Electrostatic Model for Infrared Intensities in a Spectroscopically Determined Molecular Mechanics Force Field

KIM PALMO, SAMUEL KRIMM

Biophysics Research Division and Department of Physics, University of Michigan, 930 N. University, Ann Arbor, Michigan 48109

ABSTRACT: A new electrostatic model for the calculation of infrared intensities in molecular mechanics and molecular dynamics is presented. The model is based on atomic charges, atomic charge fluxes, and internal coordinate dipoles and their fluxes. The internal coordinate dipoles are used instead of atomic dipoles, thus simplifying the derivation of parameters. The model is designed to reproduce *ab initio* dipole derivatives, and the parameters can be obtained by (iterative) transformations from these, or by linear least squares fitting to them. A first application to linear alkanes has been made. For these molecules, the intensities can be predicted with an average accuracy of 30–40%. © 1998 John Wiley & Sons, Inc. J Comput Chem 19: 754–768, 1998

Keywords: molecular mechanics; spectroscopically determined force field; infrared intensities; electrostatic model; alkanes

Introduction

The goal of our development of a spectroscopically determined force field (SDFF)[1] is to provide a molecular mechanics (MM) potential energy function, in our case based on an analytical

transformation from *ab initio* quantities, that reproduces (in addition to structures and energies) vibrational frequencies of multiple conformations of macromolecules to spectroscopic accuracy (i.e., errors in the range of 5–10 cm^{-1}). We have shown that such an approach can be implemented in the case of linear[2] and branched[3] saturated hydrocarbons, and thus by extension to other macromolecular systems such as polypeptides.

It is highly desirable that such an SDFF also be compatible with a model that could reproduce vibrational intensities. Aside from its direct utility,

this would enable competent spectroscopic information to be derived from Fourier inversion of molecular dynamics (MD) simulations. In this article, we present an electrostatic model for infrared (IR) intensities that is consistent with the MM approach and that significantly improves upon previous intensity models.

It should be noted that traditional methods of reproducing IR intensities,[4] making use of atomic polar tensors or of electro-optical parameters, are not useful for MM functions because their selection of, for example, atomic point charges is independent of the (all-quadratic) spectroscopic force field used to describe the molecular vibrations. In the MM approach, on the other hand, the point charges are explicitly included in the energy function and are initially optimized to reproduce relative energies of different conformers. Furthermore, charge fluxes, which are also needed to reproduce IR intensities,[4] must be consistent with the intermolecular force field that gives rise to significant interaction band splittings and intermolecular normal modes, such as occur in polypeptides[5] and in the formic acid dimer.[6,7] Finally, we are constrained to find a model that not only provides satisfactory IR intensities for one molecule, which can always be done, but whose parameters are usefully transferable to the many other conformers of that class of molecules.

Because we have chosen to base our intensity model on *ab initio* quantities, in this case dipole derivatives, we must also be aware of the limitation on the number of parameters that can be determined. Thus, there are too many atomic dipole fluxes to be uniquely obtained from *ab initio* dipole derivatives whereas the internal coordinate dipole fluxes that we introduce can be directly obtained by transformation from such quantities.

In the following, we develop the theory underlying our model and this transformation, and then demonstrate its implementation for one system, the *n*-alkanes. The results show that such an electrostatic model is likely to be of sufficient generality to provide good semiquantitative IR intensities of macromolecular structures.

Theory

As usual,[4] the IR absorption intensity, I_{ir}, of a vibrating normal coordinate, Q_k, of a molecule is proportional to the square of the derivative of the molecular electric dipole moment with respect to Q_k; that is:

$$I_{ir} = A\left(\frac{\partial \vec{\mu}}{\partial Q_k}\right)^2 \qquad (1)$$

where $\vec{\mu}$ is the dipole moment and A is a constant. Most MM packages currently use atomic point charges to account for electrostatic intra- and intermolecular interactions and possibly for molecular dipoles. However, it is well known that static atomic point charges alone are incapable of producing correct dipole derivatives.[4] As a result, in addition to static atomic charges, atomic charge fluxes have been introduced.[4] Charge flux means that charge actually flows from one atom to another. As was mentioned earlier, charge fluxes can sometimes be determined from force constant data, or by fitting to band splittings and intermolecular normal modes. However, there may also be cases where the charge distributions around the atoms are deformed locally, without any net flow of charge along the bonds. This can be described in terms of atomic dipoles and atomic dipole fluxes. If no independent data are available, it may be difficult to differentiate between charge fluxes and atomic dipole fluxes in regard to their effects on the dipole derivatives. In the construction of an electrostatic model aimed at reproducing IR intensities, therefore, it seems reasonable to adopt an approach whereby charge fluxes are used only when there are clear indications that local charge redistributions are inadequate or specific spectroscopic effects require them. As we shall see, the local redistributions of charge are more convenient to describe in terms of internal coordinate dipole fluxes than in terms of atomic dipole fluxes.

We can understand the development of our model by noting that, if only atomic point charges are used, the dipole moment of a neutral molecule is:

$$\vec{\mu} = \sum_i q_i \vec{r}_i \qquad (2)$$

where q_i is the charge of atom i and $\vec{r}_i$ is the radius vector of atom i relative to some arbitrary origin. In molecular mechanics it is convenient, as a simple way to ensure that the molecule stays neutral, to express the charges in terms of incremental bond dipoles, such that every bond b is associated with an intrinsic dipole moment:

$$\vec{\mu}_b = q_b \vec{R}_b \qquad (3)$$

where q_b is the incremental charge on one atom of the bond, $-q_b$ being the incremental charge on the other atom, and $\vec{R}_b$ is a vector from $-q_b$ to $+q_b$. The molecular dipole moment is then the sum of all the $\vec{\mu}_b$ values, and the net charge of an atom is the sum of all the incremental charges on that atom. If the charges are not constant, but depend on the deformation of the internal coordinates, we can write:

$$q_b = q_{b,\,min} + \sum_j a_{bj}(S_j - S_{j,\,min}) \qquad (4a)$$

where S_j is an internal coordinate, $q_{b,\,min}$ is the value of q_b at the energy minimum $\vec{S}_{min}$, and:

$$a_{bj} = \frac{\partial q_b}{\partial S_j} \qquad (4b)$$

is the charge flux along bond b due to deformation of S_j. The molecular dipole moment is then

$$\begin{aligned}
\vec{\mu} &= \sum_b q_b \vec{R}_b \\
&= \sum_b \left[q_{b,\,min} + \sum_j a_{bj}(S_j - S_{j,\,min}) \right] \vec{R}_b \quad (5)
\end{aligned}$$

and the dipole derivatives at $\vec{S}_{min}$ in Cartesian coordinates are:

$$\begin{aligned}
\frac{\partial \vec{\mu}}{\partial x_{k\alpha}} &= \sum_b q_{b,\,min} \frac{\partial \vec{R}_b}{\partial x_{k\alpha}} \\
&+ \sum_b \left(\sum_j a_{bj} \frac{\partial S_j}{\partial x_{k\alpha}} \right) \vec{R}_b \qquad (6)
\end{aligned}$$

where k runs over the atoms and $\alpha = 1, 2, 3$. This can be transformed to normal coordinates using the relation:

$$\frac{\partial \vec{\mu}}{\partial Q_n} = \sum_{k\alpha} \frac{\partial \vec{\mu}}{\partial x_{k\alpha}} \frac{\partial x_{k\alpha}}{\partial Q_n} \qquad (7)$$

