

Validation of the General Purpose QUANTA®3.2/CHARMm® Force Field*

Frank A. Momany† and Rebecca Rone

Polygen/Molecular Simulations, 200 Fifth Avenue, Waltham, Massachusetts 02254

Received 9 September 1991; accepted 24 March 1992

An evaluation of the CHARMm force field for small molecules is described. Using different force field conditions and computational techniques, a wide variety of compounds are analyzed. rms deviations of Cartesian coordinates for 49 diverse organic molecules taken from the Cambridge Crystallographic Data Base and internal coordinate geometries for 28 other molecules are reported. Results are described with different dielectrics, dihedral constraints, and crystal packing to allow analysis of deviations from experimental data and give precise statements of the reliability of the parameters used in the force field. Torsional barriers (rms = 0.4) and conformational energy differences (rms = 0.4) are examined and comparisons made to other force fields such as MM2, Tripos, and DREIDING. The results confirm that CHARMm is an internally consistent all purpose force field with energy terms for bonds, angles, dihedrals, and out-of-plane motions, as well as nonbonded electrostatic and van der Waals interactions. Reported CHARMm results (rms = 0.006 Å for bonds, rms = 1.37° for angles, and rms = 3.2° for dihedrals) are in excellent agreement with high quality electron diffraction data. © 1992 by John Wiley & Sons, Inc.

INTRODUCTION

Molecular mechanics calculations are widely used for the study of small organic molecules, proteins, sugars, lipids, and other biopolymers. More recently, calculation methods have become available for large polymeric systems, many inorganics, liquids, ionic solutions, and crystal lattices. The CHARMm force field¹ consists of a functional form for the potential energy that lends itself relatively easily to the study of a variety of different types of molecules. Individual energy contributions are obtained from functions that are simple and can be interpreted in a physical sense. These include harmonic bonding and angle terms, cosine torsional or dihedral terms, and nonbonded terms which include electrostatic Coulombic terms using partial atomic charges and van der Waals Lennard-Jones 6-12 contributions.

In this article, CHARMm results are compared to experimental data and to reported results from other force fields.^{2,3} In the evaluation of the results several features of the use of selected force fields will be noted. For example, the inclusion of full electrostatic Coulombic interactions using partial atomic charges is standard in CHARMm, and

was used in this study. However, in one study² used for comparisons in the discussion section, the authors chose not to include electrostatics in their reported results. When the electrostatic terms are not included in the force field, one must be particularly cautious in interpreting simulations on polar molecules and their intra- and intermolecular interactions.

Improving the parameters that are used within the CHARMm force field is an ongoing project and hundreds of molecules have been analyzed in the development of current parameters.⁴ The molecules used in parameter development include saturated hydrocarbons, unsaturated hydrocarbons, strained systems, conjugated and aromatic systems, all major polar functional groups, including alcohols, aldehydes, amines, etc., as well as heterocyclic and ionic systems. The result is a force field that is applicable to small organic molecules, or large polymeric molecules, as well as macromolecular biologically related molecules.

The experimental data used in parameterization includes gas phase electron diffraction and solid state molecular structures as well as high level SCF *ab initio* calculations. The *ab initio* results are very useful in giving conformationally dependent geometry trends that are not readily available from experimental measurements,⁵ energy differences between conformations in selected comparisons,⁶ and barriers to rotation about dihedral angles.⁶ Where sufficient *ab initio* basis sets are included

*QUANTA AND CHARMm are registered trademarks of Polygen Corporation. QUANTA/CHARMm parameters are from Release 3.2, Polygen Corporation.

†Author to whom all correspondence should be addressed.

and correlation (MP2, MP3, etc.) terms calculated, the data has been used to refine van der Waals and electrostatic nonbonded interactions.⁷ Parameter refinement and the determination of the charge set⁸ will not be described here since this article is directed to the preliminary validation of results of calculations using the force field.

In the analysis of crystal structures from the Cambridge Crystallographic Data Base,⁹ care is taken in this article to fully understand the cause of calculated rms deviations from experiment. Large rms deviations in atom positions can result from small variations in one or more dihedral angles when several descendent groups are moved by this small change. Because of this effect, the molecules examined here were studied using a variety of conditions and techniques. Different dielectric values were used to test the effect of the electrostatic atom-atom interactions on conformation. Dihedral angles were constrained at the X-ray values and internal coordinate geometry was energy minimized to validate geometries without the complications of conformational change. In selected cases, the molecules were simulated in their experimental crystalline environment as specified by their space group to examine the effect of the crystal forces on the conformational energetic preference.

METHODS

The CHARMM force field equations are identical to those reported in the original CHARMM publication.¹ However, the parameters used throughout this study are those distributed by Polygen in QUANTA Release 3.2, and are available upon request from Polygen/Molecular Simulations.

The parameter file for CHARMM contains 6-12 nonbonded parameter terms for all the listed atom types. This list includes 102 organic atom types and many metals in the periodic table that have had their nonbonded terms parameterized. The van der Waals terms are provided for homonuclear atom pairs and standard combination rules are applied for the energy (geometric mean) and distance (arithmetic mean) of the heteronuclear pairs. Special parameters are included for specific 1-4 interactions across a bond, such as in the cases of extended atoms. Other special 1-4 terms have also been found to be important, such as in the case of alcohols where the rotation about the hydroxyl C-O bond causes variations in the O-C-H angle of the methylene group. This and many other cases are parameterized to result in subtle but important structural enhancements.

One series of molecules used in the comparisons was taken from the Cambridge Crystallographic

Data Base.⁹ After crystal data files are read the QUANTA molecular editor is used to place hydrogens, assign atom types, and assign partial atomic charges to each atom. Other molecules studied were drawn in CHEMNOTE (a 2D drawing routine within QUANTA 3.2). The 2D structure is converted to 3D, and again atom types and partial atomic charges are assigned to every atom. Atom typing is carried out using: (1) bond types for bonds connected to the atom, (2) element types of the central atom and elements connected to it, and (3) a bonding path algorithm for searching its microenvironment. Different atom types for the same element were historically introduced to reflect variations in orbital hybridization arising from resonance or conjugated states. That reasoning has continued today with new atom types included in QUANTA 3.2 to enhance the reliability of new and uniquely different functional groups. Charges are derived from a charge template where every atom has a charge assigned to it depending upon the microenvironment of atom types as found from the bonding path algorithm. Charge smoothing algorithms are used to net the total charge to zero for neutral molecules, or to the appropriate charge of an ionic system. The partial atomic charges are designed to give correct interaction energies between molecules and groups within molecules and to give correct interaction distances such as found in hydrogen bond distances between polar molecules. No special hydrogen bond terms are included. Electrostatic energies on atoms separated by three bonds (1-4 interactions) are scaled down by a factor of 1/2. This scaling was found to be necessary to minimize the difference between conformational energies when using different dielectric methods ($C = 1$ vs. $1/R$). The nonbonded van der Waals 1-4 terms are not scaled but some specific atom types are modified as determined to be necessary by conformational energy studies.

