Spectroscopically Determined Force Fields for Macromolecules. 1. *n*-Alkane Chains

K. Palmö, N. G. Mirkin, L.-O. Pietilä,† and S. Krimm*

Biophysics Research Division and Department of Physics, University of Michigan, Ann Arbor, Michigan 48109

Received June 2, 1993; Revised Manuscript Received September 20, 1993®

ABSTRACT: Detailed spectroscopic and dynamics studies of macromolecules require force fields that not only predict structures and thermodynamic quantities correctly but also reliably reproduce, i.e., to spectroscopic standards, vibrational frequencies. We have developed a methodology for transforming *ab initio* force fields, suitably scaled to experimentally assigned bands, into a molecular mechanics potential energy function that retains the original frequency as well as structure and energy predictions. With such inherent spectroscopic constraints, we obtain what we call a spectroscopically determined force field (SDFF). We have applied this approach to a set of 14 conformers of *n*-pentane and *n*-hexane, with the condition that the intrinsic geometry and force constant parameters be common to the set. We find that, with a systematically reduced parameter set, this SDFF gives very good agreement with experimental data. By virtue of the explicit dependence of its second derivatives on conformation, this potential should be transferable to conformers of other large molecules of this class and should provide more reliable molecular dynamics simulations than are obtained with spectroscopically less accurate potentials.

Introduction

Molecular mechanics (MM) force fields have been widely developed and used to study structural, energetic, and dynamic properties of molecules.[1-21] By representing the potential energy, V, of the molecule as a sum of intrinsic bonded terms (involving displacements of bonds, angles, and dihedral angles from their equilibrium values) and explicit nonbonded interactions (van der Waals, electrostatic, hydrogen-bonding) and parametrizing this function from experimental data or *ab initio* calculations, it has been possible to probe detailed aspects of molecular and macromolecular conformation and motion. Although some basic questions still exist about the assumed decomposition of the potential energy into the chosen physical terms, the efficacy of this approach seems to depend more on the completeness of the energy representation and the accuracy of the geometry, force constant, and nonbonded parameters selected in V.

When we examine the methods that have traditionally been used to parametrize the potential from studies of small molecules, we find that the main efforts were devoted to finding the values of the constants that would reproduce structures and thermodynamic properties, such as relative conformational energies. For such purposes, a limitation mainly to diagonal force constants for bonded terms in V was thought to be adequate. Vibrational frequencies, which are much more sensitive to the accuracy of the potential (at least at the minima), initially received relatively little attention in defining the potential. Yet infrared and Raman frequencies, which are strong determinants of V, are usually known to within, say, 2 cm^{-1} and therefore ~0.006 kcal/mol, whereas energies are often uncertain to tenths of a kilocalorie per mole. Although some efforts have been made to give more weight to frequencies in optimizing the parameters,[7,9,14,16,18,22,23] necessitating the more extensive incorporation of off-diagonal bonded terms in V, such data have not been incorporated systematically as the driving force in the refinement procedure.

We have felt that reliable reproduction of vibrational frequencies and eigenvectors should be a prime criterion for a satisfactory MM potential function, and we have therefore sought to incorporate such constraints into what we call a spectroscopically determined force field (SDFF). Such a force field not only would provide more accurate structural and molecular dynamics (MD) predictions but also would be a sound basis for obtaining conformation-dependent force constants for normal-mode calculations. This is an important requirement in the application of vibrational analysis to the study of macromolecular structure.[24]

Our approach to deriving a SDFF for a given class of molecules has the following general features. First, we start from a spectroscopic force field that provides an accurate description of the normal-mode frequencies and eigenvectors of a given structure. Since empirical force fields retain a bias in the selection of off-diagonal force constants, we choose to work with the *ab initio*-derived constants, which at least provide a complete description of the force field. Such force constants are obtained by detailed scaling of the full *ab initio* force field, based on spectroscopic criteria for assigning observed bands. Second, we analytically transform the spectroscopic force field into the MM format, a transformation that we have shown to be possible.[25] This is optimally done with a starting set of nonbonded parameters; these, however, can be chosen so as to give reasonable values of intrinsic force constants and geometries and so that the parameters are maximally consistent with other data such as crystal structures and energies. We repeat this procedure for all stationary state conformers of the molecules in the class being studied. For example, for the linear alkanes considered here, we obtained scaled ab initio force fields and the equivalent MM energy functions for the 4 stable conformers of *n*-pentane and the 10 such conformers of *n*-hexane. Third, in order to achieve maximum transferability of V between structures, we optimize the nonbonded parameters using the criterion that the intrinsic force constants and the intrinsic geometry parameters should be the same for all conformers[26] and that the crystal data should still be reproduced. Finally, we do a consistent force field (CFF) reoptimization of the valence parameters for overall best agreement with the *ab initio* structures, energies, and

† Present address: Neste Oy, Innopoli, P.O. Box 356, SF-02151 Espoo, Finland.

® Abstract published in *Advance ACS Abstracts*, November 1, 1993.

0024-9297/93/2226-6831$04.00/0

frequencies. In this latter process we reduce the number of force constants, retaining only those that significantly affect the frequencies or the structures and for which we can find a consistent conformational behavior.

There are a number of advantages to this approach. Most important, it starts out with a theoretically complete spectroscopic force field that is determined (via scaling) by experimentally assigned bands. As higher level *ab initio* calculations become feasible, this part of the procedure becomes increasingly more reliable. Also, by initially incorporating all cross terms in the quadratic part of V, we avoid the possibility of biasing the nonbonded parameters to compensate for an incomplete valence force field description. The transformation to a MM energy function, of course, inherently maintains frequency agreement, but we also find that the optimization process clearly reveals which molecular properties are sensitive to which aspects of V. For example, for the linear alkanes the relative energies of the conformers are determined primarily by the partial charges, which themselves have little effect on the frequencies, these in turn depending on the intrinsic force constants and van der Waals parameters. This "decoupling" effect is also helpful in revealing the kinds of terms that must be included in V in order to obtain reproduction of vibrational frequencies to spectroscopic standards. Finally, we note that in our MM function we have used (as in the *ab initio* force field) a nonredundant coordinate basis, thus being able to avoid the nonuniqueness inherent in the use of a redundant basis and to test the true transferability of the MM parameters. Subsequent transformation to a redundant basis of choice is always possible,[27] although performing MM and MD calculations in a nonredundant basis may be a preferable alternative.

In what follows, we describe briefly the nature of the *ab initio* calculations, discuss in some detail the kinds of terms in our potential and their dependence on conformation, describe the optimization procedure, give the final parameters of our force field, and show that its predictions of structures, energies, and vibrational frequencies are in very good agreement with observation. We conclude that the SDFF procedure provides a good systematic way of developing a reliable MM potential.

Calculations

***Ab Initio* Force Field.** At the heart of our procedure is the availability of spectroscopically scaled *ab initio* force fields for a set of conformers of a given class of molecules. In the present case, this set consists of the 4 stable conformers of *n*-pentane and the 10 stable conformers of *n*-hexane. The scale factors are obtained by refinement to the well-assigned normal-mode frequencies of *all-trans* (*T*)-pentane,[28,29] and these scale factors are then applied to the *ab initio* force constants of the other conformers. Details of the *ab initio* calculations are given elsewhere;[30] here we present the salient features of the results.

In order to select a reasonable basis set for the 14 structure and force field calculations, we carried out *ab initio* calculations on *TT*-pentane at the HF/6-31G, HF/6-31G*, HF/6-31G**, HF/6-31+G*, and MP2/6-31G* levels. The fully optimized geometries are very similar for all these basis sets, with MP2/6-31G* giving the shortest CC bond length (1.527 Å) and 6-31G the longest (1.531 Å), the opposite being true for the CH bond lengths (the differences generally being about 0.01 Å), bond angle differences being no more than 0.3°, and dihedral angle differences being on average 2° (the maximum for one being 5°). The relative energies of the conformers are, as expected, significantly different. Since the calculated difference between *TT* and *trans–gauche* (*TG*)-pentane with the MP2/6-31G* basis set (670 cal/mol) agrees well with the experimental result (600 ± 100 cal/mol[28]), compared to calculated values of 960 cal/mol with 6-31G and 1000 cal/mol with 6-31G*, we have used the MP2/6-31G* values as our reference energies.

