MMFF VI. MMFF94s Option for Energy Minimization Studies

THOMAS A. HALGREN

Molecular Systems Department, Merck Research Laboratories, P.O. Box 2000, Rahway, New Jersey 07065

Received 13 May 1998; accepted 16 December 1998

ABSTRACT: This article describes the derivation of MMFF94s, which is the ''s'' (static) variant of MMFF94. MMFF94s incorporates altered out of plane bending parameters that yield planar (or nearly planar) energy-minimized geometries at unstrained delocalized trigonal nitrogen centers. Some experimental and most theoretical structures show appreciable puckering at nitrogen in isolated structures. However, condensed-phase effects or even strong intermolecular hydrogen bonding tends to reduce, but need not eliminate, such puckering; in contravention to results reported on the lower level Hartree–Fock surface, we show in the correlated LMP2/6-31G** calculations for the Watson–Crick guanine–cytosine base pair that one of the hydrogen-bonded NH_2 groups remains appreciably puckered. The resultant MMFF94s geometries emulate the ''time-averaged'' structures typically observed in crystallographic and most other experimental structure determinations. MMFF94s also employs modified torsion parameters for interactions that involve such centers, but is identical to MMFF94 for other systems. Isolated instances are found in which MMFF94s fails to locate a (probably shallow) local minimum found on the MMFF94 and reference MP2/6-31G* surfaces. In general, however, MMFF94s describes conformation energies for delocalized trigonal nitrogen systems nearly as well as MMFF94 does. © 1999 John Wiley & Sons, Inc. J Comput Chem 20: 720–729, 1999

Keywords: force field; base pair; delocalized trigonal nitrogen; puckering

Correspondence to: T. A. Halgren; e-mail: halgren@merck.com

This article includes Supplementary Material available from the author upon request or via the Internet at ftp.wiley.com/public/journals/jcc/suppmat/20/720 or http://journals.wiley.com/jcc/

© 1999 John Wiley & Sons, Inc.
CCC 0192-8651 / 99 / 070720-10

Introduction

This article, the sixth in a series, introduces MMFF94s, the ''static'' (s) variant of MMFF94.[1–5] MMFF94s was developed for use in energy-minimization studies rather than in molecular-dynamics simulations. The two force fields yield identical results for most systems but differ for amides and for unsaturated amines such as vinyl amines, anilines, guanidines, and nucleic-acid bases, all of which contain resonance-delocalized trigonal nitrogen atoms. While some experimental and most theoretical structures for compounds of these types show strongly pyramidalized geometries at nitrogen, most experimental structures show planar or nearly planar geometries that reflect time-averaged atomic positions and the effects of cooperative condensed-phase intermolecular interactions (see below). For such systems, MMFF94 typically gives a nonplanar energy-minimized geometry that emulates the puckered MP2/6-31G*-optimized geometry used in its parameterization.[3] But while MMFF94 produces relatively flat dynamically averaged structures for such species, most current pharmaceutical applications continue to rely on energy-minimization methods because of limitations in software and computational resources. As a practical matter, a computational model is needed for these applications that accurately describes conformational and intermolecular-interaction energies but yields planar or nearly planar *equilibrium* geometries that accord with what medicinal chemists and others on project teams expect to see. MMFF94s is such a model. Available in these laboratories for the past few years, it is used by applications modelers in preference to MMFF94 itself for almost all project studies.

This article describes the parameterization and performance of MMFF94s. It specifies the computational and experimental data employed, defines the procedures used to adjust key out of plane bending force constants and to derive modified torsion parameters, and documents the success with which MMFF94s reproduces time-averaged experimental geometries at delocalized trigonal nitrogen centers and *ab initio* values for conformational energies. The article restates the functional forms used for out of plane bending and torsion interaction in MMFF94 and MMFF94s and characterizes the degree of nonplanarity found in X-ray and neutron-diffraction structures for delocalized trigonal nitrogen compounds. This work also compares MMFF94 and MMFF94s geometries for such systems and describes how key MMFF94s out of plane bending parameters were determined. Then the procedures used to define and reparameterize a subset of the MMFF94 torsion parameters are described. Finally, MMFF94 and MMFF94s are compared for conformational energies in such systems and a concluding discussion follows.

We should note at the outset that all comparisons of conformational energies made in this article use high quality *ab initio* calculations to define the reference values. This use of *ab initio* reference values is a necessity because relevant experimental conformational energies are available for too few delocalized trigonal nitrogen systems to be useful either in parameterizing or validating MMFF94s.

MMFF94s, whose development was cited in the original MMFF94 articles,[1,3,4] was previously tested for conformational energies in a model peptide system.[6] The force field has been implemented in the BatchMin module of the MacroModel program suite[7] since Version 5.5 and, together with MMFF94, has now been implemented in CHARMm for access via Cerius[2] and QUANTA.[8] A companion article[9] examines the ability of MMFF94, MMFF94s, CFF95, CVFF, MSI CHARMm, AMBER*, OPLS*, MM2*, and MM3* to reproduce experimental and high-quality *ab initio* results for conformational energies and to produce realistic values and trends for intermolecular-interaction energies and geometries in hydrogen-bonded complexes, quantities that are critical for success in a host of practical molecular-modeling applications.

