How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules?

JUNMEI WANG, PIOTR CIEPLAK,* PETER A. KOLLMAN

Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94143-0446

Received 15 September 1999; accepted 21 March 2000

ABSTRACT: In this study, we present conformational energies for a molecular mechanical model (Parm99) developed for organic and biological molecules using the restrained electrostatic potential (RESP) approach to derive the partial charges. This approach uses the simple "generic" force field model (Parm94), and attempts to add a minimal number of extra Fourier components to the torsional energies, but doing so only when there is a physical justification. The results are quite encouraging, not only for the 34-molecule set that has been studied by both the highest level *ab initio* model (GVB/LMP2) and experiment, but also for the 55-molecule set for which high-quality experimental data are available. Considering the 55 molecules studied by all the force field models for which there are experimental data, the average absolute errors (AAEs) are 0.28 (this model), 0.52 (MM3), 0.57 (CHARMm [MSI]), and 0.43 kcal/mol (MMFF). For the 34-molecule set, the AAEs of this model versus experiment and *ab initio* are 0.28 and 0.27 kcal/mol, respectively. This is a lower error than found with MM3 and CHARMm, and is comparable to that found with MMFF (0.31 and 0.22 kcal/mol). We also present two examples of how well the torsional parameters are transferred from the training set to the test set. The absolute errors of molecules in the test set are only slightly larger than in the training set (differences of <0.1 kcal/mol). Therefore, it can be concluded that a simple "generic" force field with a limited number of specific torsional parameters can describe intra- and intermolecular interactions, although all comparison molecules were selected from our 82-compound training set. We also show how this effective two-body

*Permanent affiliation: Department of Chemistry, University of Warsaw, Warsaw, Poland

Correspondence to: P. A. Kollman; e-mail: Pak@cgl.ucsf.edu; web site: http://www.amber.ucsf.edu/amber/

Contract/grant sponsor: National Institutes of Health; contract/grant numbers: IH GM-29072; NIH CA-25644, NIH GM-56609

model can be extended for use with a nonadditive force field (NAFF), both with and without lone pairs. Without changing the torsional parameters, the use of more accurate charges and polarization leads to an increase in average absolute error compared with experiment, but adjustment of the parameters restores the level of agreement found with the additive model. After reoptimizing the Ψ, Φ torsional parameters in peptides using alanine dipeptide (6 conformational pairs) and alanine tetrapeptide (11 conformational pairs), the new model gives better energies than the Cornell et al. (*J Am Chem Soc* 1995, 117, 5179–5197) force field. The average absolute error of this model for high-level *ab initio* calculation is 0.82 kcal/mol for alanine dipeptide and tetrapeptide as compared with 1.80 kcal/mol for the Cornell et al. model. For nucleosides, the new model also gives improved energies compared with the Cornell et al. model. To optimize force field parameters, we developed a program called *parmscan*, which can iteratively scan the torsional parameters in a systematic manner and finally obtain the best torsional potentials. Besides the organic molecules in our test set, *parmscan* was also successful in optimizing the Ψ, Φ torsional parameters in peptides to significantly improve agreement between molecular mechanical and high-level *ab initio* energies. J Comput Chem 21: 1049–1074, 2000

Keywords: additive force field; nonadditive force field; restrained electrostatic potential (RESP); torsional angle parameterization

Introduction

Molecular mechanics models are useful for simulations of conformational energies and noncovalent interactions of complex molecular systems. Nevertheless, they do suffer from not having as firm a physical underpinning as quantum-mechanically-based approaches. The use of empirical parameters enables them to fit experimental data better than all but the high-level and most expensive quantum-mechanical approaches. However, there is a wide variety of functional forms, as well as many different approaches for derivation of parameters for molecular mechanical methods[1–19]; therefore, it is very difficult to establish unequivocally which approach and functional form is best.

In fact, the answer to what is "best" is clearly dependent on what properties one wishes to calculate. A further consideration is the desirability to have the simplest, most extendable model consistent with a satisfactory performance. By simplest and most extendable, one means, of course, fewer atom types and fewer empirical parameters.

We hypothesize that if one's research goal is the accurate description of structures and nonbonded energies for organic and bioorganic systems, a simple functional form, such as eq. (1), is adequate. The lack of anharmonic and cross-terms is critical only when accurately fitting energies for highly strained molecules and for fitting vibrational frequencies. Our second hypothesis is that, in order to accurately fit conformational and nonbonded energies, one should use restrained electrostatic potential (RESP) charges[20–23] for the partial charge, q_i, and that choice should lead to the need for fewer torsional potentials, V_i, than in models that have an empirical scheme for derivation of q_i:

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

$$E_{\text{pol}} = -\frac{1}{2}\sum_i \mu_i E_i^0 \quad (2)$$

$$\mu_i = \alpha_i E_i \quad (3)$$

$$E_i = E_i^0 + \sum_{j \neq i} T_{ij}\mu_j \quad (4)$$

$$E_i^0 = \sum_{j \neq i} q_j \frac{\mathbf{r}_{ij}}{r_{ij}^2} \quad (5)$$

$$T_{ij} = \frac{1}{r_{ij}^3}\left(3\mathbf{r}_{ij}\frac{\mathbf{r}_{ij}}{r_{ij}^2} - 1\right) \quad (6)$$

$$E_{\text{total}} = E_{\text{pair}} + E_{\text{pol}} \quad (7)$$

In what follows, we demonstrate that these two hypotheses seem justified in our molecular mechanical model, which reproduces experimental conformational energies for a limited set of molecules better than previously derived models, albeit only marginally better than MMFF. However, given that our model is simpler in functional form than the MMFF model, it is encouraging that it performs as well.

We also show how one can extend the additive force field described by eq. (1), which includes the nonadditive effect implicitly in its use of 6-31G* RESP charges, to a nonadditive model. The nonadditive force field applies a high-level *ab initio* model to derive the partial charges, q_i. The potential energy is the sum of the pairwise additive energy and the polarization potential energy described by eqs. (2)–(7), where α_i is the polarizability of atom I; $\mathbf{r}_{ij}$ is the vector for atom I, and q_j is the charge of atom j. Eqs. (3) and (4) are iterated to self-consistency during the minimization and dynamics simulations. It is our assumption that the electrostatic energy is not tightly coupled with the torsional energy; therefore, the nonadditive model can share the same set of torsional parameters developed for the additive model. Naturally, further adjustments of the V_i torsional potential will make the nonadditive model more accurate. In this way, our RESP-based molecular mechanical model allows for natural extension to nonadditive models.

Methods

QM METHOD

The *ab initio* relative energies of equilibrium conformers listed in Table I have been published in refs. 19 and 24–26. For the 34-molecule set (nos. 1–4, 15–18, 20–21, 25–26, 36–39, 44–47, 49, 55–60, 72–74, 76, 78–80), both the MP4/TZP and GVB-LMP2 energies are available; for 1,3-dioxane and its derivatives (nos. 63–71), the conformational energies have been calculated at a lower level (MP2/6-31G*).