The charge and charge flux parameters in eq. (6) can be determined via eqs. (1) and (7) by fitting to observed IR absorption intensities. This method has been used by Machida and coworkers to compute parameters for alkanes,[8] aliphatic ethers,[9] and amides,[10,11] and also by Lii and Allinger in deriving some charge flux parameters for MM3.[12]

We believe, however, that a more fundamental approach, as well as better accuracy and transferability, is achieved if all parameters are determined from *ab initio* calculations. We have, there-fore, developed an electrostatic model designed to reproduce *ab initio* dipole derivatives. This is done in the context of an SDFF,[1] which is designed to provide accurate frequencies and therefore should produce more accurate eigenvectors. This is an important requirement if we are to obtain meaningful model parameters. The force field used in this work, SDFF97,[3] is an upgraded and extended version of our original SDFF for linear alkanes.[2]

In a first attempt to implement this approach, we optimized charge and charge flux parameters to *ab initio* HF/6-31G dipole derivatives of *n*-pentane. As expected, the resulting parameters reproduce the IR intensities satisfactorily. Using these parameters, though, we were unable to reproduce the relative energies and torsion barriers of *n*-pentane and *n*-hexane,[2] because the static charges, $q_{b,\,min}$, came out too small compared with those required by our energy function.[2,3] To yield correct energies, the charges on the hydrogens in our successful SDFF[2,3] should be close to 0.1e, whereas the above fit to the dipole derivatives gave 0.02e or less. (It is of interest that the value of 0.1e is close to the "effective charge," calculated from an experimental dipole derivative matrix, found for H atoms in CH_4, CH_3CH_3, and $CH_3CH_2CH_3$,[13] and close to Mulliken charges that we find in MP2/6-31 + + G** calculations.) On the other hand, fixing these charges to the larger values resulted in poor IR intensities. Because some of the relative energies and barriers in *n*-pentane and *n*-hexane are almost entirely due to electrostatic interactions (in the context of our torsions and van der Waals interactions[2,3]), we therefore concluded that the charge–charge flux model was not adequate to account for IR intensities, and that at least atomic dipoles and atomic dipole fluxes also had to be added in some form to the model.

A general linear expression for an atomic dipole moment $\vec{m}_i$ is:

$$\vec{m}_i = \vec{m}_{i,\,min} + \sum_j c_{ij}(S_j - S_{j,\,min})\hat{e}_{ij} \qquad (8)$$

where $\vec{m}_{i,\,min}$ is the atomic dipole moment at $\vec{S}_{min}$ and c_{ij} are atomic dipole flux parameters that together with the unit vectors, $\hat{e}_{ij}$, determine how the atomic dipole moment changes upon deformation of the internal coordinates. Unfortunately, eq. (8) is impractical to use directly because it contains too many new parameters (three parameters per atom for the first term and three parameters per atom per internal coordinate for the second term). We also tested the approximation that $\vec{m}_{i,\,min}$ and

$\hat{e}_{ij}$ only have components along the bonds of atom i, but found this to be too crude. However, in order to reduce the number of new parameters and avoid the inevitably strong correlations between the parameters of eq. (5) and virtually any meaningful approximation of eq. (8), it is possible to replace the atomic dipoles and dipole fluxes with *internal coordinate* dipoles and dipole fluxes. For any neutral molecule, the molecular dipole moment $\vec{\mu}$ in the neighborhood of a local energy minimum $\vec{S}_{min}$ can then be described by a set of bond increment charges q_b and a set of internal coordinate dipoles $\vec{\mu}_j$, such that:

$$\vec{\mu} = \sum_b q_b \vec{R}_b + \sum_j \vec{\mu}_j \tag{9}$$

In the linear approximation, $\vec{\mu}_j$ depends on the internal coordinate S_j according to:

$$\vec{\mu}_j = \vec{\mu}_{j,min} + \vec{d}_j (S_j - S_{j,min}) \tag{10a}$$

where the vector:

$$\vec{d}_j = \frac{\partial \vec{\mu}_j}{\partial S_j} \tag{10b}$$

is the internal coordinate dipole flux associated with S_j. Taking into account that the charges q_b are not constants but depend on the internal coordinates as given by eq. (4), the dipole derivatives at the energy minimum can be written:

$$\frac{\partial \vec{\mu}}{\partial x_{k\alpha}} = \sum_b q_{b,min} \frac{\partial \vec{R}_b}{\partial x_{k\alpha}}$$

$$+ \sum_{b,j} a_{bj} \frac{\partial S_j}{\partial x_{k\alpha}} \vec{R}_b$$

$$+ \sum_j \left(\frac{\partial \vec{\mu}_{j,min}}{\partial x_{k\alpha}} + \vec{d}_j \frac{\partial S_j}{\partial x_{k\alpha}} \right) \tag{11}$$

(Note that even though $\vec{\mu}_{j,min}$ is a static internal coordinate dipole moment, its direction depends on the Cartesian coordinates.) Eq. (11) can be interpreted as a transformation of the derivatives of a dipole moment from internal to Cartesian coordinates. To see this we write the equation in the following form:

$$\frac{\partial}{\partial x_{k\alpha}} \left[\vec{\mu} - \sum_b q_{b,min} \vec{R}_b - \sum_{b,j} a_{bj}(S_j - S_{j,min}) \vec{R}_b \right.$$

$$\left. - \sum_j \vec{\mu}_{j,min} \right] = \sum_j \vec{d}_j \frac{\partial S_j}{\partial x_{k\alpha}} \tag{12}$$