Derivation of MMFF94s: Form for Out of Plane Bending and Torsion Interactions

In MMFF94[1,3] and MMFF94s the out of plane bending interactions use the form

$$\mathrm{EOOP}_{ijk;l} = 0.043844 \frac{k\,\mathrm{oop}_{IJK:L}}{2} \chi_{ijk;l}^2, \quad (1)$$

where $k\,\mathrm{oop}_{IJK:L}$ is the force constant (mdyn Å/rad^2) and $\chi_{ijk;l}$ is the Wilson angle[10] (°) between the bond j-l and the plane i–j–k. The three angles that arise at a given center j are assigned the same $k\,\mathrm{oop}_{IJK:L}$ force constant. In MMFF94, the angle-bending potential uses reference values that aver-

age less than 120° to generate pyramidalized trigonal centers that reproduce the reference MP2/6-31G* geometries for amides and unsaturated amines. This approach allows the out of plane term to be adjusted to improve the fit to inversion barriers. For MMFF94s we chose to retain the same angle-bending reference values and simply to modify certain out of plane bending force constants to enforce planarity in the target compounds.

Both force fields use a threefold representation for torsion interactions, where Φ is the i–j–k–l torsion angle:

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

The constants V_1, V_2, and V_3 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 a defined torsion-type index TT_{IJKL}. As previously noted,[1,4] the latter index allows torsion interactions within four-membered rings and certain five-membered rings to be given special torsion constants; it also allows special values to be employed for torsion interactions in which either the central or a peripheral bond is a delocalized single bond between sp- or sp^2-hybridized carbon atoms. Central bonds of this type occur in biphenyl, butadiene, and styrene, for example.

Other terms in the MMFF94 energy expression are unchanged in MMFF94s.

Parameterization of Out of Plane Bending Interactions

The question of how to model the geometry of amides and unsaturated amines (vinyl amines, aromatic amines, guanidines, nucleic acid bases, etc.) was one of the most vexing we encountered when developing MMFF94. Although moderately nonplanar structures for amides have been described, particularly for medium-ring lactams,[11] time-averaged crystallographic structures can be planar if distinct energy minima of opposite handedness are equally populated. In contrast, the microwave structure for aniline[12] is strongly puckered. Good quality theoretical structures for amides and unsaturated amines also tend to be puckered. For example, the MP2/6-31G* structure for aniline has virtually the same dihedral angle between the planes of the NH_2 group and the aromatic ring (42.8°) as does the microwave structure (42.3°). Moreover, the related MP2/6-31G* Wilson angle for aniline (47.7°) is not much smaller than the typical Wilson out of plane angle for saturated amines (57.5°).[3] For urea, *ab initio* geometries obtained at levels up to QCISD(T)/TZP and MP2/TZ(2*df*, 2*pd*) all show average Wilson angles of 40–43°,[3] making it nearly as strongly pyramidalized as aniline. In addition, MP2/TZP calculations find guanidine to be as strongly puckered as aniline and formamide and acetamide to be puckered as well, although much less strongly.[3]

However, several lines of evidence indicate that planarity is significantly promoted in the condensed phase by cooperative intermolecular effects. Such effects, for example, can strengthen the C—N amide-bond resonance in amides[13] and can favor locally planar geometries that maximize intermolecular hydrogen-bonding interactions. For example, while good-quality theoretical calculations indicate that individual nucleic-acid bases are appreciably puckered in gas phase,[14] the cytosine ⋯ cytosine and the Watson–Crick (WC) thymine ⋯ adenine and guanine ⋯ cytosine (GC) base pairs become perfectly planar in theoretical calculations carried out at the HF/6-31G* and/or HF/6-31G** levels.[15] This finding, however, appears to be an artifact of the use of HF calculations, which typically decrease the tendency toward puckering. In particular, in Jaguar[16] calculations carried out on the correlated LMP2/6-31G** surface[17] we find that the GC-WC complex retains appreciable puckering at the hydrogen-bonded NH_2 group of guanine (27° Wilson angle versus 46° for the isolated LMP2/6-31G** structure), although the corresponding NH_2 group of cytosine, which is slightly less puckered on the LMP2/6-31G** surface (33° Wilson angle), becomes very nearly planar in the complex (2° Wilson angle).

