
Merck Molecular Force Field. III. Molecular Geometries and Vibrational Frequencies for MMFF94*

THOMAS A. HALGREN

*Department of Molecular Design and Diversity, Merck Research Laboratories, P.O. Box 2000,
Rahway, New Jersey 07065*

Received 20 March 1995; accepted 31 August 1995

ABSTRACT

This article describes the parameterization and performance of MMFF94 for molecular geometries and deformations. It defines the form used for the valence-coordinate terms that represent variations in bond lengths and angles, and it describes the derivation of quadratic force constants from HF/6-31G* data and the derivation of reference bond lengths and angles from fits to MP2/6-31G*-optimized geometries. Comparisons offered show that MMFF94 accurately reproduces the computational data used in its parameterization and demonstrate that its derivation from such data simultaneously confers the ability to reproduce experiment. In particular, MMFF94 reproduces experimentally determined bond lengths and angles for 30 organic molecules with root mean square (rms) deviations of 0.014 Å and 1.2°, respectively. MM3 reproduces bond angles to the same accuracy, but reproduces experimental bond lengths more accurately, in part because it was fit directly to thermally averaged experimental bond lengths; MMFF94, in contrast, was fit to (usually shorter) energy-minimum values, as is proper for an anharmonic force field intended for use in molecular-dynamics simulations. The comparisons also show that UFF and a recent version of CHARMM (QUANTA 3.3 parameterization) are less accurate for molecular geometries than either MMFF94 or MM3. For vibrational frequencies, MMFF94 and MM3 give comparable overall rms deviations versus experiment of 61 cm⁻¹ and 57 cm⁻¹, respectively, for 15 small, mostly organic molecules. In a number of instances, MM3's derivation employed observed frequencies that differ substantially—by nearly 400 cm⁻¹ in one case—from other published frequencies which had themselves been confirmed theoretically by good-quality *ab initio* calculations. Overall, the comparisons to experimental geometries and vibrational frequencies demonstrate that MMFF94 achieves MM3-like accuracy for organic systems for which MM3

* This article includes Supplementary Material available from the author upon request or via the Internet at <ftp.wiley.com/public/journals/jcc/suppmat/17/553> or <http://journals.wiley.com/jcc>

has been parameterized. Because MMFF94 is derived mainly from computational data, however, it has been possible to parameterize MMFF94 with equal rigor for a wide variety of additional systems for which little or no useful experimental data exist. Equally good performance can be expected for such systems. © 1996 by John Wiley & Sons, Inc.

Introduction

This is part III in a series introducing MMFF94.¹⁻⁴ This article describes MMFF94's parameterization and characterizes its performance for molecular geometries and vibrational frequencies, particularly in relation to the widely used and highly regarded MM3 force field.⁵

The second section gives the form for the valence-coordinate terms in the MMFF94 energy expression that arise from variations in bond lengths and angles. The third section then focuses on the parameterization of the quadratic force constants required to describe bond stretching, angle bending, stretch-bend interaction, and out-of-plane bending, and the fourth section describes the derivation of reference values for bond lengths and angles; the fifth section specifies how these and other aspects of the parameterization have been combined to generate a mutually consistent set of parameters. The sixth section characterizes MMFF94's performance for molecular geometries and vibrational frequencies. Comparisons to experiment are also given for MM3, UFF,⁶ and CHARMM⁷ for geometries and for MM3 for vibrational frequencies. The seventh section presents a subset of the MMFF94 parameters and specifies how the full set may be obtained (the related Appendix A, Supplementary Material, discusses the format and use of the parameter tables). The final discussion summarizes this work and notes possible future improvements in functional form and parameterization methodology.

Form of MMFF94 for Interactions Derived From Bond Stretching and Angle Bending

The total MMFF94 energy can be written as:

$$E_{\text{MMFF}} = \sum E_{B_{ij}} + \sum E_{A_{ijk}} + \sum E_{B_{A_{ijk;l}}} + \sum E_{T_{ijk;l}} + \sum E_{\text{vdW}_{ij}} + \sum E_{Q_{ij}} \quad (1)$$

where the summations run over appropriate pairs, triplets, or quadruplets of atomic indices. This section considers the first four terms in eq. (1). The numerical constants cited in what follows are such that all terms are computed in units of kilocalories per mole when distances and angles are measured in angstroms and degrees, respectively.

QUARTIC BOND STRETCHING

The form used in MMFF94 is:

$$E_{B_{ij}} = 143.9325 \text{ k} b_{ij} / 2 * \Delta r_{ij}^2 * (1 + cs * \Delta r_{ij} + 7/12 cs^2 * \Delta r_{ij}^2) \quad (2)$$

where $k b_{ij}$ is the bond stretching force constant ($\text{md}/\text{\AA}$), $\Delta r_{ij} = r_{ij} - r_{ij}^0$ is the difference (\AA) between the actual and reference bond lengths for bonded atoms i and j , and $cs = -2 \text{\AA}^{-1}$ is the "cubic-stretch" constant. This function corresponds to an expansion through fourth order of a Morse function with an "alpha" of 2\AA^{-1} .⁸ This quartic expansion is applicable to much larger displacements from the equilibrium bond length than is the simpler, commonly used, quadratic form. It is preferable to the cubic expansion used in MM2⁹ and MM2X,^{1,10} which unphysically diverges to negative infinity (for $cs < 0$) as the bond length r_{ij} is progressively lengthened. The value chosen for cs represents the average value of roughly 2\AA^{-1} we obtained for the Morse function "alpha" in exploratory force-field optimizations using the PROBE program developed by the Biosym Consortium.¹¹ It is identical to the value used in MM2 but significantly smaller than the value of 2.55\AA^{-1} used in MM3.⁵ A very recent study that used multiconfigurational wavefunctions in conjunction with double-zeta plus polarization basis sets found values close to 2\AA^{-1} for a variety of bond types.¹² The value for cs used in MMFF94 is therefore a representative one that should capture the essential part of the bond anharmonicity.

The force constants $k b_{ij}$ and reference bond lengths r_{ij}^0 depend on the MMFF atom types I and J to which atoms i and j correspond. To indicate this dependence, we use $i, j, k \dots$ to refer to atoms

and $I, J, K \dots$ to denote the corresponding atom types (or, in some cases, the atomic species). The current MMFF atom types are defined in part I.¹ It is important to note that the atom types do not always suffice to fully define the nature of the bond in question. Consider, for example, the case of 1,3-butadiene, where each of the four sp^2 -hybridized carbon atoms is assigned a MMFF atom type of 2. To distinguish the full double bonds made to the terminal carbons from the delocalized single bond between the two central carbons, some force fields invoke additional atom types.¹³ This approach quickly leads to a multiplication of equivalent but differently labeled parameters. MMFF94, in contrast, uses the formal bond order to assign a *bond type index* BT_{IJ} that, together with the atom types, unambiguously relates the chemical environment to the appropriate parameters for bond stretching. Consequently, each stretching parameter is labeled by *three* numerical indices—one for the bond type and two for the atom types. In effect, this approach enables special sets of reference bond lengths and force constants to be employed for “conjugated single bonds,” such as those found in butadiene and biphenyl. To provide maximum flexibility, MMFF94 currently defines a special value for the bond type index for most formally single bonds between sp - or sp^2 -hybridized nitrogen and/or carbon atoms.⁴

CUBIC ANGLE BENDING

Most angle-bending interactions are described by the cubic expansion

$$EA_{ijk} = 0.043844 \frac{ka_{IJK}}{2} * \Delta\vartheta_{ijk}^2 (1 + cb * \Delta\vartheta_{ijk}) \quad (3)$$

where ka_{IJK} is the angle-bending force constant ($\text{md } \text{\AA}/\text{rad}^2$), $\Delta\vartheta_{ijk} = \vartheta_{ijk} - \vartheta_{IJK}^0$ is the difference between actual and reference i - j - k bond angles in degrees, and $cb = -0.007 \text{ deg}^{-1}$ (or, more precisely, -0.4 rad^{-1}) is the “cubic-bend” constant. This expression allows for anharmonicity in angle bending. The cubic-bend constant is small enough to ensure that EA_{ijk} does not behave unphysically as ϑ approaches 180° . This value was obtained by considering the effect on the angle-bending force constants of changes in bond angles in HF/6-31G* calculations on water, ammonia, and ethylene; in the last two cases, a previously described formalism⁸ was used to extract the force constants. While this data set is small, all three systems behaved similarly, suggesting that this value for cb should be a reasonable one.

For linear or near-linear bond angles such as those that occur in alkynes, nitriles, isonitriles, azides, and diazo compounds, MMFF94 employs the form used in DREIDING¹⁴ and UFF:⁶

$$EA_{ijk} = 143.9325 ka_{IJK} * (1 + \cos \vartheta_{ijk}) \quad (4)$$

where ka_{IJK} and ϑ_{ijk} are as defined above. Unlike eq. (3), eq. (4) possesses well-defined first and second derivatives at $\vartheta_{ijk} = 180^\circ$.¹⁴

Angle terms i - j - k are included in all cases in which atoms i and k are bonded to a common central atom j . The procedure used to assign the force constants ka_{IJK} and the reference angles ϑ_{IJK}^0 is analogous to that used for bond stretching. In the present case, an angle-type index, AT_{IJK} , is defined to allow a smaller set of atom types to be used. Here, however, these special indices are needed not only when bond i - j , bond j - k , or both involve single bonds between sp - or sp^2 -hybridized atoms, but also when i - j - k is an angle in a three- or four-membered ring (cf. Appendix A). Accordingly, the angle-bending parameters are characterized by one angle-type and three atom-type indices. The angle-type index and the procedure used to relate angle-bending interactions to entries in the parameter file are described in Appendix A (Supplementary Material).

STRETCH-BEND INTERACTION

The form used in MMFF94 is:

$$EBA_{ijk} = 2.51210 (kba_{IJK} \Delta r_{ij} + kba_{KJI} \Delta r_{kj}) * \Delta\vartheta_{ijk} \quad (5)$$

where kba_{IJK} is the force constant (md/rad) for i - j stretching coupled to i - j - k bending, kba_{KJI} is that for k - j stretching coupled to i - j - k bending, and Δr and $\Delta\vartheta$ are as defined above. When the force constant is positive, as is usually the case, the stretch-bend term allows the bond lengths r_{ij} and r_{jk} to increase as the angle ϑ_{ijk} falls and to decrease as ϑ_{ijk} rises, emulating the effect of a Urey-Bradley 1,3 term.⁵ Equation (5) is closely related to the form used in MM2⁹ and MM3.⁵ One difference is that MMFF94 retains all terms (except those involving linear bond angles), whereas MM2 (but not MM3) omits contributions arising from the stretching of bonds to hydrogen. A second difference is that MM2 and MM3, but not MMFF94, take kba_{IJK} as being equal to kba_{KJI} when both contributions are retained. The tabulated parameters (cf. Appendix A) show that quite different values are required in some cases.

The matching of stretch-bend interactions with tabulated parameters is described in Appendix A. Currently, stretch-bend interactions are omitted when the i - j - k interaction corresponds to a linear bond angle.

OUT-OF-PLANE BENDING

MMFF94 uses the quadratic form:

$$\text{EOOP}_{ijk;l} = 0.043844 \text{koop}_{ijk;l} / 2 * \chi_{ijk;l}^2 \quad (6)$$

where $\text{koop}_{ijk;l}$ is the out-of-plane bending force constant (md Å/rad²), and $\chi_{ijk;l}$ is the Wilson angle¹⁵ (degrees) between the bond j - l and the plane i - j - k , where j is the central atom. Note that three such angles arise at each tricoordinate center j , as any of the attached atoms i , k , and l can serve as the "out-of-plane" atom. In MMFF94, all three such angles are assigned the same $\text{koop}_{ijk;l}$ force constant. Formally, an out-of-plane bending contribution is defined for all trigonal centers. However, the parameter file provides zero force constants for most cases in which the equilibrium geometry is strongly nonplanar, for example, for nitrogen in saturated aliphatic amines. The procedure used to match out-of-plane interactions to the tabulated parameters is discussed in Appendix A.

To make it feasible to use an out-of-plane bending term for nonplanar, as well as planar, trigonal centers, MMFF94 uses eq. (3) for the normal angle-bending interactions at tricoordinate centers. In contrast, when MM2 and MM3 assign an out-of-plane term, they also define the associated "in-plane" angles in a manner that requires that these angles sum to 360°. Though it is suitable for planar trigonal centers, this approach does not lend itself to describing nonplanar centers such as those that occur in enamines, in sulfonamides, and even in amides, because it does not allow reference values that average less than 120° to be used to make the center pyramidal. The MMFF94 formulation, in contrast, allows the angle-bending reference values to be used for this purpose; a residual "out-of-plane" term can then be employed to improve the fit to the inversion barrier, if this is needed.

Derivation of Quadratic Force Constants

For the core, computationally derived, portion of MMFF94, the quadratic force constants $k b_{ij}$ for

bond stretching, $k a_{ijk}$ for angle bending, $k b a_{ijk}$ and $k b a_{kjl}$ for stretch-bend interaction, and $\text{koop}_{ijk;l}$ for out-of-plane bending were determined by least-squares fitting a modified version of the MMFF energy expression to Cartesian first and second derivatives of the energy for a set of 349 molecular conformations. These fits employed the versatile program PROBE.¹¹ The requisite energy derivatives were obtained at the HF/6-31G* level for HF/6-31G*-optimized geometries using Gaussian 88,¹⁶ Gaussian 90,¹⁷ or Gaussian 92.¹⁸ The principal modifications to eq. (1) required to use PROBE consisted in: (i) replacing MMFF94's Buf-14-7 van der Waals (vdW) and buffered electrostatic terms by Lennard-Jones 10-6 and simple coulombic terms; and (ii) using a symmetrized (averaged) value for the Wilson angle at center j [eq. (6)] in place of the three slightly nonequivalent angles used in MMFF94.

Only the quadratic force constants were optimized in the PROBE fits; all other classes of parameters were held constant. To be consistent with the geometries used in calculating the *ab initio* derivatives, the PROBE runs employed reference MMFF bond lengths and bond angles that had been determined from fits to HF/6-31G*-optimized geometries. It would undoubtedly have been better to use *ab initio* derivatives evaluated at the MP2/6-31G* level for MP2/6-31G*-optimized geometries, but insufficient computational resources were available.

To account for anharmonic bond stretching and angle bending in a self-consistent fashion, the PROBE fits were carried out iteratively. Initial force constants were assigned using precursor MM2X force constants¹⁰ for bond stretching and empirical-rule values⁸ for angle bending. Application of the cubic-stretch parameter of -2 Å^{-1} and of the cubic bend parameter of -0.4 rad^{-1} then yielded estimates for the requisite cubic and quartic bond-stretching and cubic angle-bending force constants. These estimates for the higher order force constants were employed in evaluating the least-squares-fit criterion but were held fixed. The new values for the quadratic force constants were then used to update the cubic and quartic force constants for the next PROBE fit. Generally, three to five such iterations produced a converged set of quadratic force constants.

Three other aspects of the PROBE fits merit comment. (1) Because PROBE did not allow bond-order information to be used in matching force-field interactions to force-field parameters, a new atom type for sp^2 -hybridized carbon had to be

invoked for butadiene and substituted butadienes to distinguish double from formally single C—C bonds. With some difficulty, procedures were developed for successfully mapping between the PROBE and MMFF parameter representations. (2) The coupling between torsion and out-of-plane bending force constants caused some difficulties. Particularly for aromatic systems, the initial values used for the twofold torsion parameters sometimes inappropriately led to significantly negative values for the associated out-of-plane bending force constants. In such cases, we reduced the twofold torsion parameters enough to yield positive or, at worst, slightly negative out-of-plane force constants; to reproduce molecular geometries, some of the latter eventually had to be reset to slightly positive values. (3) Because HF/6-31G* frequencies are too large, by $\sim 10\%$ on average,¹⁹ the derived force constants were scaled for use in MMFF94 (see "Determination of Scale Factors" below).

