
Merck Molecular Force Field. IV. Conformational Energies and Geometries for MMFF94*

THOMAS A. HALGREN[†] and ROBERT B. NACHBAR

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 conformational energies, rotational barriers, and equilibrium torsion angles. It describes the derivation of the torsion parameters from high-quality computational data and characterizes MMFF94's ability to reproduce both computational and experimental data, the latter particularly in relation to MM3. The computational data included: (i) ~ 250 comparisons of conformational energy based on "MP4SDQ/TZP" calculations (triple-zeta plus polarization calculations at a defined approximation to the highly correlated MP4SDQ level) at MP2/6-31G* geometries; and (ii) ~ 1200 MP2/TZP comparisons of "torsion profile" structures at geometries derived from MP2/6-31G* geometries. The torsion parameters were derived in restrained least-squares fits that used the complete set of available computational data, thereby ensuring that a fully optimal set of parameters would be obtained. The final parameters reproduce the "MP4SDQ/TZP" and MP2/TZP computational data with root mean square (rms) deviations of 0.31 and 0.50 kcal/mol, respectively. In addition, MMFF94 reproduces a set of 37 experimental gas-phase and solution conformational energies, enthalpies, and free energies with a rms deviation of 0.38 kcal/mol; for comparison, the "MP4SDQ/TZP" calculations and MM3 each gives a rms deviation of 0.37 kcal/mol. Furthermore, MMFF94 reproduces 28 experimentally determined rotational barriers with a rms deviation of 0.39 kcal/mol. Given the diverse nature of the experimental conformational energies and rotational barriers and the clear indications of experimental error in some cases, the MMFF94 results appear excellent. Nevertheless, MMFF94 encounters somewhat greater difficulty in handling multifunctional compounds that place highly polar functional groups in close proximity, probably because it, like other commonly

*This article includes Supplementary Material available from the authors upon request or via the Internet at ftp.wiley.com/public/journals/jcc/suppmat/17/587 or <http://journals.wiley.com/jcc>

[†]Author to whom all correspondence should be addressed.

used force fields, too greatly simplifies the description of electrostatic interactions. Some suggestions for enhancements to MMFF94's functional form are discussed. © 1996 by John Wiley & Sons, Inc.

Introduction

This is the fourth article in a series introducing MMFF94.¹⁻⁴ This study describes MMFF94's parameterization and performance for conformational energies and geometries. It specifies the computational data employed, defines the procedure used to derive the torsion parameters, and documents the success with which MMFF94 (i) reproduces theoretical and experimental values for conformational energies and barriers, and (ii) reproduces *ab initio* conformational geometries. As in previous articles, some consideration will also be given to the performance of the predecessor MM2X force field.^{1,5}

The second section of this article gives the form of the torsion potential used in MMFF94. The third section presents the computational data used in deriving the torsion parameters and specifies the procedure used to define the high-level "MP4SDQ/TZP" model used for conformational energies. The fourth section then summarizes the procedures used to derive mutually consistent values for the parameters; Appendix A (Supplementary Material) gives a more detailed account. The fifth section characterizes the ability of MMFF94 to reproduce *ab initio* and experimental conformational energies and barriers. This section also compares MMFF94 to MM3 and MM2X for experimental conformational energies and to MM3 for rotational barriers, and shows how well MMFF94 and MM2X reproduce reference MP2/6-31G* torsion angles for equilibrium conformers. The sixth section discusses two specific elements in the formulation of MMFF94—the differential scaling of 1,4-electrostatic interactions, and the assignment of zero partial atomic charge for hydrogens attached to aliphatic (*sp*³-hybridized) carbon. The seventh section examines problematic cases in more detail and suggests how MMFF94's functional form might be enhanced. The eighth section presents a subset of the MMFF94 torsion parameters and specifies how the full set may be obtained. This section also defines the manner in which torsion interactions are matched with tabulated torsion parameters. The final section summarizes this work and provides some conclusions.

Form of the Merck Molecular Force Field for Torsion Interactions

MMFF94 employs the threefold representation also used in MM2,⁶ MM2X, and MM3⁷:

$$ET_{ijkl} = 0.5(V_1(1 + \cos \Phi) + V_2(1 - \cos 2\Phi) + V_3(1 + \cos 3\Phi)) \quad (1)$$

where Φ is the *i-j-k-l* torsion angle in degrees. The constants V_1 , V_2 , and V_3 , given in kilocalories per mole, depend on the atom types *I*, *J*, *K*, and *L*, where *i-j*, *j-k*, and *k-l* are bonded pairs, as well as on the "torsion-type" index TT_{IJKL} defined later. The latter index allows torsion interactions within four-membered rings and within five-membered rings containing at least one saturated carbon to be given special torsion constants; it also allows special values to be employed for interactions in which either the central or a peripheral bond is a single bond between pairs of *sp*- or *sp*²-hybridized atoms that are capable of participating in multiple or aromatic bonds. Central bonds of this type occur, for example, in biphenyl, butadiene, and styrene.

Computational Models and Data Used in Parameterizing MMFF94

COMPUTATIONAL DATA FOR CONFORMATIONAL ENERGIES

Using Gaussian 88,⁸ Gaussian 90,⁹ or Gaussian 92,¹⁰ we carried out calculations at what we shall call the "MP4SDQ/TZP" level for ~ 380 MP2/6-31G*-optimized geometries. Most of these calculations were for equilibrium conformers, but about 110 employed torsional constraints to enforce planarity at one or more atomic centers or to probe the torsion potential at intermediate points between conformational minima. These calculations, for example, allowed comparisons to be made of planar versus puckered conformations for four- and five-membered ring systems and of planar versus pyramidal conformations for amides, vinyl amines, and related systems. Taken together, these

calculations afforded 249 comparisons of relative "MP4SDQ/TZP" energies.

The "MP4SDQ/TZP" calculations are composite calculations that add third-order and fourth-order MP3 and MP4SDQ perturbative corrections¹¹ evaluated using a 6-31G*-type basis set¹² to the MP2/TZP energy. The designation "TZP" indicates use of a polarized triple-zeta basis set. We used the triple-zeta basis sets of Dunning¹³ for hydrogen [5s/3s] and first-row atoms [10s, 6p/5s, 3p], and of McLean and Chandler¹⁴ for second-row atoms [12s, 9p/6s, 5p]. To these basis sets we added a single set of *p*-type Gaussian polarization functions with an exponent of 0.80 for hydrogen, and single *d* set with exponents of 0.60, 0.85, 1.20, and 1.65 for C, N, O, and F, and of 0.50, 0.65, and 0.80 for P, S, and Cl. These polarization exponents are regularized values constructed from polarization exponents used in Dunning's correlation-consistent basis sets,¹⁵ Roos and Siegbahn's basis sets,¹⁶ and Pople's 6-311G** basis sets.¹⁷ For consistency, for nonhydrogen atoms, we used the same polarization exponents with the 6-31G basis set in making the MP3 and MP4SDQ perturbative correlation corrections employed in the "MP4SDQ/TZP" model. We denote this modified basis set as 6-31G# in this article.

COMPUTATIONAL DATA FOR TORSION PROFILES

The "MP4SDQ/TZP" data also afforded a number of comparisons of conformational energies at discrete points along torsion profiles. To extend these comparisons in a computationally less burdensome fashion, we carried out single-point MP2/TZP calculations at ~ 1450 torsionally incremented geometries. These geometries were derived from base MP2/6-31G* equilibrium geometries by rotating one torsion bond by a specified amount (e.g., $\pm 30^\circ$, $\pm 60^\circ$, ...) and, to minimize steric clashes, by then optimizing the resultant structure subject to strong restraints applied to all torsion angles.¹⁸ To make the comparisons of relative energies meaningful, we also optimized the base MP2/6-31G* geometries in the same manner. These optimizations employed OPTIMOL,¹ the host molecular-mechanics platform for MMFF and MM2X, and used refined but not yet fully optimized values of the MMFF parameters. These precursor MMFF parameters were themselves obtained using an earlier set of data on conformational energies. That data, some aspects of which have been employed in another context,¹⁹ utilized

MP3/6-31 + G** calculations¹² at HF/6-31G* geometries. The MP2/TZP single-point calculations provided data both on torsional energy profiles and rotational barriers and yielded 1192 comparisons of relative energies. Single-point *ab initio* energies at force-field optimized geometries have previously been used effectively in force-field development by Allinger and coworkers.²⁰

COMPARISON OF COMPUTATIONAL MODELS FOR CONFORMATIONAL ENERGIES

Benchmark *ab initio* calculations for 34 of the conformational-energy comparisons, most for systems also used in the comparisons against experiment discussed later in this study, showed that a rms error of just 0.04 kcal/mol is made, relative to unapproximated MP4SDQ/TZP calculations, by evaluating the third- and fourth-order perturbative corrections with the 6-31G# basis set. In contrast, the rms difference between the MP2/TZP and MP4SDQ/TZP relative energies was 0.24 kcal/mol (for comparison, that between HF/TZP and MP4SDQ/TZP was 0.91 kcal/mol)—large enough, we felt, to justify the use of the fourth-order corrections for electron correlation. We also considered using full or approximated MP4SDTQ/TZP calculations that add fourth-order contributions from triple excitations.¹¹ For eight systems, each of which had a relatively large triples contribution, we compared conformational energies computed at the MP4SDQ/TZP and MP4SDTQ/TZP levels to full QCISD(T)/TZP calculations, again evaluated at MP2/6-31G*-optimized geometries. We found that adding the MP4 triples actually slightly worsened the agreement with the QCISD(T)²¹ results for these demanding systems. Though we later examined a ninth system for which the MP4 triples contribution significantly improved the agreement, we elected to retain the simpler, and roughly threefold less costly, MP4SDQ model. Finally, we also computed conformational energies for most of the 34 test systems at various levels of electron correlation for the 6-31G*, 6-31 + G**, DZP,²² and 6-311G** basis sets. Except for the (somewhat more costly) calculations using the similar 6-311G** basis set, however, variations relative to the TZP model were large enough to lead us to adopt the latter. The composite "MP4SDQ/TZP" model appeared to represent the best computationally tractable model, given the available computational resources.²³ As a further validation of this model, we later show that "MP4SDQ/TZP" calculations predict experimen-

tally determined conformational energies better than do simpler MP2/TZP, MP2/6-31G*, or HF/6-31G* calculations—in most cases, significantly better.

Summary of the Derivation of MMFF94 Torsion Parameters

A detailed discussion of the procedures used to derive the MMFF94 torsion parameters is given in Appendix A (Supplementary Material), which can be accessed online.²⁴ The present section provides a briefer account of those procedures.

BASIC APPROACH

We used the program TORFIT (developed in these laboratories) to derive the torsion parameters via least-squares fits to the computational data described previously. To reproduce both energies and geometries, we used an iterative approach in which we first fit the relative “MP4SDQ/TZP” conformational energies at geometries very close to the reference MP2/6-31G*-optimized geometries. Then, as the torsion parameters became better defined, we fit them at force-field-optimized geometries that increasingly resembled the final MMFF-optimized geometries. We implemented this procedure by applying gradually diminishing quadratic torsion-angle restraints during successive MMFF optimizations of the structures involved in the “MP4SDQ/TZP” comparisons. The comparisons based on the torsionally incremented structures used the geometries described previously in both the MP2/TZP and the successive MMFF energy evaluations.

INITIAL VALUES FOR TORSION PARAMETERS

To construct an unbiased initial set of torsion parameters, we set to zero all except certain twofold parameters that were expected to be large in value (cf. Appendix A). Twofold parameters for rotation about C—N bonds in amides, for example, were given initial values of 6 kcal/mol in the latter stages of this work.

LEAST-SQUARES PENALTY-FUNCTION RESTRAINTS

All the conformational-energy data were used simultaneously in each of the least-squares fits.

These fits minimized the sum of squares shown in eq. (2):

$$\text{SOS} = \sum w_i (\text{CE}_i^{\text{MMFF}} - \text{CE}_i^{\text{ref}})^2 + p \sum (V_n(IJKL) - V_n^0(IJKL))^2 \quad (2)$$

The first summation runs over the set of 249 “MP4SDQ/TZP” and 1192 MP2/TZP conformational energy differences, where “ref” denotes either a “MP4SDQ/TZP” or a MP2/TZP energy difference. The quantity w_i is a weight factor, normally chosen as 1.0 for the “MP4SDQ/TZP” comparisons and 0.2 for the more numerous MP2/TZP comparisons. The second summation adds penalty-function restraints that serve to keep each $V_n(IJKL)$ torsion parameter²⁵ close to its input value. In their absence, correlations between parameters typically led to large changes that produced parameters of questionable physical validity while reducing the sum of squares only slightly. The penalty-function restraints tended to hold parameters to zero, where possible, or to equal magnitudes where simple parameter correlations are involved.

TORFIT

After reading in the input torsion parameters and noting which are to be optimized and which are to be held constant, TORFIT sequentially processes a list of structure files. For the i th entry, it reads the molecular structures for the associated pair of conformers and computes the contribution to the MMFF conformational energy difference from the input torsion parameters to be optimized. By adding this quantity, T_i^0 , to the difference between the reference and input MMFF conformational energy differences, TORFIT determines the total torsional contribution²⁶:

$$T_i = T_i^0 + \text{CE}_i^{\text{ref}} - \text{CE}_i^{\text{MMFF}} \quad (3)$$

that MMFF would have to produce to make the adjusted value for $\text{CE}_i^{\text{MMFF}}$ (i.e., $\text{CE}_i^{\text{MMFF}} - T_i^0 + T_i$) equal CE_i^{ref} . TORFIT then treats the T_i as the observables to be fit. This formulation makes it possible for TORFIT to impose Lagrangian constraints to control parameter correlations. For MMFF94, however, it proved equally efficacious and far more

convenient simply to impose penalty-function restraints.

LEAST-SQUARES REFINEMENTS

Using the initial set of torsion parameters, we optimized the "MP4SDQ/TZP" conformers, starting from the MP2/6-31G* geometries, while employing a strong torsion-restraint force constant (100 kcal/mol/rad²) to all torsion angles. Next, we carried out a first linear-least-squares refinement to derive new estimates for the torsion parameters. We then repeated this procedure with successively smaller torsion-restraint force constants (cf. Appendix A). Finally, we eliminated small "nuisance-value" parameters by zeroing all torsion parameters smaller than 0.05 kcal/mol in magnitude and reoptimizing the remaining nonzero parameters.

This procedure gave rise to what we call the "MMFF93" parameter set. We subsequently became aware of a few areas in which improvements needed to be made. One, for example, concerned the conformational energies of the medium-sized cycloalkyl rings discussed later in this study. Another related to MMFF93's tendency to overestimate the degree of pyramidalization at nitrogen shown in the reference MP2/6-31G* structures for amides. To remedy such deficiencies, we rederived subsets of the torsion parameters, using all affected conformational energies in each case. To reduce the tendency of amides to pyramidalize, we also changed the out-of-plane force constants for secondary and tertiary amides (MMFF atom types of 3, 1, 28 and of 3, 1, 1 connected to amide nitrogen, itself of type 10) from -0.030 and -0.034 to -0.02 md Å/deg²; the original values had been derived in fits to HF/6-31G* second derivatives.³ These modifications resulted in rms values for Wilson out-of-plane angles in amides that reflect much the same overall degree of puckering as that found in the reference MP2/6-31G* structures.³

Of 906 resultant one-, two- or three fold torsion parameters for 302 IJKL atom-type and torsion-type combinations, the combined protocol yielded zero values for 248 torsion parameters (27%). In addition, 39 twofold parameters that could not be optimized (cf. Appendix A) were fixed at initially assigned values. Thus, 619 nonzero torsion parameters were fit to ~ 1440 conformational energy differences. While the ratio of 2.3 observations per torsional degree of freedom is not high, the use of penalty-function weights in the least-squares pro-

cedure nevertheless appears to have produced a parsimonious set of well-defined parameters.