What we think is a fair statement is that intermolecular interactions, as well as more general condensed-phase cooperative effects, reduce, but do not necessarily eliminate, puckering at trigonal unsaturated nitrogen. This factor complicates the parameterization of molecular force fields. Some, such as AMBER[18] and CHARMM,[19] make such nitrogens planar in all circumstances. This is also the approach taken in MMFF94s, which, like them, also produces planar or nearly planar equilibrium geometries that mimic typical condensed-phase structures. Even though we were generally aware of the tendency of condensed-phase interactions to promote planarity,[3] MMFF94, in contrast, was parameterized to reproduce the appreciable nonplanarity found in reference MP2/6-31G* cal-

culations on the isolated species. Clearly, neither limiting approach is completely correct.

To obtain experimental data for the parameterization of MMFF94s, we reexamined the set of ~ 2800 good quality crystallographic structures previously used to extend the parameterization of MMFF94.[5] From this set we extracted all structures containing one or more nitrogen atoms of MMFF atom type 10 (amide nitrogen), 40 (unsaturated amine nitrogen), or 43 (sulfonamide nitrogen). We then divided the selected structures into subsets classified according to the type of nitrogen, the presence or absence of cyclization to it, and the presence or absence of attached hydrogens; only neutron-diffraction structures were retained when one or more hydrogen atoms were attached to nitrogen.

Table SM-I (Supplementary Material)[20] lists the Cambridge Structural Database reference codes for the extracted structures, together with the experimentally determined Wilson out of plane angle at the nitrogen in question (obtained by averaging the absolute values of the three Wilson angles made to each such center). This table also reports the optimized MMFF94 and MMFF94s Wilson angles. Table I summarizes the results by listing the average observed and calculated Wilson angles for each of the structural classes.

The tables show that crystallographic structures for sulfonamides are appreciably puckered, whereas those for amides and unsaturated amines tend to be more nearly planar. Structures for ureas and cyclic amides also show some degree of pyramidalization, but usually to a lesser extent than do

TABLE I.
Average Pyramidalization at Nitrogen.

Category	No. Centers	Average Wilson Angle (°)			Average Dev. (°)	
		Expt	MMFF94	MMFF94s	MMFF94	MMFF94s
1° amides	6	6	21	4	15	−2
1° ureas	4	4	41	3	37	−1
1° acyclic thioamides	2	2	16	0	14	−2
1° aromatic amines and nucleic-acid bases	7	27	46	9	19	−18
1° amidines and guanidines	5	8	46	12	38	4
2° amides and ureas	4	13	30	10	17	−3
2° amidines and guanidines	2	22	30	11	8	−11
2° hydroxamic acids	1	15	32	10	17	−5
3° acyclic amides	10	4	10	3	6	−1
3° 5-ring amides	20	8	11	6	3	−2
3° 6-ring amides	21	10	17	7	7	−3
3° 8-ring amides	2	4	15	3	11	−1
3° bicyclic amides	11	15	22	13	7	−2
3° acyclic ureas	3	11	22	2	11	−9
3° cyclic ureas	2	20	13	13	−7	−7
3° bicyclic ureas	3	24	22	17	−2	−7
3° 6-ring hydroxamic acids	1	3	5	3	2	0
3° acyclic thioamides	3	3	8	2	5	−1
3° acyclic thioureas	3	1	14	5	13	4
3° cyclic thioamides	4	2	11	3	9	
3° bicyclic thioureas	1	44	40	33	−4	−11
3° acyclic enamines	7	5	12	5	7	0
3° aromatic amines	24	10	26	4	16	−6
3° cyclic / bicyclic aromatic amines	7	31	38	27	7	−4
3° cyclic enamines	16	26	28	15	2	−11
3° guanidines / amidines	10	6	19	8	13	2
3° acyclic sulfonamides	3	30	32	32	2	2
3° 6-ring sulfonamides	2	41	39	39	−2	−2
3° large-ring sulfonamides	2	25	26	26	1	1
Overall summary	186	13	23	9	10	−4

the energy-minimized MMFF94 structures. By contrast, the MMFF94s structures on average are closer to planarity than are the experimental structures, particularly when cyclization constraints give rise to an appreciable pyramidalization in the experimental structure that cannot be fully maintained by the modified MMFF94s potential. This enforcement of planarity was achieved principally by substituting positive out of plane bending force constants of 0.015 mdyn Å/rad^2 for type 10 (amide) and of 0.03 mdyn Å/rad^2 for type 40 (unsaturated amine) nitrogen for the negative, but comparably small, force constants used in MMFF94. These positive values represent the smallest values that produced optimized structures having suitably small Wilson angles in test calculations. The cited MMFF94s results use these force constants in conjunction with the modified MMFF94s torsion parameters described in the next section.