Derivation of Reference Values for Bond Lengths and Angles

To derive the reference values, we assembled a set containing 342 MP2/6-31G*-optimized conformers (a slightly larger set was used in their validation) and used OPTIMOL/MMFF¹ to optimize each conformer in turn and to write a comparison of initial (MP2) and final (MMFF) bond lengths or angles to an output file. We took the initial reference bond lengths and angles to be the average of the values observed in the MP2/6-31G*-optimized structures, and took the initial quadratic force constants from MM2X for bond stretching and from an empirical rule for angle bending.^{4,8} To ensure that the *ab initio*-optimized and MMFF-optimized conformations remained similar, we applied a weak harmonic restraint (penalty function) to each torsion angle.²⁰

Next, we collated information about the comparison between the *ab initio*- and MMFF-optimized geometries for each characteristic type of bond or angle. To update the reference values for bond lengths, we subtracted the average signed difference between the MMFF-optimized and *ab initio*-optimized bond lengths from the corresponding reference value and wrote the updated reference values to a new MMFF parameter file for use as input for the next iteration. Normally, the reference bond lengths converged rapidly, typically in about three iterative cycles.

We found that some bond angles responded less directly to changes in reference values than did others, in most cases because they were subject to constraints such as that the sum of the three angles at a trigonal planar center or of the four angles in a planar four-membered ring must equal 360° . We were, however, able to accelerate the convergence by multiplying the average-deviation correction by a factor of ~ 1.5 for the first few iterative cycles. Because angles in small rings are especially noncompliant, we periodically reinitialized their reference values to the average of the *ab initio*-optimized values. Roughly six to eight iterative cycle produced "effective convergence," defined as meaning that little further improvement in the accuracy of the MMFF geometries occurred during the last few iterations.

Before writing each set of updated reference angles to the output parameter file, we made a further adjustment to enforce "embedding" in an averaged sense. A trigonal planar center is embeddable in two-space, for example, if the three bond angles at that center sum to 360° . For each trigonal-planar center type *J*, we therefore uniformly adjusted the reference values ϑ_{IJK}^0 so that the resultant reference values, when weighted by the number of occurrences for each contributing *I-J-K* angle type, averaged 120° . This adjustment minimizes the coupling between angle reference values and out-of-plane force constants; it can be shown that nonembeddable reference angles induce an out-of-plane bending term with a positive or negative force constant that is proportional to the respective "circular excess" or "circular deficit,"²¹ this quantity being the amount by which the sum of the reference values exceeds or falls below 360° . Analogously, for tetracoordinate centers we required that the average reference angle for a given center type equal the average of the *ab initio*-optimized angles.

Three additional comments need to be offered. First, we jointly determined the bond and angle reference values by interleaving iterations on the former with those on the latter. Second, we applied this iterative procedure both to determine "HF" and "MP2" reference values. The latter used the previously cited set of 342 MP2/6-31G*-optimized molecular geometries and yielded the reference values used in MMFF94. The former employed a similar set of HF/6-31G*-optimized molecular geometries and yielded the reference values used in the derivation of quadratic force constants described in the previous section. Third, by analyzing the quality of the fit, we were able to

identify cases in which a given MMFF atom type could not properly represent its assigned range in chemical structure. One instance concerned nitrogen atoms in amides (MMFF numerical atom type 10) and in enamines, amidines, guanidines, and other unsaturated nitrogen compounds (now represented by numerical atom type 40).

Derivation of Mutually Consistent MMFF94 Parameters

A central objective in developing MMFF94 has been to derive as mutually consistent a set of parameters as possible. Self-consistency within a given class of parameters—for example, within the set of quadratic force constants—follows immediately from the fact that all such core MMFF94 parameters were derived simultaneously from the full set of computational data used in the parameterizations. Achieving self-consistency between parameters belonging to different classes required a two-stage approach. First, we determined the non-bonded electrostatic and vdW parameters in fits to computational data (e.g., on molecular dipole moments and on intermolecular-interaction energies and geometries)² that are determined mainly by these parameters. Second, we then carried out roughly four full iterations over: (i) procedures for deriving quadratic force constants and reference values described in this article; and (ii) the procedure for deriving torsion parameters described in part IV.³ This approach allowed parameters within each of these classes to be determined in the context of successively refined values for those belonging to other classes.

Performance of MMFF94

MP2/6-31G* GEOMETRIES

As reported in part I,¹ MMFF94 reproduces 4205 *ab initio*-optimized bond lengths for 358 MP2/6-31G* equilibrium conformers, most of which were used in its derivation, with a rms error of 0.006 Å. Similarly, it reproduces the 7021 MP2/6-31G* bond angles with an rms deviation of 1.16°. Comparable MM2X^{1,10} errors are 0.018 Å and 1.70°, respectively. For out-of-plane bending at nitrogen, MMFF94 accurately predicts the geometries of saturated amines (rms deviation 0.91° vs. a rms Wilson angle of 57.5° for the MP2/6-31G* structures) and of unsaturated amines (rms deviation and

value 2.05° and 43.8°, respectively). For amides, however, MMFF94 gives a rms deviation of 9.38°, roughly two fifths of the rms MP2/6-31G* value of 22.6°. Even here, though, the force field correctly reproduces the degree of pyramidalization found in the MP2/6-31G* structures in an overall sense.

EXPERIMENTAL MOLECULAR GEOMETRIES

In this subsection, we compare force-field and experimental geometries for 30 organic molecules. In particular, we show that MMFF94 geometries reproduce experiment nearly as well as do published MM3 geometries—a significant accomplishment, we feel, given that MM3 is widely regarded as being the best available force field for small-molecule modeling. MMFF94's performance is also particularly encouraging because it demonstrates that the approach taken in fitting MMFF94 to theoretical MP2/6-31G* geometries, many of which have recently been employed in comparisons to experiment elsewhere,²² simultaneously confers the ability to reproduce experiment.

Experimental, MM3 and MMFF94 geometries are compared in Table I. Also included are published UFF geometries^{6,23} and newly calculated QUANTA 3.3/CHARMm geometries,²⁴ some for molecules previously characterized using the earlier QUANTA 3.2/CHARMm parameterization.⁷ For reference, we cite the five-character MMFF conformational index¹ for those systems for which MP2/6-31G*-optimized geometries were used in the core MMFF94 parameterization. In most cases, we list only the structural information published in the cited MM3, UFF, or CHARMm study. The atom labeling usually corresponds to that used in a reference cited below.

For acetamide, the condensed-phase neutron-diffraction results cited in Table I reflect a nonplanar amide group.²⁵ Wong and Wiberg have shown that high-quality quantum-mechanical calculations also yield a nonplanar amide group both in the gas phase and when solute-solvent interactions are included.²⁶ MMFF94 predicts a nonplanar amide group with H—N—C—O' torsion angles of 4.3° (Table I) and 167.9° (the MP2/6-31G* angles are 11.4° and 164.6°, respectively). In contrast, MM3²⁷ yields a planar nitrogen, as evidenced by the fact that the three bond angles made to nitrogen sum to 360°. The incompletely reported UFF structure²³ is probably also planar, as, we have found, is the QUANTA 3.3/CHARMm structure²⁴; the reported gas-phase electron diffraction structure is also planar, but was derived under this assumption. In

TABLE I.
Comparison of MMFF94, MM3, UFF or CHARMM, and Experimentally Determined Molecular Geometries
(in Angstroms and Degrees).

Acetamide (AM03a)				
Bond or angle	ND ²⁵	MMFF94	MM3 ²⁷	UFF ²³
C'—N	1.335	1.362	1.377	1.364
C'=O	1.247	1.225	1.219	1.219
C'—C	1.509	1.502	1.520	1.497
N—H (av)	1.023	1.011	1.028	1.045
N—C'=O	122.4	121.2	122.8	119.8
O=C'—C	121.1	123.2	121.9	119.4
C'—N—H1	120.5	116.9	117.9	—
C'—N—H2	120.2	122.7	120.3	—
H1—N—H2	118.9	118.3	121.8	—
H1—N—C'—O	8.2	4.3	0.0	—
<i>trans</i> -N-Methylacetamide (AM04a)				
Bond or angle	ED ²⁹	MMFF94	MM3 ²⁷	CHARMM ²⁴
C'—N	1.386	1.377	1.381	1.346
C'=O	1.225	1.226	1.219	1.223
C'—C	1.520	1.505	1.520	1.529
N—C	1.469	1.441	1.453	1.456
N—H	—	1.011	1.029	0.998
N—C'=O	121.8	123.8	123.5	124.8
N—C'—C	114.1	114.1	114.9	113.5
C'—N—C	119.7	120.8	121.3	119.7
C'—N—H	—	118.7	119.1	119.5
C—N—H	—	118.2	119.6	120.8
Pyridine (AR02a)				
Bond or Angle	MW ³²	MMFF94	MM3(Pi) ³¹	UFF ²³
C2—N	1.338	1.351	1.347	1.361
C2—C3	1.394	1.386	1.396	1.398
C3—C4	1.392	1.391	1.390	1.397
C2—H	1.086	1.086	1.103	1.082
C2—C3—C4	118.5	118.2	119.4	119.3
C3—C4—C5	118.4	119.3	117.9	119.8
C2—N—C6	116.9	116.6	117.2	121.6
C3—C2—N	123.8	123.8	123.0	120.0
Biphenyl				
Bond or Angle	ED ³⁷	MMFF94	MM3(Pi) ³⁶	UFF ²³
C1—C1'	1.507	1.472	1.488	1.485
C1—C2	1.404	1.405	1.403	1.407
C2—C3	1.395	1.398	1.396	1.399
C3—C4	1.396	1.392	1.396	1.397
C—H (av)	1.102	1.087	—	1.083
C2—C1—C6	119.4	117.9	—	118.9
C2—C1—C1'—C2'	44.4	53.9	46.2	42.7

(Continues on next page)

TABLE I.
(continued)

Furan (AR07a)				
Bond or angle	MW ³⁴	MMFF94	MM3(Pi) ³³	UFF ²³
C2—O	1.362	1.358	1.372	1.321
C2—C3	1.361	1.376	1.361	1.375
C3—C4	1.430	1.416	1.431	1.363
C—H (av)	1.076	1.081	1.095	1.082
C2—O—C5	106.6	106.8	106.4	104.6
H—C2—O	116.0	115.6	—	123.3
O—C2—C3	110.6	110.4	110.4	112.4
C2—C3—C4	106.1	106.1	106.4	105.3
Imidazole (AR09a)				
Bond or angle	ND ³⁵	MMFF94	MM3(Pi) ³¹	UFF ²³
N1—C5	1.381	1.367	1.375	1.341
N1—C2	1.358	1.366	1.366	1.345
C4—C5	1.378	1.375	1.377	1.376
C4—N3	1.389	1.371	1.382	1.348
N3—C2	1.333	1.313	1.328	1.349
N1—H	1.053	1.011	1.051	1.044
C—H (av)	1.087	1.081	1.098	1.082
C5—N1—C2	107.2	107.5	107.7	105.4
N1—C5—C4	106.0	105.0	105.0	108.2
N3—C4—C5	109.8	110.7	111.0	109.5
C2—N3—C4	105.3	105.3	104.7	104.2
Acetic acid (CA02a)				
Bond or angle	MW ³⁸	MMFF94	MM3 ³⁹	CHARMm ²⁴
C—O	1.357	1.346	1.364	1.345
C=O	1.209	1.219	1.210	1.190
O—H	0.970	0.980	0.952	0.947
C—C	1.494	1.493	1.489	1.527
O=C—O	121.8	121.0	121.8	123.4
C—C—O	112.0	112.4	112.0	111.8
C—O—H	105.9	104.1	107.4	110.6
C—C=O	126.2	126.6	126.2	124.8
Methyl formate (CE01a)				
Bond or angle	ED ⁴⁰	MMFF94	MM3 ³⁹	UFF ⁶
C'—O	1.343	1.356	1.334	1.401
C'=O	1.208	1.220	1.213	1.219
O—C	1.447	1.428	1.440	1.425
H—C'	1.103	1.101	1.106	1.088
O=C'—O	126.8	126.7	125.5	—
H—C'—O	109.3	109.0	111.3	—
C'—O—C	114.3	113.8	115.7	113.6

(Continues on next page)

TABLE I.
(continued)

<i>s-trans</i> -1,3-Butadiene (CJ01b)				
Bond or angle	ED ⁴¹	MMFF94	MM3(Pi) ³⁶	UFF ⁶
C=C	1.341	1.338	1.344	1.330
C—C	1.463	1.442	1.468	1.471
C=C—C	123.3	123.0	122.4	121.1
Acrolein, trans (CJ07b)				
Bond or angle	MW ⁴²	MMFF94	MM3(Pi) ⁴³	UFF ⁶
C1—C2	1.468	1.476	1.486	1.472
C1=O4	1.214	1.224	1.216	1.220
C1—H5	1.113	1.103	1.117	1.086
C2=C3	1.340	1.337	1.340	1.335
C3—H (av)	1.085	1.085	1.102	1.086
C2—C1=O4	124.0	121.3	124.1	—
C2—C1—H5	114.7	116.4	116.2	120.6
O4=C1—H5	121.3	122.3	119.7	—
C1—C2=C3	120.4	120.8	121.2	121.2
C1—C2—H6	117.2	116.6	117.8	—
C3=C2—H6	122.4	122.6	121.1	—
C2=C3—H7	119.7	120.2	121.6	—
C2=C3—H8	122.2	122.6	120.6	—
H7—C3—H8	118.1	117.2	117.8	—
Acetaldehyde (CO02a)				
Bond or angle	ED + MW ⁴⁵	MMFF94	MM3 ⁴⁴	CHARMm ²⁴
C'=O	1.207	1.227	1.209	1.207
C'—C	1.515	1.498	1.512	1.530
C'—H	1.114	1.103	1.119	1.087
C—H (av)	—	1.092	1.116	1.090
C—C'—O	123.8	123.4	124.2	123.7
O—C'—H	—	121.0	119.1	122.4
C—C'—H	117.5	115.6	116.7	114.0
O=C'—C—H	—	0.0	0.0	0.0
Acetone (CO04a)				
Bond or angle	MW ⁴⁵	MMFF94	MM3 ⁴⁴	UFF ⁶
C=O	1.222	1.230	1.211	1.222
C—C	1.507	1.505	1.516	1.498
C—H (av)	1.085	1.093	1.116	1.111
C—C—O	121.4	121.6	122.0	119.9
C—C—H (av)	—	109.6	110.9	—
C—C—C	117.2	116.7	116.1	120.1
O=C—C—H	—	0.0	0.0	—
Cyclobutanone (CO11a)				
Bond or angle	ED + MW ⁴⁶	MMFF94	MM3 ⁴⁴	CHARMm ²⁴
C1=O	1.202	1.212	1.204	1.203
C1—C2	1.534	1.541	1.529	1.512
C2—C3	1.567	1.548	1.574	1.550
C—H (av)	1.100	1.093	1.119	1.090

(Continues on next page)

TABLE I.
(continued)

Cyclobutanone (CO11a) (Continued)				
Bond or angle	ED + MW ⁴⁶	MMFF94	MM3 ⁴⁴	CHARMm ²⁴
C2—C1=O	—	134.2	133.2	134.4
C2—C1—C4	92.8	91.6	93.5	91.2
C1—C2—C3	88.3	86.9	88.2	87.7
C2—C3—C4	90.3	91.1	90.1	88.3
C4—C1—C2—C3	> 0	13.9	0.0	17.4

Formaldehydeimine (IM03a)				
Bond or angle	MW ⁴⁷	MMFF94	MM3 ³¹	UFF ²³
C=N	1.273	1.287	1.276	1.283
N—H	1.021	1.027	1.021	1.031
C—H (av)	1.09	1.102	1.102	1.086
C—N—H	110.5	110.1	111.4	112.6
H—C—N (cis)	123.3	121.9	122.3	121.0
H—C—N (trans)	119.7	120.6	120.6	119.7
H—C—H	116.9	117.5	117.2	119.3