NONSTANDARD VALUES FOR CONFORMATIONAL ENERGY WEIGHTS

In a few cases, we tentatively employed larger weights (w_i) for particular conformational energy comparisons. Such enlarged weights were used, for example, in cases in which poor MMFF estimates for small "MP4SDQ/TZP" barriers to planarity in amides were associated with large differences in MMFF and MP2/6-31G* geometries at nitrogen. Enlarged weights were also used in a few cases in which a poor reproduction of energies for torsionally incremented structures relative to the local minimum led to MMFF-optimized torsion angles that disagreed significantly with the reference MP2/6-31G* values. Still other cases in which MMFF-optimized torsion angles disagreed significantly from the *ab initio* values required, instead, that additions be made to the set of conformational comparisons (cf. Appendix A). A final category consisted of cases in which part of the computational data could not be fit satisfactorily without severely compromising the fit to other data. In two such cases, we could not resolve the incompatibility, and set 21 weight factors for torsionally incremented comparisons (17 for ethylenediamine, 4 for *N*-methylformaldehydeimine) to small values to remove these comparisons from the torsion-parameter fitting (see "Limitations" section).

The changes to weight factors, the generation of additional computational data, as well as the modifications to MMFF's functional form made during the course of this work (see "Two Aspects" section) show more clearly than at perhaps any other point in MMFF94's derivation that the development of a complex force field needs to be an adaptive process.

MUTUALLY CONSISTENT DETERMINATION OF MMFF94 PARAMETERS

We interleaved the foregoing procedure for deriving torsion parameters with those described in part III³ for deriving quadratic force constants and reference bond lengths and angles. We also reevaluated, or where necessary rederived, nonbonded parameters involved in polar intermolecular interactions² prior to the final derivation of the MMFF94 torsion parameters. This iterative approach allowed parameters of each type to be optimized in

the context of increasingly well-refined values for parameters of other types.

Performance of MMFF94 in Relation to Other Force Fields

CALCULATED CONFORMATIONAL ENERGIES AND BARRIERS

In the listing below, the term "Conformational energies" refers to the 249 "MP4SDQ/TZP"-based comparisons, whereas "Torsion-profile energies" refers to the 1171 MP2/TZP torsionally incremented comparisons used in deriving the parameters. All results are stated as rms deviations in kilocalories per mole:

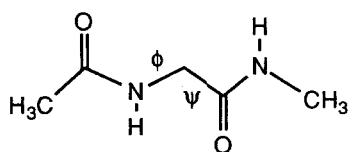
	MMFF94	MM2X
Conformational energies	0.31	1.12
Torsion-profile energies	0.50	1.57

For reference, the rms values are 3.88 kcal/mol for the conformational energies and 4.37 kcal/mol for the torsion-profile energies used in testing MMFF94, and 2.30²⁷ and 4.38 kcal/mol, respectively, for the somewhat smaller *ab initio* data sets used in assessing MM2X. Thus, MMFF94 accounts for about 90% of the variation in the *ab initio* relative energies in each case, whereas MM2X accounts for only about 50%.

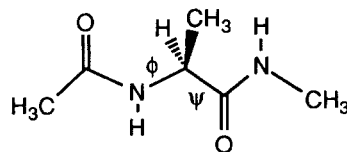
The ability of MMFF94 to reproduce the "MP4SDQ/TZP" conformational energies is shown in Table I (the five-character conformational indices used to identify the conformers are defined in part I¹). The general quality of the fit is evident in the fact that MMFF94 reproduces the sign of the "MP4SDQ/TZP" energy difference in 239 of 249 cases. Moreover, all ten errors in sign involve small absolute errors in relative energy. The largest absolute error for a qualitatively incorrect prediction, 0.42 kcal/mol, occurs for the comparison of

the axial- C_1 and equatorial- C_1 conformers of methoxycyclohexane (OR04c and OR04b, respectively), for which comparisons discussed in the next subsection suggest that it is the reference "MP4SDQ/TZP" calculations, not the MMFF94 calculations, that are faulty.

Some comparisons that relate to MMFF94's ability to describe protein backbone conformations are especially relevant to MMFF94's intended use in molecular-dynamics simulations. In particular, the AM09 and AM10 series explore conformational energies for the glycine and alanine dipeptide analogs studied computationally by Pople and coworkers,²⁸ and the related AM16 and AM17 series (cf. Fig. 1) extend the study to the full (methyl-capped) dipeptides. Table I shows that MMFF94 places the C5 conformer above the C7 by 1.14 kcal/mol for the glycine dipeptide analogs (AM09b and AM09a) and by 1.62 kcal/mol for the methyl-capped dipeptides (AM16b and AM16a, respectively). Other comparisons reveal that MMFF94 places the C7_{eq} conformers lowest in energy for both the alanine dipeptide analog (AM10a) and the dipeptide (AM17a), and places the C5 and C7_{ax} forms (conformers b and c) higher by 1.12 and 1.83 kcal/mol and by 1.68 and 2.26 kcal/mol, respectively. Furthermore, MMFF94 places the α' , β_2 , and α_L conformers (d, e, and f) from 2.3 to 5.5 kcal/mol above the C7_{eq} in each series. All of these MMFF94 conformational energies are in excellent agreement with the reference "MP4SDQ/TZP" values. The latter, it might be noted, differ significantly from some of the lower level HF/6-31 + G*//HF/6-31 + G* conformational energies reported by Pople and coworkers, but agree fairly well with their MP2/6-31 + G**//HF/6-31 + G* values.²⁸ Also important is that the torsionally incremented structures used in deriving the MMFF94 parameters included an extensive subset based on bond rotations for each of the AM09 and AM10 conformers. These compar-



AM16a, AM16b



AM17a — AM17f

FIGURE 1. Structures for the glycyl (AM16) and alanyl (AM17) dipeptides. The related glycyl (AM09) and alanyl (AM10) dipeptide analogs differ in structure only in that the two terminal methyl groups are replaced by hydrogen atoms.

isons helped to define equilibrium torsion angles and probed regions representing barriers to conformational interconversion; these data, too, are well described by MMFF94.

Though MMFF94 performs well in an overall sense, some areas of concern remain. Thus, while 227 of the 249 MMFF94 conformational energies agree with the reference "MP4SDQ/TZP" values to within 0.5 kcal/mol, 18 show deviations of 0.5–1 kcal/mol, and 4 are in error by up to 2 kcal/mol. Most of the larger discrepancies occur in highly polar, multifunctional compounds. Similar difficulties also affect some conformational comparisons within the set of MP2/TZP torsionally incremented structures. Some especially troublesome cases of this type are discussed later in this article.

EXPERIMENTAL CONFORMATIONAL ENERGIES

To see how well MMFF94 and MM2X predict experiment, especially in relation to the well-regarded MM3 force field, we examined a set of 37 conformational energies (Table II). Many of the same comparisons have been used in deriving or validating other molecular force fields.^{29–31} It is important to recognize at the outset that the experimental determinations form a diverse mixture of observables. Only in rare cases, such as for 1,3-butadiene and for 2-methyl-1,3-butadiene, does the experiment measure the same quantity as that obtained from the theoretical calculations, the energy difference between two potential minima. While a few determinations correspond to energy differences taken between ground vibrational states, most give relative enthalpies at, or sometimes below, room temperature; some, such those for methylcyclohexane, cyclohexanol, methoxycyclohexane, and cyclohexylamine, represent relative free energies. Furthermore, while most of the experimental measurements were made in the gas phase, some are for interconversions in solution. Most importantly, all are subject to error—and occasionally to significant error, as the divergent sets of experimental results cited below clearly show. For these reasons, even an exact theoretical calculation of differences between energy minima would not agree perfectly with experiment. To allow some check on experiment, particularly for cases in which two or more experimental results are available, all comparisons in Table II are ones for which we have also computed theoretical conformational energies using the "MP4SDQ/TZP"

model in conjunction with MP2/6-31G*-optimized geometries. For some, unapproximated MP4SDQ/TZP, or even QCISD(T)/TZP, conformational energies were obtained.

When more than one experimental result is listed in Table II, the reported rms deviations use the result that best agrees with the "MP4SDQ/TZP" model, or, in the case of MM3, the result that best agrees with MM3. For cyclohexylamine, we used 1.1 kcal/mol, the lower end of the listed experimental range. We first note that the "MP4SDQ/TZP" energies differ from experiment by a rms deviation of just 0.37 kcal/mol, about one sixth the rms value of 2.33 kcal/mol for the experimental relative energies. Other *ab initio* methods we have examined fare more poorly. Specifically, MP2/TZP calculations at the same MP2/6-31G* geometries show a slightly higher rms deviation of 0.44 kcal/mol, while MP2/6-31G*//MP2/6-31G* and HF/6-31G*//HF/6-31G* calculations show substantially higher rms deviations of 0.69 and 0.73 kcal/mol, respectively. These comparisons support the choice of the "MP4SDQ/TZP" model. It seems unlikely that a version of MMFF fit even to the next best model, MP2/TZP, would predict experiment as well, given that MMFF94 fits experiment *better* than does MP2/TZP (see below). *Especially clear is that a force field fit solely to MP2/6-31G* or to HF/6-31G* conformational energies would be expected to be substantially less accurate than MMFF94*, because these theoretical methods themselves reproduce the experimental conformational energies far less satisfactorily.

The comparisons presented in Table II also show that MM3 and MMFF94, with rms deviations of 0.37 and 0.38 kcal/mol, respectively, are comparable to "MP4SDQ/TZP" in accuracy, whereas MM2X (rms deviation, 0.84 kcal/mol) performs much less satisfactorily. Furthermore, they demonstrate that fitting MMFF94 to high quality computational data simultaneously confers the ability to reproduce experiment. Whether one considers MMFF94 or MM3 to be the more accurate depends on whether one believes that the experimental determinations or the *ab initio* calculations provide the better measure of the true conformational energies. Thus, while MM3 reproduces experiment slightly better than does MMFF94, Table II shows that MMFF94 (rms deviation, 0.33 kcal/mol) reproduces the "MP4SDQ/TZP" conformational energies much better than does MM3 (rms deviation, 0.49 kcal/mol). Given that significant discrepancies arise between "MP4SDQ/TZP" and experi-

TABLE I.
Comparison of "MP4SDQ/TZP" and MMFF94
Relative Conformational Energies (kcal/mol).

Conf_A ^a	Conf_B	E(A) – E(B) ^b		Δ
		"MP4"	MMFF94	
AM01t	AM01a	0.06	0.00	–0.06
AM02a	AM02j	1.21	1.16	–0.05
AM02b	AM02j	0.18	–0.12	–0.30
AM02k	AM02j	0.06	–0.11	–0.17
AM02l	AM02j	0.18	–0.02	–0.20
AM02t115	AM02j	18.02	18.17	0.15
AM02t120	AM02j	18.25	18.22	–0.03
AM02t125	AM02j	18.30	18.15	–0.15
AM02t130	AM02j	18.17	17.92	–0.25
AM02u55	AM02j	16.99	17.29 ^c	0.30
AM02u60	AM02j	17.21	17.39	0.18
AM02u65	AM02j	17.13	16.97	–0.16
AM02u70	AM02j	16.77	16.26	–0.49
AM03t	AM03a	0.29	0.01	–0.28
AM04a	AM04l	0.07	0.00	–0.07
AM04b	AM04l	1.96	2.18	0.22
AM04j	AM04l	0.41	0.13	–0.28
AM04k	AM04l	0.13	0.01	–0.12
AM04m	AM04l	1.97	2.17	0.20
AM04s	AM04b	0.08	0.02	–0.06
AM04t	AM04a	0.06	0.17	0.11
AM06t	AM06a	2.70	2.50	–0.20
AM07b	AM07a	0.41	0.40	–0.01
AM09b	AM09a	0.91	1.14	0.23
AM09s	AM09a	0.15	0.41	0.26
AM09t	AM09b	0.10	0.00	–0.10
AM10b	AM10a	1.08	1.12	0.04
AM10c	AM10a	2.20	1.83	–0.37
AM10d	AM10a	4.88	4.85	–0.03
AM10e	AM10a	2.76	2.47	–0.29
AM10f	AM10a	4.10	3.87	–0.23
AM12j	AM12a	0.59	0.73	0.14
AM13b	AM13a	0.04	0.01	–0.03
AM13s	AM13a	2.54	2.48	–0.06
AM13t	AM13b	1.16	1.13	–0.03
AM14a	AM14d	0.24	0.14	–0.10
AM14b	AM14d	1.50	1.56	0.06
AM14c	AM14d	0.08	0.27	0.19
AM15a	AM15c	0.04	0.08	0.04
AM15b	AM15c	0.96	1.07	0.11
AM15d	AM15c	0.72	1.09	0.37
AM16b	AM16a	1.66	1.62	–0.04
AM17b	AM17a	1.64	1.68	0.04
AM17c	AM17a	2.20	2.26	0.06
AM17d	AM17a	5.35	5.47	0.12
AM17e	AM17a	3.20	3.06	–0.14
AM17f	AM17a	4.25	4.75	0.50
AR23t	AR23a	1.10	1.19	0.09
CA01b	CA01a	4.79	4.90	0.11
CA02b	CA02a	5.85	5.87	0.02
CA03b	CA03a	0.85	0.84	–0.01
CA04a	CA04b	0.35	1.91	1.56
CA05b	CA05a	1.64	0.80	–0.84

TABLE I.
(continued)

Conf_A ^a	Conf_B	E(A) – E(B) ^b		Δ
		"MP4"	MMFF94	
CA07a	CA07b	0.21	0.25	0.04
CA08b	CA08a	2.22	0.97	–1.25
CA08c	CA08a	1.65	1.90	0.25
CA09b	CA09a	1.67	2.42	0.75
CE01b	CE01a	5.65	5.28	0.37
CE01j	CE01a	5.42	5.28	–0.16
CE01k	CE01a	5.55	5.50	–0.05
CE01l	CE01a	5.65	5.75	0.10
CE02a	CE02b	8.21	8.27	0.06
CE05b	CE05a	2.79	2.83	0.04
CE06b	CE06a	0.34	0.44	0.10
CE07b	CE07a	2.43	2.43	0.00
CE08b	CE08a	4.30	4.51	0.21
CE10b	CE10a	1.48	2.43	0.95
CJ01a	CJ01b	2.39	2.46	0.07
CJ01t	CJ01b	3.50	3.43	–0.07
CJ02a	CJ02b	2.20	2.12	–0.08
CJ03a	CJ03b	2.08	2.16	0.08
CJ04a	CJ04b	1.05	1.00	–0.05
CJ05b	CJ05a	0.61	0.60	–0.01
CJ06a	CJ06b	0.43	0.28	–0.15
CJ07a	CJ07b	2.03	2.05	0.02
CJ08a	CJ08b	3.14	3.15	0.01
CJ09b	CJ09a	0.60	0.62	0.02
CJ12j	CJ12a	1.35	1.24	–0.11
CJ13b	CJ13a	2.43	2.46	0.03
CJ13c	CJ13a	1.27	2.11	0.84
CO03b	CO03a	0.83	0.53	–0.30
CO05b	CO05a	0.99	0.83	–0.16
CO05j	CO05a	0.00	0.00	0.00
CO05k	CO05j	0.45	0.44	–0.01
CO05l	CO05j	1.10	1.12	0.02
CO05m	CO05j	0.99	0.87	–0.11
CO05n	CO05j	1.08	0.92	–0.16
CO05o	CO05j	1.48	1.58	0.10
CO05p	CO05j	2.37	2.42	0.05
CO06b	CO06a	0.69	0.89	0.20
CO07b	CO07a	0.14	0.11	–0.03
CO08a	CO08c	0.20	–0.05	–0.25
CO08b	CO08c	0.29	0.98	0.69
CO09a	CO09b	0.43	0.72	0.29
CO10b	CO10a	0.37	0.56	0.19
CO10j	CO10a	1.54	1.56	0.02
CO10k	CO10a	0.72	0.82	0.10
CO10l	CO10a	0.01	0.01	0.00
CO10m	CO10a	0.75	0.62	–0.13
CO10n	CO10a	1.39	1.47	0.08
CO10o	CO10a	0.85	0.99	0.14
CO10p	CO10a	0.37	0.56	0.19
CO11t	CO11a	0.25	0.16	–0.09
CO12a	CO12b	0.18	0.14	–0.04
CO13a	CO13b	0.70	0.77	0.07
CO14b	CO14a	5.40	5.46	0.06

(Continues on next page)

TABLE I.
(continued)

