# COMPASS: An ab Initio Force-Field Optimized for Condensed-Phase Applications—Overview with Details on Alkane and Benzene Compounds

## H. Sun

*Molecular Simulations Inc., 9685 Scranton Road, San Diego, California 92121-3752*

*Received: January 29, 1998; In Final Form: May 22, 1998*

A general all-atom force field for atomistic simulation of common organic molecules, inorganic small molecules, and polymers was developed using state-of-the-art ab initio and empirical parametrization techniques. The valence parameters and atomic partial charges were derived by fitting to ab initio data, and the van der Waals (vdW) parameters were derived by conducting MD simulations of molecular liquids and fitting the simulated cohesive energies and equilibrium densities to experimental data. The combined parametrization procedure significantly improves the quality of a general force field. Validation studies based on large number of isolated molecules, molecular liquids and molecular crystals, representing 28 molecular classes, show that the present force field enables accurate and simultaneous prediction of structural, conformational, vibrational, and thermophysical properties for a broad range of molecules in isolation and in condensed phases. Detailed results of the parametrization and validation for alkane and benzene compounds are presented.

## I. Introduction

Force-field development, as a fundamental issue underlying all atomistic simulations, has drawn considerable attention in recent years. This is marked by publications of several revised or newly developed general force fields in the last 10 years. Among many of them, MM3,[1] MM4,[2] Dreiding,[3] SHARP,[4] VALBON,[5] UFF,[6] CFF93,[7] AMBER,[8] CHARMM,[9] OPLS,[10] and MMFF[11] are a few examples.

Roughly speaking, three trends can be classified in the developments of these contemporary force fields. In one direction, the force fields were made to be very generic so that great coverage could be achieved. At the extreme, UFF[6] was designed to cover molecules of any combination of elements on the periodic table. Simple functional forms are used for the diagonal terms of the force-constant matrix. Because of the generality of parametrization, these force fields are normally expected to yield reasonable predictions of molecular structures only. In another direction, emphasis was made to improve the quality of prediction in a rather focused area of applications (mostly in biochemistry). Recently, much attention has been given to the prediction of condensed-phase properties. This trend is clearly seen in the new versions of AMBER[8] and CHARMM.[9] In particular, Jorgensen and co-workers published OPLS/AMBER force field[10] in which the authors extended their well-known OPLS force-field approach from a united atom model[12] to an explicit all-atom force field. Similar to the first category, these force fields use simple functional forms. In the last category, attention was paid to achieve high accuracy in predicting various molecular properties with a fairly broad coverage. The properties of interest generally included molecular structures, conformational properties, vibration frequencies, and heats of formation. To achieve this goal, complicated functional forms including off-diagonal cross-coupling terms and high-order (cubic and quartic) force constants were used. Force fields such as MM3,[1] MM4,[2] CFF93,[7] and MMFF[11] belong to this category. The parameters were derived using high-quality experimental data (MM3/MM4) or quantum mechanics ab initio data (CFF93, MMFF). With the great

flexibility of the functional forms and immense quantity of data used in the training set, these force fields were parametrized accurately. In many cases, the calculation errors are within the experimental precision. However, applications of these force fields for molecules in condensed phases have been limited to the energy minimization of molecular crystals.

The force-field method uses a set of empirical formulas to mimic the interatomic interactions in an average fashion.[13,14] Atoms in different chemical environments are classified into different 'atom types'. By ignoring the details of electron—electron and electron—nucleon interactions, the force-field method works at the atomic level. In principle, the force-field method can be used when details of the electron distribution are not required for describing the properties of interest. In the past, force fields were widely used for predicting structures, vibration frequencies, and conformational properties for molecules in isolation. Today, with the fast progress of computer hardware and calculation algorithms (in particular, the density functional theory), quantum mechanics methods have been increasingly used for these applications. It is for large molecular systems and molecules in condensed phases (from a few hundred to million atoms) that the force-field method clearly has an incomparable advantage over ab initio methods. This is not only because the force-field method is several order of magnitude faster than any ab initio method, but also because, fundamentally, an ab initio method is often not necessary for these applications. The properties of interest in large-scale simulations are usually relevant to the statistics of atomic motion in a much longer time scale than the rapid electron motion that an ab initio method describes. In addition, the most important interaction terms in the condensed-phase simulations—the nonbond (dispersion in particular) forces—are extremely difficult to describe accurately using ab initio methods. Consequently, a force-field method should emphasize the prediction of properties of large molecules and molecules in condensed phases.

Most of the well-parametrized force fields, such as MM3, MM4, CHARMM, AMBER, CFF93, and MMFF, were de-

COMPASS—Details on Alkane and Benzene Compounds

*J. Phys. Chem. B, Vol. 102, No. 38, 1998* **7339**

signed mainly for biologically interesting molecules. Although there is no fundamental difference in the force-field parametrization between synthetic and natural polymers, an urgent need to deploy force-field methods in materials science prompted us to develop a force field especially for organic materials and polymers at the beginning of this decade.[15] Starting with the protein CFF91 force field that was later developed into CFF93,[7] a dozen functional groups of most common organic and inorganic polymers were parametrized. Some of these force-field developments have been published.[16] The force field was named polymer consistent force field (PCFF).[17]

Like the CFF93 force field,[7] PCFF is an ab initio force field. Most parameters were derived based on ab initio data using a least-squares-fit technique developed by Hagler and co-workers.[18] Many of the nonbond parameters of PCFF, which include atomic partial charges and Lennard—Jones 9-6 (LJ-9-6) parameters, were taken from the CFF91 force field. Similar to many other force fields in this category, the nonbond parameters were derived by fitting to molecular crystal data,[19] based on energy minimization calculations.[20−22] Although these parameters perform reasonably well in various respects, it has been shown, based on numerous applications[23] of CFF91 and PCFF force fields, that these parameters are not suitable for molecular dynamics simulations at finite temperatures. Specifically, systematic errors in the pressure—volume—temperature (P—V—T) relation have been observed for liquids and polymers using MD simulations. Often, the calculated densities are too low in comparison with the experimental data.

The origin of these discrepancies is clear now. The parameters were developed based on static simulations corresponding to a classical state at 0 K, but the experimental data used to determine the parameters were measured at finite temperatures. The resulting parameters effectively contain factors such as thermal expansion and vibrational displacements at the experimental conditions. Consequently, good agreement between subsequent calculations and experiments can only be expected when (1) the calculations are performed using an energy minimization method and (2) the experimental data are measured under a condition that closely approximates those used in the parametrization.

To construct a force field generally suitable for condensed-phase applications, it was necessary to modify the nonbond parameters, and consequently, the valence parameters must also be changed due to the coupling between the valence and nonbond parameters. This paper summarizes a project that has been accomplished recently. Basically, a hybrid approach consisting of both ab initio and empirical methods was employed to derive a new general force field based on the PCFF force field. In addition to those molecular classes covered in the PCFF force field, a number of new molecular classes were parametrized. Most significantly, nonbond parameters were completely re-parametrized. The outcome is a new, condensed-phase-optimized ab initio force field. This force field is named COMPASS (condensed-phase optimized molecular potentials for atomistic simulation studies).[17]

Three papers[24−26] featuring the work of the COMPASS force field have been published, and more are in preparation.[27,28] In each of those papers, a specific functional group is discussed and detailed results of parametrization, validation, and extended application are presented. Together with the validation results, the relevant parameters are published in these papers. However, in the process of preparing these manuscripts, it becomes apparent that an article covering the generic issues relevant to this work was necessary since those issues cannot be presented well in the specific papers. Since the primary goal of this work was to prepare a general, high-quality force field for a broad range of applications, topics such as coverage of the force field, parametrization procedure, transferability of parameters, and overall quality of prediction have to be presented in one place. The present paper serves this purpose.

Although a large part of the validation data are presented in a summary style in this paper, details of parametrization and validation for the two most common classes—alkane and benzene compounds—are presented in detail. This is because some of the key points to be addressed can be well-illustrated with explicit data. It is also important because the parameters for those two functional groups are the most widely transferred parameters in any general force field. It is of great interest to ensure that these two groups are well-parametrized.

## II. The Model

**The Functional Forms.** The functional forms used in this force field are the same as those used in CFF-type force fields[7,16]

$$
\begin{aligned}
E_{\text{total}} = &\sum_b [k_2(b - b_o)^2 + k_3(b - b_o)^3 + k_4(b - b_o)^4] + \\
&\sum_\theta [k_2(\theta - \theta_o)^2 + k_3(\theta - \theta_o)^3 + k_4(\theta - \theta_o)^4] + \\
&\sum_\phi [k_1(1 - \cos \phi) + k_2(1 - \cos 2\phi) + k_3(1 - \cos 3\phi)] + \\
&\sum_\chi k_2\chi^2 + \sum_{b,b'} k(b - b_o)(b' - b'_o) + \\
&\sum_{b,\theta} k(b - b_o)(\theta - \theta_o) + \sum_{b,\phi} (b - b_o)[k_1 \cos \phi + \\
&k_2 \cos 2\phi + k_3 \cos 3\phi] + \sum_{\theta,\phi} (\theta - \theta_o)[k_1 \cos \phi + \\
&k_2 \cos 2\phi + k_3 \cos 3\phi] + \sum_{b,\theta} k(\theta' - \theta_o')(\theta - \theta_o) + \\
&\sum_{\theta,\theta,\phi} k(\theta - \theta_o)(\theta' - \theta'_o)\cos \phi + \sum_{i,j} \frac{q_i q_j}{r_{ij}} + \\
&\sum_{i,j} \epsilon_{ij} \left[ 2\left(\frac{r_{ij}^o}{r_{ij}}\right)^9 - 3\left(\frac{r_{ij}^o}{r_{ij}}\right)^6 \right] \quad (1)
\end{aligned}
$$

The functions can be divided into two categories—valence terms including diagonal and off-diagonal cross-coupling terms and nonbond interaction terms. The valence terms represent internal coordinates of bond ($b$), angle ($\theta$), torsion angle ($\phi$), and out-of-plane angle ($\chi$), and the cross-coupling terms include combinations of two or three internal coordinates. The cross-coupling terms are important for predicting vibration frequencies and structural variations associated with conformational changes. Among the cross-coupling terms given in eq 1, the bond—bond, bond—angle and bond—torsion angle are the most frequently used terms. The nonbond interactions, which include a LJ-9-6 function[21] for the van der Waals (vdW) term and a Coulombic function for an electrostatic interaction, are used for interactions between pairs of atoms that are separated by two or more intervening atoms or those that belong to different molecules. In comparison with the common LJ-12-6 function, which is known to be too 'hard' in the repulsion region, the LJ-9-6 function is softer but may be too attractive in the long separation range.[29] However, this difference appears to be unimportant to the properties of interest in this work based on a comparative study carried out at the beginning of this project.

The LJ-9-6 parameters ($\epsilon$ and $r^o$) are given for like atom pairs. For unlike atom pairs, a 6th order combination law[30] is used to

calculate the off-diagonal parameters:

$$r_{i,j}{}^\circ = \left( \frac{(r_i^\circ)^6 + (r_j^\circ)^6}{2} \right)^{1/6} \tag{2}$$

$$\epsilon_{i,j} = 2\sqrt{\epsilon_i \cdot \epsilon_j} \left( \frac{(r_i^\circ)^3 \cdot (r_j^\circ)^3}{(r_i^\circ)^6 \cdot (r_j^\circ)^6} \right) \tag{3}$$

The electrostatic interaction is represented using atomic partial charges. To make the charge parameters transferable, bond-increments $\delta_{ij}$, which represent the charge separation between two valence-bonded atoms $i$ and $j$, are used in the force field as parameters. For atom $i$, the partial charge is the sum of all charge bond increments $\delta_{ij}$

$$q_i = \sum_j \delta_{ij} \tag{4}$$

where $j$ represents all atoms that are valence-bonded to atom $i$.

In condensed-phase simulations of liquids and crystals, the nonbond interactions are usually truncated at a selected cutoff value (normally around 10 Å). A sharp cutoff is assumed for the present force field. However, the long-range interaction, which is the total contribution of nonbond interactions beyond the cutoff, is critically important to be considered for calculating energies and pressures. This issue is to be addressed later in this paper.

**The Atom Types.** A simple nomenclature rule is followed to systematically label atom types in this force field. The name string consists of the element symbol first, then a number indicating the coordination number or the number of bonds attached, and, if necessary, an additional number or character to label a special circumstance. For example, c4o indicates a carbon with four bonds attached next to an oxygen atom (an α-carbon in ethers or alcohols). All atom types that are used for the molecules calculated and reported in this paper are listed in Table 1.

Atom types are defined based on chemical intuition. It is also an empirical-based, trial-and-error practice. Basically, a new atom type is introduced when strong evidence shows that existing atom types are not adequate to describe the properties of the molecules of interest. The benefit of introducing a new atom type has to be carefully evaluated against the danger of having too many atom types. The number of parameters increases rapidly as a function of the number of atom types used: $O(N)$ for nonbond terms, $O(N^2)$ for bonds, $O(N^3)$ for angles, $O(N^4)$ for torsions and out-of-plane angles. The number of parameters gets out of control quickly as the number of atom types increases. More atom types also means more specific but less transferable parameters, so that the coverage of the force field decreases.

Since some parameters are more transferable than the others, the concept of *formal* and *actual* atom types is introduced to enable using more generic atom types for certain interaction terms. In Table 1, the *formal* types are listed in the first column, which represent all atom types that are used formally in the present force field. The *actual* types, given in columns 2−6, are those used in the definition of energy functions. There are five categories: nonbond (vdW term only), bond (including bond increments), angle, torsion, and out-of-plane. The cross-coupling terms are classified based on the number and connectivity of the atoms involved in the functions. Hence, bond−bond and bond−angle coupling terms are considered as 'angle', angle−angle is treated as 'out-of-plane', and bond−torsion,

angle−torsion, angle−torsion−angle are classified as 'torsion'. As listed in the table, the nonbond terms use the most specific definition of atom types—one-to-one mappings are found between *formal* types and *actual* types for the nonbond terms. A few generic types are used for the bond term, and many more generic types are used for angle, torsion, and out-of-plane terms. Since most parameters are related to the last three categories, using generic *actual* types for these terms significantly reduces the total number of parameters.

Similar to many other force fields,[1,2,7,11] a large number of atom types are used for the three elements C, O, N. This reflects the wide variation of the organic chemistry of these elements. Several hydrogen types are introduced, based on the polar strength of the atom that the hydrogen is attached to, from nonpolar (h1) to modest polar (h1n) and highly polar (h1o). This classification is necessary and appears to be sufficient for modeling various hydrogen bonds using the simple nonbond functions (Coulombic and LJ-9-6). For halogen atoms (F, Cl) in halogenated alkanes, the atom types are defined based on how many halogen atoms are attached to the same carbon atom. This is due to the strong interaction (anomeric effect) between the adjacent halogen atoms.[27]
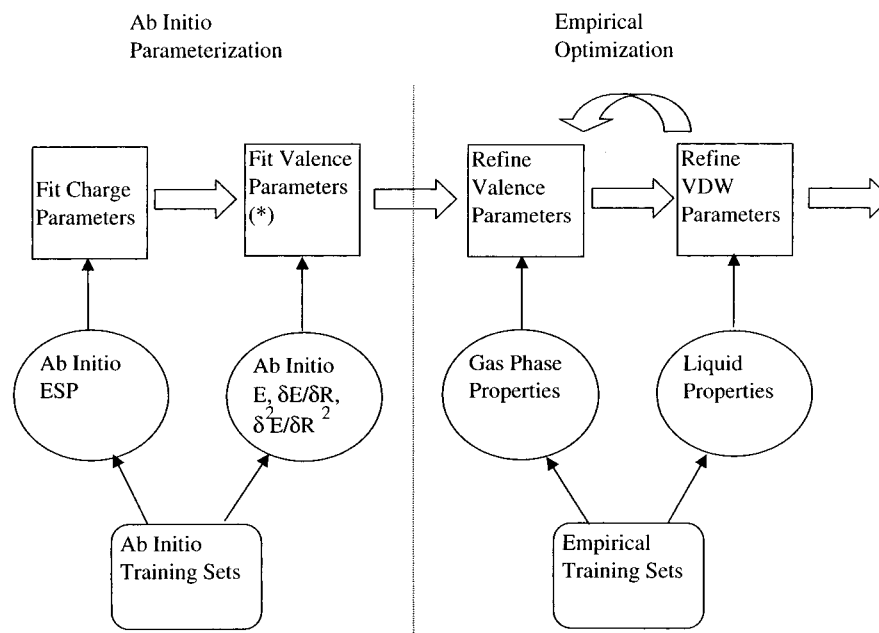
## III. Parametrization

**Ab Initio vs Empirical Parametrization.** Using ab initio data to derive force-field parameters is a more direct process than the empirical method since the energy surfaces are measured rather than 'probed' using molecular properties. However, there are two major limitations in the ab initio approach. One is the efficiency of getting accurate results using ab initio calculation methods; another is the ambiguity of the least-squares fitting of massive data to many parameters. Because of these limitations, empirical adjustments are often required. The proportion of the empirical components varies depending on the nature of the interactions.

Modest level ab initio methods are normally adequate in accuracy and efficiency for describing intramolecular energy surfaces. It is not a trivial task, however, to fit the complex energy surfaces due to the large number of degrees of freedom in the parameter space. Usually, an arbitrary 'best fit' is not necessarily the right answer (An analogue is the problem of many local minima in the conformational space of a complex molecule). Often, the obtained parameters, which depend on the initial conditions and the minimization method used, may contain terms that are physically unreasonable. Therefore, empirical control of the parametrization is required. In addition, empirical modifications are used to correct the systematic errors of the theory. This can be done by systematically scaling some or all of the force constants.[7a]

For nonbond interactions, the fitting is often not as serious a problem as that in the valence parametrization because a small number of variables is involved. However, difficulty arises from the inaccuracy in evaluating the nonbond interaction using the ab initio method. Characterizing weak nonbond interactions requires a much higher level of theory than predicting molecular structures or conformational energies. With a modest basis set and an appropriate level of theory (e.g. MP2) to incorporate the electron correlation effect, the calculated uncertainty in binding energies can be as large as the values to be characterized.[31−32] In addition, there is a problem associated with the condensed-phase behavior of molecules even if the level of theory is sufficient for characterizing the weak van der Waals interactions. Since the high-level ab initio calculations are normally performed on small molecular clusters, these calcula-

**TABLE 1: Definitions of Atom Types and Equivalence Table**

| formal type | actual type | | | | | description |
|---|---|---|---|---|---|---|
| | nonb. | bond. | angl. | tors. | oop | |
| ar | ar | ar | ar | ar | ar | argon |
| br1 | br1 | br1 | br1 | br1 | br1 | bromine, one bond |
| c1o | c1o | c1o | c1o | c1o | c1o | carbon, in carbon monoxide CO |
| c2= | c2= | c2= | c2= | c2= | c2= | carbon, sp, two double bonds O=C=O, S=C=S |
| c2t | c2t | c2t | c2t | c2t | c2t | carbon, sp, triple bond |
| c3 | c3 | c3 | c3 | c3 | c3 | carbon, $sp^2$, generic |
| c3″ | c3″ | c3″ | c3″ | c3″ | c3″ | carbon, $sp^2$, carbonyl, two polar subst |
| c3# | c3# | c3# | c3# | c3# | c3# | carbon, $sp^2$, in $CO_3^-$ anion |
| c3′ | c3′ | c3′ | c3′ | c3′ | c3′ | carbon, $sp^2$, carbonyl, one polar subst |
| c3− | c3− | c3− | c3− | c3− | c3− | carbon, $sp^2$, carboxylate |
| c3= | c3= | c3 | c3 | c3 | c3 | carbon, $sp^2$, double bond to C (−C=C−) |
| c3a | c3a | c3a | c3a | c3a | c3a | carbon, $sp^2$, aromatic |
| c3n | c3n | c3n | c3n | c3n | c3n | carbon, $sp^2$, double bond to N(−C=N−) |
| c3o | c3o | c3o | c3o | c3o | c3o | carbon, $sp^2$, carbonyl |
| c4 | c4 | c4 | c4 | c4 | c4 | carbon, $sp^3$, generic |
| c43 | c43 | c4 | c4 | c4 | c4 | carbon, $sp^3$, with 3 heavy atoms |
| c44 | c44 | c4 | c4 | c4 | c4 | carbon, $sp^3$, with 4 heavy atoms |
| c4o | c4o | c4 | c4 | c4 | c4 | carbon, $sp^3$, bond to oxygen |
| c4x | c4x | c4x | c4 | c4 | c4 | carbon, $sp^3$, bond to chlorine |
| cl1 | cl1 | cl1 | cl1 | cl1 | cl1 | chlorine, one bond |
| cl12 | cl12 | cl12 | cl1 | cl1 | cl1 | chlorine, to a carbon that has 2 halogen atoms |
| cl13 | cl13 | cl13 | cl1 | cl1 | cl1 | chlorine, to a carbon that has 3 halogen atoms |
| cl14 | cl14 | cl14 | cl1 | cl1 | cl1 | chlorine, to a carbon that has 4 halogen atoms |
| cl1p | cl1p | cl1p | cl1 | cl1 | cl1 | chlorine, in phosphazenes |
| f1 | f1 | f1 | f1 | f1 | f1 | fluorine, one bond |
| f12 | f12 | f12 | f1 | f1 | f1 | fluorine, to a carbon that has 2 halogen atoms |
| f13 | f13 | f13 | f1 | f1 | f1 | fluorine, to a carbon that has 3 halogen atoms |
| f14 | f14 | f14 | f1 | f1 | f1 | fluorine, to a carbon that has 4 halogen atoms |
| f1p | f1p | f1p | f1 | f1 | f1 | fluorine, in phosphazene |
| h1 | h1 | h1 | h1 | h1 | h1 | hydrogen, nonpolar |
| h1+ | h1+ | h1+ | h1+ | h1+ | h1+ | hydrogen, proton |
| h1h | h1h | h1h | h1 | h1 | h1 | hydrogen, in $H_2$ |
| h1n | h1n | h1 | h1 | h1 | h1 | hydrogen, bond to N,Cl |
| h1o | h1o | h1 | h1 | h1 | h1 | hydrogen, bond to O,F |
| he | he | he | he | he | he | helium |
| i1 | i1 | i1 | i1 | i1 | i1 | iodine, with one bond |
| kr | kr | kr | kr | kr | kr | krypton |
| n1n | n1n | n1n | n1t | n1t | n1t | nitrogen, in N2 |
| n1o | n1o | n1o | n1t | n1t | n1t | nitrogen, in NO |
| n1t | n1t | n1t | n1t | n1t | n1t | nitrogen, SP, 1 triple bond |
| n2= | n2= | n2= | n2= | n2= | n2= | nitrogen, $sp^2$, 1 double bond, nonaromatic |
| n2a | n2a | n2a | n2a | n2a | n2a | nitrogen, $sp^2$, 2 partial double bond, aromatic |
| n2t | n1t | n1t | n1t | n1t | n1t | nitrogen, sp, 1 triple bond, nonaromatic |
| n3 | n3 | n3 | n3 | n3 | n3 | nitrogen, $sp^3$, in amines |
| n3* | n3* | n3 | n3 | n3 | n3 | nitrogen, $sp^3$, in $NH_3$ |
| n3a | n3a | n3a | n3a | n3a | n3a | nitrogen, $sp^2$, aromatic |
| n3h1 | n3h1 | n3 | n3 | n3 | n3 | nitrogen, $sp^3$, in amines with 1 H |
| n3h2 | n3h2 | n3 | n3 | n3 | n3 | nitrogen, $sp^3$, in amines with 2 H |
| n3m | n3m | n3m | n3m | n3m | n3m | nitrogen, $sp^3$, in amides without H |
| n3mh | n3mh | n3m | n3m | n3m | n3m | nitrogen, $sp^3$, in amides with H |
| n3o | n3o | n3o | n3o | n3o | n3o | nitrogen, $sp^2$, in nitro group |
| n4+ | n4+ | n4+ | n4+ | n4+ | n4+ | nitrogen, $sp^3$, in protonated amines |
| n4o | n4o | n4o | n4o | n4o | n4o | nitrogen, $sp^3$, in amine oxides |
| ne | ne | ne | ne | ne | ne | neon |
| o1− | o1− | o1− | o1− | o1− | o1− | oxygen, $sp^2$, in carboxylate |
| o12 | o12 | o1= | o1= | o1= | o1= | oxygen, $sp^2$, in nitro group ($-NO_2$) |
| o1= | o1= | o1= | o1= | o1= | o1= | oxygen, $sp^2$, in carbonyl |
| o1=* | o1=* | o1= | o1= | o1= | o1= | oxygen, in $CO_2$ |
| o1c | o1c | o1c | o1= | o1= | o1= | oxygen, in CO |
| o1n | o1n | o1n | o1= | o1= | o1= | oxygen, in NO |
| olo | olo | olo | o1= | o1= | o1= | oxygen, in $O_2$ |
| o2 | o2 | o2 | o2 | o2 | o2 | oxygen, $sp^3$, generic |
| o2* | o2* | o2* | o2* | o2* | o2* | oxygen, $sp^3$, in water |
| o2a | o2a | o2a | o2a | o2a | o2a | oxygen, $sp^2$, aromatic, in 5 mem. ring |
| o2b | o2b | o2e | o2 | o2 | o2 | oxygen, $sp^3$, bridge atom in anhydrides |
| o2c | o2c | o2c | o2 | o2 | o2 | oxygen, $sp^3$, in acid |
| o2e | o2e | o2e | o2 | o2 | o2 | oxygen, $sp^3$, in ethers |
| o2h | o2h | o2h | o2 | o2 | o2 | oxygen, $sp^3$, in alcohol |
| o2s | o2s | o2e | o2 | o2 | o2 | oxygen, $sp^3$, in esters |
| o2z | o2z | o2z | o2 | o2 | o2 | oxygen, in siloxanes and zeolites |
| o3 | o2* | o3 | o3 | o2* | o2* | oxygen, in $H_3O^+$ |
| p4= | p4= | p4= | p4= | p4= | p4= | phosphorus, in phosphazenes |
| s2 | s2 | s2 | s2 | s2 | s2 | sulfur, 2 single bonds (−S−) |
| si4 | si4 | si4 | si4 | si4 | si4 | silicon, generic with 4 bonds |
| si4c | si4c | si4 | si4 | si4 | si4 | silicon, in siloxane with heavy atoms only |
| xe | xe | xe | xe | xe | xe | xenon |

(*) with best-guessed initial VDW parameters fixed

**Figure 1.** Flowchart of the COMPASS force-field parametrization. The large arrows indicate the direction of the process, and the small arrows indicate data flow. The entire parametrization procedure consists of two separated phases—ab initio parametrization and empirical optimization. In the first phase, only ab initio data calculated for selected molecules in the training set are used to derive the charge and valence parameters while the vdW parameters are fixed at the best-estimated initial values. In the second phase, empirical data of a separate training set are used to optimize the valence parameters and to derive the vdW parameters. Experimental data of the molecules in both gaseous and condensed phases are used in the second phase.

tions cannot capture the polarization and many-body interactions that are known to be important in condensed phases. Consequently, the parameters derived using the small-cluster approach are only good for gas-phase applications. It should be noted that the same concern also applies to empirical methods based on gas-phase properties (e.g., gas-phase second-viral coefficients). Therefore, the empirical method has to be used at least in part to derive the nonbond parameters for condensed-phase application.[33] Basically, the nonbond parameters can be determined empirically by performing condensed-phase (liquid and crystal) simulations and subsequently fitting the simulated thermophysical properties to experimental data. This method has been used by several research groups[34,35] for studying individual molecular classes. It has also been successfully used in the general force field development of OPLS.[10,12]

**Procedure.** A hybrid procedure consisting of ab initio and empirical methods was used in the development of the present force field. In Figure 1, a flowchart illustrating the parametrization procedure is given. The whole process can be divided into two phases—ab initio parametrization and empirical optimization.

