Development and Testing of a General Amber Force Field

JUNMEI WANG,[1] ROMAIN M. WOLF,[2] JAMES W. CALDWELL, PETER A. KOLLMAN, DAVID A. CASE[3]

[1]Encysive Pharmaceuticals Inc., 7000 Fannin, Houston, Texas 77030
[2]Novartis Institutes for Biomedical Research, Basle, WSJ-88.10.14, P.O. Box, CH-4002 Basle, Switzerland
[3]Department of Molecular Biology, TPC15, The Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, California 92037

Received 26 August 2003; Accepted 16 February 2004

Abstract: We describe here a general Amber force field (GAFF) for organic molecules. GAFF is designed to be compatible with existing Amber force fields for proteins and nucleic acids, and has parameters for most organic and pharmaceutical molecules that are composed of H, C, N, O, S, P, and halogens. It uses a simple functional form and a limited number of atom types, but incorporates both empirical and heuristic models to estimate force constants and partial atomic charges. The performance of GAFF in test cases is encouraging. In test I, 74 crystallographic structures were compared to GAFF minimized structures, with a root-mean-square displacement of 0.26 Å, which is comparable to that of the Tripos 5.2 force field (0.25 Å) and better than those of MMFF 94 and CHARMm (0.47 and 0.44 Å, respectively). In test II, gas phase minimizations were performed on 22 nucleic acid base pairs, and the minimized structures and intermolecular energies were compared to MP2/6-31G* results. The RMS of displacements and relative energies were 0.25 Å and 1.2 kcal/mol, respectively. These data are comparable to results from Parm99/RESP (0.16 Å and 1.18 kcal/mol, respectively), which were parameterized to these base pairs. Test III looked at the relative energies of 71 conformational pairs that were used in development of the Parm99 force field. The RMS error in relative energies (compared to experiment) is about 0.5 kcal/mol. GAFF can be applied to wide range of molecules in an automatic fashion, making it suitable for rational drug design and database searching.

 J Comput Chem 25: 1157–1174, 2004

Key words: general AMBER force field; additive force field; force field parameterization; restrained electrostatic potential (RESP)

Introduction

Molecular mechanics force fields are a key component underlying many investigations of the protein–ligand structure for rational drug design and other tasks.[1–4] The use of empirical parameters enables them (in favorable cases) to model conformational changes and noncovalent interaction energies quite accurately. A successful force field in drug design should work well both for biological molecules and the organic molecules that interact with them. For example, the "Amber" force fields[5–8] were primarily developed for protein and nucleic acid systems. The fact that Amber only has limited parameters for organic molecules has prevented it from being widely used in drug design and other studies of ligand–protein or ligand–DNA interactions. Here, we describe a general Amber force field that works for most of the pharmaceutical molecules, and which is designed to be as compatible as possible to the traditional Amber protein force fields.

The desirability of a general purpose force field to describe a wide variety of organic molecules has been recognized before. Several widely used general force fields include MMFF94,[9] MM3,[10] MM4,[11] Tripos 5.2,[12] and the AMBER* parameters in the Macromodel program.[13,14] In addition, the OPLS force fields[15] are based on functional groups, so that extensions to many organic molecules are straightforward. For various reasons, those force fields have not been as widely used in studying biological systems as more specifically parameterized force fields such as AMBER, CHARMM, and OPLS; the latter, in turn, have no automatic way of assigning parameters for arbitrary organic molecules. There is arguably a significant advantage in developing an organics force field that is consistent in its form and parameterization with those

Correspondence to: D. A. Case; e-mail: case@scripps.edu

Contract/grant sponsor: NIH; contract/grant number: GM56531 (to P.A.K. and D.A.C.)

used for proteins and nucleic acids. Here we look at extending the AMBER force field to cover more organic species. This is possible because an extensible strategy was used for development of the biomolecular parameters, and some extension to organics has already been described.[8] Here, we describe such an extension to a much wider class of molecules.

A simple functional form, such as the so-called "Class I" model of eq. (1), is often adequate to describe the structures and nonbonded energies for organic and bioorganic systems:

$$E_{\text{pair}} = \sum_{\text{bonds}} k_r(r - r_{\text{eq}})^2 + \sum_{\text{angles}} k_\theta(\theta - \theta_{\text{eq}})^2 + \sum_{\text{dihedrals}} \frac{v_n}{2} \times [1 + \cos(n\phi - \gamma)] + \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}}\right] \quad (1)$$

Here, r_{eq} and θ_{eq} are equilibration structural parameters; K_r, K_θ, V_n are force constants; n is multiplicity and γ is the phase angle for the torsional angle parameters. The A, B, and q parameters characterize the nonbonded potentials.

Parameterization plays a crucial role in molecular mechanics. A good parameter set should reproduce the experimental data not only for the molecules in a training set, but also for molecules outside of the training set. Moreover, the parameter set should be complete and self-consistent. In GAFF, the basic parameterization philosophy, which follows ideas outlined earlier for the "Parm99" parameters,[8] is as follows. For the nonbonded part, van der Waals parameters are inherited from traditional Amber force fields directly; partial charges are assigned using a restrained electrostatic potential fit (RESP) model,[16,17] because of its clear physical picture and straightforward implementation scheme. For the internal terms [the first three terms in eq. (1)], parameterizations were first performed on bond lengths and bond angles that are weakly coupled to other parts in the energy function. Typically, equilibrium bond lengths and bond angles come from experiment and high-level *ab initio* calculations; the force constants are estimated through an empirical approach (presented below) and optimized to reproduce experimental and high-level *ab initio* vibrational frequencies.

Unlike the "hard parameters" of bond length and bond angle, torsional angle parameters are soft, and are highly coupled to the nonbonded energy terms. Torsional angle parameters are parameterized last, to cover other effects that cannot be considered in a simple functional form (such as polarization, charge transfer, and many body effects), in addition to intrinsic bond torsion preferences. In practice, torsional angle parameters are derived to reproduce the energy differences of two conformations and rotational profiles, based on experimental or high-level *ab initio* data. *Parmscan,* a force field parameterization program developed in our group,[18] was extensively used to derive the bond length, bond angle, and torsional angle parameters in this work. Some of our model molecules (such as No. 90 in Fig. 4), look strange because they are not stable species; however, such examples are necessary to make our force field complete (for the basic atom types) and make our empirical rules for deriving missing parameters work smoothly. In the following parts of this article, details are given on force field parameterization and on the performance of GAFF in the three test cases.

It is worth noting that the procedure described here requires as input a complete three-dimensional structure (including hydrogen atoms) that is close enough to an energy minimum to be optimized by standard semiempirical quantum mechanical methods. Generating such structures for libraries of potential ligands is itself a nontrivial task, and is beyond the scope of this article.

Generation of Parameters

Atom Type Definition

In molecular mechanics force fields, "similar" chemical environments are encoded as atom types, and these definitions are crucial to success and transferability. By using more atom types, one can describe subtle chemical environments more accurately, but this also leads to a bigger parameterization burden. Therefore, atom types should be used as economical as possible, and new atom types introduced only to significantly improve the force field performance.

For GAFF, we have introduced 35 *basic* atom types: five carbon, eight nitrogen, three oxygen, five sulfur, four phosphorus, six hydrogen, and one atom type for each of the four common halogens; these are listed as types 1–35 in Table 1. For these basic atom types, GAFF is a complete force field, which means all the parameters are available or can be estimated according to our empirical rules. Atom types were chosen according to the following atom properties, from more general to more specific: element type, hybridization, aromaticity, and chemical environment. For example, fluorine, chlorine, bromine, and iodine have only one atom type (element type), whereas carbon has five atom types: c3 (sp3 hybridization), c1 (sp1 hybridization), c (sp2 hybridization, chemical environment, i.e., bonded to S or O), ca (sp2 hybridization, aromatic property), and c2 (sp2 hybridization).

We have then included another 22 *special* atom types (Nos. 36–57), as described in Table 1. Specific atom types were defined mainly based on the chemical environment. The special atom types are classified into four groups, and were introduced for different purposes.

Group I (Nos. 53–57) consists five hydrogen types (h1–h5) that are attached to carbons in different chemical environments, and have different van der Waals parameters. This need for multiple van der Waals parameters for hydrogens, depending on what they are bonded to, is inherited from AMBER force fields, and its justification has been discussed earlier.[8]

Group II (Nos. 39–42) contains cx and cy (sp^3 carbons in three- and four-membered rings), and cu and cv (sp^2 carbons in three- and four-membered rings). When the Group II atom types occur in a bond angle, the equilibrium parameter is usually much different from that of the corresponding angle parameter for the general carbon types. For example, cx–cx–cx has an equilibrium bond angle of 60° compared to 109.5° for c3–c3–c3. Although the minimized structures with and without special atom types are quite similar (the dominant driving force are bond length parameters, which are similar for the special and general atom types), the calculated vibrational frequencies are quite different.

Table 1. Atom Types and Their Definitions in GAFF.

No.	Atom type	Description	No.	Atom type	Description
1	c	sp^2 carbon in C═O, C═S	2	c1	sp^1 carbon
3	c2	sp^2 carbon, aliphatic	4	c3	sp^3 carbon
5	ca	sp^2 carbon, aromatic	6	n	sp^2 nitrogen in amides
7	n1	sp^1 nitrogen	8	n2	sp^2 nitrogen with 2 subst., real double bonds
9	n3	sp^3 nitrogen with 3 subst.	10	n4	sp^3 nitrogen with 4 subst.
11	na	sp^2 nitrogen with 3 subst.	12	nh	amine nitrogen connected to aromatic rings
13	no	Nitrogen in nitro groups	14	o	sp^2 oxygen in C═O, COO^-
15	oh	sp^3 oxygen in hydroxyl groups	16	os	sp^3 oxygen in ethers and esters
17	s2	sp^2 sulfur (p═S, C═S, etc.)	18	sh	sp^3 sulfur in thiol groups
19	ss	sp^3 sulfur in —SR and S—S	20	s4	hypervalent sulfur, 3 subst.
21	s6	hypervalent sulfur, 4 subst.	22	p2	sp^2 phosphorus (C═P, etc.)
23	p3	sp^3 phosphorus, 3 subst.	24	p4	hypervalent phosphorus, 3 subst.
25	p5	hypervalent phosphorus, 4 subst.	26	hc	hydrogen on aliphatic carbon
27	ha	hydrogen on aromatic carbon	28	hn	hydrogen on nitrogen
29	ho	hydrogen on oxygen	30	hs	hydrogen on sulfur
31	hp	hydrogen on phosphorus	32	f	any fluorine
33	cl	any chlorine	34	br	any bromine
35	i	any iodine			
36	cc(cd)	inner sp^2 carbon in conjugated ring systems	37	ce(cf)	inner sp^2 carbon in conjugated chain systems
38	cp(cq)	bridge aromatic carbon in biphenyl systems	39	cu	sp^2 carbon in three-membered rings
40	cv	sp^2 carbon in four-membered rings	41	cx	sp^3 carbon in three-membered rings
42	cy	sp^3 carbon in four-membered rings	43	nb	aromatic nitrogen
44	nc(nd)	inner sp^2 nitrogen in conjugated ring systems, 2 subst.	45	ne(nf)	inner sp^2 nitrogen in conjugated chain systems, 2 subst.
46	pb	aromatic phosphorus	47	pc(pd)	inner sp^2 phosphorus in conjugated ring systems, 2 subst.
48	pe(pf)	inner sp^2 phosphorus in conjugated chain systems, 2 subst.	49	px	conjugated phosphorus, 3 subst.
50	py	conjugated phosphorus, 4 subst.	51	sx	conjugated sulfur, 3 subst.
52	sy	conjugated sulfur, 4 subst.	53	h1	hydrogen on aliphatic carbon with 1 electron-withdrawal group
54	h2	hydrogen on aliphatic carbon with 2 electron-withdrawal groups	55	h3	hydrogen on aliphatic carbon with 3 electron-withdrawal groups
56	h4	hydrogen on aromatic carbon with 1 electron-withdrawal group	57	h5	hydrogen on aromatic carbon with 2 electron-withdrawal groups

Nos. 1–35 are the basic atom types and Nos. 36–55 are the special atom types.