The situation with respect to the frequencies is of a different nature. In brief, we find that only the 6-31G basis set permits satisfactory scaling to give acceptable frequencies and eigenvectors over the entire range.[30] We initially tried scaling all basis sets with either 9 diagonal internal coordinate scale factors, these 9 diagonal plus 4 off-diagonal scale factors, or 12 diagonal local symmetry coordinate scale factors. In all cases except for 6-31G and MP2/6-31G* we found that well-assigned CH_2 wag (at 1365 cm^{-1}) and CH_3 symmetric bend (at 1374 cm^{-1}) modes came out 20–35 cm^{-1} too high (and sometimes in reversed order, even compared to correctly ordered unscaled modes), and/or some CH_2 and CH_3 rock modes below 900 cm^{-1} were 15–20 cm^{-1} too high, and/or CCC deformation and skeletal torsion modes showed significant discrepancies. While the first two problems could be corrected with the MP2/6-31G* basis set and symmetry coordinate scale factors, the third could not (deformation modes were in error by ~30 cm^{-1}). With the 6-31G basis set and local symmetry coordinate scale factors, all of these problems were corrected and the final root-mean-square discrepancy with experimental bands below 1500 cm^{-1} was a respectable 4.9 cm^{-1}, with only one mode in this case having an error as large as 11 cm^{-1}. Since we are matching calculated harmonic frequencies with observed anharmonic modes, this level of agreement can be considered quite satisfactory, and we therefore used the 6-31G basis set with the scale factors of *TT*-pentane for the force fields of all conformers. When applied to *TTT*-hexane, for example, well-assigned bands below 1500 cm^{-1} were reproduced with a root-mean-square error of 4.7 cm^{-1}.

Spectroscopically Determined Force Field. The SDFF that we seek to derive for the linear alkanes is a MM potential that should be transferable between the 14 conformers of *n*-pentane and *n*-hexane. Its intrinsic geometry parameters and intrinsic force constants together with the nonbonded parameters are required to reproduce the *ab initio* geometries and energies as well as the vibrational frequencies of the scaled *ab initio* force fields. This result is achievable because it is possible, first, to transform analytically the *ab initio* geometry and force field for a given structure into a MM energy function[25] and, second, to optimize the nonbonded parameters under the condition of common intrinsic parameters.[26] To make such a function compatible with all conformers, however, we must be sure that all appropriate conformation-dependent terms are included in the potential. Our insistence on spectroscopic reproducibility has led to important insights into new terms that must be included in V, and we consider first their incorporation into the potential.

We model the potential energy by the function

$$V = \frac{1}{2}\sum_i F_{ii}(q_i - q_{i0})^2 + \sum_{i<j} F_{ij}(q_i - q_{i0})(q_j - q_{j0}) + \sum V_{\mathrm{tor}} + \sum V_{q,\mathrm{tor}} + \sum V_{\mathrm{tor,tor}} + \sum V_{\mathrm{nb}} \quad (1)$$

Here the q_i's are internal coordinates with reference values q_{i0}. In the SDFF method the force field is repeatedly transformed between Cartesian and internal coordinates, and it is therefore necessary to use a set of nonredundant valence coordinates. To make it easier to study in detail

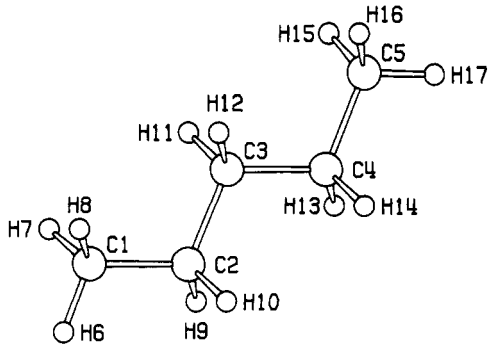

Figure 1. Structure of *TT*-pentane, with atom numbering to define coordinates of Table I.

the structure dependence of certain force constants, the internal coordinates were made nonredundant simply by defining only one torsion coordinate for each CC bond and by leaving out one valence angle at each C(sp^3) group. (This is an arbitrary choice, but the force field can always be transformed into any other nonredundant (or redundant) basis.[27]) Thus, for torsions not involving methyl groups, the CCCC dihedral angles were chosen as coordinates, and for each methyl torsion the CCCH dihedral angle closest to 180° was used. As for the angle coordinates, the methylene HCH angles were dropped, and in each methyl group the HCH angle sharing no CH bond with the selected methyl torsion coordinate was dropped.

F_{ii} and F_{ij} are intrinsic quadratic force constants, which, of course, are not equal to the scaled *ab initio* force constants but are obtained after transforming the latter set into the format of eq 1. It should be noted that the transformation does not depend on using only quadratic terms: it is entirely feasible, for example, to replace harmonic bond-stretching terms by Morse functions. The torsion potential, V_{tor}, is composed of ordinary cosine components

$$V_{tor} = \frac{1}{2}V_N(1 \pm \cos N\phi) \qquad (2)$$

where N is the periodicity and ϕ the torsion angle. The nonbonded potential, V_{nb}, contains Lennard-Jones functions that account for the van der Waals interactions and Coulomb terms that describe the electrostatic interactions. Test calculations were made with van der Waals 12–6 potentials, but the results proved to be better using the 9–6 form. Thus, for the nonbonded potential energy we have the following expression

$$\sum V_{nb} = \sum_{(ij)} \frac{A_iA_j}{r_{ij}^{9}} - \frac{B_iB_j}{r_{ij}^{6}} + \frac{Q_iQ_j}{4\pi\kappa\epsilon_0 r_{ij}} \qquad (3)$$

where r_{ij} is the distance between atoms i and j, A_i and B_i are respectively the repulsive and attractive van der Waals parameters of atom i, Q_i and Q_j are the partial charges on atoms i and j, ϵ_0 is the permittivity of free space, and κ is the dielectric constant (taken as 1). As usual, the summation runs over all atom pairs in positions 1,4 and higher.

The F_{ij} can be conformation dependent, and we discuss now the nature of such terms. As has long been known,[31] the force constant for interaction between two angle coordinates that form a dihedral angle is modulated by the cosine of the torsion angle. In a spectroscopic force field this effect is often blurred by the implicit nonbonded interactions, but it becomes explicit in MM force fields. In the present force field we have three such force constants: F(CCC–CCC#2), F(CCC–CCH#2), and F(CCH–CCH#2). (The coordinates are defined in Table I and Figure 1.) Thus, for example

$$F(\text{C1C2C3–C2C3C4}) = F^0(\text{C1C2C3–C2C3C4})\cos(\text{C1C2C3C4}) \qquad (4)$$

We also found a similar torsion cosine modulation for the interaction between an angle and a bond in a configuration where the bond shares one of the angle's end atoms but not the apex atom. Two force constants of this type were found to be significant, namely, F(CC–CCH#2) and F(CC–CCC#2). Thus, for example

$$F(\text{C1C2–C2C3C4}) = F^0(\text{C1C2–C2C3C4})\cos(\text{C1C2C3C4}) \qquad (5)$$

Although the cosine modulation does not completely explain the conformation dependence of these force

Table I. Internal Coordinates, Instrinsic Force Constants, Reference Geometry, and Nonbonded Parameters in SDFF for Linear Hydrocarbons

coordinate[a]	example[b]	force constant[c]	ref geometry[d]
CM	C2C1	599.4	1.526
MH	C1H7	687.9	1.083
CC	C2C3	580.2	1.528
CH	C2H9	653.7	1.084
CMH	C2C1H7	166.1	110.5
CMH2	C2C1H6	175.7	110.8
HMH	H6C1H7	162.1	108.3
CCC	C1C2C3	229.3	113.0
CCH	C1C2H9	186.6	109.1
CCt	H6C1C2C3	12.2	3-fold
	C1C2C3C4		
CC–CC	C1C2–C2C3	6.4	
MH–MH	C1H6–C1H8	5.8	
CH–CH	C2H9–C2H10	7.2	
CC–CCH	C1C2–C1C2H9	35.4	
	C1C2–C2C1H6(7,8)		
CC–CCH#2	C1C2–C2C3H11	−8.1[e]	
CC–CCC	C1C2–C1C2C3	28.7	
CC–CCC#2	C1C2–C2C3C4	−14.6[e]	
CC–CCt	C2C3–C1C2C3C4	−6.0[f]	
MH–XMH	C1H6–C2C1H6	20.7	
	C1H6–H7C1H6		
CH–CCX	C1H6–C2C1H7	−23.7	
	C2H9–C1C2C3		
CMH–CMHwC	C2C1H7–C2C1H8	81.7	
CMH–CMH2wC	C2C1H7–C2C1H6	86.5	
CMH–HMHwH	C2C1H7–H6C1H7	87.7	
CMH–HMHaC	C2C1H7–H6C1H8	80.4	
HMH–HMHwH	H6C1H7–H6C1H8	78.0	
CCC–CCHwC	C1C2C3–C1C2H9	103.8	
CCC–CCH#2	C1C2C3–C2C3H11	−18.1[g]	
CCC–CCC#2	C1C2C3–C2C3C4	−26.0[g]	
CCH–CCHwC	C1C2H9–C1C2H10	84.8	
CCH–CCHwH	C1C2H9–C3C2H9	101.4[h]	
CCH–CCHaC	C1C2H9–C3C2H10	86.4	
CCH–CCH#2	C3C2H9–C2C3H11	−14.0[g]	
CCH–CCtin[i]	C2C1H7–H6C1C2C3	3.6	
	C3C2H9–C1C2C3C4		
CCH–CCtout[i]	C1C2H9–C1C2C3C4	4.9	
HCH–CCtout[i]	H6C1H7–H6C1C2C3	4.9	
CCt–CCt	C1C2C3C4–C2C3C4C5	−1.5	

