02 and for MV are ±0.62. Parameter set 1 (**boldfaced**) is selected as the CHARMM96 L-J parameters for the methylene group.

TABLE VI.
Comparison of Interaction Energies and Geometries Between Propane and Helium or Neon Calculated at MP3 / 6-311 + + G(3d,3p) Level of Theory and Empirical Force Field for Selected Parameters.

		R_{min}				E_{min}			
		Ab initio	Emp.	δ	γ	*Ab initio*	Emp.	δ	γ
Prop_helium	1	4.2	3.9	−0.3	0.93	−136	−182	−46	1.34
	2	3.7	3.3	−0.4	0.91	−129	−210	−81	1.63
	3	2.9	2.6	−0.3	0.93	−132	−160	−28	1.22
	4	4.2	3.8	−0.4	0.91	−175	−273	−97	1.55
	5	4.3	4.0	−0.3	0.92	−194	−312	−118	1.61
Prop_neon	1	4.0	4.0	0.0	0.99	−820	−367	453	0.45
	2	3.5	3.4	−0.1	0.97	−821	−426	395	0.52
	3	2.7	2.7	0.0	1.00	−829	−327	502	0.39
	4	4.0	3.9	−0.1	0.96	−1003	−549	454	0.55
	5	4.1	4.0	−0.1	0.99	−1166	−628	538	0.54

R_{min}: angstroms; E_{min}: kilocalories per mole. δ = empirical − *ab initio*; γ = empirical / *ab initio*. See Figure 4 for interaction orientations.

TABLE VII.
Comparison of Empirical and Experimental[27,44] Pure Solvent Properties and Solvation Free Energies for Methane, Ethane, Propane, and Butane.

	ΔH_{vap}	Diff. %	MV	Diff. %	ΔG_{solv^3}	Diff. %
Methane						
Exp.	1.96		62.8		2.00	
OPLS[1]	2.19 ± 0.00	11.6	57.2 ± 0.1	−9.1	2.20 ± 0.22	10.0
CHARMM94	2.06 ± 0.01	5.1	59.4 ± 0.2	−5.4	2.10 ± 0.25	5.0
CHARMM96	1.94 ± 0.00	−1.0	62.9 ± 0.1	0.2	1.66 ± 0.40	−17.0
Ethane						
Exp.	3.52		91.5		1.83	
OPLS[a]	3.44 ± 0.01	−2.3	92.5 ± 0.2	1.1	1.95 ± 0.23	6.6
CHARMM94[b]	3.43	−2.6	92.6	1.2	2.67 ± 0.35	45.9
CHARMM96	3.57 ± 0.01	1.4	91.5 ± 0.2	0.0	1.96 ± 0.86	7.1
Propane						
Exp.	4.49		126.0		1.96	
OPLS[a]	4.55 ± 0.01	1.3	125.2 ± 0.3	−0.6	2.18 ± 0.24	11.2
CHARMM94[b]	4.54	1.1	126.7	0.6	—	—
CHARMM96	4.51 ± 0.02	0.4	126.0 ± 0.6	0.0	2.20 ± 0.86	12.2
Butane						
Exp.	5.35		160.3		2.08	
OPLS[a]	5.43 ± 0.01	1.5	161.3 ± 0.3	0.8	2.47 ± 0.27	19.8
CHARMM94[b]	5.59 ± 0.05	4.3	155.4 ± 0.8	−3.1	—	—
CHARMM96	5.26 ± 0.02	−1.7	162.9 ± 0.6	1.6	2.33 ± 0.87	12.0

Heats of vaporization, kilocalories per mole; molecular volume: cubic angstroms; solvation free energy: kilocalories per mole. Diff. % = [(emp. − exp.) / exp.] ∗ 100%. Number of MC configurations: 3.0 M. Omitted errors were not reported.
[a]Reported by Jorgensen and coworkers.[26]
[b]A. D. MacKerell Jr., unpublished results.
[c]Errors for the free energy of solvation were the square roots of the sum of the square of each of the individual perturbation errors from the 10 windows used to sample lambda.