Alanine dipeptide and tetrapeptide have been well studied by different level *ab initio* methods,[27–30] from simple MP2/6-31G* to MP4-BSSE/cc-pVTZ (-f). The relative energies with high-level electron

TABLE I.
Conformational Energies (kcal/mol) Found in Experiment, *Ab Initio* Calculations, and Different Force Fields.

Comp. no.	Compound name and conformational comparison[a]	Expt.[b]	MP4/ TZP	GVB-LMP2	AMBER	MMFF 94	MM3	MSI CHARMm
1	Butane, g–a	0.75	0.65	0.86	0.86	0.79	0.81	0.78
2	Cyclohexane, twist boat–chair	5.50	6.14	5.85	6.58	5.93	5.76	6.72
3	Methylcyclohexane, ax–eq	1.75	1.69	1.73	1.66	1.37	1.77	1.80
4	2,3-Dimethylbutane, H—C2—C3—H, g–a	0.05	0.04	0.07	−0.15	−0.23	0.38	0.40
5	Cyclooctane, D4d–Cs boat–chair	1.90	2.00	—	1.59	1.44	1.11	0.45
6	Cyclononane, [255]C2–[333]D3	1.00	0.98	—	1.20	1.22	0.84	1.12
7	Phenylcyclohexane, ax–eq	2.87	—	—	3.25	2.99	5.14	3.28
8	*Trans*-1,2-dimethylcyclohexane, ax,ax–eq,eq	2.58	—	—	2.29	1.8	2.57	2.70
9	*Cis*-1,3-dimethylcyclohexane, ax,ax–eq,eq	5.50	—	—	5.67	5.08	5.70	4.87
10	Cyclopenta, plane-pucker	5.20	—	—	4.03	—	—	—
11	Butane, cis–trans	4.89	—	—	5.09	—	—	—
12	Ethane, eclipsed-staggered	2.88	—	—	2.93	—	—	—
13[c]	Propane, conf2–conf1	3.30	—	—	3.30	—	—	—
14[c]	Propane, conf3–conf2	3.90	—	—	3.73	—	—	—
15	1,3-Butadiene, g^9–t	2.89	2.39	2.69	3.33	2.47	1.72	2.39
16	2-Methyl-1,3-butadiene, g^9–t	2.65	2.20	2.39	2.36	2.12	1.63	1.95
17	1-Butene, cis–skew	0.22	0.26	0.22	0.18	0.26	0.68	0.46
18	2-Butene, cis–trans	1.20	1.27	1.15	1.25	1.35	1.02	1.61
19	Propene, H—C—C—C, t–c	2.00	—	—	1.96	—	—	—

TABLE I.
(Continued)

Comp. no.	Compound name and conformational comparison[a]	Expt.[b]	MP4/ TZP	GVB-LMP2	AMBER	MMFF 94	MM3	MSI CHARMm
20	1,2-Difluoroethane, a–g	0.56	0.58	0.58	0.34	0.63	0.05	1.05
21	1-Fluoropropane, a–g	0.35	0.03	0.14	−0.03	0.05	−0.10	−0.12
22	Fluorocyclohexane, ax–eq	0.16	—	—	0.46	−0.37	0.22	0.36
23	*Trans*-1,2-Difluorocyclohexane, ax,ax–eq,eq	0.59	—	—	0.31	−0.22	0.21	0.77
24	*Trans*-1,4-difluorocyclohexane, eq,eq–ax,ax	1.14	—	—	1.13	2.59	0.78	1.02
25	1,2-Dichloroethane, g–a	1.08	1.29	1.35	0.99	1.24	1.95	1.47
26	1-Chloropropane, a–g	0.37	−0.03	0.21	−0.21	0.01	−0.29	−0.15
27	1,3-Dichloropropane ga–gg	1.10	—	—	0.29	0.39	0.18	−0.24
28	1,3-Dichloropropane aa–gg	1.50	—	—	0.97	1.13	0.60	−0.16
29	Chlorocyclohexane, ax–eq	0.50	—	—	0.46	−0.35	0.58	0.53
30	*Trans*-1,2-dichlorocyclohexane, eq,eq–ax,ax	0.93	—	—	0.91	2.20	0.91	0.94
31	*Trans*-1,4-dichlorocyclohexane, eq,eq–ax,ax	0.80	—	—	0.75	2.01	−0.28	−0.09
32	1,3,5-Trineopentyl-2,4,6-tribromo benzene, two sym–all sym	1.04	—	—	1.55	2.07	1.62	0.64
33	Bromocyclohexane, ax–eq	0.70	—	—	0.11	−0.01	0.65	0.32
34	*Trans*-1,2-dibromocyclohexane, eq,eq–ax,ax	1.50	—	—	1.03	1.69	0.80	2.29
35	*Trans*-1,4-dibromocyclohexane, eq,eq–ax,ax	0.88	—	—	0.94	0.80	−0.66	−0.41
36	Iso propylamine, LP—N—C—H, a–g	0.45	0.50	0.41	0.87	0.45	0.22	−0.07
37	Cyclohexylamine, ax–eq	1.15	0.69	0.78	1.22	0.67	1.23	1.78
38	Piperidine, ax–eq	0.53	0.78	0.56	0.41	0.90	0.31	0.13
39	*N*-Methylpiperidine, ax–eq	3.15	3.58	3.38	3.19	3.29	2.30	2.70
40	2-Methylpiperidine, NH eq, Me ax–eq	2.50	—	—	2.71	2.38	2.34	2.36
41	3-Methylpiperidine, NH eq, Me ax–eq	1.60	—	—	1.87	1.09	1.48	1.36
42	4-Methylpiperidine, NH eq, ME ax–eq	1.93	—	—	1.99	1.37	1.73	2.15
43	*N*, *N*-Dimethylcyclohexane, ax Cs–eq C1	1.31	—	—	1.50	0.80	1.18	2.14
44	*N*-Methylformamide, c–t	1.40	1.04	1.35	1.74	1.09	1.50	3.44
45	*N*-Methylacetamide, c–t	2.30	1.90	2.14	2.57	1.96	2.31	2.46
46	Ethanol, g–a	0.12	−0.06	0.19	0.79	0.18	0.41	0.23
47	Isopropanol, H—C—O—H, a–g	0.28	0.20	0.28	0.32	0.17	0.68	0.02
48	Propanol, C—C—O—H, t–g	0.18	0.10	—	0.00	—	—	—
49	Cyclohexanol, ax, C1–eq, C1	0.58	0.33	0.61	0.70	0.32	0.74	0.41
50	Cyclohexanol, eq, Cs–eq, C1	—	0.18	—	0.67	0.20	—	0.06
51	Cyclohexanol, ax, Cs–eq, C1	—	1.14	—	0.01	1.01	—	0.58
52	Cyclopentaol, eq, Cs–ax, C1	—	1.11	—	0.77	0.82	—	—
53	Cyclopentaol, ax, Cs–ax, C1	—	1.05	—	0.63	0.60	—	−0.94
54	Cyclopentaol, eq, C1–ax, C1	—	1.14	—	0.18	0.47	—	−0.62
55	2-Methoxytetrahydropyrane, OCOME g, OME, eq–ax	1.05	1.30	1.22	1.11	1.81	1.62	2.37
56	2,5-Dimethyl-1,3-dioxane, 2eq, 5ax–2eq, 5eq	0.92	0.63	1.02	0.96	0.50	0.76	−0.65
57	Methylethylether, g–a	1.50	1.41	1.53	1.50	1.50	1.48	1.53
58	Methylvinylether, skew[g]–cis	1.70	2.27	2.10	1.75	2.22	0.60	2.10
59	Diethylether, g–a	1.14	1.48	1.39	1.55	1.52	1.49	1.55