The atomic partial charges were first derived based on ab initio electrostatic potential energies (ESP)[36] calculated for a set of model compounds representing the molecular class to be parametrized. The electrostatic potential (ESP) energy surfaces were calculated with the optimized structures for each of the molecules at the HF/6-31G*[37] level. The charge bond-increment parameters, $\delta_{ij}$, were derived by fitting to the ab initio electrostatic potentials using a constrained-fit scheme. Normally, the electrostatic potentials were sampled by 500−1000 grid points (depending on the size and symmetry of the molecules) laid evenly on 8−10 extended van der Waals surfaces separated by a 1.0 Å interval.
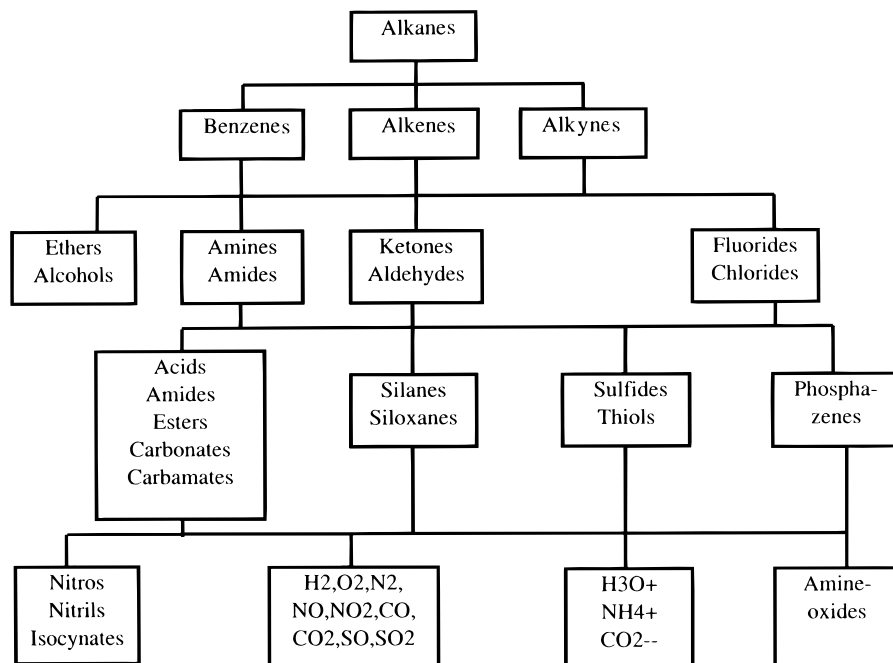
Many valence parameters of the present force field were transferred from the PCFF force field. For those functional

groups that were not available, new parameters were derived using the same CFF development method.[18] The ab initio data include total energies and first and second derivatives of the total energies for the model compounds in the training set. The optimized ab initio charge parameters were fixed during this step. To complete the functional terms, vdW parameters were also transferred from PCFF and fixed. Details of the parametrization of the valence parameters using ab initio data can be found in previous publications.[7,16,24−26]

The resulting intermediate force field, which consists of optimized charge and valence parameters, was then subjected to empirical validation and modification. Since most parameters (valence force constants and cross-coupling terms) were well-defined using the ab initio data, only a few parameters were subject to modification. The initially guessed vdW parameters, however, were reevaluated and parametrized in this phase of the work.

The valence parameters were validated based on the following intramolecular properties: molecular structures, molecular dipole moments, vibrational frequencies, and conformational energies. These calculations were performed on isolated molecules. In most cases, the calculations were full-energy minimization using a general Newton−Raphson algorithm. For conformational properties, torsion-force minimization was performed in which one or more dihedral angles were fixed while all other degrees of freedom were relaxed. The calculated results were compared with high-quality experimental data or ab initio data calculated at the MP2 level[38] with various basis sets (DZ2P to TZ2PF)[39] or DFT method[40] with the VWN[41] functional and DNP[42] basis set.

At this point, a procedure that was applied in the development of empirical force fields, such as MM3,[1] was used to refine the intermediate force field. Selecting the largest systematic discrepancy found in the comparison of the calculated results with the experimental data, the parameters that are responsible

COMPASS—Details on Alkane and Benzene Compounds

*J. Phys. Chem. B, Vol. 102, No. 38, 1998* **7343**

Alkanes

Benzenes — Alkenes — Alkynes

Ethers / Alcohols — Amines / Amides — Ketones / Aldehydes — Fluorides / Chlorides

Acids / Amides / Esters / Carbonates / Carbamates — Silanes / Siloxanes — Sulfides / Thiols — Phospha-zenes

Nitros / Nitrils / Isocynates — H2,O2,N2, NO,NO2,CO, CO2,SO,SO2 — H3O+ / NH4+ / CO2-- — Amine-oxides

**Figure 2.** Parametrization precedence tree. The parametrization started from the top. All parameters determined at one level were fixed and transferred as many times as possible to the next level. New atom types and parameters were introduced only with strong evidence indicating that it was necessary.

for the largest portion of the error were identified and adjusted to improve the agreement. Then the next largest error was considered and so on. Among the properties calculated, the molecular structures (bond lengths and angles) are the most fundamental properties that have a strong impact on other properties. The most critical parameters for the molecular structures are usually the reference values of the internal coordinates ($b_o$, $\theta_o$). The molecular dipole moments are normally well-reproduced with the ab initio partial charges if the structures are right. The conformational energies are mostly influenced by torsion parameters. The vibrational frequencies are sensitive to the force constant ($k_2$) and some coupling terms. Due to the coupling among the different terms, this process had to be iterated in order to fit all properties consistently. Because of the small number of adjustable parameters, this step was a relatively straightforward task in most cases.

After the intermediate force field was optimized to yield good agreement with the experimental data for molecules in the gas phase, the LJ-9-6 parameters were subjected to refinement using MD simulations of molecules in condensed phases. In almost all cases, this was done by calculating and comparing two physical properties of the molecular liquids—densities and cohesive energies at given temperatures and pressures. Crystal data were used for parametrization only when liquid data were not available. Modification of the adjustable parameters was performed in a trial-and-error manner. The parameters were subjected to stepwise modifications, and for each set of parameters, MD simulations were performed to calculate the properties. The final parameters were determined based on all samples simulated. To enhance the ratio between the observables and parameters, several molecules were treated at the same time for each of the molecular classes.

After the LJ-9-6 parameters were optimized, intramolecular properties were calculated again to check if any adjustment was required for the valence parameters. It was found that only a few valence parameters, mostly torsion terms, needed to be adjusted if the initial estimates of the LJ-9-6 parameters were reasonably close to the optimized values. Furthermore, small

changes of the torsion parameters normally had a negligible impact on the condensed-phase properties. In other words, modification of vdW parameters may change the torsion energy profiles but changes in the torsion parameters have very little influence on the densities and cohesive energies of molecular liquids or crystals.

**Compatibility and Transferability of Parameters.** To make a general force field, it is critical to ensure compatibility of the parameters among different functional groups. This is normally not a problem in parametrization of a force field for a particular functional group but is highly challenging for a general force field development. The central issue is to use as few parameters as possible so that a broad coverage can be achieved with approximately the same level of quality.

During this work, the comparability of parameters was handled by carefully transferring parameters among chemical classes as much as possible. A path of parametrization, as sketched in Figure 2, was followed. The most common functional groups, hydrocarbons that can be found in many chemical compounds, were parametrized first. After the parameters were optimized at any given level, they were fixed and transferred to the next level. During a later stage, only new parameters that were unique to the functional groups of interest were relaxed and optimized. This normally involved using new atom types, which were introduced only with strong evidence indicating that previously optimized parameters were not adequate.

Because the parameters were always transferred first and then tested, the transferability was ensured during the entire development process. The alkane and benzene parameters, which are the most widely transferred, are used in many other chemical classes. Generally speaking, there is no need to introduce new parameters for alkyl and phenyl groups in different chemical compounds. It should be noted that the transferability is a variable depending on the accuracy level one sets. In this work, the parameters were transferred as long as the calculated deviations for a particular molecular class (not an individual molecule due to the possibility of having accidental errors in

**TABLE 2: Comparison of Charge Bond Increments (in Electrons) Calculated Using ESP and CESP Methods**

| $\delta_{ij}$ | ESP | CESP |
|---|---|---|
| Methanol | | |
| o2−c4 | −0.226 | −0.206 |
| o2−h1 | −0.371 | −0.420[a] |
| c4−h1 | −0.018 | −0.053[a] |
| RMS (kcal/mol) | 0.444 | 0.595 |
| Phenol | | |
| o2−c3a | −0.039 | −0.052 |
| o2−h1 | −0.365 | −0.420[a] |
| c3a−h1 | −0.129 | −0.127[a] |
| c3a−c3a | 0.005 | 0.000[a] |
| RMS (kcal/mol) | 0.515 | 0.656 |
| Methylamine | | |
| n3−h1 | −0.333 | −0.355 |
| n3−c4 | −0.253 | −0.185 |
| c4−h1 | 0.042 | 0.053[a] |
| RMS (kcal/mol) | 0.431 | 0.462 |
| Phenylamine | | |
| n3−h1 | −0.357 | −0.355 |
| n3−c3a | −0.013 | −0.010 |
| c3a−h1 | −0.118 | −0.127[a] |
| c3a−c3a | 0.005 | 0.000[a] |
| RMS (kcal/mol) | 0.578 | 0.604 |

**TABLE 3: Comparison of Total Energies and Densities Obtained Using 100 ps NPT Simulations with Different Cut-Off Values[a]**

| cutoff | energy (kcal/mol) | density (g/cm$^3$) | CPU[b] |
|---|---|---|---|
| Liquid Propane ($N_{atom.} = 1320$) | | | |
| 6.5 | −292.5 ± 19.5 | 0.567 ± 0.008 | 0.131 |
| 8.5 | −296.4 ± 19.9 | 0.572 ± 0.009 | 0.200 |
| 10.5 | −298.4 ± 19.2 | 0.572 ± 0.008 | 0.331 |
| 12.5 | −302.2 ± 22.8 | 0.574 ± 0.009 | 0.506 |
| 14.5 | −301.6 ± 20.6 | 0.573 ± 0.009 | 0.759 |
| Ewald | −301.3 ± 18.3 | 0.573 ± 0.007 | 7.566 |
| Liquid Ethanol ($N_{mol} = 900$) | | | |
| 6.5 | −1156.8 ± 26.8 | 0.734 ± 0.015 | 0.088 |
| 8.5 | −1214.0 ± 26.0 | 0.776 ± 0.015 | 0.142 |
| 10.5 | −1210.4 ± 23.9 | 0.785 ± 0.014 | 0.229 |
| 12.5 | −1215.6 ± 23.6 | 0.785 ± 0.011 | 0.352 |
| 14.5 | −1214.8 ± 23.9 | 0.782 ± 0.013 | 0.524 |
| Ewald | −1207.7 ± 24.3 | 0.780 ± 0.013 | 4.379 |
| Crystal Hexane ($N_{mol} = 1500$) | | | |
| 8.5 | −1225.5 ± 13.7 | 0.883 ± 0.006 | 0.299 |
| 10.5 | −1252.5 ± 13.0 | 0.885 ± 0.006 | 0.492 |
| 12.5 | −1252.3 ± 13.7 | 0.885 ± 0.006 | 0.788 |
| 14.5 | −1256.9 ± 13.6 | 0.886 ± 0.006 | 1.208 |
| 20 | −1253.9 ± 13.1 | 0.886 ± 0.005 | 10.299 |
| Crystal Acetic Acid ($N_{mol} = 1536$) | | | |
| 8.5 | −966.4 ± 14.6 | 1.288 ± 0.007 | 0.233 |
| 10.5 | −1001.7 ± 12.6 | 1.299 ± 0.007 | 0.391 |
| 12.5 | −1005.3 ± 12.2 | 1.298 ± 0.007 | 0.587 |
| 14.5 | −1024.2 ± 12.6 | 1.300 ± 0.007 | 0.901 |
| 20 | −1013.1 ± 11.7 | 1.297 ± 0.007 | 10.646 |

[a] The standard deviations are values corresponding to running average (see text). [b] CPU time on an IBM 6000 workstation, in s/step.

the reference data) were in the ranges given later in the validation section of this paper. If the errors were too large, a new atom type was introduced and subsequently linked to a set of *actual* types. The *actual* types may be new or transferred. For example, to fit the liquid data of alcohols and ethers, a new atom type (c4o) is used to replace the normal alkyl carbon (c4) for the α carbon. However, this change only affects nonbond parameters (see Table 1).

**Constrained ESP Charges.** The same parametrization path as that illustrated in Figure 2 was followed to derive the charge parameters. This requires use of a constrained ESP fitting scheme in which all previously determined charge parameters are fixed and only the 'new' parameters are relaxed during the fit. The restriction inevitably causes a poorer fit to the ESP surfaces than the unconstrained fit. In Table 2, two sets of calculated charge bond increments are compared between the 'free' (ESP) fit and the constrained (CESP) fit. Four molecules—methanol, phenol, methylamine, and phenylamine—are presented as examples to illustrate the point. In the CESP fit, parameters that are fixed during the fit are labeled. The quality of fit can be measured by the root mean square displacement (rms) calculated between the ab initio and fitted electrostatic potential energies (in kcal/mol). The values obtained are given in the table for each set of data. Although the absolute values of the charge parameters can be quite different between the CESP and ESP fittings, the quality of the overall fit is not very different between them. The rms values of the CESP fit are slightly greater than those obtained from the ESP fit.

It should be noted that another restriction was implicitly applied to both the ESP and CESP fits in this work. Using the charge bond increment instead of atomic point charges as the adjustable parameters, an 'equivalence constraint' was enforced. Depending on the definition of atom types, all bonds that have the same atom types are treated with one parameter. For example, all C−H bonds in alkanes or alkyl groups share one bond increment $\delta_{c4,h1}$, although the 'real' charge distributions on these hydrogen atoms may be slightly different.

In most cases, the ab initio charge parameters were directly used so that only the LJ-9-6 parameters were adjusted to fit the experimental data using the MD simulations. However, this led to difficulties in fitting data for highly polarizable molecules

such as alcohols and acids. It was found that the ab initio charge bond increment $\delta_{OH}$ had to be increased in order to obtain an good overall fit of the liquid data for these molecules. Most likely, this is due to a strong polarization effect produced by hydrogen bonding. This led to fitting the charge parameters with $\delta_{OH} = -0.42$ being fixed, as given in Table 2.

**Simulation Conditions.** The MD simulations were carried out using a Verlet velocity integrator. Two types of ensembles—constant volume (NVT) and constant pressure (NPT)—were used. In the NPT simulations of liquids, pressures were controlled using a method developed by Berendsen et al.[43] For crystals, the Parrinello−Rahman[44] method was also used so that the effect of cell deformation could be studied. In all MD simulations, temperatures were controlled using the stochastic collision method proposed by Andrea et al.[45] The time step was normally 1.0 fs. Testing of smaller time steps was conducted, and no significant difference was found in the properties of interest. The preequilibration took about 50−100 ps, which is usually adequate for liquids of small molecules. The average periods were 50 ps for NVT simulations and 100 ps for NPT simulations. The cell dimensions of the MD simulations are in the range of 20−30 Å, consisting of 1000−2000 atoms. Testing on a larger cell did not show any detectable differences in the properties measured. For crystals, the simulation cells are super-cells consisting of a number of unit cells so that the cell dimensions are in the range of 20−30 Å.

In computer simulations of liquids, it is a common practice to truncate the atom−atom pairwise interactions at a selected separation (normally around 10 Å) and neglect long-range interactions. To parametrize the nonbond interaction with the cutoff scheme would associate the parameters with the simulation conditions and consequently lose the generality of the force field. Therefore, an accurate evaluation of the nonbond interactions including long-range contributions is required.

COMPASS—Details on Alkane and Benzene Compounds

*J. Phys. Chem. B, Vol. 102, No. 38, 1998* **7345**

Several general methods have been developed to treat the long-range forces in light of the concept of a periodic boundary condition. The Ewald summation[47,48] was originally formulated for ionic crystals. It has been widely used by many investigators for simulations of molecular liquids and crystals. The cell multipole expansion method, which was proposed recently,[49−51] appears to be more efficient than the Ewald summation in some applications. For nonpolar liquids, a method using a finite cutoff with long-range tail correction[52] works well. Generally speaking, the tail-correction method is much less time consuming than either the Ewald summation or cell multipole methods. It was not clear if this method could be used accurately for polar molecules in liquids and for polar and nonpolar molecules in crystals. Since the efficiency of using MD simulations to parametrize the nonbond parameters was a critical issue, the possibility of using a cutoff method in this project was investigated.

For the vdW interactions, as given in text books[52] a system with $n$ different nonbond atom types interacting with pair potential $E_{ij}(r)$, the long-range corrections to the total energy and pressure are given by

$$E_{\text{LTC}} = \frac{1}{2}\sum_{i=1}^{n} N_i \sum_{j=1}^{n} \rho_j 4\pi \int_{r_c}^{\infty} g_{ij}(r) E_{ij}(r) r^2 \, dr$$

$$P_{\text{LTC}} = \frac{1}{6}\sum_{i=1}^{n} \rho_i \sum_{j=1}^{n} \rho_j 4\pi \int_{r_c}^{\infty} g_{ij}(r) \left( r \frac{dE_{ij}(r)}{dr} \right) r^2 \, dr \qquad (5)$$

where $g_{ij}(r)$ is the pair radial distribution function and $r_c$ is the cutoff value. For liquids, $g(r) = 1$ is a good approximation for a sufficiently large value of separation $r_c$. Hence, the contribution of a long-range vdW interaction to the total energy and pressure of liquids can be analytically evaluated using eq 5.

Although the radial distribution function $g(r)$ does not converge to unity for molecular crystals, the definition of $g(r)$ implies that at large separation, the mean value of $g(r)$ must be 1. Since the vdW function and its first derivative are very flat in the region of long separation, it is plausible to assume $g(r) = 1$ and use eq 5 to calculate the tail corrections.

The remaining issue centers on the electrostatic interactions. It is now clear that if charge-neutral groups can be defined (which is always true for neutral molecules) and the sizes of the groups are small in comparison with the cutoff value, it is not necessary to calculate the long-range electrostatic interactions in both liquids and crystals. This argument is explained as follows.

For long-range interactions between two charge-neutral groups, which consist of a number of point charges, the electrostatic interactions can be expanded in a multipole series to terms of dipole−dipole, dipole−quadrupole, etc. The leading dipole−dipole interaction

$$E_{\text{dipole}-\text{dipole}} = \frac{\vec{M}_1 \vec{M}_2 - 3(\vec{n}\vec{M}_1)(\vec{n}\vec{M}_2)}{|\vec{r}_1 - \vec{r}_2|^3} \qquad (6)$$

is the dominant force. Here, $\vec{n}$ is a unit vector in the direction of $\vec{r}_1 - \vec{r}_2$. If the molecules in the liquid state are randomly oriented, for any given molecule, the time and space average of the sum of eq 6 over all other molecules is zero. For regularly arranged dipoles in crystals, it is less obvious but has been shown[53] that the dipole−dipole interaction vanishes completely or approximately due to symmetry related cancellation. Hence, there is no need to calculate the long-range dipole−dipole electrostatic interactions for both completely disordered or regularly arranged molecular systems. Since high-order moments can be represented by a set of distributed dipoles, the above reasoning seems plausible for all electrostatic interaction terms.

To numerically verify these arguments, a comparative study was carried out with four molecular systems—liquid propane and ethanol and crystalline hexane and acetic acid. These molecules represent both polar and nonpolar molecules in both the liquid and crystalline states. Charge-neutral groups were defined for the molecules as the smallest possible fragments in these molecules. NPT simulations with different cutoff values were performed in which the tail correction was performed for the LJ-9-6 terms only while the electrostatic interactions beyond the cutoff were completely ignored. For comparison, calculations with complete Ewald summation were performed with the same simulation conditions. The average total energies and densities and their standard deviations obtained from the 100 ps trajectory for each of the simulations are given in Table 3. As shown by these data, both energy and density converge quickly at the cutoff, ranging from 8.5 to 12.5 Å. The polar molecules need slightly larger cutoff values than the nonpolar molecules, as expected due to the electrostatic interactions. There is no significant difference in the convergence between the liquid and crystal simulations. In the last column, the CPU times (s/step, on an SGI-R10K workstation) for each of the simulations is given. On average, the calculations with Ewald summation are about 20 times slower than those with a 10.5 Å cutoff.

For any given system, it is a matter of choosing a proper cutoff value. The cutoff separates the integration space into two regions. In the short-range region (less than 10−15 Å), the multipole expansion of the electrostatic interaction is not valid nor is the assumption of $g(r) = 1$. Hence, all interactions, electrostatic and vdW, have to be explicitly summed. In the large-separation region that ranges from the cutoff to infinity, the vdW interaction can be corrected using eq 5; the electrostatic interactions approximately cancel. The actual cutoff value is influenced by several factors. The most important one is the size of the charge-neutral group. The strength of the interaction between molecules and the polarizability of the molecule is also relevant. A quick check of the radial-distribution function normally provides a good hint as to where the cutoff should be set. For the molecules studied here, the charge groups defined were normally smaller than 5 Å and the selected cutoff values were in the range 8.5−12.5 Å.

The standard deviations given in Table 3 correspond to the means calculated from 100 000 samples in the simulations. Since the successive measurements in the simulation are not independent, in general, these values do not represent the true statistical variances in the means. To obtained a quantitative estimate of the magnitude of the resulting uncertainties, analyses of the time-correlation functions of fluctuation $<\Delta A(0)\cdot\Delta A(0 + t)>$ on several sample liquids were performed. For example, the results obtained from liquid propanol, which represents an averaged case of the systems studied, are given in Figures 3−5. The time-correlation functions of fluctuation in pressure, potential energy (both from a 50 ps NVT simulations), and density (from a 100 ps NPT simulation) are plotted in these figures, respectively. The time-correlation function of the pressure fluctuation obtained from the NVT simulation (Figure 3) shows a pattern of two superimposed 'frequencies', a fast one that oscillates with a period of approximately 10 fs and a slow damping factor that reaches convergence (zero) around 250 fs. If we assume that the decay time of the correlation is where the slow damping profile first reaches the $x$ axis, in this
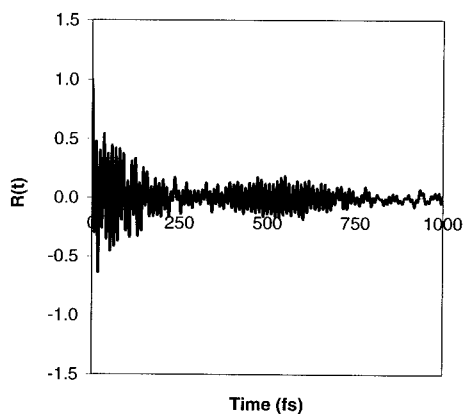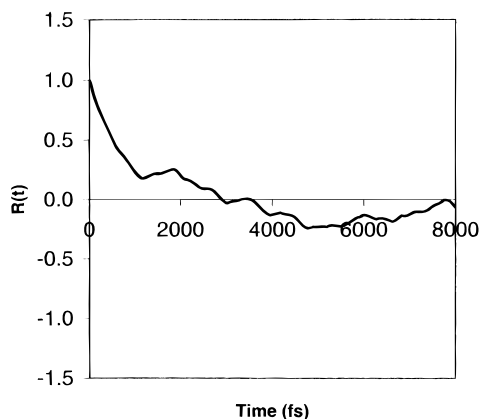
**Figure 3.** Time-correlation function of pressure fluctuation, $R(t) = \langle \Delta P(0) \Delta P(0 + t) \rangle$, in NVT simulation of liquid propanol at 293.2 K and experimental density 0.804 g/cm$^{-3}$.
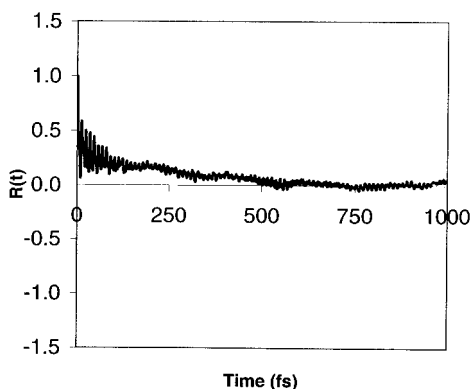


**Figure 4.** Time-correlation function of the total potential-energy fluctuation, $R(t) = \langle \Delta E_p(0) \Delta E_p(0 + t) \rangle$, in NVT simulation of liquid propanol at 293.2 K and experimental density 0.804 g/cm$^{-3}$.