Methylamine (NH01a)				
Bond or angle	MW ⁴⁹	MMFF94	MM3 ⁴⁸	CHARMm ²⁴
C—H (av)	1.099	1.094	1.110	1.090
C—N	1.471	1.452	1.463	1.457
N—H	1.010	1.019	1.016	1.012
H—C—H (av)	108.0	108.6	108.2	109.9
C—N—H	110.3	109.4	112.3	110.4
H—N—H	107.1	106.1	106.4	107.0
N—C—H (av)	110.9	110.3	110.6	109.0

Trimethylamine (NH08a)				
Bond or angle	ED ⁵⁰	MMFF94	MM3 ⁴⁸	UFF ⁶
C—N	1.458	1.462	1.460	1.471
C—N—C	110.9	110.4	111.2	110.0
N—C—H (av)	110.2	111.1	110.7	—

Methanol (OH01a)				
Bond or angle	mMW ⁵¹	MMFF94	MM3 ⁵²	CHARMm ²⁴
C—H (av)	1.094	1.093	1.111	1.094
O—H	0.945	0.972	0.948	0.954
C—O	1.425	1.416	1.431	1.419
H—C—H (av)	108.6	109.6	108.9	108.7
C—O—H	108.5	107.1	108.3	110.9

Methyl ethyl ether, anti (OR01a)				
Bond or angle	ED ⁵⁵	MMFF94	MM3 ⁵²	UFF ⁶
C1—C2	1.520	1.516	1.526	1.521
C2—O3	1.422	1.425	1.421	1.411
O3—C4	1.413	1.421	1.418	1.410
C—H (av)	1.118	1.094	1.112	1.112

(Continues on next page)

TABLE I.
(continued)

Methyl ethyl ether, anti (OR01a) (Continued)				
Bond or angle	ED ⁵⁵	MMFF94	MM3 ⁵²	UFF ⁶
C2—O3—C4	111.9	111.6	112.1	109.3
C1—C2—O3	109.4	108.5	108.7	—
H—C—H (av)	109.0	108.7	108.0	—
Oxetane (OR05a)				
Bond or angle	X-ray ⁵⁶	MMFF94	MM3 ⁵²	CHARMm ²⁴
C—O	1.460	1.446	1.451	1.405
C—C	1.534	1.524	1.548	1.539
C—O—C	90.2	91.0	91.6	87.5
C—C—C	84.8	85.1	84.4	78.3
C—C—O	92.0	91.9	91.5	96.9
C—C—C—O	7.6	2.5	7.1	4.7
C—C—O—C	7.9	2.6	7.6	5.1
Propane (RA03a)				
Bond or angle	ED ⁵⁷	MMFF94	MM3 ⁵	CHARMm ²⁴
C—C	1.534	1.519	1.534	1.532
C—C—C	112.0	111.7	112.4	112.3
Isobutane (RA05a)				
Bond or angle	MM + ED ⁵⁸	MMFF94	MM3 ⁵	UFF ²³
C—C	1.535	1.527	1.538	1.532
C (central)—H	1.122	1.097	1.118	1.112
C—C—C	110.8	110.6	110.4	110.6
Cyclobutane (RA06a)				
Bond or angle	ED ⁵⁹	MMFF94	MM3 ⁵	CHARMm ²⁴
C—C	1.554	1.543	1.557	1.552
C—H (av)	1.109	1.094	1.112	1.090
C—C—C	—	87.2	—	87.7
C—C—C—C	—	25.1	22.5	22.5
Ring pucker	35	35.8	32.3	32.1
Cyclohexane (RA08a)				
Bond or angle	ED ⁶⁰	MMFF94	MM3 ⁵	CHARMm ²⁴
C—C	1.536	1.526	1.536	1.535
C—H	1.121	1.097	1.114	1.090
C—C—C	111.4	111.6	111.3	111.0
H—C—H	107.5	107.0	106.7	106.5
C—C—C—C	54.9	54.3	55.3	55.9
Norbornane				
Bond or angle	ED ⁶³	MMFF94	MM3 ⁶²	UFF ²³
C1—C2	1.536	1.538	1.548	1.539
C1—C7	1.546	1.528	1.540	1.511

(Continues on next page)

TABLE I.
(continued)

Norbornane (Continued)				
Bond or angle	ED ⁶³	MMFF94	MM3 ⁶²	UFF ²³
C2—C3	1.573	1.536	1.557	1.538
C1—C2—C3	102.7	103.0	103.3	102.7
C2—C1—C6	109.0	108.8	107.9	108.6
C2—C1—C7	102.0	101.8	101.3	101.8
C1—C7—C4	93.4	93.5	95.0	94.4
C6—C1—C2—C3	71.6	71.3	71.0	71.6
C2—C1—C7—C4	56.3	56.2	55.5	56.1
C7—C1—C2—C3	35.8	-35.8	34.9	35.3
C1—C2—C3—C4	0.0	0.0	0.0	0.0
1-Butene, skew (Re03b)				
Bond or angle	MW + ED ⁶⁴	MMFF94	MM3 ⁶⁶	CHARMm ²⁴
C1=C2	1.340	1.339	1.339	1.344
C2—C3	1.502	1.501	1.506	1.505
C3—C4	1.535	1.521	1.534	1.531
C2—H7	—	1.088	1.104	1.075
C4—H10	1.114	1.095	1.113	1.090
C1=C2—C3	125.6	124.1	124.4	124.6
C2—C3—C4	111.7	111.4	111.9	112.0
C1=C2—H7	122.6	119.0	119.1	119.8
C3—C4—H10	111.5	111.3	111.4	111.1
C1=C2—C3—C4	119.9	118.4	114.5	119.5
<i>cis</i> -2-Butene (RE08b)				
Bond or angle	ED ⁶⁵	MMFF94	MM3 ⁶⁶	CHARMm ²⁴
C2=C3	1.346	1.342	1.342	1.347
C1—C2	1.506	1.494	1.505	1.506
C1—C2=C3	125.4	127.9	127.4	127.5
C2—C1—H10	—	113.6	113.3	112.0
C2—C1—H11	110.5	109.6	110.7	110.3
C2=C3—C4—H7	5 ± 2	0.0	1.0	7.5
C3=C2—C1—H10	—	0.0	0.6	7.5
C1—C2=C3—C4	—	0.0	0.0	5.0
Ethanethiol, gauche (SR02a)				
Bond or angle	MW ⁶⁷	MMFF94	MM3 ⁶⁹	CHARMm ²⁴
C—C	1.528	1.517	1.534	1.530
C—S	1.814	1.814	1.833	1.810
S—H	1.336	1.341	1.343	1.336
C—C—S	113.7	110.7	111.7	112.5
C—S—H	96	96.7	97.0	109.8
C—C—S—H	61.8	62.2	61.8	71.7
Dimethyl sulfide (SR03a)				
Bond or angle	MW ⁶⁸	MMFF94	MM3 ⁶⁹	CHARMm ²⁴
C—S	1.802	1.808	1.808	1.801
C—S—C	98.9	98.6	98.1	98.5

(Continues on next page)

TABLE I.
(continued)

Dimethyl sulfide (SR03a) (Continued)				
Bond or angle	MW ⁶⁸	MMFF94	MM3 ⁶⁹	CHARMm ²⁴
S—C—Ha	110.8	110.4	109.8	110.1
S—C—Hs	106.6	109.2	109.2	110.5
Ethyl methyl disulfide, gauche (SR04a)				
Bond or angle	ED ⁷⁰	MMFF94	MM3 ⁷²	CHARMm ²⁴
S—S	2.031	2.058	2.032	2.039
S—C (av)	1.817	1.817	1.822	1.810
C—C	1.540	1.519	1.534	1.530
S—S—C (av)	103.2	102.6	103.5	102.9
S—C—C	112.4	112.5	112.1	113.2
C—S—S—C	84.4	84.3	83.7	88.2
C—C—S—S	66.8	68.5	69.9	61.6
Methyl hydrogen disulfide, gauche (SR05a)				
Bond or angle	MW ⁷¹	MMFF94	MM3 ⁷²	CHARMm ²⁴
S—S	2.038	2.054	2.052	2.041
S—C	1.823	1.810	1.817	1.811
S—H	1.339	1.342	1.337	1.336
S—S—H	98.3	99.6	92.0	101.4
S—S—C	102.5	101.8	102.9	103.4
S—C—H (av)	—	110.1	109.6	110.2
C—S—S—H	88.9	87.9	88.6	93.9

considering other aspects of the structure, we note that hydrogen-bonding interactions in the condensed-phase neutron-diffraction structure appreciably shorten the C'—N amide bond (1.335 Å) and lengthen the C'—O double bond (1.247 Å) relative to the gas-phase electron-diffraction results of 1.380 and 1.220 Å, respectively.²⁸ For comparison, MP2/6-31G* optimization yields 1.373 Å for the C'—N bond length and 1.228 Å for C'—O.

For *trans*-N-methylacetamide,²⁹ MMFF94 predicts a similarly nonplanar structure, as shown by the fact that, as for acetamide, the three bond angles to nitrogen again sum to ~358°. Here, however, the MP2/6-31G* structure is almost planar. Thus, as discussed further later in this study, MMFF94 is only partially successful in predicting nonplanarity at amide nitrogen. Most other force fields, however, invariably give planar amides. For example, Meier and Coussens report that the CVFF force field of Hagler and Lifson gives a planar structure for urea.³⁰ In sharp contrast, the MP2/6-31G* structure for urea is decidedly nonplanar (we find Wilson out-of-plane angles of ~42°) and has an appreciable barrier to planarity of 2.6

kcal/mol.³⁰ For *trans*-N-methylacetamide, CHARMm²⁴ gives a shorter C'—N bond length of 1.346 Å, comparable to the value of 1.335 Å observed in the neutron-diffraction structure for acetamide. Evidently, CHARMm seeks to model the geometry expected for a hydrogen-bonded amide, not an unreasonable choice for a force field intended for use in condensed-phase simulations. Unfortunately, present-day force fields do not allow a smooth transition between the gas-phase and aqueous-phase structures to be obtained as the environment is varied from one extreme to the other. How large the difference in geometry should be, however, is unclear.

Next, we consider the aromatic molecules pyridine, biphenyl, furan, and imidazole. For pyridine, MMFF94, MM3(Pi),³¹ and UFF²³ all reasonably well reproduce the observed microwave geometry,³² though UFF is less successful for bond angles. MMFF94 and MM3(Pi)^{31,33} reproduce the alternation in the experimentally determined C—C bond lengths for furan³⁴ and in the C—N bond lengths for imidazole,³⁵ but UFF²³ does not. Like MMFF94, MM3(Pi),³⁶ and UFF²³ both yield nonplanar ge-

ometries for biphenyl, in qualitative agreement with the experimental electron diffraction structure.³⁷ For imidazole, MMFF94 predicts a length of 1.011 Å for the N—H bond, whereas experiment, MM3(Pi) and UFF all give values in the range 1.04–1.05 Å. In this case, the strong intermolecular hydrogen bonding found in the experimental neutron-diffraction structure apparently significantly lengthens the N—H bond. For comparison, the MP2/6-31G*-optimized bond length is 1.012 Å, nearly identical to the MMFF94 value. Evidently, MM3 and UFF in this case have been parameterized to reproduce a perturbed experimental bond length whose value differs considerably from the gas-phase values ordinarily used in their derivations.

For acetic acid,³⁸ both MMFF94 and MM3³⁹ predict the experimental microwave geometry well. An earlier parameterization for CHARMM⁷ gave the C—O—H angle accurately, but the QUANTA 3.3 parameters give too large an angle and overestimate the C—C distance.²⁴ For methyl formate, studied by electron diffraction,⁴⁰ MMFF94 is more accurate than MM3³⁹ for bond angles, but is less accurate for bond lengths. UFF⁶ predicts much too large a value for the length of the C'—O bond that connects the carbonyl carbon to the ester oxygen.

For the conjugated systems *trans*-1,3-butadiene⁴¹ and acrolein,⁴² MMFF94 predicts the geometries about as well as MM3(Pi).^{36,43} For the former, MMFF94 intentionally calculates too short a length for the central C—C bond to better account for bonds of similar type found in the X-ray data used to extend MMFF's parameterization.⁴ Most of the crystallographic structures we examined that had a C_{sp^2} — C_{sp^2} single bond also possessed an extended "push-pull" conjugation that further increased the pi-bond order and reduced the bond length. Consequently, we elected to replace the ideal bond length of 1.447 Å derived from the MP2/6-31G* data with a smaller, compromise value of 1.430 Å. For bond angles, the largest discrepancy for MMFF94 is the too-small value for the C2—C1=O4 bond angle in acrolein (121.3°). That the MP2/6-31G* value (124.0°) is the same as the experimental value indicates a deficiency in the force-field model. For UFF, no value is reported in this particular case, but the reported C2—C1—H5 angle to the same center⁶ is much too large. MM3⁴³ appears to fail to reproduce the qualitative trend in the C2=C3—H7/C2=C3—H8 bond angles for acrolein.

MMFF94 and MM3⁴⁴ reproduce the experimen-

tal structures for acetaldehyde and acetone⁴⁵ well, as does CHARMM for the former²⁴ except for a too-short C'—H bond length. For the latter, UFF gives too large a value for the C—C—C angle.⁶ For cyclobutanone,⁴⁶ MMFF94 predicts too small a value for the C2—C3 bond length (1.548 Å), as does MP2/6-31G* (1.551 Å), MM3⁴⁴ does well in most respects, but predicts a planar ring structure. For reference, MP2/6-31G* optimization gives 12.4° for the cited C4—C1—C2—C3 torsion angle, comparable to the MMFF94 value of 13.9°. We should note, however, that the calculated barrier to planarity is small (~0.2 kcal/mol),³ making the time-averaged structure that MM3 seeks to emulate planar. As a force field primarily intended for use in molecular-dynamics calculations, MMFF94, in contrast, seeks to reproduce the (nonplanar) minimum-energy structure on the potential-energy surface.

For formaldehydeimine ($H_2C=NH$),⁴⁷ MMFF94, MM3(Pi),³¹ and UFF²³ all perform fairly well. None, however, describes the variation between the H—C—N cis and H—C—N trans angles accurately; the MP2/6-31G* angles (125.5° and 118.4°, respectively) confirm the experimental result. MMFF94, MM3,⁴⁸ CHARMM,^{7,24} and UFF⁶ all accurately describe methylamine⁴⁹ and trimethylamine.⁵⁰

MMFF94 accounts for the experimental millimeter-wave structure for methanol⁵¹ well except for the O—H bond length, for which the experimental value of 0.945 Å, though closely reproduced by MM3,⁵² may be too short. Other cited experimental determinations⁵¹ gave bond lengths as long as 0.960 Å, and a microwave determination for ethanol gave an O—H bond length of 0.971 Å⁵³ (a second determination⁵⁴ assumed the methanol value⁵¹). For comparison, the MP2/6-31G* bond length is 0.970 Å.

For methyl ethyl ether⁵⁵ and for oxetane,⁵⁶ MMFF94 and MM3⁵² predict the experimental geometries about equally well, though the MMFF94 structure for oxetane is nearly planar. For comparison, the MP2/6-31G* structure used in MMFF94's parameterization gives dihedral angles of approximately $\pm 12^\circ$ for each of the ring torsion angles listed in Table I. UFF⁶ does well for methyl ether. For oxetane, CHARMM²⁴ gives much too short a value for the C—O bond length and reproduces the bond angles within the four-membered poorly.