TABLE I.
(Continued)

Comp. no.	Compound name and conformational comparison[a]	Expt.[b]	MP4/ TZP	GVB-LMP2	AMBER	MMFF 94	MM3	MSI CHARMm
60	Methoxycyclohexane, ax, C1–eq, C1	0.55	−0.01	0.61	0.54	0.42	0.76	0.70
61	Tetrahydrofuran, C2–Cs	0 ± 0.3	—	—	0.48	—	0.09	—
62	Tetrahydrofuran, C2v–C2	3.5	—	—	2.76	—	4.41	—
63[d,e]	1,3-Dioxane, comp3[e], twist boat–chair	2.20	2.10	—	2.19	—	5.40	—
64[d,e]	1,3-Dioxane, comp4a, twist boat–chair	0.85	1.99	—	0.93	—	3.21	—
65[d,e]	1,3-Dioxane, comp4b, twist boat–chair	0.26	1.47	—	0.56	—	3.85	—
66[d,e]	1,3-Dioxane, comp4c, twist boat[g]–chair[g]	0.22	0.99	—	−0.26	—	−2.85	—
67[d,e]	1,3-Dioxane, comp4d, twist boat–chair	1.20	0.79	—	1.78	—	−3.00	—
68[d,e]	1,3-Dioxane, comp4e, twist boat–chair	1.80	2.59	—	3.22	—	−0.01	—
69[d,e]	1,3-Dioxane, comp2, twist boat–chair	—	2.02	—	1.92	—	7.29	—
70[d,e]	1,3-Dioxane, comp1, twist boat–chair	—	4.83	—	3.90	—	5.23	—
71[d,e]	1,3-Dioxane, comp5, twist boat–chair	—	2.74	—	2.37	—	—	—
72	Formic acid, t–c	3.90	4.79	4.52	4.82	4.89	4.89	5.97
73	Glyoxylic acid, c–t	1.20	0.35	0.93	1.52	1.91	−0.49	0.89
74	Methylformate, t–c	4.75	5.65	5.09	6.12	5.28	2.18	5.54
75[f]	Methylformate, O—C—O—C, 90°[g]–0°	—	13.50	—	13.44	—	—	—
76	Methylacetate, t–c	8.50	8.21	7.91	8.02	8.27	7.85	9.09
77	Ethylformate, g–a	0.19	0.34	0.56	0.38	0.44	0.23	—
78	Propionaldehyde, skew[g]–cis	0.67	0.84	0.80	0.76	0.53	1.11	0.32
79	2-Butanone, skew[g]–cis	1.07	0.98	1.04	0.92	0.83	1.61	0.19
80	Acrolein, c–t	1.70	2.03	2.25	1.71	2.04	1.98	2.47
81[f]	Acetic acid, t–c	—	5.86	—	5.03	—	—	—
82[f]	Acetic acid, O—C—O—H, 90°[g]–0°	—	12.80	—	12.77	—	—	—

[a] For brevity, the conformational abbreviations a, g, t, and c are sometimes used for anti, gauche, trans, and cis, respectively. These designations correspond to torsional angles of 180°, 60° (or −60°), 180°, and 0°. For the skew conformation, the desired torsional angle is 120°.

[b] The experimental data are mainly from refs. 1, 19, and 24. The special sources of some data are as the following: no. 1 (ref. 42), no. 10 (refs. 43, 44), no. 11 (ref. 45), no. 12 (ref. 46), nos. 13–14 (ref. 47), no. 19 (ref. 48), no. 48 (ref. 17), nos. 63–68 (refs. 25 and 26). In total, there are six kinds of data according to the measurements: (I) ΔH, gas phase; (II) ΔE, gas phase; (III) ΔG, gas phase; (IV) ΔH, solution; (V) ΔG, solution; (VI) ΔG, solution, low temperature. Nos. 1, 4, 15–16, 19, 26, 48, 50–54, and 78 belong to class I; nos. 9, 18, 20, 22, 25, 27–30, 33, 36, 38, and 57 belong to class II; nos. 2–3, 5–6, 17, 23, 31, 34–35, 44–45, 49, 55–56, 59, 63–68, and 79 belong to class III; nos. 39, 46–47, 72–73, and 76–77 belong to class IV; nos. 8, 24, 37, 40–43, belong to class V; Nos. 7, 32, and 60 belong to class VI. For literature citation of special cases, see the Supplementary Material of ref. 19.

[c] Conf1 is the all-staggered conformation; Conf2 has the methylene group to eclipse one methyl group; Conf3 has the methylene group to eclipse both methyl groups.

[d] Compounds 1–3, 4a–4e, and 5 are defined in ref. 25.

[e] The *ab initio* method is MP2/6-31G*.

[f] The *ab initio* method is B3LYP/6-311+G(2d,p).

[g] Torsional restraint was applied.

correlation were obtained by single-point calculation for the minimized structures at the HF/6-31G* or the MP2/6-31G* level. The data (Table II) used in this work were carefully selected from refs. 27–30.

Five nucleosides were studied in this work. For each, geometry optimization was first performed at the HF/6-31G* level and then the electron correlation energy was evaluated with the same basis set at the MP2 level.

In order to develop the generic torsional parameters X—C—OS—X and X—C—OH—X (X can be any atom type), which are necessary for the studies of organic acids and esters, quantum-mechanical calculations were carried out for two model molecules, acetic acid and methylformate. Potential energy surface scans were performed using four different *ab initio* methods, HF/6-31G*, MP2/6-31G*, MP2/6-311+G (2d,p), and B3LYP/6-311+G (2d,p).