The expression inside the brackets on the left-hand side:

$$\vec{M}_s = \vec{\mu} - \sum_b q_{b,min} \vec{R}_b - \sum_{b,j} a_{bj}(S_j - S_{j,min}) \vec{R}_b$$

$$- \sum_j \vec{\mu}_{j,min} \tag{13}$$

is the remaining dipole moment when the effects of static charges, charge fluxes, and static internal coordinate dipoles are subtracted from the total molecular dipole moment. Thus, in our current model, $\vec{M}_s$ must be:

$$\vec{M}_s = \sum_j \vec{d}_j (S_j - S_{j,min}) \tag{14}$$

Obviously:

$$\vec{d}_j = \frac{\partial \vec{M}_s}{\partial S_j} \tag{15}$$

so that eq. (12) can be written

$$\frac{\partial \vec{M}_s}{\partial x_{k\alpha}} = \sum_j \frac{\partial \vec{M}_s}{\partial S_j} \frac{\partial S_j}{\partial x_{k\alpha}} \tag{16}$$

This is the transformation of the derivatives of $\vec{M}_s$ from internal to Cartesian coordinates. If we invert the transformation (possible, because, with the Eckart conditions, the matrix is square), we get:

$$\frac{\partial \vec{M}_s}{\partial S_j} = \sum_{k\alpha} \frac{\partial \vec{M}_s}{\partial x_{k\alpha}} \frac{\partial x_{k\alpha}}{\partial S_j} \tag{17}$$

or, using eqns. (13) and (14):

$$\vec{d}_j = \sum_{k\alpha} \left(\frac{\partial \vec{\mu}}{\partial x_{k\alpha}} - \sum_b q_{b,min} \frac{\partial \vec{R}_b}{\partial x_{k\alpha}} \right.$$

$$\left. - \sum_{b,i} a_{bi} \frac{\partial S_i}{\partial x_{k\alpha}} \vec{R}_b - \sum_i \frac{\partial \vec{\mu}_{i,min}}{\partial x_{k\alpha}} \right) \frac{\partial x_{k\alpha}}{\partial S_j} \tag{18}$$

Thus, if we know the static charges, the charge fluxes, and the static internal coordinate dipoles, we can directly compute the internal coordinate dipole fluxes $\vec{d}_j$ from the *ab initio* dipole derivatives. This is useful if the static charges and dipoles can be determined from energy considerations, or from *ab initio* electrostatic potentials, and (at least the most important) charge fluxes can be determined from other effects, such as shifts in vibrational frequencies and intermolecular modes.[6,14] Even doing the transformation [eq. (18)] using static charges and dipoles only, could give valu-

able clues as to what kind of charge fluxes are needed to reproduce the dipole derivatives properly. For any molecule, values for the $\vec{d}_j$ can always be directly computed from an *ab initio* dipole derivative matrix. Any set of $\vec{d}_j$, $q_{b,min}$, a_{bj}, and $\vec{\mu}_{j,min}$ compatible with eq. (18) will reproduce the *ab initio* dipole derivatives, and hence the IR intensities, exactly. However, we also wish the parameters to be determined so as to be as transferable as possible between a set of conformers of a given class of molecules. The specific kinds of parameters optimized for such a class may well depend on which are found to be important in reproducing IR intensities, and we should be aware that this may differ between different kinds of molecules.

Finally, we note that in molecular mechanics we want to be able to compute molecular dipole moments not only at the energy minima but for any geometry. This is especially important if we want to calculate IR intensities in MD simulations. Thus, to achieve greater generality, we need to change the model slightly and replace the quantities $S_{j,min}$ in eqs. (4) and (10) with the MM intrinsic geometry parameters S_{j0}. The molecular dipole moment is then:

$$\vec{\mu} = \sum_b \left[q_{b0} + \sum_j a_{bj}(S_j - S_{j0}) \right] \vec{R}_b + \sum_j \left[\vec{\mu}_{j0} + \vec{d}_j(S_j - S_{j0}) \right] \quad (19)$$

where:

$$q_{b0} = q_{b,min} + \sum_j a_{bj}(S_{j0} - S_{j,min}) \quad (20)$$

and:

$$\vec{\mu}_{j0} = \vec{\mu}_{j,min} + \vec{d}_j(S_{j0} - S_{j,min}) \quad (21)$$

For the dipole derivatives we get, instead of eq. (11):

$$\frac{\partial \vec{\mu}}{\partial x_{k\alpha}} = \sum_b \left[q_{b0} + \sum_j a_{bj}(S_j - S_{j0}) \right] \frac{\partial \vec{R}_b}{\partial x_{k\alpha}}$$
$$+ \sum_{bj} \left[a_{bj} \frac{\partial S_j}{\partial x_{k\alpha}} + \frac{\partial a_{bj}}{\partial x_{k\alpha}}(S_j - S_{j0}) \right] \vec{R}_b$$
$$+ \sum_j \left[\frac{\partial \vec{\mu}_{j0}}{\partial x_{k\alpha}} + \vec{d}_j \frac{\partial S_j}{\partial x_{k\alpha}} \right]$$
$$+ \frac{\partial \vec{d}_j}{\partial x_{k\alpha}}(S_j - S_{j0}) \right] \quad (22)$$

where we have also included the possibility that the charge fluxes a_{bj} may depend on conformation. For a given set of parameters, eq. (22) should be valid over a much larger range of internal coordinate deformations than eq. (11), because eq. (22) takes the deformations explicitly into account. However, eq. (22) does not lead to a direct expression for the $\vec{d}_j$ in the case of known charges, charge fluxes, and static internal coordinate dipoles. Thus, instead of eq. (18), we now get:

$$\vec{d}_j = \sum_{k\alpha} \left[\frac{\partial \vec{\mu}}{\partial x_{k\alpha}} \right.$$
$$- \sum_b \left[q_{b0} + \sum_i a_{bi}(S_i - S_{i0}) \right] \frac{\partial \vec{R}_b}{\partial x_{k\alpha}}$$
$$- \sum_{bi} \left[a_{bi} \frac{\partial S_i}{\partial x_{k\alpha}} + \frac{\partial a_{bi}}{\partial x_{k\alpha}}(S_i - S_{i0}) \right] \vec{R}_b$$
$$- \sum_i \frac{\partial \vec{\mu}_{i0}}{\partial x_{k\alpha}} - \frac{\partial \vec{d}_i}{\partial x_{k\alpha}}(S_i - S_{i0}) \right] \frac{\partial x_{k\alpha}}{\partial S_j} \quad (23)$$

which does not yield $\vec{d}_j$ directly because the last term contains the derivatives of the $\vec{d}_i$. Therefore, in this case, the $\vec{d}_j$ have to be computed iteratively. This requires that a functional form of the $\vec{d}_j$, such as eq. (24) (next section), first be established. Alternatively, the parameters can be determined in a linear least-squares fit to the *ab initio* dipole derivatives of a sufficiently large set of model molecules.

Even though eqs. (22) and (23) are more general, eqs. (11) and (18) are much simpler to use. Especially for nonstrained molecules, eq. (18) provides a quick method to evaluate the transferability of the $\vec{d}_j$, and to obtain information about the nature of charge fluxes needed to improve the transferability.