atom	A^j	B^j	Q^k
C	177.00	30.51	−0.2;[l] −0.3[m]
H	25.73	5.00	0.1

[a] M: methyl C. t: torsion. #2: one atom shared (not apex atom), cosine modulation. w: two atoms (including apex atom) shared. a: one atom (apex atom) shared. [b] See Figure 1. [c] Units: energy in kcal/mol, length in Å, angle in rad. [d] Bond lengths in Å; bond angles in deg. [e] F^0 in eq 5. [f] $F^0_{R\phi}$ in eq 8. [g] F^0 in eq 4. [h] F_0 in eq 6. [i] The force constant is taken to be positive if the following kind of triple product is greater than zero (cf. Figure 1): C3C2 × C3C4·C3H12. [j] Coefficients in eq 3. [k] In units of the electron charge. [l] For CH_2. [m] For CH_3.

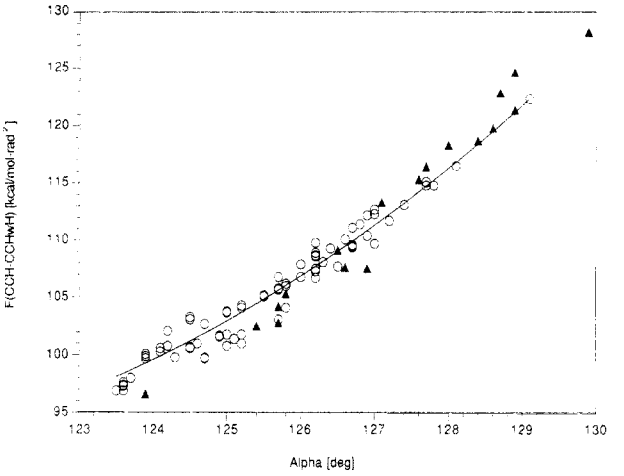

Figure 2. Dependence of F(CCH–CCHwH) on α, the angle between the CCH planes of the CC(H)C group: (○) equilibrium conformations; (▲) nonequilibrium conformations.

constants, it does so sufficiently well to warrant the presence of these force constants in the MM force field.

A more complicated form of structural dependence was found for the angle force constants in the methylene group. We noticed that in the sterically strained conformers some of the angle–angle interaction force constants were up to 20% larger than their average and that some of the diagonal angle force constants were up to 10% larger than the average. Neglecting to use such larger force constants caused frequency errors of up to 50 cm^{-1} in the 1450-cm^{-1} region. Attempts were made to attribute this increase to anharmonicity in the angle bending potential, but there was no clear trend relating the increased force constant values to the individual angle deformations. Since such large changes in the force constants should also be reflected in the geometry, a systematic study was undertaken to find out whether the behavior of the force constants could be related to changes in the angles between various planes in the methylene group. The study revealed that the force constant F(CCH–CCHwH) for interaction between two CCH angles sharing a CH bond has an almost linear dependence on the angle α between the CCH planes involved. This is shown in Figure 2. The data points represent all the stable conformations of pentane and hexane and three additional nonequilibrium conformations of pentane. The solid curve in the figure has the equation

$$F(\text{CCH–CCHwH}) = F(\alpha) = F_0 + C_1(\alpha - \alpha_0) + C_2(\alpha - \alpha_0)^2 \quad (6)$$

and was obtained in a least-squares fit to the data points from the stable conformations only. If the force constant is given in units of kcal/mol·rad^2 and the angle is given in deg, the constants in eq 6 have the values F_0 = 101.4 kcal/mol·rad^2, α_0 = 124.5°, C_1 = 3.375, and C_2 = 0.267. In the *all-trans* configuration $\alpha = \alpha_0$ and $F(\alpha)$ is then equal to F_0. It should be noted that the other angle–angle interaction force constants in the C–CH_2–C group did not exhibit such dependence on the angles between the respective planes. Nor did we find any other consistent pattern for the force constants. However, the values of the other force constants can be satisfactorily reproduced in every conformation by using coefficients derived from eq 6. We found in this connection that, whenever $F(\alpha)$ is large, both of the associated diagonals F(CCH) and F(CCC) are also large. F(CCH) can be obtained to a good approximation from its nonstrained value by multiplying by the coefficient

$$K_H(\alpha) = \left(\frac{F(\alpha)}{F_0}\right)^{1/2} \quad (7a)$$

In each C–CH_2–C configuration there are two α angles, so two coefficients, $K_H(\alpha_1)$ and $K_H(\alpha_2)$, are needed to compute the values of each pair of CCH force constants. The CCC angle force constant of the group is then obtained by multiplication of the nonstrained value by the coefficient

$$K_C(\alpha_1,\alpha_2) = [K_H(\alpha_1)\, K_H(\alpha_2)]^{1/2} \quad (7b)$$

If the coefficients K_H and K_C are thought of as correction factors associated with the angle coordinates and the correction factor for an interaction force constant is then the product of the coefficients for the two interacting coordinates, all the diagonal and interaction force constants come out close enough to their transformed *ab initio* values to give a good frequency agreement. Although the diagnoal angle force constants are strongly affected by the nonbonded interactions, the interaction force constant F(CCH–CCHwH) is not, so the correction factors are virtually independent of the nonbonded parameters.

Interaction between a torsion and its central bond represents still another case where the force constant depends on the torsion angle. We found this force constant to have the form

$$F_{R\phi} = F^0_{R\phi} \sin N(\phi - \phi_0) \quad (8)$$

where $\phi - \phi_0$ is the deformation of the torsion coordinate. This has a form different from that proposed by Allinger[9] and by Dinur and Hagler.[7] However, eq 8 is in both qualitative and quantitative agreement with the transformed *ab initio* force fields in which $F_{R\phi}$ is zero or very small whenever the torsion coordinate is at an intrinsic maximum (or minimum). According to the previously suggested forms,[7,9] the bond–torsion interaction potential energy is at its (negative) maximum at the intrinsic torsion maxima.

Cross terms involving torsion coordinates have to receive special treatment because a periodic torsion is not even approximately quadratic for large deformations from an intrinsic minimum. Even for moderate deformations, the curvature of a cosine torsion potential differs from that of a parabola, which should be taken into account in the cross terms of a realistic energy function. For very small deformations of a torsion coordinate, however, the torsion potential is almost quadratic and the cross terms involving it should then behave like ordinary quadratic cross terms. Thus, what is needed is a cross-term potential that is approximately quadratic for small deformations of a participating torsion but that for larger deformations correctly accounts for the torsion's nonquadratic curvature and for the periodicity.

Consider a cross term involving a quadratic coordinate q and a torsion coordinate ϕ of periodicity N. The criteria mentioned above are satisfied for deformations up to $\pm\pi/2N$ if the interaction potential energy is written as

$$V_{q\phi} = F_{q\phi}(q - q_0)g(\phi) \quad (9a)$$

where

$$g(\phi) = s\left\{\frac{2}{N^2}[1 - \cos N(\phi - \phi_0)]\right\}^{1/2} \quad (9b)$$

In eq 9b, ϕ_0 is the nearest intrinsic minimum of the torsion potential and s is the sign of $(\phi - \phi_0)$. The form of the function $g(\phi)$ is analogous to the case of a quadratic

coordinate. For a quadratic coordinate q the diagonal potential energy term is $V_{qq} = (1/2)F_q(q - q_0)^2$. For a cosine torsion potential the "diagonal" term can be written

$$V_{\phi\phi} = \frac{1}{2}V_N(1 \pm \cos N\phi) = \frac{F_\phi}{N^2}[1 - \cos N(\phi - \phi_0)] = \frac{1}{2}F_\phi g^2(\phi) \quad (10)$$

where V_N is the intrinsic barrier height, and F_ϕ is the torsion's intrinsic force constant at ϕ_0. It is also easy to see that, for small deformations, $g(\phi) \approx \phi - \phi_0$. This follows directly from eq 9b if $\cos N(\phi - \phi_0)$ is replaced by the Taylor expansion $1 - {}^1/_2N^2(\phi - \phi_0)^2$.