In *Group III* (Nos. 36, 37, 44, 45, 47, 48), cc (cd), nc (nd), and pc (pd) are inner sp^2 atoms in conjugated ring systems; ce (cf), ne (nf), and pe (pf) are inner sp^2 atom in conjugated chain systems; and cg (ch) are inner sp carbons in conjugated systems. Because both cc and cd are inner sp^2 carbons in conjugated ring systems, they are assigned according to a simple rule: both cc—cc and cd—cd are conjugated single bonds, and cc—cd is conjugated double bond. Other pairs of atom types in Group III behave in a similar fashion: bonds between the same type are single bonds, whereas bonds between different types are double bonds. Group III atoms are designed to describe alternating conjugated systems, without adding an explicit concept of bond order. The bond parameters and torsional potentials of the two bond types are different. For example, c2—ce or c2—cf or ce—cf double bonds have a v_2 of 11.2 kcal/mol, whereas the ce—ce or cf—cf single bonds have a v_2 of 5.6 kcal/mol. Figure 1 shows some examples of molecules having atom types in this group.

In *Group IV* (Nos. 38, 43, 46, 49–52), cp (cq) is the bridge atoms of two aromatic rings. A rule similar to that for conjugated systems is also applied for cp(cq): both cp—cp and cq—cq are single bond and other types (cp—cq, ca—cp, ca—cq) are aromatic bonds. Atom types nb and pb are aromatic nitrogen and phosphate; sx and sy are sulfur in conjugated sulfoxide and sulfone systems; px and py are phosphorus in conjugated phosphite and phosphate systems. Figure 1 shows some examples of molecules having atom types in Group IV. The atom types in Group IV have similar function as atom types in Group III. For example, without py, it would be very difficult

to discriminate single vs. double p—p bonds in compounds 24 and 25 in Figure 1.

Bond type information (such as bond orders) can be very helpful in classifying and discriminating among similar chemical environments. For various reasons, many force fields, including AMBER, only apply atom type information, and do not separately name or keep track of bond orders or types. To be consistent with the existing AMBER force fields and codes, we have used the sets of identical atom type pairs described above (cc/cd, cp/cq, ce/cf, etc.) instead of explicit bond orders to discriminate conjugated/aromatic single and double bonds. It is notable that although our scheme works for most of the molecules, there are still some special molecules that cannot be properly handled. We think that most of the failures happen to conjugated/aromatic rings attached to large aliphatic rings [$10 + 4n$ ($n = 0, 1, 2$) membered rings]. Figure 1(c) lists two examples of the kind of molecules for which our current scheme would fail. In our experience, such failures are only rarely encountered, but future extensions of the GAFF force field will have to consider these sorts of molecules.

We have developed an atom-type perception program, which is part of the *antechamber* suite of Amber, to assign the atom types described here, based only on an input geometry. Details of the algorithms involved will be presented in a separate article.[19]

Charges

To accurately fit conformational and nonbonded energies in a transferable fashion, one should choose consistent charge approach. The restrained electrostatic potential (RESP)[16,20] at HF/6-31G* is the default charge approach applied in the Amber protein force fields. Although RESP is expensive compared to empirical schemes such as Gasteiger charges, it has many desirable features, and allows one to use fewer torsional terms than might otherwise be required.[8] It has worked well in tests of small molecules[21,22] as well as proteins. This is the default charge scheme in GAFF parameterization. Unfortunately, the fact that this charge scheme needs to run *ab initio* optimization at the HF/6-31G* level has prevented it from being widely used in handling large numbers of molecules. In this situation, one may apply an alternative charge scheme called AM1-BCC (bond charge correction),[23,24] which is much cheaper than HF/6-31G* RESP. The basic idea of AM1-BCC is to first carry out a semiempirical AM1 calculation to get Mulliken charges, followed by a bond charge correction scheme to obtain results that are compatible with RESP charges. We use the BCC parameters derived by Jakalian et al.,[24] which are designed to make AM1-BCC charges match the electrostatic potential at the HF/6-31G* level.

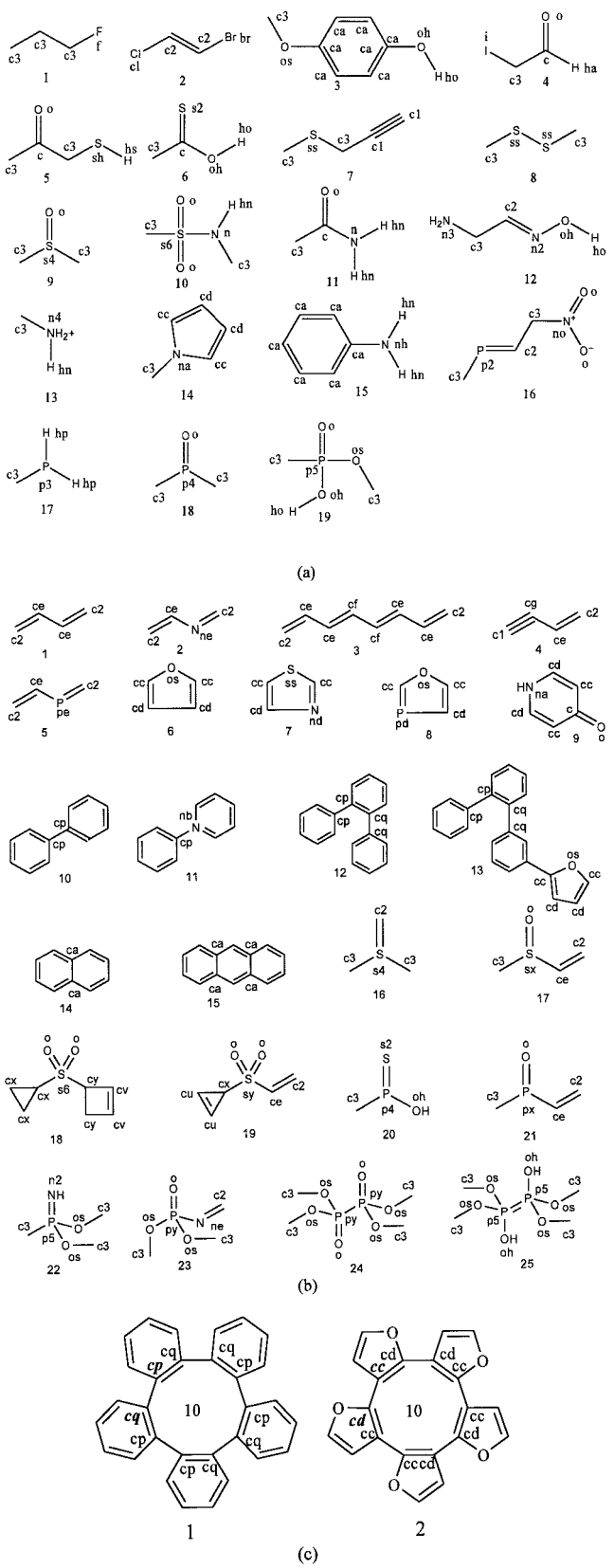

Figure 1. Example molecules that elucidate the definitions of atom types introduced in GAFF. (a) basic atom types; (b) special atom types; (c) examples of failed molecules that cannot be properly handled with our atom type scheme. In I(b), unmarked aromatic carbon in No. 11–15 have an atom type of "ca"; in I(c), atom types that causes failure are marked with bold italic font.