For the alkanes propane,⁵⁷ isobutane,⁵⁸ cyclobutane,⁵⁹ and cyclohexane,⁶⁰ MMFF94 usually predicts C—C and C—H bond lengths that are too

short by about 0.010–0.015 Å, whereas MM3⁵ closely reproduces the experimental r_g value. The MP2/6-31G* bond lengths are nearly identical to the listed MMFF94 values, indicating that the discrepancy is either to be accounted for by deficiencies in the MP2/6-31G* model or by the intrinsic difference between an energetically minimized and a thermally averaged (usually r_g) structure. As a check, we optimized ethane at the MP2 level using a triple-zeta + polarization (TZP) basis set.³ We found C—C and C—H bond lengths of 1.531 and 1.093 Å. For comparison, the experimental electron-diffraction values⁶¹ are 1.534 and 1.112 Å, respectively, and MP2/6-31G* gives 1.524 and 1.093 Å. These comparisons suggest that the MP2/6-31G* and the derived MMFF94 C—C bond lengths probably are ~0.01 Å smaller than the experimental energy-minimum values would be found to be. In contrast, the experimental values for the C—H bond lengths evidently are lengthened, relative to the energy-minimum values, by corrections due to anharmonicity.

For norbornane, MM3⁶² reproduces part of the large variation in the electron-diffraction C—C bond lengths⁶³ but MMFF94 and UFF²³ do not. In the MP2/6-31G*-optimized structure, the C2—C3 bond (along the CH₂—CH₂ bridge) lengthens to 1.555 Å, almost exactly the MM3 value, but the C1—C2 and C1—C7 bond lengths remain at 1.536 and 1.534 Å, respectively. Clearly, the much larger experimental variation, and particularly the lengthening of the C2—C3 bond to 1.573 Å, is not in accord with the MP2/6-31G* model. All the force-field methods give the torsion angles quite accurately (but note the sign of the C7—C1—C2—C3 angle, which presumably should be negative).

For *skew*-1-butene⁶⁴ and for *cis*-2-butene,⁶⁵ MMFF94, MM3,⁶⁶ and CHARMM^{7,24} for the most part, give excellent results. For the latter, however, the methods disagree as to whether the carbon framework is planar and/or whether the terminal methyl groups are twisted. CHARMM²⁴ supports the experimental suggestion of a nonplanar structure, whereas MMFF94 and MP2/6-31G* each gives an optimized structure of C_{2v} symmetry, even when the lower-symmetry CHARMM structure is used as the initial geometry. MM3 gives a planar carbon framework but slightly skews the methyl groups.

Finally, the geometries of ethanethiol⁶⁷ and dimethyl sulfide⁶⁸ are described very well both by MMFF94 and MM3,⁶⁹ but CHARMM gives a very poor value for the C—S—H bond angle in the former.²⁴ Ethyl methyl disulfide⁷⁰ and methyl hy-

drogen disulfide⁷¹ are also described well by MMFF94 except for a systematic overestimation of the S—S bond length, an overestimation present in an equal degree in the MP2/6-31G*-optimized structures.²² CHARMM²⁴ does well, for the most part, but overestimates the S—S torsion angles. MM3⁷² describes the former very accurately, but gives an anomalously small S—S—H bond angle of 92.0° for the latter; the MP2/6-31G* value of 99.6° (identical to the MMFF94 value) confirms the experimental angle of 98.3°.

To place the foregoing comparisons on a quantitative footing, we list below the root mean square (rms) deviations from the experimental geometries found for the four methods considered here:

	MMFF94	MM3	UFF	CHARMM
Bonds (Å)	0.014	0.010	0.021	0.016
Angles (deg.)	1.2	1.2	2.5	3.1

These rms deviations are based on all 95 bond lengths and 86 bond angles for MMFF94 listed in Table I, on the 94 bond lengths and 84 bond angles for MM3, on the 51 bond lengths and 32 bond angles for UFF, and on the 44 bond lengths and 39 bond angles for CHARMM. Substitution of the previously cited MP2/6-31G* values for the C'—N and C'—O bonds in acetamide, the H—N bond in imidazole, the O—H bond in methanol, and the C2—C3 bond in norbornane (the cases in which we suggest that the experimental bonds lengths may be questionable) lowers the rms deviation for MMFF94 to 0.012 Å, substantially closer to the rms deviation of 0.010 Å found for MM3 (even with these substitutions), but leaves the other values cited unchanged.

A number of specific comparisons to MM3 are of interest. First, while MM3 reproduces the experimentally determined bond lengths better than MMFF94, we need to keep in mind that MMFF94 yields energy-minimum bond lengths rather than the thermally averaged experimental r_g values MM3 emulates. Particularly for bonds to hydrogen atoms, the latter already include increases due to anharmonicity that MMFF94 can be expected to replicate when it is used in molecular-dynamics calculations—its intended domain. (We note that the comparisons may not properly reflect the capabilities of MM3, as not all of the reference experimental structures give r_g bond lengths. Moreover,

further insight might be given by separating bonds involving hydrogen from the others. However, the comparisons as given do allow us to show that MMFF94 and MM3 bond lengths fit experiment comparably and therefore do not differ greatly.) Second, though both force fields accurately reproduce the experimental geometries for aromatic and conjugated systems, MMFF94 does so without requiring that a separate molecular-orbital calculation be carried out for the pi-electron system. Third, we note that MM3 gives a planar structure for the second experimentally observed ("cisoid") conformer for 1,3-butadiene,³⁶ whereas all recent high-level quantum-mechanical calculations predict a gauche structure having a C=C—C=C torsion angle of 30° to 41°. ^{73,74} In good agreement with the *ab initio* results, MMFF94 gives a nonplanar gauche structure having a C=C—C=C angle of 43° and a barrier to planarity of ~1 kcal/mol.^{3,74}

On the other hand, MM3's use of electronegativity-related modifications to reference bond lengths confers a significant advantage, because it allows MM3 to properly describe certain trends in experimental bond lengths that MMFF94 cannot reproduce. MM3's advantage came to our attention in preparing an earlier version of this article that included comparisons to experiment for fluoroethane⁷⁵ and chloroethane,⁷⁶ systems for which CHARMM results have also been described.⁷ We decided not to include these systems because MM3 structures for fluorides and chlorides have not yet been described. The salient comparison is that for the C—F bond in fluoroethane, for which the experimental electron-diffraction determination⁷⁵ gave a bond length of 1.399 Å. In contrast, MMFF94 gives a bond length of 1.359 Å, shorter by fully 0.04 Å. For comparison, the published QUANTA 3.2/CHARMM calculation gave a C—F bond length of 1.379 Å, a calculation using the newer QUANTA 3.3 parameters²⁴ gives 1.389 Å, and the MM1 calculation gave 1.387 Å. The short MMFF94 bond length reflects an intentional compromise. Thus, we found the crystallographic structures used in extending MMFF's parameterization⁴ usually contained CF₃ groups or other multiply fluorinated aliphatic carbons rather than isolated C_{sp³}—F bonds. The C—F bonds in CF₃ groups, in particular, tend to be quite short—on the order of 1.31 to 1.32 Å. As a result, the extended parameterization gave a C—F reference bond length of 1.333 Å, far shorter than the reference value of 1.383 Å we obtained by fitting to less highly fluorinated MP2/6-31G* structures. Given this disparity, we

chose 1.360 Å as a compromise value, leading to the too-short optimized bond length for fluoroethane. While complete geometries have not been described, MM2 has been parameterized to analytically relate reference bond lengths in chlorides and fluorides to the number of electronegative substituents.⁷⁷ As expected, MM2 (and presumably MM3) accurately represents the range of C—F (and C—Cl) bond lengths. For fluorides, the newer QUANTA 3.3 and QUANTA 4.0 parameterizations for CHARMM accomplish much the same objective, but in a different way; they define separate "CF1," "CF2," and "CF3" atom types and use distinct reference C—F bond lengths of 1.388, 1.335, and 1.320 Å, respectively. Only the latter approach could be used for MMFF as it is presently constituted.

The comparisons given above show that MMFF94 performs better than UFF for common organic systems. This is to be expected, as UFF does not attempt a detailed parameterization of chemical structures and properties but rather addresses the whole periodic table and aims for a broad, comprehensive coverage. For organic structures, however, UFF's coverage is not necessarily appreciably broader than MMFF94's.⁴ Finally, CHARMM performs somewhat more accurately than UFF for bond lengths, but is the least accurate of the four methods for bond angles. Particularly troublesome is its poor prediction of the C—S—H bond angle in ethanethiol, where the QUANTA 3.2, QUANTA 3.3, and QUANTA 4.0 parameter files all contain a clearly erroneous reference bond angle of 109.47°. If this instance is excepted, CHARMM's rms deviation for bond angles falls to 2.1°, comparable to that for UFF.

VIBRATIONAL FREQUENCIES

In this subsection, we compare vibrational frequencies for MMFF94 to frequencies determined experimentally for 15 small, mostly organic molecules. As in the previous subsection, we also compare MMFF94 to MM3. In order of appearance in Table II and with reference to the experimental determinations, the molecules compared are formamide,^{78–80} benzene,⁸¹ formic acid,⁸² formaldehyde,⁸³ acetaldehyde,^{84,85} methylamine,^{86,87} ammonia,⁸⁸ methanol,⁸⁹ water,⁹⁰ methane,⁹⁰ ethane,⁹¹ ethylene,⁹² hydrogen sulfide,⁹³ gauche-ethanethiol,⁹⁴ and dimethyl disulfide.⁹⁵ Because experimental determinations of vibrational frequencies are sometimes subject to error, we chose, wherever possible, systems for which high level *ab*

initio calculations and/or scaled quantum-mechanical force fields had been published, and for which a particular set of experimental frequencies could therefore be selected as most likely to be accurate where two or more incompatible sets had been published. As is commonly done, we compared measured anharmonic frequencies to calculated harmonic frequencies, because no better comparison can readily and reliably be made.

The overall rms deviations found for MMFF94 and for MM3 using the data in Table II are:

	MMFF94	MM3
Vibrational frequencies (cm^{-1})	61	57

These rms deviations are based on 171 tabulated frequencies for MMFF94 and 157 for MM3; when restricted to the MM3 subset, the rms deviation for MMFF94 is 60 cm^{-1} .

The comparisons indicate that MMFF94 and MM3 perform comparably in predicting vibrational frequencies. We should note, however, that the MM3 rms deviations listed in Table II, in some cases, are higher than those reported in the literature. The differences relate to the choice of the experimental frequencies. In the following subsections, we explain the basis for the choices we have made and note their effect on our assessment of MM3. We also examine the nature of the errors MMFF94 makes in predicting vibrational frequencies. Finally, we describe the derivation and use of force-constant scale factors.

Selection of Experimental Frequencies

Let us begin with formamide. With one exception, we have used the "observed" vibrational frequencies cited by Lii and Allinger.²⁷ As Table II shows, however, several of these frequencies (as well as several additional frequencies used in parameterizing MM3 for other amides) actually are frequencies calculated from a scaled quantum-mechanical (SQM) force field. These frequencies are prefaced by a "*" in the table to connote their different origin. The SQM force field in question was developed by Fogarasi and Balazs⁷⁹ some time ago on the basis of single-determinant HF/4-21G calculations that lacked polarization functions. The HF/4-21G level of theory is now known to be inadequate for systems containing heteroatoms. The predicted frequencies are particularly likely to be unreliable for the lowest-frequency A" mode,

the one that mainly describes out-of-plane bending at nitrogen. Lii and Allinger parameterized MM3 to reproduce the SQM value of 507 cm^{-1} for this mode. Fogarasi and Balazs, however, also cite a number of lower frequencies, including a frequency of 196 cm^{-1} from an *ab initio* calculation that used a polarized basis set that included *d* functions⁹⁶ and a frequency of 289 cm^{-1} from a vapor and matrix infrared study.⁸⁰ We list the latter value in Table II, prefaced by a "#" to connote its still-different origin. For reference, MP2/6-31G* at the MP2/6-31G*-optimized geometry gives a frequency of 162 cm^{-1} for this mode, and a MP2/TZP calculation³ at the more strongly puckered MP2/TZP geometry (Wilson out-of-plane angles are $\sim 19^\circ$ vs. $\sim 12^\circ$ for MP2/6-31G*) gives 302 cm^{-1} . The cited experimental value therefore appears reasonable. The recalculated MM3 error of more than 200 cm^{-1} for this mode accounts for the relatively high rms deviation of 74 cm^{-1} in Table II.

For benzene, the choice of the experimental frequency for the B_{2u} C=C stretching mode is critical. Allinger et al.³⁶ cite a frequency of 1693 cm^{-1} for this mode,⁹⁷ whereas Pulay et al.,⁹⁸ in work predating the development of MM3, had cited the much lower frequency of 1309 cm^{-1} listed in Table II. The lower value⁹⁹ has considerable support. Thus, Thakur et al.¹⁰⁰ independently report an observed frequency of 1310 cm^{-1} ; Guo and Karplus¹⁰¹ calculate 1352 and 1377 cm^{-1} at the HF/6-31G* and MP2/6-31G levels, respectively; and Ozkabak et al.¹⁰² calculate a range of frequencies compatible with this assignment using larger basis sets and/or more highly correlated wavefunctions. MM3 erroneously reproduces the higher value and thereby incurs an error of $\sim 350 \text{ cm}^{-1}$.

MM3's treatment of benzene illustrates one of the inescapable hazards of parameterizing a force field against experimental data: such data, at unexpected times, can contain large errors that then become part of the empirical model. A second example arises in the case of acetaldehyde, where Allinger et al.⁴⁴ cite an experimental value¹⁰³ of 919 cm^{-1} for the third of the five A" modes. In contrast, Wiberg et al. cite an experimental frequency of 1107 cm^{-1} for this mode,⁸⁴ in close agreement with their scaled HF/6-31G* frequency of 1111 cm^{-1} . A few other instances of large discrepancies in cited experimental frequencies occur elsewhere in the parameterization of MM3. To give one example, Lii and Allinger¹⁰⁴ cite 862 cm^{-1} for the second lowest frequency A_{2u} mode in cyclohexane, whereas Wiberg et al.¹⁰⁵ and Gough and

TABLE II. Comparison of Experimental, MMFF94, and MM3 Vibrational Frequencies (cm^{-1}).