Thus, for torsion deformations smaller than or equal to $\pm\pi/2N$, eq 9a,b should give a reasonable description of the interaction potential energy. For still larger deformations, however, the potential energy becomes discontinuous unless certain modifications are made. This is seen as follows. Suppose that $q - q_0 > 0$. If ϕ is deformed in the positive direction to $\phi = \phi_0 + \pi/N$ (top of the barrier), the interaction potential energy according to eq 9a,b is $2F_{q\phi}(q - q_0)/N$. On the other hand, if the torsion coordinate starts out from the next intrinsic minimum, i.e., $\phi_0' = \phi_0 + 2\pi/N$, and is deformed in the negative direction to $\phi = \phi_0' - \pi/N$, the torsion coordinate arrives at exactly the same value as when it was deformed in the positive direction from ϕ_0, but now the interaction potential energy is $-2F_{q\phi}(q - q_0)/N$, making the potential energy discontinuous at $\phi = \phi_0 + \pi/N$.

In order to correct this, the following modifications must be made when the torsion deformation is larger than $\pm\pi/2N$. First, the sign of the interaction force constant must be changed, and, second, the deformation of the torsion coordinate must be measured from the nearest intrinsic maximum instead of the nearest intrinsic minimum; that is, ϕ_0 in eq 9b must be replaced by $\phi_{max} = \phi_0 + s\pi/N$. The sign factor in eq 9b must then also be replaced by $s' = -s$. If these substitutions are made, the interaction potential energy is zero at $\phi = \phi_0 \pm \pi/N$ and is everywhere continuous. Such behavior of the interaction energy is analogous to the case of two interacting quadratic coordinates, because near an intrinsic maximum the torsion potential resembles a parabola of negative curvature. This is also in agreement with the behavior of the force constants F(CCH–CCtout) and F(CCH–CCtin), which were indeed found to have changed their signs in the *cis* and 120° conformations of pentane.

The angle–torsion cross terms are also fairly important. Their effect on the frequencies is 15–30 cm^{-1} in the 1250- and 1460-cm^{-1} regions, and in the deformed hexane conformers they affect torsion angles by several degrees and carry a potential energy of up to 0.5 kcal/mol.

The $V_{tor,tor}$ terms in eq 1 behave like $V_{q,tor}$ with the exception that the special treatment described above must now be applied to both interacting coordinates.

Optimization of Spectroscopically Determined Force Field Parameters. The optimized nonbonded parameters in the SDFF were developed as follows. A significance test was first run on the force constant matrices in order to determine which elements are important to the vibrational frequencies. Geometrically equivalent parameters were then combined into groups, and the groups containing parameters significantly affected by the nonbonded interactions were included in the optimization. Altogether, 43 such groups were used (33 force constants and 10 reference geometry parameters). In the 4 conformers of pentane and 10 conformers of hexane this comprised a total of 3797 data points. The torsion force constants and intrinsic minima were also included, except for the strongly deformed torsion angles in the highest energy conformation of pentane and the three highest energy conformations of hexane. The intrinsic equilibrium angles of these torsions were left out of the fit because the SDFF transformation does not handle them correctly. However, the force constants were included as a separate group (CCtdef), since all the strongly deformed torsion coordinates in the studied molecules are in the range 94–96° (+ or −) and therefore should have about the same local force constant (second derivative). Although only part of the complete valence force field was explicitly used in the goodness-of-fit function, no force constants were dropped at this stage, and in the computation of the reference geometry parameters the full force constant matrices were used.

Early test calculations showed that the repulsive and attractive van der Waals parameters of the two different atoms defined (carbon and hydrogen) could not be optimized simultaneously, the reason being that the A_i's and B_i's are strongly correlated and that the valence parameters are much more sensitive to the A_i's than to the B_i's in the SDFF transformation. We therefore decided to optimize the repulsive parameters using the SDFF procedure and to optimize the attractive parameters to crystal structure and lattice energy data using the CFF method.[5,32,33] This was done in a cycle of alternating SDFF and CFF optimizations. As initial values for the parameters of the van der Waals interactions we took the Lennard-Jones 9–6 parameters given by Ermer and Lifson.[34] Atom-centered point charges were kept fixed at the values given in Discover (Version 2.7.0, Biosym Technologies Inc., March 1991), i.e., +0.1e for hydrogen and −0.3e and −0.2e for methyl and methylene carbons, respectively. From the initial Ermer–Lifson values, the repulsive parameters were first optimized using the SDFF procedure. The attractive parameters were then optimized to crystal data of ethane and pentane keeping the new repulsive parameters fixed. The optimized B_i's were thereafter fed back into the SDFF calculation, and the A_i's were reoptimized. This cycle was repeated until no significant changes resulted in the van der Waals parameters.

The optimization of the van der Waals parameters also yielded the pertinent values for the complete sets of valence force constants and reference geometry parameters that reproduce the *ab initio* Hessians and structures of each molecule. Although it is possible, at least in principle, to include the complete valence force field in the MM energy function, inclusion of very small cross terms is not required for most purposes. In practice, it is also difficult to find a consistent pattern for how small non-nearest-neighbor interaction force constants change with conformation. Primarily for the latter reason, the number of interaction force constants was reduced to include only those that clearly are important to the vibrational frequencies or the structures. Using the average values of the selected valence parameters, the *ab initio* structures, frequencies, and relative energies and barriers were all reasonably well reproduced. No CFF optimization to structures or energies was therefore made, with the exception that the transformed value, −12.1 kcal/mol·rad·Å, of the bond–torsion force constant F(CC–CCt) was replaced by the value −6.0 kcal/mol·rad·Å, since otherwise it turned out to increase the affected CC bond lengths too much. This force constant has no effect on the vibrational frequencies, which were calculated with a root-mean-square deviation of 8.9 cm^{-1}. However, a subsequent CFF optimization to the

Table II. *Ab Initio* Geometries and SDFF Reference Geometry Parameters and Their Ranges for Conformers of *n*-Pentane and *n*-Hexane

para-meter[a]	ab inito mean	min	max	st dev[b]	SDFF mean	min	max	st dev[b]
CM	1.532	1.531	1.534	0.001	1.526	1.526	1.528	0.001
MH	1.085	1.082	1.086	0.001	1.083	1.079	1.085	0.001
CC	1.537	1.533	1.545	0.003	1.528	1.526	1.534	0.002
CH	1.087	1.085	1.089	0.001	1.084	1.081	1.086	0.001
CMH	111.3	110.5	112.3	0.5	110.5	110.0	111.0	0.2
CMH2	111.0	110.6	111.3	0.2	110.8	110.5	111.0	0.1
HMH	107.7	107.4	107.8	0.1	108.3	108.2	108.6	0.1
CCC	114.4	112.4	117.1	1.2	113.0	111.9	113.9	0.5
CCH	109.0	106.8	110.2	0.6	109.1	108.3	110.1	0.3
CCt	−0.4	−7.7	7.3	3.8	−0.5	−4.8	6.2	1.7

[a] M: methyl C. Bond lengths in Å; bond angles in deg. [b] Standard deviation.

frequencies was done for some of the force constants in order to compensate for the many omitted cross terms. The force constants then changed very little, but the root-mean square deviation dropped to 6.9 cm^{-1}.

Results

The optimization procedure described above led to a reasonable reduction in the *ab initio* force constant set (1035 for *n*-pentane conformers and 1485 for *n*-hexane conformers) to a common set of 35 diagonal and off-diagonal intrinsic quadratic force constants. These plus the 14 geometry, nonbonded, and partial charge parameters define our SDFF. The definitions of the nonredundant coordinates are given in Table I, together with the final intrinsic geometry parameters, force constants, and nonbonded parameters.