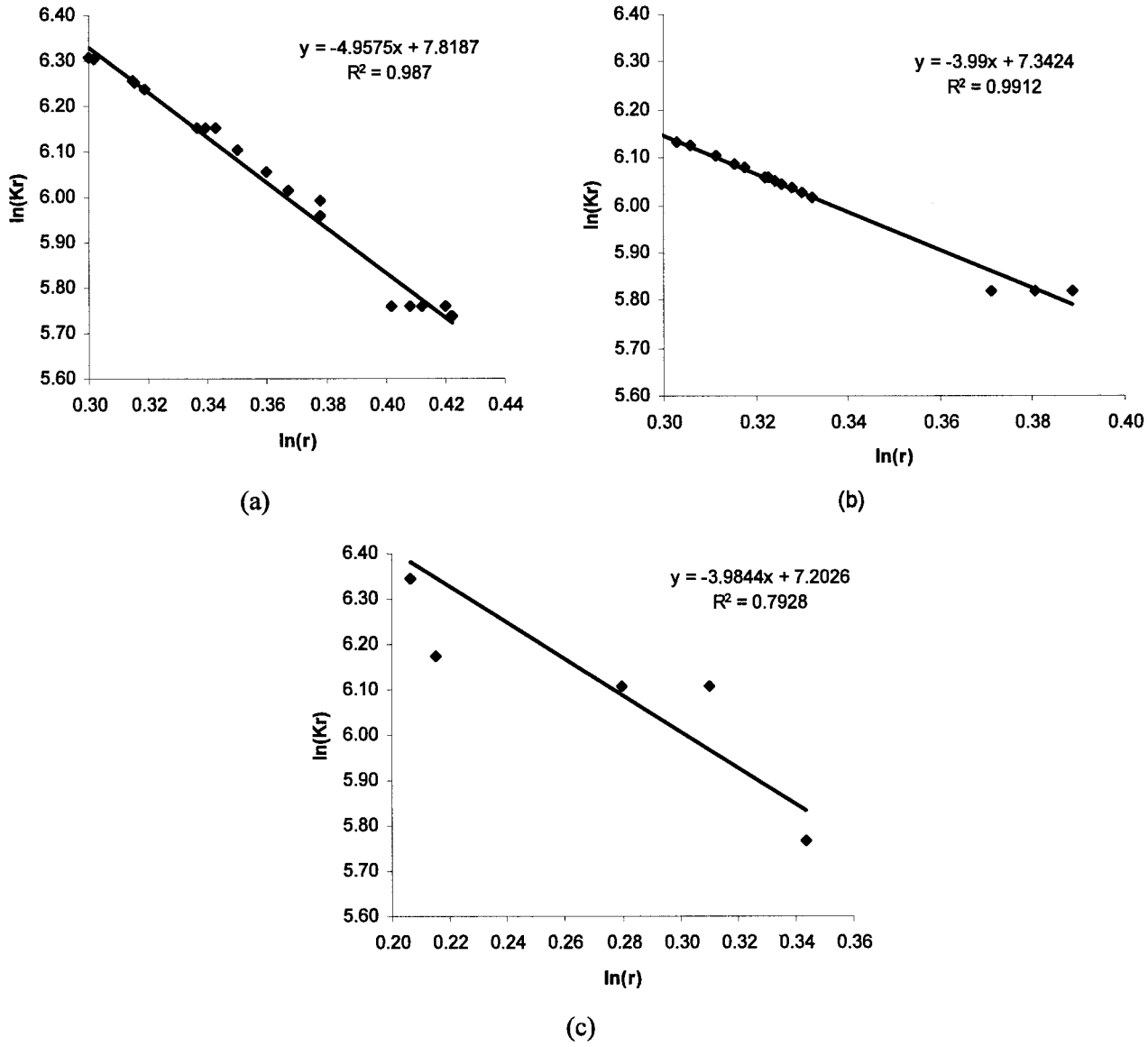

Figure 2. Least-squares fits of bond stretching force constants (K_r) and equilibrium bond lengths (r). For each bond type broadly defined by elements, n is the number of data, m is the negative of slope and c is the intercept. (a) C—C bonds (n = 25, m = 4.96, c = 7.82); (b) C—N bonds (n = 24, m = 3.99, c = 7.34); (c) C—O bonds (n = 7, m = 3.98, c = 7.20).

Lennard–Jones Parameters

The van der Waals parameters of GAFF are as same as those used by the Amber parm94 or parm99 force fields. There is only one set of parameters (the internuclear separation of the ij pair at the potential minimum, r^*_{ij}, and the potential well, ε_{ij}) for each element except hydrogen. Therefore, we believe the van der Waals parameters can be reliably transferred to new introduced atom types in GAFF.

Bond Parameters

GAFF uses three sources of information about equilibrium bond lengths r_{eq}: the Amber protein force fields, *ab initio* calculations (MP2/6-31G*), and crystal structures. Most of the experimental data comes from mean values of r_{eq} obtained from X-ray and neutron diffraction.[25,26] Data were carefully classified based on the atom types, and the most precise mean values from these three compilations were chosen as the r_{eq} in GAFF.

Table 2. Comparisons of Vibrational Frequencies.

No	Compound name	Vibrational mode	Experimental frequencies (cm^{-1})	*Ab initio* frequencies (cm^{-1})	Scaled *ab initio* frequencies (cm^{-1})	GAFF frequencies (cm^{-1})
1	H_2	H—H stretch	4401			4401
2	F_2	F—F stretch	916			917
3	Cl_2	Cl—Cl stretch	560			559
4	Br_2	Br—Br stretch	325			326
5	I_2	I—I stretch	214			215
6	CH_3—NH—NH—CH_3	N—N—C sym bend		482	458	394
7		N—N stretch		857	814	868
8		C—N stretch		1152	1094	1174
9	CH_3—O—O—CH_3	C—O torsion		273	259	273
10		O—O—C sym Bend		489	464	405
11		O—O sym Stretch		820	779	891
12		C—O sym Stretch		1094	1039	1011
13		C—O stretch + C—H asym stretch		1274	1210	1175
14	CH_3—PH—PH—CH_3	P—P—C sym Bend		256	243	295
15		P—P stretch		463	440	477
16		C—P stretch		703	668	775
17	CH_3—S—S—CH_3	S—S—C sym Bend	240	246	234	223
18		S—S—C asym Bend	272	275	261	255
19		S—S stretch	509	512	486	490
20		S—C asym Stretch	691	746	708	706
21		S—C sym stretch	694	748	710	711

Ab initio frequencies were computed at the MP_2/6-311G+(d,p) level and then scaled down by 0.9496 as suggested by Scott and Radom (Scott, A. P.; Radom, L. J Phys Chem 1996, 100, 16502). Experimental frequencies are taken from ref. 9.

As does MMFF 94,[9] GAFF applies an empirical rule to estimate the missing parameters of bond length force constants. MMFF 94 applied an inverse sixth-power dependence in Badger's formula to relate the desired force constants to tabulated reference values:

$$K_r = K_{ij}^{\text{ref}}\left(\frac{r_{ij}^{\text{ref}}}{r_{ij}}\right)^6 \tag{2}$$

Because the functional form of GAFF is different from that of MMFF 94, we experimented with a more general power law:

$$K_r = K_{ij}\left(\frac{1}{r_{ij}}\right)^m \tag{3}$$

where

$$K_{ij} = \frac{K_{ii}|r_{ij}^{\text{ref}} - r_{jj}^{\text{ref}}| + K_{jj}|r_{ij}^{\text{ref}} - r_{ii}^{\text{ref}}|}{|r_{ij}^{\text{ref}} - r_{jj}^{\text{ref}}| + |r_{ij}^{\text{ref}} - r_{ii}^{\text{ref}}|} \tag{4}$$

Here, m is the new power order for GAFF; K_r is the calculated force constant; r_{ij} is the actual bond length; K_{ij} is empirical parameter of element i and j, and it is equivalent to K_{ij}^{ref} times the sixth power of r_{ij}^{ref} in eq. (2). m and $\ln(K_{ij})$ can be parameterized through linear least-squares fitting. We performed such fittings for C—C (25 data), C—N (24 data), and C—O (7 data) bonds using the Amber bond length parameters as training set. Taking C—O as an example, 7 out of 11 C—O unique bond length parameters are represented by the following bond length types: C—O, C—O2, C—OH, C2—OH, C2—OS, C3—OH, CT—OS.

Figure 2 shows the results, with regression coefficients m of 5.0 for C—C, and 4.0 for C—N and C—O. Based on these results, a mean value of m of 4.5 was chosen. The three $\ln(K_{ij})$ constants were then refit using the final value of m.

To derive the K_{ij} parameters for X—X, (X = H, F, Cl, Br, I, N, O, and P), we have designed some model molecules and performed high-level *ab initio* vibrational frequencies analysis, as described in the Methods section. The force constants were then optimized using *Parmscan* to reproduce the vibration frequencies. For sulfur, K_{ij} of S—S was estimated using bond length parameters of disulfides in the Amber protein force fields. [If one knows r_{ij} and K_r, then K_{ij} can be computed according to eq. (3).] Table 2 lists the *ab initio* as well as GAFF vibrational frequencies for the eight model molecules and CH_3SSCH_3. For the 21 frequencies, the unsigned average error (UAE) and the root-mean-square error (RMSE) between *ab initio* and GAFF are 37 and 48 cm^{-1}, respectively. If the comparisons were made to scaled *ab initio* frequencies, the UAE and RMSE are 48 and 70 cm^{-1}.

Table 3. Parameters to Estimate the Bond Stretching Force Constants in GAFF.

No.	*i*	*j*	r_{ij}^{ref}	ln K_{ij}
1	H	H	0.738	4.661
2	C	C	1.526	7.643
3	N	N	1.441	7.634
4	O	O	1.460	7.561
5	F	F	1.406	7.358
6	Cl	Cl	2.031	8.648
7	Br	Br	2.337	9.012
8	I	I	2.836	9.511
9	P	P	2.324	8.805
10	S	S	2.038	8.316
11	H	C	1.090	6.217
12	H	N	1.010	6.057
13	H	O	0.960	5.794
14	H	F	0.920	5.600
15	H	Cl	1.280	6.937
16	H	Br	1.410	7.301
17	H	I	1.600	7.802
18	H	P	1.410	7.257
19	H	S	1.340	7.018
20	C	N	1.470	7.504
21	C	O	1.440	7.347
22	C	F	1.370	7.227
23	C	Cl	1.800	8.241
24	C	Br	1.940	8.478
25	C	I	2.160	8.859
26	C	P	1.830	8.237
27	C	S	1.820	8.117
28	N	O	1.420	7.526
29	N	F	1.420	7.475
30	N	Cl	1.750	8.266
31	N	Br	1.930	8.593
32	N	I	2.120	8.963
33	N	P	1.720	8.212
34	N	S	1.690	8.073
35	O	F	1.410	7.375
36	O	Cl	1.700	8.097
37	O	Br	1.790	8.276
38	O	I	2.110	8.854
39	O	P	1.640	7.957
40	O	S	1.650	7.922
41	F	Cl	1.648	7.947
42	Cl	I	2.550	9.309
43	Br	I	2.671	9.380
44	F	P	1.500	7.592
45	F	S	1.580	7.733
46	Cl	P	2.040	8.656
47	Cl	S	2.030	8.619
48	Br	P	2.240	8.729
49	Br	S	2.210	8.728
50	I	P	2.490	9.058
51	I	S	2.560	9.161
52	P	S	2.120	8.465

Table 4. Parameters to Estimate the Bond Angle Bending Force Constants in GAFF.

Element	C	Z
H	—	0.784
C	1.339	1.183
N	1.300	1.212
O	1.249	1.219
F	—	1.166
Cl	—	1.272
Br	—	1.378
I	—	1.398
P	0.906	1.620
S	1.448	1.280

Finally, force constants were calculated for 103 bond lengths in AMBER force field using the above model. The average percent error is 1.6%, which shows a good consistency with the Amber bond stretching parameters that were empirically fit to vibrational spectra of peptide fragments. Table 3 lists the final 47 K_{ij} parameters. This strategy of using the force constants in the AMBER protein force field as a references has the obvious limitation that the results are dependent upon the quality of the previous parameterization; on the other hand, it helps to ensure that GAFF is consistent with the AMBER parameters. Further studies are planned, aimed at improving both small molecule and protein values.