Formamide (AM01a)				
Symm	Exp / SQM	MMFF94	MM3 ²⁷	Assignment
A'	3535 ⁷⁸	3570	3557	NH ₂ asym stretch
A'	3450 ⁷⁸	3458	3440	NH ₂ sym stretch
A'	2852 ⁷⁸	2918	2827	C—H stretch
A'	1740 ⁷⁸	1813	1784	C=O stretch
A'	1572 ⁷⁸	1581	1579	NH ₂ scissor
A'	*1388 ⁷⁹	1516	1357	C—H rock
A'	1253 ⁷⁸	1179	1197	C—N stretch
A'	*1055 ⁷⁹	1038	1062	NH ₂ rock
A''	1030 ⁷⁸	1020	1023	C—H oop wag
A''	*672 ⁷⁹	648	632	C—N torsion
A'	*567 ⁷⁹	535	570	N—C—O bend
A''	#289 ⁸⁰	216	528	NH ₂ oop wag
rms dev.		57	74	
Benzene (AR01a)				
Symm	Exp ⁸¹	MMFF94	MM3 ³⁶	Assignment
A _{1g}	3073	3106	3058	C—H sym stretch
E _{1u}	3064	3094	3051	C—H stretch
B _{1u}	3057	3065	3038	C—H stretch
E _{2g}	3056	3077	3043	C—H stretch
E _{2g}	1599	1560	1637	C—C stretch
E _{1u}	1482	1504	1482	C—C stretch
A _{2g}	1350	1438	1368	C—C—H ip bend
B _{2u}	1309	1423	1657	C—C "Kekule" stretch
E _{2g}	1178	1262	1185	C—C—H ip bend
B _{2u}	1146	1319	1249	C—C—H ip bend
E _{1u}	1037	1032	982	C—C stretch + C—C—H bend
B _{1u}	1010	893	947	C—C—C bend
A _{1g}	993	850	911	C—C sym breathe
B _{2g}	990	1180	1119	C—H oop wag
E _{2u}	967	1117	1030	C—H oop wag
E _{1g}	846	889	773	C—H oop wag
B _{2g}	707	677	581	C—C—C oop wag
A _{2u}	673	699	569	C—H oop wag
E _{2g}	606	534	596	C—C—C ip bend
E _{2u}	398	431	361	C—C—C oop wag
rms dev.		91	101	
Formic acid, cis (CA01a)				
Symm	Exp ⁸²	MMFF94	MM3 ³⁹	Assignment
A'	3570	3659	3577	O—H stretch
A'	2944	2940	2953	C—H stretch
A'	1776	1819	1782	C=O stretch
A'	1387	1440	1378	O=C—H ip bend
A'	1223	1167	1218	C—O—H bend
A'	1105	1100	1144	C—O stretch
A''	1033	949	1003	C—H oop wag

(Continues on next page)

TABLE II.
(continued)

Formic acid, cis (CA01a) (Continued)				
A''	642	684	615	C—O torsion
A'	625	566	648	O—C=O bend
rms. dev.		56	21	
Formaldehyde (CO01a)				
Symm	Exp ⁸³	MMFF94	MM3 ⁴⁴	Assignment
B2	2843	2959	2879	CH ₂ asym stretch
A1	2783	2871	2806	CH ₂ sym stretch
A1	1747	1700	1789	C=O stretch
A1	1500	1660	1424	H—C—H bend
B2	1250	1158	1258	CH ₂ rock
B1	1167	1096	1112	CH ₂ oop wag
rms dev.		102	46	
Acetaldehyde (CO02a)				
Symm	Exp ^{84,85}	MMFF94	MM3 ⁴⁴	Assignment
A'	3014	2982	2973	CH ₃ asym stretch
A''	2964	2983	2973	CH ₃ asym stretch
A'	2923	2876	2877	CH ₃ sym stretch
A'	prtbd	2908	2821	C(=O)—H stretch
A'	1746	1760	1757	C=O stretch
A''	1435	1431	1425	H (methyl) wag
A'	1430	1435	1421	H—C—H scissor
A'	1395	1534	1380	H—C=O ip bend
A'	1353	1404	1331	CH ₃ umbrella bend
A'	1114	1076	1016	CH ₃ rock + C—C stretch
A''	1107	1100	973	H—C oop + H ₂ C twist
A'	877	926	897	C—C stretch + CH ₃ rock
A''	764	922	776	H ₂ C twist + H—C oop
A'	506	555	525	C—C=O bend
A''	143	181	155	C—C torsion
rms dev.		65	49	
Methylamine (NH01a)				
Symm	Exp ⁸⁶	MMFF94	MM3 ⁴⁸	Assignment
A''	3424	3451	3422	NH ₂ asym stretch
A'	3360	3385	3361	NH ₂ sym stretch
A''	2985	2978	2970	CH ₂ asym stretch
A'	2962	2978	2969	CH ₃ sym stretch
A'	2820	2869	2868	H—CH ₂ stretch
A'	1623	1606	1617	H—N—H scissor
A''	1476	1438	1446	H—CH ₂ wag
A'	1474	1440	1486	H—C—H scissor
A'	1430	1396	1460	CH ₃ umbrella bend
A''	1337 ⁸⁷	1231	1180	NH ₂ twist
A'	1130	1086	1133	H—CH ₂ rock in Cs plane
A'	1044	1048	1047	N—C stretch

(Continues on next page)

TABLE II.
(continued)

Methylamine (NH01a) (Continued)				
A''	995 ⁸⁷	988	1039	CH ₂ twist
A'	780	938	815	NH ₂ wag
A''	264	331	264	C—N torsion
rms. dev.		58	47	
Ammonia (NH10a)				
Symm	Exp ⁸⁸	MMFF94		Assignment
E	3444	3447		N—H stretch
A1	3337	3551		NH ₃ sym stretch
E	1627	1558		NH ₃ bend
A1	950	1167		NH ₃ umbrella bend
rms dev.		114		
Methanol (OH01a)				
Symm	Exp ⁸⁹	MMFF94	MM3 ⁵²	Assignment
A'	3682	3732	3679	O—H stretch
A'	2999	2988	2976	H—CH ₂ stretch
A''	2970	2989	2970	CH ₂ asym stretch
A'	2844	2872	2873	CH ₂ sym stretch
A'	1478	1433	1484	H—C—H scissor
A''	1465	1444	1446	CH ₂ twist
A'	1454	1414	1430	CH ₃ umbrella bend
A'	1334	1294	1287	C—O—H bend (+ CH ₃ rock)
A''	1145	1098	1106	CH ₂ rock
A'	1074	1043	1052	C—O stretch
A'	1034	1069	1086	CH ₃ rock (+ COH bend)
A''	272	319	263	C—O torsion
rms dev.		34	24	
Water (OH09a)				
Symm	Exp ⁹⁰	MMFF94		Assignment
B1	3756	3783		O—H asym stretch
A1	3657	3723		O—H sym stretch
A1	1595	1582		H—O—H bend
rms dev.		42		
Methane (RA01a)				
Symm	Exp ⁹⁰	MMFF94		Assignment
T2	3019	2982		C—H stretch
A1	2917	2840		C—H sym stretch
E	1534	1471		H—C—H bend
T2	1306	1305		H—C—H bend
rms dev.		53		

(Continues on next page)

TABLE II.
(continued)

Ethane (RA02a)				
Symm	Exp ⁹¹	MMFF94	MM3 ¹⁰⁴	Assignment
Eu	2978	2972	2968	CH ₃ d str
Eg	2955	2970	2963	CH ₃ d str
A1g	2920	2873	2882	CH ₃ s str
A2u	2915	2861	2864	CH ₃ s str
Eu	1472	1439	1455	CH ₃ d def
Eg	1468	1451	1444	CH ₃ d def
A1g	1388	1417	1437	CH ₃ s def
A2u	1379	1437	1360	CH ₃ s def
Eg	1190	1102	1064	CH ₃ rock
A1g	995	977	960	C—C stretch
Eu	822	923	908	CH ₃ rock
A1g	289	327	283	C—C torsion
rms. dev.		51	52	
Ethylene (RE01a)				
Symm	Exp ⁹²	MMFF94	MM3 ⁶⁶	Assignment
B2u	3105	3104	3107	C—H stretch
B1g	3083	3090	3103	C—H stretch
Ag	3026	3018	3023	C—H sym stretch
B3u	3021	2988	2995	C—H stretch
Ag	1630	1649	1596	C=C stretch
B3u	1444	1439	1449	CH ₂ asym scissor
Ag	1342	1322	1227	CH ₂ sym scissor
B1g	1220	1140	1111	CH ₂ ip rock
Au	1023	1135	1089	C=C torsion
B1u	949	927	931	CH ₂ oop wag
B2g	940	1060	1067	CH ₂ oop wag
B2u	826	936	923	CH ₂ ip rock
rms dev.		63	69	
Hydrogen Sulfide (SR01a)				
Symm	Exp ⁹³	MMFF94		Assignment
B1	2627	2643		S—H asym stretch
A1	2615	2637		S—H sym stretch
A1	1183	1191		H—S—H bend
rms dev.		16		
Ethanethiol, gauche (SR02a)				
Symm	SQM ⁹⁴	MMFF94	MM3 ⁶⁹	Assignment
C1	2973	2978	2930	C—H stretch
C1	2952	2961	2930	C—H stretch
C1	2939	2960	2930	C—H stretch
C1	2921	2904	2930	Methylene C—H sym stretch
C1	2885	2858	2930	Methyl C—H sym stretch
C1	2613	2640	2586	S—H stretch
C1	1479	1438	1468	Methyl CH ₂ scissor

(Continues on next page)

TABLE II.
(continued)

Ethanethiol, gauche (SR02a) (Continued)				
C1	1473	1466	1457	Methyl CH ₂ def
C1	1464	1458	1493	Methylene CH ₂ scissor
C1	1408	1446	1442	Methyl umbrella bend
C1	1310	1283	1376	Methylene CH ₂ def
C1	1265	1193	1269	CH ₂ def
C1	1114	1066	1035	CH ₃ def
C1	1048	1039	1016	CH ₃ def
C1	954	1007	975	C—C stretch
C1	862	957	850	C—S—H bend
C1	719	800	801	Methylene CH ₂ rock
C1	648	656	662	C—S stretch
C1	318	340	334	S—C—C bend
C1	248	280	268	C—C torsion
C1	206	219	205	C—S torsion
rms dev.		40	36	
Dimethyl disulfide (SR06a)				
Symm	Exp ⁹⁵	MMFF94	MM3 ⁷²	Assignment
A	2990	2982	2984	CH ₃ asym stretch
B	2990	2981	2983	CH ₃ asym stretch
A	2983	2977	2982	CH ₃ asym stretch
B	2983	2876	2982	CH ₃ asym stretch
B	2915	2867	2876	CH ₃ sym stretch
A	2913	2868	2876	CH ₃ sym stretch
B	1430	1442	1430	CH ₃ asym bend
A	1426	1444	1432	CH ₃ asym bend
A	1419	1418	1415	CH ₃ asym bend
B	1415	1418	1414	CH ₃ asym bend
A	1311	1346	1356	CH ₃ sym bend
B	1303	1344	1354	CH ₃ sym bend
A	955	939	1019	CH ₃ rock
B	955	935	1014	CH ₃ rock
A	949	927	1013	CH ₃ rock
B	949	923	1011	CH ₃ rock
A	694	692	703	S—C sym stretch
B	691	687	701	S—C asym stretch
A	509	490	514	S—S stretch
B	272	332	279	S—S—C asym bend
A	240	292	241	S—S—C sym bend
A	[134]	190	189	C—C torsion
B	[134]	183	184	C—C torsion
A	117	120	116	S—S torsion
rms dev.		37	35	

Murphy¹⁰⁶ cite 1039 cm⁻¹ and calculate 1014 and 1040 cm⁻¹, respectively; MM3 gives 869 cm⁻¹, a value apparently in error by ~170 cm⁻¹.

A second impediment to the development of an experimentally determined force field arises from

the incompleteness of many sets of published frequencies. In such cases, the derived force field may well produce erroneous values for unobserved or unassigned frequencies. For the systems considered here, examples of missing frequencies

that might have affected MM3's parameterization occur for methylamine, methanol, and dimethyl disulfide, though only for methylamine does any serious error seem to have resulted. In that case, the experimental data⁸⁷ cited by Schlegel et al.,¹⁰⁷ as listed in Table II, include frequencies of 1337 and 995 cm^{-1} for the second and third lowest A'' modes, for which they calculate values of 1391 and 992 cm^{-1} , respectively. Our HF/6-31G* calculations yield frequencies of 1332 and 947 cm^{-1} when scaled by a rather standard factor of 0.9.¹⁹ Thus, the cited experimental frequencies, even though taken from solid-state measurements, appear to be well founded. Moreover, the higher frequency mode appears to correspond to a weak mode observed at 1335 cm^{-1} in the vapor-phase spectrum.¹⁰⁸ No frequencies were reported for these modes, however, in the experimental determination¹⁰⁹ cited by Schmitz and Allinger.⁴⁸ As Table II shows, MM3 reproduces the lower frequency mode well, but predicts 1180 cm^{-1} for the higher, a value in error by roughly 150 cm^{-1} . When these A'' modes are included, the rms deviation for methylamine increases to 46 cm^{-1} for MM3, twice the figure reported in the original publication. A similar problem may also occur for *n*-butane, where Lii and Allinger¹⁰⁴ calculate a MM3 frequency of 1053 cm^{-1} for the third lowest B_g mode. They cite no experimental frequency for this mode, but Murphy et al.¹¹⁰ calculate a significantly higher frequency of 1202 cm^{-1} from their scaled HF/6-31G force field.

A third, more subtle, hazard in using experimental data to parameterize vibrational frequencies stems from the difficulty in matching the character of calculated and observed vibrational modes, especially when the molecular symmetry is low. While experimental studies usually give a qualitative description of the vibrational modes, in cases of low symmetry one often must assume that the calculated modes of given symmetry come in the same order and reflect the same composition as do the experimentally determined modes. In constructing Table II, in contrast, we have established the correspondence for MMFF by comparing the composition of its vibrational modes to that of the modes calculated at the HF/6-31G* level.¹¹¹ While not perfect, the HF/6-31G* model is accurate enough to allow its computed modes to be matched against experiment relatively easily and reliably. For MM3, on the other hand, we have assumed the most favorable match consistent with the mode symmetries.

Errors in MMFF94 Frequencies

Most stretching modes are described fairly well. However, as Table II shows, MMFF94 performs unevenly in reproducing frequencies for the stretching of bonds to hydrogen. In particular, the calculated frequencies tend to be too high for O—H bonds and too low for saturated C—H bonds. Remarkably, while the C(=O)—H stretches in formic acid and formaldehyde are much too high, olefinic and aromatic C—H stretches, which also are based on hydrogen attached to sp^2 -hybridized carbon, are well described. A similar pattern emerges for C=O stretching modes, where the calculated frequency is too high in formamide and formic acid, about right in acetaldehyde, and too low for formaldehyde. It would, to be sure, be easy to adjust the associated C=O and X—H force constants to improve the fit, as MM3 (which uses several atom types for carbonyl oxygen) has done with considerable success. For the sake of generality, however, we have chosen simply to uniformly scale the force constants obtained from the HF/6-31G energy surface.

Bending modes show considerable variation in accuracy but little obvious pattern. Frequencies for out-of-plane bending, however, usually are overestimated, for example, in five of six cases for benzene. For these modes, the MMFF frequencies could not be uniformly adjusted downward because the out-of-plane deformation frequency is significantly determined, to varying extents, by contributions from torsional interactions. Such interactions contribute strongly in benzene, for example, but play no role in formaldehyde. Thus, reducing the scale factor for out-of-plane bending sufficiently to bring the benzene deformations in line would further reduce the already underestimated out-of-plane frequency in formaldehyde.

In saturated systems, the MMFF94 torsional frequencies are usually too high, as comparison of the lowest frequency modes for acetaldehyde, methylamine, methanol, ethane, and ethanethiol shows. This overestimation appears to be a consequence of parameterizing MMFF94 primarily to reproduce conformational energies rather than torsional frequencies. By way of comparison, MM3 accurately predicts the torsional frequency in ethane, but yields a torsional barrier of 2.41 kcal/mol, about 20% lower than the experimentally determined barrier of 2.88 kcal/mol.¹¹² Increasing MM3's threefold H—C—C—H torsion parameter would improve the barrier, but would also raise the calculated torsion frequency.⁵ Inclusion of sixfold tor-

sion terms would probably be required to fit both the barrier and the frequency simultaneously.

MMFF94 incurs particularly large errors in individual vibrational frequencies for benzene one of the most complex systems examined here, and for formaldehyde and ammonia, two of the simplest. These deficiencies almost certainly reflect the simplicity of the MMFF94 force field. Closer reproduction of the observed frequencies would probably require a broad selection of "cross terms" akin to that employed in the Biosym Consortium force field of Hagler and co-workers.¹¹ We would not be surprised to find that heteronuclear aromatic systems are even less well described by MMFF94, as the recent papers from Allinger's group show to be the case for recent versions of MM3.^{31,33} But whether improvements in the treatment of vibrational frequencies over the level of accuracy attained in MMFF94 and in MM3 would be of practical importance in modeling organic and bio-organic systems is less certain. Moreover, the case of ammonia presents a sobering picture of what level of complexity might be needed to achieve uniformly high accuracy. Thus, at an early stage in this work we employed PROBE¹¹ to fit a series of force-field models to the HF/6-31G* first and second derivatives of the energy for several of the molecules in Table II, and monitored how well the derived force-field model reproduced the *ab initio* vibrational frequencies. For ammonia, we found that an elaborate force-field model that included a cross term between bond stretching and *out-of-plane bending* (not simply between bond stretching and angle bending) could reproduce all the HF/6-31G* frequencies to within 2 cm⁻¹. To our surprise, however, when we omitted this single cross term we found that the (rederived) force field now gave 1092 cm⁻¹ for the frequency of the "umbrella" bending mode, a value in error by a very substantial 118 cm⁻¹ with respect to the reference *ab initio* frequency of 1216 cm⁻¹.