Application to Linear Alkanes

As a first application we have determined electrostatic parameters for a series of linear alkanes using *ab initio* HF/6-31G dipole derivatives. The reason for using HF/6-31G derivatives is that our SDFF for alkanes is based on structures and scaled force fields computed with this basis set.[15] The following 12 molecular structures were included: ethane, propane, both conformations of butane, all four conformations of pentane, three conforma-

tions of hexane (*ttt*, *ttg*, *tgg*), and one conformation of heptane (*tttt*). The *ab initio* dipole derivatives were calculated at energy-minimized geometries using Gaussian-94.[16]

Because all the molecules in the series are relatively nonstrained, we started the calculations by an evaluation of the internal coordinate dipole fluxes, $\vec{d}_j$, using eq. (18). For the static atomic charges we used the ones proven valid for relative energies and barriers;[2,3] that is, 0.1e for hydrogen, and −0.3e and −0.2e for methyl and methylene carbons, respectively. As expected, the IR intensities produced by only these atomic charges are extremely crude and in many cases deviate from (exceed) the *ab initio* values by an order of magnitude. The static internal coordinate dipoles in this case were assumed to be negligible, because the energies of, and barriers between, conformers could be well accounted for by these charges. We note that this may not be true in other systems.[6] The usual redundant internal coordinates were used. Thus, all six valence angles around an sp^3 carbon atom were included. The local redundancies, which have to be taken into account when using eq. (18), were treated with the method described by Pulay and Fogarasi.[17] The values for the $\vec{d}_j$ thus obtained were then transformed into local internal coordinate systems. For each bond and angle three mutually perpendicular unit vectors were defined. For bonds, these consist of one vector along the bond, and two vectors (one arbitrary) perpendicular to the bond. For angles, the first vector bisects the angle, the second vector lies in the plane of the angle (perpendicular to the first vector), and the third vector is the vector product of the first and the second vectors. Torsion dipole flux was found to be insignificant in this class of molecules and was therefore neglected. In each case the internal coordinate dipole flux can then be written:

$$\vec{d}_j = \sum_{i=1}^{3} d_{ji}\hat{e}_{ji} \qquad (24)$$

where d_{ji} are scalar constants and $\hat{e}_{ji}$ are the internal coordinate unit vectors.

The initial calculations showed that no bonds other than C—H bonds had significant dipole flux, and that the components perpendicular to a bond were small compared to the parallel component but still non-negligible. All angles were found to have significant dipole flux, the biggest components always being along the bisector of the angle, but with nonzero contributions also in the other directions. Internal coordinate dipole flux compo-

nents that are perpendicular to the bond or the angle are inconvenient to use in molecular mechanics, because such components are bound to depend significantly on atoms other than those forming the bond or the angle. This makes it more difficult to obtain transferable parameters. We therefore decided to have the charge flux terms account for the perpendicular dipole flux components of bonds and angles. This is a somewhat arbitrary choice, but for alkanes there is no easy way to distinguish between dipole derivative contributions stemming from charge fluxes on the one hand, and from internal coordinate (or atomic) dipole fluxes on the other. (For other molecules, charge fluxes may be constrained by vibrational properties such as band splittings and intermolecular modes; in such cases, it may be necessary to include some perpendicular internal coordinate fluxes.) Test calculations on the small molecules (ethane, propane, butane) showed that the following three types of charge flux were needed: (1) nearest-neighbor bond–bond; (2) bond–angle, where the bond shares the central atom of the angle, but is not part of the angle; and (3) bond–angle, where the bond shares one of the angle's end atoms, but is not part of the angle. Charge flux terms of type (3) also turned out to be conformation-dependent. If the angle is formed by the atoms ABC and the bond is between atoms C and D, then the charge flux parameter depends on the dihedral angle ϕ = ABCD according to:

$$a_{bj} = a_{bj,0} \cos^2 \phi \qquad (25)$$

Charge flux terms beyond these types do not improve the transferability of the dipole fluxes. The following sign convention is used in the charge flux terms. For a charge increment of $\Delta q = a_{bj}(S_j - S_{j0})$, the atom that the bond $\vec{R}_b$ shares with the internal coordinate S_j [cf. eq. (19)] is given a charge increment of $+\Delta q$, whereas the other atom of $\vec{R}_b$ is given a charge increment of $-\Delta q$.

After the initial evaluation, eq. (22) was used to further optimize some of the bond and angle dipole fluxes, together with charge flux terms of the types just mentioned. The optimization was a linear least squares fit to the *ab initio* dipole derivatives. To ensure proper weighting of the data points, both sides of eq. (22) were transformed to normal coordinates using eq. (7), and the transformed dipole derivatives were the ones actually used in the fit. Molecular mechanics eigenvectors were used, calculated at the HF/6-31G equilibrium geometries.

TABLE I.
SDFF Electrostatic Parameters for Linear Alkanes.[a]

Bond[b]	Static charge (q_{b0})	Bond dipole flux (d_j)
C—H	0.1	−0.3207

	Angle dipole flux	
Angle[b]	Along bisector (d_{j1})	Perpendicular to bisector (d_{j2})[c]
C—C—C	0.0770	0.0
C_3—C_3—H	0.0450	0.0
C_2—C_3—H	0.0535	0.0
C_3—C_2—H	0.0780	0.0191
C_2—C_2—H	0.0657	0.0
H—C_3—H	0.0550	0.0
H—C_2—H	0.0640	0.0

Bond–bond charge flux (a_{bj})[d]

C—C / C_3—C_2	−0.0444
C—C / C_2—C_2	−0.0170
C_3—H / C_3—H	0.0389
C_2—H / C_2—H	0.0296

Bond–angle charge flux

Central atom shared (a_{bj})

C_3—C_2 / H—C_3—H	0.0086
C_2—H / C—C—H	0.0106
C_3—H / H—C_3—H	0.0222

End atom shared ($a_{bj,0}$)[e]

C—C / C—C—C	0.0410
C—C / C—C—H	0.0148
C_2—H / C—C—H	−0.0106

[a] The units are: static charge, e; bond dipole flux, e; angle dipole flux, eÅ / rad; bond–bond charge flux, e / Å; bond–angle charge flux, e / rad.
[b] C = methylene or methyl carbon atom; C_2 = methylene carbon atom; C_3 = methyl carbon atom.
[c] If the angle is given as A—B—C, the positive direction points away from atom C.
[d] The first mentioned bond dipole changes when the second bond is deformed.
[e] Depends on conformation as given by eq. (25).

Test calculations made with eigenvectors corresponding to the full (scaled) ab initio force field[15] showed that the optimization process is not very sensitive to the eigenvectors (although the intensities themselves are). The final electrostatic parameters are given in Table I. It should be noted that this set of parameters is not unique. The goal was to make it as simple as possible, but still compatible with ab initio dipole derivatives and relative energies and barriers.