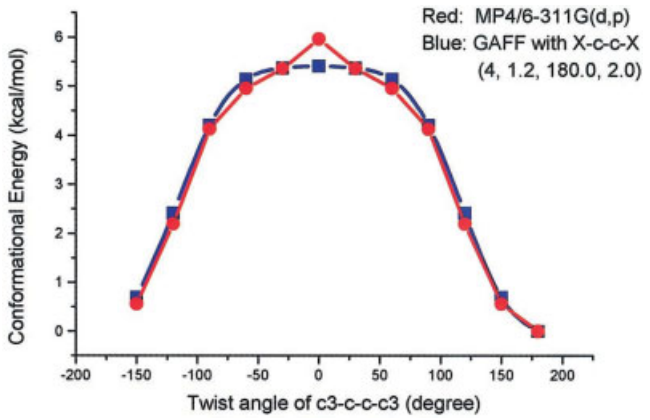

Figure 3. Principle of torsional angle parameterization using Parmscan: red and blue curves are rotational profiles produced by MP4/6-311G(d,p) and GAFF, respectively. The model compound is Molecule 1 shown in Figure 4(a) and the torsional angle in question is X–c–c–X. For this molecule, the *ab initio* torsional angle scanning was carried out from 0 to 330° with a step of 30°. With a V_2 term of 1.2 (phase angle is 180° and multiplicity is 4) kcal/mol, the two curves are well matched, and the RMSD of relative energies is 0.24 kcal/mol. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com]

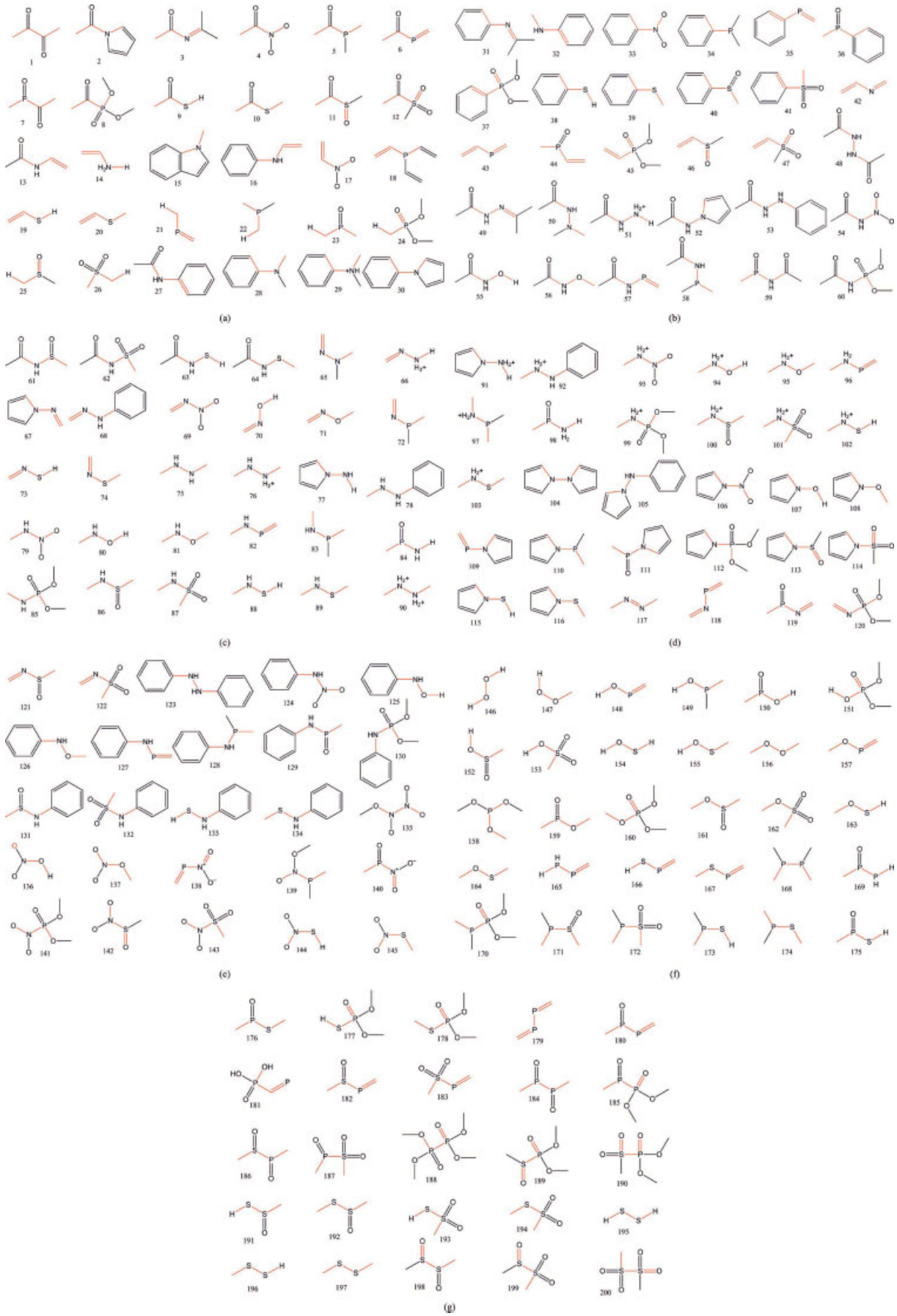

Figure 4

Angle Parameters

GAFF uses the same three sources of information (previous force fields, MP2/6-31G* calculations, and crystal data) for equilibrium bond angles as for bond lengths. Experimental data mainly comes from ref. 25. Data from all three resources were carefully classified based on the atom types, and the most precise mean values were chosen for θ_{eq}.

It is a very heavy burden to work out all the bond angle parameters. If all the combinations are considered, there are roughly 20,000 bond angle parameters only for the basic atom types, and most of them are not available. We have found an empirical approach that works well in most situations. For bond angle θ_{A-B-C}, the equilibrium bond angle is taken as the mean value of θ_{A-B-A} and θ_{C-B-C}. Therefore, if we have parameterized all bond angles of the form θ_{A-B-A}, θ_{eq} of any bond angle can be estimated from this simple rule. In total we have carried out minimizations on 1800 model molecules at the MP2/6-31G* level, with an aim to get the equilibrium angle of θ_{A-B-A} in all kinds of combinations.

As with bond lengths, GAFF applies empirical formulas to estimate bond angle force constants. A similar formula was also used by MMFF 94 to estimate the missing bond angle force constants.

$$K^{\theta}_{ijk} = 143.9 Z_i C_j Z_k (r^{eq}_{ij} + r^{eq}_{jk})^{-1} (\theta^{eq}_{ijk})^{-2} \exp(-2D) \quad (5)$$

$$D = \frac{(r^{eq}_{ij} - r^{eq}_{jk})^2}{(r^{eq}_{ij} + r^{eq}_{jk})^2} \quad (6)$$

Here, Z_i and Z_k are empirical parameters for the first and the third atoms in a angle; C_j is empirical parameter for the second atom in a angle, r^{eq}_{ij} and r^{eq}_{jk} are the equilibrium bond lengths, and θ^{eq}_{ijk} is the equilibrium bond angle. Table 4 lists the parameters of C and Z derived using 252 bond angle parameters in AMBER force fields.

With this model, we have calculated bond angle force constants for 252 bond angles in the Amber protein force field. The average percent error is 9.0%. One should realize that bond angle force constants in Amber force fields are really simple. Typically for bond angle θ_{A-B-C}, if A and C are both hydrogen atoms, the force constant is roughly 30–35 kcal/mol · rad^{-2}; if A or C is hydrogen, the force constant is roughly 50 kcal/mol · rad^{-2}; in other cases, the force constant is roughly 70 kcal/mol · rad^{-2}. As with the comparable parameterization of bond length force constants above, our procedure mainly helps to ensure a consistent pattern of bond angle distortion energies, compared to the AMBER protein force field.

Torsional Angle Parameters

In GAFF, torsional angle parameterizations were performed using the following strategy. First, torsional angle scanning was carried out at the MP4/6-311G(d,p)//MP2/6-31G* level; then *Parmscan* was applied to derive torsional angle potential V_i to reproduce the *ab initio* rotational profile. The basic idea was shown in Figure 3. The red curve is the MP4/6-311G(d,p) rotational profile and the blue one is the molecular mechanical profile. A V_2 term with force constant of 1.2 (phase angle is 180° and multiplicity of 4) makes the MM curve match the *ab initio* one perfectly except at range −15 to 15°.

In total, we have derived 200 torsional angle parameters using this strategy. Figure 4 shows the 200 model molecules and the torsional angle used for scanning was colored in red. Table 5 lists the unsigned average error as well as the RMS error for each torsional angle parameters. Overall, the average UAE and RMS error of 200 torsional angles are 0.54 and 0.72 kcal/mol, respectively.

To make parameters more transferable, it is not a good idea to include more terms just for the purpose of reducing UAE and RMS error. Typically, most of the 200 parameters have only one general term, either V_1 or V_2 or V_3. Special torsional angle parameters were added only when UAE and RMS can be significantly reduced.

Test Cases

The performance of GAFF is illustrated here with three test cases. Test I was designed to test how well GAFF predicts the molecular structures, and Tests II and III were used to test how well GAFF predicts the inter- and intramolecular energies.

Test I: Molecular Structures

In this test, we have performed minimizations for 74 molecules that have crystallographic structures. Those molecules were also studied by other widely used force field such as MMFF 94, Tripos 6.8, etc. None of the 74 molecules were in the training set to parameterize GAFF. Minimization were performed using the AM1-BCC scheme both in gas phase (dielectric constants $\varepsilon = 1$) and a medium with $\varepsilon = 4$. The minimized structures were than compared to the crystal ones. Table 6 lists the root-mean-square of displacement of the two structures, the root-mean-square deviations of bond lengths and bond angles.

For gas phase dielectric constant, a RMS displacement of 0.256 Å was achieved. This performance is comparable to that of Tripos 5.2 force field[12] (0.25 Å) and better than those of MMFF 94 and CHARMm (0.474 and 0.438 Å, respectively).[9] The root-mean-square deviation of bond length and bond angle were 0.023 Å and 2.201°, which are comparable to those of MMFF 94 (0.021 Å and 1.97°, respectively) and better than those of Tripos 5.2 (0.025 Å and 2.5°) and DREIDING (0.035 Å and 3.22°, respectively).[9]

Most of the large deviations happen in molecules containing S and P. We found the atomic charges of S and P seem too large in GAFF. Some molecules (Compounds 6, 66, and 68) give poor geometries, presumably because of so strong an electrostatic interaction in the gas phase. However, applying a larger dielectric constant can scale down the electrostatic interaction, and we found a dielectric constant of 4 works well for such situations (data not

Figure 4. Model molecules for used to derive angle parameters using Parmscan. The torsional angles in question are colored in red. (a) Molecule 1–30; (b) Molecule 31–60; (c) Molecule 61–90; (d) Molecule 91–120; (e) Molecule 121–145; (f) Molecule 146–175; (g) Molecule 176–200.

Table 5. Torsion Angle Parameterizations.