Determination of Scale Factors for Quadratic Force Constants

The fit to the HF/6-31G* energy derivatives was discussed previously. Here we describe the calibration of the scale factors used to relate the derived "HF/6-31G*" force constants to the values employed in MMFF94.

During the first few cycles of the grand-iterative procedure described in the previous section, we used the following scale factors: 0.88 for bond stretching; 0.84 for angle bending; 0.86 for

stretch-bend interactions; and 0.78 for out-of-plane bending. Before the final derivation of reference values for bond lengths and angles and of torsion parameters, however, we determined optimal values for these scale factors by computing the MMFF94 vibrational frequencies and comparing them to the experimental values shown in Table II. In so doing, we divided the experimental frequencies into sets describing bond stretching to hydrogen, other bond stretching, angle bending, and out-of-plane bending. We then determined both the rms deviation and the average ratio of experimental to computed frequencies within each set. Next, we multiplied each of the previously assumed scale factors by the square of the average ratio of frequencies for the corresponding set of vibrational modes. This procedure led to the following MMFF94 scale factors: 0.84 for bond stretching; 0.78 for angle bending; 0.81 for stretch-bend interaction (assigned as the average of the previous two values); and 0.72 for out-of-plane bending. These scale factors are consistent with the known propensity of HF/6-31G* calculations to overestimate observed frequencies by about 10%,¹⁹ and therefore to overestimate force constants by about 20%. We considered applying separate bond-stretching scale factors for multiple bonds, for bonds to hydrogen, and for other single bonds, but found that relatively little further improvement in the fit to the experimental frequencies would have resulted.

GEOMETRIES FOR AMIDES AND UNSATURATED AMINES

The question of how to model the geometry of amides and unsaturated amines (vinyl amines, aromatic amines, guanidines, etc.) is one of the most vexing we encountered in the development of MMFF94. Thus, crystallographically determined structures for systems of both types tend to be relatively planar at nitrogen, though moderately nonplanar amides have been described, particularly for medium-ring lactams.¹¹³ On the other hand, the experimental microwave structure for aniline is strongly nonplanar.¹¹⁴ Such differences appear to arise because crystallographic structures often yield time averages over distinct energy minima, whereas other experimental determinations, such as that cited for aniline, can yield structures that reflect local energy-minimum geometries. Because MMFF94 has been developed primarily for use in molecular-dynamics calculations, it seeks to reproduce the latter. For use in still-prevalent en-

ergy-minimization (as opposed to molecular-dynamics) studies, however, we are developing a modified force field, currently called "MMFF94s," that yields planar or nearly planar equilibrium geometries for trigonal delocalized nitrogen centers.¹¹⁵

In the remainder of this section, we shall review the evidence that *ab initio* structures involving such nitrogen centers are nonplanar in most cases and will characterize the ability of MMFF94 to reproduce them. We shall begin by noting that MP2/6-31G*-optimized geometries usually find amides to be moderately nonplanar and consistently make unsaturated amines quite strongly puckered. Thus, while high-level *ab initio* calculations find *trans*-*N*-methylacetamide to be nearly planar,¹¹⁶ as do the MP2/6-31G* calculations used in this work, recent calculations by Sponer and Hobza¹¹⁷ find that the nucleic acids bases, like aniline, are distinctly puckered at this level. For aniline, for example, our MP2/6-31G*-optimized equilibrium geometry has virtually the same dihedral angle between the plane of the NH₂ group and the plane of the aromatic ring (42.8°) as does the experimental microwave structure (42.3°). Notably, the Wilson out-of-plane angle at nitrogen is 47.7° for the MP2/6-31G* structure, not much smaller than the rms angle for saturated amines of 57.5° reported earlier.

Evidence that this nonplanarity is not an artifact comes from *ab initio* calculations we have carried out at higher levels of theory. For urea, for example, *ab initio* geometries obtained at the MP2/6-31G*, MP2/TZP, MP4SDQ/TZP, QCISD(T)/TZP, and MP2/TZ(2d,f) levels all show average Wilson out-of-plane angles of 40° to 43°,¹¹⁸ thus making urea nearly as strongly pyramidalized as is aniline. We also examined three other systems in some detail. For formamide, as previously noted, MP2/6-31G* optimization yields a slightly puckered structure having Wilson out-of-plane angles of ~12°; here, MP2/TZP optimization affords a structure having out-of-plane angles of ~19°, even more strongly puckered. For acetamide, on the other hand, the average out-of-plane angles (23° and 24° for MP2/6-31G* and MP2/TZP, respectively) change little as the basis set is improved. Finally, for guanidine, MP2/6-31G* and MP2/TZP out-of-plane angles are about 48° and 47°, respectively.

On the presumption that the lack of influence of higher polarization functions and of electron correlation beyond the MP2 level found for urea is general, we conclude that the MP2/6-31G* mini-

mum-energy structures are representative of the actual isolated (gas-phase) systems. However, it is important to recognize that the corresponding condensed-phase geometries need not be comparably strongly pyramidalized. Thus, hydrogen bonding to carbonyl oxygen should make the oxygen more electronegative, thereby increasing the resonance delocalization of the nitrogen lone pair into the C=O double bond and promoting planarity. To assess how large this effect might be, we examined HF/6-31G*-optimized structures for cyclic (doubly hydrogen bonded) urea...water and urea...urea dimers. We found that the average Wilson out-of-plane angle of 35° at nitrogen for the (somewhat less strongly puckered) HF/6-31G* geometry fell slightly, to 33° for the hydrogen-bonded NH₂ group in the urea...water dimer and to 29° for the hydrogen-bonded NH₂ groups in the cyclic urea...urea dimer. To probe the strong interaction limit, we protonated urea at the carbonyl oxygen and optimized the MP2/6-31G* geometry. Here, the Wilson out-of-plane angles at nitrogen fell to 4° for the anti NH₂ group and to 14° for the syn group.

Our tentative conclusion is that condensed-phase energy-minimum structures probably are less strongly puckered than are either the gas-phase experimental or the MP2/6-31G*-optimized structures. For the present, however, we shall measure the performance of MMFF94 against the MP2/6-31G* structural model. For unsaturated amines, the comparisons for out-of-plane angles summarized previously showed that MMFF94 accurately reproduces the degree of pyramidalization reflected in the MP2/6-31G* structures. While larger errors occur for out-of-plane angles in amides, MMFF94 nevertheless correctly reproduces the MP2/6-31G* pyramidalization in an overall sense. In particular, MMFF94 gives rms Wilson angles of 23° for 18 primary amide nitrogens and of 16° for 30 secondary amide nitrogens, satisfactorily close to the rms MP2/6-31G* angles of 25° and 14°, respectively. The tertiary amides we have examined—*N,N*-dimethylformamide, *N,N*-dimethylacetamide, *N*-methyl- γ -butyrolactam, and *N*-methyl- δ -valerolactam—are predicted to be somewhat too pyramidal, but usually not grossly so; the Wilson out-of-plane angles are about 7°, 14°, 19°, and 26°, respectively, for MMFF94 versus 0°, 13°, 13°, and 13° for MP2/6-31G*. Among other pertinent comparisons, we note that MMFF94 makes the doubly delocalized imide nitrogen in *N*-formylformamide planar, as does MP2/6-31G*, and that both MMFF94 and MP2/6-31G* give rms out-of-plane angles of 37° for ten *N*-hydroxylamide

(hydroxamic acid) nitrogens. Finally, we note that we achieved this good overall performance only by rederiving some torsion parameters used in an earlier "MMFF93" set that consistently overestimated the degree of pyramidalization at nitrogen in amides. This rederivation, and the concomitant adjustment of two key force constants for out-of-plane bending at amide nitrogen (cf. Appendix A), is described in part IV.³

Interestingly, Allinger and coworkers, using a combination of experimental and *ab initio* data, have recently chosen to parameterize MM2 and MM3 to yield substantially nonplanar structures for unsaturated amines,¹¹⁹ rather than time-averaged planar structures.

MMFF94 PARAMETERS FOR BOND AND ANGLE DEFORMATIONS

Tables III to VI present a small subset of the MMFF94 parameters for the energy terms described earlier in this article. The complete set is, however, available as supplementary material to part I¹ and is obtainable in computer-readable form.¹²⁰ The parameters for the predecessor MM2X force field are given in the supplementary material

to a previous study.¹⁰ The parameters listed in the tables arise from various combinations of the following numerical atom types: 1, 2, 3, 5, 7, 10, 28, 31, and 70. The first five atom types describe saturated carbon, olefinic carbon, carbonyl carbon, hydrogen attached to carbon, carbonyl oxygen, and amide nitrogen, respectively. The next two describe polar hydrogens in amides and water. The last, atom type 70, is used for oxygen in water. These atom types cover simple alkanes, alkenes, aldehydes, ketones, and amides, in addition to water. Within this limited set of compound types, the tabulated parameters are complete: they cover all computationally derived parameters and all parameters defined in the extended parameterization presented in part V.⁴ For consistency, the parameters for other force-filled terms given in parts II² and IV³ cover identical sets of atom and compound types.

A detailed discussion of the format and usage of the parameter files may be found in Appendix A (Supplementary Material).¹²¹

Concluding Discussion

In this article, we fully define the bond stretching, angle bending, stretch-bend, and out-of-plane

TABLE III.
Partial List of MMFF94 Bond-Stretching Parameters [Eq. (2)].

BT _{IJ} ^a	Atom types		kb _{IJ} ^b	r _{IJ} ^{0c}	Origin / comment ^d
	I	J			
0	1	1	4.258	1.508	C94
0	1	2	4.539	1.482	C94
0	1	3	4.190	1.492	C94
0	1	5	4.766	1.093	C94
0	1	10	4.664	1.436	C94
0	2	2	9.505	1.333	C94
1	2	2	4.814	1.430	#C94
1	2	3	4.565	1.468	C94
0	2	5	5.170	1.083	C94
0	2	10	6.441	1.358	E94
1	3	3	4.418	1.489	C94
0	3	5	4.650	1.101	C94
0	3	7	12.950	1.222	C94
0	3	10	5.829	1.369	C94
0	10	10	3.995	1.373	E94
0	10	28	6.663	1.015	C94
0	31	70	7.880	0.969	C94

^a Bond-type index (cf. Appendix A, Supplementary Material). ^b Force constant in millidynes / angstrom. ^c Reference bond length in angstroms. ^d Label indicating origin of parameter (cf. Appendix A, Supplementary Material).

TABLE IV.
Partial List of MMFF94 Angle-Bending Parameters [Eqs. (3) and (4)].

AT _{ijk} ^a	Atom types			ka _{ijk} ^b	ϑ_{ijk}^0 ^c	Origin / comment ^d
	I	J	K			
0	0	1	0	0.000	108.100	0:*-1-* Default angle
0	1	1	1	0.851	109.608	C94
0	1	1	2	0.736	109.445	C94
0	1	1	3	0.777	107.517	C94
0	1	1	5	0.636	110.549	C94
0	1	1	10	1.050	109.960	C94
0	2	1	2	1.113	111.453	C94
0	2	1	3	0.667	104.829	C94
0	2	1	5	0.632	110.292	C94
0	2	1	10	1.183	107.067	E94
0	3	1	3	0.973	111.721	E94
0	3	1	5	0.650	108.385	C94
0	3	1	10	0.634	102.655	C94
0	5	1	5	0.516	108.836	C94
0	5	1	10	0.740	107.646	C94
0	10	1	10	1.198	111.521	E94
0	0	2	0	0.000	118.400	0:*-2-* Default angle
1	0	2	0	0.000	117.000	1:*-2-* Default angle
2	0	2	0	0.000	120.200	2:*-2-* Default angle
3	0	2	0	0.000	61.700	3:*-2-* Default angle
0	1	2	1	0.752	118.043	C94
0	1	2	2	0.672	122.141	C94
1	1	2	2	0.684	116.929	C94
1	1	2	3	0.698	116.104	C94
0	1	2	5	0.446	120.108	C94
0	1	2	10	1.049	115.172	E94
1	2	2	2	0.747	121.550	C94
2	2	2	2	0.834	124.429	E94
1	2	2	3	0.545	111.297	C94
2	2	2	3	0.925	117.690	E94
5	2	2	3	0.205	56.646	E94
0	2	2	5	0.535	121.004	C94
1	2	2	5	0.463	118.442	C94
0	2	2	10	1.000	121.192	E94
1	2	2	10	1.065	115.771	E94
2	3	2	3	0.850	122.318	E94
1	3	2	5	0.487	117.291	C94
1	3	2	10	1.055	115.778	E94
0	5	2	5	0.365	119.523	C94
0	5	2	10	0.712	122.228	E94
0	0	3	0	0.000	116.300	0:*-3-* Default angle
1	0	3	0	0.000	115.100	1:*-3-* Default angle
4	0	3	0	0.000	89.600	4:*-3-* Default angle
7	0	3	0	0.000	90.400	7:*-3-* Default angle
8	0	3	0	0.000	88.300	8:*-3-* Default angle
0	1	3	1	1.151	118.016	C94
1	1	3	2	1.106	116.853	C94
1	1	3	3	1.214	114.612	C94
0	1	3	5	0.808	117.280	C94
0	1	3	7	0.938	124.410	C94

(Continues on next page)

TABLE IV.
(continued)

AT _{JK} ^a	Atom types			ka _{JK} ^b	ϑ_{JK}^0 ^c	Origin / comment ^d
	I	J	K			
0	1	3	10	0.984	112.735	C94
2	2	3	2	0.922	117.408	E94
2	2	3	3	0.960	114.588	E94
1	2	3	5	0.901	115.350	C94
1	2	3	7	0.936	122.623	C94
1	2	3	10	1.042	111.721	C94
2	3	3	3	0.899	117.893	E94
8	3	3	3	1.312	89.970	E94
1	3	3	5	0.943	113.762	C94
1	3	3	7	0.919	117.024	C94
1	3	3	10	1.179	109.105	E94
0	5	3	5	0.594	116.699	C94
0	5	3	7	0.670	123.439	C94
0	5	3	10	0.874	111.761	C94
0	7	3	10	0.907	127.152	C94
0	10	3	10	1.612	114.923	C94
0	0	10	0	0.000	116.800	0:-10-* Default angle
3	0	10	0	0.000	58.300	3:-10-* Default angle
4	0	10	0	0.000	91.400	4:-10-* Default angle
0	1	10	1	1.117	117.909	C94
0	1	10	2	1.022	117.859	E94
0	1	10	3	0.821	119.600	C94
0	1	10	10	1.273	109.809	E94
0	1	10	28	0.552	120.066	C94
0	2	10	2	1.129	113.908	E94
0	2	10	3	1.001	121.103	E94
0	2	10	28	0.650	118.117	E94
0	3	10	3	0.709	120.274	C94
0	3	10	10	1.216	114.236	E94
0	3	10	28	0.575	120.277	C94
0	10	10	28	0.755	113.694	E94
0	28	10	28	0.435	115.630	C94
0	31	70	31	0.658	103.978	C94

^a Angle-type index (cf. Appendix A, Supplementary Material). ^b Force constant in millidynes · angstrom / radian². ^c Reference bond angle in degrees. ^d Label indicating origin of parameter (cf. Appendix A, Supplementary Material).

bending terms in the MMFF94 energy expression and fully specify the computational data and procedures used in deriving the parameters. In addition, we described illustrative subsets of the parameters employed for bond stretching, angle bending, stretch–bend interaction and out-of-plane deformation, and define how force-field interactions are matched to entries in the parameter tables (Appendix A).¹²¹ The full parameter tables are included in the supplementary material for part I¹ and are available in computer-readable form.¹²⁰

We are confident that the final set of MMFF94 parameters describes the computational data about as well as is possible for MMFF as presently con-

stituted. However, some limitations in MMFF94's performance are likely to follow from its simple functional form. For example, we noted that MMFF94 does not allow for the appreciable variation in C—F (and other) bond lengths with multiple electronegative substitution at carbon observed experimentally. The omission of cross terms connecting torsional motion with bond stretching and angle bending also has readily discernible consequences. Such terms are needed, for example, to reproduce the significant increase in bond length of ~ 0.1 Å that occurs when a C—N rotation of ~ 90° in an amide breaks the amide-group resonance; similar, though generally smaller, increases

TABLE V.
Partial List of MMFF94 Stretch – Bend Parameters [Eq. (5)].