**Figure 5.** Time-correlation function of density fluctuation, $R(t) = \langle \Delta D(0) \Delta D(0 + t) \rangle$, in NPT simulation of liquid propanol at 293.2 K and 1 atm pressure.

case, the correlation time is estimated to be 250 fs. On the basis of the same trajectory, the time correlation function of fluctuation in the total potential energy (Figure 4) shows a similar pattern: a fast oscillation with a similar frequency as that found in the pressure plot (Figure 3) and an outline gradually converges to zero at ca. 500 fs. These results suggest that a 50 ps NVT simulation may provide approximately 200 statistically independent measurements for the pressures and 100 statistically independent measurements for the potential energies. In Figure 5, the time-correlation function of the density fluctuation is given. This is obtained from a 100 ps NPT simulation of liquid propanol. Apparently, the density estimated in the NPT

simulation converges much slower than the pressure and energy in the NVT simulations. Using the same criteria stated above, the decay time of the correlation in density measurement is approximately 3000 fs. Hence, a 100 ps NPT simulation provides about 33 statistically independent measurements for densities.

Provided that the decay time $\tau_A$ is known, the true variance of the independent measurements can be estimated using the following formula[52]

$$\sigma = \sqrt{\frac{\tau_A}{\tau_{run}}} \sigma_{run} \tag{7}$$

where $\tau_{run}$ is the total run time (50 000 fs for NVT and 100 000 fs for NPT, in this work) and $\sigma_{run}$ is the variance obtained in the calculation. Using the estimated correlation time, $\tau_A = 500$ fs for the energies and $\tau_A = 3000$ fs for the densities, the true uncertainties estimated for the data given in Table 3 are in the range of 0.84–1.91 kcal/mol for the energies and 0.001–0.003 g/cm$^3$ for the densities. The same treatment is used later in this paper for the liquid simulation data.

## IV. Overview of Validation Results

A total of 28 molecular classes have been parametrized and validated so far. In addition to the most common organic molecules, efforts focused on common polymer materials and small gas molecules that are often required for diffusion studies. Calculations were performed on 178 molecules in isolation, 102 molecular liquids, and 69 molecular crystals, representing the 28 molecular classes. In this section, highlights of the validation results are presented.

**Intramolecular Properties.** Intramolecular properties were performed on isolated molecules as given in Appendix A. It should be noted that these molecules are not the ab initio parametrization training sets which are given in separate publications.[7,16,24–26] The criterion for selecting the validation molecules was primarily based on the availability of gas-phase experimental data. For molecules whose experimental data were not available in the literature, high-level ab initio calculations were performed and the calculated results were used for the validation. For this reason, the combined experimental and theoretical data are referred to 'reference' data.

In Figures 6 and 7, charts of correlations between the calculated and reference data for bond lengths and angles are presented. There is a total of 1296 data points for the bond lengths and 931 for bond angles. These data points represent symmetrically independent internal coordinates whose reference data are available for comparison. Detailed comparisons for each of the molecules are published elsewhere.[24–28] As shown in these figures, excellent agreement between the calculated and the reference data is obtained. For the bond lengths, the data-set ranges from ca. 0.7 to 2.4 Å. Overall, the maximum absolute percentage deviation is 4.1%, the average percentage deviation is −0.1%, and the root mean squares (rms) deviation is 0.9%. Deviations obtained for the bond angles are slightly higher than those obtained for the bond lengths. Of the data points ranging from ca. 60° to 175°, the maximum absolute deviation is 9.7%, the average percentage deviation is −0.1%, and the rms deviation is 1.8%.

A comparison of the molecular vibration frequencies is given in Figure 8. This correlation chart is based on calculations performed for 50 small molecules whose experimental data are available for comparison. Ab initio frequencies are not used in these comparisons because of the systematic errors in the ab

COMPASS—Details on Alkane and Benzene Compounds

*J. Phys. Chem. B, Vol. 102, No. 38, 1998* **7347**
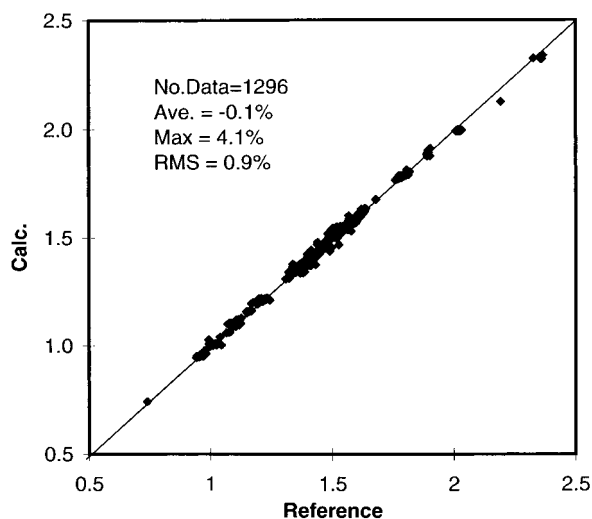


**Figure 6.** Comparison of bond lengths between the force-field-calculated results and the reference values. The total number of data points (no. of data), maximum absolute percentage deviation (max), average percentage deviation (av), and root mean squares (rms) percentage deviation are listed.



**Figure 8.** Comparison of vibrational frequencies (in $cm^{-1}$) between the force-field-calculated results and the reference values. The total number of data points (no. of data), maximum absolute deviation (max), average deviation (av), and root mean squares (rms) deviation are listed.
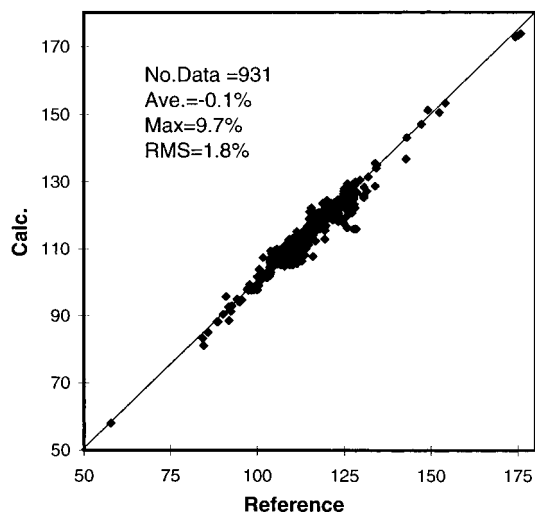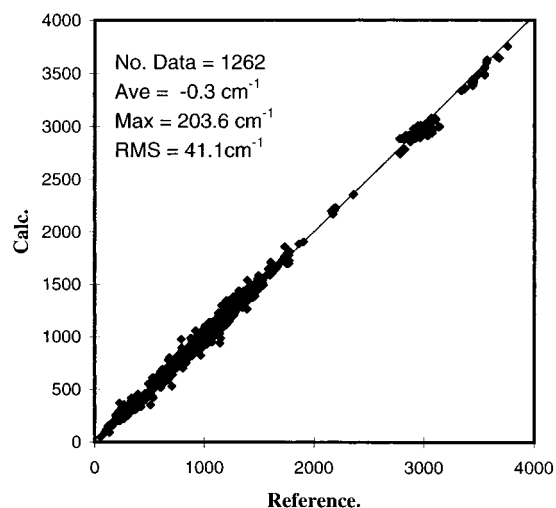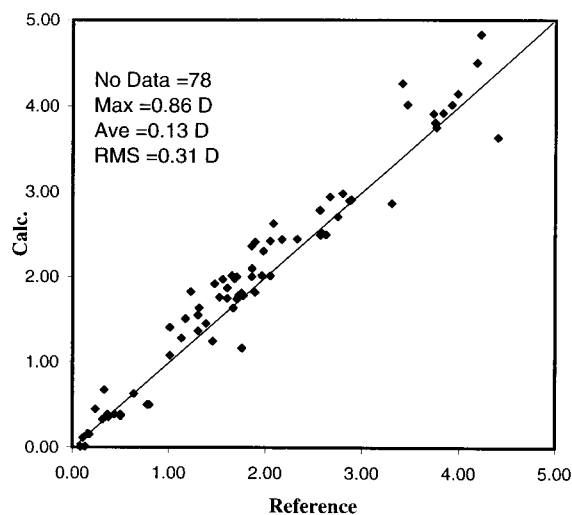


**Figure 7.** Comparison of bond angles between the force-field-calculated results and the reference values. The total number of data points (no. of data), maximum absolute percentage deviation (max), average percentage deviation (av), and root mean squares (rms) percentage deviation are listed.



**Figure 9.** Comparison of molecular dipole moments (in Debye) between the force-field-calculated results and the reference values. The total number of data points (no. of data), maximum absolute deviation (max), average deviation (av), and root mean squares (rms) deviation are listed.

initio calculations. The agreement between the calculated and experimental data is reasonably good. There are 1262 data points, the average deviation is $-0.3$ $cm^{-1}$, the rms deviation is 41.1 $cm^{-1}$, and the maximum absolute deviation is 203.6 $cm^{-1}$. The absolute maximum deviations are larger than what one could achieve (normally less than 100 $cm^{-1}$)[7,16] using the CFF approach. However, the data given here represent a broad range of molecules, which are covered with many generic parameters for angle, torsion, out-of-plane, and cross-coupling terms. Although the accuracy could be improved by using more specific parameters, this approach was not pursued due to lack of strong justification considering the number of parameters to be introduced. The current accuracy is adequate for most force-field applications of molecules in condensed phases (e.g., free-energy evaluation). For spectroscopic identification, high-level quantum mechanics calculations on small molecules in isolation appear to be the preferable choice.

A comparison of the molecular dipole moments is given in Figure 9. The reference values are measured[54] or calculated[55]

for 78 small molecules among those listed in Table 16 in the gas phase. The agreement between the calculation and reference data is reasonably good. With data ranging from 0 to 5 D, the average deviation between the calculated and reference data is 0.13 D, the rms deviation is 0.28 D, and the maximum absolute deviation is 0.86 D. As shown by the distribution of data points in the figure, the values obtained are mostly above the diagonal line, which means that the calculated values are systematically larger than those measured. This is not surprising since the charge parameters were derived from the HF/6-31G* calculations, which generally overestimate the electrostatic interactions by ca. 10–20%. In previous PCFF parametrization,[16] the calculated HF/6-31G* charges were scaled by a factor of 0.9 in order to obtain a better agreement with the experimental values. In this work, it was realized that the systematic errors in the ab initio calculations should not cause much concern. This is because the dipole moments of molecules in condensed phases are generally 10–20% larger than those of molecules in gas phases due to polarization.[8]
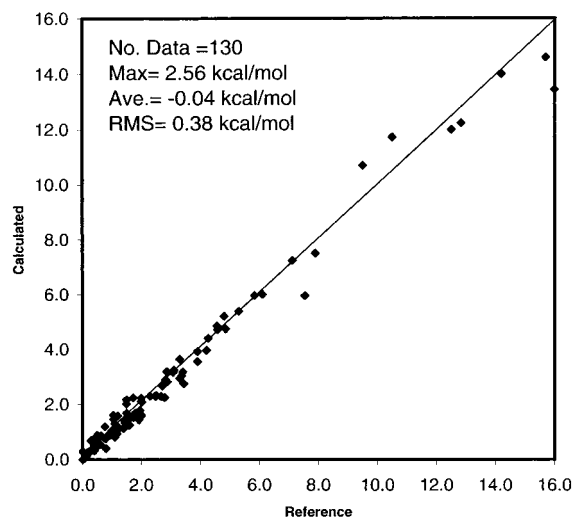
**Figure 10.** Comparison of conformational energies (in kcal/mol) between the force-field-calculated results and the reference values. The total number of data points (no. of data), maximum absolute deviation (max), average deviation (av), and root mean squares (rms) deviation are listed.

Among the intramolecular properties calculated in this work, conformational properties are mostly influenced by the nonbond parameters due to the coupling between the torsion function and the '1−4 nonbond interaction' that describes forces between any two atoms separated by three intervening valence bonds. A correlation chart of the calculated conformational energies and the reference values is given in Figure 10. There are 130 data points, representing 66 small molecules. Most of the reference values are experimental data that have been used in the parametrization of MM3,[1] CFF93,[7] and PCFF[16] force fields. Some of these data are high-level ab initio results taken from the literature or calculated in this work. Overall, the calculated results agree very well with the reference data. The statistical deviations are summarized in the figure. With the data ranging from 0 to 16.0 kcal/mol, the average deviation is −0.03 kcal/mol, the rms deviation is 0.38 kcal/mol, and the maximum deviations is 2.56 kcal/mol.

**Molecular Liquids.** NVT simulations of molecular liquids were used primarily to derive the vdW parameters in light of efficiency. In these simulations, pressures and cohesive-energy densities were measured and compared with the experimental values. The cohesive-energy density $E_{CED}$ is the total intermolecular energy, related to the heat of vaporization $\Delta H_v$ by

$$E_{CED} = \frac{\rho}{M}(\Delta H_v - RT) \qquad (8)$$

where $M$ is the molecular weight and $\rho$ is the density. Equation 8 implies two assumptions—the intramolecular energy is the same in both the liquid and gas phases and the gas obeys the ideal gas law.

The final results of the NVT simulations of liquids are summarized in Table 4. The experimental[55−64] densities, temperatures, and heat of vaporization are listed for comparison. The calculated pressures and cohesive-energy densities averaged over 50 ps simulations, together with the variances estimated using eq 7, are given in the table.

The calculated heats of vaporization and the percentage errors with respect to the experimental data are given in Table 4. For most molecules, the calculated data agree well with the experimental values within a few percent errors, although a few large errors (up to 14.6%) are obtained for which the causes

are not immediately clear. Further investigations on the quality of the experimental measurements and the simulation conditions for individual molecules are required to resolve these questions. The statistic analysis of the percent errors is summarized at the bottom of this table. For a total of 100 data points (the experimental data of the heat of vaporization are missing for two liquids), the average percent error is −0.2%, the rms percent error is 4.1%, and the maximum errors are 14.6% and −14.5%.

The average pressures obtained in the NVT simulations are within a range of ±300 bar from the experimental values that are 1 atm, 5.2 atm for $CO_2$ (at the triplet point). Since the isothermal compressibility ($k_T$) values of most molecular liquids are in the order of $10^{-4}$ (1/bar), a deviation of 300 bar in the pressure corresponds to a few percent error in the densities.

Direct comparison of the densities was made at the end of the parametrization using NPT simulations. After equilibrium, a 100 ps NPT simulation was performed to estimate the densities for each of the molecular liquids studied. The results are summarized in Table 5. The average densities and their standard deviations are listed. The variances are calculated using eq 7 with the estimated correlation time $\tau_A = 3000$ fs for the density measurement. The average densities obtained agree well with the experimental data. In the last column, the percent errors of the calculated densities with respect to the experimental data are given. Overall, for the 102 data points calculated, the average percent error is only −0.4%, a rms deviation is 1.9%, and the maximum deviations are 5.2% and −4.0%.

**Molecular Crystals.** Three types of simulations were performed for molecular crystals to validate the present force field. To compare with early parametrization work on nonbond parameters,[19−22] energy minimizations with full relaxation of all degrees of freedom were performed. More rigorous constant-pressure MD simulations were performed for direct comparison with the experimental data. Two pressure control methods, due to Berendsen et al.[43] and Parrinello and Rahman,[44] respectively, were used in the MD simulations. The first method, which preserves the cell shape, works fine for liquid simulations but the symmetry constraint may be questionable for crystal predictions. To validate the force field under the true experimental conditions, the Parrinello−Rahman method[44] that fully relaxes all cell parameters was used. All MD simulations were performed at experimental temperatures and pressures (1 atm). Some of the early measurements were reported without explicit temperature. In these cases, room temperature, $T = 298.2$ K, was assumed.

In Table 6, the average densities obtained from these three simulations and the experimental values[65−132] are given for comparison. For each of the calculations, percent errors with respect to the experimental data are calculated. The overall statistical analyses of the calculation errors are given at the bottom of this table. As expected, the densities obtained using energy minimization are systematically too large (5.9% average deviation) in comparison with the experimental values measured at finite temperature. Both MD simulations, carried out at the same conditions as the experimental measurements, yield good agreements with the experimental data, with average percent errors of −0.6% and −1.0%. This reveals the origin of the problems of early development based on the energy minimization method. If the parameters were optimized using the energy minimization method, they could not yield good results with the MD simulations under the experimental conditions.

Both the full-energy minimization and Parrinello−Rahman NPT simulation yield good agreement in cell the shape with the experimental data for most crystals. In Figures 11 and 12,

COMPASS—Details on Alkane and Benzene Compounds

J. Phys. Chem. B, Vol. 102, No. 38, 1998   7349

**TABLE 4: Comparison of Experimental and NVT-Simulated (50 ps) Results of Liquids[a]**

| molecule | exp | | | calcd[b] | | | | ref |
|---|---|---|---|---|---|---|---|---|
| | $T$ | $D$ | $H_v$ | $P$ | CED | $H_v$ | $H_v$ % | |
| cyclohexane | 293.2 | 0.779 | 7.96 | 135 ± 106 | 6.85 ± 0.07 | 7.99 | 0.4 | 56 |
| ethane | 184.5 | 0.546 | 3.51 | 193 ± 72 | 5.24 ± 0.06 | 3.25 | −7.4 | 56 |
| isopentane | 293.2 | 0.620 | 6.01 | −100 ± 81 | 4.76 ± 0.05 | 6.12 | 1.8 | 56 |
| methane | 111.0 | 0.424 | 1.96 | −147 ± 61 | 4.26 ± 0.04 | 1.83 | −6.6 | 56 |
| pentane | 293.2 | 0.626 | 6.39 | 53 ± 80 | 4.96 ± 0.06 | 6.3 | −1.4 | 56 |
| propane | 231.1 | 0.581 | 4.49 | 63 ± 78 | 5.05 ± 0.05 | 4.3 | −4.2 | 56 |
| 2-methylheptane | 303.2 | 0.690 | 9.49 | −66 ± 125 | 5.29 ± 0.07 | 9.36 | −1.4 | 57 |
| 2,5-dimethylhexane | 303.2 | 0.685 | 8.93 | −108 ± 132 | 4.93 ± 0.08 | 8.82 | −1.2 | 57 |
| 2,2-dimethylhexane | 303.2 | 0.687 | 8.93 | −61 ± 125 | 4.91 ± 0.06 | 8.77 | −1.8 | 57 |
| 2,2,4-trimethylpentane | 303.2 | 0.684 | 8.42 | −181 ± 127 | 4.57 ± 0.07 | 8.24 | −2.1 | 57 |
| butene | 266.9 | 0.628 | 5.28 | 530 ± 102 | 5.43 ± 0.09 | 5.39 | 2.1 | 55, 56 |
| ethylene | 169.4 | 0.568 | 3.23 | 391 ± 71 | 5.82 ± 0.06 | 3.21 | −0.6 | 56 |
| propene | 225.4 | 0.611 | 4.42 | 517 ± 95 | 5.75 ± 0.07 | 4.41 | −0.2 | 56 |
| benzene | 298.2 | 0.872 | 8.09 | 48 ± 112 | 8.47 ± 0.08 | 8.18 | 1.1 | 56 |
| toluene | 298.2 | 0.865 | 9.09 | −103 ± 111 | 7.97 ± 0.08 | 9.08 | −0.1 | 56 |
| ethylacethylene | 273.2 | 0.678 | 5.86 | 39 ± 104 | 6.71 ± 0.07 | 5.9 | 0.7 | 55, 56 |
| acetylene | 198.2 | 0.612 | 3.64 | −17 ± 82 | 7.48 ± 0.08 | 3.57 | −1.9 | 56 |
| methylacetylene | 248.2 | 0.673 | 5.62 | 222 ± 106 | 7.62 ± 0.10 | 5.03 | −10.5 | 55, 56 |
| dimethylacetylene | 293.2 | 0.691 | 6.44 | 296 ± 112 | 7.33 ± 0.10 | 6.32 | −1.9 | 55, 56 |
| ethanol | 293.2 | 0.789 | 10.20 | −22 ± 121 | 16.74 ± 0.26 | 10.36 | 1.6 | 56 |
| isopropyl alcohol | 293.2 | 0.785 | 10.96 | −309 ± 125 | 14.27 ± 0.19 | 11.51 | 5.0 | 56 |
| methanol | 293.2 | 0.791 | 9.01 | 310 ± 112 | 20.74 ± 0.23 | 8.98 | −0.3 | 56 |
| phenol | 323.2 | 1.050 | 13.36 | −82 ± 152 | 14.11 ± 0.21 | 13.29 | −0.5 | 56 |
| propanol | 293.2 | 0.804 | 11.44 | 113 ± 125 | 14.57 ± 0.17 | 11.48 | 0.3 | 56 |
| diethyl ether | 293.2 | 0.714 | 6.55 | 56 ± 112 | 5.89 ± 0.10 | 6.7 | 2.3 | 56 |
| dimethyl ether | 248.3 | 0.735 | 5.14 | 82 ± 109 | 7.30 ± 0.10 | 5.07 | −1.4 | 55, 56 |
| isopropyl methyl ether | 288.2 | 0.724 | 6.46 | −98 ± 114 | 5.91 ± 0.09 | 6.63 | 2.6 | 55, 56 |
| acetaldehyde | 293.2 | 0.783 | 6.16 | 44 ± 121 | 10.02 ± 0.13 | 6.22 | 1.0 | 55, 56 |
| n-butyraldehyde | 293.2 | 0.802 | 8.08 | 71 ± 127 | 8.67 ± 0.12 | 8.38 | 3.7 | 57, 118 |
| formaldehyde | 254.0 | 0.812 | 5.54 | 322 ± 117 | 13.01 ± 0.33 | 5.32 | −4.0 | 55, 56 |
| propionaldehyde | 293.2 | 0.792 | 7.16 | −41 ± 122 | 9.20 ± 0.09 | 7.33 | 2.4 | 57, 118 |
| acetone | 329.3 | 0.750 | 7.03 | 15 ± 125 | 8.01 ± 0.13 | 6.86 | −2.4 | 56 |
| methyl etheyl ketone | 352.8 | 0.743 | 7.48 | 22 ± 126 | 7.08 ± 0.11 | 7.57 | 1.2 | 55, 56 |
| acetic acid | 391.2 | 0.939 | 5.66 | 2 ± 145 | 14.46 ± 0.26 | 6.01 | 6.2 | 56 |
| propionic acid | 298.2 | 0.988 | 7.68 | −99 ± 134 | 15.98 ± 0.24 | 7.54 | −1.8 | 55, 56 |
| formic acid | 373.8 | 1.108 | 5.42 | 123 ± 138 | 20.96 ± 0.31 | 5.67 | 4.6 | 55, 56 |
| ethyl acetate | 350.3 | 0.831 | 7.63 | 41 ± 131 | 6.76 ± 0.10 | 7.86 | 3.0 | 56 |
| methyl benzoate | 298.2 | 1.083 | 13.28 | 192 ± 155 | 10.54 ± 0.11 | 13.85 | 4.3 | 55, 56 |
| methyl acetate | 330.3 | 0.885 | 7.25 | 155 ± 130 | 7.90 ± 0.13 | 7.27 | 0.3 | 56 |
| acetic anhydride | 298.2 | 1.075 | 11.38 | −174 ± 125 | 13.10 ± 0.12 | 13.04 | 14.6 | 55, 56 |
| N,N-dimethylacetamide | 298.2 | 0.937 | 12.01 | 56 ± 141 | 12.19 ± 0.14 | 11.93 | −0.7 | 55 |
| N,N-dimethylformamide | 298.2 | 0.944 | 11.21 | 21 ± 138 | 13.99 ± 0.12 | 11.43 | 2.0 | 55 |
| formamide | 298.2 | 1.107 | 14.38 | −628 ± 142 | 28.75 ± 0.25 | 12.29 | −14.5 | 55, 56 |
| N-methylformamide | 298.2 | 1.005 | 13.43 | −173 ± 139 | 20.29 ± 0.18 | 12.52 | −6.8 | 55 |
| aniline | 457.6 | 0.874 | 10.14 | −119 ± 172 | 9.43 ± 0.16 | 10.96 | 8.1 | 56 |
| methylamine | 266.8 | 0.691 | 6.12 | 197 ± 111 | 12.30 ± 0.20 | 6.06 | −1.0 | 55, 56 |
| ammonia | 239.8 | 0.682 | 5.58 | −150 ± 105 | 21.87 ± 0.34 | 5.94 | 6.5 | 56 |
| propylamine | 320.0 | 0.689 | 7.06 | 78 ± 130 | 7.60 ± 0.17 | 7.15 | 1.3 | 55, 56 |
| trimethylamine | 276.0 | 0.655 | 5.48 | −74 ± 123 | 5.53 ± 0.07 | 5.54 | 1.1 | 55, 56 |
| 1-fluorobutane | 293.2 | 1.024 | 8.32 | −116 ± 140 | 8.29 ± 0.12 | 8.36 | 0.5 | 55 |
| 1,1-difluoroethane | 264.5 | 0.979 | 5.15 | 82 ± 98 | 6.53 ± 0.10 | 4.93 | −4.3 | 55 |
| 1,1,1-trifluoroethane | 223.2 | 1.176 | 4.65 | 66 ± 90 | 5.98 ± 0.09 | 4.72 | 1.5 | 55, 56 |
| 1,1,1,2-tetrafluoroethane | 263.5 | 1.329 | | −18 ± 89 | 6.71 ± 0.08 | 5.68 | | 55 |
| difluoromethane | 260.0 | 1.099 | 4.68 | 169 ± 42 | 8.29 ± 0.12 | 4.44 | −5.1 | 55, 56 |
| 1,1,1,3,3,3-hexafluoropropane | 262.1 | 1.470 | | −141 ± 104 | 6.44 ± 0.09 | 7.18 | | 55 |
| 1-chlorobutane | 293.2 | 0.897 | 8.01 | 241 ± 107 | 7.20 ± 0.12 | 8.02 | 0.1 | 55, 56 |
| chlorobenzene | 293.2 | 1.106 | 9.79 | −129 ± 124 | 9.29 ± 0.09 | 10.04 | 2.6 | 55, 56 |
| 1-chloropropane | 293.2 | 0.890 | 6.78 | 163 ± 98 | 6.97 ± 0.07 | 6.74 | −0.6 | 55, 56 |
| 1,3-dichloropropane | 293.2 | 1.188 | 9.74 | 79 ± 96 | 10.10 ± 0.11 | 10.2 | 4.7 | 55 |
| 1,2-dichloropropane | 293.2 | 1.156 | 9.96 | −282 ± 93 | 9.75 ± 0.12 | 10.11 | 1.5 | 55 |
| 1,2,3-trichloropropane | 293.2 | 1.389 | 12.54 | −686 ± 89 | 12.75 ± 0.14 | 14.12 | 12.6 | 55 |
| acetonitrile | 293.2 | 0.783 | 7.87 | −51 ± 104 | 13.95 ± 0.14 | 7.9 | 0.4 | 55, 56 |
| butyronitrile | 293.2 | 0.790 | 9.40 | 72 ± 116 | 10.07 ± 0.11 | 9.39 | −0.1 | 55, 56 |
| propionitrile | 293.2 | 0.782 | 8.61 | 27 ± 116 | 11.20 ± 0.15 | 8.47 | −1.6 | 55, 56 |
| nitromethane | 293.2 | 1.136 | 9.15 | −66 ± 133 | 16.68 ± 0.16 | 9.54 | 4.3 | 55, 56 |
| nitrobenzene | 293.2 | 1.204 | 13.15 | 224 ± 150 | 11.76 ± 0.13 | 12.61 | −4.1 | 55, 56 |
| isotetrasilane | 293.0 | 0.792 | 8.31 | −20 ± 43 | 5.01 ± 0.08 | 8.32 | 0.1 | 59 |
| isopentasilane | 293.0 | 0.820 | 10.98 | 60 ± 87 | 5.14 ± 0.07 | 10.14 | −7.7 | 59 |
| trisilane | 293.0 | 0.739 | 7.11 | 54 ± 79 | 4.73 ± 0.07 | 6.5 | −8.6 | 59 |
| tetrasilane | 293.0 | 0.795 | 9.09 | 123 ± 92 | 5.13 ± 0.07 | 8.48 | −6.7 | 59 |
| pentasilane | 293.0 | 0.827 | 11.30 | 101 ± 84 | 5.34 ± 0.08 | 10.42 | −7.8 | 59 |
| hexasilane | 293.0 | 0.847 | 13.12 | 10 ± 78 | 5.45 ± 0.07 | 12.33 | −6.0 | 59 |