No.	Torsional angle	No. of conf.	UAE	RMSE	No.	Torsional angle	No. of conf.	UAE	RMSE
1	X–c–c–X	14	0.1649	0.2394	47	X–ce–sy–X or X–cf–sy–X	7	0.5964	0.8844
2	X–c–na–X	7	1.6509	2.6026					
3	X–c–ne–X or X–c–nf–X	7	2.3056	2.783	48	X–n–n–X	7	1.2296	1.7180
					49	X–n–n2–X	7	1.4559	1.9685
4	X–c–no–X	7	0.1759	0.2486	50	X–n–n3–X	7	0.6168	0.9310
5	X–c–p3–X	7	0.3915	0.5918	51	X–n–n4–X	7	0.3400	0.4919
6	X–c–pe–X or X–c–pf–X	7	0.8514	1.1017	52	X–n–na–X	7	0.5733	0.7916
					53	X–n–nh–X	6	0.7868	0.8980
7	X–c–px–X	7	0.071	0.0969	54	X–n–no–X	7	0.7202	0.8854
8	X–c–py–X	3	0.2970	0.3883	55	X–n–oh–X	7	1.5664	1.8988
9	X–c–sh–X c3–c–sh–hs	7	0.3682	0.5417	56	X–n–os–X	7	1.0857	1.2498
					57	X–n–p2–X	7	0.4518	0.5484
10	X–c–ss–X	7	0.5854	0.7267	58	X–n–p3–X	7	1.0356	1.5099
11	X–c–sx–X	7	1.4433	1.7906	59	X–n–p4–X	5	0.3747	0.4949
12	X–c–sy–X	5	0.3786	0.4713	60	X–n–p5–X	5	0.3143	0.4094
13	X–c2–n–X c–n–c2–c2	7	0.8837	1.2102	61	X–n–s4–X	7	0.8008	1.0519
					62	X–n–s6–X	7	0.8440	1.1573
14	X–c2–n4–X	7	0.8027	1.018	63	X–n–sh–X	7	0.6584	0.8607
15	X–c2–na–X	5	0.3833	0.6314	64	X–n–ss–X	7	0.7577	0.9103
16	X–c2–nh–X	7	0.3964	0.4654	65	X–n2–n3–X	7	1.1041	1.5665
17	X–c2–no–X	7	0.205	0.2525	66	X–n2–n4–X	5	0.2192	0.2728
18	X–c2–p3–X	7	1.4872	2.2239	67	X–n2–na–X	7	0.2445	0.2949
19	X–c2–sh–X	7	0.3988	0.5041	68	X–n2–nh–X	7	1.0014	1.2859
20	X–c2–ss–X c2–c2–ss–c3	7	0.4856	0.6310	69	X–n2–no–X	7	0.2609	0.2963
					70	X–n2–oh–X	7	0.3967	0.4742
21	X–c3–p2–X	7	0.4531	0.5784	71	X–n2–os–X	7	0.6858	0.8952
22	X–c3–p3–X	5	0.4177	0.5743	72	X–n2–p3–X	6	0.7170	0.8680
23	X–c3–p4–X	4	0.0814	0.0982	73	X–n2–sh–X	7	0.5857	0.8611
24	X–c3–p5–X	4	0.2564	0.348	74	X–n2–ss–X	7	0.7068	0.9469
25	X–c3–s4–X	7	0.0228	0.0324	75	X–n3–n3–X	6	0.6755	0.8119
26	X–c3–s6–X	7	0.0407	0.0527	76	X–n3–n4–X	6	0.3593	0.4909
27	X–ca–n–X	7	0.0877	0.1050	77	X–n3–na–X	7	0.3112	0.3781
28	X–ca–n3–X	7	0.2633	0.3230	78	X–n3–nh–X	7	0.6448	0.7711
29	X–ca–n4–X	5	0.3677	0.56	79	X–n3–no–X	3	0.6028	0.7623
30	X–ca–na–X	7	0.3222	0.4074	80	X–n3–oh–X	7	0.7214	0.8752
31	X–ca–ne–X or X–ca–nf–X	7	1.0585	1.3077	81	X–n3–os–X	7	0.8600	1.0886
					82	X–n3–p2–X	7	0.7801	1.1569
32	X–ca–nh–X	7	0.1854	0.2691	83	X–n3–p3–X	7	0.4412	0.5675
33	X–ca–no–X	7	0.098	0.1572	84	X–n3–p4–X	7	0.7148	1.3278
34	X–ca–p3–X	6	0.0485	0.0696	85	X–n3–p5–X	7	1.0947	1.3306
35	X–ca–pe–X or X–ca–pf–X	7	0.3187	0.4453	86	X–n3–s4–X	7	0.7126	0.8621
					87	X–n3–s6–X	4	0.1986	0.2672
36	X–ca–px–X	6	0.1520	0.2532	88	X–n3–sh–X	7	1.1239	1.6974
37	X–ca–py–X	7	0.1082	0.1520	89	X–n3–ss–X	7	1.0435	1.3692
38	X–ca–sh–X	7	0.1724	0.2463	90	X–n4–n4–X	7	0.5497	0.66430
39	X–ca–ss–X	7	0.0679	0.0943	91	X–n4–na–X	6	0.0845	0.1140
40	X–ca–sx–X	7	0.9675	1.1595	92	X–n4–nh–X	7	0.3832	0.4680
41	X–ca–sy–X	7	0.1924	0.2154	93	X–n4–no–X	7	0.5606	0.7750
42	X–ce–ne–X or X–cf–nf–X	7	0.7232	1.018	94	X–n4–oh–X	7	0.4798	0.5384
					95	X–n4–os–X	7	0.7807	0.9072
43	X–ce–pe–X or X–cf–pf–X	7	0.4658	0.7620	96	X–n4–p2–X	7	0.3801	0.4473
					97	X–n4–p3–X	7	0.3761	0.4583
44	X–ce–px–X or X–cf–px–X	3	0.5512	0.7518	98	X–n4–p4–X	7	0.0245	0.0344
					99	X–n4–p5–X	7	0.2688	0.3755
45	X–ce–py–X or X–cf–py–X	5	0.5660	0.7028	100	X–n4–s4–X	7	1.0123	1.3855
					101	X–n4–s6–X	7	0.7639	1.2253
46	X–ce–sx–X or X–cf–sx–X	7	0.6322	1.0069	102	X–n4–sh–X	7	0.4657	0.5475
					103	X–n4–ss–X	7	0.3767	0.4407

Table 5. (*Continued*)

No.	Torsional angle	No.of conf.	UAE	RMSE	No.	Torsional angle	No. of conf.	UAE	RMSE
104	X–na–na–X	7	0.2027	0.2281	152	X–oh–s4–X	3	0.1716	0.2180
105	X–na–nh–X	4	0.4287	0.5504	153	X–oh–s6–X	7	0.8162	0.9887
106	X–na–no–X	7	0.4227	0.6913	154	X–oh–sh–X	7	0.1271	0.1616
107	X–na–oh–X	7	0.4747	0.5694	155	X–oh–ss–X	7	0.2496	0.3248
108	X–na–os–X	7	0.1946	0.2621	156	X–os–os–X	7	0.7587	0.996
109	X–na–p2–X	7	0.2228	0.3314	157	X–os–p2–X	7	0.4006	0.5745
110	X–na–p3–X	7	0.1441	0.2363	158	X–os–p3–X	7	0.6269	0.7231
111	X–na–p4–X	3	0.3061	0.5257	159	X–os–p4–X	7	0.6449	0.8986
112	X–na–p5–X	4	0.2684	0.3100	160	X–os–p5–X	7	0.3542	0.4185
113	X–na–s4–X	7	0.3543	0.4262	161	X–os–s4–X	4	0.8268	0.9630
114	X–na–s6–X	6	0.2563	0.3208	162	X–os–s6–X	7	0.3873	0.5322
115	X–na–sh–X	6	0.3381	0.4304	163	X–os–sh–X	6	0.3849	0.4600
116	X–na–ss–X	7	0.9906	1.3332	164	X–os–ss–X	7	0.6176	0.8007
117	X–ne–ne–X or X–nf–nf–X	7	0.0633	0.0895	165	X–p2–p3–X	7	0.3468	0.4041
					166	X–p2–sh–X	7	0.2102	0.2510
118	X–ne–pe–X or X–nf–pf–X	7	0.4072	0.6282	167	X–p2–ss–X	7	0.6204	0.8506
					168	X–p3–p3–X	7	0.7789	0.9502
119	X–ne–px–X or X–nf–py–X	7	1.2603	1.5511	169	X–p3–p4–X	7	0.4028	0.4590
					170	X–p3–p5–X	7	0.6971	0.9927
120	X–ne–py–X or X–nf–py–X	7	0.6422	0.8433	171	X–p3–s4–X	6	0.6580	0.9206
					172	X–p3–s6–X	6	0.1763	0.2338
121	X–ne–sx–X or X–nf–sx–X	7	0.2744	0.3510	173	X–p3–sh–X	7	0.3161	0.3543
					174	X–p3–ss–X	7	0.5412	0.6533
122	X–ne–sy–X or X–nf–sy–X	7	0.7875	1.1072	175	X–p4–sh–X	7	0.5252	0.7549
					176	X–p4–ss–X	7	0.4878	0.7500
123	X–nh–nh–X	3	0.3533	0.4739	177	X–p5–sh–X	7	0.5499	0.7603
124	X–nh–no–X	5	0.6420	0.9939	178	X–p5–ss–X	7	0.6439	0.7348
125	X–nh–oh–X	7	0.5054	0.6130	179	X–pe–pe–X or X–pf–pf–X	7	0.4970	0.7763
126	X–nh–os–X	7	0.7320	1.1544					
127	X–nh–p2–X	7	0.5782	0.8419	180	X–pe–px–X or X–pf–px–X	7	1.1071	1.3709
128	X–nh–p3–X	5	0.3962	0.4638					
129	X–nh–p4–X	6	0.6608	0.8601	181	X–pe–py–X or X–pf–py–X	7	0.6416	0.8767
130	X–nh–p5–X	5	0.4796	0.6078					
131	X–nh–s4–X or X–nh–s4–X	7	1.1270	1.5026	182	X–pe–sx–X or X–pf–sx–X	7	0.4887	0.5953
132	X–nh–s6–X	4	0.0784	0.1091	183	X–pe–sy–X or X–pf–sy–X	7	0.1058	0.1489
133	X–nh–sh–X	7	0.6109	0.8032					
134	X–nh–ss–X	7	0.7635	1.0002	184	X–px–px–X	3	0.0857	0.1141
135	X–no–no–X	7	0.8926	1.3938	185	X–px–py–X	4	0.2617	0.3672
136	X–no–oh–X	7	0.8870	1.1501	186	X–px–sx–X	7	0.4176	0.5657
137	X–no–os–X	7	0.3754	0.4670	187	X–px–sy–X	7	0.3734	0.4452
138	X–no–p2–X	7	0.0472	0.0699	188	X–py–py–X	5	0.3407	0.4452
139	X–no–p3–X	7	0.3209	0.4286	189	X–py–sx–X	7	0.4693	0.6087
140	X–no–p4–X	4	0.2007	0.2703	190	X–py–sy–X	7	0.5907	0.9946
141	X–no–p5–X	6	0.4410	0.5300	191	X–s4–sh–X	7	0.8816	1.0561
142	X–no–s4–X	7	1.1026	1.6405	192	X–s4–ss–X	7	0.6932	0.8306
143	X–no–s6–X	7	0.1483	0.2410	193	X–s6–sh–X	4	0.3285	0.4279
144	X–no–sh–X	7	0.4371	0.5776	194	X–s6–ss–X	7	0.5587	0.6621
145	X–no–ss–X	7	0.3661	0.4709	195	X–sh–sh–X	2	0.0110	0.0156
146	X–oh–oh–X	7	0.7674	0.8761	196	X–sh–ss–X	7		
147	X–oh–os–X	7	0.7674	0.8761	197	X–ss–ss–X	7	0.6821	1.0387
148	X–oh–p2–X	7	0.3579	0.5216	198	X–sx–sx–X	4	0.7632	0.8931
149	X–oh–p3–X	7	0.5840	0.9075	199	X–sx–sy–X	7	0.6008	0.7195
150	X–oh–p4–X	7	1.3064	1.5871	200	X–sy–sy–X	7	0.3829	0.5046
151	X–oh–p5–X	6	1.3124	1.6399					

For each molecule, *ab initio* torsional angle scanning was performed at the MP4/6-311G(d,p)//MP2/6-31G* level from 0 to 180° with a step of 30° except No. 1, which was from 0 to 330 with a step of 30°. If an optimization failure happened, that conformation was taken out. UAE and RMSE are unsigned-average and root-mean-square errors of the molecular mechanical energies compared to the *ab initio* for the $n - 1$ energy pairs in kcal/mol. Here, n is the number of successfully optimized conformations listed in column 3.