The method used to study the torsional barrier heights and conformational energy differences, as shown in Tables I and II, includes full energy minimization of each conformation. The CHARMM energy differences include all contributions to the energy, including bond stretch, angle bend, dihedral, out-of-plane, electrostatic 1-4, electrostatic 1-5 or greater, and nonbonded interactions. The method for defining the dihedral barriers to include torsional terms for every set of four atoms across the bond has recently been adopted in QUANTA 3.2. The torsional energy barrier study reports the CHARMM, DREIDING,³ and Tripos² results obtained by holding the dihedral angle constrained and relaxing the geometry by energy minimization (see the relaxed column under CHARMM). The column designated "rigid" under CHARMM is obtained by simply rotating the dihe-

Table I. Comparisons of experimental energy differences between conformers with those obtained from the MM2, Tripos, and CHARMM force fields.

Molecule	Experimental (11)	MM2 (2)	TRIPOS (2)	DRIEDING (3)	CHARMM
Butane, <i>g-t</i>	0.8	0.9	0.6	0.8	0.8
2-Methylbutane, <i>g-t</i>	0.6	0.7	0.4	— ^b	1.0
1-Butene (120-0) deg	0.2 (−0.5) ^a	−0.5	1.0	— ^b	−0.7
Butadiene (40-180) deg	2.1	79.4 ^b	3.4	— ^b	2.2
Methyl-cyclohexane, <i>a-e</i>	1.9	1.6	1.4	1.3	1.6
Phenyl-cyclohexane, <i>a-e</i>	3.0	1.9	4.2	4.5	2.4
Fluoro-cyclohexane, <i>a-e</i>	0.2	0.2	0.0	0.3	0.4
Chloro-cyclohexane, <i>a-e</i>	0.4	0.4	0.2	0.4	0.4
Bromo-cyclohexane, <i>a-e</i>	0.5	0.5	0.3	0.5	0.4
Nitrocyclohexane, <i>a-e</i>	1.2	— ^b	0.6	1.6	2.4 ^b
Cyclohexanol, <i>a-e</i> (CCOH,180)	0.5	0.9	0.0	0.3	1.3
1,4-Dichlorocyclohexane, <i>a-e</i>	0.2	0.3	0.4	1.7	0.2
Methyl-ethyl-ether, <i>g-t</i>	1.5	1.7	0.8	1.6	1.2
Propionaldehyde (120-0)	0.9	0.7	−0.1	— ^b	0.5
Cyclohexane, twist-boat-chair	6.0	5.4	7.7	7.7	6.0
rms deviation		0.4	0.8	0.9	0.4

^aAb initio value in parentheses is from reference 10.^bOmitted from calculation of rms. Nitro-cyclohexane was omitted in the CHARMM calculation since data was not available for comparison in the MM2 results. The rms when this value is included is 0.5.**Table II.** Comparisons of experimental torsional barrier heights (kcal/mol) with those obtained from Tripos, BIOGRAF, and CHARMM force fields.^a

Molecule ^b	Experimental	TRIPOS (2)	BIOGRAF (3)	CHARMm	
				Rigid	Relaxed
Set 1					
Ethane	2.9 ^c	3.8	2.9	2.9	2.9
Propane	3.4 ^d	4.0	3.4	3.1	3.1
Butane, Me—Me eclipsed	6.0 ^d	5.9	—	6.7	4.7
Butane, Me—H eclipsed	3.4 ^d	5.1	3.8	3.3	3.1
2-Methyl-butane	3.9 ^d	4.5	4.0	3.4	3.3
Neopentane	4.7 ^d	4.7	5.1	3.8	3.7
Methylsilane	1.7 ^d	1.9	2.3	1.6	1.6
Methylamine	2.0 ^d	2.8	2.1	1.9	1.9
Dimethylamine	3.6 ^d	4.9	2.9	4.0	4.0
Fluoroethane	3.3 ^c	3.8	3.2	3.1	3.1
Chloroethane	3.7 ^c	4.0	3.5	3.5	3.3
Bromoethane	3.7 ^c	4.0	3.3	3.8	3.5
Iodoethane	3.6 ^d	4.1	3.3	3.4	3.5
Methanol	1.1 ^c	4.1	2.1	1.8	1.3
Dimethylether	2.7 ^c	4.2	3.0	2.8	2.6
rms deviations		1.1	0.2	0.4	0.4
Set 2					
1,1,1-Trifluoroethane	3.2 ^c	—	3.8	3.6	3.5
1,1,1-Trichloroethane	5.1 ^c	—	4.9	5.0	4.7
Hexafluoroethane	3.9 ^d	—	5.6	5.3	5.0
Dimethylthiol	2.1 ^c	—	2.9	1.9	1.8
Propene (CH3—CHCH2)	2.0 ^d	—	0.8	2.3	2.3
Acetaldehyde	1.1 ^c	—	0.9	1.0	1.0
Acetic acid (CH3—COOH)	0.5 ^c	—	1.0	1.3	1.1
rms deviations		—	0.9	0.6	0.5

^aAll comparisons are made with the experimental values in column 1.^bSet 1 includes molecules in the Tripos article,² while set 2 includes more examples found in the BIOGRAF³ article.^cReference 11.^dReference 12.

dral angle from the lowest energy minimized structure and calculating the energy at particular values of the dihedral angle close to the energy maximum.

Crystal packing studies on selected molecules were carried out using space groups reported in the Cambridge Crystallographic Data Base⁹ for the particular molecule. From the space group information, symmetry copies within 15 Å of any atom in the reference molecule were generated and used. Lattice parameters were held constant and the reference molecule and its symmetry copies allowed to change during energy minimization.