Parameterization of Torsional Interactions

SELECTION OF *AB INITIO* CONFORMATIONAL DATA

The parameterization of MMFF94s used a subset of the data previously employed in parameterizing MMFF94.[4] These data consist of ''MP4SDQ/TZP'' calculations carried out at MP2/6-31G*-optimized equilibrium geometries (set *A*) and single-point MP2/TZP calculations carried out at a series of torsionally incremented geometries (set *B*). The ''MP4SDQ/TZP'' calculations are composite calculations that add third-order and fourth-order MP3 and MP4SDQ perturbative corrections evaluated using a 6-31G*-type basis set to the MP2/TZP energy, where TZP indicates use of a polarized triple-zeta basis set. We previously showed that evaluating the third-order and fourth-order perturbative corrections with the smaller basis set introduces very little error relative to unapproximated MP4SDQ/TZP calculations.[4] Moreover, this level of theory accounts for experimentally determined relative conformational energies, enthalpies, and free energies much better than do calculations that employ smaller basis sets and/or make lesser provision for the effects of electron correlation.[4, 9] The comparisons in sets *A* and *B* use, or are based on, the MMFF94 conformations cited in Table II. The geometries for set *B* were derived from base MP2/6-31G* equilibrium geometries by rotating one torsion bond by a specified amount (e.g., $\pm 30°$, $\pm 60°$, etc.) and (to minimize steric clashes) then minimizing the resultant structure, subject to strong restraints applied to all torsion angles, using an early version of MMFF; to make the force-field comparisons of relative energies meaningful, we also minimized the base MP2/6-31G* geometries in the same manner.[4]

GENERAL APPROACH

The torsion reparameterization used the program TORFIT in conjunction with the previously described iterative approach.[4] This approach, which fits all the conformational-energy data simultaneously, minimizes the sum of squares shown in eq. (3):

$$\mathrm{SOS} = \sum w_i (\mathrm{CE}_i^{\mathrm{MMFF}} - \mathrm{CE}_i^{\mathrm{ref}})^2 + p \sum (V_n(IJKL) - V_n^0(IJKL))^2. \quad (3)$$

The first summation extends over the conformational energies of sets *A* and *B*, where ref denotes either an ''MP4SDQ/TZP'' or a MP2/TZP energy difference. The quantity w_i is a weight factor that was set to 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 keep each torsion parameter reasonably close to its input value. In the absence of such restraints, correlations between parameters typically lead to large changes in value that produce parameters of questionable physical validity while minimally reducing the sum of squares. The penalty-function restraints tend to hold parameters to zero, where possible, or to equal magnitudes when simple parameter correlations are involved.

SELECTION OF TORSION PARAMETERS TO BE OPTIMIZED

To determine which MMFF94 torsion parameters needed to be reoptimized, we collected all the MMFF94 torsion parameters that contribute to the MMFF energy expressions for the molecules of sets *A* and *B*. Using the modified out of plane bending force constants described above and employing the MMFF94 torsion parameters as input values, we then performed a linear least-squares fit to the conformational-energy data. In this refinement, we applied weak penalty-function restraints using $p = 0.1$ [see eq. (3)] to suppress variations in the torsion parameters while allowing the conforma-

tional energies to be fit almost as well as in a free fit. Next we fixed all torsion parameters that did not change appreciably at their MMFF94 values (i.e., removed them from the optimization set) and repeated the least-squares fit. Proceeding in this way, we fixed additional sets of torsional parameters at their MMFF94 values and ultimately determined that reoptimizing only those parameters that explicitly referenced a type 10 or a type 40 nitrogen would yield conformational energies practically identical to those that would be obtained from a full torsional reparameterization.

LEAST-SQUARES REFINEMENT

As previously described,[4] we began by generating an unbiased initial set of torsion parameters in which all except certain twofold MMFF94 parame-

TABLE II.
Conformers Considered in Parameterizing MMFF94s.