**TABLE 4: (Continued)**

| molecule | exp | | | calcd[b] | | | | ref |
| | T | D | $H_v$ | P | CED | $H_v$ | $H_v$ % | |
|---|---|---|---|---|---|---|---|---|
| hexamethyldisiloxane | 298.2 | 0.759 | 9.13 | 273 ± 88 | 3.99 ± 0.07 | 9.14 | 0.1 | 60 |
| PDMS (n = 3) | 298.2 | 0.851 | 14.89 | 37 ± 93 | 3.95 ± 0.05 | 15.03 | 0.9 | 61, 62 |
| disiloxane | 193.2 | 0.881 | 5.84 | 189 ± 74 | 6.13 ± 0.06 | 5.83 | −0.2 | 59 |
| 1,3-disilyldisiloxane | 273.2 | 0.876 | 9.20 | −208 ± 80 | 5.68 ± 0.08 | 9.53 | 3.6 | 63 |
| argon | 87.3 | 1.393 | 1.54 | −3 ± 5 | 4.73 ± 0.03 | 1.53 | −0.6 | 56 |
| helium | 4.2 | 0.125 | 0.02 | 3 ± 0 | 0.04 ± 0.00 | 0.02 | 0.0 | 56 |
| krypton | 59.8 | 2.155 | 2.17 | 1 ± 4 | 4.90 ± 0.03 | 2.14 | −1.4 | 56 |
| neon | 27.1 | 1.205 | 0.41 | 26 ± 2 | 2.12 ± 0.02 | 0.41 | 0.0 | 56 |
| xenon | 165.0 | 3.520 | 3.02 | −25 ± 6 | 7.23 ± 0.04 | 3.02 | 0.0 | 56 |
| carbon monoxide (CO) | 81.7 | 0.789 | 1.44 | 28 ± 46 | 3.59 ± 0.03 | 1.44 | 0.0 | 56 |
| carbon dioxide (CO$_2$) | 216.6 | 1.179 | 3.65 | −32 ± 103 | 8.69 ± 0.10 | 3.67 | 0.5 | 56 |
| hydrogen (H$_2$) | 20.4 | 0.071 | 0.22 | 5 ± 9 | 0.65 ± 0.01 | 0.22 | 0.0 | 56 |
| nitrogen (N$_2$) | 77.3 | 0.807 | 1.32 | −1 ± 45 | 3.41 ± 0.03 | 1.34 | 1.5 | 56 |
| nitric oxide (NO) | 123.0 | 1.269 | 3.31 | −4 ± 67 | 12.32 ± 0.07 | 3.16 | −4.5 | 56 |
| nitric dioxide (NO$_2$) | 293.2 | 1.446 | 9.11 | −44 ± 107 | 26.42 ± 0.15 | 8.99 | −1.3 | 56 |
| oxygen (O$_2$) | 90.2 | 1.136 | 1.62 | 22 ± 51 | 5.11 ± 0.04 | 1.62 | 0.0 | 56 |
| sulfur dioxide (SO$_2$) | 223.2 | 1.557 | 6.51 | −60 ± 84 | 14.65 ± 0.10 | 6.47 | −0.6 | 56 |
| ethanethiol | 298.2 | 0.833 | 6.58 | −12 ± 97 | 7.85 ± 0.10 | 6.45 | −2.0 | 55, 56 |
| diethyl sulfide | 298.2 | 0.831 | 8.55 | −16 ± 107 | 7.47 ± 0.09 | 8.7 | 1.8 | 55, 56 |
| methanethiol | 279.2 | 0.888 | 5.87 | 285 ± 89 | 9.37 ± 0.12 | 5.63 | −4.1 | 55, 56 |
| dimethyl sulfide | 298.2 | 0.842 | 6.65 | 117 ± 100 | 8.22 ± 0.10 | 6.66 | 0.2 | 55, 56 |
| methylethe yl sulfide | 298.2 | 0.837 | 7.61 | 19 ± 115 | 7.79 ± 0.12 | 7.68 | 0.9 | 55, 56 |
| (NPF$_2$)$_5$ | 393.3 | 1.826 | 9.80 | 612 ± 96 | 3.50 ± 0.06 | 9.64 | −1.6 | 64 |
| (NPF$_2$)$_7$ | 443.9 | 1.850 | 11.60 | 354 ± 112 | 3.12 ± 0.06 | 11.92 | 2.8 | 64 |
| (NPF$_2$)$_8$ | 466.0 | 1.857 | 12.00 | 234 ± 115 | 2.77 ± 0.05 | 12.53 | 4.4 | 64 |
| (NPF$_2$)$_9$ | 487.6 | 1.859 | 12.70 | 87 ± 104 | 2.41 ± 0.05 | 12.77 | 0.6 | 64 |
| (NPF$_2$)$_{10}$ | 504.0 | 1.864 | 13.50 | 195 ± 109 | 2.58 ± 0.06 | 13.59 | 0.7 | 64 |
| N$_3$P$_3$C$_{14}$F$_2$ | 454.8 | 1.785 | 11.00 | 476 ± 93 | 4.78 ± 0.07 | 10.93 | −0.6 | 64 |
| N$_4$P$_4$C$_{14}$F$_4$ | 478.2 | 1.834 | 12.20 | 265 ± 80 | 4.46 ± 0.07 | 12.23 | 0.2 | 64 |
| N$_4$P$_4$C$_{15}$F$_3$ | 505.2 | 1.842 | 13.00 | 220 ± 42 | 4.63 ± 0.07 | 13.25 | 1.9 | 64 |
| no. of data | | | | | | 100 | | |
| max dev | | | | | | 14.6 | | |
| min dev | | | | | | −14.5 | | |
| av dev | | | | | | −0.2 | | |
| rms dev | | | | | | 4.1 | | |

[a] Average pressures and cohesive-energy densities are listed. The uncertainties are calculated using eq 6 with estimated correlation time of the measurement (see text). [b] T = temperature in kelvin. $H_v$ = heat of vaporization in kcal/mol. D = density in g/cm³. P = pressure in bar. CED = cohesive-energy density in $10^7$ cal/m³.

the percent errors in the cell-edge parameters (*a*, *b*, *c*) and angles ($\alpha$, $\beta$, $\gamma$) are plotted, respectively. In each of these figures, two histogram curves, denoting the distribution of percent errors obtained from the energy minimizations and NPT simulations, respectively, are presented for comparison. Each of the data points in the figure represents the total number of data that fall into the corresponding 'bin' with a range of ±1.0% for cell-edge parameters and ±0.5% for the angles.

As shown in Figure 11, most cell-edge parameters obtained are within ±5% error in both the energy minimizations and MD simulations. With the energy minimization, however, the peak of the distribution curve is slightly shifted to the negative side. On average, the deviation is −1.7%, with a rms deviation of 3.7%. This is consistent with the observation of a −5.9% average deviation in the densities obtained with these calculations. The distribution of Parrinello−Rahman NPT simulations, on the other hand, has its peak located at the center. Statistical analysis yields an average error of 0.7%, with a rms deviation of 5.3%.

Distribution curves of the percent errors (Figure 11) in the angle parameters are much narrower than those found for the cell-edge parameters. It is of interest to point out that both energy minimizations and MD simulations yield very similar error distribution curves. Most data points are within ±1% error with the experimental data. Statistical analysis reveals rms deviations are 1.5% for the energy minimization and 2.5% for the MD simulation.

The lattice energies reported in the literature are derived from the observed enthalpy of sublimation by[22]

$$E_L = -\Delta H - 2RT \tag{9}$$

where the last term represents an approximate correction for the difference between gas-phase enthalpy, $PV + 3RT$ (ideal gas), and the estimated vibrational contribution of $6RT$. Due to the approximation embedded in eq 9 and the experimental uncertainty, the error bar was estimated to be in the range of 3−4 kcal/mol.[22] Clearly, the lattice energy given in eq 9 corresponds to an idealized potential energy at zero temperature. Therefore, a reasonable validation is to compare the calculated lattice energies using the energy minimization method against the 'experimental' data. Calculations were performed for a number of crystals whose lattice energies were found in the literature.[20−21,133−136] The comparisons are given in Table 7, which shows excellent agreement between the calculations and the measurements. Most absolute differences between the calculated and experimental values are less than 2 kcal/mol. The maximum error, which is obtained for a rather complicated molecule, diketopiperazine, is only 4.5 kcal/mol.

**V. Parametrization and Validation of Alkane and Benzene Compounds**

Most of the valence parameters of the alkyl and phenyl groups are transferred from the PCFF force field, which can be further

COMPASS—Details on Alkane and Benzene Compounds

*J. Phys. Chem. B, Vol. 102, No. 38, 1998* **7351**

**TABLE 5: Comparison of NPT-Simulated (100 ps) Liquid Densities and Experimental Values**[a]

| | exp | | calcd | | | exp | | calcd | |
|---|---|---|---|---|---|---|---|---|---|
| molecule | T | D | D | %D | molecule | T | D | D | %D |
| cyclohexane | 29.3.2 | 0.779 | 0.774 ± 0.002 | −0.6 | 1,1,1,3,3,3-hexafluoropropane | 262.1 | 1.470 | 1.515 ± 0.004 | 3.1 |
| ethane | 184.5 | 0.546 | 0.522 ± 0.002 | −4.4 | 1-chlorobutane | 293.2 | 0.897 | 0.870 ± 0.004 | −3.0 |
| isopentane | 293.2 | 0.620 | 0.633 ± 0.002 | 2.1 | chlorobenzene | 293.2 | 1.106 | 1.115 ± 0.003 | 0.8 |
| methane | 111.0 | 0.424 | 0.446 ± 0.001 | 5.2 | 1-chloropropane | 293.2 | 0.890 | 0.864 ± 0.003 | −2.9 |
| pentane | 293.2 | 0.626 | 0.619 ± 0.002 | −1.1 | 1,3-dichloropropane | 293.2 | 1.188 | 1.182 ± 0.002 | −0.5 |
| propane | 231.1 | 0.581 | 0.570 ± 0.002 | −1.9 | 1,2-dichloropropane | 293.2 | 1.156 | 1.180 ± .003 | 2.1 |
| 2-methylheptane | 303.2 | 0.690 | 0.702 ± 0.002 | 1.7 | 1,2,3-trichloropropane | 293.2 | 1.389 | 1.442 ± .002 | 3.8 |
| 2,5-dimethylhexane | 303.2 | 0.685 | 0.693 ± 0.002 | 1.2 | acetonitrile | 293.2 | 0.783 | 0.786 ± 0.002 | 0.4 |
| 2,2-dimethylhexane | 303.2 | 0.687 | 0.697 ± 0.003 | 1.5 | butyronitrile | 293.2 | 0.790 | 0.783 ± 0.002 | −0.9 |
| 2,2,4-trimethylpentane | 303.2 | 0.684 | 0.701 ± 0.002 | 2.5 | propionitrile | 293.2 | 0.782 | 0.776 ± 0.002 | −0.8 |
| butene | 266.9 | 0.628 | 0.627 ± 0.002 | −0.2 | nitromethane | 293.2 | 1.136 | 1.130 ± 0.003 | −0.5 |
| ethylene | 169.4 | 0.568 | 0.571 ± 0.002 | 0.5 | nitrobenzene | 293.2 | 1.204 | 1.173 ± 0.003 | −2.6 |
| propene | 225.4 | 0.611 | 0.609 ± 0.002 | −0.3 | isotetrasilane | 293.2 | 0.792 | 0.800 ± 0.003 | 1.0 |
| benzene | 298.2 | 0.872 | 0.862 ± 0.002 | −1.1 | isopentasilane | 293.2 | 0.820 | 0.825 ± 0.003 | 0.6 |
| toluene | 298.2 | 0.865 | 0.855 ± 0.002 | −1.2 | trisilane | 293.2 | 0.739 | 0.727 ± 0.003 | −1.6 |
| ethylacethylene | 273.2 | 0.678 | 0.666 ± 0.003 | −1.8 | tetrasilane | 293.2 | 0.795 | 0.795 ± 0.003 | 0.0 |
| acetylene | 198.2 | 0.612 | 0.612 ± 0.002 | 0.0 | pentasilane | 293.2 | 0.827 | 0.825 ± 0.002 | −0.2 |
| methylacetylene | 248.2 | 0.673 | 0.648 ± 0.002 | −3.7 | hexasilane | 293.2 | 0.847 | 0.851 ± 0.002 | 0.5 |
| dimethylacetylene | 293.2 | 0.691 | 0.664 ± .002 | −3.9 | hexamethyldisiloxane | 298.2 | 0.759 | 0.743 ± 0.006 | −2.1 |
| ethanol | 293.2 | 0.789 | 0.783 ± 0.003 | −0.8 | PDMS (n = 3) | 298.2 | 0.851 | 0.862 ± 0.003 | 1.3 |
| isopropyl alcohol | 293.2 | 0.785 | 0.809 ± 0.002 | 3.1 | disiloxane | 193.2 | 0.881 | 0.880 ± 0.002 | −0.1 |
| methanol | 293.2 | 0.791 | 0.761 ± 0.002 | −3.8 | 1,3-disilyldisiloxane | 273.2 | 0.876 | 0.863 ± 0.002 | −1.5 |
| phenol | 323.2 | 1.050 | 1.039 ± .003 | −1.0 | argon | 87.3 | 1.393 | 1.399 ± 0.003 | 0.4 |
| propanol | 293.2 | 0.804 | 0.794 ± .003 | −1.2 | helium | 4.2 | 0.125 | 0.121 ± 0.000 | −3.2 |
| diethyl ether | 293.2 | 0.714 | 0.706 ± 0.003 | −1.1 | krypton | 119.8 | 2.155 | 2.170 ± 0.004 | 0.7 |
| dimethyl ether | 248.3 | 0.735 | 0.714 ± 0.003 | −2.9 | neon | 27.1 | 1.205 | 1.199 ± 0.002 | −0.5 |
| isopropyl methyl ether | 288.2 | 0.724 | 0.721 ± 0.003 | −0.4 | xenon | 165.0 | 3.520 | 3.536 ± 0.004 | 0.5 |
| acetaldehyde | 293.2 | 0.783 | 0.769 ± 0.004 | −1.8 | carbon monoxide (CO) | 81.7 | 0.789 | 0.786 ± 0.002 | −0.4 |
| n-butyraldehyde | 293.2 | 0.802 | 0.795 ± 0.004 | −0.9 | carbon dioxide (CO₂) | 216.6 | 1.179 | 1.167 ± 0.003 | −1.0 |
| formaldehyde | 254.0 | 0.812 | 0.836 ± 0.003 | 3.0 | hydrogen (H₂) | 20.4 | 0.071 | 0.069 ± 0.000 | −2.8 |
| propionaldehyde | 293.2 | 0.792 | 0.782 ± 0.002 | −1.3 | nitrogen (N₂) | 77.3 | 0.807 | 0.804 ± 0.002 | −0.4 |
| acetone | 329.3 | 0.750 | 0.745 ± 0.002 | −0.7 | nitric oxide (NO) | 123.0 | 1.269 | 1.269 ± 0.002 | 0.0 |
| methyl etheyl ketone | 352.8 | 0.743 | 0.733 ± 0.003 | −1.3 | nitric dioxide (NO₂) | 293.2 | 1.446 | 1.446 ± 0.002 | 0.0 |
| acetic acid | 391.2 | 0.939 | 0.941 ± 0.005 | 0.2 | oxygen (O₂) | 90.2 | 1.136 | 1.122 ± 0.002 | −1.2 |
| propionic acid | 298.2 | 0.988 | 0.994 ± .004 | 0.6 | sulfur dioxide (SO₂) | 223.2 | 1.557 | 1.577 ± 0.003 | 1.3 |
| formic acid | 373.8 | 1.108 | 1.080 ± 0.006 | −3.4 | ethanethiol | 298.2 | 0.833 | 0.829 ± 0.002 | −0.5 |
| ethyl acetate | 350.3 | 0.831 | 0.823 ± .003 | −1.0 | diethyl sulfide | 298.2 | 0.831 | 0.833 ± 0.003 | 0.2 |
| methyl benzoate | 298.2 | 1.083 | 1.066 ± .004 | −1.6 | methanethiol | 279.2 | 0.888 | 0.856 ± 0.002 | −3.6 |
| methyl acetate | 303.3 | 0.885 | 0.867 ± 0.003 | −2.0 | dimethyl sulfide | 298.2 | 0.842 | 0.835 ± 0.002 | −0.8 |
| acetic anhydride | 298.2 | 1.075 | 1.091 ± 0.002 | 1.5 | methyletheyl sulfide | 298.2 | 0.837 | 0.836 ± 0.003 | −0.1 |
| N,N-dimethylacetamide | 298.2 | 0.937 | 0.912 ± 0.003 | −2.7 | (NPF₂)₅ | 293.2 | 1.826 | 1.806 ± 0.006 | −1.1 |
| N,N-dimethylformamide | 298.2 | 0.944 | 0.932 ± 0.002 | −1.3 | (NPF₂)₇ | 293.2 | 1.850 | 1.862 ± 0.006 | 0.6 |
| formamide | 298.2 | 1.107 | 1.155 ± .002 | 4.3 | (NPF₂)₈ | 293.2 | 1.857 | 1.869 ± 0.004 | 0.7 |
| N-methylformamide | 298.2 | 1.005 | 1.003 ± .002 | −0.2 | (NPF₂)₉ | 293.2 | 1.859 | 1.867 ± 0.009 | 0.4 |
| aniline | 457.6 | 0.874 | 0.881 ± 0.004 | 0.8 | (NPF₂)₁₀ | 293.2 | 1.864 | 1.865 ± 0.003 | 0.1 |
| methylamine | 266.8 | 0.691 | 0.668 ± 0.002 | −3.3 | N₃P₃C₁₄F₂ | 293.2 | 1.785 | 1.773 ± 0.004 | −0.7 |
| ammonia | 239.8 | 0.682 | 0.691 ± .002 | 1.3 | N₄P₄C₁₄F₄ | 293.2 | 1.834 | 1.853 ± 0.005 | 1.0 |
| propylamine | 320.0 | 0.689 | 0.681 ± 0.002 | −1.2 | N₄P₄C₁₅F₃ | 293.2 | 1.842 | 1.830 ± 0.003 | −0.7 |
| trimethylamine | 276.0 | 0.655 | 0.660 ± .003 | 0.8 | | | | | |
| 1-fluorobutane | 293.2 | 1.024 | 1.018 ± 0.002 | −0.6 | no. of data | | | | 102 |
| 1,1-difluoroethane | 264.5 | 0.979 | 0.988 ± 0.005 | 0.9 | max dev | | | | 5.2 |
| 1,1,1-trifluoroethane | 223.2 | 1.176 | 1.144 ± 0.005 | −2.7 | min dev | | | | −4.4 |
| 1,1,1,2-tetrafluoroethane | 263.5 | 1.329 | 1.344 ± 0.005 | 1.1 | av dev | | | | −0.4 |
| difluoromethane | 260.0 | 1.099 | 1.055 ± 0.003 | −4.0 | rms dev | | | | 1.8 |

[a] The uncertainties of calculations are calculated using eq 6 with estimated correlation time of the measurement (see text). [b] References are given in Table 4. $T$ = temperature in kelvin. $D$ = density in g/cm³. $P$ = pressure in bar.

traced back to the CFF91 force field. Originally, the valence parameters were derived from the HF/6-31G* data using the CFF parametrization technique[7,18] and the vdW and charge parameters were derived based on energy minimization calculations and fitting of the calculated lattice energies and structures to experimental data.[19] Parametrization and validation of the CFF93 force field for alkanes, which was developed based on the CFF91 force field, has been published.[7a] Due to later modifications, the published CFF93 parameters for alkanes are slightly different from those used in the CFF91. However, validation performed at the beginning of this study indicates that very similar results to those reported[7a] can be obtained using

the PCFF parameters for alkanes and benzenes. In addition, the charge parameters used in the PCFF (the same as those in the CFF93) are in good agreement with the CESP charges. Hence, modification of the nonbond vdW parameters was the major concern for the alkyl and phenyl groups, although a few valence parameters had to be modified due to the coupling with the vdW parameters. The final parameters are given in Appendix B.