Table 6. Comparisons of Crystallographic and GAFF Minimized Structures.

No.	System name	RMSDAD	RMSDBL	RMSDBA	No.	System name	RMSDAD	RMSDBL	RMSDBA
1	AAXTHP	0.189	0.025	2.320	39	ACRAMS	0.692	0.028	2.148
2	ABAXES	0.595	0.047	3.855	40	ACSALA01	0.057	0.019	2.649
3	ABBUMO10	0.075	0.018	1.414	41	ACSESO10	0.167	0.022	3.534
4	ABINOR02	0.089	0.017	2.447	42	ACTAND	0.170	0.022	0.992
5	ABTOET	0.267	0.021	2.624	43	ACTHBZ	0.308	0.010	1.505
6	ABZTCX	—	—	—	44	ACTOLD	0.098	0.013	1.196
7	ACADOS	0.146	0.024	2.367	45	ACTYSN	0.241	0.016	2.068
8	ACAFLR	0.175	0.012	1.409	46	ACURID	0.488	0.027	2.726
9	ACANIL01	0.084	0.011	0.998	47	ACVCHO	0.061	0.017	1.529
10	ACARAP	0.327	0.023	2.007	48	ACXMOL	0.338	0.024	2.039
11	ACBNZA01	0.553	0.019	2.404	49	ACXMPR	0.218	0.025	3.275
12	ACBUOL	0.204	0.020	1.779	50	ACYGLY11	0.080	0.022	2.980
13	ACCITR10	0.052	0.023	1.695	51	ACYTID	0.207	0.026	2.685
14	ACDXUR	0.195	0.028	2.550	52	ADELOX10	0.142	0.023	1.932
15	ACFPCH	0.210	0.022	1.199	53	ADENOS10	0.183	0.024	2.720
16	ACFUCN	0.451	0.022	2.470	54	ADFGLP	0.061	0.035	1.740
17	ACGLSP	0.464	0.030	2.153	55	ADGSMH	0.398	0.025	2.228
18	ACGLUA11	0.214	0.019	1.946	56	ADHELA10	0.096	0.017	2.875
19	ACHGAL	0.252	0.021	2.116	57	ADMANN	0.176	0.024	2.518
20	ACHIST20	0.902	0.041	2.517	58	ADMHEP	0.094	0.016	1.903
21	ACHNAP10	0.060	0.028	2.169	59	ADMINA	0.077	0.027	2.475
22[a]	ACHTAR10a	0.152	0.022	1.162	60	ADMOPM	0.735	0.028	3.889
23[a]	ACHTAR10b	0.405	0.029	3.775	61*	ADRTARa	0.515	0.017	1.613
24	ACIMDC	0.034	0.037	1.559	62*	ADRTARb	0.464	0.026	4.120
25	ACINDN	0.060	0.020	1.755	63	ADYPNL	0.362	0.013	1.121
26	ACINST	0.320	0.023	1.995	64	AEBDOD10	0.220	0.013	1.325
27	ACKYNU	0.452	0.022	2.109	65	AENLAN10	0.154	0.028	2.074
28	ACMBPN	0.157	0.023	2.037	66	AFCYDP	—	—	—
29	ACMEBZ	0.155	0.017	2.639	67	AFURPO10	0.099	0.025	3.342
30	ACMTDE	0.450	0.020	1.561	68	AFUTDZ10	—	—	—
31	ACNORT	0.143	0.021	2.091	69	AFUTHU	0.077	0.029	2.721
32	ACNPAC10	0.058	0.029	2.581	70	AGALAM10	0.140	0.024	1.863
33	ACNPEC	0.795	0.023	2.497	71	AGLUAM10	0.229	0.021	2.571
34	ACONTN10	0.187	0.021	1.292	72	AHARFU	0.248	0.030	2.948
35	ACPENC10	0.617	0.034	2.474	73	AHCDLA	0.056	0.022	1.775
36	ACPPCA	0.462	0.028	2.147	74	AHDITX	0.089	0.019	1.638
37	ACPRET03	0.175	0.022	1.411	Average		0.256	0.023	2.201
38	ACPYNS	0.359	0.032	2.098					

RMSDAD (root-mean-square deviation of atomic displacement) and RMSDBL (root-mean-square deviation of bond length) are in Å; RMSDBA (root-mean-square deviation of bond angle) is in degrees.
[a]For ACHTAR10 and ADRTAR, each one has two different compounds, which are further discriminated by using the system Name followed by a and b.

shown). In a condensed phase simulation, the environment might provide such a screening of electrostatic interactions. Nevertheless, future studies aimed at finding improved charges for some of these systems seem warranted.

Test II: Hydrogen-Bond Energies

In Test II, 22 base pairs were studied using GAFF with two charge schemes (AM1-BCC and RESP charge) and three dielectric constants (ε = 1, 2, and 4). Then the minimized structures and intermolecular energies were compared to those obtained at the MP2/6-31G* level. Intermolecular energies were always calculated using dielectric constant of 1. The results are very encouraging: for GAFF/RESP with the gas dielectric constant, the RMS of displacements, and relative energies (energetic differences between the MM and QM values) are 0.255 Å and 1.21 kcal/mol, respectively. These data are only marginally larger than those in Parm99/RESP, which are 0.16 Å and 1.18 kcal/mol, respectively. As to GAFF/AM1-BCC, the RMS of displacements and relative energies are still acceptable, which are 0.286 Å and 2.06 kcal/mol.

When minimizations were performed with a dielectric constant of 2, the performances were improved significantly for Parm99/RESP, GAFF/RESP, and GAFF/AM1-BCC. The RMS of displacements are 0.169, 0.194, and 0.207 Å, respectively; the RMS of relative energies

Table 7. Comparisons of Structures and Intermolecular Energies of 22 Base Pairs Between *Ab Initio* and Molecular Mechanics.

No.	Conf. name	$\Delta E_{\text{ab initio}}$ (kcal/mol)	RMSD (Å)	ΔE_{calc} (kcal/mol)	$\Delta\Delta E$ (kcal/mol)
1	aa1	−11.50	0.267	−11.17	−0.33
2	aa2	−11.00	0.192	−10.77	−0.23
3	aa3	−9.80	0.065	−10.12	0.32
4	ac2corr	−14.10	0.017	−14.14	0.04
5	atrwc	−12.40	0.145	−13.35	0.95
6	cc	−18.80	0.114	−19.51	0.71
7	ga1	−15.20	0.205	−15.36	0.16
8	ga2	−10.30	0.408	−11.13	0.83
9	ga3	−13.80	0.176	−15.67	1.87
10	ga4	−11.40	0.242	−11.66	0.26
11	gc1corr	−14.30	0.031	−15.43	1.13
12	gcwc	−25.80	0.040	−27.85	2.05
13	gg1	−24.70	0.265	−26.73	2.03
14	gg3	−17.80	0.146	−17.74	−0.07
15	gg4	−10.00	0.233	−10.57	0.57
16	gt1	−15.10	0.168	−16.44	1.34
17	gt2	−14.70	0.246	−16.44	1.74
18	tc1	−11.40	0.046	−11.95	0.55
19	tc2	−11.60	0.037	−11.97	0.37
20	tt1	−10.60	0.127	−12.33	1.73
21	tt2	−10.60	0.163	−12.03	1.43
22	tt3	−10.60	0.232	−12.69	2.09
UAE			0.162		1.18

Ab initio is at the MP_2/6-31G* Level and MM is at Parm99/RESP.

are 0.76, 0.96, and 1.18 kcal/mol, respectively. Tables 7, 8, and 9 list the RMS of displacement and relative energies for 22 base pairs.

Test III: Relative Conformational Energies

Test III was designed to test how well GAFF performs in reproducing the relative energies between two conformation pairs. In total, 71 conformational pairs were studied with the gas phase dielectric constant using GAFF. All of the 71 conformational pairs have experimental data, and they were extensively studied in the Parm99 project.[8] Table 10 lists the root-mean-square deviation of the relative energies in a gas-phase dielectric. The results were very encouraging. The RMS relative energies to experiments are 0.47 and 0.52 kcal/mol for GAFF/RESP and GAFF/AM1-BCC, respectively. This performance is only marginally larger than that of Parm99/RESP, which has an RMS relative energy of 0.44 kcal/mol. Moreover, the RMS displacement of minimized structures to the *ab initio* ones are quite small, typically smaller than 0.1 Å, and very few of them are larger than 0.2 Å.

Methods

Ab Initio *Calculations*

To parameterize the bond force constants of X—X (X = H, F, Cl, Br, I, N, O, and P), high-level *ab initio* minimization and frequency analysis were carried out for the following model molecules: H_2, F_2, Cl_2, and Br_2 at MP2/6-311+G(2d,p); I_2 at MP2/CEP-121G; CH_3OOCH_3, CH_3SSCH_3, $CH_3NHNHCH_3$ and $CH_3PHPHCH_3$ at the MP2/6-31G+(d,p) level. For the last four molecules, optimizations and ESP calculations were also performed at the HF/6-31G* level to get the RESP charges. The density of points per unit area in ESP fitting was increased by using overlay option of "iop(6/42) = 6." We found this option is necessary to get reliable RESP partial charges that are independent of the molecule's orientation and coordinates.

MP2/6-31G* optimizations were also performed for nearly 2000 model molecules, to get reliable and sufficient equilibrium bond lengths and bond angles. Most of those molecules are simple small-sized molecules that are widely used in force field developments, such as molecules listed in Table 1 of ref. 27. To make GAFF parameter set complete, we also designed a certain mount of model molecules. For instance, butane-2,3-dione was used to estimate c—c bond length and pentane-2,3,4-trione to estimate c—c—c bond angle parameter.