Conf_A ^a	Conf_B	E(A) – E(B) ^b		Δ
		"MP4"	MMFF94	
CO14t	CO14a	5.97	5.87	–0.10
HL04a	HL04b	0.58	0.63	0.05
HL05b	HL05a	1.29	1.24	–0.05
HL08j	HL08a	3.52	3.52	0.00
HL09a	HL09b	0.03	0.06	0.03
HL10b	HL10a	0.03	0.00	–0.03
IM01a	IM01b	2.03	2.27	0.24
IM01t	IM01b	2.91	2.72	–0.19
IM02t	IM02a	2.06	0.03	–2.03
IM04a	IM04b	2.00	2.12	0.12
IM04t	IM04a	0.90	0.55	–0.35
IM05t	IM05a	5.39	5.19	–0.20
IM06t	IM06a	5.59	5.80	0.21
IM07b	IM07a	1.73	1.94	0.21
KT02a	KT02b	1.30	1.81	0.51
KT03b	KT03a	2.43	2.11	–0.33
KT04b	KT04a	0.63	0.50	–0.13
KT05b	KT05a	2.25	2.97	0.70
NC03t	NC03a	2.57	2.72	0.15
NC05a	NC05b	0.07	0.03	–0.04
NC07b	NC07a	0.62	0.38	–0.24
NC11b	NC11a	1.26	1.18	–0.08
NH02b	NH02a	0.33	0.55	0.22
NH03a	NH03b	0.50	0.45	–0.05
NH03j	NH03b	2.74	2.67	–0.07
NH03k	NH03b	1.43	1.56	0.13
NH03l	NH03b	0.50	0.45	–0.05
NH03m	NH03b	1.64	1.57	–0.07
NH03n	NH03b	2.73	2.82	0.09
NH03o	NH03b	1.30	1.38	0.08
NH03p	NH03b	0.01	0.00	0.01
NH04b	NH04a	0.69	0.67	–0.02
NH06j	NH06a	2.32	2.32	0.00
NH07b	NH07a	0.78	0.90	0.12
NH09b	NH09a	3.58	3.28	–0.30
NH10t	NH10a	5.51	6.64	1.13
NH11b	NH11a	0.08	0.44	0.36
NH14t	NH14a	2.22	1.85	–0.37
NH15j	NH15a	5.23	5.26	0.03
NH16b	NH16a	0.50	0.51	0.01
NH16c	NH16a	0.15	0.65	0.50
NH17b	NH17a	0.67	0.61	–0.06
NH18a	NH18b	1.39	1.45	0.06
NH19t	NH19a	1.49	1.43	–0.06
NH20b	NH20a	1.38	1.41	0.03
NH21b	NH21a	1.69	2.25	0.56
NH22b	NH22a	0.68	0.73	0.05
NH23b	NH23a	0.19	0.05	–0.14
OH02a	OH02l	0.02	0.00	–0.02
OH02b	OH02l	0.08	–0.18	–0.26
OH02j	OH02l	1.47	1.18	–0.29
OH02k	OH02l	0.65	0.59	–0.06
OH02m	OH02l	0.73	0.62	–0.11

TABLE I.
(continued)

Conf_A ^a	Conf_B	E(A) – E(B) ^b		Δ
		"MP4"	MMFF94	
OH02n	OH02l	1.22	1.19	–0.03
OH02o	OH02l	0.55	0.47	–0.08
OH02p	OH02l	0.08	–0.18	–0.26
OH03b	OH03a	0.23	0.59	0.36
OH03c	OH03a	0.05	0.24	0.19
OH03d	OH03a	0.19	–0.17	–0.36
OH03e	OH03a	0.10	0.12	0.02
OH04b	OH04a	0.19	0.17	–0.02
OH04j	OH04a	1.40	1.32	–0.08
OH04k	OH04a	0.82	0.67	–0.15
OH04l	OH04a	0.03	0.00	–0.03
OH04m	OH04a	0.44	0.62	0.18
OH04n	OH04a	1.36	1.33	–0.03
OH04o	OH04a	0.80	0.80	0.00
OH04p	OH04a	0.19	0.17	–0.02
OH06a	OH06d	1.11	0.82	–0.29
OH06b	OH06d	1.05	0.60	–0.45
OH06c	OH06d	1.14	0.47	–0.67
OH06j	OH06d	6.70	6.32	–0.38
OH07a	OH07c	0.18	0.20	0.02
OH07b	OH07c	1.14	1.01	–0.13
OH07d	OH07c	0.33	0.32	–0.01
OH10a	OH10b	1.43	1.43	0.00
OH11b	OH11a	1.13	1.60	0.47
OH12a	OH12c	1.24	1.10	–0.14
OH12b	OH12c	1.11	1.13	0.02
OH13b	OH13a	0.38	0.47	0.09
OH14b	OH14a	0.10	0.33	0.23
OH14c	OH14a	0.18	0.11	–0.07
OH14d	OH14a	0.48	0.47	–0.01
OH14e	OH14a	0.29	0.46	0.17
OH14f	OH14a	0.56	0.59	0.03
OH14g	OH14a	0.75	0.72	–0.03
OH14h	OH14a	0.57	0.66	0.09
OH14i	OH14a	1.13	1.14	0.01
OH14r	OH14a	5.40	5.48	0.08
OH14s	OH14a	6.08	5.50	–0.58
OH14t	OH14a	1.69	1.55	–0.14
OH14u	OH14a	1.19	1.26	0.07
OH14v	OH14a	1.23	1.37	0.14
OH14w	OH14a	3.18	4.15	0.97
OH14x	OH14a	3.74	4.30	0.56
OH14y	OH14a	5.08	5.28	0.20
OH14z	OH14a	5.06	5.54	0.48
OH15a	OH15b	2.70	2.91	0.21
OH15c	OH15b	0.49	1.15	0.66
OH15d	OH15b	1.41	1.58	0.17
OR01b	OR01a	1.41	1.50	0.09
OR02b	OR02a	2.27	2.22	–0.05
OR03b	OR03a	1.47	1.52	0.05
OR04a	OR04b	2.10	1.94	–0.16
OR04c	OR04b	0.01	–0.41	–0.42
OR05a	OR05t	0.08	0.00	–0.08

(Continues on next page)

TABLE I.
(continued)

Conf_A ^a	Conf_B	E(A) - E(B) ^b		Δ
		"MP4"	MMFF94	
OR07t	OR07a	4.20	4.15	-0.05
OR11t	OR11a	3.61	3.56	-0.05
OR13b	OR13a	1.91	1.88	-0.03
OR14j	OR14a	1.91	1.90	-0.01
RA02t	RA02a	2.91	3.21	0.30
RA04b	RA04a	0.65	0.78	0.13
RA04t	RA04a	5.79	5.21	-0.58
RA06t	RA06a	2.42	2.42	0.00
RA07t	RA07a	5.17	5.51	0.34
RA08b	RA08a	6.14	5.93	-0.21
RA10b	RA10a	1.69	1.37	-0.32
RA12a	RA12b	0.04	-0.23	-0.27
RA14a	RA14b	2.00	1.44	-0.56
RA14c	RA14b	1.71	1.99	0.28
RA14d	RA14b	2.97	3.27	0.30
RA15b	RA15a	0.77	0.78	0.01
RA15j	RA15a	2.47	2.47	0.00
RA16a	RA16b	1.97	1.41	-0.56
RA16c	RA16b	0.99	1.21	0.22
RA16d	RA16b	3.65	3.81	0.16
RE03a	RE03b	0.26	0.26	0.00
RE04b	RE04a	0.61	0.51	-0.10
RE05a	RE05b	0.02	-0.03	-0.05
RE07a	RE07b	0.33	0.43	0.10
RE08b	RE08a	1.27	1.35	0.08
RE10b	RE10a	1.31	1.33	0.02
SR02b	SR02a	0.74	0.68	-0.06
SR04b	SR04a	0.76	0.76	0.00
SR07t	SR07a	0.17	0.08	-0.09
SR08t	SR08a	0.19	0.32	0.13
SR09b	SR09a	0.30	0.37	0.07
SR09c	SR09a	1.23	0.95	-0.28
SR11a	SR11c	0.86	0.94	0.08
SR11b	SR11c	1.77	2.01	0.24
SR12b	SR12a	0.09	0.05	-0.04
SR12c	SR12a	0.66	0.60	-0.06
SR12d	SR12a	0.25	0.32	0.07
conformational energy, error:		3.88	0.31	

^aSee Table I in ref. 1 for the notation used or the conformational indices and for a qualitative description of the conformations. Note: the *N*-methylformamide conformers labeled AM02t115 through AM02u70 are denoted as conformers AM02s through AM02z in ref. 1. The present labeling explicitly shows the value of the H—N—C=O torsion angle that was constrained in the *ab initio* and was restrained in the MMFF94 optimizations.

^bMP4SDQ/TZP energies at MP2/6-31G* optimized geometries, and MMFF94 energies at MMFF94 optimized geometries for 249 conformational comparisons. All energies are given in kilocalories per mole.

^cTo prevent "collapse," the MMFF94 optimization for this conformer restrained the C—N—C=O angle to the *ab initio* value of -63.8°. The H—N—C=O angle then optimized to 53.4°, not appreciably different than the angle of 55° imposed in the MP2/6-31G* optimization.

ment, as well as among different experimental determinations for the same conformational difference (see below), it is not immediately obvious that experiment is more accurate than theory for these systems.

A survey of comparisons to theory and experiment for MMFF94 and MM3 will now be given. For *N*-methylacetamide, the MMFF94 and the *ab initio* "MP4SDQ/TZP" results for the cis-trans difference agree more closely with the "rough" ΔH of 2.3 kcal/mol obtained in a nitrogen matrix by temperature-dependent infrared intensities³² than with the larger ΔG of 2.8 kcal/mol from NMR line-shape analysis³³; MM3 more nearly reproduces the latter.³⁴ The two force fields perform comparably for *N*-methylformamide.

Relatively large discrepancies between experiment and *ab initio* theory arise for the trans versus cis comparisons for formic acid and methyl formate.³⁵ It is not known whether these discrepancies mainly reflect limitations in theory or experiment. For glyoxalic acid, the higher lying conformer places one COOH moiety in an unfavorable trans conformation, but favorably positions the outwardly splayed hydroxyl to hydrogen bond to the second carbonyl oxygen. Here, the microwave determination for the difference in conformational energy falls neatly between the "MP4SDQ/TZP" and MMFF94 values. This discrepancy is one of the largest for MMFF94 in the fit to the *ab initio* "MP4SDQ/TZP" relative energies. It may well indicate that MMFF94's functional form (and MM3's, etc.) is too simple to handle highly polar, multifunctional compounds consistently well.

2-Butanone presents an especially interesting case, as the experimental determinations have found the skew conformer to lie above the cis conformer by 1.0–1.3 kcal/mol in the liquid phase but by 2.0–2.2 kcal/mol in the gas phase.³⁶ The higher gas-phase values, which have been used in developing and evaluating force-field models,^{29,30} had seemed to be supported by Wiberg and Martin's MP2/6-31G* calculations at HF/3-21G-optimized geometries (cf. Table III).³⁷ When MP2/6-31G*-optimized geometries are employed, however, the cis conformer is favored by just 1.497 kcal/mol, an amount only modestly above the range of the liquid-phase measurements. Moreover, a wide variety of still higher level calculations at MP2/6-31G* and at the nearly identical MP2/TZP-optimized geometries all strongly support the lower range. As Table III shows, for basis

sets of DZP and higher quality, the calculated conformational energy difference is relatively insensitive to the valence basis set, to the polarization basis set, and to the electron-correlation level. These calculations yield a best estimate of ca. 1.2 kcal/mol, which we obtain as the MP2/TZP(3d, 2f; 3p, 2d) triple first-polarization, double second-polarization difference of 1.091 kcal/mol augmented by allowance of ca. 0.1 kcal/mol for electron correlation beyond MP2. The rms deviations cited in Table II use the alternative experimental estimate, which we list in the table as the average of the stated range of 1.0–1.3 kcal/mol. We suggest that the measured gas-phase difference of 2.0–2.2 kcal/mol is in error.

The haloalkanes are a story in themselves. As Dixon and Smart observe, the anti, gauche difference for 1,2-difluoroethane has been the subject of intense scrutiny, with experimental determinations ranging from –1.41 kcal/mol to +1.98 kcal/mol.³⁸ This enormous range, and especially the fact that even the sign of the anti, gauche difference is in dispute, demonstrate that experimental determinations can incur very large errors, even for seemingly simple systems. Dixon and Smart favor (see Table II) the recent estimate of 0.8 kcal/mol from

gas-phase ¹⁹F and ¹H high-resolution NMR. That value agrees well with their very large basis set (TZ + 2D + P) calculations as well as with our more modest (but more highly correlated) “MP4SDQ/TZP” calculations—and with MMFF94. For 1,2-dichloroethane, an even wider variety of experimental determinations have been reported, though the measured differences in conformational energy fall into a relatively narrow range. The gauche–anti difference of 1.16 kcal/mol listed in Table II, which represents the average of 13 such determinations cited by Dixon et al.,³⁹ agrees well with our “MP4SDQ/TZP” results and with MMFF94. For the 1-halopropanes, experiment indicates that the fluoro,⁴⁰ chloro,^{40,41} and, remarkably, even the bromo⁴² compound prefer the gauche conformation. The “MP4SDQ/TZP” and MMFF94 results do not fully agree, but the discrepancies are not large. MM3 studies for bromides, chlorides, and fluorides have not yet been published.

For 2-methoxytetrahydropyran, MMFF94 correctly places the equatorial methoxy conformer above the “anomeric” stabilized axial methoxy conformer, but considerably exaggerates the difference. The “MP4SDQ/TZP” and MM3 results, in

TABLE II.
Experimental and Calculated Conformational Energies (kcal/mol).

Conformational comparison	Relative energy, enthalpy, or free energy				
	Exp.	“MP4” ^a	MMFF94	MM2X	MM3
<i>N</i> -methylformamide, cis–trans	1.45 ^b	1.04	1.28	0.06	1.84 ^e
<i>N</i> -methylacetamide, cis–trans	2.3, ^c 2.8 ^d	1.90	2.18	1.46	2.94 ^e
Formic acid, trans–cis	3.90 ^f	4.79	4.90	4.40	3.98 ^f
Glyoxalic acid, (O=C–C=O t, O=C–O–H c) – (O=C–C=O t, O=C–O–H t)	1.2 ^g	0.35	1.91	—	—
Methyl formate, trans–cis	3.85, 4.75 ^f	5.65	5.28	5.13	4.94 ^f
Methyl acetate, trans–cis	8.5 ^f	8.21	8.27	4.83	8.71 ^f
Ethyl formate, (O=C–O–C c, C–O–C–C g) – (O=C–O–C c, C–O–C–C a)	0.19 ^f	0.34	0.44	0.23	0.18 ^f
Propionaldehyde, skew–cis	0.67, ^h 0.95 ⁱ	0.84	0.53	0.53	1.16 ⁱ
Butanone, skew–cis	1.15, 2.1 ^j	0.98	0.83	1.50	1.39 ⁱ
1,3-Butadiene, gauche–trans	2.5 ^k	2.39	2.46	2.30	1.85 ^l
2-Methyl-1,3-butadiene, gauche–trans	2.65 ^m	2.20	2.12	2.08	1.66 ^l
Acrolein, cis–trans	1.9 ^k	2.03	2.05	0.91	1.45 ⁿ
1,2-Difluoroethane, anti–gauche	0.8 ^o	0.58	0.63	0.03	—
1,2-Dichloroethane, gauche–anti	1.16 ^p	1.29	1.24	1.59	—
1-Fluoropropane, anti–gauche	0.35 ^q	0.03	0.06	–0.17	—
1-Chloropropane, anti–gauche	0.05, ^r 0.36 ^q	–0.03	0.00	–0.36	—
2-Methoxytetrahydropyran, (OCOMe g, OMe eq) – (OCOMe g, OMe ax)	1.05 ^s	1.30	1.81	1.99	0.88 ^s
2,5-Dimethyl-1,3-dioxane (2 eq, 5 ax) – (2 eq, 5 eq)	0.9 ^s	0.63	0.50	0.84	0.67 ^s

(Continues on next page)

TABLE II.
(continued)