The most significant differences between the PCFF(CFF9X) and COMPASS parameters for alkanes and benzenes are in the vdW LJ-9-6 parameters, which are listed in Table 8. The new vdW radii ($r_o$) are smaller and the well depths ($\epsilon_o$) are larger

**TABLE 6: Comparison of Experimental and Calculated Crystal Densities (g/cm³)[a]**

| molecules | exp | | calcd | | | | | | ref |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | $D$ | $T$ | MM | %MM | MD(V) | %MD(B) | MD(P) | %MD(P) | |
| acetic acid | 1.321 | 133 | 1.393 | 5.5 | 1.293 | −2.1 | 1.288 | −2.5 | 65 |
| adipic acid | 1.366 | 298 | 1.442 | 5.6 | 1.363 | −0.2 | 1.313 | −3.9 | 66 |
| β-oxalic acid | 1.906 | 298 | 1.927 | 1.1 | 1.906 | 0.0 | 1.791 | −6.0 | 67 |
| β-succinic acid | 1.562 | 298 | 1.629 | 4.3 | 1.540 | −1.4 | 1.512 | −3.2 | 68 |
| butyric acid | 1.135 | 230 | 1.253 | 10.4 | 1.100 | −3.1 | 1.122 | −1.1 | 69 |
| formic acid | 1.573 | 98 | 1.561 | −0.8 | 1.498 | −4.8 | 1.490 | −5.3 | 70 |
| phenylbenzoate | 1.272 | 298 | 1.358 | 6.8 | 1.251 | −1.7 | 1.243 | −2.3 | 71 |
| propionic acid | 1.219 | 138 | 1.314 | 7.8 | 1.234 | 1.2 | 1.234 | 1.2 | 72 |
| subacic acid | 1.216 | 298 | 1.312 | 7.9 | 1.215 | −0.1 | 1.180 | −3.0 | 73 |
| suberic acid | 1.270 | 298 | 1.364 | 7.4 | 1.286 | 1.3 | 1.242 | −2.2 | 74 |
| valeric acid | 1.140 | 298 | 1.229 | 7.8 | 1.179 | 3.4 | 1.162 | 1.9 | 75 |
| D-glucitol | 1.540 | 298 | 1.625 | 5.5 | 1.573 | 2.1 | 1.572 | 2.1 | 76 |
| D-mannitol | 1.471 | 298 | 1.528 | 3.9 | 1.446 | −1.7 | 1.461 | −0.7 | 77 |
| catechol | 1.378 | 298 | 1.465 | 6.3 | 1.347 | −2.2 | 1.378 | 0.0 | 78 |
| erythritol | 1.488 | 22.6 | 1.533 | 3.0 | 1.532 | 3.0 | 1.526 | 2.6 | 79 |
| ethanol | 1.025 | 87 | 1.079 | 5.3 | 1.013 | −1.2 | 0.990 | −3.4 | 80 |
| galactitol | 1.496 | 298 | 1.548 | 3.5 | 1.503 | 0.5 | 1.486 | −0.7 | 81 |
| γ-hydroquinone | 1.381 | 298 | 1.462 | 5.9 | 1.418 | 2.7 | 1.368 | −0.9 | 82 |
| pentaerythritol | 1.392 | 298 | 1.490 | 7.0 | 1.445 | 3.8 | 1.413 | 1.5 | 83 |
| ribitol | 1.456 | 298 | 1.563 | 7.3 | 1.456 | 0.0 | 1.481 | 1.7 | 84 |
| xylitol | 1.515 | 298 | 1.577 | 4.1 | 1.509 | −0.4 | 1.505 | −0.7 | 85 |
| adamantane | 1.179 | 298 | 1.204 | 2.1 | 1.135 | −3.7 | 1.133 | −3.9 | 86 |
| anthracene | 1.291 | 95 | 1.319 | 2.2 | 1.307 | 1.2 | 1.285 | −0.5 | 87 |
| anthracene | 1.245 | 290 | 1.319 | 5.9 | 1.251 | 0.5 | 1.230 | −1.2 | 88 |
| benzene | 1.114 | 295 | 1.125 | 1.0 | 1.094 | −1.8 | 1.090 | −2.2 | 89 |
| benzene | 1.024 | 78 | 1.125 | 9.9 | 1.020 | −0.4 | 0.964 | −5.9 | 90 |
| chrysene | 1.291 | 298 | 1.353 | 4.8 | 1.273 | −1.4 | 1.262 | −2.2 | 91 |
| n-hexane | 0.888 | 158 | 0.960 | 8.1 | 0.888 | 0.0 | 0.883 | −0.6 | 92 |
| naphthalene | 1.185 | 298 | 1.258 | 6.2 | 1.166 | −1.6 | 1.130 | −4.6 | 93 |
| octane | 0.891 | 193 | 0.992 | 11.3 | 0.903 | 1.3 | 0.905 | 1.6 | 94 |
| ovalene | 1.479 | 298 | 1.552 | 4.9 | 1.496 | 1.1 | 1.484 | 0.3 | 95 |
| pentane | 0.867 | 123 | 0.931 | 7.4 | 0.866 | −0.1 | 0.867 | 0.0 | 96 |
| perylene | 1.361 | 298 | 1.432 | 5.2 | 1.346 | −1.1 | 1.332 | −2.1 | 97 |
| phenanthrene | 1.209 | 298 | 1.291 | 6.8 | 1.213 | 0.3 | 1.206 | −0.2 | 98 |
| triphenylene | 1.308 | 298 | 1.375 | 5.1 | 1.292 | −1.2 | 1.277 | −2.4 | 99 |
| polyethylene | 1.063 | 213 | 1.102 | 3.7 | 1.014 | −4.6 | 1.044 | −1.8 | 100 |
| benzenamide | 1.290 | 298 | 1.374 | 6.5 | 1.252 | −2.9 | 1.283 | −0.5 | 101 |
| diketopiperazine | 1.592 | 298 | 1.614 | 1.4 | 1.533 | −3.7 | 1.504 | −5.5 | 102 |
| formamide | 1.333 | 90 | 1.408 | 5.6 | 1.329 | −0.3 | 1.371 | 2.9 | 103 |
| malonamide | 1.428 | 298 | 1.523 | 6.7 | 1.413 | −1.1 | 1.508 | 5.6 | 104 |
| oxamide | 1.668 | 298 | 1.711 | 2.6 | 1.628 | −2.4 | 1.635 | −2.0 | 105 |
| succinamide | 1.444 | 298 | 1.501 | 3.9 | 1.421 | −1.6 | 1.427 | −1.2 | 106 |
| triethylenediamine | 1.206 | 298 | 1.257 | 4.2 | 1.132 | −6.1 | 1.141 | −5.4 | 107 |
| hexmethylenetetramine | 1.346 | 298 | 1.453 | 7.9 | 1.336 | −0.7 | 1.334 | −0.9 | 108 |
| 3-azabicyclo[3.2.2]nonane | 1.124 | 123 | 1.180 | 5.0 | 1.068 | −5.0 | 1.066 | −5.2 | 109 |
| 1,5,9,13-tetraazacyclohexadecane | 1.070 | 298 | 1.141 | 6.6 | 1.071 | 0.1 | 1.061 | −0.8 | 110 |
| methylamine | 0.874 | 88 | 0.904 | 3.4 | 0.859 | −1.7 | 0.883 | 1.0 | 111 |
| ethylenediamine | 1.106 | 213 | 1.153 | 4.2 | 1.091 | −1.4 | 1.093 | −1.2 | 112 |
| trimethylamine | 0.879 | 118 | 0.956 | 8.8 | 0.888 | 1.0 | 0.893 | 1.6 | 113 |
| hexamethylendiamine | 1.003 | 298 | 1.104 | 10.1 | 1.018 | 1.5 | 1.010 | 0.7 | 114 |
| ethyl carbamate | 1.189 | 168 | 1.307 | 9.9 | 1.197 | 0.7 | 1.150 | −3.3 | 115 |
| PC (C₂₉O₆H₂₄) | 1.308 | 298 | 1.350 | 3.2 | 1.283 | −1.9 | 1.276 | −2.4 | 116 |
| PU (C₁₇H₁₈N₂O₂) | 1.316 | 258 | 1.370 | 4.1 | 1.306 | −0.8 | 1.278 | −2.9 | 117 |
| diethyl ether | 0.952 | 128 | 1.036 | 8.8 | 0.974 | 2.3 | 0.972 | 2.1 | 118 |
| trioxane | 1.456 | 103 | 1.535 | 5.4 | 1.418 | −2.6 | 1.446 | −0.7 | 119 |
| PVF(I) | 1.972 | 298 | 2.194 | 11.3 | 2.028 | 2.8 | 2.018 | 2.3 | 120 |
| PVF(II) | 1.925 | 298 | 2.045 | 6.2 | 1.956 | 1.6 | 1.946 | 1.1 | 121 |
| PVF(III) | 1.930 | 298 | 2.037 | 5.5 | 1.944 | 0.7 | 1.958 | 1.5 | 122 |
| PVF(IV) | 1.925 | 298 | 2.075 | 7.8 | 1.971 | 2.4 | 1.989 | 3.3 | 123 |
| α-D-glucose | 1.562 | 298 | 1.645 | 5.3 | 1.602 | 2.6 | 1.585 | 1.5 | 124 |
| β-D-glucose | 1.546 | 298 | 1.617 | 4.6 | 1.564 | 1.2 | 1.551 | 0.3 | 125 |
| β-lactose | 1.535 | 298 | 1.591 | 3.6 | 1.476 | −3.8 | 1.497 | −2.5 | 126 |
| sucrose | 1.590 | 298 | 1.650 | 3.8 | 1.586 | −0.3 | 1.572 | −1.1 | 127 |
| disiloxane | 1.104 | 108 | 1.180 | 6.9 | 1.062 | −3.8 | 1.038 | −6.0 | 128 |
| hexamethyldisiloxane | 0.989 | 148 | 1.088 | 10.0 | 0.988 | −0.1 | 0.987 | −0.2 | 128 |
| (NPC₁₂)₃ | 2.018 | 298 | 2.219 | 10.0 | 2.032 | 0.7 | 2.022 | 0.2 | 129 |
| (NPC₁₂)₄ | 2.172 | 298 | 2.278 | 4.9 | 2.105 | −3.1 | 2.104 | −3.1 | 130 |
| PDCP | 2.168 | 298 | 2.281 | 5.2 | 2.103 | −3.0 | 2.129 | −1.8 | 131 |
| PDMP | 1.243 | 298 | 1.430 | 15.0 | 1.279 | 2.9 | 1.274 | 2.5 | 132 |
| no. of data | | | | 69 | | 69 | | 69 | |
| max dev | | | | 15.0 | | 3.8 | | 5.6 | |
| min dev | | | | −0.8 | | −6.1 | | −6.0 | |
| av dev | | | | 5.9 | | −0.6 | | −1.0 | |
| rms dev | | | | 2.8 | | 2.2 | | 2.5 | |

[a] Results of three types of simulations are listed: the energy minimization (MM), NPT simulation with pressure control method proposed by berendsen et al. MD(B), and NPT simulations with pressure control method of parrinello−Rahman MD(P). The percentage errors with respect to the experimental data are given for each of these calculations.

COMPASS—Details on Alkane and Benzene Compounds

*J. Phys. Chem. B, Vol. 102, No. 38, 1998* **7353**
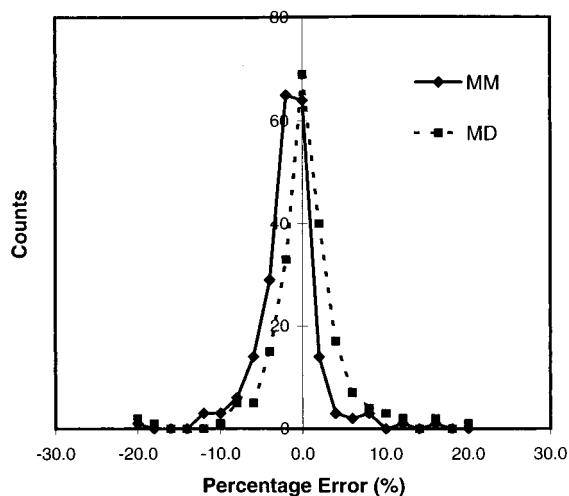


**Figure 11.** Distribution of percentage errors of calculated cell-edge parameters for molecular crystals listed in Table 6. The solid line represents the distribution of percentage errors of the energy minimization calculations, and the dashed line represents the distribution of percentage errors of the NPT simulation using the Parrinello—Rahman pressure control method.
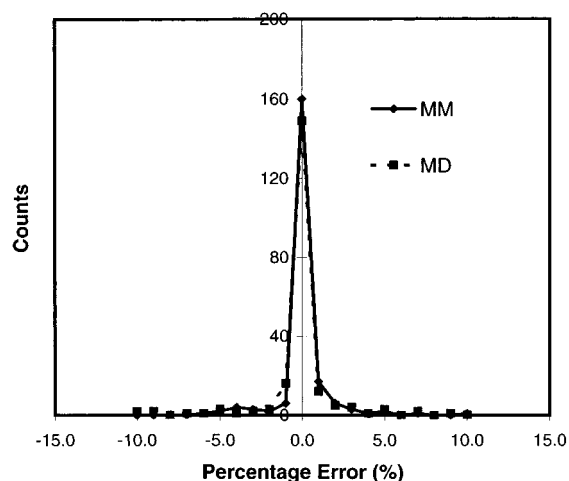


**Figure 12.** Distribution of percentage errors of calculated cell-angle parameters for molecular crystals listed in Table 6. The solid line represents the distribution of percentage errors of the energy minimization calculations, and the dashed line represents the distribution of percentage errors of the NPT simulation using the Parrinello—Rahman pressure control method.

than those in the PCFF(CFF9X). The difference in the radii reflects the difference in the parametrization. The new parameters are optimized based on molecular dynamics simulations at finite temperature. Due to the thermal expansion, which is completely missing in the energy minimization, the radii should be smaller than what was 'measured' in the MD simulation. The larger values of the well depths can be understood by considering London's dispersion formula:

$$\epsilon_o = \frac{C_6}{r_o{}^6} \tag{10}$$

Table 9 gives a series comparisons of the structural parameters for linear alkanes, cycloalkanes, and aromatic hydrocarbons. For alkanes and cycloalkanes, the published CFF93 results are also listed for comparison. Good agreements are obtained between the COMPASS and CFF93 force fields and between the force-field calculation and the experimental measurement.[137-147]

**TABLE 7: Comparison of Lattice Energies (in kcal/mol)**

| molecule | exp | calcd | Δ | ref |
|---|---|---|---|---|
| pentane | 9.9 | 10.7 | 0.8 | 133 |
| hexane | 12.6 | 12.9 | 0.3 | 133 |
| octane | 15.9 | 16.9 | 1.0 | 133 |
| benzene | 12.5 | 12.0 | −0.5 | 134 |
| naphthalene | 17.3 | 18.6 | 1.3 | 134 |
| anthracene | 24.4 | 25.7 | 1.3 | 134 |
| pheranthrene | 20.7 | 22.7 | 2.0 | 134 |
| triphenylene | 27.4 | 30.4 | 3.0 | 134 |
| perylene | 31.0 | 33.8 | 2.8 | 135 |
| ovalene | 50.6 | 49.4 | −1.2 | 136 |
| foramide | 17.5 | 16.9 | −0.6 | 20 |
| oxamide | 28.2 | 27.4 | −0.8 | 20 |
| malonamide | 28.8 | 31.4 | 2.6 | 20 |
| diketopiperazine | 26.0 | 30.5 | 4.5 | 20 |
| succinamide | 32.3 | 34.7 | 2.4 | 20 |
| formic acid | 15.2 | 13.6 | −1.6 | 21 |
| acetic acid | 16.3 | 15.3 | −1.0 | 21 |
| propionic acid | 17.7 | 16.4 | −1.3 | 21 |
| butyric acid | 19.2 | 17.9 | −1.3 | 21 |
| β-succinic acid | 29.3 | 29.1 | −0.2 | 21 |
| adipic acid | 32.1 | 31.4 | −0.7 | 21 |
| suberic acid | 35.4 | 35.4 | 0.0 | 21 |
| subaric acid | 39.6 | 39.3 | −0.3 | 21 |
| valeric acid | 20.2 | 20.0 | −0.2 | 21 |

**TABLE 8: Comparison of LJ-9-6 Parameters**

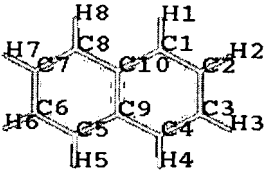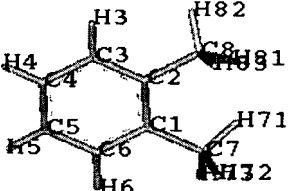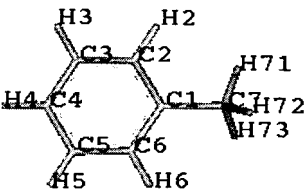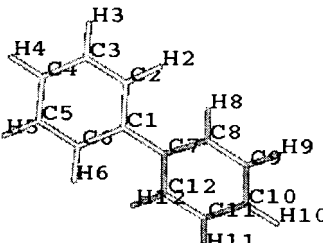| | PCFF/CFF93 | | COMPASS | |
|---|---|---|---|---|
| | $r_0$ | $\epsilon_0$ | $r_0$ | $\epsilon_0$ |
| c4 | 4.010 | 0.054 | 3.854 | 0.062 |
| c3a | 4.010 | 0.064 | 3.915 | 0.068 |
| h1 | 2.995 | 0.020 | 2.878 | 0.023 |

For aromatic hydrocarbons, several compounds are used to validate the parameters. For benzene and alkyl benzenes, such as toluene and *o*-xylene, the force field yields good agreement with the experimental data. However, for extended conjugation systems and fused rings, such as biphenyl and naphthalene, the current functional form of the valence bond is not flexible enough. This is because the $\pi$ electrons are partially localized in certain region so that all C—C bonds are not equivalent. The only information to define a bond in the present force field is the atom types of the two atoms that are bonded together. For example, all C—C bonds in and between aromatic rings are treated using the same atom type of c3a. One way to solve this problem is to use the bond order in the formulas. The cross-coupling terms in the current functional form can be used to represent the alternation to some extent, but the effectiveness is limited. This is demonstrated by some large discrepancies in the bond lengths obtained for biphenyl, as given in Table 9.

Comparisons of the vibrational frequencies are listed in Table 10. Since the original force constants were well-defined based on the ab initio Hessian matrices and subsequently scaled empirically, very little adjustment was required after the vdW parameters were modified in order to obtain good agreement with the experimental frequencies. Overall, the calculated frequencies agree well with the experimental data.[148-150] There are 168 data points. The maximum deviations are 57.7 and −121.9 cm⁻¹, the average deviation is −8.2 cm⁻¹, and the rms deviation is 30.2 cm⁻¹. These results are comparable with other high-quality force fields (MM3[1], CFF93,[7] MMFF[11]).

Comparisons of the conformation energies between the calculated and experimental data[151-160] are given in Table 11. Although the vdW parameters have a strong influence on the conformational energies of these molecules, this impact can be easily compromised by adjusting the torsion parameters only. As indicated in this table, excellent agreement between the

**TABLE 9: Comparison of Structural Parameters of Alkane and Benzene Compounds**

| property | obsd | CFF93 | COMPASS | property | obsd | CFF93 | COMPASS | structure |
|---|---|---|---|---|---|---|---|---|
| | | | | Methane[137] | | | | |
| C—H | 1.107 | 1.108 | 1.099 | | | | | |
| | | | | Ethane[138] | | | | |
| C—H | 1.112 | 1.112 | 1.102 | H—C—H | | 108.0 | 108.1 | |
| C—C | 1.534 | 1.526 | 1.525 | C—C—H | | 111.0 | 111.0 | |
| | | | | Butane[139] | | | | |
| C1—C2 | 1.531 | 1.534 | 1.532 | C1—C2—C3 | 113.8 | 113.2 | 112.8 | |
| C2—C3 | | 1.538 | 1.529 | H1—C1—C2 | 110.0 | 111.6 | 111.4 | |
| C1—H1 | 1.117 | 1.112 | 1.102 | H2—C2—H2 | | 106.7 | 106.5 | |
| C2—H2 | 1.088 | 1.115 | 1.104 | | | | | |
| | | | | Isobutane[140] | | | | |
| C—C | 1.535 | 1.537 | 1.532 | C'—C—H | 111.4 | 111.6 | 111.4 | |
| C—H | 1.122 | 1.116 | 1.106 | C'—C—H | 110.1 | 110.6 | 110.6 | |
| C'—H[a] | 1.113 | 1.113 | 1.102 | H—C—H | 108.7 | 107.4 | 107.6 | |
| C—C'—C | 110.8 | 110.7 | 110.7 | H—C—H | 106.5 | 107.7 | 107.8 | |
| H—C'—C | 108.1 | 108.2 | 108.2 | | | | | |
| | | | | Cyclopropane[141] | | | | |
| C—C | 1.514 | 1.503 | 1.506 | C—C—H | 117.9 | 118.7 | 118.57 | |
| C—H | 1.099 | 1.104 | 1.093 | H—C—H | 114.5 | 112.7 | 112.97 | |
| | | | | Cyclobutane[142] | | | | |
| C—C | 1.552 | 1.549 | 1.543 | C—C—H(e) | | 118.0 | 118.1 | |
| C—H(e)[b] | 1.093 | 1.107 | 1.097 | C—C—H(a) | | 111.6 | 111.4 | |
| C—H(a)[b] | | 1.111 | 1.101 | H—C—H | 106.4 | 108.3 | 108.6 | |
| C—C—C | | 88.4 | 88.2 | | | | | |
| | | | | Cyclohexane[143] | | | | |
| C—C | 1.536 | 1.543 | 1.537 | C—C—H(e) | 110.1 | 110.5 | 110.4 | |
| C—H(e) | 1.121 | 1.114 | 1.104 | C—C—H(a) | 109.2 | 109.2 | 109.2 | |
| C—H(a) | 1.121 | 1.115 | 1.105 | H—C—H | 107.5 | 106.4 | 106.5 | |
| C—C—C | 111.4 | 111.0 | 111.0 | C1—C2—C3—C4 | 54.9 | 56.2 | 56 | |
| | | | | Benzene[144] | | | | |
| C—C | 1.399 | | 1.398 | C—C—C | | | 120.0 | |
| C—H | 1.100 | | 1.099 | C—C—H | | | 120.0 | |
| | | | | Naphthalene[145] | | | | |
| C2—C3 | 1.417 | | 1.400 | C1—C2—C3 | 120.0 | | 120.6 | |
| C1—C2 | 1.381 | | 1.383 | C10—C1—C2 | 120.0 | | 120.6 | |
| C1—C10 | 1.422 | | 1.418 | C9—C10—C1 | 120.0 | | 118.8 | |
| C9—C10 | 1.412 | | 1.444 | C10—C1—H1 | 120.0 | | 119.6 | |
| | | | | *o*-Xylene[146] | | | | |
| C1—C7 | 1.509 | | 1.511 | H72—C7—H72 | 108.9 | | 107.9 | |
| C1—C6 | 1.414 | | 1.389 | H71—C7—H73 | 106.0 | | 107.2 | |
| C1—C2 | 1.394 | | 1.400 | C7—C1—C2 | 121.1 | | 121.5 | |
| C7—H71 | 1.080 | | 1.099 | C6—C1—C2 | 119.8 | | 119.1 | |
| C7—H72 | 1.095 | | 1.102 | C2—C3—H3 | 118.9 | | 119.0 | |
| C7—H73 | 1.095 | | 1.102 | C3—C4—H4 | 119.7 | | 120.4 | |
| C3—H3 | 1.072 | | 1.100 | H71—C7—C1 | 111.2 | | 111.3 | |
| C4—H4 | 1.079 | | 1.099 | H72—C7—C1 | 111.0 | | 111.3 | |
| H7—C7—H72 | 108.9 | | 107.9 | H73—C7—C1 | 111.0 | | 111.8 | |
| | | | | Toluene[54] | | | | |
| C1—C2 | 1.399 | | 1.394 | C7—H71 | 1.110 | | 1.099 | |
| C2—C3 | 1.399 | | 1.393 | C7—H72 | 1.110 | | 1.103 | |
| C3—C4 | 1.399 | | 1.397 | C7—H73 | 1.110 | | 1.100 | |
| C1—C7 | 1.524 | | 1.509 | C2—C1—C7 | 120.0 | | 121.0 | |
| C2—H2 | 1.110 | | 1.099 | C3—C2—C1 | 120.0 | | 121.4 | |
| C3—H3 | 1.110 | | 1.100 | C4—C3—C2 | 120.0 | | 119.9 | |
| C4—H4 | 1.110 | | 1.099 | | | | | |
| | | | | Biphenyl[147] | | | | |
| C1—C7 | 1.489 | | 1.437 | C1—C2—C3 | 121.3 | | 121.3 | |
| C1—C2 | 1.403 | | 1.430 | C2—C3—C4 | 119.8 | | 120.2 | |
| C2—C3 | 1.396 | | 1.381 | C3—C4—C5 | 119.8 | | 120.2 | |
| C3—C4 | 1.398 | | 1.398 | C2—C1—C7—C8 | −44.5 | | −30.9 | |
| C2—C1—C6 | 117.9 | | 116.9 | | | | | |

[a] C' denotes carbon at branch point. [b] H(a) = axial hydrogen; H(e) = equaborial hydrogen.