Energy minimization was carried out using the Newton-Raphson method for single molecules and conjugate gradient method for crystal studies. A gradient over all atoms of better than 1×10^{-6} kcal/mol-Å was obtained for all single molecule studies and 1×10^{-3} for the crystal studies. In the single molecule calculations cutoff distances were taken to be greater than the longest intramolecular distance. Crystal calculations used interaction cutoffs of 15 Å with a switching function¹ between 11-14 Å. The dielectric constant was maintained at $\epsilon = 1$ for all calculations (Tables I-III) with the exceptions noted in Table IV.

RESULTS AND DISCUSSION

Conformational Energy Differences and Torsional Barrier Heights

The Tripos (and their reported MM2 results),² DREIDING,³ and CHARMM conformational energy differences are compared to the experimental energy differences for *gauche-trans* and *axial-equatorial* conformations in Table I. This limited list of compounds was chosen for comparison purposes with other published force fields.^{2,3} The comparisons show clearly that MM2 and CHARMM results are comparable. An anomaly is found for the case of 1-butene, where the lowest energy conformation was found to occur near a dihedral angle of 120° according to recent *ab initio* results.¹⁰ Results using CHARMM and the MM2 results are in agreement with this *ab initio* data. Nevertheless, experimental values were used in computing rms differences. If the 1-butene experimental data is in error, the rms of MM2 and CHARMM will be improved.

Experimental dihedral barriers are compared with calculated results of Tripos,² DREIDING,³ and CHARMM in Table II. The data set is limited, but was chosen to allow comparison with previously published calculations.^{2,3} The newly implemented multiple periodicities of torsional terms in CHARMM results in very reasonable barriers for both geometry relaxed and rigid structures. This suggests that there are no major strain forces act-

ing across 1-4 interactions, and that it is a balance of forces that contributes to these torsional barriers. In the few cases where the differences between the relaxed and rigid values are significant, one finds real chemically significant reasons for the difference, and the release of strain energy is in these cases important. For example, in the case of *cis*-butane where two methyl groups are *cis* across the bond, one finds steric hindrance between methyl groups. The energy difference between relaxed and rigid geometries in this case is significant (4.7-6.7 kcal/mol) and the relaxation of geometry and release of strain takes place primarily by increasing the bond angles on each side of the bond of interest.

These results of Table II are also of interest for comparison purposes. For example, the BIOGRAF DREIDING³ dihedral barriers are all set to 2.0 kcal/mol for single bonds. Thus, in the first set of data shown, their results are very good. As shown in the second set of more complex molecules, where the atoms 1-4 across the bond contribute to the torsional barrier significantly, BIOGRAF has a higher rms than the CHARMM results. It is because CHARMM includes the microenvironment, scaled 1-4 nonbonded and electrostatic interactions, plus multiple periodicities for the torsional barrier terms, that the overall agreement with experiment is better than force fields missing these terms for the different types of complex molecules studied.

Internal Coordinate Geometries

Internal coordinates (bond lengths, angles, and dihedrals) obtained from energy minimization of the molecules of Tables I and II are shown in Table III. These coordinates are compared to results from MM2³⁴, MM3¹³, CFF³² calculations, and electron diffraction or microwave data (see Table III). Where possible the comparisons to CHARMM were made with the newer MM3 results of Allinger. In some cases MM2 comparisons were necessary since the newer MM3 results were not yet available in the literature. Geometries of these molecules were not available in the Tripos² or DREIDING³ articles.

The saturated and unsaturated hydrocarbons are well represented and give excellent agreement with experimental studies and compare well with MM3 results. For example, considering those structures where comparisons with MM3 and experiment can be made, the CHARMM calculated bond lengths are found to be within several thousandths of an Å (rms of 0.006 Å, 0.005 Å for MM3), while the angles are of the order of a degree (rms of 1.37°, 0.81 for MM3). Dihedral angles, which are the hardest internal coordinates to obtain correctly from calculations, are found to be calculated to within several degrees using CHARMM (rms of

Table III. Calculated CHARMm, MM2, MM3, CCF, or *ab initio* geometries and experimental gas phase electron diffraction (ED) bond lengths, bond angles, and dihedral angles of molecules of Tables I and II.

Molecule	Bond angle dihedral	CHARMm	MM2, MM3, or CCF	Experimental
Saturated hydrocarbons				
<i>n</i> -Butane	C-C(anti)	#1-2 2-3	MM3 ¹³	ED ¹⁴
	C-C(<i>g</i>)	1.532; 1.534	1.534; 1.536	1.531(ave)
	C-C-C(anti)	1.533; 1.535	1.535; 1.538	1.531(ave)
	C-C-C(<i>g</i>)	112.4	112.4	113.4(ave)
	C-C-C-C	113.3	113.7	113.4(ave)
Ethane	C-C-C-C	67.0	64.5	71
	C-C(<i>t</i>)		MM3 ¹³	ED ¹⁵
	C-C(<i>e</i>)	1.530	1.531	1.534
	C-C-H(<i>t</i>)	1.533	1.544	
2-Methylbutane	C-C-H(<i>e</i>)	111.0	111.4	111.0
	C-C-C	111.1	111.5	
		#1-2 2-3 4-3 5-3		
	C-C	1.533; 1.539; 1.535		
	C-C-C	#123; 234(<i>g</i>); 234(<i>t</i>); 435		
Neopentane		113.5; 112.1; 110.8; 110.5		
	C-C	1.538	MM3 ¹³	ED ¹⁶
	C-C		1.541	1.537
	C-C-C	1.531	MM3 ¹³	ED ¹⁷
	C-C-C	112.4	1.534	1.533
Cyclohexane			112.4	112.0
	C-C	1.534	MM3 ¹³	ED ¹⁸
	C-C-C	111.2	1.536	1.536
	C-C-C-C	55.4	110.4	111.4
	C-C-C-C		55.3	54.9
Methyl-cyclohexane(<i>e</i>)			MM3 ¹³	ED ¹⁹
	C-C	#1-2 2-3 3-4		
	C(Me)-C	1.534	1.535; 1.536; 1.540	1.536
	C-C-C	#123 234 345	1.538	1.536
	C-C-C-C	111.0; 111.2; 111.4	111.3; 111.6; 109.9	111.4; 112.1
Unsaturated hydrocarbons		55.9		54.8
	C=C	1.344	MM3 ²⁰	ED ²¹
	C-C	1.461	1.344	ED ²²
	C-C-C		1.468	(1.340)
	C-C-C	122.6	122.4	(1.461)
Butadiene				(122.9)
	C=C			ED ²³
	C-C			(1.340)
	C-C-C			(1.461)
	C-C-C			(123.6)