The values and ranges of the *ab initio* and SDFF reference geometry parameters for all the conformers of *n*-pentane and *n*-hexane are given in Table II. The computed reference geometry parameters are consistent with the MM model. The non-methyl CC reference bond lengths have a standard deviation of 0.002 Å from the mean value of 1.528 Å, the largest deviation being 0.006 Å. Similarly, the computed reference values for the CCC angles have a standard deviation of 0.5° from the mean value of 113.0°, and there are no deviations larger than 1.1° while in the *ab initio* structures the CCC angles cover a range of 4.7°. The computed torsion reference angles are more scattered and have a standard deviation of 1.7° from the fixed values given by the periodicity (±60° or 180°), and the deviations range from −4.8° to +6.2°. In comparison, the *ab initio* torsion angles have a standard deviation of 3.8° from the intrinsic minima, the differences ranging from −7.7° to +7.3°.

Some of the SDFF energy-minimized structures are compared with the *ab initio* structures in Table III. The agreement is good, with the SDFF CC bond lengths within 0.002 Å, the CCC bond angles within 1°, and the CC torsion angles within 6° of the corresponding HF/6-31G values. The MM bond lengths and bond angles depend largely on the reference geometry parameters, which were obtained in the SDFF transformation from the HF/6-31G structures and force fields. Even if a discrepancy were present in the nonbonded parameters, the transformation would try to change the reference geometry so as to make the complete energy function still yield the *ab initio* geometry. It is therefore not surprising that the bonds and angles are close to the HF/6-31G values. The torsions, on the other hand, have a fixed periodicity and thus always have the same "reference angles". The SDFF transformation actually produces reference torsion angles that differ somewhat from the obvious ones (see Table II), but in the final MM force field the true periodicity is, of course, used. The computed torsion angles are therefore less directly dependent on the SDFF reference geometry but are to a very large extent determined by the nonbonded interactions. This is especially true for the strongly deformed torsion coordinates (the ones near 90°) since they were not included in the fit of the nonbonds. The good agreement between the SDFF and *ab initio* torsion angles is therefore an indication, together with the reproduced crystal data, that the Lennard-Jones 9–6 potential, and the parameters optimized in this work, give a good description of the true van der Waals interactions.

The relative energies of the stable conformers of pentane and hexane are compared in Table IV, where we give the component contributions and relative energies at the

Table III. Comparison of Geometries for Some Conformers of *n*-Pentane and *n*-Hexane from SDFF and *Ab Initio* Calculations

Pentane

	TT			*GG'*		
parameter[a]	SDFF	HF/6-31G	MP2/6-31G*	SDFF	HF/6-31G	MP2/6-31G*
C1C2	1.532	1.531	1.527	1.531	1.532	1.527
C2C3	1.534	1.533	1.527	1.538	1.538	1.531
C3C4	1.534	1.533	1.527	1.543	1.544	1.538
C4C5	1.532	1.531	1.527	1.533	1.534	1.529
C1C2C3	112.8	113.0	112.8	115.9	115.6	115.0
C2C3C4	113.4	113.4	113.4	116.3	115.8	115.0
C3C4C5	112.8	113.0	112.8	114.5	114.8	114.3
HC1C2C3	180.0	180.0	180.0	174.1	174.4	174.8
C1C2C3C4	180.0	180.0	180.0	61.4	62.8	60.7
C2C3C4C5	180.0	180.0	180.0	−90.4	−93.5	−94.2
C3C4C5H	180.0	180.0	180.0	178.6	−179.4	180.0

Hexane

	GGG			*TGG'*		
parameter[a]	SDFF	HF/6-31G	MP2/6-31G*	SDFF	HF/6-31G	MP2/6-31G*
C1C2	1.532	1.532	1.527	1.531	1.532	1.527
C2C3	1.537	1.537	1.531	1.534	1.534	1.527
C3C4	1.539	1.538	1.532	1.539	1.538	1.532
C4C5	1.537	1.537	1.531	1.543	1.545	1.538
C5C6	1.532	1.532	1.527	1.533	1.534	1.529
C1C2C3	114.5	114.2	113.7	112.2	112.4	112.2
C2C3C4	116.2	115.8	114.9	116.2	115.7	115.6
C3C4C5	116.2	115.8	114.9	116.3	115.5	115.1
C4C5C6	114.5	114.2	113.7	114.5	114.5	114.3
HC1C2C3	177.1	175.1	175.1	178.9	179.6	179.4
C1C2C3C4	62.3	61.7	59.4	174.3	175.2	174.6
C2C3C4C5	60.2	59.6	56.7	60.6	63.9	61.0
C3C4C5C6	62.3	61.7	59.4	−90.0	−96.4	−93.6
C4C5C6H	177.1	175.1	175.1	178.6	−179.9	180.0

	GGG'			*GG'G*		
parameter[a]	SDFF	HF/6-31G	MP2/6-31G*	SDFF	HF/6-31G	MP2/6-31G*
C1C2	1.532	1.532	1.527	1.532	1.534	1.529
C2C3	1.536	1.537	1.530	1.542	1.544	1.537
C3C4	1.540	1.540	1.533	1.541	1.540	1.534
C4C5	1.543	1.544	1.537	1.542	1.544	1.537
C5C6	1.532	1.534	1.529	1.532	1.534	1.529
C1C2C3	114.5	114.1	113.7	114.4	114.6	114.3
C2C3C4	117.6	116.9	116.7	117.7	117.1	117.0
C3C4C5	116.5	115.5	115.3	117.7	117.1	117.0
C4C5C6	114.9	114.9	114.9	114.4	114.6	114.3
HC1C2C3	178.0	176.3	176.3	−177.9	179.9	180.0
C1C2C3C4	62.4	64.9	62.2	91.0	93.9	91.3
C2C3C4C5	63.9	65.1	63.1	−60.5	−65.8	−62.0
C3C4C5C6	−89.7	−95.6	−91.6	91.0	93.9	91.2
C4C5C6H	179.9	−178.4	−177.6	−177.9	179.8	179.9

[a] Bond lengths in Å; bond angles in deg.

Table IV. Relative Energies of Conformers of *n*-Pentane and *n*-Hexane

conformer[a]	SDFF (Hartree–Fock)[b] V_q[c]	V_{tor}[d]	V_{LJ}[e]	V_Q[f]	rel	SDFF (CFF)[g]	MP2[h]
			Pentane				
TT	0.16	0.00	1.98	5.31	0.00	0.00	0.00
TG	0.21	0.15	2.00	5.79	0.70	0.69	0.67
GG	0.36	0.11	1.91	6.44	1.37	1.34	1.09
GG'	0.39	1.72	1.93	6.50	3.09	2.89	3.19
			Hexane				
TTT	0.20	0.00	2.44	5.36	0.00	0.00	0.00
TTG	0.25	0.14	2.45	5.85	0.69	0.68	0.63
TGT	0.26	0.16	2.47	5.72	0.60	0.59	0.64
TGG	0.41	0.11	2.34	6.34	1.21	1.19	1.03
GTG	0.28	0.37	2.43	6.34	1.41	1.37	1.23
GTG'	0.33	0.29	2.49	6.32	1.43	1.40	1.41
GGG	0.54	0.11	2.24	6.95	1.84	1.80	1.42
G'GT	0.41	1.92	2.42	6.26	3.00	2.69	3.06
GGG'	0.57	1.93	2.18	6.96	3.63	3.34	3.57
GG'G	0.66	3.45	2.22	6.99	5.33	4.85	5.70

[a] *T*: *trans* ($\phi \cong 180°$). *G*: *gauche* ($\phi \cong 60°$). *G'*: *gauche'* ($\phi \cong -60°$). [b] SDFF energy (in kcal/mol) calculated at Hartree–Fock minimum structure (with CFF-optimized parameters). [c] Valence energy (see eq 1). [d] Torsion energy (see eq 2). [e] van der Waals energy (see eq 3). [f] Coulomb energy (see eq 3). [g] SDFF energy (in kcal/mol) calculated at CFF-minimized geometries. [h] *Ab initio* energy (in kcal/mol) calculated with the MP2/6-31G* basis set.