Conformational comparison	Relative energy, enthalpy, or free energy				
	Exp.	"MP4" ^a	MMFF94	MM2X	MM3
Isopropylamine, Ip — N — C — H anti — gauche	0.45 ^t	0.50	0.45	0.10	0.22 ^t
Cyclohexylamine, ax — eq	1.1 — 1.8 ^u	0.69	0.67	1.17	—
Piperidine, ax — eq	0.4 ^t	0.78	0.90	— 0.30	0.29 ^t
N-methylpiperidine, ax — eq	3.15 ^v	3.58	3.28	2.23	—
Ethanol, gauche — anti	0.12, ^w 0.4 ^x	— 0.06	0.18	0.62	0.40 ^s
Isopropanol, H — C — O — H anti — gauche	0.28 ^y	0.20	0.17	0.60	—
Cyclohexanol, ax C ₁ — eq C ₁	0.52 ^s	0.33	0.32	0.75	0.74 ^s
Methyl ethyl ether, gauche — anti	1.5 ^s	1.41	1.50	1.75	1.49 ^s
Methyl vinyl ether, C=C — O — C skew — cis	1.7 ^z	2.27	2.22	—	2.44 ^z
Diethyl ether, (C — C — O — C a, C — O — C — C g) — (C — C — O — C a, C — O — C — C a)	1.1 ^s	1.48	1.52	1.77	1.51 ^s
Methoxycyclohexane, ax C ₁ — eq C ₁	0.45 ^{aa}	— 0.01	0.41	0.74	0.77 ^s
Butane, gauche — anti	0.75, ^{bb} 0.97 ^v	0.65	0.78	0.36	0.81 ^{bb}
Cyclohexane, twist — boat — chair	5.5 ^{cc}	6.14	5.93	5.52	5.76 ^{bb}
Methylcyclohexane, ax — eq	1.75 ^{dd}	1.69	1.37	1.86	1.77 ^{bb}
2,3-Dimethylbutane, H — C ₂ — C ₃ — H gauche — H — C ₂ — C ₃ — H anti	0.05, 0.17 ^{dd}	0.04	— 0.23	0.17	0.38 ^{bb}
Cyclooctane, D _{4d} — C _s boat chair	1.9 ^{bb}	2.00	1.44	0.97	1.12 ^{bb}
Cyclononane, [255] C ₂ — [333] D ₃	0.95 ^{ee}	0.98	1.21	0.67	0.84 ^{bb}
1-Butene, cis — skew	0.53 ^{ff}	0.26	0.26	1.10	0.69 ^{ff}
2-Butene, cis — trans	1.0 ^{ff}	1.27	1.35	1.55	—
(rms value), rms deviation vs. exp.	(2.33)	0.37 ^{gg}	0.38 ^{gg}	0.84 ^{gg}	0.37 ^{hh}
rms deviation vs. "MP4SDQ / TZP"	—	—	0.33	0.81	0.49

^a"MP4SDQ / TZP" calculations (see text).^bAverage of values cited in: D. M. Schnur, Y. H. Yuh, and D. R. Dalton, *J. Org. Chem.*, **54**, 3779–3785 (1989).^cRef. 32.^dRef. 33.^eRef. 34.^fAs cited in ref. 35.^gB. P. van Eijck and F. B. van Duijneveldt, *J. Mol. Struct.*, **39**, 157–163 (1977).^hJ. R. Durig, D. A. C. Compton, and A. Q. McArver, *J. Chem. Phys.*, **73**, 719–724 (1980).ⁱAs cited in N. L. Allinger, K. Chen, M. Rahman, and A. Pathiaseril, *J. Am. Chem. Soc.*, **113**, 4505–4517 (1991).^jRef. 36.^kAverage of values cited in K. B. Wiberg, P. R. Rablen, and M. Marquez, *J. Am. Chem. Soc.*, **114**, 8654–8668 (1992).^lRef. 57.^mY. N. Panchenko, V. I. Pupyshev, A. V. Abramnikov, M. Traetteberg, and S. J. Cyvin, *J. Mol. Struct.*, **130**, 355–359 (1985).ⁿN. L. Allinger, S. Rodriguez, and K. Chen, *J. Mol. Struct.*, **260**, 161–178 (1992).^oPreferred value cited in ref. 38. Other cited values range from –1.41 to +1.98 kcal/mol.^pAverage of 13 experimental determinations cited in ref. 39.^qRef. 40.^rRef. 41.^sAs cited in ref. 43.^tAs cited in L. R. Schmitz and N. L. Allinger, *J. Am. Chem. Soc.*, **112**, 8307–8315 (1990).^uAs reported by and cited in ref. 44.^vAs cited in ref. 30.^wRef. 46.^xRef. 45.^yE. Hirota, as cited in W. A. Latham, L. Radom, W. J. Hehre, and J. A. Pople, *J. Am. Chem. Soc.*, **95**, 699 (1973).^zAs cited in N. L. Allinger and L. Yan, *J. Am. Chem. Soc.*, **115**, 11918–11925 (1993).^{aa}As cited in N. L. Allinger and D. Y. Chung, *J. Am. Chem. Soc.*, **98**, 6798 (1976).^{bb}As cited in ref. 7.^{cc}M. Squillacote, R. S. Sheridan, O. L. Chapman, and F. A. L. Anet, *J. Am. Chem. Soc.*, **97**, 3244–3246 (1975).^{dd}As cited in ref. 29b.^{ee}Ref. 51.^{ff}As cited in N. L. Allinger, F. Li, and L. Yan, *J. Comput. Chem.*, **11**, 848–867 (1990).^{gg}Computed using the experimental result closest to the MP4SDQ / TZP value when more than one experimental value is listed.^{hh}Computed using the experimental result closest to the MM3 value.

TABLE III.
Relative Conformational Energies for 2-Butanone
(kcal/mol).

Geometry	Method ^a	$\Delta E(\text{skew} - \text{cis})$
HF/3-21G	MP2/6-31G*	1.83 ^b
HF/6-31G*	MP3/6-31 + G**	1.177
MP2/6-31G*	MP2/6-31G*	1.497
MP2/6-31G*	MP2/DZP	1.015
MP2/6-31G*	MP4SDTQ/DZP	1.053
MP2/6-31G*	MP2/TZP	0.852
MP2/6-31G*	MP4SDQ/TZP	0.961
MP2/6-31G*	QCISD(T)/TZP	0.930
MP2/TZP	MP2/TZP	0.865
MP2/TZP	MP4SDQ/TZP	0.984
MP2/TZP	QCISD(T)/TZP	0.987
MP2/TZP	MP2/TZ(d, f; p) ^c	0.929
MP2/TZP	MP4SDQ/TZ(d, f; p) ^c	1.063
MP2/TZP	QCISD(T)/TZ(d, f; p) ^c	1.036
MP2/TZP	MP2/TZ(2d, f; 2p, d) ^d	1.035
MP2/TZP	MP2/TZ(3d, 2f; 3p, 2d) ^e	1.091

^a See text for basis sets and correlation methods used.

^b Ref. 37.

^c Basis set adds a set of f-type Gaussian atomic orbitals on carbon and oxygen with exponents of 0.78 and 1.56, respectively.

^d Basis set splits the TZP polarization exponents by factors of $3^{1/2}$ and $3^{-1/2}$, adds the f-type functions of the (d, f; p) set to C and O, and adds d-type functions to H with an exponent of 0.78.

^e Basis set adds two additional first-polarization functions to the TZP set with exponents threefold larger and smaller than the single polarization function of that set, and splits the second-polarization functions of the (2d, f; 2p, d) set by factors of $3^{1/2}$ and $3^{-1/2}$.

contrast, agree well with experiment ($\Delta H = 1.05$ kcal/mol in tetralin).⁴³ For cyclohexylamine, all the experimental results for the axial – equatorial difference (ΔG , in solution)⁴⁴ are larger than the “MP4SDQ/TZP” and MMFF94 differences. However, the *ab initio* calculations also underestimate, by 0.2–0.5 kcal/mol, the corresponding, even better established, difference in conformational energies for cyclohexanol and for methoxycyclohexane. As it happens, MMFF94 closely reproduces the “MP4SDQ/TZP” result for the amine and the alcohol, and hence, also underestimates the experimental difference, but diverges from the *ab initio* result for the methyl ether in just the manner required to bring MMFF94 into line with experiment. Methoxycyclohexane would seem to be a case in which the “MP4SDQ/TZP” calculations are in error. Interestingly, the underlying MP2/TZP calculations wrongly favor the axial conformer even more strongly, by 0.21 kcal/mol.

The MP3 and MP4SDQ perturbative corrections adjust the conformational difference in the proper direction, but not far enough; perhaps further correlative corrections are needed, or perhaps a conformationally dependent “basis set superposition error” favors the axial form, which volume and surface-area calculations show to be slightly more compact.

For ethanol, Allinger and coworkers⁴³ parameterize MM2 and MM3 to reproduce an experimental preference for the anti conformation of ~ 0.4 kcal/mol (ΔG) based on measured gas-phase infrared intensities of the O–H stretching bands and on the assumption that the oscillator strengths are equal.⁴⁵ A microwave determination, however, places the gauche conformer just 0.12 kcal/mol higher.⁴⁶ The “MP4SDQ/TZP” calculations agree better with the latter determination, but actually slightly favor the gauche. Remarkably, we find that MP2/TZP calculations at MP2/TZP-optimized geometries also place the gauche conformer lower, by 0.032 kcal/mol, and that QCISD(T)/TZP calculations at the MP2/TZP geometries further increase the difference to 0.111 kcal/mol in favor of *gauche*-ethanol. With respect to experiment, MMFF94 more nearly reproduces the microwave result, which we believe to be the better founded, while MM3 reproduces the infrared result. It is disconcerting, though, that very good theory continues to differ qualitatively from experiment.

For the alkanes and alkenes examined, the experimental, “MP4SDQ/TZP,” MMFF94, MM2X, and MM3 results are generally in good accord. MMFF94 orders the two conformers of 2,3-dimethylbutane incorrectly, though the absolute error is not large. For cyclooctane, all the methods correctly identify the lowest-energy conformer. For this system, Anet⁴⁷ places a crown family conformer having D_{4d} (“crown”), D_2 (“twist-chair-chair”), or C_{2v} (“chair-chair”) symmetry 1.9 ± 0.2 kcal/mol above the lowest-lying “boat-chair” conformer of C_s symmetry (conformer RA14b in Table I). In good agreement, the “MP4SDQ/TZP” calculations place the highly symmetric D_{4d} crown conformer (RA14a) 2.00 kcal/mol above the C_s boat-chair and place the D_2 twist-chair-chair 2.13 kcal/mol above the boat-chair. MMFF94 places the twist-chair-chair 1.44 kcal/mol above the boat-chair, in qualitative agreement with theory and experiment, but finds the D_{4d} crown to be a double-partial maximum that lies 0.18 kcal/mol above the D_2 twist-chair-chair.

Vibrational analysis⁴⁸ predicts the MMFF94 population of the twist-chair-chair conformation to be 21% at room temperature, roughly threefold larger than the 6% measured by Anet for the crown family population. Evidently, MMFF94 gives too small a difference in steric energy, while that for the "MP4SDQ/TZP" model is probably about right. However, MMFF94 does find the entropy of the twist-chair-chair to be 1.50 eu higher than that of the boat-chair, in qualitative accord with Anet's estimate of 1 ± 1 eu. As Allinger and coworkers point out,⁷ the fact that the twist-chair-chair has a symmetry number of 4 but is chiral *disfavors* this conformer by $R \ln 2 = -1.38$ eu. For MMFF94, however, we find that the vibrational contribution to the entropy more than cancels this symmetry-number contribution. MM3 predicts an even lower relative steric energy for the twist-chair-chair than does MMFF94 (cf. Table II). Nevertheless, using the unfavorable entropy difference of -1.38 eu just cited, Allinger and coworkers calculate 7% twist-chair-chair at room temperature, in good agreement with Anet's result. Their calculation, however, appears to have neglected the vibrational contributions to the enthalpy and entropy that led to the discrepancy for MMFF94. Finally, we note that the "MP4SDQ/TZP" and MMFF94 calculations find the C_2 twist-boat-chair (conformer RA14c) to lie 1.71 and 1.99 kcal/mol above the C_s boat-chair. The fourth stable conformer, RA14d, has S_4 symmetry⁴⁹ and lies above the boat-chair by 2.97 kcal/mol for "MP4SDQ/TZP" and by 3.27 kcal/mol for MMFF94. At 298 K, the MMFF94 free energies relative to the twist-chair-chair (RA14b) are 0.72, 1.42, and 3.22 kcal/mol for RA14a, RA14c, and RA14d, respectively.

For cyclononane, all four methods surveyed agree that the [333] conformer⁵⁰ of D_3 symmetry (RA16b in Table I) is lowest in energy and would predominate at low temperatures. From integrated NMR peak areas, Anet and Krane⁵¹ find that a second conformer, which they argue is the [225] conformer of C_2 symmetry (RA16c), comprises 12% of the mixture at 128 K. Assuming a force-field-calculated relative entropy of $+3.5$ eu, they conclude that this second conformer lies 0.95 kcal/mol above the [333] (ΔH) and would be slightly favored at room temperature. Based on a departure from linearity found when the ^{13}C -NMR linewidth is plotted against reciprocal temperature, they infer that a third conformation, which they assign as the [144] conformer of C_2 symmetry (RA16a), is present at a relative concentration of

0.5–1.5% at 178 K, corresponding to a ΔG of 1.6 ± 0.2 kcal/mol. Assuming a force-field-calculated relative entropy of 3.3 eu, they place this structure 2.2 ± 0.2 kcal/mol above the [333] in enthalpy. These values for ΔH and ΔS , if taken to be temperature-independent, yield populations for the [333], [225], and [144] conformers of 95%, 5%, and 0.01% at 100 K and of 43%, 51%, and $6\% \pm 2\%$ at 298 K (Anet and Krane round the latter values to 40%, 50%, and 10% in their study).

MMFF94 places the [225] and [144] conformers 1.21 and 1.41 kcal/mol above the [333] in energy (ΔE), qualitatively, but not quantitatively, in accord with the "MP4SDQ/TZP" relative energies of 0.99 and 1.97 kcal/mol, respectively. As can be seen from Table IV, MMFF94 assigns relative entropies for the [225] and [144] conformers that are very similar to those assumed by Anet and Krane. For MMFF94, however, enthalpy corrections at 298 K disfavor the [225] by 0.47 kcal/mol (wrongly), but favor the [144] by 0.11 kcal/mol relative to the [333], with the result that room-temperature populations of 49%, 17%, and 30% are predicted for [333], [225], and [144]. These populations differ somewhat from those obtained from Anet and Krane's data. When entropy and enthalpy contributions⁴⁸ obtained using MP2/6-316* vibrational frequencies are applied to the "MP4SDQ/TZP" relative energies, however, predicted populations of 26%, 60%, and 12% are found (Table IV), in far better agreement. Moreover, the *ab initio* model places the [225] and [144] conformers 0.84 and 1.76 kcal/mol above the [333] in enthalpy at 298 K, in reasonable agreement with Anet and Krane's lower temperature estimates of 0.95 and 2.2 ± 0.2 kcal/mol.