Internal parameters, to be consistent with the new set of L-J parameters, were optimized by reproducing the experimental internal geometries, vibrational frequencies, and torsional barriers of the selected alkanes. The calculated results and corresponding experimental data are included in the Supplementary Materials. The internal parameters and partial atomic charges used in the calculations are listed in Table S6 and Figure S1, respectively. The calculated geometries reproduced all bond lengths and angles in ethane, propane, and butane to within 0.01 Å and 1° of the experimental values, respectively.[45–48] Compared with the results from the MM3 force field,[30] CHARMM reproduced the experimental geometries of the selected alkanes to a similar or better degree. Low frequency regions of the vibrational spectrum of the selected alkane reproduced the experimental values within 0.3–0.7%. The overall percent errors for the vibrational frequencies of ethane, propane, and butane were 2.3%, 2.7%, and 3.5%, respectively, whereas the rms differences were 47, 38, 37 cm^{-1}. The corresponding values from the MM3 force field were 1.9%, 2.7%, and 3.3% and 54, 41, and 35 cm^{-1}.[31] The calculated ethane rotational barrier (2.90 kcal/mol), propane V1 and V2 rotational barriers (3.12 and 3.58 kcal/mol), and the butane torsional CCCC barrier (120°, 3.47 kcal/mol) were all in good agreement with the available experimental or *ab initio* values, as shown in Table S5 of the Supplementary Materials. For butane, the potential energy surface of the dihedral angle CCCC rotation is shown in Figure S2. *Ab initio* data from the HF/6-31G(d) level are also shown. Empirical results satisfactorily reproduced the *ab initio* data. The relative empirical gauche energy of butane was 0.83, which compares with experimental values ranging from 0.75 to 0.97 kcal/mol. The main contribution to the empirical energy surface was from the dihedral angle term. Contributions from electrostatic, L-J, and angle terms were significant but smaller. To gauge the relative contributions of the LJ and electrostatic terms to the butane surface, *ab initio* HF/6-31G(d) and empirical values for the terminal CCH and CCC angles are shown in Figure S3 as a function of the terminal HCCC and CCCC dihedral, respectively. The empirical values adequately reproduced the *ab initio* data indicating that the nonbonded contributions to the butane surface, based on changes in the internal geometry of the molecule, were appropriate. The overall agreement of the calculated internal geometries, vibrational frequencies, and torsional surfaces with respect to the experimental and *ab initio* results were satisfactory.

Discussion

A summary of the results from the present parametrization project (CHARMM96) along with data from the OPLS[26] and CHARMM22[49] force fields are presented in Table VII. The CHARMM22 parameters will be referred to as CHARMM94 to be consistent with previous work,[26] some of the reported results in Table VII are those reported in that same work. For the pure solvent properties of methane, ethane, and propane, the CHARMM96 parameters represent an improvement over currently used parameter sets. With the free energies of solvation, however, the present results are similar to the published OPLS data, with the largest discrepancy occurring in the case of methane. For butane, which was not explicitly parameterized in the present study, the CHARMM96 parameters yield pure solvent properties that are in slightly poorer agreement than the OPLS values, although they are still within 2% of experiment; the CHARMM96 free energy of solvation is in better agreement with experiment. Differences between the CHARMM96 and experimental aqueous solvation free energies are all within the estimated errors for the computed values. Overall, the ability of the CHARMM96 parameters to reproduce experimental condensed-phase properties supports the validity of the methodology developed for L-J parametrization. Due to the extensive amount of work on alkane force field parametrization it was not expected that significant improvements over present force fields would be obtained. Data from those previous studies, however, make the alkanes the optimal class of compounds to develop and test a new method for the optimization of L-J parameters.

Comparison of the OPLS, CHARMM94, and CHARMM96 alkane parameters is made in Table VIII. OPLS uses a single set of L-J parameters for all aliphatic carbons, whereas the CHARMM94 and CHARMM96 parameter sets vary the carbon L-J parameters according to the extent of substitution of the carbon. CHARMM96 also allows the hydrogen L-J parameters to vary as a function of substitution. It should be noted that OPLS and CHARMM use different combination rules for the L-J radii (geometric vs. arithmetic means), which could contribute to differences in the parameters.