For some of the molecules, ab initio and MM vibrational frequencies and IR intensities are given in Table II. The MM frequencies and intensities (labeled "SDFF") should be compared with those produced by the ab initio dipole derivatives and scaled force fields (labeled "ab initio"). The SDFF columns are stand-alone MM results, calculated at the SDFF energy-minimized geometries. At this stage, however, we have not included the charge fluxes and internal coordinate dipole fluxes in the energy function. (In the alkanes, such inclusions have minimal effect on the charges and the energies, and only a modest effect on the intrinsic force constants. In any case, because the second derivatives of the potential energy are required to reproduce the scaled ab initio force constants,[1] the eigenvectors would be essentially unchanged. However, in peptides and other polar molecules these terms may give rise to much larger energy contributions and must therefore be included in the energy function.) Because the intensities are very sensitive, not only to the electrostatic model but also to the vibrational normal modes, we have included, in Table II, the column labeled "ab initio/SDFF," which gives the intensities produced by the SDFF electrostatic model when the eigenvectors from the scaled ab initio force fields are used. Thus, by comparing the "ab initio" and "ab initio/SDFF" columns, one can assess the quality of the electrostatic model under the assumption that the ab initio normal modes are reproduced perfectly.

There are, indeed, some large differences between the ab initio and SDFF intensities that are caused by changes in the eigenvectors rather than by failure of the electrostatic model. Examples of this are the 1333 cm^{-1} band in propane and the 1296 cm^{-1} band in t-butane. For these bands the electrostatic model predicts the intensities to within 20% of the ab initio values when used with the ab initio eigenvectors, but is off by more than a factor of 2 when the SDFF eigenvectors are used. This shows the critical importance of good eigenvectors when calculating vibrational intensities. In cases where several similar normal modes have approximately the same frequency, the composition of the eigenvectors is extremely sensitive to small changes in geometry and/or force field. This is clearly seen in the C—H stretching region for the molecules larger than butane. The SDFF eigenvectors here are significantly different from the ab initio eigenvectors. In the SDFF intensity column in Table II we therefore show intensities integrated

TABLE II.
Ab Initio and SDFF Vibrational Frequencies and IR Intensities.

Frequency (cm^{-1})		Infrared intensity (10^2 m / mol)		
Ab initio[a]	SDFF[b]	*Ab initio*[c]	*Ab initio* / SDFF[d]	SDFF[b]
Ethane				
312	283	0	0	0
808	811	103	91	83
991*	996	0	0	0
1192*	1199	0	0	0
1379	1406	46	40	40
1380*	1368	0	0	0
1464*	1457	0	0	0
1467	1446	213	200	203
2901	2883	720	713	710
2908*	2890	0	0	0
2952*	2944	0	0	0
2976	2949	1928	1668	1673
Propane				
227*	217	0	0	0
282	277	0	0	0
370	360	1	2	2
739	744	38	34	29
861	857	10	15	15
894*	885	0	0	0
913	922	28	13	12
1055	1046	7	10	6
1159	1162	18	17	19
1189	1186	2	4	4
1284*	1283	0	0	0
1333	1323	25	30	61
1368	1374	58	41	49
1384	1375	42	57	13
1455*	1454	0	0	0
1457	1459	0	8	5
1460	1455	29	20	22
1469	1460	143	117	117
1471	1451	61	62	74
2868	2860	459	409	518
2893	2893	183	163	155
2895	2892	466	417	464
2900	2896	171	203	389
2952*	2951	0	0	0
2960	2951	406	299	254
2963	2953	1254	1151	930
2963	2955	706	642	577
t-Butane				
123	118	0	0	0
231	231	0	0	0
262	253	1	5	4
267*	264	0	0	0
425*	424	0	0	0
726	730	39	36	30
797*	795	0	0	0

(continued)

TABLE II.
(Continued).

Frequency (cm^{-1})		Infrared intensity (10^2 m / mol)		
Ab initio[a]	SDFF[b]	Ab initio[c]	Ab initio / SDFF[d]	SDFF[b]
t-Butane				
828*	825	0	0	0
945	935	12	9	8
964	968	85	77	73
1009	1008	1	0	1
1062*	1056	0	0	0
1157*	1157	0	0	0
1186*	1173	0	0	0
1263	1265	1	1	1
1296	1289	11	13	26
1300*	1308	0	0	0
1352*	1341	0	0	0
1376*	1378	0	0	0
1377	1371	92	91	79
1453*	1452	0	0	0
1456	1458	24	22	22
1462*	1454	0	0	0
1463	1452	152	118	111
1465*	1456	0	0	0
1471	1458	93	107	118
2857*	2866	0	0	0
2862	2863	816	782	944
2875*	2893	0	0	0
2897	2892	768	697	618
2899*	2898	0	0	0
2899	2904	485	470	836
2956*	2952	0	0	0
2960	2953	1390	1324	966
2961*	2952	0	0	0
2962	2952	1090	942	836
g-Butane				
115	112	0	0	0
221	222	0	0	0
265	265	1	1	1
328	321	0	0	0
435	427	4	3	2
740	745	45	30	26
781	780	18	19	17
821	823	1	1	1
953	946	48	49	42
957	953	4	9	12
977	990	3	1	1
1085	1078	1	1	2
1139	1141	28	16	16
1177	1175	1	2	3
1260	1254	3	1	1
1284	1278	4	5	5
1338	1338	7	1	6
1346	1342	15	7	38

(continued)

TABLE II.
(Continued).

Frequency (cm^{-1})		Infrared intensity (10^2 m / mol)		
Ab initio[a]	SDFF[b]	*Ab initio*[c]	*Ab initio* / SDFF[d]	SDFF[b]
g-Butane				
1375	1376	71	36	11
1379	1379	56	55	51
1452	1458	0	14	0
1453	1466	4	12	18
1462	1456	82	55	63
1466	1462	50	32	50
1468	1452	133	117	68
1472	1451	20	6	18
2856	2858	345	311	359
2861	2895	252	219	251
2888	2864	655	567	617
2894	2894	282	251	377
2900	2899	616	547	632
2903	2896	83	126	179
2958	2952	2	1	36
2959	2952	1144	1019	778
2965	2957	703	587	570
2974	2959	517	586	416
tt-Pentane				
109*	109	0	0	0
115	114	0	0	0
181	174	0	2	2
245	247	0	0	0
254*	254	0	0	0
396	398	0	0	0
399	396	0	3	4
722	728	42	40	33
752*	753	0	0	0
852	846	14	12	11
861	860	13	22	21
916	919	37	19	20
977*	965	0	0	0
1024	1025	28	13	12
1042	1040	4	4	3
1073	1066	11	6	3
1153	1150	13	15	18
1185	1168	1	2	2
1242*	1242	0	0	0
1269	1263	6	6	9
1297	1300	7	0	0
1301*	1315	0	0	0
1334	1323	2	2	10
1356	1347	23	19	27
1377	1374	45	52	41
1378	1376	30	24	17
1452	1450	0	3	0
1454	1452	11	4	27

(continued)

TABLE II.
(Continued).