TABLE II.
Conformational Energies (kcal/mol) of Alkanes for Different Torsional Parameter Sets.[a]

Compound no.	Compound name and conformational comparison	Expt.[b]	AMBER (set I)	AMBER (set II)	AMBER (set III)
1	Butane, g–a	0.75	0.86	0.72	0.72
2	Cyclohexane, twist boat–chair	5.50	6.58	7.53	8.0
3	Methylcyclohexane, ax–eq	1.75	1.66	1.34	1.44
4	2,3-Dimethylbutane, H—C2—C3—H, g–a	0.05	−0.15	−0.38	−0.34
5	Cyclooctane, D4d–Cs boat–chair	1.90	1.59	0.13	0.51
6	Cyclononane, [255]C2–[333]D3	1.00	1.20	0.01	0.29
7	Phenylcyclohexane, ax–eq	2.87	3.25	3.22	3.22
8	*Trans*-1,2-dimethylcyclohexane, ax,ax–eq,eq	2.58	2.29	1.67	1.90
9	*Cis*-1,3-dimethylcyclohexane, ax,ax–eq,eq	5.50	5.67	4.95	5.26
10	Cyclopenta, plane-pucker	5.20	4.03	4.67	5.22
11	Butane, cis–trans	4.89	5.09	5.45	5.68
12	Ethane, eclipsed-staggered	2.88	2.93	3.03	2.88
13	Propane, conf2–conf1	3.30	3.30	3.34	3.30
14	Propane, conf3–conf2	3.90	3.73	3.77	3.90

[a] Set I: Parm99; set II: Parm94; set III: Parm99; however, for CT—CT—CT—CT, only V_3 is included, and the optimized force constant is 0.28.
[b] For sources and measure types of data, see Table I.

The relative energies of three stationary points on the energy surfaces were selected to derive the corresponding torsional parameters.

RESP CHARGES

The atom-centered point monopole charges used for the molecular-mechanical calculations were derived from the electrostatic potential. First, quantum-mechanical optimizations were performed for all the compounds using the 6-31G* basis set with the GAUSSIAN-94 software package.[31] For the AFF model, electrostatic potentials were then calculated at the same level for the minimized geometry. However, for the NAFF model, electrostatic potentials used for RESP charge fitting were calculated at a higher level (B3LYP/cc-pVTZ). For each compound, RESP charges were derived using only the electrostatic potential of the lower energy conformer.

MOLECULAR MECHANICS METHOD

All minimization calculations reported for this work were carried out using the AMBER-5 package.[32] Scale factors of 1/1.2 and 1/2 were applied to the 1–4 electrostatic and van der Waals interactions, respectively. The dielectric constant was set to 1.0 and no cutoff was used for the nonbonded interactions. When the minimized structure deviated far away from the reference one, torsional angle restraints were applied. The torsional restraint we employed is a well with a square bottom and two parabolic sides out to a defined distance, with linear sides beyond that. If R was the value of the torsional angle restraint in question, two force constants (K_1 and K_2) and four sequential torsional angles ($R_1 < R_2 < R_3 < R_4$) were used to define this flexible restraint. If R was between R_2 and R_3, no restraint was applied; if R was between R_1 and R_2, or between R_3 and R_4, restraint was exerted with a force constant of K_1; if R was smaller than R_1, or larger than R_4, restraint was applied with another force constant of K_2. In our cases, K_1 and K_2 were set to 50 kcal/mol·rad^2 and 500 kcal/mol·rad^2, respectively. The reference torsional angle was at the midpoint between R_2 and R_3, and the distances between R_1 and R_2, R_2 and R_3, and R_3 and R_4 were set to 5°. Tables I and II list the compounds and conformations for which torsional restraint was applied during the minimization.

MOLECULAR MECHANICAL ENERGIES OF OTHER FORCE FIELDS

The energy differences of the other force fields were selected from refs. 19 and 24–26.

MOLECULAR DYNAMICS SIMULATIONS

Molecular dynamics simulations of ubiquitin were carried out using the SANDER module of the AMBER-5 program[32] for both the Cornell et al. model and the new model. The simulations were carried out at 300 K with a time step of 2.0 fs. The nonbonded cutoff was set to 9.0 Å and the particle-mesh Ewald (PME) method[33] was used. A periodic box of water was added using the TIP3P potential[34] and approximately 10,000 waters were added. The dimensions of the box were 61.0, 62.0, and 54.6 Å. SHAKE[35] was applied to bonds involving hydrogen atoms.

Molecular dynamics simulations were also performed for the Dickerson dodecamer, $d(CGCGAATTCGCG)_2$,[36] in aqueous solution. A similar protocol was used for molecular dynamics simulations as that for the ubiquitin molecule; that is: a 9.0-Å nonbonded cutoff; the PME method for electrostatic interactions; the SHAKE procedure for bonds involving hydrogen atoms; a 2.0-fs time step; and the periodic box of water contained approximately 3800 TIP3P water molecules. The total time of the molecular dynamics simulation for this system was 1 ns. Two trajectories were generated. One started from the crystallographic B-form[36] and the other was initiated from the canonical A-form. The second trajectory was used to estimate the A-to-B transition time in aqueous solution for comparison with previous data obtained using the Cornell et al.[1] force field.

FITTING PROCEDURE

Parmscan, an automatic force field parameter optimization program developed by our group was used to derive new torsional parameters. Unlike other automated parameterization programs,[37–39] *parmscan* primarily attempts to find the best Fourier series and force constants so as to reproduce precisely the energy differences of the training set. The main purpose of *parmscan* is to change the force constants systematically, with a certain step, to find the optimum torsional parameters that give the smallest absolute error of molecular mechanical (MM) energy differences when compared with experimental or *ab initio* data. For each step, the file of modified force field parameters (frcmod) was first regenerated; molecular mechanics minimization was then performed for both of the two conformers of the compounds that share the same parameters. Eq. (8) or eq. (9) was applied to evaluate the score for each step, where E_{conf1} is the MM energy of the first conformer, E_{conf2} is that of the second conformer, and n is the number of conformational pairs. In this work, the scores are represented by the average absolute error as determined by eq. (8). The least-squares geometry fitting procedure was also necessary to ensure that the minimized structure did not deviate too far from the starting geometry. The root mean square of distance (RMSD) should be smaller than a certain criterion (we used 0.2 Å). If the RMSD was larger than the criterion, reoptimization with a torsional constraint from the initial structure was triggered automatically:

$$\text{Score} = \sum \text{abs}(E_{\text{conf1}} - E_{\text{conf2}})/n \tag{8}$$

$$\text{Score} = \text{sqrt}\left(\sum (E_{\text{conf1}} - E_{\text{conf2}})^2/n\right) \tag{9}$$

Results

The 82 conformational pairs examined herein are listed in Table I. They are further grouped into ten classes: alkanes (14); alkenes (3); conjugated compounds (2); fluorides (5); chlorides (7); bromides (4); amide and amino compounds (10); alcohols (9); ethers (17); and compounds with a carbonyl functional group (11). The numerals in parentheses refer to the number of conformational pairs for each class. Two special molecular sets were extracted from the whole set for comparison purposes. For the 34-molecule set (nos. 1–4, 15–18, 20–21, 25–26, 36–39, 44–47, 49, 55–60, 72–74, 76, 78–80), the experimental relative energies, high-level *ab initio* energies (GVB/LMP2 and MP4/TZP), as well as the conformational energy differences of AMBER, MMFF, MM3, and CHARMm (MSI), are all available. For the 55-molecule set (nos. 1–9, 15–18, 20–47, 49, 55–60, 72–74, 76, 78–80), both the experimental data and the energy differences of the four mentioned force fields are available.