1-Butene

C=C(1-2)	1.345	MM3 ²⁴	ED ²⁵
C-C(2-3)	1.505	1.339	1.340
C-C(3-4)	1.531	1.506	1.502
C=C-C	123.3	1.534	1.535
C-C-C	112.3	124.4	125.6
C=C-C-C	119.6	111.9	111.7
		114.5	119.9

Phenyl-cyclohexane

C-C	#12 23 34
C-C(Ar)	1.534; 1.542
C(Ar)-C(Ar)	1.517
C-C-C	1.410; 1.405
C-C-C(Ar)	111.3 (ave)
C-C-C(Ar)	111.9
C-C(Ar)-C(Ar)	120.4
C-C-C-C	55.6
C-C-C-C(Ar)	178.9
C-C-C(Ar)-C(Ar)	62.5

Alcohols

Cyclohexanol(e)

C-C	#12 23 34	CFE ²⁶
C-O	1.535; 1.538	1.530
O-H	1.428	1.438
C-C-C	0.964	0.970
C-C-C	111.3; 110.9	110.3
C-C-O	107.7	108.6
C-O-H	108.4	
C-C-C-C	55.8; 56.4	
C-C-C-O	-176.2	
C-C-O-H	80.1	

Methanol

C-O	1.423	MM3 ²⁷	ED ²⁸	MW ²⁹
O-H	0.964	1.431	1.428	1.421
C-O-H	107.5	0.948	0.960	0.945
		108.3	109.0	108.5

Ethers

Dimethylether

C-O	1.412	MM3 ²⁷	ED ³⁰
C-O-C	114.4	1.418	1.415
Me-tilt	1.0	1.6	111.8
			3.6

Ethyl-methyl-ether

C-C	1.531t; 1.533g	MM3 ²⁷	ED ³¹
C(Me)-O	1.412t; 1.414g	1.526	1.520
C(CH ₂)-O	1.412t; 1.415g	1.418	1.413
C-O-C	114.4t; 115.1g	1.421	1.422
O-C-C	109.3t; 111.6g		111.9
C-C-O-C	79.0g		109.4
			84.0

Table III. (continued)

Molecule	Bond angle dihedral	CHARMm	MM2, MM3, or CFF	Experimental
Dimethyl-thiol	C-S C-S-C	1.821 98.5	CFF ³² 1.816 98.9	ED ³³ 1.805 99.1
Halogens				
Fluoroethane	C-C C-F C-C-F	1.530 1.371 109.4	MM ²³⁴ 1.530 1.387 109.4	ED ³⁵ 1.503 1.399 109.5
Chloroethane	C-C C-Cl C-C-Cl	1.532 1.786 111.3	MM ²³⁴ 1.532 1.787 111.1	MW ³⁴ 1.520 1.788 111.0
Bromoethane	C-C C-Br C-C-Br	1.534 1.929 110.8	MM ²³⁴ 1.533 1.949 111.3	MW ³⁴ 1.518 1.950 111.2
Iodoethane	C-C C-I C-C-I	1.531 2.169 110.8	MM ²³⁴ 1.532 2.151 111.5	ED ³⁵ 1.497 1.340 112.3
1,1,1-Trifluoroethane	C-C C-F C-C-F F-C-F	1.531 1.371 109.6 109.4		
1,1,1-Trichloroethane	C-C C-Cl C-C-Cl Cl-C-Cl	1.538 1.786 109.9 109.0		
1,4-Dichlorocyclohexane	C-C C-Cl C-C-Cl C-C-C-C C-C-C-Cl	1.535; 1.538 1.788 110.1 55.7 177.7		
Aldehyde				
Acetaldehyde	C(Me)-C C=O C(Me)-C=O	1.512 1.194 124.4	4-21G ³⁶ 1.510 1.210 124.5	ED ³⁷ 1.512 1.207 124.2

Propanal(<i>skew</i>)	C(Me)-C	1.531	4-21G ³⁶	ED ³⁸
	C-C	1.514	1.532	1.521(syn); 1.569(skew)
	C=O	1.194	1.514	1.515
	C(Me)-C-C	111.2	1.223	1.209
Acid	C-C=O	124.5	111.9	113.8(syn); 110.2(skew)
			122.9	124.5
Acetic acid	C-C	1.514	CFF ³⁹	ED ⁴⁰
	C-O	1.328	1.506	1.520
	C=O	1.190	1.337	(1.494)
	H-O	0.960	1.201	1.364
	C-C-O	112.1		(1.357)
	C-C=O	125.4		(1.209)
Amine	H-O-C	108.4	108.4	110.6
			125.9	(112.0)
				126.6
Methylamine	C-N	1.453	MM3 ⁴²	MW ⁴³
	H-N	1.020	1.463	ED ⁴⁴
	C-N-H	112.0	1.106	1.474
	H-N-H	107.0	112.3	1.014
Silane			106.4	112.1
				105.9
Methylsilane	C-Si	1.877	3-21G ⁴⁵	
	Si-H	1.484	1.896	
	C-Si-H	110.6	1.486	
Nitrate				
Nitrocyclohexane	N=O	1.210		ED ⁴⁶ (CH ₂) ₂ CHNO ₂
	C-N	1.515		1.226
	C-C	1.535; 1.537		1.517
	O=N=O	125.8		
	O=N-C	117.5		
	C-C-N	109.6		
	C-C-N=O	-61.1		
				125.4

Table IV. rms movements (Å) for: (a–f) all heavy atoms, (a, b, c) single molecules, (d) 4R dielectric, (e) crystal (CRYSTAL) environment, and (f) constrained (CONSR) dihedrals, and (g, h) rings only.