Frequency (cm^{-1})		Infrared intensity (10^2 m / mol)		
Ab initio[a]	SDFF[b]	Ab initio[c]	Ab initio / SDFF[d]	SDFF[b]
tt-Pentane				
1460	1455	1	9	2
1462	1457	152	118	118
1463*	1457	0	0	0
1466	1458	24	23	35
1472	1453	92	112	95
2849	2871	137 ⎫	151 ⎫	34 ⎫
2855	2864	6 ⎬ 1234	4 ⎬ 1260	3 ⎬ 1406
2863	2865	1091 ⎭	1105 ⎭	1369 ⎭
2866	2892	0 ⎫	0 ⎫	146 ⎫
2882*	2901	0 ⎪	0 ⎪	0 ⎪
2898	2892	557 ⎬ 1647	426 ⎬ 1571	470 ⎬ 1922
2899	2900	213 ⎪	255 ⎪	238 ⎪
2903	2908	877 ⎭	890 ⎭	1068 ⎭
2958*	2952	0 ⎫	0 ⎫	0 ⎫
2959	2952	1420 ⎬ 2538	1380 ⎬ 2330	978 ⎬ 1833
2962	2952	440 ⎪	279 ⎪	232 ⎪
2962	2952	678 ⎭	671 ⎭	623 ⎭
gg′-Pentane				
83	81	0	0	0
135	128	0	0	0
218	219	0	0	0
269	280	2	1	1
303	299	0	0	0
359	343	3	2	1
451	449	2	4	3
731	744	33	22	20
754	752	17	15	12
813	814	16	16	20
865	867	13	17	15
889	886	28	32	31
997	991	19	9	9
1013	1014	12	6	13
1023	1023	8	8	5
1106	1100	4	3	1
1131	1132	17	11	12
1168	1169	2	4	8
1237	1237	6	5	4
1269	1267	11	12	12
1289	1290	5	6	8
1339	1337	14	4	20
1341	1353	4	7	11
1353	1364	11	4	10
1377	1377	34	7	24
1378	1386	90	79	44
1453	1462	8	9	5
1454	1466	2	6	1

(continued)

TABLE II.
(Continued).

Frequency (cm^{-1})		Infrared intensity (10^2 m / mol)			
Ab initio[a]	SDFF[b]	Ab initio[c]	Ab initio / SDFF[d]	SDFF[b]	

gg′-Pentane

Ab initio[a]	SDFF[b]	Ab initio[c]		Ab initio / SDFF[d]		SDFF[b]	
1458	1472	19		28		7	
1463	1453	36		21		22	
1466	1452	120		61		81	
1470	1457	25		32		31	
1474	1478	120		84		87	
2850	2862	278 ⎫		286 ⎫		773 ⎫	
2857	2856	275 ⎬	716	231 ⎬	673	404 ⎬	1443
2860	2852	163 ⎭		156 ⎭		266 ⎭	
2882	2896	373 ⎫		256 ⎫		219 ⎫	
2892	2899	600 ⎪		608 ⎪		442 ⎪	
2894	2901	894 ⎬	2531	834 ⎬	2321	769 ⎬	1905
2902	2893	260 ⎪		196 ⎪		214 ⎪	
2906	2900	404 ⎭		427 ⎭		261 ⎭	
2957	2950	422 ⎫		363 ⎫		441 ⎫	
2959	2963	774 ⎪		673 ⎪		474 ⎪	
2972	2952	743 ⎬	2313	663 ⎬	2151	451 ⎬	1799
2986	2977	374 ⎭		452 ⎭		433 ⎭	

[a] Calculated at HF / 6-31G-energy-minimized geometry using scaled force field (scale factors from ref. 11). Frequencies marked with an asterisk are infrared inactive by symmetry.
[b] Molecular mechanics calculation at SDFF-energy-minimized geometry.
[c] HF / 6-31G dipole derivatives and geometry. Eigenvectors calculated using scaled force field.
[d] Eigenvectors calculated at HF / 6-31G-energy-minimized geometry using scaled force field. SDFF electrostatic parameters.

TABLE III.
Intensity Deviations (rms) in Different Ranges (10^2 m / mol).

Molecule	Non-C—H stretching (intensity range[a])				C—H stretching (all ranges)
	0–10	10–50	50–100	100+	
Ethane	–	6	–	16	181
Propane	3	20	12	26	168
t-Butane	2	9	8	41	285
g-Butane	5	13	37	65	135
tt-Pentane	4	11	3	34	448
tg-Pentane	4	7	34	41	471
gg-Pentane	3	14	26	53	461
gg′-Pentane	3	8	46	36	629
ttt-Hexane	3	18	21	21	486
ttg-Hexane	1	11	23	65	516
tgg-Hexane	4	9	30	69	544
tttt-Heptane	3	15	12	21	712
All molecules	3	12	27	42	405
rms dev%[b]	66	39	37	21	30
Ave dev%[b]	40	30	32	18	23
No. of freqs.	153	111	26	16	48

[a] Intensities that are zero by symmetry are not included.
[b] The dev% values are defined as $100 * \text{dev} / I_{mid}$, where I_{mid} is the average of the highest and the lowest intensity found in the range.

over the respective C—H stretching frequency intervals. Also, in this case, as can be seen from the ''*ab initio*/SDFF'' column, the electrostatic model itself is in good agreement with *ab initio*. The region near 1450 cm^{-1} is another crowded area. Here, many methyl and methylene bending modes occur with closely spaced frequencies.

The root-mean-square (rms) intensity deviations of SDFF compared with *ab initio* in different intensity ranges are given in Table III, separately for each molecule. With respect to the average deviations, the *ab initio* IR intensities are reproduced with an accuracy of 30–40%, which in many cases is sufficient for semiquantitative predictions. The electrostatic model as such is capable of significantly better predictions, but its performance is limited by the less-than-perfect SDFF eigenvectors (and probably by the restriction in the kinds and number of parameters in the model). This is seen by comparing the dev% values in Table III with those obtained using *ab initio*/SDFF data (defined as in Table II) instead of pure SDFF data. The corresponding rms dev% (ave dev%) values are 59(36), 31(24), 28(23), 14(12), and 8(7) for the 0–10, 10–50, 50–100, 100 + , and C—H stretching intensity ranges, respectively, which represents an average improvement of about 22% for non-C—H stretching modes.