TORSIONAL PARAMETER DEVELOPMENT

The following paragraphs describe the procedures and results of parameter development based on the functional groups. A torsional parameter is usually depicted as: $V_n(c_1, c_2)$, where V_n signifies the n-fold Fourier component, and c_1 and c_2 are the force constants (in kcal/mol) and phase angles (in degrees), respectively. For each compound mentioned in what follows, its two conformers and the experimental relative energy are given, with the second conformer, in all cases, having the more favorable energy than the first.

ALKANES

Key studies for understanding the use of quantum-mechanical electronic structure calculations for analysis of molecular torsional preferences included those by Radom et al.[40] and Brunck and Weinhold.[41] As shown in ref. 41, the single most important feature in the conformational preferences in single-bonded molecules, X—Y—Z—W, is the strength of the bond–antibond interaction, which is largest in a trans conformation, with cis next, and smallest in a skew orientation. The greater the difference in electronegativity between X and W, the stronger the stabilizing bond–antibond interactions. The bond–antibond interaction is caused in part by the electronic delocalization from the bond involving the atom with the lower electronegativity to the bond involving the atom with the higher electronegativity. This simple analysis can rationalize most of the rotational barrier in ethane (where only ~0.3 kcal/mol is due to nonbond H. . .H interactions), because staggered ethane has three trans H—C—C—H interactions and eclipsed ethane has only cis and skew H—C—C—H interactions. It can also rationalize the fact that 1,2-difluoroethane prefers the gauche to trans conformation, because the former has two trans H—C—C—F interactions and the latter has trans F—C—C—F and H—C—C—H interactions. The conceptual basis underlying the Cornell et al. force field is that one should only add extra torsional potential in the case of specific electron delocalization effects. For example, for aliphatic hydrocarbons, given the similar electronegativity of H and C, Cornell et al. showed that a single V_3 (X—C—C—X), where X = C or H, could accurately reproduce the conformational energies of ethane, propane, and butane, and the assumption was that this result would extrapolate to other hydrocarbons.

With the single V_3 parameter, the Cornell et al. model shows a reasonable performance in most of the cases. The average absolute error for 14 compounds is 0.63 kcal/mol. However, this simple model does not accurately predict the conformational preferences for cyclohexane (twist boat vs. chair, $\Delta E_{\text{twist boat–chair}} = 5.5$ kcal/mol) and cyclooctane (D4d vs. Cs boat chair, $\Delta E_{\text{D4d–Cs boat chair}} = 1.9$ kcal/mol) as well as cyclononane (C2 vs. D3, $\Delta E_{\text{C2–D3}} = 1.0$ kcal/mol). The absolute errors are 2.03, 1.77, and 0.99 kcal/mol, respectively. Therefore, a more complex model is necessary to improve the energies.

There are two ways to extend this simple model. The first approach is to break a generic torsional parameter into several parameters, each with a specific force constant. The second approach is to include other Fourier components in addition to the V_3 term. In the case of hydrocarbons, both approaches are necessary to achieve the best performance.

First, we broke the generic torsional parameter (X—C_{sp3}—C_{sp3}—X) into three specific ones: C_{sp3}—C_{sp3}—C_{sp3}—C_{sp3}; C_{sp3}—C_{sp3}—C_{sp3}—H; and H—C_{sp3}—C_{sp3}—H. Considering ethane (eclipsed vs. staggered), which involves only H—C_{sp3}—C_{sp3}—H, the force constant was set directly to 0.15 kcal/mol (V_3), which can predict the energy difference (2.93 kcal/mol) as essentially the same as that of experiment (2.88 kcal/mol). Then the force constant of C_{sp3}—C_{sp3}—C_{sp3}—H was determined by fitting the energy differences of propane (Conf1 vs. Conf2 and Conf2 vs. Conf3, with the definitions of Conf1, Conf2, and Conf3 described in Table I), which involves only C_{sp3}—C_{sp3}—C_{sp3}—H and H—C_{sp3}—C_{sp3}—H torsions. The final parameter was set to 0.16 kcal/mol (V_3), which can reproduce the two energy differences very well. Finally, the other 11 molecules that involve all of the three parameters were used to fit the force constant of C_{sp3}—C_{sp3}—C_{sp3}—C_{sp3} with *parmscan*. We found that the force constant of 0.25 kcal/mol (V_3) gave the minimum average absolute error, which is 0.54 kcal/mol. Although this model is a great improvement over the Cornell et al. model, it cannot predict the energy differences of the three aforementioned cases very well. Inspired by MM3, which uses all of V_1, V_2, and V_3 for C_{sp3}—C_{sp3}—C_{sp3}—C_{sp3}, we performed a systematic search for the 11 compounds to find the best force constants of the three Fourier components. *Parmscan* suggested that the force constants be 0.20, 0.25, and 0.18 kcal/mol for V_1, V_2, and V_3, respectively. Using this model, the AAE is reduced to 0.38 kcal/mol and the RMS deviation is only 0.47 kcal/mol. Moreover, it leads to a much better representation of the energies of the three cycloparaffins. The absolute errors are 1.08, 0.31, and 0.20 kcal/mol for cyclohexane, cyclooctane, and cyclononane, respectively. Although we made our model slightly more complex, it is encouraging that the V_3 term for all three torsional types are nearly identical, and the V_1 and V_2 terms for C_{sp3}—C_{sp3}—C_{sp3}—C_{sp3} are also rather small. Table II lists the conformational energies for each model discussed earlier.

ALKENES

Three simple compounds were studied in this work: (1) propene (H—C_{sp3}—C_{sp2}—C_{sp2} trans vs. cis, $\Delta E_{\text{trans–cis}} = 2.00$ kcal/mol); (2) 1-butene (cis

vs. skew, $\Delta E_{cis-skew} = 0.22$ kcal/mol); and (3) 2-butene (cis vs. trans, $\Delta E_{cis-trans} = 1.20$ kcal/mol). The energy differences in the Cornell et al. model are −0.32, 1.87, and 0.49 kcal/mol, respectively. It is well known that, in propene, the C_{sp3}—H prefers to eclipse the C_{sp2}—C_{sp2} bond rather than the C_{sp2}—H bond. As for 1—butene (cis vs. skew), the relative conformational energy is a compromise between C_{sp3}—H and C_{sp3}—C_{sp3}, eclipsing the C=C. We were able to fit these two compounds with two specific Fourier terms for H—C_{sp3}—C_{sp2}—C_{sp2} (V_1 [1.15, 0.0] and V_3 [0.38, 180.0]). When we consider 2-butene (cis vs. trans), a V_1 (1.90, 180.0) term is required to correctly represent this preference. With this complex model, very encouraging results are obtained and the energy differences are 1.96, 0.18, and 1.25 kcal/mol for the three compounds mentioned earlier, respectively.