Regardless, the OPLS carbon radii are smaller than the CHARMM96 parameter values, whereas the well depth falls within the range of the CHARMM96 values. The smaller radius of the OPLS carbon appears to be compensated by a larger radius on the hydrogens as compared with the CHARMM96 parameters, yet the OPLS hydrogen well depth is more favorable than the CHARMM96 parameters. Comparison of CHARMM94 and CHARMM96 shows the values to be similar. However, upon going from the methyl to methylene groups, the carbon radius in CHARMM94 increases, whereas that in CHARMM96 decreases. In both cases, the carbon well depths decrease, although the change in CHARMM96 is smaller. The opposite trends in the carbon radii for the two CHARMM sets may be due, in part, to the increasing well depth on the hydrogens in CHARMM96 as well as the larger change in the carbon well depth in CHARMM94. Although differences between the three parameter sets are difficult to interpret in detail, two effects may be contributing, in addition to the combination rules. Parameter correlation may still be allowing for the variations in the parameter sets listed in Table VIII while yielding experimental values that are all in reasonable agreement with experiment, and the use of different water models (see subsequent text) may contribute.

Parameter correlation is most likely responsible for the difference between the CHARMM94 and CHARMM96 parameters. Comparison of the ethane parameters show them to be similar. However, significant differences in the aqueous solvation free energy occur (Table VII). The relatively small difference in L-J parameters leading to the significant difference in the free energy of solvation emphasizes the importance of proper L-J parameters for condensed-phase calculations. The difference between the carbon methyl and methylene parameters for CHARMM94 versus CHARMM96 is related to inclusion of the *ab initio* data and allowing the hydrogen L-J parameters to vary in the present study. In parametrization of CHARMM94, optimization of the methyl parameters included methane dimer and ethane dimer *ab initio* data along with the ethane pure solvent properties;[38] optimization of the methylene parameters, however, was based only on the propane pure solvent properties. The additional *ab initio* data in the present study lets the problem become better determined and, therefore, may be assumed to yield more accurate parameters. It should be noted that although the hydrogen radii were allowed to vary in the optimization procedures (see Tables I, III, and V), they were shown to be equivalent for the selected methane, methyl, and methylene parameter sets.

TABLE VIII.
L-J Parameters from OPLS, CHARMM94 and New CHARMM96 Parameters for Alkanes.

Parameter set	Carbon		Hydrogen	
	$R_{min}/2$	ε	$R_{min}/2$	ε
OPLS	1.965	0.066	1.403	0.030
CHARMM94				
Methyl	2.060	0.080	1.320	0.022
Methylene	2.175	0.055	1.320	0.022
CHARMM96				
Methane	2.110	0.095	1.340	0.017
Methyl	2.040	0.078	1.340	0.024
Methylene	2.010	0.056	1.340	0.028

$R_{min}/2$: angstroms; ε: kilocalories per mole. OPLS σ values were converted to R_{min} values via multiplication by $2^{1/6}$.

Upon going from OPLS to CHARMM96 the degree of flexibility in the number of L-J parameters being optimized increased. Accordingly, it is expected that the extent of agreement with experiment would also increase. For the pure solvent properties of methane, ethane, and propane, this is approximately true. With free energies of solvation, however, for given molecules the best parameter set varies, although the differences are generally within the error estimates for the calculations. For methane, the OPLS and CHARMM94 pure solvent properties are significantly worse than CHARMM96, whereas the aqueous solvation results are better. This indicates that a trade-off is occurring between the methane pure solvent and aqueous solvation properties that may not be reconciled within the context of the present form of the potential energy function (see [eq. (1)]).[26] Such a discrepancy represents a possible test for improving empirical force field calculations based on extended forms of the potential energy function.