To develop 200 missing torsional angle parameters in GAFF, torsional angle scanning calculations were performed using the following scheme: the whole search range is 180° with a step size of 30°; for each point, minimization was performed at MP2/6-31G* level then followed by a single calculation at the MP4/6-311G(d,p) level. Similarly, HF/6-31G* optimization followed by the high-density ESP calculation with the Merz–Singh–Kollman scheme, was performed for each model molecule to derive the RESP charges.

Table 8. Comparisons of Structures and Intermolecular Energies of 22 Base Pairs between *Ab Initio* and Molecular Mechanics.

No.	Conf. name	$\Delta E_{\text{ab initio}}$ (kcal/mol)	RMSD (Å)	ΔE_{calc} (kcal/mol)	$\Delta\Delta E$ (kcal/mol)
1	aa1	−11.50	0.609	−11.76	0.26
2	aa2	−11.00	0.333	−11.11	0.11
3	aa3	−9.80	0.358	−10.29	0.49
4	ac2corr	−14.10	0.406	−15.08	0.98
5	Atrwc	−12.40	0.211	−13.13	0.73
6	Cc	−18.80	0.205	−20.14	1.34
7	ga1	−15.20	0.482	−17.54	2.34
8	ga2	−10.30	0.393	−11.72	1.42
9	ga3	−13.80	0.281	−14.18	0.38
10	ga4	−11.40	0.124	−12.11	0.71
11	gc1corr	−14.30	0.257	−16.68	2.38
12	Gcwc	−25.80	0.240	−27.64	1.84
13	gg1	−24.70	0.243	−25.47	0.77
14	gg3	−17.80	0.216	−17.97	0.17
15	gg4	−10.00	0.155	−11.16	1.16
16	gt1	−15.10	0.143	−15.82	0.72
17	gt2	−14.70	0.207	−15.62	0.92
18	tc1	−11.40	0.113	−12.00	0.60
19	tc2	−11.60	0.116	−12.01	0.41
20	tt1	−10.60	0.112	−12.18	1.58
21	tt2	−10.60	0.170	−11.82	1.22
22	Tt3	−10.60	0.241	−12.53	1.93
UAE			0.255		1.21

Ab initio is at MP_2/6-31G* level and MM is at GAFF/RESP.

All the above quantum mechanical calculations were performed using the Gaussian 98 package[28] on SGI origin 2000 computers at NCSA.

Molecular Mechanics Methods

RESP partial charges were derived by fitting the *ab initio* ESP using the Resp program implemented in the Amber7 package.[29] Default hyperbolic restraint parameters were applied for the two-stage RESP fitting. RESP two-stage input files were prepared using a program called *respgen,* which can automatically recognize the equivalent atoms. (The *respgen* program is part of the *antechamber* package.)

All minimizations in this work were also carried out using Amber7. Scale factors of 1/1.2 and 1/2 were applied to the 1–4 electrostatic and van der Waals interactions, respectively. A dielectric constant of 1 was applied except as noted. No cutoff was used for the nonbonded interactions. When the minimized structure deviated far away from the reference one (*ab initio* or crystallographic structure), a harmonic restraint was applied for torsional angles, with a force constant of 50–100 kcal/mol-rad^2. All such restraints happened in Test III: the *gauche* conformations of Compound 15 and 16, the skew conformations of Compounds 58, 78, and 79, and the twist boat and chair conformations of Compound 66.

Fitting Procedure

The procedure of the force field parameterization is to optimize a merit function χ^2, which is a sum of the square of all deviations of calculated values, y_i, from the reference values, y_i^0:

$$\chi^2 = \sum_i \omega_i^2 (y_i - y_i^0)^2 \tag{7}$$

Here, ω_i are the weighting factors, which balance the contributions of different kinds of molecular properties. In this work, molecular properties are vibrational frequencies and relative energy pairs calculated from high-level *ab initio* calculations.

To get more reliable and transferable torsional angle parameters, we only applied one generic torsional angle parameters for each model molecule, and specific parameters were added only when it can dramatically reduce the χ^2.

We applied *parmscan,*[18] an automatic force field parameter optimization program developed in our group to optimize χ^2. ω_i was set to 1, because we only consider one kind of molecular property for each optimization. Parmscan provides two optimization algorithms, systematic search, and the genetic algorithm (GA). GA works efficiently if one wants to optimize several force field parameters simultaneously. On the other hand, a systematic search is more efficient in the case of only one or two parameters being

Table 9. Comparisons of Structures and Intermolecular Energies of 22 Base Pairs between *Ab Initio* and Molecular Mechanics.

No.	Conf. name	$\Delta E_{\text{ab initio}}$ (kcal/mol)	RMSD (Å)	ΔE_{calc} (kcal/mol)	$\Delta\Delta E$ (kcal/mol)
1	aa1	−11.50	0.740	−13.25	1.75
2	aa2	−11.00	0.326	−11.55	0.55
3	aa3	−9.80	0.342	−10.01	0.21
4	ac2corr	−14.10	0.378	−15.61	1.51
5	atrwc	−12.40	0.253	−13.56	1.16
6	cc	−18.80	0.196	−21.47	2.67
7	ga1	−15.20	0.525	−16.90	1.70
8	ga2	−10.30	0.647	−14.73	4.43
9	ga3	−13.80	0.590	−18.19	4.39
10	ga4	−11.40	0.157	−13.44	2.04
11	gc1corr	−14.30	0.205	−17.74	3.44
12	gcwc	−25.80	0.191	−27.90	2.10
13	gg1	−24.70	0.304	−24.08	−0.62
14	gg3	−17.80	0.171	−17.91	0.11
15	gg4	−10.00	0.182	−13.21	3.21
16	gt1	−15.10	0.143	−15.21	0.11
17	gt2	−14.70	0.191	−14.44	−0.26
18	tc1	−11.40	0.123	−12.26	0.86
19	tc2	−11.60	0.109	−12.65	1.05
20	tt1	−10.60	0.106	−11.87	1.27
21	tt2	−10.60	0.191	−11.65	1.05
22	tt3	−10.60	0.229	−12.07	1.47
UAE			0.286		2.06

Ab initio is at MP_2/6-31G* level and MM is at GAFF/AM1-BCC.

involved in an optimization. In this work, systematic search was extensively used to derive the 200 generic torsional angle parameters V_i.

The following strategies were applied to improve the efficiency of systematic search. First of all, if the phase angle is hard to determine, the searching range was purposely set to cross 0.0. If the optimized V_i is negative, the phase angle should converse to 180° from 0° or vice versa. Second, a "focus strategy," which decreases the step size and searching range gradually, can be efficiently, located the optimal V_i from a broad searching space. Empirically, for a bond X—Y, if it is a triple bond, the force constant is usually set to 0.0 kcal/mol; if it is a double bond, a V_2 term is adequate; if it is a single bond, either or all of V_1, V_2, and V_3 may be required. In terms of the range of the unified force constant (constant divided by multiplicity), V_3 is usually from −1 to 1 kcal/mol; V_2 and V_3 are from −6 to 6 kcal/mol.

Summary and Conclusions

In this work, we have developed a general AMBER force field. We hope it is a useful molecular mechanical force field for most of the organic molecules, especially the drug-like molecules. We have based the parameterization on more than 3000 MP2/6-31G* optimizations and 1260 MP4/6-311G(d,p) single-point calculations.

GAFF uses 33 basic atom types and 22 special atom types to cover almost all the chemical space composed of H, C, N, O, S, P, F, Cl, Br, and I. For the basic atom types, all the bond length, bond angle, and torsional angle parameters are available or can be calculated with empirical rules. Special atom types were introduced to describe certain chemical environments accurately, such as conjugated single and double bonds. Unlike most conventional force fields, parameters for all combinations of atom types are not contained in an exhaustive table, but are determined algorithmically for each input molecule, based on the bonding topology (which determines the atom types) and the geometry [which is used in eqs. (2) through (6) to determine force constants]. A completely automated, table-driven procedure (called *antechamber*) is available to assign atom types, charges, and force field parameters to almost any organic molecule.[19]

From the three test cases, encouraging results were achieved, but the real value of this effort will only become clear as it is used in a variety of projects. Even given inevitable specific weaknesses, the current procedure can be used to get starting values for hand-optimization of parameters. Furthermore, as we gain experience with GAFF, improvements can be fed back into *antechamber,* so that future investigators gain the benefits of improved understanding. Even with its limitations, this general strategy has many advantages over the current situation, where the development of force fields for new ligands is often fairly subjective, even when a general outline of how to proceed is available. Furthermore, an automated procedure also makes possible systematic surveys of large numbers of ligands, as in database mining and docking.

Table 10. Comparisons of GAFF Relative Energies to Experiments for 71 Conformational Pairs.