CONJUGATED COMPOUNDS

Two conjugated compounds, 1,3-butadiene (gauche vs. trans, $\Delta E_{gauche-trans} = 2.89$ kcal/mol) and 2-methyl-1,3-butadiene (gauche vs. trans, $\Delta E_{gauche-trans} = 2.65$ kcal/mol), were studied in this work. Although glyoxylic acid (no. 73 in Table I) and acrolein (no. 80 in Table I) are also conjugated systems, they are discussed with other carbonyl-containing compounds. Considering that the inner two carbons of butadiene are not pure sp^2 carbons, and the bond lengths between them are slightly longer than a pure single C_{sp2}—C_{sp2}, a new atomic type (CD) can be introduced. The van der Waals parameter is the same as that of sp^2 carbon and the equilibrium bond length of CD—CD is set to 1.40 Å, similar to that of C_{sp2}—C_{sp2} bonds in pure benzene. A new generic torsional parameter, X—CD—CD—X, was optimized by *parmscan* and, finally, the force constant was set to 4.0 kcal/mol with a phase angle of 180.0°. The average absolute error for the two compounds is 0.36 kcal/mol.

FLUORIDES, CHLORIDES, AND BROMIDES

Five fluorides, seven chlorides, and four bromides were studied for halides. The Cornell et al. model gives average absolute errors of 1.32, 1.11, and 1.86 kcal/mol for the three kinds of compounds, respectively. Given the fact that halogens in 1,2-difluoroethane prefer to be trans to hydrogen atoms, rather than to each other, two kinds of torsional parameters, H—C_{sp3}—C_{sp3}—X and X—C_{sp3}—C_{sp3}—X, were introduced. Instead of using only the default threefold Fourier component, we applied the onefold Fourier component for these two torsions with phase angles of 0.0° and 180.0°, respectively. For the former, the force constants change from 0.19, to 0.25, and then to 0.55 kcal/mol for fluorides, chlorides, and bromides, respectively. For the latter torsion parameter, the sequence order of force constants is still qualitatively consistent with the periodic table, changing from 1.20 to 0.45 and then to 0.0 kcal/mol for fluorides, chlorides, and bromides, respectively. With the new model, the mean absolute errors are 0.24, 0.30, and 0.41 kcal/mol for fluorides, chlorides, and bromides, respectively. We also tried to apply twofold Fourier components to improve the model, which occurs to some degree. However, the model applying the onefold Fourier terms shows a better performance.

AMIDES AND AMINO COMPOUNDS

Two amides and eight amino molecules were investigated in this study. A new atomic type, NT, that is, the neutral sp^3 nitrogen, was generated. This atom type shares the same van der Waals parameter as other nitrogen atomic types. Bond stretching and bending parameters are also shared with the N3 atom type, which is the charged ammonium nitrogen in the Cornell et al. force field. For the torsional parameter of X—C_{sp3}—N_{sp3}—X (X can be any atomic type), a generic threefold Fourier component was applied, just as that in the Cornell et al. model for X—CT—N3—X (V_3 [1.80, 0.0]). The average absolute error of this simple model is 0.58 kcal/mol. Encouraged by the great improvement made for alkanes after applying all of the first three Fourier components for C_{sp3}—C_{sp3}—C_{sp3}—C_{sp3}, we added a small twofold component for C_{sp3}—C_{sp3}—N_{sp3}—C_{sp3} (V_3 [0.3, 0.0], V_2 [0.48, 180.0]). We found that the average absolute error is further reduced to 0.20 kcal/mol and the RMS deviation is only 0.23 kcal/mol.

ALCOHOLS

For alcohols, nine energy conformational pairs were considered for optimizing the torsional parameters. For the first four conformational pairs (nos. 46–49 in Table I), experimental data are available, whereas only high-level *ab initio* (MP4SDQ/TZP) data are available for the last five conformational pairs (nos. 50–54 in Table I).

The average absolute error for the Cornell et al. model is 0.74 kcal/mol, and therefore it is necessary to add specific torsional parameters to further

reduce the error. Similar to the H—C_{sp3}—C_{sp3}—X (X = F, Cl, or Br) torsional parameters noted earlier, we included the onefold Fourier component for H—C_{sp3}—C_{sp3}—O V_1 (0.25, 0.0). The force constant for this torsion is the same as that of H—C_{sp3}—C_{sp3}—Cl, which is consistent with the similar electronegativities of oxygen and chlorine. Another specific torsional angle, C_{sp3}—C_{sp3}—O—H, was also introduced to regenerate the relative energies to the extent possible. After optimization by *parmscan*, both onefold component V_1 (0.25, 0.0) and threefold component V_3 (0.16, 0.0) were included for this torsion. This new model reduces the average absolute error to 0.48 kcal/mol. Although the AAE is still slightly larger than those of other organic species, we still believe that the new model works quite well because only MP4SDQ/TZP data are available for the last five conformational pairs. It is known that this *ab initio* method has an RMS deviation of ~0.4 kcal/mol, similar to the experimental data.

ETHERS

Seventeen ethers, including nine 1,3-dioxane derivatives (nos. 63–71 in Table I), were investigated in this work. For the last three 1,3-dioxane derivatives (nos. 69–71 in Table I), only MP2/6-31G* energies are available. The Cornell et al. model applies only a generic torsional parameter for X—C_{sp3}—O—X (X can be any atom type), and it works well for the most ethers except those with OCO units, such as 1,3-dioxane and its derivatives. The AAE is 0.44 kcal/mol without considering 1,3-dioxane and its derivatives. However, the AAE is 2.08 kcal/mol for 1,3-dioxane and its eight derivatives. Howard et al. found that a onefold Fourier component (V_1) instead of a threefold one (V_3) for C_{sp3}—O—C_{sp3}—O reproduces the relative energies very well. Moreover, the relative energies of their model show very good correlation with those of *ab initio* MP2/6-31G*.[25] Although the AAE of this model is only ~0.5 kcal/mol compared with the MP2/6-31G* energies, it is still significantly larger, when compared with the experimental data. In our opinion, adding V_2 and V_3 torsional parameters was necessary to depict the energies more accurately.