Differences between the CHARMM96 and OPLS/CHARMM94 free energies of solution may also be attributed to the use of the TIP3P water model in the present study as compared with the use of the TIP4P model in the OPLS/CHARMM94 work.[26] Comparison of the two water models shows the TIP4P model to have generally improved properties with respect to experiment as compared with the TIP3P model.[50] Because the experimental free energies of solvation of alkanes are dominated by entropic effects,[27] these dif-

ferences may be significant. For example, the interaction energy per molecule of TIP3P is less favorable (−9.86 kcal/mol) than that of TIP4P (−10.07 kcal/mol).[50] Furthermore, the second peak in the O—O radial distribution function observed experimentally[51] is nearly absent in the TIP3P model, while being well represented by the TIP4P model. These differences suggest that the TIP4P model may more readily form the ''caged'' water structures around the nonpolar solutes that are responsible for the entropic contributions to the free energies of solvation. Proper ''cage'' structures may also contribute to the energetic interactions between water and the solutes that are related to enthalpic contributions to the free energies of solvation. It has previously been noted that the values of optimized parameters are dependent on the water model used in the parametrization.[7]

During the parameter optimization procedure several interesting trends were observed. In the determination of L-J parameters that minimized the relative interaction energies and distances of different orientations the energetic term dominated the selection criteria. Furthermore, the energetic criteria based on the ratio rather than the differences gave better agreement concerning the pure solvent properties. The greater influence of the energetics versus the distance may be related to the 0.1-Å increments at which the distance data were obtained and to the interaction energy surfaces being relatively flat in the region of the minima.[25] The improved agreement with condensed-phase properties when the energetic criteria based on ratios is used, as opposed to differences, is difficult to interpret. A simple explanation is that the use of ratios led to smaller deviations in the selected parameter sets as compared with the input parameter set, which already yielded reasonable agreement with experiment. This would produce parameter sets that yield better agreement with the experiment than the use of the differences. This, however, does not explain the trend (see Table I) whereby the use of differences leads to heats of vaporization that are too large and molecular volumes that are too small. Work is in progress to clarify these results.

Trends were also observed with respect to the agreement between the empirical and *ab initio* energy data for different interaction orientations involving helium and neon. Analysis of interaction orientations 1 and 3 in Figures 2, 3, and 4 and the minimum interaction energies in Tables II, IV, and VI shows those ratios to be small as compared with orientations 2, 4, and 5. Orientations 1 and 3 represent the most direct interactions between the hydrogen and carbon atoms, respectively. Orientations 2, 4, and 5 are such that secondary interactions from either hydrogens or methyl groups make significant contributions to the energy surfaces. The less favorable empirical interaction energies, relative to the *ab initio* values, for the direct interactions may represent limitations in the spherical model of atoms used in the present potential energy function [eq. (1)]. Deformation of the electron distribution in the direct interactions will tend to make the interaction energies more favorable in the *ab initio* calculations. Such deformation cannot occur in the empirical results. The omission of such deformations, however, will not have as large an effect on the interaction energy surfaces for orientations 2, 4, and 5 due to the significant secondary contributions. The present parametrization methodology is designed to obtain L-J parameters that are a compromise between the different types of interaction orientations. As may be seen the extent of the compromise is limited by the form of the L-J equation and the spherical model of atoms. Further advances may require altering these assumptions.

Of note is the consistent agreement of the minimum interaction distances between the empirical and *ab initio* data. The empirical data for helium tended to be shorter than the *ab initio* values, whereas the neon values were in good agreement. Such agreement supports the previous assertion based on *ab initio* calculations of the helium and neon homodimers, that the obtained distances at the MP3/6-311++G(3d,3p) level of theory are representative of the experimental regime.[25] It should be noted that the *ab initio* calculations are being performed without corrections for basis-set superposition error (BSSE). Omission of BSSE corrections was chosen to avoid problems due to outward shifts in the position of minimum interaction distances observed in the *ab initio* calculations. Although this omission appears to be responsible for the overestimation of the *ab initio* minimum interaction energies, as compared with the empirical values for neon, we consider this sacrifice appropriate to minimize the problems induced by its inclusion with respect to the position of the minima. As discussed in the Introduction, work by Tsuzuki et al. using the MP3/6-311(3d,3p) level of theory with BSSE correction yielded parameters for methane that produce poor condensed-phase properties.[14] Accordingly, the need to include experimental data along with *ab initio* results, as performed in the present study, is evi-

dent. Future increases in computational resources will allow for improvements in the quality of the *ab initio* level of theory, but presently the use of experimental data for determination of the absolute values of the L-J parameters is assumed to compensate for limitations in the *ab initio* data.