To give an impression of what the current SDFF predictions would look like, we have simulated a few spectra assuming Lorentzian band shapes with half-widths of 10 cm^{-1}. Figures 1–4 show a comparison between *ab initio* and SDFF-simulated spectra for ethane, *t*-butane, *g*-butane, and *tgg*-hexane. The spectral features are reasonably well accounted for, especially below 1300 cm^{-1}. Of course, a condition for a simulation to be even qualitatively successful is that the positions of the bands, that is, the vibrational frequencies, are accurately reproduced. The rms frequency deviation below 1500 cm^{-1} for the molecules treated in this work (SDFF compared with *ab initio*) is 6.9 cm^{-1}.

Conclusions

We have presented a new electrostatic model, and parameters for linear alkanes, which make semiquantitative predictions of IR spectra possible in molecular mechanics and molecular dynamics calculations. Such predictions give significantly better results than previous MM models.
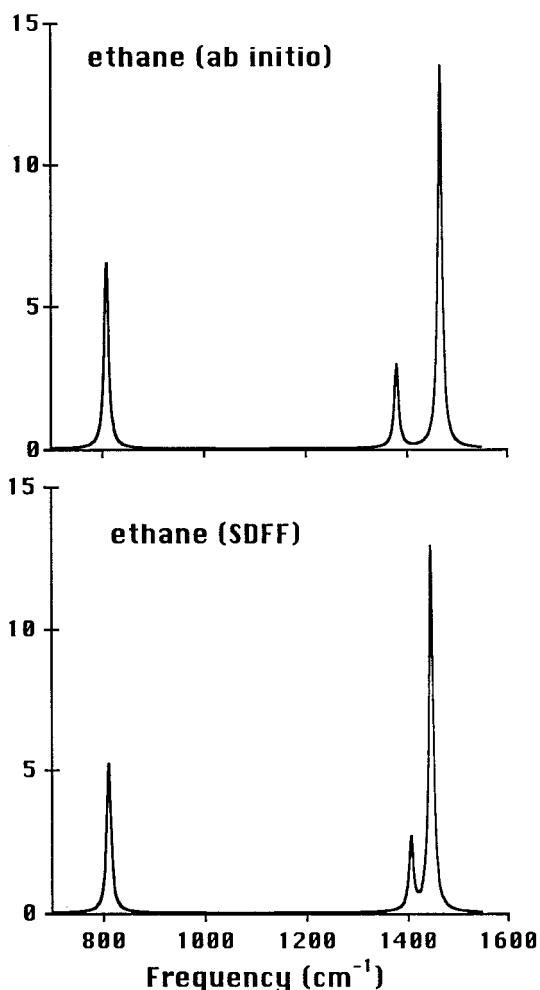

FIGURE 1. Simulated *ab initio* (top) and SDFF (bottom) infrared spectra of ethane. The bands consist of Lorentzian profiles with half-widths of 10 cm^{-1}. The *ab initio* intensities were calculated using HF / 6-31G dipole derivatives and scaled force fields. Intensities are in units of 10^2 m / mol.

Our model of course differs from traditional (i.e., atomic polar tensor or electro-optical parameter) models in that these are based on spectroscopic force fields. Also, we have included angle dipole fluxes in addition to bond dipole fluxes to account for dipole derivatives. The parameters in our model can conveniently be obtained from *ab initio* dipole derivatives by linear least-squares fitting. The main new quantities presented here, internal coordinate dipole fluxes, can also be determined by direct transformation (or iteratively) from *ab initio* dipole derivatives if the other parameters are known. In planar molecules or groups, internal coordinate dipole fluxes probably provide the easi-
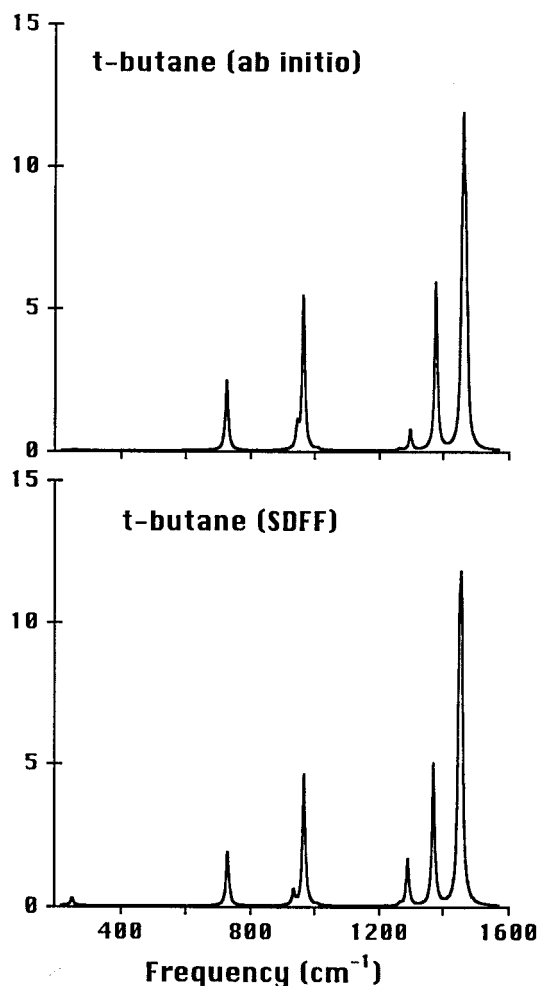
FIGURE 2. Simulated *ab initio* (top) and SDFF (bottom) infrared spectra of *t*-butane. Description as in Figure 1.
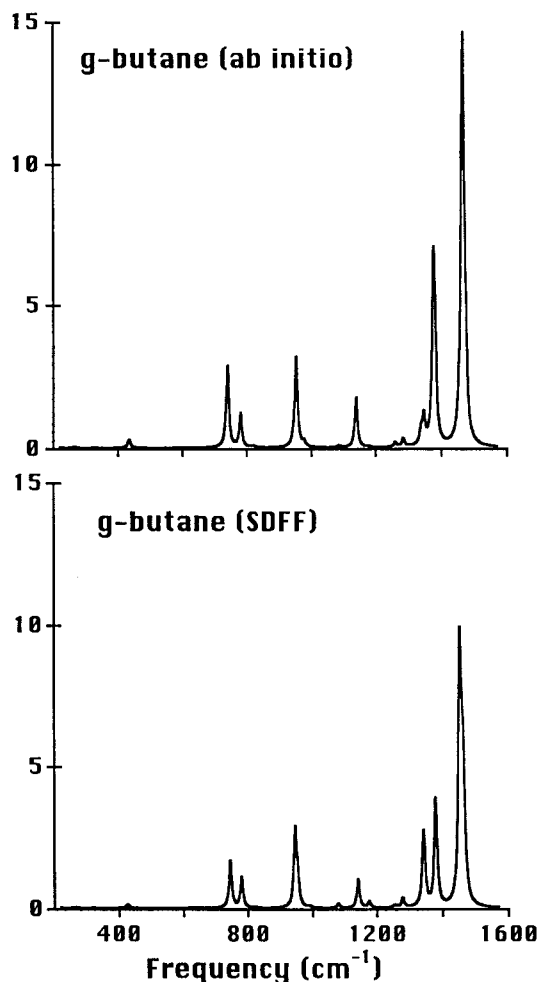
FIGURE 3. Simulated *ab initio* (top) and SDFF (bottom) infrared spectra of *g*-butane. Description as in Figure 1.

est way to describe out-of-plane charge redistribution, because charge fluxes do not have out-of-plane components. To be of any value, such an electrostatic model must be combined with force fields, such as SDFFs,[2,3] which yield very good vibrational frequencies and eigenvectors.