SBT _{ijk} ^a	Atom types			kba _{ijk} ^b	kba _{kji} ^c	Origin / comment ^d
	<i>I</i>	<i>J</i>	<i>K</i>			
0	1	1	1	0.206	0.206	C94
0	1	1	2	0.136	0.197	C94
0	1	1	3	0.211	0.092	C94
0	1	1	5	0.227	0.070	C94
0	1	1	10	0.187	0.338	C94
0	2	1	2	0.282	0.282	C94
0	2	1	3	0.206	0.022	C94
0	2	1	5	0.234	0.088	C94
0	3	1	5	0.157	0.115	C94
0	3	1	10	0.038	0.195	C94
0	5	1	5	0.115	0.115	C94
0	5	1	10	0.043	0.261	C94
0	1	2	1	0.250	0.250	C94
0	1	2	2	0.203	0.207	C94
2	1	2	2	0.222	0.269	C94
2	1	2	3	0.244	0.292	C94
0	1	2	5	0.215	0.128	C94
1	2	2	2	0.250	0.219	C94
2	2	2	3	0.155	0.112	C94
0	2	2	5	0.207	0.157	C94
1	2	2	5	0.267	0.159	C94
1	3	2	5	0.264	0.156	C94
0	5	2	5	0.140	0.140	C94
0	1	3	1	0.358	0.358	C94
2	1	3	2	0.246	0.409	C94
2	1	3	3	0.303	0.145	C94
0	1	3	5	0.321	0.183	C94
0	1	3	7	0.154	0.856	C94
0	1	3	10	0.223	0.732	C94
1	2	3	5	0.407	0.159	C94
1	2	3	7	0.214	0.794	C94
1	2	3	10	0.298	0.600	C94
1	3	3	5	0.251	0.133	C94
1	3	3	7	-0.093	0.866	C94
0	5	3	5	0.126	0.126	C94
0	5	3	7	0.032	0.805	C94
0	5	3	10	0.169	0.619	C94
0	7	3	10	0.771	0.353	C94
0	10	3	10	1.050	1.050	C94
0	1	10	1	0.063	0.063	C94
0	1	10	3	-0.021	0.340	C94
0	1	10	28	0.155	-0.051	C94
0	3	10	3	-0.219	-0.219	C94
0	3	10	28	0.137	0.066	C94
0	28	10	28	0.081	0.081	C94
0	31	70	31	0.210	0.210	C94

^a Stretch-bend type index (cf. Appendix A, Supplementary Material). ^b Force constant coupling *i-j* stretch to *i-j-k* bend, in millidynes/radian. ^c Force constant coupling *j-k* stretch to *i-j-k* bend. ^d The label "C94" indicates that all listed stretch-bend parameters are core MMFF94 values derived in fits to HF/6-31G* energy derivatives.

TABLE VI.
Partial List of MMFF94 Out-of-Plane Bending Parameters [Eq. (6)].

Atom Types <i>I, J, K, L</i>				$k_{oop_{IJK;L}}$ ^a	Origin / comment ^b
0	2	0	0	0.020	*-2-* C94 DEF
1	2	1	2	0.030	C94
1	2	2	2	0.027	C94
1	2	2	3	0.026	C94
1	2	2	5	0.013	C94
2	2	2	5	0.013	C94
2	2	3	5	0.012	C94
2	2	5	5	0.006	C94
0	3	0	0	0.130	*-3-* C94 DEF
1	3	1	7	0.146	C94
1	3	2	7	0.138	C94
1	3	3	7	0.134	C94
1	3	5	7	0.122	C94
1	3	7	10	0.129	C94
2	3	5	7	0.113	C94
2	3	7	10	0.116	C94
3	3	5	7	0.113	C94
5	3	5	7	0.103	C94
5	3	7	10	0.102	C94
7	3	10	10	0.113	C94
0	10	0	0	-0.020	*-10-* C94 DEF
1	10	1	3	-0.020	#C94 (MMFF93 = -0.034)
1	10	3	28	-0.020	#C94 (MMFF93 = -0.030)
3	10	3	28	-0.030	C94
3	10	28	28	-0.019	C94

^a Wilson-angle force constant, in millidynes · angstrom / rad². ^b Label showing origin of parameter (cf. Appendix A, Supplementary Material).

in bond length occur in other delocalized systems. An improved version of MMFF that incorporates a broader range of cross terms would be able to describe geometries and vibrational frequencies more accurately than MMFF94. The case of ammonia, cited previously, however, provides a sobering reminder of the difficulties encountered in accurately predicting frequencies in a method that attempts to fit the energy surface in a rigorously detailed way. That case showed accurate reproduction of the "umbrella" bending frequency for ammonia requires the inclusion of an unusual cross term coupling bond stretching with *out-of-plane bending*. No such term is included in any empirical force field of which we are aware.

Despite these limitations, which we regard for the most part as being minor, the comparisons to experimental molecular geometries and vibrational frequencies presented demonstrate that this initial version of MMFF performs quite well by current standards. In particular, MMFF94 reproduces experimentally determined geometries for a set of 30 organic molecules much better than does either

UFF or CHARMM (QUANTA 3.3 parameterization) and nearly as well as does MM3. Root mean square deviations for bond lengths and angles were found to be 0.014 Å and 1.2° for MMFF94, respectively.

In examining the fit to the experimental data, we noted that MMFF94 and MM3 predict qualitatively different equilibrium torsion angles for 1,3-butadiene. Here, MM3(Pi), in accord with an early experimental suggestion,⁷⁴ predicts a *cis* structure for the second equilibrium conformer, whereas MMFF94, in conformity with a variety of high-level quantum-mechanical calculations and with later interpretations of the experimental data,^{74,122-124} predicts a *gauche* structure. We should note, however, that the *ab initio* barrier to planarity is low enough (~1 kcal/mol^{3,74}) to make the time-averaged structure planar. As a force field defined primarily for use in molecular-dynamics studies, however, MMFF94 seeks to reproduce the minimum-energy structure on the potential energy surface. In the present case, we take the minimum-energy structure to be the nonplanar *gauche* con-

former; while *ab initio* calculations are subject to error, we think it unlikely that calculations of the quality we and others have performed are qualitatively incorrect for so simple a system.¹²⁵

The accuracy—and completeness—of the experimental data figured especially strongly in comparisons of the performance of MMFF94 and MM3 in reproducing observed vibrational frequencies. We found overall rms deviations for a series of 15 small molecules of 61 cm⁻¹ for MMFF94 versus 57 cm⁻¹ for MM3, showing that the two force fields perform comparably in an overall sense. For MM3, however, we noted several instances in which its parameterization had employed observed frequencies that differed significantly from other published, and theoretically confirmed, values. The most prominent case concerned the assignment of a B_{2u} mode in benzene, for which MM3 reproduces an experimental frequency that differs by nearly 400 cm⁻¹ from an alternative experimental assignment that itself had been shown to be compatible with theoretically computed frequencies. We also noted one particularly significant instance in which MM3 employed an incomplete set of vibrational frequencies rather than an alternative set that was complete and had been theoretically validated. In this instance, MM3, unknowingly, produced a frequency for an A'' mode in methylaniline that appears to be in error by roughly 150 cm⁻¹.

As examples given for both geometries and vibrational frequencies show, an experimentally derived force field can be limited by the availability of suitable experimental data as well as by the incompleteness of such data and/or the presence of occasional large errors therein. Errors in experimental determinations can easily arise from understandable difficulties encountered for molecular geometries in interpreting the raw experimental data¹²⁶ or, for vibrational frequencies, from difficulties in unambiguously assigning the complete vibrational spectrum, particularly in systems of low symmetry. Indeed, in recognition of limitations in experimental vibrational spectra, Allinger and coworkers in developing their successor MM4 force field routinely check vibrational assignments against the results of scaled HF/6-31G* (or better) calculations, and generally favor the calculated results when large differences in frequency are found.¹²⁷ On the former count, even CFF93, which ultimately parameterizes molecular geometries against experimental data to compensate for systematic errors introduced at the HF/6-31G* level,^{11b} may have some limitations in this regard.

MMFF94, in contrast, uses MP2/6-31G* geometries and makes no further adjustment; MP2/6-31G* geometries have been shown to be comparable in accuracy to experiment.^{22,128}

The uniformity and breadth of performance conferred by parameterization against data produced by a standard theoretical model of good quality is a strong mark in favor of a computationally derived force field. MMFF94 can be expected to perform equally well for all chemical structures covered in its parameterization from high-quality *ab initio* data.

Acknowledgments

I thank Robert Nachbar for making available for implementation in OPTIMOL/MMFF code that he had developed for his program VIBRATION, which assigns symmetries to normal modes of vibration and determines contributions of internal-coordinate interactions. I also thank Chris Culberson for the suggestion that amides may be less strongly pyramidalized when hydrogen bonded than when in the gas phase.

Supplementary Material

Computer-readable parameter files for the MMFF94 force field¹²⁰ and Appendix A (Detailed discussion of the format and usage of the parameter files excerpted in Tables III–VI)¹²¹ are available in Supplementary Material.

References

1. Part I: T. A. Halgren, *J. Comput. Chem.* (this issue).
2. Part II: T. A. Halgren, *J. Comput. Chem.* (this issue).
3. Part IV: T. A. Halgren and R. B. Nachbar, *J. Comput. Chem.* (this issue).
4. Part V: T. A. Halgren, *J. Comput. Chem.* (this issue).
5. N. L. Allinger, Y. H. Yuh, and J.-H. Lii, *J. Am. Chem. Soc.*, **111**, 8551–8566 (1989). See also: N. L. Allinger and L. Yan, *J. Am. Chem. Soc.*, **115**, 11918–11925 (1993), and references therein.
6. A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard III, and W. M. Skiff, *J. Am. Chem. Soc.*, **114**, 10024–10035 (1992), and references therein.
7. F. A. Momany and R. Rone, *J. Comput. Chem.*, **13**, 888–900 (1992).
8. T. A. Halgren, *J. Am. Chem. Soc.*, **112**, 4710–4723 (1990).
9. (a) N. L. Allinger, *J. Am. Chem. Soc.*, **89**, 8127 (1977); (b) U. Burkert and N. L. Allinger, *Molecular Mechanics*, American Chemical Society, Washington, DC, 1982; (c) N. L. Allinger and Y. Yuh, *QCPE*, **12**, 395 (1980).

10. M. K. Holloway, J. M. Wai, T. A. Halgren, P. M. D. Fitzgerald, J. P. Vacca, B. D. Dorsey, R. B. Levin, W. J. Thompson, L. J. Chen, S. J. deSolms, N. Gaffin, A. K. Ghosh, E. A. Giuliani, S. L. Graham, J. P. Guare, R. W. Hungate, T. A. Lyle, W. M. Sanders, T. J. Tucker, M. Wiggins, C. M. Wiscount, O. W. Woltersdorf, S. D. Young, P. L. Darke, and J. A. Zugay, *J. Med. Chem.*, **38**, 305–317 (1995).
11. (a) J. R. Maple, M.-J. Hwang, T. P. Stockfish, U. Dinur, M. Waldman, C. S. Ewig, and A. T. Hagler, *J. Comput. Chem.*, **15**, 163–182 (1994); (b) M.-J. Hwang, T. P. Stockfish, and A. T. Hagler, *J. Am. Chem. Soc.*, **116**, 2515–2525 (1994). See also: J. R. Maple, U. Dinur, and A. T. Hagler, *Proc. Natl. Acad. Sci. USA*, **85**, 5350–5354 (1988); A. T. Hagler, J. R. Maple, T. S. Thacher, G. B. Fitzgerald, and U. Dinur, In *Computer Simulation of Biomolecular Systems*, W. F. van Gunsteren and P. K. Weiner, Eds., ESCOM, Leiden, 1989, pp. 149–167; U. Dinur and A. T. Hagler, In *Reviews in Computational Chemistry*, K. B. Lipkowitz and D. B. Boyd, Eds., VCH Publishers, New York, 1991, Vol. 2, pp. 99–163. See also: J. Palca, *Nature*, **322**, 586 (1986).
12. M. Orozco and F. J. Luque, *J. Comput. Chem.*, **14**, 881–894 (1993).
13. For example, the CHARMM force field distributed through QUANTA 3.2 and other versions (Molecular Simulations, Inc., San Diego, CA; cf. ref. 7) takes this approach, as do the academic CHARMM force field, AMBER, and CVFF. MM2 and MM3, in contrast, make this distinction by carrying out explicit quantum-mechanical pi-orbital calculations. See ref. 1 for other literature citations.
14. S. L. Mayo, B. D. Olafson, and W. A. Goddard III, *J. Phys. Chem.*, **94**, 8897 (1990).
15. E. B. Wilson Jr., J. C. Decius, and P. C. Cross, *Molecular Vibrations*, Dover, New York, 1955, Chapter 4.
16. M. J. Frisch, M. Head-Gordon, H. B. Schlegel, K. Raghavachari, J. S. Binkley, C. Gonzalez, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, E. M. Fluder, S. Topiol, and J. A. Pople, Gaussian 88, Gaussian, Inc., Pittsburgh, PA, 1988, as modified at Merck for improved I/O performance by E. M. Fluder.
17. M. J. Frisch, M. Head-Gordon, G. W. Trucks, J. B. Foresman, H. B. Schlegel, K. Raghavachari, M. Robb, J. S. Binkley, C. Gonzalez, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, S. Topiol, and J. A. Pople, Gaussian 90 (Revision J), Gaussian, Inc., Pittsburgh, PA, 1990.
18. M. J. Frisch, G. W. Trucks, M. Head-Gordon, P. M. W. Gill, M. W. Wong, J. B. Foresman, B. G. Johnson, H. B. Schlegel, M. A. Robb, E. S. Replogle, R. Gomperts, J. L. Andres, K. Raghavachari, J. S. Binkley, C. Gonzalez, R. L. Martin, D. J. Fox, D. J. Defrees, J. Baker, J. J. P. Stewart, and J. A. Pople, Gaussian 92 (Revision C), Gaussian, Inc., Pittsburgh, PA, 1992.
19. J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, L. A. Curtiss, *J. Chem. Phys.*, **90**, 5622–5629 (1989); J. A. Pople, H. B. Schlegel, R. Krishnan, D. J. Defrees, J. S. Binkley, M. J. Frisch, R. A. Whiteside, R. F. Hout, and W. J. Hehre, *Int. J. Quantum Chem. Symp.*, **15**, 269 (1981).
20. In the final stages of this work we used a restraint force constant of 2 kcal/mol/rad², which yields an energy penalty of 0.0003 kcal/mol per squared degree of deviation between the actual and reference torsion angles. We had used larger values of 10 to 20 kcal/mol/rad² in work carried out before refined values for the torsion parameter had been obtained. Such values held torsion angles to within a degree or two of the reference values even in the most demanding cases.
21. B. L. Bush and T. A. Halgren, Merck & Co., Inc. (unpublished research).
22. A. St-Amant, W. D. Cornell, T. A. Halgren, and P. A. Kollman, *J. Comput. Chem.*, **16**, 1483–1506 (1995).
23. C. J. Caeswit, K. S. Colwell, and A. K. Rappé, *J. Am. Chem. Soc.*, **114**, 10035–10046 (1992).
24. Optimized in this work using CHARMM as accessed by QUANTA 3.3 from Molecular Simulations, Inc. (San Diego, CA). The Release 3.3 parameters differ from the 3.2 set used in ref. 7, but the molecular geometries usually are quite similar.
25. G. A. Jeffrey, J. R. Ruble, R. K. McMullan, D. J. DeFrees, J. S. Binkley, and J. A. Pople, *Acta Crystallogr.*, **B36**, 2292–2299 (1980).
26. M. W. Wong and K. B. Wiberg, *J. Phys. Chem.*, **96**, 668–671 (1992).
27. J.-H. Lii and N. L. Allinger, *J. Comput. Chem.*, **12**, 186–199 (1991).
28. M. Kitano and K. Kuchitsu, *Bull. Chem. Soc. Jpn.*, **46**, 3081 (1973).
29. M. Kitano, T. Fukuyama, and K. Kuchitsu, *Bull. Chem. Soc. Jpn.*, **46**, 384–387 (1973).
30. R. J. Meier and B. Coussens, *J. Mol. Struct.*, **253**, 25–33 (1992).
31. J. C. Tai, L. Yang, and N. L. Allinger, *J. Am. Chem. Soc.*, **115**, 11906–11917 (1993).
32. G. O. Sorensen, L. Mahler, and N. Rastrup-Andersen, *J. Mol. Struct.*, **20**, 119 (1974).
33. N. L. Allinger and L. Yan, *J. Am. Chem. Soc.*, **115**, 11918–11925 (1993).
34. F. Mata, C. M. Martin, and G. O. Sorensen, *J. Mol. Struct.*, **48**, 157 (1978).
35. R. K. McMullan, J. Epstein, J. R. Ruble, and B. M. Craven, *Acta Crystallogr.*, **B35**, 688–691 (1979).
36. N. L. Allinger, F. Li, L. Yan, and J. C. Tai, *J. Comput. Chem.*, **7**, 868–895 (1990).
37. A. Almennigen, O. Bastiansen, L. Fernholt, B. N. Cyvin, S. J. Cyvin, and S. Samdal, *J. Mol. Struct.*, **128**, 59–76 (1985).
38. B. P. Van Eijck, J. Van Opheusden, M. M. M. Van Schaik, and E. Van Zoeren, *J. Mol. Spectrosc.*, **86**, 465–479 (1981). Note: The O=C—O angle listed in Table I has been inferred by assuming planarity at the carbonyl carbon.
39. N. L. Allinger, Z. S. Zhu, and K. Chen, *J. Am. Chem. Soc.*, **114**, 6120–6133 (1992).
40. S. Cradock and D. W. H. Rankin, *J. Mol. Struct.*, **69**, 145 (1980).
41. K. Kuchitsu, T. Fukuyama, and Y. Morino, *J. Mol. Struct.*, **1**, 463–479 (1967–1968).
42. C. E. Blom, G. Grassi, and A. Bauder, *J. Am. Chem. Soc.*, **106**, 7427–7431 (1984).
43. N. L. Allinger, S. Rodriguez, and K. Chen, *J. Mol. Struct.*, **260**, 161–178 (1992).
44. N. L. Allinger, K. Chen, M. Rahman, and A. Pathiaseril, *J. Am. Chem. Soc.*, **113**, 4505–4517 (1991).