Amides and Peptide Analogs
AM01a: Formamide
AM02a: *N*-Methylformamide, *cis*
AM02b: *N*-Methylformamide, *trans*
AM03a: Acetamide
AM04a: *N*-Methylacetamide, *trans*
AM04b: *N*-Methylacetamide, *cis*
AM05a: *N*,*N*-Dimethylformamide
AM06a: Urea, puckered
AM07a: *N*-Formylformamide, both o=c—n—h *cis*
AM07b: *N*-Formylformamide o=c—n—h cis, *trans*
AM08a: Formylglycinamide
AM09a: Glycine dipeptide analog, C7
AM09b: Glycine dipeptide analog, C5
AM10a: Alanine dipeptide analog, $C7_{eq}$
AM10b: Alanine dipeptide analog, C5
AM10c: Alanine dipeptide analog, $C7_{ax}$
AM10d: Alanine dipeptide analog, α'
AM10e: Alanine dipeptide analog, β_2
AM10f: Alanine dipeptide analog, α_L
AM11a: Propionamide, c—c—c—n *anti*
AM12a: *N*-Ethylformamide, c—c—n—c *gauche*
AM12b: *N*-Ethylformamide, c—c—n—c *anti*
AM13a: *N*-OH,*N*-Methylacetamide, o—n—c=o *trans*
AM13b: *N*-OH,*N*-Methylacetamide, o—n—c=o *cis*
AM14a: *N*-OH,*N*-Ethylacetamide, o—n—c=o *trans*, c—c—n—o *gauche*
AM14b: *N*-OH,*N*-Ethylacetamide, o—n—c=o *trans*, c—c—n—o *skew*
AM14c: *N*-OH,*N*-Ethylacetamide, o—n—c=o *cis*, c—c—n—c(=o) *skew*
AM14d: *N*-OH,*N*-Ethylacetamide, o—n—c=o *cis*, c—c—n—c(=o) *gauche*
AM15a: *N*-OH,*N*-Methylpropionamide, o—n—c=o *trans*, c—c—c=o *cis*
AM15b: *N*-OH,*N*-Methylpropionamide, o—n—c=o *trans*, c—c—c=o *skew*
AM15c: *N*-OH,*N*-Methylpropionamide, o—n—c=o *cis*, c—c—c=o *cis*
AM15d: *N*-OH,*N*-Methylpropionamide, o—n—c=o *cis*, c—c—c=o *skew*
AM16a: Glycine dipeptide, C7
AM16b: Glycine dipeptide, C5
AM17a: Alanine dipeptide, $C7_{eq}$
AM17b: Alanine dipeptide, C5
AM17c: Alanine dipeptide, $C7_{ax}$
AM17d: Alanine dipeptide, α'
AM17e: Alanine dipeptide, β_2
AM17f: Alanine dipeptide, α_L
CJ04a: 2-Methylpropenamide, c=c—c=o *cis*
CJ04b: 2-Methylpropenamide, c=c—c—o *skew*
CJ05a: Propenamide, c=c—c=o *cis*
CJ05b: Propenamide, c=c—c=o *skew*
Guanidines and Amidines
IM01a: Formamidine, h—n=c—n *cis*, N puckered
IM01b: Formamidine, h—n=c—n *anti*, N puckered
IM04a: *N*-Methylformamidine, n—c=n—c *cis*, N puckered
IM04b: *N*-Methylformamidine, n—c=n—c *trans*, N puckered
IM05a: Guanidine, N puckered
IM06a: N_2-Methylguanidine, N puckered
IM07a: Butadiene Schiff base, c=c—c=n s-*trans*, h—n=c—c *cis*
IM07b: Butadiene Schiff base, c=c—c=n s-*cis*, h—n=c—c *trans*
Aromatic Amines
NH14a: Aniline, N puckered
NH19a: *N*-Methylaniline, N puckered

ters (e.g., for rotation about C—N bonds in amides) were set to zero. We then employed a strong torsion-restraint force constant (100 kcal/mol/rad^2) to all torsion angles in carrying out force-field optimizations of the reference MP2/6-31G* geometries for set *A*. Next we performed an initial linear least-squares refinement using a moderate penalty-function restraint of $p = 0.5$. We then repeated this procedure, using the updated torsion parameters as the new input values, with successively smaller torsion-restraint force constants. Finally, we zeroed all torsion parameters smaller than 0.05 kcal/mol magnitude and reoptimized the remaining nonzero parameters.

TABLE III.
MMFF94s Torsion Parameters (kcal / mol).

TT_{IJKL}	*I*	*J*	*K*	*L*	V_1	V_2	V_3	Origin / Comment[a]
0	5	1	1	10	0.000	0.000	0.418	C94S
0	1	1	3	10	−0.763	1.244	0.986	C94S
0	5	1	3	10	−0.687	1.244	0.136	C94S
0	10	1	3	7	0.530	2.905	2.756	C94S
0	10	1	3	10	0.465	−0.241	1.850	C94S
0	1	1	10	3	−0.884	0.578	0.818	C94S
0	1	1	10	6	0.000	−0.379	0.565	C94S
0	1	1	10	28	0.750	−0.404	0.369	C94S
0	3	1	10	3	3.219	−2.699	1.875	C94S
0	3	1	10	28	0.207	0.461	0.324	C94S
0	5	1	10	1	0.000	0.000	0.706	C94S
0	5	1	10	3	−2.334	1.517	−0.065	C94S
0	5	1	10	6	0.000	0.688	0.665	C94S
0	5	1	10	28	−0.982	−0.207	0.166	C94S
0	5	1	40	28	0.000	−0.105	0.000	C94S
0	5	1	40	37	0.000	0.000	0.468	C94S
1	1	2	3	10	0.000	2.237	−0.610	C94S
1	2	2	3	10	0.000	1.599	0.380	C94S
1	5	2	3	10	0.000	1.409	0.254	C94S
0	2	2	40	28	0.000	3.305	−0.530	C94S
0	5	2	40	28	0.139	3.241	0.139	C94S
0	40	3	9	1	−0.704	18.216	0.000	C94S
0	40	3	9	27	0.000	16.000	0.178	C94S
0	1	3	10	1	0.831	6.061	0.522	C94S
0	1	3	10	6	−1.152	8.588	1.511	C94S
0	1	3	10	28	−0.259	5.934	1.326	C94S
2	2	3	10	28	0.000	6.561	0.294	C94S
0	5	3	10	1	−0.195	6.304	1.722	C94S
0	5	3	10	3	−0.705	5.383	0.234	C94S
0	5	3	10	28	−0.417	5.981	0.511	C94S
0	7	3	10	1	−0.491	6.218	0.000	C94S
0	7	3	10	3	0.733	−0.543	−0.163	C94S
0	7	3	10	6	1.234	8.372	−0.539	C94S
0	7	3	10	28	1.168	4.857	−0.341	C94S
0	10	3	10	28	0.000	3.706	1.254	C94S
0	5	3	40	28	−1.355	3.964	0.800	C94S
0	9	3	40	28	1.045	3.785	−0.291	C94S
0	40	3	40	28	0.508	2.985	0.809	C94S
0	21	6	10	1	0.829	0.000	−0.730	C94S
0	21	6	10	3	0.675	−0.185	−1.053	C94S
0	37	37	40	1	0.000	4.095	0.382	C94S
0	37	37	40	28	0.698	2.542	3.072	C94S