Compound	All heavy atoms (rms)						Ring heavy atoms (rms)	
	CHARMm (a)	TRIPOS (b)	DREIDING (c)	4R-Diel (d)	CRYSTAL (e)	CONSR (f)	CHARMm (g)	TRIPOS (h)
AAXTHP	0.401	0.500	0.334	0.309	0.393	0.170	0.039	0.064
ABAXES	0.312	0.123	0.112				0.046	0.072
ABINOR02	0.106	0.125	0.090				0.047	0.036
ABTOET	0.568	0.351	0.414	0.568	0.409	0.258	0.070	0.262
ABBUM010	0.089	0.124	0.113				0.086	0.122
ABZTCX	0.268	0.480	0.351				0.125	0.170
ACADOS	0.262	0.264	0.148				0.178	0.125
ACAFRLR	0.310	0.144	0.147				0.025	0.040
ACANIL01	0.133	0.250	0.070				0.020	0.010
ACARAP	0.611	0.463	0.340	0.376	0.492	0.263	0.045	0.055
ACBNZA01	0.307	0.155	0.146				0.034	0.033
ACBUOL	1.450	0.422	0.449	0.466	0.371	0.121	0.025	0.025
ACCITR10	0.305	0.091	0.265				0.066	0.074
ACDXUR	0.344	0.246	0.141				0.049	0.120
ACENAP03	0.027	0.020	0.023				0.027	0.020
ACFPCH	0.821	0.277	0.073	0.209	0.236	0.096	0.022	0.184
ACFUCN	0.330	0.425	0.275				0.026	0.122
ACGLSP	0.573	0.448	0.341	0.569	0.341	0.082	0.046	0.056
ACGLUA11	0.216	0.092	0.190				0.035	0.030
ACHGAL	0.240	0.297	0.180				0.040	0.057
ACHIST20	1.227	0.117	0.336	0.694	0.750 ^a	0.091	0.030	0.020
ACHNAP10	0.069	0.055	0.078				0.047	0.036
ACHTAR10	0.229	0.139	0.148				—	—
ACIMDC	0.028	0.019	0.023				—	—
ACINDN	0.067	0.072	0.072				0.032	0.039
ACINST	0.349	0.548	0.224				0.067	0.048
ACKYNU	1.138	0.420	0.391	0.668	0.459 ^a	0.049	0.024	0.031
ACMBPN	0.448	0.151	0.259				0.121	0.060
ACMEBZ	0.354	0.175	0.175				0.033	0.020
ACMTDE	0.564	0.175	0.254	0.257	0.293	0.046	—	—
ACNORT	0.295	0.215	0.266				0.108	0.118
ACNPAC10	0.061	0.047	0.050				0.041	0.025
ACNPEC	1.178	0.931	0.844	1.084	0.449	0.058	0.035	0.028
ACPENC10 ^b	1.192	0.511	0.288	0.795	0.553	0.145	0.087	0.201
ACPPCA	0.219	0.140	0.225				0.051	0.040
ACPRET03	0.379	0.202	0.290				0.152	0.073
ACPYNS	0.429	0.280	0.263				0.090	0.050
ACSALA01	0.312	0.075	0.097				0.029	0.018
ACSES01	0.294	0.208	0.211				0.236	0.114
ACTAND	0.317	0.257	0.379				0.054	0.070
ACTHBZ	0.439	0.466	0.297				0.077	0.220
ACTHCP	0.117	0.093	0.093				0.099	0.093
ACTOLD	0.146	0.231	0.045				0.029	0.025
ACTYSN	0.355	0.448	0.774				0.015	0.012
ACURID	0.595	0.268	0.126	0.199	0.211	0.165	0.040	0.134
ACVCHO	0.050	0.071	0.109				0.017	0.028
ACXMOL	0.800	0.412	0.408	0.432	0.334	0.127	0.148	0.088
ACXMPR	0.247	0.203	0.671				0.049	0.033
ADENOS10	0.637	0.109	0.088	0.274	0.299	0.095	0.051	0.080
Ave. dev.	0.438 (0.802) ^c	0.252 (0.338) ^c	0.233 (0.289) ^c	0.493	0.399	0.126	0.061	0.074

A total of 49 molecules were taken from the Cambridge Structural Database.⁹

^arms is average of dimer set included in unit cell.

^bRemoved improper forcing planarity on the nitrogen in the ring.

^crms includes only those molecules in column (e).

3.2°, 6.4° from MM3). This level of precision is as one should expect since the experimental results (gas phase) are average measures of a dynamic molecular situation while the energy minimized structures are static minima. Those results in which MM3 is better than CHARMm are most often those molecules where dihedral-stretch or dihedral-bend terms are responsible for the deviation. These cross-terms are missing in CHARMm, but are included in MM3.

In the case of the fluorine compounds we have taken a middle ground by not attempting to add special functions to reproduce the "fluorine effect." This effect can be described by the observation that as more fluorines are added to a carbon the C—F bond becomes significantly shorter and the C—C bond lengths vary with different fluorine occupancy. From structural results it has been observed that the C—F bond length goes from 1.399 Å for CH₂F—CH₃ to 1.326 Å in CF₃—CF₃ (see Table III for references on fluorine compounds). CHARMm uses a value intermediate between these values that gives the bond lengths shown in Table III. The C—C bond in the ethyl-halides is also modeled as an average of observed structures. The experimentally observed bond length variations, which occur as a function of number of fluorine or other halogen atoms attached, are not reproduced in these calculations.

Future comparisons will show differences between gas phase electron diffraction studies and calculated dynamically averaged structures at experimentally equivalent temperatures, as measures of the ability of the force field to simulate real molecules in their dynamic states. These studies will be described elsewhere in more detail.⁸

Cambridge Crystallographic Data Base Organic Molecules

A set of 49 organic molecules were taken from the Cambridge Crystallographic Data Base⁹ and compared to simulated results (Table IV). Molecule names are given in Table V and figures of the structures may be found in references 2 and 3 where this set was also studied. The Tripos² results are reported in columns b and h. The DREIDING³ results are given in column c. During the CHARMm calculations, the dielectric constant was taken to be 1.0 for columns a, e, f, and g, and a distance dependent form, $C = 4R$ was used and is given in column d. Each energy minimized structure from CHARMm was compared to the experimental structure using standard rms comparison techniques available in the molecular similarity module of QUANTA 3.2. All heavy atoms were used in the comparisons in columns a–f, whereas only the

heavy atoms in ring fragments were taken for the results in columns g–h.