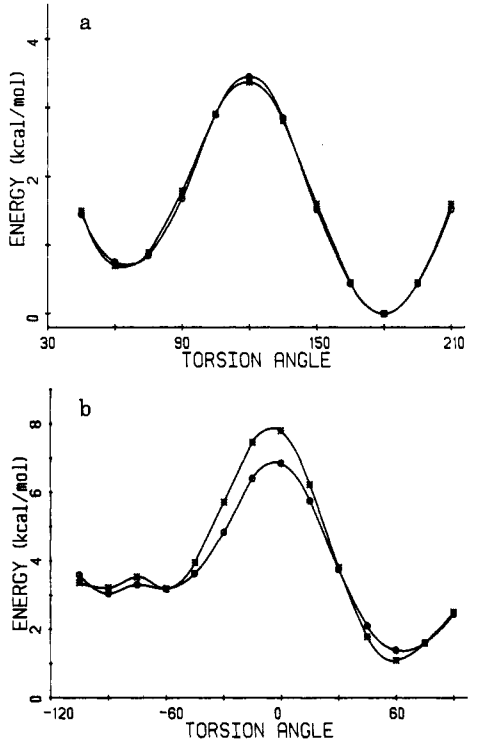

Figure 3. Comparison of *ab initio* (*) and SDFF (O) barriers of *n*-pentane. (a) *TT* (180°, 180°) to *TG* (180°, ~60°). (b) *GG* (~60°, ~60°) to *GG'* (~60°, ~ -90°; or ~90°, ~ -60°).

Hartree–Fock, the CFF-minimized, and the MP2-minimized structures. The *TT* → *TG* and *GG* → *GG'* barriers of pentane are graphically compared in parts a and b of Figure 3, respectively. (Note that the two minima in Figure 3b correspond to the equivalent *GG'* and *G'G* structures.) There is a fair agreement between the SDFF and MP2/6-31G* energies. There are a few large differences, such as the pentane barrier height in the *cis* conformation which is 12% smaller than the MP2 barrier, and the relative energy of the *GGG* conformer of hexane which is calculated 27% higher than the MP2 energy. These are significant differences that cannot be ignored. We believe, however, that they are mainly due to the charges. The charges strongly affect the relative energies and are in some cases probably in error. In *GG*-pentane, for example, the MM relative energy of 1.34 kcal/mol is 23% off the MP2/6-31G* value of 1.09 kcal/mol, but the contribution from the charges is as large as 1.13 kcal/mol. Similarly, in *GGG*-hexane, which has a MM relative energy of 1.80 kcal/mol, as much as 1.60 kcal/mol comes from the charges, the MP2/6-31G* energy being only 1.42 kcal/mol. On the other hand, in cases where the charge interactions are less significant the relative energies are in excellent agreement with *ab initio*. For instance, in the 120° (top of the barrier) conformation of pentane where the charges contribute only 0.14 kcal/mol, the MM barrier height is 3.45 kcal/mol and the MP2/6-31G* barrier is 3.37 kcal/mol. In fact, the entire *TT* → *TG* barrier is not much affected by the charges and is very well reproduced in the MM calculation. A few test calculations were made in which different charges were tried out. We then found that, if the charges on all the hydrogen atoms are increased to +0.115e, and the carbon charges are correspondingly changed to −0.230e and −0.345e for methylene and methyl carbons, respectively, the *cis* barrier that is now too low improves by ~0.5 kcal/mol but the relative energies that are already too high become even larger. Thus, no clear net improvement results from increasing (or decreasing) the point charges across the board. This may suggest that the charges depend on conformation or that atomic point charges are not sufficient to describe the electrostatic interactions. In the present study, we did not set out to find optimal electrostatic interactions, since they only seem to have a small (indirect) effect on the SDFF quantities that are currently our main interest, namely, the force constants, the intrinsic geometry, and the van der Waals parameters.

Table V. SDFF Structures[a] and Lattice Energies[b] of Crystal Structures of Ethane, *n*-Pentane, and *n*-Hexane

molecule		*a*	*b*	*c*	*α*	*β*	*γ*	*E*
ethane	calcd	4.306	5.534	5.588	90.0	89.9	90.0	−4.42
	exp[c]	4.226	5.623	5.845	90.0	90.4	90.0	−5.26[d]
n-pentane	calcd	4.13	8.690	14.852	90.0	90.0	90.0	−10.4
	exp[e]	4.10	9.076	14.859	90.0	90.0	90.0	−11.1[f]
n-hexane	calcd	4.14	4.45	8.61	96.9	88.4	102.6	−12.4
	exp[g]	4.17	4.70	8.57	96.6	87.2	105.0	−13.2[h]

[a] Unit cell lengths in Å; unit cell angles in deg. [b] In kcal/mol. [c] Reference 35. [d] Reference 32. [e] Reference 36. [f] Reference 38. [g] Reference 37. [h] Reference 39.

The computed and experimental crystal structures of ethane,[35] pentane,[36] and hexane,[37] and their lattice energies,[32,38,39] are given in Table V. The hexane crystal data were not used in the optimization but were computed as a check.

The ranges of some of the MM force constants that resulted from the SDFF transformations from *ab initio* for all the conformers of pentane and hexane are presented in Table VI. As is seen from this table, the SDFF force constants do not show large deviations among the studied conformers. It is interesting to compare these force constants with the scaled *ab initio* values. The ranges of the corresponding *ab initio* force constants are also shown in Table VI. The transformed force constants are generally smaller than the *ab initio* values. As expected, the angle bending force constants are those that changed the most. One of the biggest differences occurs for the CCC angle force constant. In the scaled *ab initio* force field it ranges from 252.8 to 311.6 and has a standard deviation of 16.3 from the mean value of 273.9 kcal/mol·rad^2. In the SDFF, on the other hand, it is between 220.3 and 233.7, with a standard deviation of only 3.5 from the mean value of 229.3 kcal/mol·rad^2. Another force constant that changed

Table VI. Variations of Some *Ab Initio* and SDFF Force Constants

force constant[a]	ab initio mean[c]	min	max	st dev[d]	SDFF[b] mean[c]	min	max	st dev[d]
CCC	273.9	252.8	311.6	16.3	229.3	220.3	233.7	3.5
CCH	196.1	186.7	218.9	6.0	185.3	180.9	189.2	1.6
CCt	16.5	12.5	37.4	4.4	12.2	9.5	18.2	1.5
CCtdef	29.8	22.0	38.9	8.5	2.9	0.3	4.1	1.6
CC–CC	17.9	14.5	22.7	2.5	6.4	4.5	7.9	0.8
CC–CCH	40.4	36.4	48.4	2.3	35.4	30.0	42.0	1.9
CC–CCH#2	0.3	−4.8	6.1	3.6	−8.1	−16.3	−2.7	2.9
CC–CCC	48.7	37.5	58.9	5.5	28.7	20.9	33.7	3.0
CC–CCC#2	18.3	8.5	44.3	6.8	−14.6	−21.7	−8.3	4.5
CC–CCt	−0.5	−9.2	5.7	7.9	−12.1	−13.2	−10.9	1.1
CCC–CCHwC	108.8	99.2	124.7	5.4	103.8	98.9	107.8	1.4
CCC–CCH#2	1.8	−7.8	18.7	9.1	−18.1	−26.4	−13.2	2.1
CCC–CCC#2	18.3	8.5	44.3	6.8	−26.0	−36.0	−22.8	2.7
CCH–CCHwC	89.6	81.9	100.1	3.9	86.3	84.3	88.9	0.8
CCH–CCHwH	106.0	96.8	126.9	7.1	101.1	98.7	103.2	1.1
CCH–CCHaC	90.8	83.7	101.1	3.8	87.5	85.7	89.6	0.7
CCH–CCH#2	2.4	−6.3	16.8	8.8	−12.9	−17.3	−7.4	1.9
CCt–CCt	−2.8	−5.4	9.8	2.7	−1.1	−1.9	4.6	1.2

[a] See Table I and Figure 1 for definitions of coordinates. CCtdef: force constant in conformers with highly deformed torsion angles. [b] Before CFF optimization. [c] Units: energy in kcal/mol, length in Å, and angle in rad. [d] Standard deviation.

radically in the SDFF transformation was CCtdef, which pertains to the strongly deformed torsion coordinates. CCtdef is not a real force constant or MM parameter but describes the local curvature of the intrinsic torsion potential at the actual geometry. In the transformation, CCtdef changed from 29.8 to 2.9 kcal/mol·rad^2 (average values). The transformed values are very small, as they should be, since the second derivatives of the cosine torsion potential are also small at the torsion angles in question.