[a] The C94S designation distinguishes these entries from core MMFF94 (C94) entries in the full MMFFSTOR.PAR torsion parameter file.

TORSION PARAMETERS

The unique MMFF94s torsion parameters are shown in Table III. The full set of MMFF94s out of plane and torsion parameters, which includes those unchanged from MMFF94, are available in ASCII format in the Supplementary Material for this article.

Conformational Energies for Delocalized Trigonal Nitrogen Systems

In the listing below, the term Conformational Energies refers to the 27 ''MP4SDQ/TZP''-based comparisons of set *A* cited in Table IV, whereas Torsion-Profile Energies refers to the 228 MP2/TZP

TABLE IV.
Comparison of ''MP4SDQ/TZP,'' MMFF94, and MMFF94s Conformational Energies (kcal/mol).

Conformational Comparison	$E(A) - E(B)$		
	''MP4''	MMFF94	MMFF94s
N-Methylformamide, *cis* − *trans*	1.04	1.28	1.09
N-Methylacetamide, *cis* − *trans*	1.90	2.16	1.97
N-Formylformamide, [o=c—n—h *cis*, *trans*] − [o=c—n—h *cis*, *cis*]	0.41	0.40	0.44
Glycine dipeptide analog, C5 − C7	0.91	1.14	1.08
Alanine dipeptide analog, C5 − $C7_{eq}$	1.08	1.12	1.11
Alanine dipeptide analog, $C7_{ax}$ − $C7_{eq}$	2.20	1.83	1.58
Alanine dipeptide analog, α' − $C7_{eq}$	4.88	4.85	5.24
Alanine dipeptide analog, β_2 − $C7_{eq}$	2.76	2.47	2.55[a]
Alanine dipeptide analog, α_L − $C7_{eq}$	4.10	3.87	4.28
N-OH,*N*-Methylacetamide, o—n—c=o *cis* − *trans*	0.04	0.01	-0.19
N-OH,*N*-Ethylacetamide, [o—n—c=o *trans*, c—c—n—o *gauche*] − [o—n—c=o *cis*, c—c—n—c(=o) *gauche*]	0.24	0.14	0.53
N-OH,*N*-Ethylacetamide, [o—n—c=o *trans*, c—c—n—o *skew*] − [o—n—c=o *cis*, c—c—n—c(=o) *gauche*]	1.50	1.56	1.05[a]
N-OH,*N*-Ethylacetamide, [o—n—c=o *cis*, c—c—n—c(=o) *skew*] − [o—n—c=o *cis*, c—c—n—c(=o) *gauche*]	0.08	0.27	1.64[a]
N-OH,*N*-Methylpropionamide, [o—n—c=o *trans*, c—c—c=o *cis*] − [o—n—c=o *cis*, c—c—c=o *cis*]	0.04	0.08	0.24
N-OH,*N*-Methylpropionamide, [o—n—c = o *trans*, c—c—c=o *skew*] − [o—n—c=o *cis*, c—c—c=o *cis*]	0.96	1.07	1.24
N-OH,*N*-Methylpropionamide, [o—n—c=o *cis*, c—c—c=o *skew*] − [o—n—c=o *cis*, c—c—c=o *cis*]	0.72	1.09	1.24
Glycine dipeptide, C5 − C7	1.66	1.62	1.40
Alanine dipeptide, C5 − $C7_{eq}$	1.64	1.68	1.42
Alanine dipeptide, $C7_{ax}$ − $C7_{eq}$	2.20	2.26	1.84
Alanine dipeptide, α' − $C7_{eq}$	5.35	5.47	5.66
Alanine dipeptide, β_2 − $C7_{eq}$	3.20	3.06	2.89[a]
Alanine dipeptide, α_L − $C7_{eq}$	4.25	4.75	4.74
2-Methylpropenamide, c=c—c=o *cis* − *skew*	1.05	1.00	1.04
Propenamide, c=c—c=o *skew* − *cis*	0.61	0.60	0.57
Formamidine, [h—n=c—n *cis*, N puckered] − [h—n=c—n *trans*, N puckered]	2.03	2.27	2.03
N-Methylformamidine, n—c=n—c *cis* − *trans*	2.00	2.12	2.00
Butadiene Schiff base, [c=c—c=n s-*cis*, h—n=c—c *trans*] − [c=c—c=n s-*trans*, h—n=c—c *cis*]	1.73	1.94	1.94
(RMS value), RMS error	(2.30)	0.20	0.41