In some cases the molecules in Table IV show large rms deviations (maximum value = 1.45 Å) in column a, indicating significant variations from the X-ray structure. In every case where unusually large rms values were found, the differences were a result of changes in dihedral angles and represent conformational rather than geometric variations. Those molecules with significant dihedral angle changes were chosen for further study to elucidate the energetic reasons for such changes. In column d, the effect of changing the dielectric from $C = 1$ to $C = 4R$ is shown. Some structures such as ACBUOL show significant improvements, going from a value of 1.450 to 0.466 rms deviation. Other structures showed only slight or no improvement in rms deviation, for example ABTOET does not change. It would appear that in some cases, the electrostatic field of the molecule is responsible for major conformational changes, while in others conformational changes occur in the vacuum calculations but are not a direct result of electrostatic forces. These molecules may have other reasons for the deviation, such as nonbonded interactions that may move groups relative to the positions found in the crystal structure.

The electrostatic driving force for conformational change in the vacuum is neutralized by other polar groups in close contact when the molecule is in the crystal environment. However, it is clear that when we remove the molecule from the crystal environment, it will change its conformation to optimize its intramolecular forces. A further study (column e) was carried out to test the effect of crystal packing on changes in molecular conformation. Those molecules with significant rms deviations in the vacuum calculation (column a) were further analyzed by building the observed crystal structures in QUANTA/CHARMm, and energy minimizing the internal coordinates of the molecules while holding the crystal lattice constants fixed. This method restricts the distances between symmetry related molecules in the crystal to be nearly constant. It is expected that by holding the crystal lattice at the measured values for the molecule of interest, interactions between symmetry generated molecules in the crystal should be closely simulated. The rms deviations (see column e) of the molecules studied in the crystal lattice are much smaller than the deviations found in the free vacuum calculations (column a) and this is in agreement with the proposal that the complete crystal environment should be simulated to retain those conformations observed from the X-ray study, for some of these flexible molecules. rms deviations from the vacuum study of just those molecules listed in column e are shown below the average

Table V. Names of molecules as they appear in Table IV.

Label	Name
AAXTHP	1,4,6-tri- <i>O</i> -acetyl-2-(<i>N</i> -acetylacetamido)-2,3-dideoxyalpha-D-threo-hex-2-enopyranose
ABAXES	2-alpha-bromo-17-beta-acetoxy-9-methyl-5-alpha,9-beta,10-alpha-estran-3-one
ABINOR02	Beta-DL-arabinose
ABTOET	Ethyl-3,7-anhydro-6,8- <i>O</i> -benzylidene-4-deoxy-2-ethylenedithio-D-talo-2-octulosonate
ABBUM010	16-Alpha,17-butanomorphinan-3-ol
ABZTCX	3-Dimethylamino-4,4-dimethyl-5,6-dihydro-4 <i>H</i> -1,2,5-benzothiazocin-6-one-1,1-dioxide
ACADOS	3'- <i>O</i> -Acetyladenosine
ACAFLR	2-Acetylaminofluorene
ACANIL01	Acetanilide
ACARAP	1,2,3,4-Tetra- <i>O</i> -acetyl-alpha-D-arabinopyranose
ACBNZA01	<i>o</i> -Acetamidobenzamide
ACBUOL	2-(2-Hydroxy-3-isopropylamino-propoxy)-5-butyrylamino-acetophenone hydrochloride
ACCITR10	Acetylcitrin
ACDXUR	Alpha-5-acetyl-2'-deoxyuridine
ACENAP03	1,2-Dihydro-acenaphthylene
ACFPCH	<i>cis</i> -1-Acetamido-2-fluoro-1-phenylcyclohexane
ACFUCN	<i>N</i> -Acetyl-furanomycin
ACGLSP	Methyl-2,3,4,5-tetra- <i>O</i> -acetyl-alpha-D-galactoseptanoside
ACGLUA11	<i>N</i> -acetyl-alpha-D-glucosamine
ACHGAL	2,3-Di- <i>O</i> -acetyl-1,6-anhydro-beta-D-galactopyranose
ACHIST20	L- <i>N</i> -Acetylhistidine monohydrate
ACHNAP10	<i>trans</i> -4a-Acetoxy-8a-chloro-1,4,4a,5,8,8a-hexahydronaphthalene
ACHTAR10	(Acetoxyethyl)-trimethylammonium hydrogen (+ -)-tartrate
ACIMDC	Acetamidinium chloride
ACINDN	2-Acetyl-indan-1,3-dione
ACINST	DL-1,4,5,6-Tetra- <i>O</i> -acetyl-3-chloro-2-C-(chloromethyl)-epi-inositol
ACKYNU	<i>N</i> -Acetyl-kynurenine
ACMBPN	2-Amino- <i>N</i> -(3-dichloromethyl-3,4,4a,5,6,7-hexahydro-5,6,8-trihydroxy-3-methyl-1-oxo-1 <i>H</i> -2-benzopyran-4-yl)-propanamide hydrobromide dihydrate
ACMEBZ	2-Acetoxy-3-methylbenzoic acid
ACMTDE	<i>N</i> -Acetyl-DL-methionine-diethylamide
ACNORT	<i>N</i> , <i>O</i> -Diacetyl-4-hydroxy-nornantenine
ACNPAC10	Acenaphthylene-1-carboxylic acid
ACNPEC	(<i>R</i> , <i>R</i>)- <i>N</i> -Acetyl- <i>S</i> -(2-nitro-1-phenylethyl)-L-cysteine
ACPENC10	(5 <i>S</i>)-Acetonyl-penem-3-carboxylate
ACPPCA	(+ -)-(<i>E</i>)- <i>N</i> -acetyl-piperidine-2-carboxylic acid
ACPRET03	21-Acetoxy-17-alpha-hydroxy-pregn-4-ene-3,11,20-trione
ACPYNS	Methyl-3,4-di- <i>O</i> -acetyl-2,6-anhydro-alpha-D-altopyranoside
ACSALA01	2-(Acetyloxy)-benzoic acid
ACSES01	(10 <i>S</i>)-17-Beta-acetoxy-3,10-cyclo-3,4-seco-4,9(11)-estradien-1-one
ACTAND	Alpha-acetylthio-5-alpha-androstan-17-one
ACTHBZ	<i>N</i> , <i>N</i> -Diacetyl-3-methylthiobenzidine
ACTHCP	7-Thia-1,3-diazabicyclo(3.3.0)octa-2,4-dione
ACTOLD	<i>p</i> -Acetotoluidine
ACTYSN	<i>N</i> -Acetyl-L-tyrosine
ACURID	Beta-5-acetyl-2'-deoxyuridine
ACVCHO	(+ -)-(1 <i>R</i> ,2 <i>R</i> ,3 <i>R</i>)-3-Acetyl-2-vinylcyclohexane-1-ol
ACXMOL	(-)-1- <i>O</i> -Acetyl-xylomollin
ACXMPR	3 <i>R</i> -(1' <i>S</i>)-Aminocarboxymethyl-2-pyrrolidone-5(<i>S</i>)-carboxylic acid
ADENOS10	Adenosine

deviations in columns a-c, to show equivalent comparisons between the same set of molecules.