COMPASS—Details on Alkane and Benzene Compounds

*J. Phys. Chem. B, Vol. 102, No. 38, 1998* **7355**

**TABLE 10: Comparison of Normal Mode Frequencies (in cm$^{-1}$)**

| no. | sym | obsd | calcd | diff | assign | no. | sym | obsd | calcd | diff | assign |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Methane[148] | | | | | | |
| 1 | $T_2$ | 3019 | 2992 | −27 | asym str | 6 | E | 1534 | 1509 | −25 | |
| 2 | $T_2$ | 3019 | 2992 | −27 | | 7 | $T_2$ | 1306 | 1313 | 7 | |
| 3 | $T_2$ | 3019 | 2992 | −27 | | 8 | $T_2$ | 1306 | 1313 | 7 | |
| 4 | $A_1$ | 2917 | 2894 | −23 | sym str | 9 | $T_2$ | 1306 | 1313 | 7 | |
| 5 | E | 1534 | 1509 | −25 | def | | | | | | |
| | | | | | Ethane[149] | | | | | | |
| 1 | $E_u$ | 2974 | 2970 | −4 | asym str | 10 | $E_g$ | 1460 | 1454 | −6 | |
| 2 | $E_u$ | 2974 | 2970 | −4 | | 11 | $A_{1g}$ | 1388 | 1423 | 35 | |
| 3 | $E_g$ | 2950 | 2967 | 17 | | 12 | $A_{2u}$ | 1370 | 1397 | 27 | |
| 4 | $E_g$ | 2950 | 2967 | 17 | | 13 | $E_g$ | 1190 | 1139 | −51 | CH$_3$ sym rocking |
| 5 | $A_{1g}$ | 2915 | 2904 | −12 | sym str | 14 | $E_g$ | 1190 | 1139 | −51 | |
| 6 | $A_{2u}$ | 2915 | 2897 | −18 | | 15 | $A_{1g}$ | 995 | 963 | −32 | CC str |
| 7 | $E_u$ | 1469 | 1459 | −10 | CH$_3$ deform | 16 | $E_u$ | 822 | 769 | −53 | CH$_3$ asym rocking |
| 8 | $E_u$ | 1469 | 1459 | −10 | | 17 | $E_u$ | 822 | 769 | −53 | |
| 9 | $E_g$ | 1460 | 1454 | −6 | | 18 | $A_{2u}$ | 279 | 312 | 33 | CH$_3$−CH$_3$ torsion |
| | | | | | Propane[149] | | | | | | |
| 1 | $B_2$ | 2965 | 2969 | 4 | CH$_3$ asym str | 15 | $A_1$ | 1370 | 1413 | 43 | CH$_2$ sci + CH$_3$ def |
| 2 | $A_2$ | 2965 | 2967 | 2 | | 16 | $B_1$ | 1332 | 1364 | 32 | CH$_2$ wag + CH$_3$ def |
| 3 | $A_1$ | 2965 | 2966 | 1 | | 17 | $A_2$ | 1278 | 1232 | −46 | CH$_2$ twist |
| 4 | $B_1$ | 2965 | 2966 | 1 | | 18 | $B_2$ | 1187 | 1138 | −49 | CH$_2$ rock + CH$_3$ def |
| 5 | $B_2$ | 2915 | 2950 | 35 | CH$_2$ asym str | 19 | $A_1$ | 1157 | 1130 | −27 | CH$_3$ rock + CCC bend |
| 6 | $A_1$ | 2875 | 2909 | 34 | CH$_2$ sym str | 20 | $B_1$ | 1049 | 1007 | −42 | CH$_3$ rock + CH$_2$ wag + CC str |
| 7 | $B_1$ | 2875 | 2899 | 24 | CH$_3$ sym str | 21 | $A_2$ | 899 | 900 | 1 | CH$_2$ twist + CH$_3$ def |
| 8 | $A_1$ | 2875 | 2897 | 22 | | 22 | $B_1$ | 921 | 866 | −55 | CH$_3$ rock + asym CC str |
| 9 | $B_1$ | 1464 | 1473 | 9 | CH$_3$ def + CH$_2$ wag | 23 | $A_1$ | 868 | 824 | −44 | CH$_3$ rock + sym CC str |
| 10 | $B_2$ | 1459 | 1472 | 13 | CH$_3$ def + CH$_2$ wag | 24 | $B_2$ | 748 | 700 | −48 | CH$_2$ rock + CH$_3$ def |
| 11 | $A_1$ | 1473 | 1464 | −9 | CH$_2$ sci + CH$_3$ def | 25 | $A_1$ | 375 | 320 | −56 | CCC bend |
| 12 | $A_2$ | 1473 | 1459 | −14 | CH$_3$ def + CH$_2$ wag | 26 | $B_2$ | 265 | 268 | 3 | CH$_3$−CH$_2$ torsion |
| 13 | $B_1$ | 1449 | 1451 | 2 | | 27 | $A_2$ | 217 | 233 | 16 | |
| 14 | $A_1$ | 1385 | 1422 | 37 | CH$_2$ sci + CH$_3$ sci | | | | | | |
| | | | | | Butane[149] | | | | | | |
| 1 | $A_u$ | 2966 | 2967 | 1 | CH$_3$ asym str | 19 | $A_g$ | 1375 | 1396 | 21 | |
| 2 | $B_g$ | 2966 | 2967 | 1 | | 20 | $B_u$ | 1300 | 1306 | 6 | CH$_2$ wag + CH$_3$ def |
| 3 | $B_u$ | 2965 | 2967 | 2 | | 21 | $A_u$ | 1293 | 1261 | −32 | CH$_2$ twist |
| 4 | $A_g$ | 2965 | 2967 | 2 | | 22 | $B_g$ | 1257 | 1224 | −34 | |
| 5 | $A_u$ | 2920 | 2948 | 28 | CH$_2$ asym str | 23 | $A_g$ | 1148 | 1140 | −8 | CC str + CH$_3$ rock |
| 6 | $B_g$ | 2912 | 2947 | 35 | | 24 | $B_g$ | | 1129 | | CH$_2$ rock + CH$_3$ def |
| 7 | $A_g$ | 2875 | 2908 | 33 | CH$_2$ sym str | 25 | $A_u$ | 1053 | 1017 | −36 | CH$_2$ twist + CH$_3$ def |
| 8 | $B_u$ | 2872 | 2904 | 32 | | 26 | $B_u$ | 1010 | 1002 | −8 | CH$_3$ rock + CC str |
| 9 | $A_g$ | 2861 | 2898 | 37 | CH$_3$ sym str | 27 | $A_g$ | 965 | 933 | −32 | CC str + CH$_3$ rock |
| 10 | $B_u$ | 2853 | 2898 | 45 | | 28 | $B_u$ | 944 | 922 | −22 | |
| 11 | $A_g$ | 1462 | 1490 | 28 | CH$_3$ def | 29 | $B_g$ | 835 | 801 | −34 | CH$_2$ rock + CH$_3$ def |
| 12 | $B_g$ | 1460 | 1474 | 14 | | 30 | $A_g$ | 733 | 760 | 27 | CC str + CH$_3$ rock |
| 13 | $A_u$ | 1468 | 1462 | −6 | | 31 | $A_u$ | | 687 | | CH$_2$ rock + CH$_3$ def |
| 14 | $B_u$ | 1459 | 1461 | 2 | CH$_2$ sci + CH$_3$ sci | 32 | $A_g$ | 427 | 378 | −49 | CCC bend |
| 15 | $A_g$ | 1459 | 1459 | −1 | | 33 | $B_u$ | | 268 | | |
| 16 | $B_u$ | 1455 | 1458 | 3 | | 34 | $B_g$ | 266 | 230 | −36 | CH$_3$−CH$_2$ torsion |
| 17 | $A_g$ | 1455 | 1433 | −22 | | 35 | $A_u$ | | 220 | | |
| 18 | $B_u$ | 1375 | 1413 | 38 | CH$_2$ sci + CH$_3$ def | 36 | $A_u$ | 121 | 127 | 6 | CH$_2$−CH$_2$ torsion |
| | | | | | Isobutane[149] | | | | | | |
| 1 | $A_1$ | 2962 | 2966 | 3.8 | CH$_3$ asym str | 19 | $A_1$ | 1394 | 1418 | 24.4 | |
| 2 | E | 2962 | 2966 | 3.8 | | 20 | E | 1330 | 1352 | 22.4 | |
| 3 | E | 2962 | 2965 | 3.3 | | 21 | E | 1330 | 1352 | 22.4 | |
| 4 | E | 2962 | 2963 | 1.4 | | 22 | E | 1166 | 1169 | 3 | CH$_3$ rock + CC str |
| 5 | E | 2962 | 2963 | 1.4 | | 23 | E | 1166 | 1148 | −18.1 | |
| 6 | $A_2$ | 2958 | 2962 | 4 | | 24 | $A_1$ | 1177 | 1148 | −29.1 | CH$_3$ rock + CCC bend |
| 7 | $A_1$ | 2904 | 2917 | 13.2 | CH str | 25 | E | 966 | 941 | −25.4 | CH$_3$ rock |
| 8 | E | 2894 | 2897 | 2.8 | CH$_3$ sym str | 26 | E | 966 | 907 | −58.6 | |
| 9 | E | 2894 | 2897 | 2.8 | | 27 | $A_2$ | | 907 | | |
| 10 | $A_1$ | 2880 | 2896 | 16.1 | | 28 | E | 918 | 900 | −18.3 | |
| 11 | $A_1$ | 1477 | 1483 | 5.5 | CH$_3$ def | 29 | E | 918 | 900 | −18.3 | |
| 12 | E | 1477 | 1483 | 5.5 | | 30 | $A_1$ | 797 | 764 | −33.2 | CC str + CH$_3$ rock |
| 13 | E | | 1468 | | | 31 | $A_1$ | 433 | 384 | −49.5 | CCC bend + CH$_3$ rock |
| 14 | E | 1475 | 1458 | −16.7 | | 32 | E | 367 | 307 | −59.9 | CCC bend |
| 15 | E | 1475 | 1457 | −17.6 | | 33 | E | 367 | 307 | −59.9 | |
| 16 | $A_2$ | 1450 | 1457 | 7.4 | | 34 | E | | 256 | | CH$_3$−CH torsion |
| 17 | E | 1371 | 1429 | 57.7 | | 35 | E | | 256 | | |
| 18 | E | 1371 | 1429 | 57.7 | | 36 | $A_2$ | | 242 | | |

**TABLE 10: (Continued)**

| no. | sym | obsd | calcd | diff | assign | no. | sym | obsd | calcd | diff | assign |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Cyclopropane[149] | | | | | |
| 1 | $A_2''$ | 3102 | 3059 | −43 | CH str | 12 | $A_1'$ | 1188 | 1228 | 40 | ring breathing |
| 2 | $E''$ | 3083 | 3057 | −26 | | 13 | $A_1''$ | 1126 | 1051 | −75 | CH₂ twist |
| 3 | $E''$ | 3083 | 3057 | −26 | | 14 | $E'$ | 1028 | 1051 | 23 | CH₂ wag |
| 4 | $A_1'$ | 3038 | 3005 | −34 | | 15 | $E'$ | 1028 | 1006 | −22 | |
| 5 | $E'$ | 3024 | 2991 | −33 | | 16 | $A_2'$ | 1070 | 948 | −122 | |
| 6 | $E'$ | 3024 | 2991 | −33 | | 17 | $E'$ | 869 | 823 | −47 | ring def |
| 7 | $A_1'$ | 1482 | 1508 | 26 | CH₂ bend | 18 | $E'$ | 869 | 823 | −47 | |
| 8 | $E'$ | 1438 | 1465 | 27 | | 19 | $E''$ | 738 | 722 | −16 | CH₂ twist |
| 9 | $E'$ | 1438 | 1465 | 27 | | 20 | $E''$ | 738 | 722 | −16 | |
| 10 | $E''$ | 1187 | 1241 | 54 | CH₂ rock | 21 | $A_2''$ | 654 | 637 | −17 | CH₂ rock |
| 11 | $E''$ | 1187 | 1228 | 41 | | | | | | | |
| | | | | | Benzene[150] | | | | | | |
| 1 | $A_{1u}$ | 3062 | 3082 | 20 | CH str | 19 | $B_{1u}$ | 1010 | 965 | −45 | CCC bend |
| 2 | $E_{1u}$ | 3080 | 3067 | −13 | | 20 | $E_{2u}$ | 985 | 952 | −33 | CH out-of-plane |
| 3 | $E_{1u}$ | 3080 | 3067 | −13 | | 21 | $E_{2u}$ | 985 | 945 | −40 | |
| 4 | $E_{2g}$ | 3048 | 3061 | 13 | | 22 | $A_{1g}$ | 993 | 945 | −48 | breathing |
| 5 | $E_{2g}$ | 3048 | 3059 | 11 | | 23 | $E_{1g}$ | 850 | 804 | −46 | CH out-of-plane |
| 6 | $B_{1u}$ | 3060 | 3059 | −1 | | 24 | $E_{1g}$ | 850 | 804 | −46 | |
| 7 | $B_{2u}$ | 1693 | 1707 | 14 | CC str | 25 | $B_{2g}$ | 685 | 659 | −26 | CCC out-of-plane |
| 8 | $E_{2g}$ | 1595 | 1598 | 3 | | 26 | $A_{2u}$ | 671 | 624 | −47 | CH out-of-plane |
| 9 | $E_{2g}$ | 1595 | 1598 | 3 | | 27 | $E_{2g}$ | 606 | 604 | −2 | CCC in-plane |
| 10 | $E_{1u}$ | 1485 | 1453 | −32 | | 28 | $E_{2g}$ | 606 | 604 | −2 | |
| 11 | $E_{1u}$ | 1485 | 1453 | −32 | | 29 | $E_{2u}$ | 400 | 384 | −16 | CCC out-of-plane |
| 12 | $A_{2g}$ | 1298 | 1286 | −13 | CCH in-plane bend | 30 | $E_{2u}$ | 400 | 384 | −16 | |
| 13 | $E_{2g}$ | 1178 | 1172 | −6 | | no. of data | | | | 168 | |
| 14 | $E_{2g}$ | 1178 | 1122 | −56 | | max | | | | 57.7 | |
| 15 | $B_{2u}$ | 1170 | 1122 | −48 | | min | | | | −121.9 | |
| 16 | $E_{1u}$ | 1035 | 1082 | 47 | | av | | | | −8.2 | |
| 17 | $E_{1u}$ | 1035 | 994 | −41 | | rms | | | | 30.2 | |
| 18 | $B_{2g}$ | 1016 | 994 | −22 | CH out-of-plane | | | | | | |

**TABLE 11: Comparison of Conformational Energies of Alkanes and Benzenes (in kcal/mol)**

| molecule | property | obsd | COMPASS | ref |
|---|---|---|---|---|
| ethane | barrier (C−C) | 2.88 | 2.82 | 151 |
| propane | barrier (C−C) | 3.40 | 3.18 | 152 |
| isobutane | barrier (C−C) | 3.90 | 3.55 | 153 |
| neopentane | barrier (C−C) | 4.20, 4.80 | 3.96 | 154 |
| butane | barrier (C−C) | 4.56 | 4.85 | 155 |
| | barrier (A,G) | 3.30 | 3.64 | 152 |
| | $\Delta E$(A,G) | 0.50, 0.89 | 0.88 | 156 |
| pentane | $\Delta E$(AA,AG) | 0.465, 0.56 | 0.84 | 157, 158 |
| biphenyl | twist | 0.00 | 0.00 | 159 |
| | planar | 2.00 | 2.23 | 159 |
| | perpendicular | 1.00 | 1.05 | 159 |
| ethylbenzene | B(Ph−C) | 1.50 | 1.68 | 160 |

calculated and experimental conformational energies and energy barrier heights are obtained.

Data for liquid and crystal alkanes and benzenes are included in Tables 4−7. More comparisons of the liquid simulation results are made between the COMPASS and PCFF force fields in Tables 12 and 13. With the experimental densities and temperatures, the average pressures estimated over 50 ps NVT simulation after the equilibrium are between ±200 bar using the COMPASS force field while the values obtained using the PCFF force field are at least several hundred bar too high. The values of the heat of vaporization of both force fields, however, are in good agreement with the experimental values when the experimental densities are used. The average densities estimated using NPT simulations over 100 ps after the equilibrium are given in Table 13. Consistent with the results of NVT simulations, the COMPASS force field yields good agreement with the experimental densities but systematically low densities are obtained using the PCFF parameters. The values of the heat of vaporization were calculated using the simulated densities as well. The agreement with the experimental data is reasonably good using the COMPASS force field. However, significantly large errors are found using the PCFF force field.

The data points used to optimize the vdW parameters are densities and cohesive energies measured at one point on the P−V−T surface for each of the molecules. By simultaneously fitting all molecules of the training set, the number of observables is hopefully large enough to fully determine the potential-energy surfaces. The validity of the force field outside of the parametrization zone needs to be tested. Indeed, validation studies show that the present force field is able to predict various thermophysical properties in a broad range of experimental conditions.[24−28] To demonstrate this point, P−V−T data calculated for liquid n-hexane and benzene are plotted in Figures 13 and 14 together with experimental results[161,162] for comparison.

Liquid n-hexane was not in the parametrization training set. Hence, none of the data points given in Figure 13 were used for parametrization. Comparison of the data shows that in a broad range of pressure (0−4000 bar) and temperature (223−333 K), the calculated isothermal compressibility data agree very well with the experimental results.[161] The largest deviation, which occurs at the high-pressure (about 3000 bar) and low-temperature (223 K) region, is only about 3%.

The P−V−T data calculated for liquid benzene are presented in Figure 14. Although liquid benzene was included in the training set, the parametrization data point used was at room temperature (298 K) and 1 atm (see Table 4), which is not plotted in this chart. In Figure 14, three isothermal compressibility curves[162] are plotted at T = 373.2, 473.2, and 673.2 K. The pressure ranges roughly from 0 to 3000 bar; the temperature spans 300 K. Despite the large variation of temperature, the agreement between the calculated and experimental data is remarkable. Overall, the largest deviation is also about 3%.

Similar results have been obtained for many other molecular

**TABLE 12: NVT Simulation Results of Alkanes and Benzenes, Comparison between COMPASS and PCFF/CFF91 Force Fields**[a]

|  | obsd[56] | | | COMPASS | | | PCFF/CFF91 | | |
|---|---|---|---|---|---|---|---|---|---|
|  | $T$ (K) | $D$ | $H_v$ | $P$ | $H_v$ | $\%H_v$ | $P$ | $H_v$ | $\%H_v$ |
| cyclohexane | 293.2 | 0.779 | 7.96 | 135 | 7.98 | 0.3 | 1023 | 7.96 | 0.0 |
| ethane | 184.5 | 0.546 | 3.51 | 193 | 3.25 | −7.3 | 884 | 3.17 | −9.7 |
| isopentane | 293.2 | 0.620 | 6.01 | −100 | 6.12 | 1.9 | 355 | 6.11 | 1.7 |
| methane | 111.0 | 0.424 | 1.96 | −149 | 1.83 | −6.5 | 339 | 1.78 | −9.2 |
| pentane | 293.2 | 0.626 | 6.39 | 53 | 6.30 | −1.4 | 588 | 6.30 | −1.4 |
| propane | 231.1 | 0.581 | 4.49 | 63 | 4.29 | −4.4 | 672 | 4.25 | −5.4 |
| benzene | 298.20 | 0.872 | 8.09 | 48 | 8.18 | 1.1 | 771 | 8.21 | 1.4 |
| toluene | 298.20 | 0.865 | 9.09 | −103 | 9.08 | −0.1 | 711 | 9.09 | 0.0 |

[a] $D$ = density in g/cm³; $H_v$ = heat of vaporization in kcal/mol, $P$ = pressure in bar. $\%H_v$ = percentage errors in $H_v$ with respect to the experimental data.

**TABLE 13: NPT Simulation Results, Comparison between COMPASS and PCFF/CFF91 Force Fields**[a]
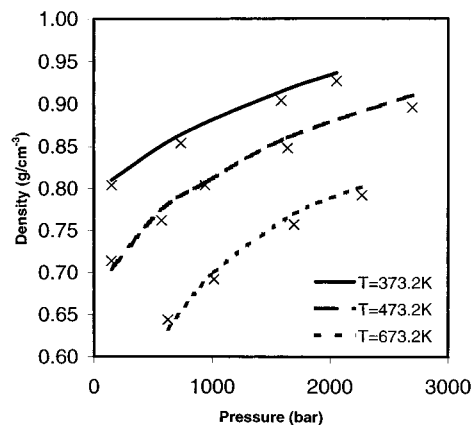
|  | obsd[36] | | | COMPASS | | | | PCFF/CFF91 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
|  | $D$ | $T$ (K) | $H_v$ | $D$ | $\%D$ | $H_v$ | $\%H_v$ | $D$ | $\%D$ | $H_v$ | $\%H_v$ |
| cyclohexane | 0.779 | 293.2 | 7.96 | 0.774 | −0.6 | 7.90 | −0.7 | 0.713 | −8.5 | 7.26 | −8.7 |
| ethane | 0.546 | 184.5 | 3.51 | 0.522 | −4.4 | 3.11 | −11.3 | 0.465 | −14.8 | 2.77 | −21.0 |
| isopentane | 0.620 | 293.2 | 6.01 | 0.633 | 2.1 | 6.26 | 4.2 | 0.585 | −5.6 | 5.75 | −4.3 |
| methane | 0.424 | 111.0 | 1.96 | 0.446 | 5.2 | 1.91 | −2.5 | 0.389 | −8.3 | 1.67 | −14.9 |
| pentane | 0.626 | 293.2 | 6.39 | 0.619 | −1.1 | 6.27 | −1.9 | 0.570 | −8.9 | 5.76 | −9.9 |
| propane | 0.581 | 231.1 | 4.49 | 0.570 | −1.9 | 4.22 | −6.0 | 0.512 | −11.9 | 3.76 | −16.3 |
| benzene | 0.872 | 298.2 | 8.09 | 0.862 | −1.1 | 8.05 | −0.5 | 0.817 | −6.3 | 7.67 | −5.2 |
| toluene | 0.865 | 298.2 | 9.09 | 0.855 | −1.2 | 8.99 | −1.1 | 0.809 | −6.5 | 8.58 | −5.6 |

[a] $D$ = density in g/cm³; $H_v$ = heat of vaporization in kcal/mol, $\%D$ and $\%H_v$ = percentage errors with respect to the experimental data.
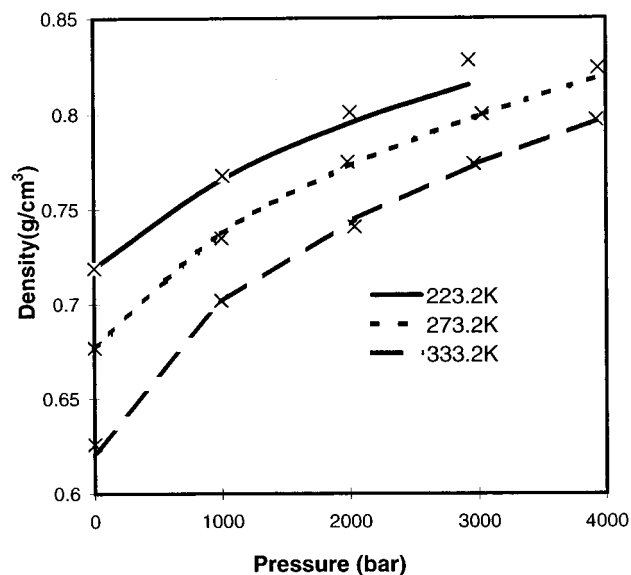


**Figure 13.** Comparison of calculated and experimental[161] isothermal compressibility for liquid *n*-hexane. The lines represent the experimental results, and the data points denote the calculated values.



**Figure 14.** Comparison of calculated and experimental[162] isothermal compressibility for liquid benzene. The lines represent the experimental results, and the data points denote the calculated values.

liquids and polymers.[24−28] In addition to the P−V−T data, cohesive-energy densities calculated at different temperatures were found to be in excellent agreement with the experimental data.[24,25] The ability to predict the P−V−T maps and cohesive energies as functions of temperature in a broad range is significant. First of all, it indicates that other properties, such as compressibility ($\kappa$) and thermal expansion coefficients ($\alpha$), can be well predicted. More importantly, it demonstrated that the simple functions (Lennard−Jones and Coulombic) for nonbond interactions are adequate to describe the thermophysical properties in a very broad range. Finally, it shows that the present force field can be used not only to reproduce the experimental data that were used in the parametrization, but also to predict properties outside of the parametrization region.

The predictive power of the COMPASS force field is also demonstrated by calculations of molecular crystals (see Tables 6 and 7) since most of these crystals were not used in the parametrization. In addition to the overall comparison given in Table 6 and Figures 11 and 12, detailed comparisons of the cell parameters for a total of 12 molecular crystals of alkanes, cycloalkanes, and aromatics are given in Table 14. The experimental data, including the temperature, cell symmetry, the number of molecules (Z) in the unit cell, and cell parameters, are listed with references. Among these molecules, benzene and anthracene are given with two temperatures. Both energy minimization (MM) and molecular dynamics (MD) data are listed for comparison. The MM calculations were performed on the unit cell with full Ewald summation of the nonbond interactions. The NPT simulations were performed on supercells consisting of a number of unit cells, using the Parrinello−Rahman pressure control method.[44] In both types of calculations, P$\bar{1}$ symmetry was used and all cell parameters (cell edges and angles) and internal coordinates were fully relaxed. Overall, good agreement was obtained between the calculated results

**TABLE 14: Comparison of Calculated and Observed Cell Parameters of Alkane and Benzene Molecules**

| molecule | | obsd | | | | | calcd | | |
|---|---|---|---|---|---|---|---|---|---|
| molecule | $T$ (K) | sym | $Z$ | par. | MM | %MM | MD | %MD | ref |
| *n*-pentane | 123 | *Pbcn* | 4 | 4.100 | 4.073 | −0.7 | 4.2 | 2.7 | 96 |
| | | | | 9.076 | 8.572 | −5.6 | 8.817 | −2.9 | |
| | | | | 14.859 | 14.747 | −0.8 | 14.891 | 0.2 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| *n*-hexane | 158 | $P\bar{1}$ | 1 | 4.170 | 4.085 | −2.0 | 4.274 | 2.5 | 92 |
| | | | | 4.700 | 4.394 | −6.5 | 4.551 | −3.2 | |
| | | | | 8.570 | 8.554 | −0.2 | 8.653 | 1.0 | |
| | | | | 96.6 | 96.6 | 0.0 | 96.8 | 0.2 | |
| | | | | 87.2 | 8.3 | 1.3 | 88.9 | 1.9 | |
| | | | | 105.0 | 102.2 | −2.7 | 104.0 | −1.0 | |
| *n*-octane | 193 | $P\bar{1}$ | 1 | 4.220 | 4.084 | −3.2 | 4.297 | 1.8 | 94 |
| | | | | 4.790 | 4.378 | −8.6 | 4.560 | −4.8 | |
| | | | | 11.020 | 11.011 | −0.1 | 11.111 | 0.8 | |
| | | | | 94.7 | 95.1 | 0.4 | 95.3 | 0.6 | |
| | | | | 84.3 | 84.6 | 0.4 | 85.2 | 1.1 | |
| | | | | 105.8 | 102.0 | −3.6 | 104.3 | −1.4 | |
| adamantane | 188 | $P42_1c$ | 2 | 6.600 | 6.556 | −0.7 | 6.620 | 0.3 | 86 |
| | | | | 6.600 | 6.556 | −0.7 | 6.620 | 0.3 | |
| | | | | 8.810 | 8.741 | −0.8 | 8.880 | 0.8 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| anthracene | 95 | $P2_1a$ | 2 | 8.443 | 8.304 | −1.7 | 8.171 | −3.2 | 87 |
| | | | | 6.002 | 6.007 | 0.1 | 6.202 | 3.3 | |
| | | | | 11.124 | 11.074 | −0.4 | 11.210 | 0.8 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| | | | | 125.6 | 125.7 | 0.1 | 125.8 | 0.2 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| anthracene | 290 | $P2_1a$ | 2 | 8.562 | 8.301 | −3.0 | 8.468 | −1.1 | 88 |
| | | | | 6.038 | 6.007 | −0.5 | 6.201 | 2.7 | |
| | | | | 11.184 | 11.071 | −1.0 | 11.331 | 1.3 | |
| | | | | 90.0 | 90.0 | 0.0 | 89.3 | −0.8 | |
| | | | | 124.7 | 125.6 | 0.7 | 125.9 | 1.0 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.5 | 0.6 | |
| benzene | 78 | *Pbca* | 4 | 7.292 | 7.478 | 2.5 | 7.521 | 3.1 | 90 |
| | | | | 9.471 | 9.133 | −3.6 | 9.217 | −2.7 | |
| | | | | 6.742 | 6.756 | 0.2 | 6.865 | 1.8 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| benzene | 295 | *Pbca* | 4 | 7.460 | 7.478 | 0.2 | 7.789 | 4.4 | 90 |
| | | | | 9.660 | 9.133 | −5.5 | 9.129 | −5.5 | |
| | | | | 7.030 | 6.756 | −3.9 | 7.610 | 8.3 | |
| | | | | 90.0 | 90.0 | 0.0 | 89.9 | −0.1 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| | | | | 90.0 | 90.0 | 0.0 | 89.7 | −0.3 | |
| chrysene | 298 | *I2/c* | 4 | 25.732 | 24.732 | −1.9 | 25.571 | 1.5 | 91 |
| | | | | 6.196 | 5.985 | −3.4 | 6.667 | 7.6 | |
| | | | | 8.386 | 8.330 | −0.7 | 8.155 | −2.8 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| | | | | 116.2 | 114.6 | −1.3 | 120.1 | 3.4 | |
| | | | | 90.0 | 90.0 | 0.0 | 89.9 | −0.1 | |
| naphthalene | 298 | $P2_1a$ | 2 | 8.235 | 7.952 | −3.4 | 8.466 | 2.8 | 93 |
| | | | | 6.003 | 5.953 | −0.8 | 5.782 | −3.7 | |
| | | | | 8.658 | 8.591 | −0.8 | 8.325 | −3.8 | |
| | | | | 90.0 | 90.0 | 0.0 | 89.6 | −0.4 | |
| | | | | 122.9 | 123.7 | 0.6 | 112.3 | −8.6 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.1 | 0.1 | |
| ovalene | 298 | $P2_1a$ | 2 | 19.470 | 19.135 | −1.7 | 18.764 | −3.6 | 95 |
| | | | | 4.700 | 4.635 | −1.4 | 4.931 | 4.9 | |
| | | | | 10.120 | 9.907 | −2.1 | 9.930 | −1.9 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.2 | 0.2 | |
| | | | | 105.0 | 104.0 | −0.9 | 103.7 | −1.2 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| perylene | 298 | $P2_1a$ | 4 | 11.277 | 10.491 | −7.0 | 12.002 | 6.4 | 97 |
| | | | | 10.826 | 11.490 | 6.1 | 10.839 | 0.1 | |
| | | | | 10.263 | 9.741 | −5.1 | 9.680 | −5.7 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| | | | | 100.6 | 94.7 | −5.9 | 91.5 | −9.0 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |

**TABLE 14: (Continued)**

| | obsd | | | | calcd | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| molecule | $T$ (K) | sym | $Z$ | par. | MM | %MM | MD | %MD | ref |
| phenanthrene | 298 | $P2_1$ | 1 | 8.472 | 8.515 | 0.5 | 8.709 | 2.8 | 98 |
| | | | | 6.166 | 5.974 | −3.1 | 5.704 | −7.5 | |
| | | | | 9.467 | 9.034 | −4.6 | 9.883 | 4.4 | |
| | | | | 90.0 | 91.5 | 1.6 | 90.1 | 0.1 | |
| | | | | 98.0 | 93.1 | −5.0 | 90.9 | −7.3 | |
| | | | | 90.0 | 91.3 | 1.5 | 90.0 | 0.0 | |
| triphenylene | 298 | $P2_12_12_1$ | 4 | 13.170 | 13.051 | −0.9 | 14.072 | 6.8 | 99 |
| | | | | 16.730 | 16.250 | −2.9 | 17.196 | 2.8 | |
| | | | | 5.260 | 5.201 | −1.1 | 4.910 | −6.6 | |
| | | | | 90.0 | 90.0 | 0.0 | 89.9 | −0.1 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| | | | | 90.0 | 90.0 | 0.0 | 90.0 | 0.0 | |
| no. of data | | | | | 84 | | 84 | | |
| max (%) | | | | | | 6.1 | | 8.3 | |
| min | | | | | | −8.6 | | −9.0 | |
| av (%) | | | | | | −1.1 | | 0.0 | |
| rms (%) | | | | | | 2.2 | | 3.2 | |

**TABLE 15: Comparison of Calculated and Observed Cell Parameters of Crystalline Polyethylene at Different Temperatures**

| | obsd | | | calcd | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| $T$ (K) | $a$ | $b$ | $c$ | $a$ | $b$ | $c$ | ref |
| 4 | 7.121 | 4.851 | 2.548 | 6.961 | 4.745 | 2.592 | 163 |
| 10 | 7.16 | 4.86 | 2.534 | 6.968 | 4.748 | 2.592 | 164 |
| 77 | 7.18 | 4.88 | 2.543 | 7.057 | 4.777 | 2.592 | 164 |
| 77 | 7.155 | 4.899 | 2.5473 | | | | 165 |
| 90 | 7.161 | 4.886 | 2.546 | 7.076 | 4.782 | 2.592 | 163 |
| 195 | 7.27 | 4.91 | 2.534 | 7.240 | 4.829 | 2.591 | 164 |
| 293 | 7.432 | 4.945 | 2.543 | 7.453 | 4.871 | 2.589 | 166 |
| 297 | 7.42 | 4.96 | 2.534 | 7.474 | 4.865 | 2.589 | 164 |
| 297 | 7.40 | 4.93 | 2.534 | | | | 167 |
| 297 | 7.36 | 4.92 | 2.534 | | | | 168 |
| 303 | 7.414 | 4.942 | 2.5473 | 7.503 | 4.862 | 2.589 | 165 |



**Figure 15.** Comparison of calculated and experimental densities of crystalline polyethylene (PE) from 4 to 303 K. The solid line represents the calculated results; the dots denote the experimental data.[163−168]

and the experimental data. Without symmetry constraints, the calculated cell shapes, especially those are not orthorhombic or cubic ($\alpha = \beta = \gamma = 90°$), agree well with the experimental data, in general. In some cases (e.g., *n*-hexane, *n*-octane), the MD results are in better agreement with experimental data than the MM results. Overall, the MM results are more stable (indicated by smaller rms deviations) than the MD simulations.

Crystalline polyethylene has been well-studied using diffraction techniques. In Table 15, the calculated cell parameters are compared with the reported experimental values.[163−168] Again, these calculations were performed using the Parrinello−Raman NPT simulations. A supercell consisting of $3 \times 5 \times 10$ unit cells was used for these simulations. The averaged *c* values were overall larger than those measured by about 2%, the *b* values are smaller than the experimental data by the same percentage, while the *a* values are in good agreement with the experimental data at the high-temperature region (>190 K). In Figure 15, the calculated densities are compared with the experimental data. It appears that between 190 and 300 K, excellent agreement is obtained between the calculated and experimental data. The deviation is larger at the lower temperature region. However, the largest deviation, which is found at the extremely low temperature of 4 K, is about 3%.

## VI. Conclusion

A combined ab initio and empirical parametrization procedure is presented after the functional from and definition of atom types are introduced. The parameters are divided into three categories—valence, charge, and vdW (LJ-9-6) terms. The charge parameters were first derived using a constrained ab initio
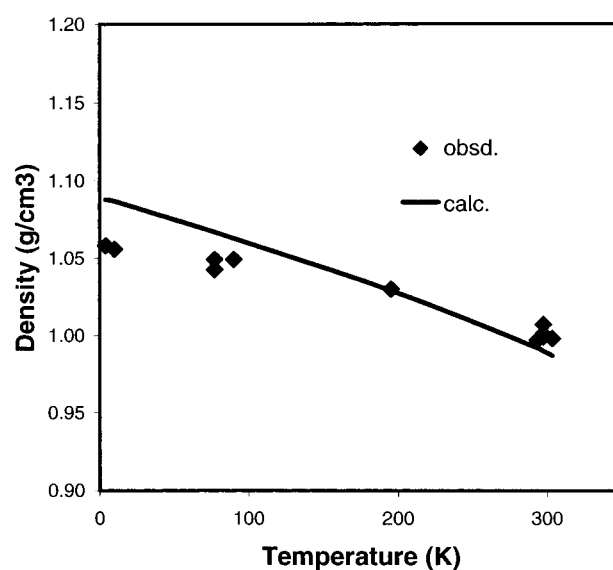
ESP fit. Then with a set of initial vdW parameters fixed, the valence parameters were derived based on ab initio data. The valence parameters were subsequently adjusted empirically to fit the experimental data. Finally, the vdW parameters were optimized using MD simulations of molecular liquids. The consistence of various parts of the force field was ensured by iterating the parametrization procedure. To maximize the coverage and minimize the number of parameters, generic atom types were introduced and parameters transferred extensively. The parametrization was conducted following a precedence tree in which the previously determined parameters were transferred, as many as possible, to the next level of parametrization.

The central part of this work was to optimize the vdW parameters using MD simulations of liquids. Details about the simulation, including results of liquids and crystals, are presented in this paper. One of the critical issues deals with the long-range interactions. It is demonstrated that the long-range electrostatic interaction is canceled out in both liquids and crystals if charge-neutral groups can be defined and the size of the groups is much smaller than the cutoff value. Therefore, it is not necessary to perform expensive calculations such as Ewald summation for such molecular systems. Together with the long-

## APPENDIX A: List of Molecules Calculated

| | | | | | |
|---|---|---|---|---|---|
| alkanes | butane | | dimethylamine | | $NH_4^+$ |
| | cyclobutane | | methylamine | asocynates | HNCO |
| | cyclohexane | | ammonia | | $CH_3NCO$ |
| | cyclopropane | | trimethylamine | | $C_6H_5NCO$ |
| | ethane | amides | *N,N*-dimethylformamide | nitriles | HCN |
| | isobutane | | *Z-N*-methylacetamide | | $CH_3CN$ |
| | methane | | *Z-N*-methylformamide | | $C_6H_5CN$ |
| | butane | | *N*-phenylformamide | nitro derivatives | $HNO_2$ |
| | propane | | acetamide | | $CH_3NO_2$ |
| alkenes | *cis*-1-butene | | diketopiperazine | | $C_6H_5NO_2$ |
| | *cis*-2-butene | | formamide | phosphazenes | $(NPBr_2)_3$ |
| | *cis*-2-pentene | | malonamide | | $(NPCl_2)_3$ |
| | cyclobutene | | oxamide | | $(NPCl_2)_3$ |
| | cyclohexene | amineoxides | $NH_3O$ | | $(NPPh_2)_3$ |
| | cyclopentene | | $CH_3NH_2O$ | | $(NPCl_2)_4$ |
| | 2,3-dimethyl-2-butene | | $(CH_3)_3NO$ | | $(NPF_2)_4$ |
| | ethylene | | $C_6H_5NH_2O$ | | $(NP(NMe_2)_2)_4$ |
| | 1-butene (gauche) | anhydrides | formanhydrides | | $(NP(OMe)_2)_4$ |
| | isobutene | | acetic anhydride | | $(NPCl_2)_5$ |
| | propene | carbamates | carbamic acid | | $(NPF_2)_5$ |
| | 2-butene (trans) | | ethylcarbamic acid | | $(NPCl_2)_7$ |
| | 2-pentene (trans) | | ethylcarbamate | | $(NPF_2)_7$ |
| alkynes | ethylacetylene | | methylcarbamic acid | sacchrides | α-D-glucose |
| | acetylene | | methylcarbamate | | α-2-deoxyribose |
| | methylacetylene | | ethylcarbamic acid | silanes | dimethylsilane |
| | dimethylacetylene | | phenylcarbamate | | dimethyltrisilane |
| benzenes | benzene | carbonates | carbonic acid | | diphenyltrisilane |
| | biphenyl (planar) | | dimethyl carbonate | | disilane |
| | biphenyl (twist) | | diphenyl carbonate | | ethyldisilane |
| | naphthalene | chloroalkanes | 1,4-dichlorobutane | | ethylsilane |
| | *o*-xylene | | 1,3-dichlorobutane | | methyldisilane |
| | toluene | | dichloromethane | | methylsilane |
| ethers | dimethyl ether | | chloromethane | | phenyldisilane |
| | diisopropyl ether | | trichloromethane | | silane |
| | ethyl methyl ether | | tetrachloromethane | | tetrasilane |
| | methylisopropyl ether | | chloroethane | | trisilane |
| | oxanorbornane | | 1,2-dichloroethane | siloxane/zeolites | cyclotrisilica acid |
| | oxetane | | 1,1-dichloroethane | | cycloterasilicic acid |
| | tetrahydrofuran | | 1,1,2,2-tetrachloroethane | | cyclopentasilicic acid |
| alcohols | ethanol | | 1,1,1-trichloroethane | | cyclohexasilicic acid |
| | methanol | | hexachloroethane | | octahydroxyoxydodecasilsesquioxane |
| | phenol | | 2-chloropropane | | dodecahydroxydodecasilasequioxane |
| | water | | 2,2-dichloropropane | | disilicic acid |
| aldehydes | formaldehyde | | 1,3-dichloropropane | | dodalite |
| | acetaldehyde | fluoroalkanes | 1,1,1,3,3,3-hexafluoropropane | | trisilic acid |
| ketones | acetone | | 1,3-difluoropropane | | disiloxane |
| | 2-butanone | | 1,4-difluorobutane | | trisiloxane |
| | cyclobutanone | | 1,1,1,4,4,4-hexafluorobutane | small molecules | CO |
| | cyclohexanone | | 1-fluoroisobutane | | $CO_2$ |
| | cyclopentanone | | 1-fluoropropane | | $CS_2$ |
| | cyclopropanone | | 2,2-difluoropropane | | $H_2$ |
| acids | acetic acid | | 2-fluoropropane | | $N_2$ |
| | formic acid | | octafluoropropane | | NO |
| | propionic acid | | tetrafluoromethane | | $NO_2$ |
| esters | *β*-butyrolactone | | difluoromethane | | $SO_2$ |
| | methyl acetate | | fluoromethane | sulfur organics | $SO_3$ |
| | methyl benzoate | | trifluoromethane | | methanethiol |
| | methyl formate | organic ions | $CO_3^-$ | | dimethyl sulfide |
| | phenyl benzoate | | $H_3O^+$ | | phenylthiol |
| amines | aniline | | | | |

range correction for vdW terms, the nonbond interaction can be accurately evaluated with a reasonably small cutoff value in the range of 8.5−12.5 Å.

This work was strongly influenced by three pioneering works in the force field development: the empirical parametrization of MM3,[1] the ab initio parametrization of CFF93,[7,18] and empirical nonbond parametrization of OPLS.[10] By combining the advantages of these developments, the present force field is parametrized accurately in predicting not only intramolecular properties for molecules in isolation, but also intermolecular properties for molecules in condensed phases. The validation studies based on 178 molecules, 102 liquids, and 69 crystals,

representing 28 molecular classes of most common organic molecules, inorganic gas molecules, and common polymers, show that the calculated structural parameters (bond lengths and angles) are within approximately 1% deviation from the experimental data. The rms deviation of the vibrational frequencies is 41 cm$^{-1}$, and the rms deviation of the conformational energies is 0.38 kcal/mol. These results are comparable with the CFF93[7] and MM3[1] force fields. The predicted liquid properties show a rms deviation of 1.9% for the densities and 4.1% for the cohesive energies (heat of vaporization). These results are similar to those reported using the OPLS force field[10] but represent a much broader coverage of molecules.

## APPENDIX B:  COMPASS Parameters for Alkanes and Benzenes

### Atoms Types

| | | | | | |
|---|---|---|---|---|---|
| c3a | aromatic carbon | c43 | sp³ carbon with three heavy atoms attached | h1 | nonpolar hydrogen |
| c4 | generic sp³ carbon | c44 | sp³ carbon with four heavy atoms attached | | |

### Equivalence Table

| type | nonbond | bond | angle | torsion | out-of-plane | type | nonbond | bond | angle | torsion | out-of-plane |
|---|---|---|---|---|---|---|---|---|---|---|---|
| c3a | c3a | c3a | c3a | c3a | c3a | c44 | c44 | c4 | c4 | c4 | c4 |
| c4 | c4 | c4 | c4 | c4 | c4 | h1 | h1 | h1 | h1 | h1 | h1 |
| c43 | c43 | c4 | c4 | c4 | c4 | | | | | | |

### Bond Increments

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| c3a | c3a | 0.0000 | c3a | h1 | −0.1268 | c4 | h1 | −0.0530 |
| c3a | c4 | 0.0000 | c4 | c4 | 0.0000 | | | |

### Quartic Bond

| I | J | $b_0$ | $k_2$ | $k_3$ | $k_4$ | I | J | $b_0$ | $k_2$ | $k_3$ | $k_4$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| c3a | c3a | 1.4170 | 470.8361 | −627.6179 | 1327.6345 | c4 | c4 | 1.5300 | 299.6700 | −501.7700 | 679.8100 |
| c3a | c4 | 1.5010 | 321.9021 | −521.8208 | 572.1628 | c4 | h1 | 1.1010 | 345.0000 | −691.8900 | 844.6000 |
| c3a | h1 | 1.0982 | 372.8251 | −803.4526 | 894.3173 | | | | | | |

### Quartic Angle

| I | J | K | $q_0$ | $k_2$ | $k_3$ | $k_4$ | I | J | K | $q_0$ | $k_2$ | $k_3$ | $k_4$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| c3a | c3a | c3a | 118.9000 | 61.0226 | −34.9931 | 0.0000 | c3a | c4 | h1 | 111.0000 | 44.3234 | −9.4454 | 0.0000 |
| c3a | c3a | c4 | 120.0500 | 44.7148 | −22.7352 | 0.0000 | c4 | c4 | c4 | 112.6700 | 39.5160 | −7.4430 | −9.5583 |
| c3a | c3a | h1 | 117.9400 | 35.1558 | −12.4682 | 0.0000 | c4 | c4 | h1 | 110.7700 | 41.4530 | −10.6040 | 5.1290 |
| c3a | c4 | c3a | 111.0000 | 44.3234 | −9.4454 | 0.0000 | h1 | c4 | h1 | 107.6600 | 39.6410 | −12.9210 | −2.4318 |
| c3a | c4 | c4 | 108.4000 | 43.9594 | −8.3924 | −9.3379 | | | | | | | |

### Torsion

| I | J | K | L | $k_1$ | $k_2$ | $k_3$ | I | J | K | L | $k_1$ | $k_2$ | $k_3$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| c3a | c3a | c3a | c3a | 8.3667 | 1.2000 | 0.0000 | c3a | c3a | c4 | h1 | −0.2802 | −0.0678 | −0.0122 |
| c3a | c3a | c3a | c4 | 0.0000 | 4.4072 | 0.0000 | c3a | c4 | c4 | h1 | −0.0228 | 0.0280 | −0.1863 |
| c3a | c3a | c3a | h1 | 0.0000 | 3.9661 | 0.0000 | c4 | c4 | c4 | c4 | 0.0000 | 0.0514 | −0.1430 |
| c4 | c3a | c3a | h1 | 0.0000 | 1.5590 | 0.0000 | c4 | c4 | c4 | h1 | 0.0000 | 0.0316 | −0.1681 |
| h1 | c3a | c3a | h1 | 0.0000 | 2.3500 | 0.0000 | h1 | c4 | c4 | h1 | −0.1432 | 0.0617 | −0.1530 |
| c3a | c3a | c4 | c3a | −0.2802 | −0.0678 | −0.0122 | * | c3a | c3a | * | 0.0000 | 4.5000 | 0.0000 |
| c3a | c3a | c4 | c4 | −0.2802 | −0.0678 | −0.0122 | * | c4 | c4 | * | 0.0000 | 0.0000 | −0.1530 |

### Out of Plane

| I | J | K | L | $k_2$ | I | J | K | L | $k_2$ |
|---|---|---|---|---|---|---|---|---|---|
| c3a | c3a | c3a | c3a | 7.1794 | c3a | c3a | c3a | h1 | 4.8912 |
| c3a | c3a | c3a | c4 | 7.8153 | | | | | |

### Nonbond (LJ-9-6)

| I | $r_0$ | $e_0$ | I | $r_0$ | $e_0$ | I | $r_0$ | $e_0$ |
|---|---|---|---|---|---|---|---|---|
| c3a | 3.9150 | 0.0680 | c43 | 3.8540 | 0.0400 | h1 | 2.8780 | 0.0230 |
| c4 | 3.8540 | 0.0620 | c44 | 3.8540 | 0.0200 | | | |

### Bond—Bond

| I | J | K | I−J/J−K | I | J | K | I−J/J−K | I | J | K | I−J/J−K |
|---|---|---|---|---|---|---|---|---|---|---|---|
| c3a | c3a | c3a | 68.2856 | c3a | c3a | h1 | 1.0795 | c4 | c4 | h1 | 3.3872 |
| c3a | c3a | c4 | 12.0676 | c3a | c4 | h1 | 2.9168 | h1 | c4 | h1 | 5.3316 |

### Bond—Bond (1,3)

| I | J | K | L | I−J/K−L | I | J | K | L | I−J/K−L | I | J | K | L | I−J/K−L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| c3a | c3a | c3a | c3a | 53.0000 | c3a | c3a | c3a | h1 | −6.2741 | h1 | c3a | c3a | h1 | −1.7077 |
| c3a | c3a | c3a | c4 | 2.5085 | c4 | c3a | c3a | h1 | 0.8743 | h1 | c3a | c3a | c4 | −3.4826 |

### Bond—Angle

| I | J | K | I−J/I−J−K | J−K/I−J−K | I | J | K | I−J/I−J−K | J−K/I−J−K | I | J | K | I−J/I−J−K | J−K/I−J−K |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| c3a | c3a | c3a | 28.8708 | 28.8708 | c3a | c4 | h1 | 26.4608 | 11.7717 | c4 | c4 | h1 | 20.7540 | 11.4210 |
| c3a | c3a | c4 | 31.0771 | 47.0579 | c4 | c4 | c4 | 8.0160 | 8.0160 | h1 | c4 | h1 | 18.1030 | 18.1030 |
| c3a | c3a | h1 | 20.0033 | 24.2183 | | | | | | | | | | |

### Angle-Angle

| I | J | K | L | I−J−K/J−K−L | I | J | K | L | I−J−K/J−K−L | I | J | K | L | I−J−K/J−K−L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| c4 | c4 | c3a | h1 | 2.0403 | c4 | c4 | c4 | h1 | −1.3199 | c4 | c4 | h1 | c4 | 0.1184 |
| h1 | c4 | c3a | h1 | 3.0118 | h1 | c4 | c4 | h1 | −0.4825 | c4 | c4 | h1 | h1 | 0.2738 |
| c3a | c4 | c4 | h1 | −1.8202 | c3a | c4 | h1 | c4 | 1.0827 | h1 | c4 | h1 | h1 | −0.3157 |
| c4 | c4 | c4 | c4 | −0.1729 | c3a | c4 | h1 | h1 | 2.3794 | | | | | |

**APPENDIX B: (Continued)**

End Bond−Torsion

| I | J | K | L | $k_1$ (I−J/I−J−K−L) | $k_2$ | $k_3$ | $k_1$ (K−L/I−J−K−L) | $k_2$ | $k_3$ |
|---|---|---|---|---|---|---|---|---|---|
| c3a | c3a | c3a | c3a | −0.1185 | 6.3204 | 0.0000 | −0.1185 | 6.3204 | 0.0000 |
| c3a | c3a | c3a | c4 | 0.0000 | −0.6918 | 0.0000 | 0.0000 | 0.2421 | 0.0000 |
| c3a | c3a | c3a | h1 | 0.0000 | −6.8958 | 0.0000 | 0.0000 | −0.4669 | 0.0000 |
| c4 | c3a | c3a | h1 | 0.0000 | −1.7970 | 0.0000 | 0.0000 | −0.4879 | 0.0000 |
| h1 | c3a | c3a | h1 | 0.0000 | −0.6890 | 0.0000 | 0.0000 | −0.6890 | 0.0000 |
| c3a | c3a | c4 | c4 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| c3a | c3a | c4 | h1 | −0.5835 | 1.1220 | 0.3978 | 1.3997 | 0.7756 | 0.0000 |
| c3a | c4 | c4 | h1 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| c4 | c4 | c4 | c4 | −0.0732 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| c4 | c4 | c4 | h1 | 0.2486 | 0.2422 | −0.0925 | 0.0814 | 0.0591 | 0.2219 |
| h1 | c4 | c4 | h1 | 0.2130 | 0.3120 | 0.0777 | 0.2130 | 0.3120 | 0.0777 |

Middle Bond−Torsion

| I | J | K | L | $k_1$ | $k_2$ | $k_3$ | I | J | K | L | $k_1$ | $k_2$ | $k_3$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| c3a | c3a | c3a | c3a | 27.5989 | −2.3120 | 0.0000 | c3a | c3a | c4 | h1 | −5.5679 | 1.4083 | 0.3010 |
| c3a | c3a | c3a | c4 | 0.0000 | 9.1792 | 0.0000 | c3a | c4 | c4 | h1 | 0.0000 | 0.0000 | 0.0000 |
| c3a | c3a | c3a | h1 | 0.0000 | −1.1521 | 0.0000 | c4 | c4 | c4 | c4 | −17.7870 | −7.1877 | 0.0000 |
| c4 | c3a | c3a | h1 | 0.0000 | 3.9421 | 0.0000 | c4 | c4 | c4 | h1 | −14.8790 | −3.6581 | −0.3138 |
| h1 | c3a | c3a | h1 | 0.0000 | 4.8228 | 0.0000 | h1 | c4 | c4 | h1 | −14.2610 | −0.5322 | −0.4864 |
| c3a | c3a | c4 | c4 | 0.0000 | 0.0000 | 0.0000 | | | | | | | |

Angle−Torsion

| I | J | K | L | $k_1$ (I−J−K/I−J−K−L) | $k_2$ | $k_3$ | $k_1$ (J−K−L/I−J−K−L) | $k_2$ | $k_3$ |
|---|---|---|---|---|---|---|---|---|---|
| c3a | c3a | c3a | c3a | 1.9767 | 1.0239 | 0.0000 | 1.9767 | 1.0239 | 0.0000 |
| c3a | c3a | c3a | c4 | 0.0000 | 3.8987 | 0.0000 | 0.0000 | −4.4683 | 0.0000 |
| c3a | c3a | c3a | h1 | 0.0000 | 2.5014 | 0.0000 | 0.0000 | 2.7147 | 0.0000 |
| c4 | c3a | c3a | h1 | 0.0000 | −0.1242 | 0.0000 | 0.0000 | 3.4601 | 0.0000 |
| h1 | c3a | c3a | h1 | 0.0000 | 2.4501 | 0.0000 | 0.0000 | 2.4501 | 0.0000 |
| c3a | c3a | c4 | c4 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| c3a | c3a | c4 | h1 | 0.2251 | 0.6548 | 0.1237 | 4.6266 | 0.1632 | 0.0461 |
| c3a | c4 | c4 | h1 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| c4 | c4 | c4 | c4 | 0.3886 | −0.3139 | 0.1389 | 0.3886 | −0.3139 | 0.1389 |
| c4 | c4 | c4 | h1 | −0.2454 | 0.0000 | −0.1136 | 0.3113 | 0.4516 | −0.1988 |
| h1 | c4 | c4 | h1 | −0.8085 | 0.5569 | −0.2466 | −0.8085 | 0.5569 | −0.2466 |

Angle-Angle Torsion

| I | J | K | L | IJ/IJKL/JKL | I | J | K | L | IJ/IJKL/JKL |
|---|---|---|---|---|---|---|---|---|---|
| c3a | c3a | c3a | c4 | −14.4097 | c3a | c3a | c4 | h1 | −5.8888 |
| c3a | c3a | c3a | h1 | −4.8141 | c4 | c4 | c4 | c4 | −22.0450 |
| c4 | c3a | c3a | h1 | 4.4444 | c4 | c4 | c4 | h1 | −16.1640 |
| h1 | c3a | c3a | h1 | 0.3598 | h1 | c4 | c4 | h1 | −12.5640 |

Detailed results of the parametrization and validation for alkane and benzene compounds are presented. The parameters of these two functional groups are most widely transferred in a general force field, so it is very important to have these groups well represented. Validation results show that both functional groups have been parametrized very well. In addition, it is demonstrated that the present force field can be extrapolated to predict thermophysical properties outside of the parametrization region. Numerous calculations and applications based on other molecules containing alkyl or phenyl groups indicate that the parameters for alkyl and phenyl groups are well transferable as well.[24−28]

This work demonstrates that the simple functional forms, particularly those used for the nonbond interactions (Coulombic and LJ function), are highly capable for the most common and basic applications in condensed phases. Although more complicated functional forms, such as polarization models, are useful for a detailed description of the molecular interactions, the power of atomistic simulation using current force-field technology has not been completely explored yet, where a better parametrization method is one of the key issues. The present work is intended to make a contribution to this end.