[a] The higher energy conformers in these comparisons were optimized on the MMFF94s surface with one torsion angle restrained (see text).

torsionally incremented comparisons of set *B*. All results are stated as root mean square deviations (RMSDs; kcal/mol):

	MMFF94	MMFF94s
Conformational Energies	0.20	0.41
Torsion-Profile Energies	0.62	0.59

To place these RMSD in context, we note that the RMS values of the reference *ab initio* relative energies are 2.30 kcal/mol for the conformational energies subset and 4.55 kcal/mol for the torsion-profile energies subset. Thus, MMFF94s accounts reasonably well for the *ab initio* data, though MMFF94 describes the conformational energies subset somewhat better. A good part of the larger MMFF94s deviation for this subset arises from the [o—n—c=o *cis*, c—c—n—c(=o) *skew*] – [o—n—c=o *cis*, c—c—n—c(=o) *gauche*] comparison for *N*-OH,*N*-ethylacetamide (Table IV), the exclusion of which would lower the RMSD to 0.28 kcal/mol. The strongly puckered MP2/6-31G* geometry of the higher energy conformer includes an internal O—H ⋯ O hydrogen bond that strongly affects the local geometry and energy surface. Because unrestrained optimization on the MMFF94s surface converts this conformer to the o—n—c=o *cis*, c—c—n—c(=o) *gauche* form, we tethered one torsion angle during its optimization (and also during those of the β_2 conformer of the alanine dipeptide analog, the o—n—c=o *trans*, c—c—n—o *skew* conformer of *N*-OH,*N*-ethylacetamide, and the β_2 conformer of alanine dipeptide, which freely optimize to the $C7_{eq}$, the [o—n—c=o *trans*, c—c—n—o *gauche*], and $C7_{eq}$ conformers, respectively). Ordinarily, the modest displacement required to achieve a locally planar geometry would not be expected to greatly affect the conformational energy. In this case, however, the *N*-OH,*N*-ethylacetamide conformers pucker differently on the MMFF94 surface and therefore have different (negative) out of plane bending energies. MMFF94s, in contrast, produces near-planar structures that have small (positive) out of plane bending energies. The net result is that the change from a negative to a positive force constant selectively raises the relative energy of the o—n—c=o *cis*, c—c—n—c(=o) *skew* form. The torsional reparameterization undoubtedly compensates to some degree, but the net result is the large overestimate of the [o—n—c=o *cis*, c—c—n—c(=o) *skew*] – [o—n—c=o *cis*, c—c—n—c(=o) *gauche*] energy difference reported in Table IV. This example shows that enforcement of planarity at a trigonal delocalized nitrogen atom can significantly damage a conformational energy difference. Fortunately, such cases appear to be rare.

Concluding Discussion

This article describes the derivation of MMFF94s, the static option to MMFF94. MMFF94s is intended for use in energy-minimization studies rather than molecular-dynamics simulations. The two force fields share most parameters and yield identical results in all cases not involving delocalized trigonal nitrogen of MMFF type 10 (amides) or 40 (unsaturated amines). MMFF94 typically yields optimized geometries that are appreciably puckered at such nitrogen centers, but MMFF94s uses altered out of plane bending parameters to produce planar or more nearly planar geometries that mimic the time-averaged structures commonly found via crystallographic techniques. We did not include sulfonamides (MMFF type 43 nitrogen) in the reparameterization because the crystallographic structures show a degree of puckering at nitrogen similar to that found in the reference MP2/6-31G* structures used in parameterizing MMFF94.

To compensate for the changed potential for out of plane bending, the derivation of MMFF94s also required a torsional reparameterization. We found that only interactions that involved MMFF types 10 (amides) or 40 (enamines, aromatic amines, guanidines, and nucleic-acid bases) needed to be reevaluated. Comparison to MMFF94 shows that conformational energies (except, obviously, for barriers to inversion at delocalized nitrogen) are still well described. However, a few cases were noted in which MMFF94 successfully located a conformation close to the reference MP2/6-31G* geometry whereas MMFF94s did not.

MMFF94 is clearly appropriate for molecular-dynamics simulations, but MMFF94s can be recommended for use in energy-minimization studies. Comparisons for conformational and intermolecular-interaction energies and geometries for MMFF94, MMFF94s, and a wide variety of other force fields[6, 9] suggest that both can be expected to perform well in general molecular simulations.