The SDFF IR intensities can be improved by calculating better normal modes or by a more comprehensive electrostatic model or parameters. For example, there may be room for improvement in the conformation dependence of the parameters. Although torsion dipole fluxes in alkanes are negligible for small deformations, it is still possible that the static charges have some systematic conformation dependence that is not taken properly into account by the current set of charge fluxes. *Ab initio* investigations of electrostatic potentials of

alkanes even seem to indicate that a net atomic charge model is not adequate and that additional charge centers are needed.[18,19] However, there are also indications that the alkanes are an exception in this regard, because, although we find that atomic point charges in *N*-methylacetamide can reproduce the *ab initio* (HF/6-31 + G**) electrostatic potential (on a Merz–Singh–Kollman grid) with a relative rms error of 7%, the corresponding value for *tt*-pentane is 85%. Another difficulty in improving the electrostatic model for alkanes is that charge fluxes and internal coordinate (or atomic) dipole fluxes affect the dipole derivatives in the same way and cannot easily be distinguished.

A variation of the method described in this work can be used for the determination of parame-
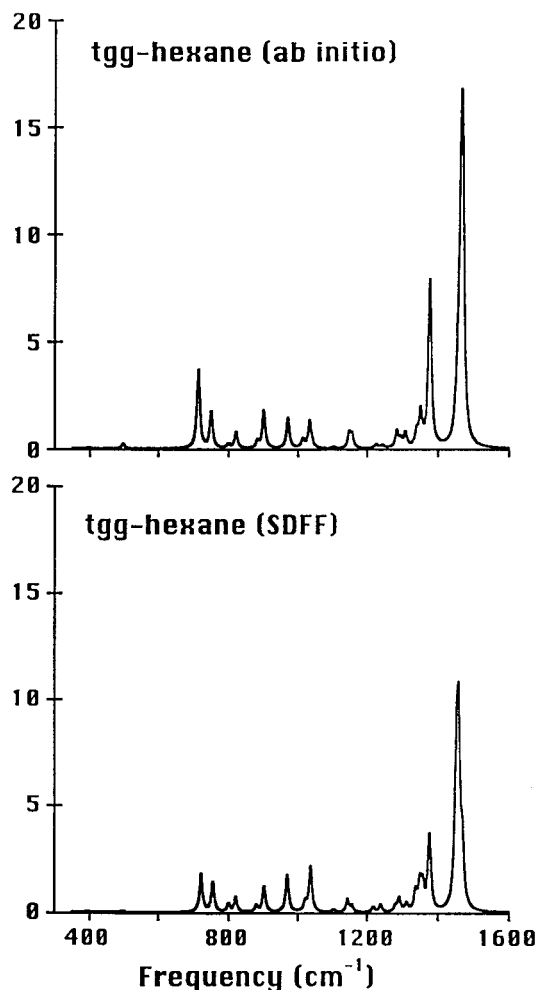
FIGURE 4. Simulated *ab initio* (top) and SDFF (bottom) infrared spectra of *tgg*-hexane. Description as in Figure 1.

ters for the calculation of conventional Raman spectra,[20] preliminary results of which (including those of the present work) have been presented.[21]

Acknowledgment

We are grateful for the helpful discussions with Weili Qian.

References

1. K. Palmo, L-O. Pietila, and S. Krimm, *J. Comput. Chem.*, **12**, 385 (1991).

2. K. Palmo, N. G. Mirkin, L-O. Pietila, and S. Krimm, *Macromolecules*, **26**, 6831 (1993).

3. K. Palmo, N. G. Mirkin, and S. Krimm, to be published.

4. W. Person and G. Zerbi, Eds., *Vibrational Intensities in Infrared and Raman Spectroscopy*, Elsevier, Amsterdam, 1982.

5. S. Krimm and J. Bandekar, *Adv. Prot. Chem.*, **38**, 181 (1986).

6. W. Qian and S. Krimm, *J. Phys. Chem.*, **100**, 14602 (1996).

7. W. Qian and S. Krimm, *J. Phys. Chem.*, **102**, 659 (1998).

8. Y. Miwa and K. Machida, *J. Am. Chem. Soc.*, **110**, 5183 (1988).

9. Y. Miwa and K. Machida, *J. Am. Chem. Soc.*, **111**, 7733 (1989).

10. E. Ganeshsrinivas, D. N. Sathyanarayana, K. Machida, and Y. Miwa, *J. Mol. Struct. (Theochem)*, **361**, 217 (1996).

11. E. Ganeshsrinivas, D. N. Sathyanarayana, K. Machida, and Y. Miwa, *J. Mol. Struct.*, **403**, 153 (1997).

12. J-H. Lii and N. L. Allinger, *J. Comput. Chem.*, **13**, 1138 (1992).

13. W. T. King, ref. 4, p. 122.

14. I. Yokoyama, Y. Miwa, and K. Machida, *J. Am. Chem. Soc.*, **113**, 6458 (1991).

15. N. G. Mirkin and S. Krimm, *J. Phys. Chem.*, **97**, 13887 (1993).

16. M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, *Gaussian-94, Revision D.4*, Gaussian Inc., Pittsburgh, PA, 1995.

17. P. Pulay and G. Fogarasi, *J. Chem. Phys.*, **96**, 2856 (1992).

18. C. Chipot, J. G. Angyan, G. G. Ferenczy, and H. A. Scheraga, *J. Phys. Chem.*, **97**, 6628 (1993).

19. D. E. Williams, *J. Comput. Chem.*, **15**, 719 (1994).

20. K. Palmo and S. Krimm, to be published.

21. K. Palmo, N. G. Mirkin, and S. Krimm, In *XVth International Congress on Raman Spectroscopy*, S. A. Asher and P. B. Stein, Eds., Wiley, New York, 1996, p. 208.