No.	Compound name	ΔE_{expt}	GAFF/RESP				GAFF/AM1-BCC			
			E_{conf1}	E_{conf2}	ΔE_{calc}	$\Delta E_{calc} - \Delta E_{expt}$	E_{conf1}	E_{conf2}	ΔE_{calc}	$\Delta E_{calc} - \Delta E_{expt}$
1	comp1	0.75	1.27	2.06	0.80	0.05	1.42	2.23	0.81	0.06
2	comp2	5.50	5.20	11.65	6.45	0.95	5.56	12.04	6.48	0.98
3	comp3	1.75	5.79	7.28	1.48	−0.27	6.51	7.99	1.48	−0.27
4	comp4	0.05	6.59	6.51	−0.08	−0.13	5.89	5.81	−0.08	−0.13
5	comp5	1.90	19.82	21.21	1.39	−0.52	20.23	21.68	1.45	−0.45
6	comp6	1.00	23.90	25.07	1.16	0.16	24.31	25.56	1.25	0.25
7	comp7	2.87	10.95	14.06	3.11	0.24	10.43	13.75	3.32	0.45
8	comp8	2.58	10.02	11.94	1.92	−0.66	9.13	11.05	1.92	−0.66
9	comp9	5.50	4.92	10.30	5.39	−0.11	7.45	12.77	5.32	−0.18
10	comp10	5.20	12.54	15.38	2.85	−2.36	12.82	15.73	2.92	−2.28
11	comp11	4.89	1.27	6.49	5.23	0.34	1.42	6.69	5.27	0.38
12	comp12	2.88	0.09	2.96	2.88	0.00	0.42	3.30	2.88	0.00
13	comp13	3.30	0.24	3.44	3.20	−0.10	0.71	3.91	3.20	−0.10
14	comp14	3.90	3.44	7.08	3.65	−0.25	3.91	7.57	3.66	−0.24
15	comp15	2.89	2.62	5.68	3.06	0.17	1.02	3.90	2.88	−0.01
16	comp16	2.65	−1.41	0.95	2.36	−0.29	0.82	3.04	2.21	−0.44
17	comp17	0.22	6.12	5.97	−0.15	−0.37	4.04	3.98	−0.07	−0.29
18	comp18	1.20	12.13	13.20	1.07	−0.13	11.93	12.85	0.92	−0.28
19	comp19	2.00	3.80	5.89	2.10	0.10	4.77	6.89	2.12	0.12
20	comp20	0.56	2.87	3.84	0.96	0.40	2.25	3.29	1.04	0.48
21	comp21	0.35	1.21	1.18	−0.03	−0.38	1.32	1.26	−0.06	−0.41
22	comp22	0.16	4.97	5.23	0.26	0.10	5.73	5.91	0.17	0.01
23	comp23	0.59	8.25	9.18	0.93	0.34	9.41	10.61	1.20	0.61
24	comp24	1.14	5.66	6.45	0.79	−0.35	6.15	6.55	0.40	−0.74
25	comp25	1.08	2.10	2.10	0.00	−1.08	1.95	1.58	−0.38	−1.46
26	comp26	0.37	1.10	1.03	−0.07	−0.44	1.55	1.51	−0.04	−0.41
27	comp27	1.10	1.47	1.75	0.28	−0.82	1.15	1.41	0.26	−0.84
28	comp28	1.50	1.47	2.26	0.79	−0.71	1.15	1.86	0.72	−0.78
29	comp29	0.50	6.07	6.13	0.06	−0.44	6.42	6.42	0.00	−0.50
30	comp30	0.93	7.10	7.47	0.37	−0.56	8.36	8.44	0.08	−0.85
31	comp31	0.80	6.66	7.58	0.92	0.12	6.60	7.33	0.73	−0.07
32	comp32	1.04	37.48	39.22	1.74	0.70	38.20	39.96	1.76	0.72
33	comp33	0.70	8.44	8.11	−0.34	−1.04	8.74	8.46	−0.28	−0.98
34	comp34	1.50	14.02	15.51	1.49	−0.01	10.96	12.36	1.40	−0.10
35	comp35	0.88	10.92	12.27	1.35	0.47	10.67	11.86	1.19	0.31
36	comp36	0.45	−24.04	−23.29	0.76	0.31	−4.76	−4.19	0.57	0.12
37	comp37	1.15	0.24	2.11	1.88	0.73	2.04	3.61	1.57	0.42
38	comp38	0.53	5.63	5.87	0.25	−0.28	4.27	4.50	0.23	−0.31
39	comp39	3.15	5.95	9.05	3.10	−0.05	5.77	9.15	3.38	0.23
40	comp40	2.50	−7.71	−5.11	2.60	0.10	1.23	3.78	2.55	0.05
41	comp41	1.60	9.45	11.12	1.67	0.07	5.85	7.26	1.40	−0.20
42	comp42	1.93	−4.73	−3.07	1.66	−0.27	5.47	6.88	1.41	−0.52
43	comp43	1.31	10.35	11.56	1.21	−0.10	11.03	12.49	1.46	0.15
44	comp44	1.40	−0.21	2.10	2.31	0.91	−4.18	−1.95	2.23	0.83
45	comp45	2.30	−0.34	2.18	2.52	0.22	−10.12	−6.90	3.22	0.92
46	comp46	0.12	−3.54	−2.89	0.65	0.53	−0.66	−0.13	0.53	0.41
47	comp47	0.28	−13.36	−13.11	0.26	−0.02	−2.27	−1.67	0.60	0.32
48	comp48	0.18	2.61	2.21	−0.40	−0.58	1.38	1.52	0.14	−0.04
49	comp49	0.58	3.82	4.26	0.43	−0.15	4.26	4.46	0.21	−0.37
50	comp55	1.05	6.39	7.78	1.39	0.34	4.55	5.94	1.39	0.34
51	comp56	0.92	−3.18	−2.67	0.52	−0.40	−0.47	−0.51	−0.04	−0.96
52	comp57	1.50	1.11	2.29	1.19	−0.31	0.15	1.44	1.29	−0.21
53	comp58	1.70	0.57	1.66	1.09	−0.61	0.00	0.00	0.00	−1.70
54	comp59	1.14	−0.35	0.85	1.20	0.06	−1.03	0.23	1.27	0.13
55	comp60	0.55	5.81	6.04	0.23	−0.32	6.43	6.52	0.08	−0.47
56	comp61	0.00	11.35	11.69	0.34	0.34	10.99	11.32	0.33	0.33

Table 10. (*Continued*)

No.	Compound name	ΔE_{expt}	GAFF/RESP				GAFF/AM1-BCC			
			E_{conf1}	E_{conf2}	ΔE_{calc}	$\Delta E_{calc} - \Delta E_{expt}$	E_{conf1}	E_{conf2}	ΔE_{calc}	$\Delta E_{calc} - \Delta E_{expt}$
57	comp62	3.50	11.69	13.57	1.88	−1.62	11.32	13.27	1.95	−1.55
58	comp63	2.20	−9.92	−7.79	2.13	−0.07	−5.46	−3.11	2.35	0.15
59	comp64	0.85	−8.92	−7.42	1.50	0.65	−3.49	−1.98	1.51	0.66
60	comp65	0.26	−12.73	−11.10	1.63	1.37	−6.42	−5.08	1.34	1.08
61	comp66	0.22	−3.37	−4.92	−1.55	−1.77	4.50	3.19	−1.32	−1.54
62	comp67	1.20	−1.89	−1.80	0.09	−1.11	−4.65	−4.30	0.35	−0.85
63	comp68	1.80	−8.90	−7.50	1.40	−0.40	−4.31	−3.29	1.02	−0.78
64	comp72	3.90	−11.26	−6.62	4.64	0.74	−8.99	−4.91	4.08	0.18
65	comp73	1.20	1.62	3.25	1.63	0.43	4.24	6.24	2.00	0.80
66	comp74	4.75	2.78	7.40	4.61	−0.14	−0.08	4.78	4.86	0.11
67	comp76	8.50	−8.42	−1.15	7.26	−1.24	−5.68	2.17	7.85	−0.65
68	comp77	0.19	0.00	0.00	0.00	−0.19	0.00	0.00	0.00	−0.19
69	comp78	0.67	2.58	2.78	0.20	−0.47	1.73	1.86	0.13	−0.54
70	comp79	1.07	1.40	1.66	0.26	−0.81	−0.04	0.18	0.22	−0.85
71	comp80	1.70	5.37	6.19	0.82	−0.88	1.74	2.31	0.57	−1.13
UAE						0.47				0.52

Two charge schemes, RESP and AM1-BCC, were studied and the RMS deviations to experiments are 0.47 and 0.52 kcal/mol for the two charge approaches, respectively. All the experimental data come from ref. 8. All the energies are in kcal/mol.

In summary, GAFF is an extension of the AMBER force fields, and it was parameterized for most of the organic molecules. GAFF it is a complete force field, and all parameters are available for the basic atom types. We believe that the combination of GAFF with traditional AMBER offers a useful molecular mechanical tool for rational drug design and other areas where protein–ligand or DNA–ligand simulations are employed.

Machine-readable copies of all of the parameters developed here, along with the source code for the *antechamber* program, are available at http://amber.scripps.edu/antechamber.

Acknowledgments

This article is dedicated to Peter A. Kollman, who initiated the project and directed the development of GAFF prior his death in May 2001. J.M.W. is grateful to acknowledge computer time at NCSA (MCB000013N).

References

1. Charifson, P. S.; Kuntz, I. D. In Practical Application of Computer-Aided Drug Design; Charifson, P. S., Ed., Dekker: New York, 1997, 1.
2. Leach, A. R. Molecular Modelling. Principles and Applications, 2nd ed. Prentice-Hall: Harlow, England, 2001.
3. Cramer, C. J. Essentials of Computational Chemistry: Theories and Models; John Wiley & Sons: New York, 2002.
4. Kollman, P. A.; Case, D. A. In Burger's Medicinal Chemistry and Drug Discovery, 6th ed., vol. 1: Drug Discovery; Abraham, D. J., Ed.; John Wiley & Sons: New York, 2003, p. 169.
5. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. J Am Chem Soc 1984, 106, 765.
6. Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. J Comput Chem 1986, 7, 230.
7. 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.
8. Wang, J.; Cieplak, P.; Kollman, P. A. J Comput Chem 2000, 21, 1049.
9. Halgren, T. A. J Comput Chem 1996, 17, 490.
10. Lii, J.-H.; Allinger, N. J Comput Chem 1991, 12, 186.
11. Allinger, N. L. J Comput Chem 1996, 17, 642.
12. Clark, M.; Cramer, R. D., III; van Opdenbosch, N. J Comput Chem 1989, 10, 982.
13. Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. J Comput Chem 1990, 11, 440.
14. McDonald, D. Q.; Still, W. C. Tetrahed Lett 1992, 33, 7743.
15. Jorgensen, W. L. In Encyclopedia of Computational Chemistry; von R. Schleyer, P.; Allinger, N. L.; Clark, T.; Gasteiger, J.; Kollman, P. A.; H. F. Schaefer, III, Eds.; John Wiley: Chichester, 1998, p. 1986.
16. Bayly, C.; Cieplak, P.; Cornell, W.; Kollman, P. A. J Phys Chem 1993, 97, 10269.
17. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. J Am Chem Soc 1993, 115, 9620.
18. Wang, J.; Kollman, P. A. J Comput Chem 2001, 22, 1219.
19. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., in preparation.
20. Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. J Comput Chem 1995, 16, 1357.
21. Gundertofte, K.; Liljefors, T.; Norrby, P.-O.; Pettersson, I. J Comput Chem 1996, 17, 429.
22. Hobza, P.; Kabelác, M.; Sponer, J.; Mejzlik, P.; Vondrásek, J. J Comput Chem 1997, 18, 1136.
23. Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. J Comput Chem 2000, 21, 132.

24. Jakalian, A.; Jack, D. B.; Bayly, C. I. J Comput Chem 2002, 23, 1623.
25. Harmony, M. D.; Laurie, V. W.; Kuczkowski, R. L.; Schwendeman, R. H.; Ramsay, D. A.; Lovas, F. J.; Lafferty, W. J.; Maki, A. G. J Phys Chem Ref Data 1979, 8, 619.
26. Allen, F. H.; Kennard, O.; Watson, D. G.; Brammer, L.; Orpen, A. G.; Taylor, R. J Chem Soc Perkin Trans II 1987, S1.
27. Halgren, T. A. J Comput Chem 1996, 17, 490.
28. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, J. A.; Cheeseman, J. R.; Keith, T. A.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian 98 (Revision A.9); Gaussian, Inc.: Pittsburgh, PA, 1998.
29. Case, D. A.; Pearlman, D. A.; Caldwell, J. C.; Cheatham, T. E., III; Wang, J.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P.; Kollman, P. A. AMBER 7; University of California: San Francisco, 2002.