To confirm that the rms deviations described above arise from conformational changes and not from incorrect geometry, the variable dihedral angles (nonrings) were constrained to the experimental values and the geometry (bond lengths and angles) optimized. The rms deviations from this study compared to experimental results are shown in column f. These results clearly confirm that the geometry of the molecules is calculated to be in

excellent agreement with experiment. In the case of ring fragments, where the effect of partial charges upon conformation would be negligible, the rms differences between CHARMM (column g) and experimental results are very small, again indicating that the internal coordinates are well represented.

Comparing the CHARMM results with the other reported force fields is not straightforward for two major reasons. First, the Tripos calculations reported were carried out without electrostatic in-

teractions²; thus they eliminated important polar contributions which affect the conformations of many of the molecules studied here. One does not expect the conformation of polar molecules to be the same in a vacuum situation as in the crystal for this very reason. Second, one must consider the effect of premature termination of the minimization process. During minimization bonds and angles respond to harmonic forces very quickly, while dihedral angles change in response to electrostatic and nonbonded interactions much more slowly. Because of this, incomplete minimization can lead to an artificially low rms deviation when comparing the calculated to the experimental structure. For example, when we studied single molecules in vacuum, using the gradient cutoff conditions used in the Tripos article² with no electrostatic contributions, we obtain an rms deviation of 0.376 for the set of molecules listed in column e of Table IV. This is clearly less than the 0.802 value found for this molecule set when using the full CHARMM force field with a low value for the gradient, but is in reasonable agreement with both the Tripos and DREIDING values in columns b and c (values in parentheses) of Table IV. We consider that a gradient over all atoms of at least 1×10^{-3} is necessary in order that the soft variables (dihedral angles) have had ample opportunity to vary from the starting structure. These variables are notoriously slow in reaching a minimum because of the very flat surface around their energy wells.

CONCLUSIONS

The CHARMM force field has been shown to be an all purpose, highly flexible, well parameterized force field for small organic molecules. Data presented here show that highly accurate geometries and conformations are obtained using this force field. Comparisons were made to experimental results on rotational barriers, conformational energy differences, and geometries. The large number of parameters routinely available through QUANTA 3.2 allows the study of new compounds with complex microenvironments. That is, many different atom type combinations for angles, dihedrals, etc., have refined parameters and are easily accessible. This allows broad molecular variability without the difficulties incurred when one defaults to generic parameters, the result of which may be unrealistic.

The diversity of types of molecules that can be studied using QUANTA/CHARMM is impressive and becoming greater every day as new partial atomic charges, atom types, bonding parameters, and nonbonded parameters are added. These studies will be reported in detail elsewhere.⁸

Simulations of cyclic-polypeptides and comparisons with X-ray results have been recently published⁶ and are not included here. The geometries and conformations produced using CHARMM on these polypeptides show extremely close correspondence to structures found experimentally from X-ray diffraction data.

References

1. (a) B.R. Gelin and M. Karplus, *Biochemistry*, **18**, 1256 (1979); (b) B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus, *J. Comp. Chem.*, **4**, 187 (1983).
2. M. Clark, R.D. Cramer III, and N. Van Opdenbosch, *J. Comp. Chem.*, **10**, 982 (1989).
3. S.L. Mayo, B.D. Olafson, and W.A. Goddard III, *J. Phys. Chem.*, **94**, 8897 (1990).
4. *Parameter Handbook*, QUANTA Release 3.0, June 1990, and *Supplement to Parameter Handbook*, QUANTA Release 3.2, Polygen Corporation, Waltham, MA, July 1991.
5. (a) L. Schäfer, C. Van Alsenoy, and J.N. Scarsdale, *J. Chem. Phys.*, **76**, 1439 (1982); (b) K. Siam, V.J. Klimkowski, C. Van Alsenoy, J.D. Ewbank, and L. Schäfer, *J. Mol. Struct.*, **152**, 261 (1987); (c) L. Schäfer, S.Q. Newton, F.A. Momany, and V.J. Klimkowski, *J. Mol. Struct.*, in press.
6. (a) J.N. Scarsdale, C. Van Alsenoy, V.J. Klimkowski, L. Schäfer, and F.A. Momany, *J. Am. Chem. Soc.*, **105**, 3438 (1983); (b) L. Schäfer, V.J. Klimkowski, F.A. Momany, H. Chuman, and C. Van Alsenoy, *Biopolymers*, **23**, 2335 (1984); (c) V.J. Klimkowski, L. Schäfer, F.A. Momany, and C. Van Alsenoy, *J. Mol. Struct.*, **124**, 143 (1985); (d) F.A. Momany, V.J. Klimkowski, and L. Schäfer, *J. Comp. Chem.*, **11**, 654 (1990); (e) L. Schäfer, S.Q. Newton, F.A. Momany, and V.J. Klimkowski, *J. Mol. Struct.*, in press; (f) F.A. Momany, V.J. Klimkowski, and L. Schäfer, *Expanding Frontiers in Polypeptide and Protein Structural Research*, V. Renugopalakrishnan, P.R. Carey, S.G. Huang, A. Storer, and I.C.P. Smith, Eds., in press.
7. (a) G. Karlstrom, P. Linse, A. Wallqvist, and B. Jonsson, *J. Am. Chem. Soc.*, **105**, 3777 (1983); (b) W.L. Jorgensen and C.J. Swenson, *J. Am. Chem. Soc.*, **107**, 569 (1985); (c) P.G. Jasien and W.J. Stevens, *J. Chem. Phys.*, **84**, 3271 (1986); (d) F.J. Lovas, R.D. Suenram, G.T. Fraser, C.W. Gillies, and J. Zozom, *J. Chem. Phys.*, **88**, 722 (1988); (e) W.L. Jorgensen and D.L. Severance, *J. Am. Chem. Soc.*, **112**, 4768 (1990); (f) C.E. Dykstra, *J. Am. Chem. Soc.*, **112**, 7540 (1990); (g) C.H. Reynolds, *J. Am. Chem. Soc.*, **112**, 7903 (1990); (h) H. Basch and W.J. Stevens, *Chem. Phys. Lett.*, **169**, 275 (1990); (i) W.L. Jorgensen and D.L. Severance, *J. Am. Chem. Soc.*, **112**, 4768 (1990); (j) S. Tsuzuki, L. Schäfer, H. Goto, E.D. Jemmes, H. Hosoya, K. Siam, K. Tanabe, and E. Osawa, *J. Am. Chem. Soc.*, **113**, 4665 (1991).
8. F.A. Momany and R. Rone, in preparation.
9. F.H. Allen, S. Bellard, M.D. Brice, B.A. Cartwright, A. Doubleday, H. Higgs, T. Hummelink, B.G. Hummelink-Peters, O. Kinnard, W.D.S. Motherwell, J.R. Rogers, and D.G. Watson, *Acta Crystallog. B*, **B35**, 2331 (1979).
10. Personal communication, Tom Halgren, Merck Sharpe and Dohm, Molecular Systems Department.
11. J.P. Lowe, *Prog. Phys. Org. Chem.*, **6**, 1 (1968).