The MM3 calculations reported for cyclononane reproduce Anet and Krane's more accurately,⁷ but appear to have used relative entropies favoring the [225] and [144] conformers that include the symmetry-number correction ($R \ln 6 - R \ln 2 = 2.19$ eu), but omit vibrational contributions that the MMFF94 and MP2/6-31G* vibrational analyses find to be significant. Very recently, Allinger et al. reported MM4 calculations for cyclononane that properly include all thermochemical contributions.⁵² These calculations give 79%, 10%, and 11% at 100 K and 20%, 34%, and 37% at 298 K for [333], [225], and [144], respectively. Especially at 100 K, where the experimental data imply that the third conformer would be present to only about 0.01%, these results differ considerably from experiment. Allinger et al., however, suggest that the NMR resonances for [144] are buried under those

TABLE IV.
Thermochemistry for Cyclononane.^a

MMFF94	ΔE	ΔH_{298}^b	S_{298}	ΔG_{298}^c
RA16a [144]	1.408	166.08	92.93	138.37
RA16b [333]	0.000	164.79	89.61	138.07
RA16c [225]	1.211	166.47	93.18	138.69
RA16d [234]	3.810	168.71	94.35	140.58
RA16e [9a] ^d	3.389	168.05	95.15	139.68
"MP4SDQ/TZP"	ΔE	$\Delta H_{298}^{b,e}$	S_{298}	ΔG_{298}^f
RA16a [144]	1.964	173.83	95.01	145.50
RA16b [333]	0.000	172.07	90.68	145.04
RA16c [225]	0.985	172.91	95.16	144.53
RA16d [234]	3.645	175.78	95.51	147.30
RA16e [9a] ^d	3.373	175.27	96.66	146.45

^aIn kcal/mol and eu.^bObtained from the listed ΔE by adding translational, rotational, and zero-point and thermal vibrational contributions (cf. ref. 48).^cThe calculated percent composition at 298 K is 30% [144], 49% [333], 17% [225], 0.7% [234], and 3% [9a]. A similar analysis gives 0.9% [144], 99% [333], 0.3% [225], and vanishing small amounts of [234] and [9a] at 100 K.^dSee text.^eBased on "MP4SDQ/TZP" relative energies at MP2/6-31G*-optimized geometries and on MP2/6-31G* vibrational frequencies.^fThe calculated percent composition at 298 K is 12% [144], 26% [333], 60% [225], 0.5% [234], and 2% [9a]. A similar analysis gives 0.1% [144], 85% [333], 15% [225], and vanishing small amounts of [234] and [9a] at 100 K.

for the [333] conformer and were unobserved, and that a conformer not considered by Anet and Krane corresponds to the third (minor) species detected experimentally. We denote this conformer, which we call RA16e, as [9a] in Table IV⁵³; it is an asymmetric version of Dale's conformer, [9] = [144],⁵⁰ with which it shares the same qualitative pattern of ring torsion angles. The MM4 populations of 90%, 10%, and 0% at 100 K and of 58%, 34%, and 7% at 298 K for [333] + [144], [225], and [9a] compare reasonably closely with those inferred experimentally for the three conformers detected by Anet and Krane. Our *ab initio* results, however, are consistent with Anet and Krane's assignment of [333], [225], and [144]. Moreover, the "MP4SDQ/TZP"-MP2/6-31G* model places the [9a] conformer above the [333] at 298 K by 3.37 kcal/mol in energy, 3.20 kcal/mol in enthalpy, and 1.41 kcal/mol in free energy, leading to a predicted composition containing just 2% [9a] at 298 K; MMFF94 gives comparable results. Furthermore, the *ab initio* model predicts a population of

only about 0.2% [9a] at 178 K, probably too small to explain the observed NMR line broadening. We believe that Anet and Krane's original interpretation is correct and suggest that the relative steric energies obtained by MM4 for the [144] and [9a] conformers are too low.

To return now to a discussion of Table II, we note that the results cited in the table also permit some comparisons to be made to UFF³⁰ and to CHARMM.³¹ The recent study of conformational energies and rotational barriers by Gundertofte et al.^{29a} allows an even more extensive set of comparisons to be made to the MM2*, MM3*, and AMBER* force fields implemented in MacroModel; to the modified MM2 force field implemented in Chem3D Plus; to the DREIDING, UFF, and MMFF93 force fields implemented in CERIUSS², to the version of the TRIPOS force field implemented in Alchemy, to the MMX force field implemented in PCModel, and to the CVFF and CFF(91) force fields implemented in Insight/Discover; Gundertofte et al. also examined MM2(85), MM2(91), and MM3(92),^{29a} and in earlier work assessed the semiempirical AM1 and PM3 models (both of which fared poorly).^{29b}

For the systems in Table I, CHARMM appears to do well,³¹ though too few comparisons can be made to allow firm conclusions to be drawn. UFF also does fairly well in many cases, but reverses the order of stability of the axial and equatorial conformers in the anomeric 2-methoxyhydropyran system, inverts the stabilities of the cis and skew conformers for 2-butanone, incorrectly predicts the ordering of the anti and gauche conformers for 1,2-difluoroethane, and considerably overestimates the relative energy for the twist-boat conformer of cyclohexane.³⁰ These UFF calculations, we should note, ignore electrostatic interactions.

Gundertofte et al. found comparable problems for UFF in calculations in which electrostatic interactions were ignored, and encountered even more pronounced difficulties when charges based on the recommended QEq charge-equilibration procedure were used; the resultant UFF model produced an rms error that *exceeded* the rms value of the conformational energies being calculated! DREIDING, the Tripos force field, CVFF, and CFF(91) also fared relatively poorly—though CFF(91) encountered significant difficulties only for halogen-containing compounds, where fall-back generic parameters had to be used. Of the 15 force-field implementations surveyed, MMFF93—the immediate precursor to MMFF94—stood among the very best; MMFF93 finished in a statistical dead heat

with the various MM2 and MM3 implementations considered, and in actual test score surpassed all except MM2* (which it may even have edged, had the alternative experimental value for 2-butanone discussed above been used).^{29a} Given that MMFF93 and MMFF94 are not expected to differ greatly (and sometimes not at all) for the systems considered by Gundertofte et al., these comparisons confirm that MMFF94 stands with MM2 and MM3 as being generally successful in predicting experimental conformational energies.

EXPERIMENTAL ROTATIONAL BARRIERS

In Table V, we compare experimental and MMFF94 rotational barriers for 30 molecular systems. Where possible, we have drawn these comparisons from compilations previously used in force-field development.^{29,54} We also list published MM3 barriers where these are available. For a few cases, we list for comparison the *ab initio* barrier that we obtained from "MP4SDQ/TZP" calculations at constrained MP2/6-31G*-optimized geometries or that we more crudely estimated from MP2/TZP calculations at torsionally incremented geometries. The MMFF94 barriers are based on energy comparisons involving

transition-region structures that were optimized subject to an imposed torsion-angle restraint; in cases in which the value of the restrained torsion angle at the transition state was not determined by symmetry considerations, a succession of closely incremented values was used to locate the transition state.

Examination of Table V shows that MMFF94 reproduces the experimental barriers very well. The listed overall rms deviation is 0.39 kcal/mol, just over 10% of the rms value of 3.67 kcal/mol for the 28 experimentally determined barriers included in the comparison. Larger discrepancies are found in some cases, however, and the experimental barrier seems to be poorly established in others, particularly in the two cases excluded from the rms comparison. Comments on such instances follow.

Among the methyl-group rotations, a few cases stand out. In neopentane, the experimental value for the methyl-rotation barrier seems a bit high, as the *ab initio* calculations show a gentler rise with increasing congestion. The MMFF94 barrier is also lower than the experimental value, but the MM3 barrier, intentionally,⁷ is lower still. In contrast, the *ab initio* results are in agreement with the experimental finding that the methyl barrier is substan-

TABLE V.
Comparison of Experimental^a and Calculated Rotational Barriers about Single Bonds (kcal / mol).

Torsion bond	Exp.	MMFF94	MM3	<i>Ab initio</i>
CH ₃ —CH ₃	2.88	3.22	2.41 ^b	2.91 ^c
CH ₃ —CH ₂ CH ₃	3.4	3.41	—	3.3 ^{d,e}
CH ₃ —C(CH ₃) ₃	4.7 ^f	4.12	3.35 ^b	4.1 ^d
CH ₃ —CH=CH ₂	2.0	1.96	1.74 ^g	2.2 ^d
CH ₃ —CH=O	1.17 ^h	2.02	1.08 ⁱ	1.4 ^d
CH ₃ —C(=O)CH ₃	0.78	0.83	0.74 ⁱ	—
CH ₃ —COOH	0.48	0.36	0.49 ^j	—
CH ₃ —CH ₂ F	3.33	3.22	—	—
CH ₃ —CH ₂ Cl	3.68	4.04	—	—
CH ₃ —CF ₃	3.25	3.16	—	—
CH ₃ —CCl ₃	5.75	6.07	—	—
CH ₃ —NH ₂	1.98	2.36	1.45 ^k	—
CH ₃ —NHCH ₃	3.62	3.52	—	—
CH ₃ —OH	1.07	1.23	0.78 ^l	1.1 ^d
CH ₃ —OCH ₃	2.7	2.43	2.45 ^l	2.7 ^d
CH ₃ —OCH=O	1.19	0.78	1.13 ^j	—
CH ₃ —OCH=CH ₂ ^m	3.83	3.74	—	3.7 ^d
CH ₃ —SH	1.27	1.38	1.28 ⁿ	—
CH ₃ —SCH ₃	2.11	1.83	2.25 ⁿ	—
CH ₃ CH ₂ —CH ₂ CH ₃ ^o	3.4 ^f	3.96	3.30 ^b	—
CH ₃ CH ₂ —CH ₂ CH ₃ ^p	4.54 ^q	5.21	4.83 ^b	5.79 ^c

(Continues on next page)

TABLE V.
(continued)

Torsion bond	Exp.	MMFF94	MM3	Ab initio
H ₂ C=CH—CH=CH ₂ ^r	5.96 ^s	6.19	7.31 ^t	
H ₂ C=CH—CH=O ^u	7.9 ^v	7.81	7.76 ^v	8.0 ^d
Phenyl—CH ₂ CH ₃	1.16 ^h	1.19	1.9 ^w	1.10 ^c
Phenyl—OH	3.3 (av)	3.56	3.30 ^x	3.6 ^d
Phenyl—OCH ₃	3.6, < 3.1 ^y	1.90 ^z	1.1 ^x	1.91 ^c
Phenyl—SCH ₃	1.0 ^h	0.48	—	
Phenyl—C(=O)CH ₃	3.15, 5.35 ^{aa}	4.42	3.90 ^v	5.2 ^d
H ₂ C=CH—OCH ₃ ^{bb}	6.33 ^{cc}	5.86	6.95 ^x	6.0 ^d
CH ₃ S—SCH ₃ ^{dd}	6.8 ^{ee}	8.14 ^y	5.53 ^{ff}	8.6 ^d
rms value	3.67			
rms deviation		0.39		

^aUnless otherwise specified, the experimental barriers are from: J. P. Lowe, *Prog. Phys. Org. Chem.*, **6**, 1 (1968).^bRef. 7.^cBased on "MP4SDQ/TZP" calculations at optimized MP2/6-31G* geometries.^dEstimated from single-point MP2/TZP calculations at torsionally incremented geometries computed as described in the text.^eCalculated for the terminal methyl group in *anti*-butane.^fK. H. Hellwege, Ed., *Molecular Constants*, Landolt-Boerstein Group II, Springer-Verlag, New York, 1982, Vol. 14, Subvol. A.^gN. L. Allinger, F. Li, and L. Yan, *J. Comput. Chem.*, **11**, 848–867 (1990).^hAs cited in ref. 54.ⁱN. L. Allinger, K. Chen, M. Rahman, and A. Pathiaseril, *J. Am. Chem. Soc.*, **113**, 4505–4517 (1991).^jRef. 35.^kL. R. Schmitz and N. L. Allinger, *J. Am. Chem. Soc.*, **112**, 8307–8315 (1990).^lRef. 43.^mMethyl rotation barrier in the *cis* conformer.ⁿN. L. Allinger, M. Quinn, M. Rahman, and K. Chen, *J. Phys. Org. Chem.*, **4**, 647–658 (1991).^oBarrier relative to the *anti* conformer; the MMFF94 C—C—C angle at the transition state is 121.0°.^p*syn* (C—C—C—C = 0°)–*anti*.^qAs cited in ref. 55.^rBarrier relative to the *trans* conformer; the MMFF94 C=C—C=C angle at the transition state is 107.0°.^sRef. 60.^tRef. 57.^uBarrier relative to the *trans* conformer; the MMFF94 C=C—C=O angle at the transition state is 92.9°.^vN. L. Allinger, S. Rodriguez, and K. Chen, *J. Mol. Struct.*, **260**, 161–178 (1992).^wRef. 29a.^xN. L. Allinger and L. Yan, *J. Am. Chem. Soc.*, **115**, 11918–11925 (1993).^yAs cited in ref. 63.^zNot included in the calculation of the rms deviation; see text.^{aa}Refs. 66 and 67.^{bb}Barrier relative to the *cis* conformer; the MMFF94 C=C—O—C angle at the transition state is 77.4°.^{cc}J. R. Durig and D. A. Compton, *J. Chem. Phys.*, **69**, 2028–2035 (1978). A barrier of 7.02 kcal/mol was subsequently reported (J. F. Sullivan, T. J. Dickson, and J. R. Durig, *Spectrochim. Acta A*, **42**, 113 [1986]), but we have chosen to cite the earlier work, which agrees far better with the MP2/TZP result.^{dd}Barrier for the *trans* transition state (C—S—S—C = 180°).^{ee}Ref. 68.^{ff}Ref. 69.

tially higher in propene than in acetaldehyde, though MMFF94 fails to reproduce this trend. On the other hand, MMFF94 reproduces significant experimental differences in the CH₃—O barriers very nicely. MMFF94 also accurately reproduces both the experimental values and trends for the other listed CH₃—X barriers.

Among other C—C barriers, the *syn* barrier in butane has been the subject of numerous theoretical studies.⁵⁵ Both MMFF94 and the "MP4SDQ/TZP" calculations give a larger value for this barrier than the experimental estimate of 4.54 kcal/mol cited by Allinger and coworkers. The "MP4SDQ/TZP" barrier of 5.79 kcal/mol,

however, is representative of good quality *ab initio* calculations, which apparently tend to overestimate this quantity.⁵⁵ For comparison, the MM3 barrier is 4.8 kcal/mol¹⁷ and the MMFF94 barrier is 5.21 kcal/mol, only modestly higher.

The entries for 1,3-butadiene and acrolein show that partial conjugation across a formally single bond between *sp*²-hybridized carbon atoms leads to a substantial barrier. MMFF94 reproduces the cited experimental barriers very well, though other, sometimes discordant, experimental results are also available.⁵⁶ MM3's barrier for butadiene of 7.31 kcal/mol relative to the trans conformer⁵⁷ is significantly higher than MMFF94's barrier of 6.19 kcal/mol. The difference in this case arises because MM3 was parameterized to reproduce the experimental barrier of 7.16 kcal/mol deduced by Carrieri from an analysis of the Raman overtone spectrum.⁵⁸ We note, however, that Carrieri's analysis assumed a torsional potential that took the second butadiene conformer to be the cis conformer (C=C—C=C torsion angle of zero). In contrast, a variety of good quality *ab initio* calculations have identified this second conformation as being gauche.^{3,59} Bock and coworkers,⁶⁰ in particular, have questioned Carrieri's analysis and have used similar experimental data in conjunction with *ab initio* calculations to assign a gauche conformation and to determine the barrier of 5.96 kcal/mol cited in Table V; the same research group subsequently employed a similar approach to obtain a barrier of 6.05 kcal/mol.⁵⁹ Similarly, Wiberg and Rosenberg, using *ab initio* calculations in conjunction with new experimental data, have concluded that the gauche form better accounts for the very experimental observations on which the early suggestion of a planar cis structure had been based.⁶¹ Furukawa and coworkers,⁶² in deriving a spectroscopic force field, also concluded that the second conformer is probably the gauche. The MMFF94 barrier, which ultimately derives from high-quality *ab initio* calculations, therefore appears to us to be more trustworthy than the MM3 barrier.

Even more importantly, MMFF94 gives a gauche structure for the second conformer of 1,3-butadiene, whereas the MM3(Pi) calculation yields a cis structure.⁵⁷ While the thermally averaged structure may indeed be cis, we believe that the weight of the evidence indicates that the true energy-minimum structure is gauche. A further discussion is given in part IV.³

The next five entries in Table V pertain to rotations about bonds attached to phenyl rings. Here,

the picture is complicated by the fact that large discrepancies arise between theory and experiment—and between experiment and experiment, for that matter. Thus, while experiment, theory, MMFF94, and MM3 all agree that phenol is planar and has a barrier to rotation of the hydroxyl group out of the plane of the ring of ~ 3.5 kcal/mol, the barrier for anisole has been difficult to characterize experimentally. One experimental result cited in a theoretical study by Kollman and coworkers⁶³ places the gas-phase barrier at 3.6 kcal/mol, slightly higher than the barrier for phenol. Another they cite finds a gas-phase barrier of 5.7 kcal/mol, and a third is said to lead to a barrier of "less than 3.1 kcal/mol." Clearly, large experimental errors are at play here. As do Vincent and Hillier in another recent *ab initio* study,⁶⁴ Kollman and coworkers suggest that the neglect of a substantial *V*₄ contribution in the experimental analysis leads to a considerable overestimation of the barrier. The "MP4SDQ/TZP" barrier of 1.91 kcal/mol, in contrast to the cited experimental barriers, is smaller than that for phenol, as might be expected on the basis that the larger methyl group should encounter greater steric repulsion with the ortho hydrogen of the phenyl ring, thereby destabilizing the planar ground state. We believe that the MMFF94 barrier, which closely reproduces this *ab initio* result, is a reasonable one. However, we have omitted anisole from the calculation of the rms deviation because no definitive experimental result can be cited for comparison. The MM3 barrier of 1.1 kcal/mol⁶⁵ appears to be too small.