45. T. Iijima and M. Kimura, *Bull. Chem. Soc. Jpn.*, **42**, 2159 (1969).
46. K. Tamagawa and R. L. Hilderbrandt, *J. Phys. Chem.*, **87**, 5508 (1983).
47. R. Pearson Jr. and F. J. Lovas, *J. Chem. Phys.*, **66**, 4149 (1977).
48. L. R. Schmitz and N. L. Allinger, *J. Am. Chem. Soc.*, **112**, 8307–8315 (1990).
49. K. Takagi and T. Kojima, *J. Phys. Soc. Jpn.*, **30**, 1145–1157 (1971). Note that “HCN” angle of $110^{\circ}16'$ listed in Table XX of this reference is actually the “HNC” angle; the “NCH” angle of 110.9° listed in Table I of this work has been computed from the other structural parameters given by Takagi and Kojima.
50. B. Beagley and A. R. Medwid, *J. Mol. Struct.*, **38**, 229 (1977).
51. R. M. Lees and J. G. Baker, *J. Chem. Phys.*, **48**, 5299–5318 (1968).
52. N. L. Allinger, M. Rahman, and J.-H. Lii, *J. Am. Chem. Soc.*, **112**, 8293–8307 (1990).
53. J. P. Culot, *4th Austin Symp. Gas Phase Mol. Struct.*, paper T8 (1972).
54. Y. Sasada, M. Takano, and T. Satoh, *J. Mol. Spectrosc.*, **38**, 33–42 (1971).
55. K. Oynagi and K. Kuchitsu, *Bull. Chem. Soc. Jpn.*, **51**, 2237–2242 (1978).
56. P. Luger and J. Buschmann, *J. Am. Chem. Soc.*, **106**, 7118 (1984).
57. T. Iijima, *Bull. Chem. Soc. Jpn.*, **45**, 1291–1293 (1972).
58. R. L. Hilderbrandt and J. D. Wieser, *J. Mol. Struct.*, **15**, 27 (1973).
59. Y. Egawa, S. Yamamoto, H. Kambara, and K. Kuchitsu, personal communication from K.K. to N.L.A., as cited in ref. 5.
60. O. Bastiansen, L. Fernholt, H. M. Seip, H. Kambara, and K. Kuchitsu, *J. Mol. Struct.*, **18**, 163–168 (1973).
61. L. S. Bartell and H. K. Higginbotham, *J. Chem. Phys.*, **42**, 851 (1965).
62. N. L. Allinger, H. J. Geise, W. Pyckhout, L. A. Paquette, and J. C. Gallucci, *J. Am. Chem. Soc.*, **111**, 1106–1114 (1989).
63. L. Doms, L. Van den Enden, H. J. Geise, and C. Van Alsenoy, *J. Am. Chem. Soc.*, **105**, 158–162 (1983).
64. D. Van Hemelrijk, L. Van den Enden, H. J. Geise, H. L. Sellers, and L. Schafer, *J. Am. Chem. Soc.*, **102**, 2189–2195 (1980).
65. J. R. Durig and C. W. Hawley, *J. Chem. Phys.*, **57**, 1426 (1972); A. Almenninger, I. M. Afinsen and A. Haaland, *Acta. Chem. Scand.*, **24**, 43 (1970).
66. N. L. Allinger, F. Li, and L. Yan, *J. Comput. Chem.*, **7**, 848–867 (1990).
67. J. Nakagawa, K. Kuwada, and M. Hayashi, *Bull. Chem. Soc. Jpn.*, **49**, 3420 (1976).
68. C. O. Kadzhar, A. A. Abbasov, and L. M. Imanov, *Opt. Spectrosc. (USSR)*, **24**, 334 (1968).
69. N. L. Allinger, M. Quinn, M. Rahman, and K. Chen, *J. Phys. Org. Chem.*, **4**, 647–658 (1991). Note that only a single, average C—H stretching frequency has been reported for ethanethiol (cf. Table II). Exclusion of the C—H stretches, however, does not appreciably change the listed results.
70. A. Yokozeki and S. H. Bauer, *J. Phys. Chem.*, **80**, 618–625 (1976).
71. M. Tyblewski, T.-K. Ha, and A. Bauder, *J. Mol. Spectrosc.*, **115**, 353 (1986).
72. K. Chen and N. L. Allinger, *J. Phys. Org. Chem.*, **4**, 659–666 (1991).
73. H. Guo and M. Karplus, *J. Chem. Phys.*, **94**, 3679–3699 (1991), and references therein.
74. K. Wiberg and R. E. Rosenberg, *J. Am. Chem. Soc.*, **112**, 1509–1519 (1990).
75. B. Beagley, M. O. Jones, and M. A. Zanjanchi, *J. Mol. Struct.*, **56**, 215 (1979); B. Beagley, M. O. Jones, and P. Yavari, *J. Mol. Struct.*, **71**, 203 (1981).
76. R. H. Schwendeman and G. D. Jacobs, *J. Chem. Phys.*, **26**, 1245 (1962).
77. N. L. Allinger, M. R. Inman, M. R. Frierson, and Y. Yuh, In *Computational Concepts in Chemistry*, N. Trinajstić, Ed., E. Horkand, London, 1986, Chapter 3.
78. J. C. Evans, *J. Chem. Phys.*, **22**, 1228 (1965); *J. Chem. Phys.*, **31**, 1435 (1959).
79. G. Fogarasi and A. Balazs, *J. Mol. Struct.*, **133**, 105–123 (1985).
80. S. T. King, *J. Phys. Chem.*, **75**, 405 (1971).
81. S. Brodersen and A. K. Langseth, *Dan Vidensk. Selsk. Mat.-Fys. Sk.*, **1** (1956), as cited in P. Pulay, G. Fogarasi, and J. E. Boggs, *J. Chem. Phys.*, **74**, 3999 (1981). Note that the symmetry of the mode observed at 993 cm^{-1} should be A_{1g} , not A_{2g} as listed by Pulay and coworkers.
82. R. L. Redington, *J. Mol. Spectrosc.*, **65**, 171–189 (1977).
83. M. M. Wohar and P. W. Jagodzinski, *J. Mol. Spectrosc.*, **148**, 13–19 (1991).
84. K. B. Wiberg, V. Walters, and S. D. Colson, *J. Phys. Chem.*, **88**, 4723–4728 (1984).
85. H. Hollenstein, *Mol. Phys.*, **39**, 1013 (1980).
86. A. Y. Hirakawa, M. Tsuboi, and T. Shimanouchi, *J. Chem. Phys.*, **57**, 1236 (1972).
87. H. Woelff and H. Ludwig, *Ber. Bunsenges, Phys. Chem.*, **70**, 474 (1966).
88. J. L. Duncan and I. M. Mills, *Spectrochim. Acta*, **20**, 523 (1964).
89. A. Serrallach, R. Meyer, and H. H. Gunthard, *J. Mol. Spectrosc.*, **52**, 94–129 (1974).
90. As cited in W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986, Chapter 6.
91. J. L. Duncan, R. A. Kelly, G. D. Nivellini, and F. Tullini, *J. Mol. Spectrosc.*, **98**, 87–110 (1983).
92. J. L. Duncan, D. C. McKean, and P. D. Mallinson, *J. Mol. Spectrosc.*, **45**, 221–246 (1973). See also: J. L. Duncan, F. Hegelund, R. B. Foster, G. W. Hills, and W. J. Jones, *J. Mol. Spectrosc.*, **61**, 470–473 (1976).
93. H. C. Allen Jr. and E. K. Plyler, *J. Chem. Phys.*, **25**, 1132–1136 (1956).
94. Theoretical frequencies obtained in this work by scaling HF/6-31G* frequencies computed at the optimized HF/6-31G* geometry by a factor of 0.90 (cf. ref. 19). We used these values because the available sets of experimental frequencies (cf. ref. 69) were incomplete.

95. S. G. Frankiss, *J. Mol. Struct.*, **3**, 89–101 (1969). The results cited are liquid-phase frequencies from Table 3. The frequencies bracketed in Table II are values inferred from overtone bands.
96. Y. Sugawara, Y. Hamada, A. Hirakawa, M. Tsuboi, and K. Morokuma, *Chem. Phys.*, **50**, 105 (1980).
97. G. Varsanyi and A. Szoke, *Vibrational Spectra of Benzene Derivatives*, Academic Press, 1969, pp. 68–71.
98. P. Pulay, G. Fogarasi, and J. E. Boggs, *J. Chem. Phys.*, **74**, 3999 (1981).
99. S. Brodersen and A. K. Langseth, *Dan Vidensk. Selsk. Mat.-Fys. Sk.*, **1** (1956).
100. S. N. Thakur, L. Goodman, and A. G. Ozkabak, *J. Chem. Phys.*, **84**, 6642–6656 (1986).
101. H. Guo and M. Karplus, *J. Chem. Phys.*, **89**, 4235–4245 (1988).
102. A. G. Ozkabak, L. Goodman, and K. B. Wiberg, *J. Chem. Phys.*, **92**, 4115 (1990).
103. P. Cossee and J. H. Schachtsneider, *J. Chem. Phys.*, **44**, 97 (1966).
104. J.-H. Lii and N. L. Allinger, *J. Am. Chem. Soc.*, **111**, 8566–8575 (1989).
105. K. B. Wiberg, V. A. Walters, and W. P. Dailey, *J. Am. Chem. Soc.*, **107**, 4860–4867 (1985).
106. K. M. Gough and W. F. Murphy, *J. Chem. Phys.*, **87**, 1509–1519 (1987).
107. B. Schlegel, S. Wolfe, and F. Bernardi, *J. Chem. Phys.*, **67**, 4181–4193 (1977).
108. P. Pulay and F. Torok, *J. Mol. Struct.*, **29**, 239–246 (1975).
109. A. P. Gray and R. C. Lord, *J. Chem. Phys.*, **26**, 690 (1957).
110. W. F. Murphy, J. M. Fernandez-Sanchez, and K. Raghavachari, *J. Phys. Chem.*, **95**, 1124–1139 (1991).
111. These comparisons were made by expressing the normal coordinates in the redundant internal coordinate framework using code developed by Dr. Robert B. Nachbar of the Merck Research Laboratories. We thank Dr. Nachbar for making his code available.
112. E. Hirota, Y. Endo, S. Saito, and J. L. Duncan, *J. Mol. Spectrosc.*, **89**, 285 (1981).
113. J. D. Dunnitz and F. K. Winkler, *Acta Cryst.*, **B31**, 251–263 (1975), and references therein.
114. G. Roussy and A. Nonat, *J. Mol. Spectrosc.*, **118**, 180–188 (1986).
115. T. A. Halgren and R. B. Nachbar (work in progress).
116. N. G. Mirkin and S. Krimm, *J. Mol. Struct. (Theochem)*, **334**, 1–6 (1995).
117. J. Sponer and P. Hobza, *J. Phys. Chem.*, **98**, 3161–3164 (1994); J. Sponer and P. Hobza, *J. Mol. Struct.*, **304**, 35–40 (1994).
118. The triple-zeta plus polarization (TZP) basis sets used in these calculations are defined in part IV (ref. 3), where references to the MP4SDQ and QCISD(T) theoretical models may also be found. The TZ(2d,f) basis set was generated from the TZP set: by (i) replacing the single uncontracted Gaussian polarization function (*d* for C, N, and O; *p* for H) by a pair of uncontracted functions having Gaussian exponents respectively $3^{1/2}$ larger and smaller; and (ii) adding a next-higher angular momentum function (*f* for C, N, and O; *d* for H) having an exponent equal to 1.3 times the single *d* or *p* exponent in the TZP basis set.
119. N. L. Allinger, L. Yan, and K. Chen, *J. Comput. Chem.*, **15**, 1321–1330 (1994).
120. The MMFF94 parameters are included in Appendix B in the Supplementary Material for part I (ref. 1) and are available in computer-readable ASCII format.
121. Appendix A is available in Supplementary Material (see footnote * on the first page of this article).
122. C. W. Bock, P. George, M. Trachtman, and M. Zanger, *J. Chem. Soc. Perkin Trans. II*, 26–34 (1979).
123. Y. Furukawa, H. Takenchi, I. Harada, and M. Tasumi, *Bull. Chem. Soc. Jpn.*, **56**, 392 (1983).
124. C. W. Bock and Y. N. Panchenko, *J. Mol. Struct.*, **187**, 69–82 (1989).
125. One *ab initio* calculation does predict a planar cis structure: cf. D. Feller and E. R. Davidson, *Theor. Chim. Acta*, **68**, 57–67 (1985). However, this calculation, which found cis and gauche structures to be virtually identical in energy, omitted polarization functions and included too few orbitals in the electron-correlation calculation to obtain more than a small fraction of the correlation energy.
126. For a recent example of a significant experimental error that was corrected on the basis of information from *ab initio* calculations, including calculations at the MP2/6-31G* level used in the present work, see: B. J. Smith and L. Radom, *J. Am. Chem. Soc.*, **112**, 7525–7528 (1990).
127. N. Nevins and N. L. Allinger, *J. Comput. Chem.* (this issue).
128. For early studies examining the accuracy of MP2 calculations, see: W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986, Chapter 6; D. J. DeFrees, B. A. Levi, S. K. Pollack, W. J. Hehre, J. S. Binkley, and J. A. Pople, *J. Am. Chem. Soc.*, **101**, 4085–4089 (1979).