12. K.H. Hellwege, Ed., *Molecular Constants*, Landolt-Boerstein, Group II, Springer-Verlag, New York, 1982, vol. 14, subvol. A.
13. N.L. Allinger, Y.H. Yuh, and J-Huei Lii, *J. Am. Chem. Soc.*, **111**, 8551, 8566, 8576 (1989).
14. (a) R.K. Heenan and L.S. Bartell, *J. Chem. Phys.*, **78**, 1270 (1983); (b) W.F. Bradford, S. Fitzwater, and L.S. Bartell, *J. Mol. Struct.*, **38**, 185 (1977).
15. L.S. Bartell and H.K. Higginbotham, *J. Chem. Phys.*, **42**, 851 (1965).
16. L.S. Bartell and W.F. Bradford, *J. Mol. Struct.*, **37**, 113 (1977).
17. T. Iijima, *Bull. Chem. Soc. Jpn.*, **45**, 1291 (1972).
18. (a) O. Bastiansen, L. Fernholt, H.M. Seip, H. Kambara, and K. Kuchitsu, *J. Mol. Struct.*, **18**, 163 (1973); (b) J.D. Ewbank, G. Kirsch, and L. Schafer, *J. Mol. Struct.*, **31**, 39 (1976).
19. A. Tsuboyama, A. Murayama, S. Konaka, and M. Kimura, *J. Mol. Struct.*, **118**, 351 (1984).
20. N.L. Allinger, L. Yan, and J.C. Tai, *J. Comp. Chem.*, **11**, 868 (1990).
21. W. Haugen and M. Traetteberg, *Acta Chem. Scand.*, **20**, 1726 (1966).
22. K. Kuchitsu, T. Fukuyama, and Y. Morino, *J. Mol. Struct.*, **1**, 463 (1968).
23. K. Kveseth, R. Seip, and D.A. Kohl, *Acta Chem. Scand.*, **A34**, 31 (1980).
24. N.L. Allinger, F. Li, and L. Yan, *J. Comp. Chem.*, **11**, 848 (1990).
25. D.V. Hemelrijk, L. Van der Enden, H.J. Geise, H.L. Seller, and L. Schäfer, *J. Am. Chem. Soc.*, **102**, 2189 (1980).
26. J.J.C. Teixeira-Dias and R. Fausto, *J. Mol. Struct.*, **144**, 199 (1986).
27. N.L. Allinger, M. Rahman, and J-H Lii, *J. Am. Chem. Soc.*, **112**, 8293 (1990).
28. K. Kimura and M. Kubo, *J. Chem. Phys.*, **30**, 151 (1959).
29. M.C.L. Gelly, R.M. Lees, and G. Winnewisser, *J. Mol. Spectrosc.*, **61**, 231 (1976).
30. K. Tamagawa, M. Takemura, S. Konaka, and M. Kimura, *J. Mol. Struct.*, **125**, 131 (1984).
31. K. Oyanagi and K. Kuchitsu, *Bull. Chem. Soc. Jpn.*, **51**, 2237 (1978).
32. R. Fausto, J.J.C. Teixeira-Dias, and P.R. Carey, *J. Mol. Struct.*, **159**, 137 (1987).
33. T. Iijima, S. Tsuchiya, and M. Kimura, *Bull. Chem. Soc. Jpn.*, **50**, 2564 (1977).
34. A.Y. Meyer and N.L. Allinger, *Tetrahedron*, **31**, 1971 (1975).
35. (a) B. Baegley, M.O. Jones, and M.A. Zanjanchi, *J. Mol. Struct.*, **56**, 215 (1979); (b) B. Baegley, M.O. Jones, and P. Yavari, *J. Mol. Struct.*, **71**, 203 (1981).
36. N.L. Allinger, L. Schäfer, K. Siam, V.J. Klimkowski, and C. Van Alsenoy, *J. Comp. Chem.*, **6**, 331 (1985).
37. T. Iijima and M. Kimura, *Bull. Chem. Soc. Jpn.*, **42**, 2159 (1969).
38. P. Van Nuffel, L. Van den Enden, C. Van Alsenoy, and H.J. Geise, *J. Mol. Struct.*, **116**, 99 (1984).
39. R. Fausto and J.J.C. Teixeira-Dias, *J. Mol. Struct.*, **144**, 215 (1986).
40. J.L. Derissen, *J. Mol. Struct.*, **7**, 67 (1971).
41. B.P. Van Eijck, J. Van Ophensden, M.M.M. Van Schaik, and E. Van Zoeren, *J. Mol. Spectrosc.*, **86**, 465 (1981).
42. L.R. Schmitz and N.L. Allinger, *J. Am. Chem. Soc.*, **112**, 8307 (1990).
43. D.R.J. Lide, *J. Chem. Phys.*, **27**, 343 (1957).
44. C.H. Higginbotham and L.S. Bartell, *J. Chem. Phys.*, **42**, 1131 (1965).
45. S. Grigoros and T.H. Lane, *J. Comp. Chem.*, **9**, 25 (1988).
46. I.F. Shichkiv, N.I. Sadova, L.V. Vilkov, and Y.A. Pankrushev, *Z. Strukt. Khim.*, **24**, 25 (1983).