The situation for acetophenone is equally convoluted, for two sets of seemingly discordant experimental data exist for para-substituted benzaldehydes and acetophenones. One set, typified by the value of 3.15 kcal/mol for acetophenone listed in Table V, uses infrared spectroscopy to measure the barrier, usually in the gas phase.⁶⁶ For acetophenone, this method extrapolates the barrier from a single observed vibrational frequency and an assumed *V*₂ form for the torsional potential. The second set uses dynamic NMR spectroscopy to directly measure the interconversion process, but does so in solvent and gives the barrier as an enthalpy and/or a free energy of activation. For acetophenone, this second approach gives a barrier of 5.35 kcal/mol in CHCl₂F—CHClF₂.⁶⁷ For benzaldehyde, it gives a barrier of about 7.7 kcal/mol in the solvents CHCl₂F—CHClF₂, dimethyl ether, and methylene chloride, as opposed to 4.61

kcal/mol in the gas phase from Durig's infrared method.⁶⁶

It is difficult to see how the very different solution and gas-phase results obtained for benzaldehyde and for acetophenone can be reconciled. That they cannot simply be attributed to differences in solvent polarity is implied by the fact that NMR barriers for the para-substituted benzaldehydes and acetophenones do not change in any substantive or consistent way on going from toluene to, say, methylene chloride or ethyl chloride as solvent.⁶⁷ For acetophenone, moreover, we find that the calculated MP2/6-31G* dipole moment of 3.48 Debye units for the planar equilibrium form decreases only very slightly, to 3.41 Debye units, in the transition-state structure. Consequently, simple theoretical considerations would suggest that solvent polarity should affect the barrier to little if any degree. Nevertheless, Durig et al. find that the infrared method gives a higher barrier of 6.68 kcal/mol for benzaldehyde in the neat liquid (vs. 4.61 kcal/mol in the gas phase), reflecting the "large shift of the torsional frequency" in the condensed phase.⁶⁶ We are not aware, however, that unimolecular rate processes for neutral systems typically slow markedly on going from the gas phase to nonpolar solvent. The *ab initio* calculations support the larger, NMR-derived barrier. We have used this barrier in computing the rms deviation for MMFF94.

For dimethyl disulfide, the MMFF94 and MP2/TZP calculations give a S—S barrier in excess of 8 kcal/mol. These barriers are in qualitative agreement with the experimental value of 6.8 kcal/mol.⁶⁸ The latter, however, is an estimate that, taken together with an estimate of 1.5 kcal/mol for the barrier to methyl-group rotation and assumed values for two other parameters, accounts for the vapor-phase entropy and heat capacity over a range of temperatures. It is difficult to know what error limits should be placed on this experimental barrier. We suspect that the *ab initio* barrier is more trustworthy, and have excluded this system in determining the rms deviation for MMFF94. The much lower MM3 barrier of 5.53 kcal/mol derives from fitting to an *ab initio* barrier of 5.55 kcal/mol obtained from relatively low-level HF/3-21G* calculations.⁶⁹

In part because Allinger and coworkers, in a few instances, used a significantly different experimental value for the rotational barrier in parameterizing MM3, we have not cited an overall rms deviation for MM3. A survey of the results listed

in Table V, however, suggests that its rms deviation would be similar to that for MMFF94. Comparisons based on published data can also be made for UFF⁵⁴ and for CHARMM.³¹ These comparisons, though limited, show that UFF and CHARMM also perform reasonably well. Comparisons of rotational barriers for alkanes can also be made in a few cases to results obtained using the recently described CFF93 force field of Hagler and coworkers.⁷⁰ The "quantum-mechanical force field," the predecessor to CFF93,^{70b} is derived from HF/6-31G* data that, as we have seen, reproduce experimental conformational energies much less well than does the "MP4SDQ/TZP" model used in deriving MMFF94. The published results show that CFF93 gives methyl rotation barriers of 2.75 kcal/mol for ethane, of 3.11 kcal/mol for propane, and of 3.95 kcal/mol for neopentane, and that it gives barriers of 3.47 and 4.95 kcal/mol for the anti→gauche and syn (gauche→gauche) conversions in butane.^{70a} These CFF93 barriers are comparable to those obtained using MMFF94 and MM3. One conformational-energy difference can also be compared: CFF93 places the butane gauche conformer 0.72 kcal/mol higher than the anti, as compared to 0.78 and 0.81 kcal/mol for MMFF94 and MM3, respectively (Table II).

Gundertofte et al. also briefly considered rotational barriers in the work previously cited.^{29a} They found MMFF93, MMFF94's precursor, to be the best of the 15 force-field implementations they examined.

TORSION ANGLES

The comparisons shown below were obtained by optimizing 348 equilibrium conformers for MMFF94 and 315 conformers for MM2X, using MP2/6-31G* geometries as input structures in each case. The comparison encompassed a total of 7974 torsion angles for MMFF94 and 7409 for MM2X. The results, cited below, are stated as rms deviations in degrees from the reference MP2/6-31G* torsion angles:

	MMFF94	MM2X
Torsion angles	5.83	11.38

Clearly, MMFF94 performs much better than MM2X.

One area of relative difficulty for MMFF94 concerns torsion angles in amides, some of which are given relatively poorly because the degree of pyra-

midalization at nitrogen found in the reference MP2/6-31G* structures was either underestimated or overestimated by MMFF94. A number of cases also arose in which the MMFF94 and MP2/6-31G* torsion angles differ appreciably, but in which the MMFF94 structure is not necessarily in error. Such cases particularly involve the orientations of methyl groups attached to oxygen in esters, to nitrogen in amides, and to carbonyl carbon in certain carbonyl compounds. These cases are ones in which the relative energies obtained on the MP2/TZP surface for the torsionally incremented structures (used in MMFF94's derivation to define the equilibrium torsion angles) suggest that the reference MP2/6-31G* structures are not themselves equilibrium conformers on the higher level MP2/TZP surface. When we exclude 24 torsion angles for documented cases of this type, the rms deviation for the remaining comparisons falls to 5.42°. This reduction probably underestimates the true effect of conformational differences on the two surfaces, however, as we have far from complete data for the higher level surface. To obtain an upper bound for the possible effect, we excluded all torsion angles involving methyl groups attached to carbonyl carbon in amides, esters, and unsaturated ketones; to nitrogen in amides; and to oxygen in esters. With these torsion angles excluded, the rms deviation for 7464 remaining torsion angles fell to 4.23°. The best estimate of the rms deviation that would be obtained with respect to conformers optimized a high-level *ab initio* surface is probably about 5°.

Two Aspects of MMFF's Functional Form and Parameterization

SCALING OF 1,4-ELECTROSTATIC INTERACTIONS

Developers of molecular force fields inevitably make many choices about what terms to include and what functional forms to employ. For its part,

MMFF94 uses a novel "buffered" form for the van der Waals (vdW) potential and modifies the coulombic form by introducing an analogous buffering constant into the denominator.² As usual, MMFF94 excludes 1,2- and 1,3-interactions—or, equivalently, uses a scale factor of zero. But, while MM2⁶ and MM3⁷ use a scale factor of one for 1,4-nonbonded interactions, other force fields choose intermediate values.^{31,71,72} As 1,4-interactions typically involve a side-on contact between a pair of vicinally related bonds rather than the more or less head-on interaction usually encountered in other cases, they might well need to be treated differently.

The need to scale 1,4-nonbonded interactions came to light midway through the development of MMFF94 when we added a larger number of highly polar, multifunctional compounds such as hydroxycarboxylic acids and esters and dicarboxylic acids to the parameterization set. We obtained larger than expected errors in some of the associated conformational energies with then-implicit scale factors of 1 for 1,4-interactions, but found that smaller values for the electrostatic scale factor reduced the largest errors. Ultimately, we chose a value of 0.75 for the 1,4-electrostatic scale factor. We also examined the effect on the conformational energies of scaling 1,4-vdW interactions, but found no advantage to using a scale factor different from unity.

CHARGES ON ALIPHATIC HYDROGEN ATOMS

In parameterizing MM2⁶ and MM3,⁷ Allinger and coworkers assigned a zero dipole moment to bonds between hydrogen and *sp*³-hybridized carbon. We have carried this idea over into MM2X and MMFF94 by assigning a zero partial charge to such hydrogens. Other force fields^{70,73-75} and other charge-assignment procedures^{2,76} usually assign positive charges of 0.05–0.15 to aliphatic hydrogens, though small negative charges are some-

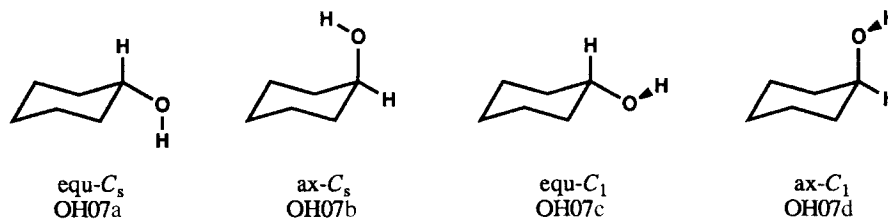


FIGURE 2. Equatorial and axial conformers for cyclohexanol (OH07a–d).

times encountered.^{2,77} Here we consider the basis for the choice made in MMFF94 and describe its effect on conformational energies.

While doing preliminary development on an early precursor to MMFF94, we found that certain conformational energies became very difficult to fit when appreciable positive charges were assigned for aliphatic hydrogens. One case of particular interest concerns cyclohexanol, in which two equatorial and two axial conformers arise (Fig. 2). Here, the "MP4SDQ/TZP" calculations, like a variety of earlier *ab initio* calculations we had carried out, found the equatorial- C_1 conformer (OH07c) to be lowest in energy. These calculations further indicated that the equatorial- C_s (OH07a) and axial- C_1 (OH07d) conformers lie a few tenths of a kilocalorie per mole higher, and found that the axial- C_s conformer (OH07b) lies more than 1 kcal/mol above the equatorial- C_1 conformer (cf. Table VI). Conformers OH07c and OH07d would thus appear to give rise to the experimentally observed preference of ~ 0.5 kcal/mol for equatorial cyclohexanol (cf. Table II). The problem that arises particularly concerns the placement of OH07b, the relatively high-lying axial conformer of C_s symmetry. In this conformation, the hydroxyl hydrogen folds back over the ring and interacts favorably with nearby methylene carbons. These carbons bear net negative charges when the aliphatic hydrogens are positively charged. Positive aliphatic hydrogens therefore selectively stabilize the axial- C_s form, and can even bring it *below* the axial- C_1 conformer. For example, MMFF94 places the axial- C_s form 1.44 kcal/mol below the axial- C_1 conformer when the aliphatic

hydrogens are assigned positive charges of 0.10 (and the carbon charges are correspondingly adjusted) but no other changes are made to the force field. That the axial- C_s conformer in fact lies higher is strongly suggested both by the *ab initio* calculations and by the experimental observation that a similar axial, equatorial separation is found in methoxycyclohexane (cf. Table II), where the axial- C_s conformer is untenable on steric grounds.

To determine the affect on the performance of the total force field of assigning nonzero charges to aliphatic hydrogens, we derived new sets of torsion parameters by refitting the full set of "MP4SDQ/TZP" and MP2/TZP conformational and torsion-profile energies using a series of modified C—H bond charge increments that conferred aliphatic-hydrogen charges of -0.05 , 0.05 , and 0.10 , respectively.⁷⁸ Some pertinent results are displayed in Table VI. These results show that MMFF cannot effectively adjust torsional contributions to offset a large electrostatic stabilization for the axial- C_s form of cyclohexanol. This finding can be understood in terms of the numbers of gauche and anti CCCO or CCOH interactions present in the four conformers. The relationships between the numbers of these interactions keep any change made to the relative energies of gauche and anti CCCO and CCOH torsions from raising the energy of the axial- C_s conformer to the extent needed without damaging other conformational comparisons. Left open is the question of whether changes to parameters that affect other terms, or contributions from additional terms included in other force fields, could produce the necessary offset. For MMFF, the rms differences cited in Table VI show

TABLE VI.
"MP4SDQ / TZP" and MMFF Conformational Energies for Cyclohexanol as a Function of the Charge Assigned to Aliphatic Hydrogen Atoms

Conformer	"MP4" ^a	Relative energy (kcal / mol)				Torsion interactions	
		Charge on aliphatic hydrogen				CCCO	CCOH
		-0.05	0.00	0.05	0.10		
OH07b (ax- C_s)	1.14	1.42	1.01	0.67	0.30	g,g	g,g
OH07d (ax- C_1)	0.33	0.00	0.32	0.62	0.90	g,g	g,a
OH07a (eq- C_s)	0.18	0.00	0.20	0.40	0.62	a,a	g,g
OH07c (eq- C_1)	0.00	0.00	0.00	0.00	0.00	a,a	g,a
rms dev., OH07 ^b		0.27	0.08	0.34	0.64		
rms dev., TORFIT ^c		0.27	0.27	0.28	0.32		

^a"MP4SDQ / TZP" relative energy.

^bRoot mean square (rms) deviation in fitting the listed relative energies for OH07a, OH07b, and OH07d.

^cWeighted rms deviation [cf. eq. (2)] in fitting all 249 conformational and 1171 torsion-profile energies using TORFIT (kcal / mol).

that a charge of zero for the aliphatic hydrogens is nearly optimal for cyclohexanol. We should note that other conformational comparisons are also damaged by positively charged aliphatic hydrogens. For example, the difference between the axial- C_1 and equatorial- C_1 conformers for methoxycyclohexane is given roughly correctly by MMFF94 when aliphatic hydrogens are uncharged. When these hydrogens carry positive charges of 0.10, however, the difference rises to about 1.6 kcal/mol for the rederived force field, considerably higher than the experimental value of 0.45 kcal/mol cited in Table II.

Even the use of zero charges on aliphatic hydrogens is insufficient to ensure that the conformational comparisons for cyclohexanol (and other systems) will come out well. In particular, we found that we could not fit the OH07 and OH06 (cyclopentanol) conformation energies unless MMFF assigned significant steric repulsion to congested interactions involving the hydroxyl hydrogen. This consideration, among others, led us to change the form of MMFF while this work was underway. At one point, we had assigned smaller radii to polar hydrogens and had used the same combination rules for all well depths, ϵ_{ij} , and minimum-energy separations, R_{ij}^* . As detailed elsewhere,² however, MMFF94 now assigns larger radii to polar hydrogens but reduces ϵ_{ij} and R_{ij}^* when such hydrogens interact with hydrogen-bond acceptors. Though perhaps unesthetic, this hybrid approach has the virtue of allowing MMFF94 to describe hydrogen-bonding interactions well while retaining steric repulsions of the size needed to fit otherwise troublesome conformational energies.

Limitations in and Possible Improvements to MMFF94's Functional Form

As noted previously, 18 of the 249 conformational comparisons listed in Table I deviate from the reference "MP4SDQ/TZP" values by 0.5–1 kcal/mol, while 4 others deviate by 1–2 kcal/mol. That many of these instances occur in compounds in which two or more strongly polar groups exist in close proximity suggests that the difficulty may stem from deficiencies in MMFF94's electrostatic representation. In this section we discuss these and other aberrant cases and suggest possible enhancements in functional form.

Particularly problematic are the CA04-, CA05-, CA08-, and CA09-series comparisons for glyoxalic,

glycolic, oxalic, and pyruvic acid, and the CE10-series comparisons for methyl glycolate, most of which involve structures having adjacent $C=O$ groups. The largest error in this set occurs for the CA04a–CA04b comparison of glyoxalic acid conformers cited in Table I. The same comparison is included in Table II, where the experimental and "MP4SDQ/TZP" conformational energies can also be seen to differ widely from one another.