The vibrational frequencies are in excellent agreement with the scaled HF/6-31G values. The root-mean-square frequency deviation for all 720 frequencies is 6.9 cm^{-1}, with 462 frequencies (64%) being within ±5 cm^{-1} of the *ab initio* values and 635 frequencies (88%) being within ±10 cm^{-1}. There are 9 frequencies that are off by 20–25 cm^{-1}. These are partly due to the fact that we have omitted all angle–angle interactions between coordinates that do not share a bond or the apex atom. Although some of these longer range interaction force constants are large enough to affect the frequencies to some degree, they have not been included in the MM force field because they depend on the structure in a nonobvious way. However, in any simplified valence force field, where hundreds of interaction force constants have been left out and average values are used for the remaining ones, a few larger frequency deviations are only to be expected, since some of the omitted or modified force constants are bound to affect some frequencies in the same direction, thus causing an accumulation of error. For *TT*-pentane, the root-mean-square error of the SDFF with respect to the observed frequencies below 1500 cm^{-1} is 10.3 cm^{-1} (the standard deviation is 7.1 cm^{-1}), with two frequencies being in the 25–30-cm^{-1} range. The comparable root-mean-square error for *TTT*-hexane is 10.7 cm^{-1} (with a standard deviation of 7.2 cm^{-1}). This is to be compared, for example, to root-mean-square errors of 30.6 cm^{-1} for an early MM function[39] (although this might be improved by subsequent inclusion of vibrational frequencies in the optimization[22]) and at least 90.1 cm^{-1} for the AMBER, 82.6 cm^{-1} for the CHARMM, and 30.0 cm^{-1} for the Discover potential functions (based on a sequential matching of assigned bands,[29] since symmetry species were not available). The MM3 potential gives a root-mean-square error of ~50 cm^{-1} for molecules in this class.[40] In Table VII we compare some of the SDFF frequencies of three conformers of *n*-hexane with the *ab initio* values and with observed bands. The conformational dependence is well characterized, and thus, in spite of the omission of so many interaction force constants, our SDFF gives a very reasonable reproduction of vibrational frequencies.

Table VII. Observed, *Ab Initio*, and SDFF Frequencies of Some *n*-Hexane Conformers

TTT			*GTG*			*GG'G*	
ν_0[a]	ν_c(AI)[a]	ν_c(SDFF)	ν_0[a]	ν_c(AI)[a]	ν_c(SDFF)	ν_c(AI)[a]	ν_c(SDFF)
1383	1378	1376	1378	1377	1380	1380	1379
1369	1376	1374	1378	1377	1378	1377	1374
1370	1357	1354	1359	1362	1360	1356	1361
1353	1349	1344	1343	1343	1349	1343	1351
1307	1311	1324	1343	1341	1339	1342	1334
1307	1307	1321	1314	1318	1337	1338	1326
1301	1300	1311	1293	1293	1291	1288	1276
1280	1281	1274	1287	1287	1273	1274	1273
1242	1252	1272	1252	1250	1245	1262	1240
1225	1228	1230	1224	1223	1222	1220	1226
1179	1184	1209	1167	1164	1158	1157	1148
1143	1151	1132	1136	1142	1148	1127	1140
1067	1069	1061	1088	1094	1091	1119	1118
1064	1068	1057	1056	1059	1062	1056	1053
1041	1043	1038	1042	1038	1051	1040	1049
1008	1007	1008	1004	1009	1019	1018	1009
996	997	1002	951	950	944	958	955
898	894	902	903	908	927	920	904
898	893	896	869	863	856	874	870
886	877	874	869	861	854	827	832
798	794	797	772	775	761	787	790
735	734	729	758	766	756	754	757
721	723	705	725	724	733	713	728
467	467	473	488	492	500	449	466
370	365	368		387	389	385	375
303	298	292		304	302	336	325
256	257	260		302	295	288	287
241	245	249		244	253	230	238
	154	151		201	201	225	233
	133	125		139	136	140	138
	101	106		113	113	123	126
	77	77		50	52	82	91

[a] See ref 30.

The vibrational spectrum can also provide a sensitive test of the nonbonded part of the force field through an examination of the low-frequency crystal lattice vibrations. These may be an even better indicator than the experimental lattice energies, which are obtained after theoretical corrections to the heats of evaporation and melting.[39] In Table VIII we present the low-frequency results of a crystal normal-mode calculations on triclinic *n*-hexane[37] using our SDFF and the 0 K minimized structure (Table V). These are compared with Raman (20 K)[41] and inelastic neutron scattering (9 K)[42] data and with calculations by others.[9,41,43,44] Our predictions of the three lowest-lying librational modes are better than the others except for Brunel and Dows,[41] in whose calculation the influence of the three lowest internal modes was accounted for by mixing through their frequency parameters. The SDFF predictions of the higher crystal frequencies are also quite reasonable, in view of the difference of the SDFF with respect to the *ab initio* isolated chain frequencies, the (unknown) difference between the 0 and 9 K crystal structures, the possible effects of anharmonicity,[44] and the density-of-states representation of the inelastic neutron scattering spectrum. Another, more demanding, test is provided by crystalline *n*-pentane, which has four molecules per unit cell[36] and therefore has translational as well as librational lattice modes. In Figure 4 we compare the mass-weighted squared hydrogen atom eigenvectors from a crystal normal-mode calculation (0 K structure)

Table VIII. Low-Frequency Modes of 0 K Crystalline n-Hexane

ν(obs)		isolated chain		crystal				
INS[a]	R[b]	ν(AI)[c]	ν(SDFF)[d]	ν(SDFF)[d]	ν(WL)[e]	ν(W)[f]	ν(BD)[b]	ν(TASD)[g]
55	53			64	49	55	59	66
71	74			68	57	33	71	80
93	87			90	74	76	89	100
105		77	77	100				114
124		101	106	135				150
142		133	125	148				161
182	174	154	151	171				189
241	252	245	249	251				270
256		257	260	277				276
303	312	298	292	322				
370	374	365	368	378				
467		467	473	481				

[a] Reference 41. [b] Reference 40. [c] Reference 30. [d] This work. [e] Reference 39. [f] Reference 43. [g] Reference 42.

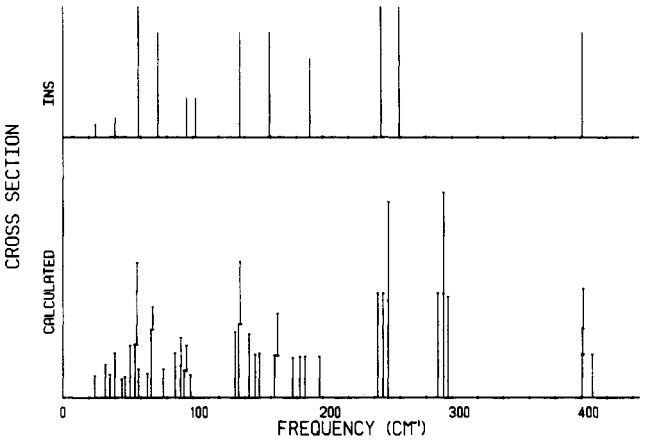

Figure 4. Comparison of inelastic neutron scattering peaks of 9 K *n*-pentane[42] with mass-weighted squared hydrogen atom eigenvectors from SDFF crystal normal-mode calculation of optimized (0 K) structure.

with the neutron scattering data.[42] The agreement is generally quite good, except for the discrepancy between the calculated ~294-cm^{-1} modes and observed 259-cm^{-1} peak (which may be due to the particular dispersion curve for this mode). These results thus indicate that we have refined a very satisfactory (if not yet perfect) nonbonded potential.

Discussion

The SDFF method constitutes an alternative to other force field developing schemes[7,45] that rely on molecular data derived from *ab initio* calculations. An advantage of the SDFF procedure is that it offers a convenient self-consistent way to obtain both van der Waals parameters and valence parameters that are refined for spectroscopic accuracy. Since the method does not employ any least-squares fitting to *ab initio* potential energy surfaces or derivatives but makes a direct transformation of these into MM quantities, correlations between the valence parameters or between the van der Waals and the valence parameters are not a problem, and the computed valence parameters are not "contaminated" by the nonbonded interactions. In fact, the implicit relations between these two types of parameters are what is utilized in the optimization of the nonbonds. Fortunately, the charges do not significantly affect the valence parameters and do not have to be optimized simultaneously with the van der Waals parameters. Further, only energy-minimized structures are needed in the calculations, although nonequilibrium conformations can be used for checking purposes. This increases the reliability of the SDFF energy function because the *ab initio* second derivatives used in the transformation can be properly scaled.