As a service to the community, a validation suite for MMFF94s, comprising input molecular coordinates in various formats and output results from OPTIMOL[1] and from BatchMin,[7] will be

posted on the Computational Chemistry List Web site[21] when this article is published. The Web site also contains an extensive validation suite for MMFF94.

Acknowledgments

I wish to thank Mary Baum, Chris Culberson, Ralph Mosley, Bob Nachbar, Kris Prendergast, Dennis Underwood, and others in the Merck Research Laboratories who encouraged me to develop MMFF94s and have routinely used it in their own work. I also thank Jan Labanowski of the Ohio Supercomputer Center for his encouragement and assistance in posting the MMFF94 and MMFF94s Validation Suites on the CCL Web site.

Supplementary Material

Table SM-I gives observed and calculated degrees of pyramidalization at trigonal delocalized nitrogen in structures retrieved from the Cambridge Structural Database. Computer readable files of the MMFF94s parameters for out of plane bending and for torsion interactions in ASCII format are also provided. These files (which are called MMFFSOOP.PAR and MMFFSTOR.PAR in the OPTIMOL,[1] BatchMin,[7] and CHARMm[8] implementations of MMFF94s) are full parameter sets that include the original MMFF94 and modified MMFF94s parameters. They are used in place of the corresponding MMFF94 files MMFFOOP.PAR and MMFFTOR.PAR. For information on accessing this Supplementary Material, see the footnote on the first page of this article.

References

1. Halgren, T. A. J Comput Chem 1996, 17, 490–519.
2. Halgren, T. A. J Comput Chem 1996, 17, 520–552.
3. Halgren, T. A. J Comput Chem 1996, 17, 553–586.
4. Halgren, T. A.; Nachbar, R. B. J Comput Chem 1996, 17, 587–615.
5. Halgren, T. A. J Comput Chem 1996, 17, 615–641.
6. Beachy, M. D.; Chassman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. J Am Chem Soc 1997, 119, 5908–5920.
7. MacroModel and BatchMin were developed in the laboratories of Professor W. Clark Still and are available from Schrodinger, Inc. (Portland, OR). Version 5.5 was released in 1996.
8. Cerius[2], QUANTA, and CHARMm are available from Molecular Simulations, Inc., 9685 Scranton Road, San Diego, CA 92121-2777.
9. Halgren, T. A. J Comput Chem 1999, 20, 730–748.
10. Wilson, E. B., Jr.; Decius, J. C.; Cross, P. C. Molecular Vibrations; Dover: New York, 1955; Chapter 4.
11. Dunnitz, J. D.; Winkler, F. K. Acta Crystallogr 1975, B31, 251–263, and references therein.
12. Roussy, G.; Nonat, A. J Mol Spectrosc 1986, 118, 180–188.
13. See, for example, Guo, H.; Karplus, M. J Phys Chem 1992, 96, 7273–7287.
14. (a) Sponer, J.; Hobza, P. J Phys Chem 1994, 98, 3161–3164; (b) Leszcynski, J. J Phys Chem A 1998, 102, 2357–2361, and references therein.
15. (a) Sponer, J.; Leszcynski, J.; Hobza, P. J Phys Chem 1996, 100, 1965–1974, and references therein; (b) Gould, I. R.; Kollman, P. A. J Am Chem Soc 1994, 116, 2493.
16. Schrodinger, Inc. Jaguar v3.5; Schrodinger, Inc.: Portland, OR, 1998.
17. (a) Murphy, R. B.; Beachy, M. D.; Friesner, R. A.; Ringnalda, M. N. J Chem Phys 1995, 103, 1481; (b) Friesner, R. A.; Murphy, R. B.; Beachy, M. D.; Ringnalda, M. N.; Pollard, W. T.; Dunietz, B. D.; Cao, Y. J Chem Phys, in press.
18. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. J Am Chem Soc 1995, 117, 5179–5197.
19. (a) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph–McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz–Kuczera, J.; Yin, D.; Karplus, M. J Phys Chem B 1998, 102, 3586–3616; (b) MacKerell, A. D., Jr.; Wiórkiewicz–Kuczera, J.; Karplus, M. J Am Chem Soc 1995, 117, 11946.
20. For information on accessing this Supplementary Material, see the footnote on the first page of this article.
21. The MMFF94 validation suite can be accessed via the Web at http://ccl.osc.edu/cca/data/MMFF94 or at ftp://ccl.osc.edu/pub/chemistry/data/MMFF94. Alternatively, the MMFF94 suite can be accessed via ftp at ccl.osc.edu; cd to pub/chemistry/data/MMFF94. This location is part of the Computational Chemistry List archive site maintained by the Ohio Supercomputer Center. The MMFF94s validation suite, which will be posted when this article appears, will have a similar address, but with MMFF94 replaced by MMFF94s.