For our new model, V_1 was introduced first and then V_2 and V_3 were added in sequence. It is usually thought that bond stretching and bond bending make little contribution to the relative conformational energy. However, we found that the bond angle parameter, O—C_{sp3}—O, plays an important role in the successful modeling of 1,3-dioxane and its derivatives. The equilibrium value of this bond angle and force constant were first set to 109.5° and 50.0 kcal/mol·Å^2, respectively, as shown by Howard et al. However, we found that MM3 uses a significantly smaller equilibrium bond angle, 103.1°. We therefore sought to optimize the equilibrium bond angle as well as the force constant of O—C_{sp3}—O using *parmscan*. The equilibrium bond angle and force constant were refined at 101° and 160 kcal/mol·Å^2, respectively. It is of interest to point out that, in phosphates, the pendant oxygen group also applies a large force constant for O2—P—O2 (140.0 kcal/mol·Å^2), although it is much smaller for the ester oxygen OS—P—OS. Encouragingly, our somewhat more complex model not only reproduces the *ab initio* data better than Howard et al.'s model but also reproduces the experimental data quite well. The AAE is only 0.34 kcal/mol for all 17 ethers. Table III lists the conformational energies and the AAEs of the different models. From this table, one can estimate the contributions of different parameters.

CARBONYL COMPOUNDS

The 11 compounds with carbonyl groups were studied. First, acetic acid and methylformate were selected as model molecules to derive the generic parameters of X—C—O—H and X—C—O—X, where X can be any atomic type. Potential energy surface scanning and minimization were performed for O—C—O—H of acetic acid and O—C—O—C of methylformate. With the aim to investigate the effects of basis sets and different *ab initio* approaches, the following four methods were applied for the *ab initio* calculations: HF/6-31G*; MP2/6-31G*; MP2/6-311+G (2d,p); and B3LYP/6-311+G (2d,p). The curves of two potential energy surfaces are very similar in shape: both of them are symmetrical around 180°, and have global minima at 0.0°, a local minima at 180° and two other stationary points around ±90°. Because the profiles are not the standard cosine curves, a single twofold Fourier component cannot depict the whole curve very well. With the intention of making our model as simple as possible, we tried only to reproduce the relative energies between the stationary points. The following are the relative energies of different methods for acetic acid. For the local minimum at 180°, the relative energies for the global minimum conformation are 7.19, 6.99, 5.64, and 5.50 kcal/mol for the four methods in sequence; for the maxima at 90° and −90°, the relative energies for the global minimum conformation are 13.62, 14.47, 13.01, and 12.81 kcal/mol, respectively. As to methylformate,

TABLE III.
Conformational Energies of Ethers for Different Parameter Sets.[a]

Compound no.	Compound name and conformational comparison	Expt.[b]	AMBER (set I)	AMBER (set II)	AMBER (set III)	AMBER (set IV)
1	2-Methoxytetrahydropyrane, OCOME g, OME, eq–ax	1.05	1.11	0.28	2.81	1.71
2	2,5-Dimethyl-1,3-dioxane, 2eq,5ax–2eq,5eq	0.92	0.96	1.74	1.72	0.80
3	Methylethylether, g–a	1.50	1.50	1.54	1.54	1.50
4	Methylvinylether, skew–cis	1.70	1.75	1.75	1.75	1.75
5	Diethylether, g–a	1.14	1.55	1.59	1.59	1.55
6	Methoxycyclohexane, ax,C1–eq,C1	0.55	0.54	1.31	1.31	0.54
7	Tetrahydrofuran, C2–Cs	0 ± 0.3	0.48	−0.06	0.48	0.48
8	Tetrahydrofuran, C2v–C2	3.5	2.76	3.80	2.76	2.76
9[b]	1,3-Dioxane, comp3, twist boat–chair	2.20	2.19	5.45	3.76	1.39
10[b]	1,3-Dioxane, comp4a, twist boat–chair	0.85	0.93	0.93	0.93	0.93
11[b]	1,3-Dioxane, comp4b, twist boat–chair	0.26	0.56	0.57	0.57	0.56
12[b]	1,3-Dioxane, comp4c, twist boat–chair	0.22	−0.26	−2.64	−1.15	−0.10
13[b]	1,3-Dioxane, comp4d, twist boat–chair	1.20	1.78	−0.96	0.47	1.66
14[b]	1,3-Dioxane, comp4e, twist boat–chair	1.80	3.22	0.74	2.11	3.18
15[b]	1,3-Dioxane, comp2, twist boat–chair	2.02	1.92	5.33	3.61	1.06
16[b]	1,3-Dioxane, comp1, twist boat–chair	4.83	3.90	7.80	5.78	2.26
17[b]	1,3-Dioxane, comp5, twist boat–chair	2.74	2.37	0.0	1.40	2.42

[a] Set I: Parm99; set II: Parm94; set III: Parm99; however, for OS—CT—OS—CT only V_1 is included and the optimized force constant is 1.80; set IV: Parm99; however, the reference bond angle of OS—CT—OS is set to 109.0° and the force constant is optimized to 160 kcal/mol (conformational energies in kcal/mol).

[b] For the sources and measure types of data, see Table I.

[c] Compounds 1–3, 4a–4e, and 5 are defined in ref. 25.

for the local minimum at 180°, the relative energies for the global minimum conformation are 6.26, 6.36, 5.45, and 4.76 kcal/mol for the four methods in sequence, whereas the experimental value is 4.75 kcal/mol (no. 74 in Table I); for the maxima at 90°, and −90°, the energies relative to the global minimum conformation are 12.79, 14.44, 13.88, and 13.45 kcal/mol, respectively. Given that the first two methods are not as accurate (the AAE for the experiment was estimated to be 0.8 kcal/mol[19]), we focused on the latter two methods. One may believe that the two methods have similar accuracy, but our choice is the B3LYP/6-311+G (2d,p) model, which requires less CPU time than the MP2 method with the same basis set. In this work, the B3LYP/6-311+G (2d,p) energies were chosen as references to derive the two generic torsional parameters. With *parmscan*, we can easily obtain the best force constants to reproduce the *ab initio* relative energies, and the force constants are 4.6 and 5.4 kcal/mol for X—C—O—H and X—C—O—X, respectively.

Besides the two generic torsional parameters, the following specific ones were included to further minimize the average absolute error: H—C_{sp3}—C=O (V_1 [0.80, 0.0], V_3 [0.08, 180.0]); H—O—C=O (V_1 [1.90, 0.0], V_2 [2.30, 180.0]); C=C—C=O (V_2 [2.175, 180.0], V_3 [0.30, 0.0]); and C_{sp3}—O—C=O (V_1 [1.40, 180.0], V_2 [2.70, 180.0]). Again, *parmscan* was used to choose the Fourier components as well as the force constants. First, only the component making the largest contribution was included in our model and other components were added only when necessary, because we prefer to keep our model simple. The final model reproduced the relative energies very well for all the 11 molecules. The AAE was 0.4 kcal/mol and the RMS deviation was 0.59 kcal/mol.