**References and Notes**

(1) (a) Lii, J. H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8576. (b) Allinger, N. L.; Li, F. B.; Yan, L. Q. *J. Comput. Chem.* **1990**, *11*, 848. (c) Allinger, N. L.; Rahman, M.; Lii, J. H. *J. Am. Chem. Soc.* **1990**, *112*, 8293. (d) Allinger, N. L.; Li, F. B.; Yan, L. Q.; Tai, J. C. *J. Comput. Chem.* **1990**, *11*, 868. (e) Allinger, N. L.; Rahman, M.; Lii, J. H. *J. Am. Chem. Soc.* **1990**, *112*, 8293. (f) Schmitz, L. R.; Allinger, N. L. *J. Am. Chem. Soc.* **1990**, *112*, 8307. (g) Lii, J. H.; Allinger, N. L. *J. Comput. Chem.* **1991**, *12*, 186. (h) Lii, J. H.; Chen, K. S.; Rahman, M.; Pathiaseril, A. *J. Am. Chem. Soc.* **1991**, *113*, 4505. (i) Allinger, N. L.; Quinn, M.; Rahman, M.; Chen, K. H. *J. Phys. Org. Chem.* **1991**, *4*, 647. (j) Chen, K. H.; Allinger, N. L. *J. Phys. Org. Chem.* **1991**, *4*, 659. (k) Allinger, N. L. *Reviews in Computational Chemistry;* Lipkowitzand, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1992; Vol. 2. (l) Cui, W. L.; Li, F. B.; Allinger, N. L. *J. Am. Chem. Soc.* **1993**, *115*, 2943.

(2) (a) Allinger, N. L.; Chen, K. S.; Katzenellenbogen, J. A.; Wilson, S. R.; Anstead, G. M. *J. Comput. Chem.* **1996**, *17*, 747. (b) Allinger, N. L.; Chen, K. S.; Lii, J. H. *J. Comput. Chem.* **1996**, *17*, 642. (c) Nevins, N.;

COMPASS—Details on Alkane and Benzene Compounds

*J. Phys. Chem. B, Vol. 102, No. 38, 1998* **7363**

Allinger, N. L. *J. Comput. Chem.* **1996**, *17*, 730. (d) Nevins, N.; Chen, K. S.; Allinger, N. L. *J. Comput. Chem.* **1996**, *17*, 669. (e) Nevins, N.; Lii, J. H.; Allinger, N. L. *J. Comput. Chem.* **1996**, *17*, 695.

(3) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. *J. Phys. Chem.* **1990**, *94*, 8897.

(4) (a) Allured, V. S.; Kelly, C. M.; Landis, C. R. *J. Am. Chem. Soc.* **1991**, *113*, 1. (b) Bearpark, M. J.; Robb, M. A.; Bernardi, F.; Olivucci, M. *Chem. Phys. Lett.* **1994**, *217*, 513.

(5) (a) Root, D. M.; Landis, C. R.; Cleveland, T. *J. Am. Chem. Soc.* **1993**, *115*, 4201. (b) Cleveland, T.; Landis, C. R. *J. Am. Chem. Soc.* **1996**, *118*, 6020.

(6) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024.

(7) (a) Hwang, M. J.; Stockfisch, T. P.; Hagler, A. T. *J. Am. Chem. Soc.* **1994**, *116*, 2515. (b) Maple, J. R.; Hwang, M. J.; Stockfisch, T. P.; Dinur, U.; Waldman, M.; Ewig, C. S.; Hagler, A. T. *J. Comput. Chem.* **1994**, *15*, 162. (c) Maple, J. R.; Hwang, M. J.; Stockfisch, T. P.; Hagler, A. T. *Isr. J. Chem.* **1994**, *34*, 195. (d) Peng, Z. W.; Ewig, C. S.; Hwang, M. J.; Waldman, M.; Hagler, A. T. *J. Phys. Chem. A* **1997**, *101*, 7243.

(8) (a) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179. (b) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1996**, *118*, 2309.

(9) Mackerell A. D.; Wiorkiewiczkuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946.

(10) Jorgensen, W. L.; Maxwell, D. S.; Tiradorives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(11) (a) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490. (b) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 520. (c) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 553. (d) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 616. (e) Halgren, T. A.; Nachbar, R. B. *J. Comput. Chem.* **1996**, *17*, 587.

(12) (a) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 6638. (b) Jorgensen, W. L.; Swenson, C. J. *J. Am. Chem. Soc.* **1985**, *107*, 569. (c) Jorgensen, W. L. *J. Phys. Chem.* **1986**, *90*, 6379. (d) Jorgensen, W. L. *J. Phys. Chem.* **1986**, *90*, 1276. (e) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.

(13) Lifson, S.; Warshel, A. *J. Chem. Phys.* **1968**, *49*, 5116.

(14) Burkert, U.; Allinger, N. L. *Molecular Mechanics*; Caserio, M. C., Ed.; ACS Monograph 77; American Chemical Society: Washington, DC, 1989.

(15) This work has been supported by MSI (formerly Biosym) Polymer Consortium since 1989.

(16) (a) Sun, H. *J. Comput. Chem.* **1994**, *15*, 752. (b) Sun, H.; Mumby, S. J.; Maple, J. R.; Hagler, A. T. *J. Am. Chem. Soc.* **1994**, *116*, 2978. (c) Sun, H. *Macromolecules* **1995**, *28*, 701. (d) Sun, H. *Macromolecules* **1994**, *26*, 5924. (e) Sun, H.; Mumby, S. J.; Maple, J. R.; Hagler, A. T. *J. Phys. Chem.* **1995**, *99*, 5873.

(17) PCFF and COMPASS are distributed along with MSI's software package InsightII and Ceries2.

(18) Maple, J. R.; Dinur, U.; Hagler, A. T. *Proc. Natl. Acad. Sci.* **1988**, *85*, 5350.

(19) Thacher, T.; Hagler, A. T. Internal communication based on unpublished work.

(20) (a) Halger, A. T.; Huler, E.; Lifson, S. *J. Am. Chem. Soc.* **1974**, *96*, 5319. (b) Halger, A. T.; Lifson, S. *J. Am. Chem. Soc.* **1974**, *96*, 5327.

(21) (a) Lifson, S.; Halger, A. T.; Dauber, P. *J. Am. Chem. Soc.* **1979**, *101*, 5111. (b) Hagler, A. T.; Lifson, S.; Dauber, P. *J. Am. Chem. Soc.* **1979**, *101*, 5122. (c) Halger, A. T.; Dauber, P.; Lifson, S. *J. Am. Chem. Soc.* **1979**, *101*, 5131.

(22) For a general review on the topic of nonbond parametrization using lattice information, see: Pertsin, A. J.; Kitaigorodsky, A. I. *The Atom-Atom Potential Method—Applications to Organic Molecular Solids;* Springer Series in Chemical Physics; Springer: New York, 1987; Vol. 43.

(23) Ridgy, D. Internal communication based on unpublished work.

(24) Sun, H.; Rigby, D. *Spectrochim. Acta, Parb A* **1997**, *53*, 1301.

(25) Rigby, D.; Sun, H.; Eichinger, B. E. *Polym. Int.* **1997**, *44*, 311.

(26) Sun, H.; Fried, J. R.; Ren, P. *Comput. Theor. Polym. Sci.* **1998**, in print.

(27) Yian, G. S.; Tian, A.; Yang, J.; Wu, D. Y.; Ren, Y.; Xiue, Y.; Li, Z. R.; Sun, H. manuscript in preparation.

(28) A series of publications on parametrization and application for individual functional groups is being prepared.

(29) Halgren, T. A. *J. Am. Chem. Soc.* **1992**, *114*, 7827.

(30) Waldman, M.; Hagler, A. T. *J. Comput. Chem.* **1993**, *14*, 1077.

(31) Michael, D. W.; Dykstra, C. E.; Lisy, J. M. *J. Chem. Phys.* **1984**, *81*, 2535.

(32) Trular, D. G.; Schwenke, W. *J. Chem. Phys.* **1985**, *82*, 2418.

(33) (a) Bohm, H.-J.; Ahlrichs, R. *J. Chem. Phys.* **1982**, *77*, 2028. (b) Bohm, H.-J.; Ahlrichs, R.; Scharf, P.; Schiffer, H. *J. Chem. Phys.* **1984**, *81*, 1389. (c) Sagarik, K. P.; Ahlrichs, R. *J. Chem. Phys.* **1987**, *86*, 5117.

(34) Gough, C. A.; Debolt, S. E.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 963.

(35) Widmann, A. H.; Laso, M.; Suter, U. W. *J. Chem. Phys.* **1995**, *102*, 5761.

(36) (a) Williams, D. E. *J. Comput. Chem.* **1988**, *9*, 745. (b) Williams, D. E. *Biopolymers* **1990**, *29*, 1367.

(37) Hariharam, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

(38) (a) Pople, J. A.; Seeger, R.; Krishnan, R. *Int. Quantum Chem.* **1977**, *11*, 149. (b) Krishnan, R.; Pople, J. A. *Int. Quantum Chem.* **1978**, *14*, 91.

(39) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571.

(40) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules;* Oxford University: New York, 1989.

(41) Vosko, S. J.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.

(42) Belley, B. *J. Chem. Phys.* **1990**, *92*, 508.

(43) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

(44) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182.

(45) Andrea, T. A.; Swope, W. C.; Andersen, H. C. *J. Chem. Phys.* **1983**, *79*, 4576.

(46) Discover is a molecular mechanics software package distributed by MSI.

(47) Tosi, M. P. *Solid State Phys.* **1964**, *16,* 107.

(48) Ewald, P. P. *Ann. Phys.* **1921**, *64*, 253.

(49) Greegard, L.; Rokhlin, V. I. *J. Comput. Phys.* **1987**, *73*, 2848.

(50) Schmidt, K. E.; Lee, M. A. *J. Stat. Phys.* **1991**, *63*, 1223.

(51) Ding, H. Q.; Karasawa, N.; Goddard, W. A. *J. Chem. Phys.* **1992**, *97*, 4309.

(52) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids;* Oxford University Press: Oxford, 1989.

(53) Jackson, J. D. *Classical Electrodynamics,* 2nd ed.; John Wiley & Sons: New York, 1975.

(54) *CRC Handbook of Chemistry and Physics,* 5th ed.; Lide, D. R., Ed.; CRC Press, Inc.: Boca Raton, FL, 1994.

(55) DFT VWN/DNP results. This work.

(56) *Physical Property Data for the Design Engineer;* Beaton, C. F.; Hewitt, G. F., Eds.; Hemisphere Publishing Co.: New York, 1989.

(57) Barton, A. F. M. *CRC Handbook of Solubility Parameters and Other Cohesive Parameters*; CRC Press: Boca Raton, FL, 1983.

(58) *Physical Constants of hydrocarbon and non-hydrocarbon compounds*; ASTM data series DS 4B; American Petroleum Institute: Philadelphia, PA, 1988.

(59) *Gmelin Handbook of Inorganic Chemistry,* 8th ed.; Springer-Verlag: New York, 1982.

(60) Scott, D. W.; Messerly, J. F.; Todd, S. S.; Guthrie, G. B.; Hossenlopp, I. A.; Moore, R. T.; Osborne, A.; Berg, W. T.; McCullough, J. P. *J. Phys. Chem.* **1961**, *65*, 1320.

(61) Beevers, M. S.; Mumby, S. J.; Clarson, S. J.; Semlyen, J. A. *Polymer* **1983**, *24*, 1565.

(62) Flaningam, O. L. *J. Chem. Eng. Data* **1986**, *31*, 266.

(63) Ward, L. G.; MacDiarmid, A. G. *J. Am. Chem. Soc.* **1960**, *82*, 2151.

(64) Allcock, H. R. *Phosphorus—Nitrogen Compounds;* Academic Press: New York, 1972 and references therein.

(65) Jönsson, P.-G. *Acta Crystallogr.* **1971**, *B27*, 893.

(66) Housty, P.; Hospital, M. *Acta Crystallogr.* **1965**, *18*, 693.

(67) Derissen, J. L.; Smit, P. H. *Acta Crystallogr.* **1974**, *B30*, 2240.

(68) Broadley, J. S.; Cruickshank, D. W. J.; Morrison, J. D.; Robertson, J. M.; Shearer, H. M. M. *Proc. R. Soc. London* **1959**, *A251*, 441.

(69) Strieter, F. J.; Templeton, D. H. *Acta Crystallogr.* **1962**, *15*, 1240.

(70) Nahringbauer, I. *Acta Crystallogr.* **1978**, *B34*, 315.

(71) Adams, J. M.; Morsi, S. E. *Acta Crystallogr.* **1976**, *B32*, 1345.

(72) Strieter, F. J.; Templeton, D. H. *Acta Crystallogr.* **1962**, *15*, 1233.

(73) Housty, P.; Hospital, M. *Acta Crystallogr.* **1966**, *20*, 325.

(74) Housty, P.; Hospital, M. *Acta Crystallogr.* **1965**, *18*, 753.

(75) Scheuerman, R. F.; Sass, R. L. *Acta Crystallogr.* **1962**, *15*, 1244.

(76) Park, Y. J.; Jeffrey, G. A.; Hamilton, W. C. *Acta Crystallogr.* **1971**, *B27*, 2393.

(77) Kim, H. S.; Jeffrey, G. A.; Rosenstein, R. D. *Acta Crystallogr.* **1968**, *B24*, 1449.

(78) Brown, C. J. *Acta Crystallogr.* **1966**, *21*, 170.

(79) Ceccarelli, C.; Jeffrey, G. A.; McMullan, R. K. *Acta Crystallogr.* **1980**, *B36*, 3079.

(80) Jönsson, P.-G. *Acta Crystallogr.* **1976**, *B32*, 232.

(81) Berman, H. M.; Rosenstein, R. D. *Acta Crystallogr.* **1968**, *B24*, 435.

(82) Maartmann-Moe, K. *Acta Crystallogr.* **1966**, *21*, 979.

(83) Llewellyn, F. J.; Cox, E. G.; Goodwin, T. H. *J. Chem Soc.* **1937**, 883.

(84) Kim, H. S.; Jeffrey, G. A.; Rosenstein, R. D. *Acta Crystallogr.* **1969**, *B25*, 2223.

(85) Kim, H. S.; Jeffrey, G. A. *Acta Crystallogr.* **1969**, *B25*, 2607.

(86) Nordman, C. E.; Schmitkons, D. L. *Acta Crystallogr.* **1965**, *18*, 764.

(87) Mason, R. *Acta Crystallogr.* **1964**, *17*, 547.

(88) Mathieson, A.; Robertson, J. M.; Singlair, V. C. *Acta Crystallogr.* **1950**, *3*, 245.

(89) Cox, E. G.; Gruickshank, D. W. J.; Smith, J. A. S. *Proc. R. Soc. London* **1958**, *A247*, 1.

(90) Kozhin, V. M. *Zh. Fiz. Khim.* **1954**, *28*, 566.

(91) Burns, D. M.; Iball, J. *Proc. R. Soc. London* **1960**, *A257*, 491.

(92) Norman, N.; Mathisen, H. *Acta Chem. Scand.* **1961**, *15*, 1755.

(93) Gruickshank, D. W. J. *Acta Crystallogr.* **1957**, *10*, 504.

(94) Mathisen, H.; Norman, N.; Pedersen, B. F. *Acta Chem. Scand.* **1967**, *21*, 127.

(95) Donaldson, D. M.; Robertson, J. M. *Proc. R. Soc. London* **1953**, *A220*, 157.

(96) Mathisen, H.; Norman, N.; Pedersen, B. F. *Acta Chem. Scand.* **1967**, *21*, 127.

(97) Cammerman, A.; Trotter, J. *Proc. R. Soc. London* **1964**, *A279*, 129.

(98) Kay, M. I.; Okaya, Y.; Cox, D. E. *Acta Crystallogr.* **1971**, *B27*, 26.

(99) Ahmed, F. R.; Trotter, J. *Acta Crystallogr.* **1963**, *16*, 503.

(100) Avitabile, G.; Napolitano, R.; Pirozzi, B.; Rouse, K. D.; Thomas, H. W.; Wills, B. T. M. *J. Polym. Sci., Polym. Lett. Ed.* **1975**, *13*, 351.

(101) Blake, C. C. F.; Small, R. W. H. *Acta Crystallogr.* **1972**, *B28*, 2201.

(102) Degeilh, R.; Marsh, R. E. *Acta Crystallogr.* **1959**, *12*, 1007.

(103) Stevens, E. D. *Acta Crystallogr.* **1978**, *B34*, 544.

(104) Vhieh, P. C.; Subramanian, E.; Trotter, J. *J. Chem. Soc. A* **1970**, 179.

(105) Ayerst, E. M.; Duke, J. R. C. *Acta Crystallogr.* **1954**, *7*, 588.

(106) Davies, D. R.; Pasternak, R. A. *Acta Crystallogr.* **1956**, *9*, 334.

(107) Nimmo, J. K.; Lucas, B. W. *Acta Crystallogr.* **1976**, *B32*, 348.

(108) Andersen, A. F. *Acta Crystallogr.* **1957**, *10*, 107.

(109) Amzel, L. M.; Baggio, S.; Baggio, R. F.; Becka, L. N. *Acta Crystallogr.* **1974**, *B30*, 2494.

(110) Smith, W. L.; Ekstrand, J. D.; Raymond, K. N. *J. Am. Chem. Soc.* **1978**, *100*, 3539.

(111) Atoji, M.; Lipscomb, W. N. *Acta Crystallogr.* **1953**, *6*, 770.

(112) Samet-Delcroix, P. S. *Acta Crystallogr.* **1973**, *B29*, 977.

(113) Blake, A. J.; Ebsworth, E. A. V.; Welch, A. J. *Acta Crystallogr.* **1984**, *C40*, 413.

(114) Binnie, W. P.; Robertson, J. M. *Acta Crystallogr.* **1950**, *3*, 424.

(115) Bracher, B. H.; Small, R. W. H. *Acta Crystallogr.* **1967**, *23*, 410.

(116) Perez, S.; Scaringe, R. P. *Macromolecules* **1987**, *20*, 68.

(117) Gardner, K. H.; Blackwell, J. *Acta Crystallogr.* **1980**, *B36*, 1972.

(118) Andre, D.; Fourme, R.; Zechmeister, K. *Acta Crystallogr.* **1972**, *B28*, 2389.

(119) Busetti, V.; Del Pra, A.; Mammi, M. *Acta Crystallogr.* **1969**, *B25*, 1191.

(120) Hasagawa, R.; Takahashi, Y.; Chatani, Y.; Tadokoro, H. *Polymer* **1972**, *3*, 600.

(121) Takahashi, Y.; Matsubara, Y.; Tadokoro, H. *Macromolecules* **1983**, *16*, 1588.

(122) Weinhold, S.; Litt, M. H.; Lando, J. B. *Macromolecules* **1980**, *13*, 1178.

(123) Baachamann, M.; Gordon, W. L.; Weinhold, S.; Lando, J. B. *J. Appl. Phys.* **1980**, *51*, 5095.

(124) Brown, G. M.; Levy, H. A. *Acta Crystallogr.* **1979**, *B35*, 656.

(125) Chu, S. S. C.; Jeffrey, G. A. *Acta Crystallogr.* **1968**, *B24*, 830.

(126) Hirotsu, K.; Shimada, A. *Bull. Chem. Soc. Jpn.* **1974**, *47*, 1872.

(127) Brown, G. M.; Levy, H. A. *Acta Crystallogr.* **1973**, *B29*, 790.

(128) Barrow, M. J.; Ebsworth, E. A. V.; Harding, M. M. *Acta Crystallogr.* **1979**, *B35*, 2093.

(129) Wilson, A.; Carroll, D. F. *J. Chem. Soc. London* **1960**, 2548.

(130) Wagner, A. J.; Vos Aafje, *Acta Crystallogr.* **1968**, *B24*, 707.

(131) Chatani, Y.; Yatsuyanagi, K. *Macromolecules* **1987**, *20*, 1042.

(132) Meille, S. V.; Poletti, A. R.; Gallazzi, M. C.; Gleria, M.; Bruckner, S. *Polymer* **1992**, *33*, 2364.

(133) Shipman, L. L.; Burgess, A. W.; Sheraga, H. A. *J. Phys. Chem.* **1976**, *80*, 52.

(134) De Kruif, C. G. *J. Chem. Thermodyn.* **1980**, *12*, 243.

(135) Gigli, R.; Malaspina, L.; Bardi, G. *Ann. Chim.* **1973**, *63*, 627.

(136) Alberty, R. A.; Chung, M. B.; Reif, A. K. *J. Phys. Chem. Ref. Data* **1990**, *19*, 349.

(137) Bartell, L. S.; Kuchitsu, K.; deNeui, R. *J. Chem. Phys.* **1961**, *35*, 1211.

(138) Bartell, L. S.; Higginbotham, H. K. *J. Chem. Phys.* **1965**, *42*, 851.

(139) Bradford, W. F.; Fitzwater, S.; Bartell, L. S. *J. Mol. Struct.* **1977**, *38*, 185.

(140) Hilderbrandt, R. L.; Weiser, J. D. *J. Mol. Struct.* **1973**, *15*, 27.

(141) Yamamoto, S.; Nakata, M.; Fukuyama, T.; Kuchitsu, K. *J. Phys. Chem.* **1985**, *89*, 3298.

(142) Egawa, T.; Fukuyama, T.; Yamamoto, S.; Tabayashi, F.; Kambara, H.; Ueda, T.; Kuchitsu, K. *J. Chem. Phys.* **1987**, *86*, 6018.

(143) Bastiansen, O.; Fernholt, L.; Seip, H. M.; Kambara, H.; Kuchitsu, K. *J. Mol. Struct.* **1973**, *18*, 163.

(144) Tamagawa, K.; Iijima, T.; Kimura, M. *J. Mol. Struct.* **1976**, *30*, 243.

(145) Ketkar, S. N.; Fink, M. *J. Mol. Struct.* **1981**, *77*, 139.

(146) Rudolph, H. D.; Walzer, K.; Krutzik, I. *J. Mol. Spectrosc.* **1973**, *47*, 314.

(147) Almenningen, A. *J. Mol. Struct.* **1985**, *128*, 59.

(148) Duncan, J. L.; Mills, I. M. *Spectrchim. Acta* **1964**, *20*, 523.

(149) Schachtschneider, J. H.; Snyder, R. G. *Spectrochim. Acta* **1963**, *19*, 117.

(150) Varsanyi, G.; Szoke, S. *Vibrational Spectra of Benzene Derivatives*; Academic Press: New York, 1969.

(151) H irota, E.; Endo, Y.; Saito, S.; Duncan, J. L. *J. Mol. Spectrosc.* **1981**, *89*, 285.

(152) Pitzer, K. S. *J. Chem. Phys.* **1944**, *12*, 310.

(153) Lide, D. R.; Mann, D. E. *J. Chem. Phys.* **1958**, *29*, 914.

(154) Durig, J. R.; Craven, S. M.; Harris, W. C. *Vibrational Spectra and Structure*; Marcel Dekker: New York, 1978; Vol. I.

(155) Compton, D. A.; Montero, S. M.; Murphy, W. F. *J. Phys. Chem.* **1980**, *84*, 3587.

(156) Durig, J. R.; Compton, D. A. C. *J. Phys. Chem.* **1979**, *83*, 265.

(157) Maissara, M.; Cornut, J.C.; Devaure, J.; Lascombe, J. *J. Spectrosc. Int. J.* **1983**, *104*, 2.

(158) Kanesaka, I.; Snyder, R. G.; Strauss, H. L. *J. Chem. Phys.* **1986**, *84*, 395.

(159) Almenningen, A. *J. Mol. Struct.* **1985**, *128*, 59.

(160) Ab initio result, this work.

(161) Harris, K. R. *J. Chem. Soc., Faraday Trans. 1* **1982**, *78*, 2265.

(162) Gehrig, M.; Lentz, H. *J. Chem. Thermodyn.* **1977**, *9*, 445.

(163) Avitabile, G.; Napolitano, R.; Pirozzi, B.; Rouse, K. D.; Thomas, M. W.; Willis, B. T. M. *J. Polym. Sci. B* **1975**, *13*, 351.

(164) Shen, M.; Hansen, W. N.; Romo, P. C. *J. Chem. Phys.* **1969**, *51*, 425.

(165) Swan, P. R. *J. Polym. Sci.* **1962**, *56*, 403.

(166) Chatani, Y.; Ueda, Y.; Tadokoro, H. Annual Meeting of Society of Polymer Science, Japan, Tokyo, 1977; p 1326.

(167) Bunn, C. W. *Trans. Faraday Soc.* **1939**, *35*, 482.

(168) Walter, E. R.; Reding, F. P. *J. Polym. Sci.* **1956**, *21*, 561.