Conclusion

We have presented a novel methodology for the optimization of Lennard-Jones parameters for empirical force field calculations. The procedure includes advantages associated with *ab initio*–based data, allowing for microscopic interactions to be considered. Inclusion of experimental pure solvent properties eliminates problems inherent in *ab initio* approaches associated with limitations in currently accessible levels of theory with respect to dispersion–attraction interactions. In the present study it was shown that parameters optimized based on *ab initio* and pure solvent properties alone yield satisfactory agreement with experiment concerning the free energies of aqueous solvation. Thus, the combination of goal data used in the proposed optimization procedure appears to yield parameters that are transferable to other condensed phase environments. The availability of L-J parameters that yield quantitatively accurate condensed-phase properties will extend the usefulness of empirical force field computations. Work in our laboratory has successfully applied this methodology to determine L-J parameters for alkenes.[52] Current efforts are applying the presented methodology to develop L-J parameters for polar-neutral model compounds.

Supplementary Material

Supplemental materials include the optimized internal geometries, vibrational frequencies, and rotational barriers for the selected alkanes. The partial atomic charges and the internal parameters used in this study are also included.

Acknowledgments

Computational support for this project was provided by the NCI's Frederick Biomedical Supercomputing Center and the Pittsburgh Supercomputer Center.

References

1. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, New York, 1989.
2. W. L. Jorgensen, *J. Am. Chem. Soc.*, **106,** 6638 (1984).
3. W. L. Jorgensen and C. J. Swenson, *J. Am. Chem. Soc.*, **107,** 569 (1985).
4. A. D. MacKerell Jr. and M. Karplus, *J. Phys. Chem.*, **95,** 10559 (1991).
5. A. T. Hagler, E. Huler, and S. Lifson, *J. Am. Chem. Soc.*, **96,** 5319 (1974).
6. S. Lifson and M. Levitt, *Comput. Chem.*, **3,** 49 (1979).
7. A. D. MacKerell Jr., J. K. Wiorkiewicz, and M. Karplus, *J. Am. Chem. Soc.*, **117,** 11946 (1995).
8. A. Warshel and S. Lifson, *J. Chem. Phys.*, **53,** 582 (1970).
9. W. L. Jorgensen, *J. Phys. Chem.*, **90,** 1276 (1986).
10. D. L. Veenstra, D. M. Ferguson, and P. A. Kollman, *J. Comput. Chem.*, **13,** 971 (1992).
11. C. A. Gough, S. E. DeBolt, and P. A. Kollman, *J. Comput. Chem.*, **13,** 963 (1992).
12. A. D. MacKerell Jr., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, I. W. E. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, *FASEB J.*, 1992, 6:A143.
13. G. Liang, P. C. Fox, and J. P. Bowen, *J. Comput. Chem.*, **17,** 940 (1996).
14. S. Tsuzuki, T. Uchimaru, K. Tanabe, and S. Kuwajima, *J. Phys. Chem.*, **98,** 1830 (1994).
15. J. A. Riddick, W. B. Bunger, and T. K. Sakano, *Organic Solvents: Physical Properties and Methods of Purification, Vol. II*, Wiley, New York, 1986.
16. W. L. Jorgensen, *J. Am. Chem. Soc.*, **103,** 335 (1981).
17. B. Kukawska-Tarnawska, G. Chalasinski, and M. M. Szczesniak, *J. Mol. Struct.*, **297,** 313 (1993).
18. S. M. Chalasinski, M. M. Szczesniak, and S. Scheiner, *J. Chem. Phys.*, **97,** 8181 (1992).
19. G. Chalasinski, M. M. Szczesniak, and S. Scheiner, *J. Chem. Phys.*, **94,** 2807 (1991).
20. G. Chalasinski, S. M. Cybulski, M. M. Szczesniak, and S. Scheiner, *J. Chem. Phys.*, **91,** 7809 (1989).
21. M. M. Szczesniak, G. Chalasinski, and S. M. Cybulski, *J. Chem. Phys.*, **96,** 463 (1992).
22. G. Chalasinski, *Chem. Rev.*, **94,** 1723 (1994).
23. B. Jeziorski, R. Moszynski, and K. Szalewicz, *Chem. Rev.*, **94,** 1887 (1994).
24. F. B. van Duijneveldt, J. G. C. M. van Duijneveldt-van de Rijdt, and J. H. van Lenthe, *Chem. Rev.*, **94,** 1873 (1994).
25. D. Yin and A. D. MacKerell Jr., *J. Phys. Chem.*, **100,** 2588 (1996).
26. G. Kaminski, E. M. Duffy, T. Matsui, and W. L. Jorgenson, *J. Phys. Chem.*, **98,** 13077 (1994).
27. A. Ben-Naim and Y. Marcus, *J. Chem. Phys.*, **81,** 2106 (1984).
28. Y. Sun, D. Spellmeyer, D. A. Pearlman, and P. Kollman, *J. Am. Chem. Soc.*, **114,** 6798 (1992).