The results produced by the application of the SDFF method derived in this work are in good, although not perfect, agreement with *ab initio* and (where available) experimental data. Some of the larger discrepancies concern the relative energies. For the stable conformers that have too low SDFF relative energies, we found that significantly better agreement with the MP2/6-31G* values was obtained when the energies were computed (without minimizing) at the HF/6-31G geometries. For example, the three highest conformations of hexane then have SDFF energies of 3.00, 3.63, and 5.33 kcal/mol, respectively. This shows that the too low energies are to a large extent caused by the (small) differences between the SDFF and *ab initio* minimized structures. The too low barrier at the *cis* configuration of pentane, however, is due to other reasons. Deficiencies in the charges were previously mentioned as a possible explanation. A higher barrier could also be obtained by introducing additional torsion terms. At least 1-fold and 2-fold torsion terms would be theoretically justified.[8,9] However, the barrier between the *TT* and *TG* conformers of pentane is well reproduced with the 3-fold terms only, which indicates that the mentioned extra terms would not suffice and that the final torsion potential would become fairly complicated. Therefore, and because of the uncertainty regarding the charges, we have used only the 3-fold torsion terms. Similar conclusions were also drawn by Smith and Karplus[11] in a recent computational study of some organic molecules.

Another way of improving the energies, as well as other results, could be to alter the form of the van der Waals potential. The parameters that we have derived here are optimal for the Lennard-Jones 9–6 form (with regard to the *ab initio* HF/6-31G force fields and structures), and like Dinur and Hagler[7] we found that the 9–6 function is better than the 12–6 form. However, test calculations showed that a Lennard-Jones 8–6 potential might be still better for the purpose of reproducing the *ab initio* relative energies and structures. Buckingham potentials were not tested but would probably yield comparable results. Thus, the question of the ultimate form of the van der Waals potential remains open.

Despite the above problems, our SDFF gives better predictions than other MM potentials. In Table IX we compare the torsion angles and energies of pentane conformers predicted from *ab initio*, SDFF, and two current MM potentials—those of MM3[9] and Discover. We see that the latter, while predicting fairly well the undeformed torsion angles, does poorly, in distinction to the SDFF, in predicting the highly deformed torsion. The energies of all the conformers are predicted high by the MM functions (and would be significantly higher for *GG'* if calculated at the correct geometry), whereas the SDFF

Table IX. Comparisons of Torsion Angles and Energies of Pentane Conformers from *Ab Initio*, SDFF, MM3, and Biosym Potentials

conformer		HF/6-31G	MP2/6-31G*	SDFF	MM3[a]	Discover[b]
			Torsions			
TG	$\tau 1$	176.6	176.0	177.6		178
	$\tau 2$	66.5	63.8	65.6		65
GG	$\tau 1$	62.0	59.3	63.1		62
	$\tau 2$	61.7	59.3	63.1		62
GG'	$\tau 1$	62.8	60.7	61.4	77	66
	$\tau 2$	−93.5	−94.2	−90.4	−78	−82
			Energies			
TG		0.96	0.67	0.67	0.86	0.89
GG		1.80	1.09	1.34	1.62	1.69
GG'		3.81	3.19	2.82	3.51	3.42

[a] Reference 9. [b] CFF 91.

energies are closer to the MP2 values. Of course, the SDFF provides a much better frequency agreement than do the other MM potentials.

In conclusion, we find that the SDFF procedure is not only realizable but that it provides a MM force field that is superior in predicting structures, energies, and vibrational frequencies than previous potentials. It should, therefore, also be more reliable for molecular dynamics calculations of long alkane chains. Such applications to polymer chains will be dealt with in forthcoming publications.

Acknowledgment. This research was supported by NSF Grants DMR-9110353 and MCB-9115906 and, for one of us (K.P.), by the Academy of Finland. We thank Gull-Maj Roberts for calculating the AMBER and CHARMM *n*-hexane frequencies.

References and Notes

(1) Jacob, E. J.; Thompson, H. B.; Bartell, L. S. *J. Chem. Phys.* **1967**, *47*, 3736.
(2) Bartell, L. S.; Fitzwater, S.; Hehre, W. J. *J. Chem. Phys.* **1975**, *63*, 4750.
(3) Scheraga, H. A. *Adv. Phys. Org. Chem.* **1968**, *6*, 103.
(4) Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361.
(5) Lifson, S.; Warshell, A. *J. Chem. Phys.* **1968**, *49*, 5116.
(6) Lifson, S.; Hagler, A. T.; Dauber, P. *J. Am. Chem. Soc.* **1979**, *101*, 5111.
(7) Dinur, U.; Hagler, A. T. *Rev. Comput. Chem.* **1991**, *2*, 99.
(8) Burkert, U.; Allinger, N. L. *Molecular Mechanics*; American Chemical Society: Washington, DC, 1982.
(9) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551.
(10) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(11) Smith, J. C.; Karplus, M. *J. Am. Chem. Soc.* **1992**, *114*, 801.
(12) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.
(13) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.
(14) Sorensen, R. A.; Liau, W. B.; Boyd, R. H. *Macromolecules* **1988**, *21*, 194.
(15) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
(16) Dasgupta, S.; Goddard, W. A., III *J. Chem. Phys.* **1989**, *90*, 7207.
(17) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A., III; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024.
(18) Miwa, Y.; Machida, K. *J. Am. Chem. Soc.* **1988**, *110*, 5183.
(19) Dillen, J. L. M. *J. Comput. Chem.* **1990**, *11*, 1125.
(20) Halgren, T. A. *J. Am. Chem. Soc.* **1992**, *114*, 7827.
(21) Amodeo, P.; Barone, V. *J. Am. Chem. Soc.* **1992**, *114*, 9085.
(22) Lifson, S.; Stern, P. S. *J. Chem. Phys.* **1982**, *77*, 4542.
(23) Derreumaux, P.; Vergoten, G.; Lagant, P. *J. Comput. Chem.* **1990**, *11*, 560.
(24) Krimm, S.; Bandekar, J. *Adv. Protein Chem.* **1986**, *38*, 181.
(25) Palmö, K.; Pietilä, L.-O.; Krimm, S. *J. Comput. Chem.* **1991**, *12*, 385.
(26) Palmö, K.; Pietilä, L.-O.; Krimm, S. *Comput. Chem.* **1993**, *17*, 67.
(27) Palmö, K.; Pietilä, L.-O.; Krimm, S. *J. Comput. Chem.* **1992**, *13*, 1142.
(28) Snyder, R. G. *J. Chem. Phys.* **1967**, *47*, 1316.
(29) Shimanouchi, T.; Matsuura, H.; Ogawa, Y.; Harada, I. *J. Phys. Chem. Ref. Data* **1978**, *7*, 1323.
(30) Mirkin, N. G.; Krimm, S. *J. Phys. Chem.*, in press.
(31) Warshel, A.; Lifson, S. *Chem. Phys. Lett.* **1969**, *4*, 255.
(32) Pietilä, L.-O.; Rasmussen, Kj. *J. Comput. Chem.* **1984**, *5*, 252.
(33) Pietilä, L.-O. *J. Mol. Struct.* **1989**, *195*, 111.
(34) Ermer, O.; Lifson, S. *J. Am. Chem. Soc.* **1973**, *95*, 4121.
(35) van Ness, G. J. H.; Vos, A. *Acta Crystallogr.* **1978**, *B34*, 1947.
(36) Mathisen, H.; Norman, N.; Pedersen, B. F. *Acta Chem. Scand.* **1967**, *21*, 127.
(37) Norman, N.; Mathisen, H. *Acta Chem. Scand.* **1961**, *15*, 1755.
(38) Shipman, L. L.; Burgess, A. W.; Scheraga, H. A. *Proc. Natl. Acad. Sci.* **1975**, *72*, 543.
(39) Warshel, A.; Lifson, S. *J. Chem. Phys.* **1970**, *53*, 582.
(40) Lii, J.-H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8566.
(41) Brunel, L.-C.; Dows, D. A. *Spectrochim. Acta* **1974**, *30A*, 929.
(42) Nelligan, W. B.; LePoire, D. J.; Brun, T. O.; Kleb, R. *J. Chem. Phys.* **1987**, *87*, 2447.
(43) Takeuchi, H.; Allen, G.; Suzuki, S.; Dianoux, A. J. *Chem. Phys.* **1980**, *51*, 197.
(44) Warshel, A. *J. Chem. Phys.* **1971**, *54*, 5324.
(45) Pietilä, L.-O.; Mannfors, B.; Palmö, K. *J. Mol. Struct.* **1990**, *218*, 315.