Even larger errors of 2–3 kcal/mol occur for ethylenediamine, another polar multifunctional compound. While the anti, gauche conformational-energy difference (NH18a – NH18b in Table I) is itself given accurately by MMFF94, the torsionally incremented MP2/TZP comparisons involving $\pm 30^\circ$ and $\pm 60^\circ$ rotations about a C–N bond (results not shown) are not. For MMFF94, electrostatic interactions produce large contributions to the conformational energy that differ substantially in magnitude and pattern from the conformational profile reflected in the MP2/TZP energies. When the associated MP2/TZP conformational energies were included in the least-squares fit, some of the derived torsion parameters became so badly compromised that the force field could no longer adequately describe conformational comparisons for simple amines. In the latter stages of this work, as noted earlier, we therefore removed the 17 MP2/TZP comparisons for ethylenediamine from the least-squares fit by assigning essentially zero weight factors, w_i , in eq. (2). In so doing, we accepted the implication that MMFF94's functional form is not capable of coping with these conformational comparisons.

We suspect that the problem with ethylenediamine may largely arise from MMFF94's neglect of polarizability. Its use of simple atom-centered charges may also be a factor but does not appear to be primarily responsible. To explore the latter possibility, we created an off-center charge model by physically displacing the nitrogen charge center in each C–N bond rotamer by 0.1 Å along the pseudo-threefold axis formed by the attached hydrogen and carbon atoms. This displacement, which we chose to be similar in magnitude and direction to that made in the TIP4P model⁷⁹ for water, makes the *effective* N–N distance a function of the C–N rotation angle, and thus changes the electrostatic contribution to the C–N rotation profile. Unfortunately, however, we found that this modification improved some of the problematic relative energies for ethylenediamine but damaged others.

Another difficult case concerns the equilibrium orientation of the methyl groups in *N*-methylformaldehydeimine (IM02a) and in the anti conformer of the structurally similar *N*-methylformamidine (IM04b). As indicated in Figure 3, one of the methyl C—H bonds eclipses the N=C double bond in the former, much as a C—H bond eclipses the C=C double bond in propene, while two C—H bonds stagger the N=C double bond in the latter. We were unable to find any way to satisfactorily account for this change of $\sim 60^\circ$ in orientational preference. As it happens, the MMFF94 parameters favor the correct methyl orientations but describe an essentially free methyl-group rotation for *N*-methylformaldehydeimine and, consequently, underestimate the barrier for the IM02t-IM02a comparison by ~ 2 kcal/mol. This change in preference might possibly reflect an enhanced hyperconjugative interaction in the formamidine. If so, it is difficult to see how this effect could be accommodated in a force-field model other than by assigning different atom types to the methylated nitrogens in the two systems, despite the fact that their environments appear to be very similar.

A last notable case concerns C—N bond rotations in resonance delocalized guanidinium cations and in similar species such as amidinium cations. A 90° out-of-plane rotation of a NH_2 group in the guanidinium cation, for example, removes the partial pi-bonding for that C—N bond and confines the charged pi-system to the carbon and the remaining two NH_2 groups. As a result, the C—N bond that is made to the orthogonal NH_2 group should lengthen and the charge distribution should change markedly. A force field that includes bond-torsion cross terms, as do MM3 and CFF93, can easily reproduce the bond lengthening. It is quite another matter, however, to properly model the change in the charge distribution. MMFF94,

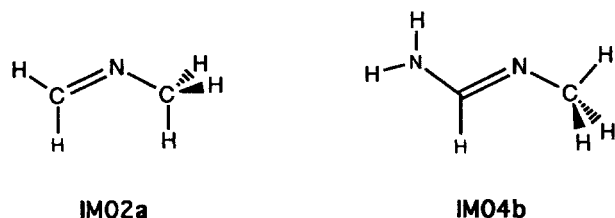


FIGURE 3. Structures for *N*-methylformaldehydeimine (IM02a) and for the anti conformation of *N*-methylformamidine (IM04b). Note the different orientation of the methyl group hydrogens relative to the N=C double bond.

like other contemporary force fields, incorrectly makes the three nitrogens in guanidinium equivalent, independently of conformation. This equivalence, though appropriate for the planar system, may contribute to a relatively poor description of the conformational energetics for C—N bond rotation.

We suspect that a more complex description of the molecular electrostatic potential than that used in MMFF94 and other contemporary force fields will be needed to accurately describe polar, multifunctional compounds. We suggest that the cyclohexanol and methoxycyclohexane conformers discussed earlier and the systems discussed in this section constitute a demanding set of test systems that any improved force-field model should be asked to address.

MMFF94 Torsion Parameters

In this section we describe a small subset of the MMFF94 torsion parameters. The full set is contained in supplementary material for Part I¹ and is accessible in computer-readable form.⁸⁰ The parameters listed in Table VII arise from various combinations of the following numerical atom types: 1, 2, 3, 5, 7, 10, and 28. These atom types describe saturated carbon, olefinic carbon, carbonyl carbon, hydrogen attached to carbon, carbonyl oxygen, amide nitrogen, and amide hydrogen, respectively. They cover simple alkanes, alkenes, aldehydes, ketones, and amides. Within this limited set of atom types, the tabulated parameters are complete.

The format of the parameter file illustrated in Table VII is as follows. The first column gives the value of the torsion type index, TT_{ijkl} . This index normally takes the value "0," but is "1" when the j - k bond has a bond type index³ BT_{jk} of 1; is "2" when BT_{jk} is "0" but BT_{il} and/or BT_{kl} is "1"; is "4" when i , j , k , and l are all members of the same four-membered ring; and is "5" when the four atoms are members of a five-membered ring and at least one is a sp^3 -hybridized carbon (MMFF atom type 1). The torsion parameters are ordered using the canonical index

$$\text{CXT} = \text{MC} * (J * \text{MA}^3 + K * \text{MA}^2 + I * \text{MA} + L) + \text{TT}_{ijkl}$$

where MC is at least the maximum permitted torsion-type index plus 1 and MA is at least the maximum permitted MMFF atom type plus 1.⁸¹

TABLE VII.
Partial List of MMFF94 Torsion Parameters [Eq. (1)]

TT_{IJKL}	I	J	K	L	V_1	V_2	V_3	Origin / Comment
0	0	1	1	0	0.000	0.000	0.300	C94 0:*-1-1-* Def
5	0	1	1	0	0.200	-0.800	1.500	C94 5:*-1-1-* Def
0	1	1	1	1	0.103	0.681	0.332	C94
5	1	1	1	1	0.144	-0.547	1.126	C94
0	1	1	1	2	-0.295	0.438	0.584	C94
0	1	1	1	3	0.066	-0.156	0.143	C94
0	1	1	1	5	0.639	-0.630	0.264	C94
0	2	1	1	5	0.321	-0.411	0.144	C94
0	3	1	1	3	0.443	0.000	-1.140	C94
0	3	1	1	5	-0.256	0.058	0.000	C94
0	5	1	1	5	0.284	-1.386	0.314	C94
0	5	1	1	10	0.000	0.000	0.427	C94
0	0	1	2	0	0.000	0.000	0.000	C94 0:*-1-2-* Def
2	0	1	2	0	0.000	0.000	0.000	E94 2:*-1-2-* Def
5	0	1	2	0	0.000	0.000	0.000	C94 5:*-1-2-* Def
0	0	1	2	2	0.000	0.000	-0.650	C94 0:*-1-2 = 2 Def
5	0	1	2	2	0.000	0.000	-0.650	C94 5:*-1-2 = 2 Def
0	1	1	2	1	0.419	0.296	0.282	C94
0	1	1	2	2	-0.494	0.274	-0.630	C94
0	1	1	2	5	0.075	0.000	0.358	C94
0	2	1	2	2	-0.293	0.115	-0.508	C94
0	2	1	2	5	0.301	0.104	0.507	C94
0	3	1	2	1	0.565	-0.554	0.234	C94
0	3	1	2	2	-0.577	-0.482	-0.427	C94
0	3	1	2	5	0.082	0.000	0.123	C94
0	5	1	2	1	0.000	-0.184	0.220	C94
0	5	1	2	2	0.501	-0.410	-0.535	C94
2	5	1	2	2	0.000	0.000	0.055	C94
2	5	1	2	3	0.000	0.000	-0.108	C94
0	5	1	2	5	-0.523	-0.228	0.208	C94
0	0	1	3	0	0.000	0.400	0.300	C94 0:*-1-3-* Def
2	0	1	3	0	0.000	0.500	0.350	C94 2:*-1-3-* Def
5	0	1	3	0	0.000	0.000	0.000	E94 5:*-1-3-* Def
0	0	1	3	1	0.000	0.000	0.550	C94 0:*-1-3-1 Def
0	0	1	3	5	0.000	0.200	0.700	C94 0:*-1-3-5 Def
0	0	1	3	7	0.000	0.400	0.400	C94 0:*-1-3-7 Def
0	1	1	3	1	0.103	0.177	0.545	C94
0	1	1	3	5	-0.072	0.316	0.674	C94
0	1	1	3	7	0.825	0.139	0.325	C94
0	1	1	3	10	-0.927	1.112	1.388	C94
0	2	1	3	5	0.663	-0.167	0.426	C94
0	2	1	3	7	-0.758	0.112	0.563	C94
0	5	1	3	1	-0.073	0.085	0.531	C94
2	5	1	3	2	0.000	0.000	0.115	C94
2	5	1	3	3	0.000	0.000	0.446	C94
0	5	1	3	5	-0.822	0.501	1.008	C94
0	5	1	3	7	0.659	-1.407	0.308	C94
0	5	1	3	10	-0.412	0.693	0.087	C94
0	10	1	3	7	0.338	2.772	2.145	C94
0	10	1	3	10	0.548	0.000	1.795	C94
0	0	1	10	0	0.000	0.000	0.300	C94 0:*-1-10* Def
5	0	1	10	0	0.000	0.000	0.000	E94 5:*-1-10-* Def
0	0	1	10	3	0.000	0.000	1.000	C94 0:*-1-10-3 Def
0	1	1	10	3	-1.027	0.694	0.948	C94
0	1	1	10	28	0.552	-0.380	0.326	C94
0	3	1	10	3	3.100	-2.529	1.494	C94

(Continues on next page)

TABLE VII.
(continued)

TT_{IJKL}	I	J	K	L	V_1	V_2	V_3	Origin / Comment
0	3	1	10	28	0.079	0.280	0.402	C94
0	5	1	10	1	0.000	0.000	0.779	C94
0	5	1	10	3	-2.099	1.363	0.021	C94
0	5	1	10	28	-0.616	0.000	0.274	C94
0	0	2	2	0	0.000	12.000	0.000	C94 0:*2 = 2-* Def
1	0	2	2	0	0.000	1.800	0.000	C94 1:* = 2-2 = * Def
5	0	2	2	0	0.000	12.000	0.000	C94 5:*2 = 2-* Def
0	1	2	2	1	-0.403	12.000	0.000	C94
0	1	2	2	2	0.000	12.000	0.000	C94
1	1	2	2	2	-0.418	2.089	-0.310	C94
0	1	2	2	5	0.000	12.000	0.000	C94
1	1	2	2	5	0.412	2.120	0.269	C94
1	2	2	2	2	0.094	1.621	0.877	C94
0	2	2	2	5	0.000	12.000	0.000	C94
1	2	2	2	5	0.317	1.421	-0.870	C94
0	3	2	2	5	0.000	12.000	0.000	C94
0	5	2	2	5	0.000	12.000	0.000	C94
1	5	2	2	5	-0.406	1.767	0.000	C94
1	0	2	3	0	0.000	1.800	0.000	C94 0:*2-3-* Def
1	1	2	3	1	0.136	1.798	0.630	C94
1	1	2	3	5	0.497	2.405	0.357	C94
1	1	2	3	7	-0.401	2.028	-0.318	C94
1	1	2	3	10	-0.084	2.214	-0.610	C94
1	2	2	3	1	-0.325	1.553	-0.487	C94
1	2	2	3	5	-0.295	2.024	-0.590	C94
1	2	2	3	7	0.362	1.978	0.000	C94
1	2	2	3	10	0.095	1.583	0.380	C94
1	5	2	3	1	0.213	1.728	-0.042	C94
1	5	2	3	5	-0.208	1.622	0.223	C94
1	5	2	3	7	0.000	2.046	0.000	C94
1	5	2	3	10	0.000	1.395	0.227	C94
0	0	2	10	0	0.000	6.000	0.000	E94 0:*2-10-* Def
2	0	2	10	0	0.000	6.000	0.000	E94 2:*2-10-* Def
5	0	2	10	0	0.000	6.000	0.000	E94 5:*2-10-* Def
1	0	3	3	0	0.000	0.600	0.000	C94 0:*3-3-* Def
4	0	3	3	0	0.000	1.800	0.000	E94 4:*3-3-* Def
1	1	3	3	1	-0.486	0.714	0.000	C94
1	1	3	3	7	1.053	1.327	0.000	C94
1	5	3	3	7	0.000	0.177	-0.412	C94
1	7	3	3	7	-0.260	1.084	0.193	C94
0	0	3	10	0	0.000	6.000	0.000	C94 0:*3-10-* Def
2	0	3	10	0	0.000	6.000	0.000	C94 2:*3-10-* Def
4	0	3	10	0	0.000	6.000	0.000	C94 4:*3-10-* Def
5	0	3	10	0	0.000	6.000	0.000	E94 5:*3-10-* Def
0	1	3	10	1	0.647	6.159	0.507	C94
0	1	3	10	28	-0.294	5.805	1.342	C94
2	2	3	10	28	-0.287	7.142	0.120	C94
0	5	3	10	1	-0.183	6.314	1.753	C94
0	5	3	10	3	-0.751	5.348	0.209	C94
0	5	3	10	28	-0.388	5.972	0.459	C94
0	7	3	10	1	-0.319	6.294	-0.147	C94
0	7	3	10	3	0.776	-0.585	-0.145	C94
0	7	3	10	28	1.435	4.975	-0.454	C94
0	10	3	10	28	0.000	3.495	1.291	C94
0	0	10	10	0	0.000	0.000	0.000	E94 0:*10-10-* Def
5	0	10	10	0	0.000	0.000	0.000	E94 5:*10-10-* Def

Thus, J changes least rapidly and K next least rapidly. For I , for L , and for TT_{IJKL} , which changes most rapidly, the effect on the ordering can be seen in Table VII.

Requests for one-, two- and threefold torsion parameters that cannot be resolved using the primary numeric MMFF atom types and the associated torsion-type index TT_{IJKL} are first handled via the step-down equivalencing procedure outlined in part 1.¹ The parameter search may terminate with a request for a fully wildcarded $M:0-J-K-0$ parameter (where M denotes the torsion-type index TT_{IJKL}) or, in a few cases, may locate a "half-wildcard" $M:0-J-K-L$ or $M:I-J-K-0$ parameter. If necessary, a carefully calibrated empirical rule is invoked.⁴

Concluding Discussion

This article completes the derivation and documentation of the core, computationally derived, portion of the new Merck Molecular Force Field (MMFF94). It describes the parameterization of torsional interactions and conformational energies using quantum-mechanically calculated data. The high-level data afforded 249 comparisons of "MP4SDQ/TZP" relative conformational energies calculated at optimized MP2/6-31G* geometries, and yielded 1192 comparisons of MP2/TZP relative energies for torsionally incremented structures that probed the energetic consequences of rotations about specific chemical bonds. The fitting of the MMFF94 torsion parameters to this computational data employed TORFIT, a powerful program developed in these laboratories that enabled all the computed conformational energies to be considered simultaneously. Importantly, the fitting procedure was an adaptive one that allowed intermediate results to be used to define the force-field model and to direct the generation and use of additional computational data.

MMFF94 performs quite well by current standards. It reproduces the "MP4SDQ/TZP" conformational energies and the MP2/TZP comparisons of torsionally incremented structures with rms deviations of 0.31 and 0.50 kcal/mol, respectively, as against rms values of 3.88 and 4.37 kcal/mol for the *ab initio* relative energies. MM2X performs significantly more poorly. Comparisons for a diverse set of 37 experimentally determined conformational energies, enthalpies, and free energies and for 28 rotational barriers further demonstrate

that MMFF94, even though parameterized solely against computational data, predicts conformational energies and rotational barriers as well as MM3. Root mean square deviations are roughly 0.4 kcal/mol for MMFF94 in each case, as opposed to rms values of 2.3 and 3.7 kcal/mol, respectively, for the experimental determinations. Torsion angles are given fairly well by MMFF94; rms errors relative to MP2/6-31G*-optimized geometries are on the order of 5°.