29. J. Nagy, D. F. Weaver, and J. V. H. Smith, *J. Phys. Chem.*, **99,** 8058 (1995).
30. N. L. Allinger, Y. H. Yuh, and J. H. Lii, *J. Am. Chem. Soc.*, **111,** 8551 (1989).
31. J. H. Lii and N. L. Allinger, *J. Am. Chem. Soc.*, **111,** 8566 (1989).
32. J. H. Lii and N. L. Allinger, *J. Am. Chem. Soc.*, **111,** 8576 (1989).
33. J. C. Smith and M. Karplus, *J. Am. Chem. Soc.*, **114,** 801 (1992).
34. J. L. M. Dillen, *J. Comput. Chem.*, **16,** 595 (1995).
35. J. L. M. Dillen, *J. Comput. Chem.*, **16,** 610 (1995).
36. M. Schlenkrich, J. Brickmann, A. D. MacKerell Jr., and M. Karplus, In *Biological Membranes, a Molecular Perspective from Computation and Experiment*, K. M. Merz Jr., and B. Roux, Eds., Birkhauser, Boston, 1996.
37. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.*, **4,** 187 (1983).
38. A. D. MacKerell Jr. and M. Karplus, manuscript in preparation.
39. R. K. Feltgen, H. Kohler, K. A. Pauly, and H. Torello, *J. Chem. Phys.*, **76,** 2360 (1982).
40. R. A. Aziz, *Chem. Phys. Lett.*, **40,** 57 (1976).
41. R. A. Aziz, W. J. Meath, and A. R. Allnatt, *Chem. Phys.*, **78,** 295 (1983).
42. W. L. Jorgensen, *BOSS, version* 3.4, Yale University, New Haven, CT, 1993.
43. M. J. Frisch, G. W. Trucks, M. Head-Gordon, P. M. W. Gill, M. W. Wong, J. B. Foresman, B. G. Johnson, H. B. Schlegel, M. A. Robb, E. S. Replogle, R. Gomperts, J. L. Andres, K. Raghavachari, J. S. Binkley, C. Gonzalez, R. L. Martin, D. J. Fox, D. J. Defrees, J. Baker, J. J. P. Stewart, and J. A. Pople, *GAUSSIAN*-92, *revision A*, Gaussian, Inc., Pittsburgh, PA, 1992.
44. B. D. Smith and R. Srivastave, *Thermodynamic Data for Pure Compounds, Part A*, Elsevier, New York, 1986.
45. J. R. Hoyland, *J. Chem. Phys.*, **50,** 2775 (1969).
46. K. Wiberg and R. H. Boyd, *J. Am. Chem. Soc.*, **94,** 5236 (1972).
47. P. K. Heenan and L. S. Bartell, *J. Chem. Phys.*, **78,** 1270 (1983).
48. W. F. Bradford, S. Fitzwater, and L. S. Bartell, *J. Mol. Struct.*, **38,** 185 (1977).
49. A. D. MacKerell Jr. and M. Karplus, manuscript in preparation.
50. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.*, **79,** 926 (1983).
51. A. K. Soper and M. G. Philips, *Chem. Phys.*, **107,** 47 (1986).
52. S. E. Feller, D. Yin, R. W. Pastor, and A. D. MacKerell Jr., *Biophys. J.*, in press.