Both the force-field model and the methodology used in MMFF94's derivation could be further improved in a number of ways. Ideally, the data on molecular geometries and conformational energies should come from a single computational model, but this was not feasible in the present work. Some significant limitations in the force-field model also need to be addressed. In particular, the relatively greater difficulty encountered for molecules containing two or more highly polar moieties in close proximity may stem from the overly simplistic electrostatic representation used in MMFF94 and other present-day force fields. While a number of pioneering studies have been carried out,⁸² polarizability needs to be included on a consistent basis in future force fields, and a more complex and accurate electrostatic representation than is obtainable when fixed, atom-centered charges are used should also be employed, perhaps by including atom-centered dipoles.⁸³ A review citing this and other possible extensions has recently appeared.⁸⁴ A variety of demanding test systems that could prove useful in developing or validating future, physically superior force fields were discussed in a previous section. Another modification that might prove beneficial would be to reformulate the torsional potential in terms of local symmetry-adapted linear combinations of torsion interactions. This modification, for example, would decouple torsional from out-of-plane motion; for a system like ethane, it would yield a single composite torsional coordinate, in conformity with spectroscopic formulations.

We view the determination of torsion parameters as a particularly strong component in the development of MMFF94. One element that serves to define the special nature of MMFF94's derivation is its simultaneous use of all available conformational-energy data in the least-squares fits. The contrasting "functional group" build-up approach has been used in most other force-field derivations. While the practical limitations of this approach are not known, we prefer an approach

that, by construction, allows the parameters to attain mutually consistent values.

A second aspect of MMFF94's parameterization confers an especially important advantage. In particular, its derivation from *ab initio* data ensures that MMFF94 will predict conformational energies and barriers equally well throughout the range of its core parameterization. In contrast, the data on conformational energies required to reliably develop an experimentally derived force field like MM2 or MM3 are limited or lacking entirely for much—indeed, most—of the range of organic structure that needs to be described accurately by a general organic/bio-organic force field such as MMFF94. Moreover, *ab initio* errors tend to be relatively modest in size and predictable in nature. As was evident in several examples discussed previously, in contrast, errors in experimental determinations of relative conformational energies and barriers can be large and unpredictable.

A third element concerns the relatively high level of theory used to obtain the computational data employed in MMFF94's torsional parameterization. The CFF93 approach, in contrast, uses relatively simple HF/6-31G* calculations.^{70b} Our results suggest that this difference in methodology is likely to have a significant effect on the ability of the derived force field to account for conformational energies. Thus, HF/6-31G* calculations predict experimentally observed relative conformational energies much more poorly (rms deviation 0.73 kcal/mol) than do the "MP4SDQ/TZP" calculations used in this work (rms deviation 0.37 kcal/mol). It seems clear that a force field derived solely on the basis of HF/6-31G* data could not predict experimental conformational energies more accurately than the HF/6-31G* calculations themselves. To achieve a level of accuracy comparable to that attained by MMFF94 (rms deviation 0.40 kcal/mol), such a force field would also need to be calibrated against experiment in some manner. But while such a calibration is feasible in some cases,^{70a} it cannot in general be carried out, because too little useful experimental data are available.

In defining the range of potential applications for MMFF94, particularly notable is a series of conformational comparisons based on "MP4SDQ/TZP" energies at MP2/6-31G* geometries for the glycine and alanine dipeptides and dipeptide analogs. When considered together with the careful attention also given to achieving a balanced parameterization for intermolecular interactions,² the good fit obtained to this high-quality data, in our view, justifies designating

MMFF94 as a "protein" force field as well as an "organic" force field. We believe that it will perform well in both domains.

Supplementary Material

Computer-readable files of torsion and other MMFF94 parameters⁸⁰ and Appendix A (detailed discussion of the derivation of the MMFF94 torsion parameters)²⁴ 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 III: T. A. Halgren, *J. Comput. Chem.* (this issue).
4. Part V: T. A. Halgren, *J. Comput. Chem.* (this issue).
5. 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).
6. (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).
7. N. L. Allinger, Y. H. Yuh, and J.-H. Lii, *J. Am. Chem. Soc.*, **111**, 8551–8566 (1989).
8. 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.
9. 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.
10. 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.
11. R. Krishnan, M. J. Frisch, and J. A. Pople, *J. Chem. Phys.*, **72**, 4244–4245 (1980); R. Krishnan and J. A. Pople, *Int. J. Quantum Chem.*, **14**, 91 (1978).
12. W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986, Chapter 4. The 6-31G* and 6-31 + G** basis sets are also known as 6-31G(d) and 6-31 + G(p, d), respectively.

13. T. H. Dunning Jr., *J. Chem. Phys.*, **55**, 716–723 (1971).
14. A. D. McLean and G. S. Chandler, *J. Chem. Phys.*, **72**, 5639–5648 (1980).
15. T. H. Dunning Jr., *J. Chem. Phys.*, **90**, 1007–1023 (1989).
16. B. Roos and P. Siegbahn, *Theoret. Chim. Acta*, **17**, 199–208 (1970).
17. M. Frisch and J. A. Pople, *J. Chem. Phys.*, **80**, 3265–3269 (1984); R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, *J. Chem. Phys.*, **72**, 650–654 (1980).
18. These penalty-function restraints used a force constant of 50 kcal/mol/deg².
19. A. St-Amant, W. D. Cornell, T. A. Halgren, and P. A. Kollman, *J. Comput. Chem.*, **16**, 1483–1506 (1995).
20. J. P. Bowen, A. Pathiaseril, S. Profeta, Jr., and N. L. Allinger, *J. Org. Chem.*, **52**, 5162–5166 (1987).
21. The QCISD(T) (quadratic configuration interaction) method as implemented in Gaussian 92 (ref. 10) includes the fourth-order triples contribution to the energy; these contributions are computed after the QCISD energy has converged.
22. T. H. Dunning Jr. and P. J. Hay, *Modern Theoretical Chemistry*, Plenum, New York, 1976, Chapter 1.
23. This work was done on the Merck Research Laboratories' Cray YMP8i 4/128.
24. Appendix A is available in Supplementary Material (see footnote * on the first page of this article).
25. As discussed in the "MMFF Torsion Parameters" section, each torsion parameter actually carries a fifth numerical index corresponding to the "torsion type." For simplicity, this dependence is suppressed in the present context.
26. Eq. (3) may be more readily understood in the form $CE_i^{MMFF} + \Delta T_i = CE_i^{ref}$, where $\Delta T_i = T_i - T_i^0$ represents the change needed in the MMFF torsional energy contribution.
27. For MM2X, we omitted eight comparisons for structures near the "syn" and "anti" transition states for C–N rotation in *N*-methylformamide used in parameterizing and evaluating MMFF. MM2X gives reasonable energies for some of these structures, but others collapse to structures not representative of the transition state region. Inclusion of these structures would increase the rms deviation and the rms value for "conformational energies" to 2.07 and 3.99 kcal/mol, respectively.
28. T. Head-Gordon, M. Head-Gordon, M. J. Frisch, C. L. Brooks III, and J. A. Pople, *J. Am. Chem. Soc.*, **113**, 5989–5997 (1991).
29. (a) K. Gundertofte, T. Liljefors, P.-O. Norrby, and I. Pettersson, *J. Comput. Chem.* (in press); (b) K. Gundertofte, J. Palm, I. Pettersson, and A. Stamvik, *J. Comput. Chem.*, **11**, 200–208 (1991).
30. C. J. Casewit, K. S. Colwell, and A. K. Rappé, *J. Am. Chem. Soc.*, **114**, 10035–10046 (1992).
31. F. A. Momany and R. Rone, *J. Comput. Chem.*, **13**, 888–900 (1992).
32. S. Ataka, H. Takeuchi, and M. Tasumi, *J. Mol. Struct.*, **113**, 147–160 (1984).
33. T. Drakenberg and S. Forsen, *J. Chem. Soc. Chem. Commun.*, 1404–1405 (1971).
34. J.-H. Lii and N. L. Allinger, *J. Comput. Chem.*, **12**, 186–199 (1991).
35. N. L. Allinger, Z. S. Zhu, and K. Chen, *J. Am. Chem. Soc.*, **114**, 6120–6133 (1992).
36. Averages for each of two disparate sets of experimental values cited by: J. P. Bowen, A. Pathiaseril, S. Profeta Jr., and N. L. Allinger, *J. Org. Chem.*, **52**, 5162–5166 (1987).
37. K. B. Wiberg and E. Martin, *J. Am. Chem. Soc.*, **107**, 5035–5041 (1985).
38. D. A. Dixon and B. E. Smart, *J. Phys. Chem.*, **92**, 2729–2733 (1988).
39. D. A. Dixon, N. Matsuzawa, and S. C. Walker, *J. Phys. Chem.*, **96**, 10740–10746 (1992).
40. J. R. Durig, S. E. Godbey, and J. F. Sullivan, *J. Chem. Phys.*, **80**, 5983–5993 (1984).
41. K. Yamanouchi, M. Sugie, H. Takeo, C. Matsumura, and K. Kuchitsu, *J. Phys. Chem.*, **88**, 2315–2320 (1984).
42. Y. Ogawa, S. Imazeki, H. Yamaguchi, H. Matsuura, I. Harada, and T. Shimanouchi, *Bull. Chem. Soc. Jpn.*, **51**, 748–767 (1978).
43. N. L. Allinger, M. Rahman, and J.-H. Lii, *J. Am. Chem. Soc.*, **112**, 8293–8307 (1990).
44. H. Booth and M. L. Jozefowicz, *J. Chem. Soc. Perkin Trans. 2*, 895–901 (1976).
45. J. R. Durig, W. E. Bucy, C. J. Wurrey, and L. A. Carreira, *J. Phys. Chem.*, **79**, 988–993 (1975).
46. R. K. Kakar and C. R. Quade, *J. Chem. Phys.*, **72**, 4300–4307 (1980).
47. F. A. L. Anet and V. J. Basus, *J. Am. Chem. Soc.*, **95**, 4424–4426 (1973).
48. Evaluated using Ermer's formulation; cf. O. Ermer, *Aspekte von Kraftfeldrechnungen*, Wolfgang Baur Verlag, München, 1981, p. 59. We note that the S_{rot} coefficient of 1.287 following eq. (10b) should be 2.287.
49. J. B. Hendrickson, *J. Am. Chem. Soc.*, **89**, 7036–7043 (1967).
50. For the notation used, see J. Dale, *Top. Stereochem.*, **9**, 199 (1976).
51. F. A. L. Anet and J. Krane, *Isr. J. Chem.*, **20**, 72–83 (1980).
52. N. L. Allinger, K. Chen, and J.-H. Lii, *J. Comput. Chem.* (this issue).
53. We thank K. Chen and N. L. Allinger for alerting us to the possible role of this conformer and for providing a MM4-optimized geometry prior to publication (ref. 52). In recent work, Anet has referred to this species as "TCTC" (cf. F. A. L. Anet, *J. Am. Chem. Soc.*, **112**, 7172 [1990]).
54. 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).
55. N. L. Allinger, R. S. Grev, B. F. Yates, and H. F. Schaefer III, *J. Am. Chem. Soc.*, **112**, 114–118 (1990), and references therein.
56. See K. B. Wiberg, P. R. Rablen, and M. Marquez, *J. Am. Chem. Soc.*, **114**, 8654–8668 (1992).
57. N. L. Allinger, F. Li, and L. Yan, *J. Comput. Chem.*, **11**, 868–895 (1990).
58. L. A. Carriera, *J. Chem. Phys.*, **62**, 3851–3854 (1975).
59. Y. N. Panchenko, A. V. Abramnikov, and C. W. Bock, *J. Mol. Struct.*, **140**, 87–92 (1986).
60. C. W. Bock, P. George, M. Trachtman, and M. Zanger, *J. Chem. Soc. Perkin Trans. II*, 26–34 (1979).
61. K. Wiberg and R. E. Rosenberg, *J. Am. Chem. Soc.*, **112**, 1509 (1990).

62. Y. Furakawa, H. Takenchi, I. Harada, and M. Tasumi, *Bull. Chem. Soc. Jpn.*, **56**, 392 (1983).
63. D. C. Spellmeyer, P. D. J. Grootenhuis, M. D. Miller, L. F. Kuyper, and P. A. Kollman, *J. Phys. Chem.*, **94**, 4483–4491 (1990).
64. M. A. Vincent and I. H. Hillier, *Chem. Phys.*, **140**, 35 (1990). Note: Allinger and co-workers (cf. ref. 65) cite an experimental barrier for anisole of 2.4 kcal/mol and reference the article by Vincent and Hillier. However, the barrier cited was in fact obtained by *ab initio* calculation.
65. N. L. Allinger and L. Yan, *J. Am. Chem. Soc.*, **115**, 11918–11925 (1993).
66. J. R. Durig, H. D. Bist, K. Furic, J. Qui, and T. S. Little, *J. Mol. Spectrosc.*, **129**, 45–56 (1985).
67. T. Drakenberg, J. Sommer, and R. Jost, *J. Chem. Soc., Perkin Trans. II*, 363–369 (1980).
68. W. N. Hubbard, D. R. Douslin, J. P. McCullough, D. W. Scott, S. S. Todd, J. F. Messerly, I. A. Hossenlopp, A. George, and G. Waddington, *J. Am. Chem. Soc.*, **80**, 3547–3554 (1958).
69. K. Chen and N. L. Allinger, *J. Phys. Org. Chem.*, **4**, 659–666 (1991).
70. (a) M.-J. Hwang, T. P. Stockfish, and A. T. Hagler, *J. Am. Chem. Soc.*, **116**, 2515–2525 (1994); (b) J. R. Maple, M.-J. Hwang, T. P. Stockfish, U. Dinur, M. Waldman, C. S. Ewig, and A. T. Hagler, *J. Comput. Chem.*, **15**, 161–182 (1994).
71. S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta Jr., and P. Weiner, *J. Am. Chem. Soc.*, **106**, 765–784 (1984).
72. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.*, **4**, 187–217 (1983).
73. S. J. Weiner, P. A. Kollman, D. T. Nguyen, and D. A. Case, *J. Comput. Chem.*, **7**, 230–252 (1986).
74. P. Dauber-Osguthorpe, V. A. Roberts, D. J. Osguthorpe, J. Wolff, M. Genest, and A. T. Hagler, *Proteins*, **4**, 31–47 (1988).
75. J. C. Smith and M. Karplus, *J. Am. Chem. Soc.*, **114**, 801–812 (1992).
76. A. K. Rappé and W. A. Goddard III, *J. Phys. Chem.*, **95**, 3558–3563 (1991).
77. Y. Sun, D. Spellmeyer, D. A. Pearlman, and P. Kollman, *J. Am. Chem. Soc.*, **114**, 6798–6801 (1992); C. I. Bayly and P. A. Kollman, *J. Am. Chem. Soc.*, **115**, 9620–9631 (1993).
78. These rederivations each employed a single least-squares fit starting from the final MMFF94 parameters. In these fits, some results of which are shown in Table VI, a weak penalty-function restraint of $p = 0.1$ in eq. (2) was used and all except the 39 twofold torsion parameters cited were allowed to vary. Alternative rederivations using $p = 0.01$ led to much larger changes in torsion parameters but did not significantly change the conformational energies listed in the table, suggesting that the listed values are representative of the best attainable for the modified force field. For these fits, as indeed for the primary MMFF94 fit, enhanced weight factors $w_i = 10$ were used in eq. (2) to ensure that the torsion parameters would reproduce the conformational energy differences for cyclohexanol as well as possible.
79. W. L. Jorgensen, J. Chandrasekhar, and J. D. Madura, *J. Chem. Phys.*, **79**, 926–935 (1983).
80. 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.
81. In the current OPTIMOL implementation, these quantities are set to 6 and 136, respectively, to insure that the maximum value computed for CXT will fit within a 32-bit integer word.
82. See ref. 21 in part I¹ for citations to some of this work.
83. U. Dinur and A. T. Hagler, *J. Chem. Phys.*, **91**, 2949–2958 (1989); U. Dinur and A. T. Hagler, *J. Chem. Phys.*, **91**, 2959–2970 (1989).
84. D. E. Williams, in *Reviews in Computational Chemistry*, Vol. 2, K. B. Lipkowitz and D. B. Boyd, Eds., VCH, New York, 1991, pp. 219–271.