ADDITIVE FORCE FIELD (AFF) MODEL FOR ORGANIC MOLECULES

We studied a total of 82 conformational pairs using the simple pairwise additive model presented in eq. (1). The average absolute error is 0.33 kcal/mol and the RMS deviation is 0.47 kcal/mol. To compare with other widely available force fields, we extracted two subsets from the 82 conformational

TABLE IV.
Conformational Energies (kcal/mol) of Additive Force Field (AFF) and Nonadditive Force Field (NAFF).[a]

Compound no.	Compound name and conformational comparison	Expt.[b]	AFF	NAFF (I)	NAFF (II)	NAFF (III)
1	Butane, g–a	0.75	0.86	0.87	—	—
2	Cyclohexane, twist boat-chair	5.50	6.58	6.58	—	—
3	Methylcyclohexane, ax–eq	1.75	1.66	1.69	—	—
4	2,3-dimethylbutane, H—C2—C3—H, g–a	0.05	−0.15	−0.10	—	—
5	Cyclooctane, D4d–Cs boat–chair	1.90	1.59	1.58	—	—
6	Cyclononane, [255]C2–[333]D3	1.00	1.20	1.20	—	—
7	Butane, cis–trans	4.89	5.09	5.47	—	—
8	1,3-Butadiene, g–t	2.89	3.33	3.42	—	—
9	2-Methyl-1,3-butadiene, g–t	2.65	2.36	2.69	—	—
10	1-Butene, cis–skew	0.22	0.18	0.22	—	—
11	2-Butene, cis–skew	1.20	1.25	1.23	—	—
12	1,2-Difluoroethane, a–g	0.58	0.34	0.70	—	—
13	1-Fluoropropane, a–g	0.33	−0.03	−0.02	—	—
14	Fluorocyclohexane, ax–eq	0.16	0.46	0.35	—	—
15	*Trans*-1,2-difluorocyclohexane, ax,ax–eq,eq	0.59	0.31	0.79	—	—
16	*Trans*-1,4-difluorocyclohexane, eq,eq–ax,ax	1.14	1.13	0.81	—	—
17	1,2-Dichloroethane, g–a	1.08	0.99	0.66	—	—
18	1-Chloropropane, a–g	0.37	−0.21	−0.26	—	—
19	1,3-Dichloropropane, ga–gg	1.10	0.29	0.07	—	—
20	1,3-Dichloropropane, aa—gg	1.50	0.97	0.43	—	—
21	Chlorocyclohexane, ax–eq	0.50	0.46	0.35	—	—
22	*Trans*-1,2-dichlorocyclohexane, eq,eq–ax,ax	0.93	0.91	0.15	—	—
23	*Trans*-1,4-dichlorocyclohexane, eq,eq–ax,ax	0.80	0.75	0.23	—	—
24	Bromocyclohexane, ax–eq	0.70	0.11	0.13	—	—
25	*Trans*-1,2-dibromocyclohexane, eq,eq–ax,ax	1.50	1.03	1.23	—	—
26	Trans-1,4-dibromocyclohexane, eq,eq–ax,ax	0.88	0.94	0.46	—	—
27	Isopropylamine, LP—N—C—H, a–g	0.45	0.87	1.52	1.43	1.47
28	Cyclohexylamine, ax–eq	1.15	1.22	1.25	1.23	1.22
29	Properidine, ax–eq	0.53	0.41	0.12	0.89	0.92
30	*N*-Methylpiperidine, ax–eq	3.15	3.19	3.24	3.95	4.06
31	2-Methylpiperidine, NH eq, Me ax–eq	2.50	2.71	2.59	2.25	2.25
32	3-Methylpiperidine, NH eq, Me ax–eq	1.60	1.87	1.91	2.15	2.01
33	4-Methylpiperidine, NH eq, ME ax–eq	1.93	1.99	1.50	2.24	1.99
34	*N*, *N*-dimethylcyclohexane, ax Cs–eq C1	1.31	1.50	1.10	0.77	0.53
35	*N*-methylformamide, c–t	1.40	1.74	2.34	2.06	2.08
36	*N*-methylacetamide, c–t	2.30	2.57	3.44	2.87	2.94
37	Ethanol, g–a	0.12	0.79	0.56	0.82	0.78
38	Isopropanol, H—C—O—H, a–g	0.28	0.32	−0.19	0.35	0.34
39	Cyclohexanol, ax C1–eq C1	0.58	0.70	0.37	0.36	0.33
40	Cyclohexanol, eq Cs–eq C1	0.18[b]	0.67	0.29	0.47	0.45
41	Cyclohexanol, ax Cs–eq C1	1.14[b]	0.014	−0.17	0.02	0.00
42	Cyclopentaol, eq Cs–ax C1	1.11[b]	0.77	0.68	0.83	0.81

TABLE IV.
(Continued)

Compound no.	Compound name and conformational comparison	Expt.[b]	AFF	NAFF (I)	NAFF (II)	NAFF (III)
43	Cyclopentaol, ax Cs–ax C1	1.05[b]	0.63	0.40	0.66	0.62
44	Cyclopentaol, eq C1–ax C1	1.14[b]	0.18	0.27	0.10	0.08
45	2-Methoxytetrahydropyran, OCOME g, OME, eq–ax	1.05	1.11	0.67	0.65	0.67
46	2,5-Dimethyl-1,3-dioxane, 2eq, 5ax–5eq	0.92	0.96	1.05	1.23	1.07
47	Methylethylether, g–a	1.50	1.50	1.50	1.96	2.00
48	Methylvinylether, skew–cis	1.70	1.75	1.62	1.70	2.17
49	Diethylether, g–a	1.14	1.55	1.52	1.96	1.98
50	Methoxycyclohexane, ax–eq	0.55	0.54	0.31	0.51	0.47
51	Tetrahydrofuran, C2–Cs	0 ± 0.3	0.48	0.49	0.00	0.00
52	Tetrahydrofuran, C2v–C2	3.5	2.76	2.78	3.21	3.22
53	Formic acid, t–c	3.90	4.82	3.20	4.06	3.84
54	Glyoxalic acid, c–t	1.20	1.52	2.08	1.82	2.16
55	Methylformate, t–c	4.75	6.12	5.57	6.04	6.52
56	Methylacetate, t–c	8.50	8.02	7.37	8.10	8.21
57	Ethylformate, g–a	0.19	0.38	0.38	0.55	0.47
58	Propionaldehyde, skew–cis	0.67	0.76	0.83	0.66	0.63
59	2-Butanone, skew–cis	1.07	0.92	0.97	0.79	0.76
60	Acrolein, c–t	1.70	1.71	1.87	2.30	2.23

[a] Three schemes of NAFF are: I—polarizable model without lone pairs; II—polarizable model with lone pairs having distances 0.2 Å from heavy atoms; III—polarizable model with lone pairs having distances 0.35 Å from heavy atoms. The polarizability parameters for NAFF models were developed by Applequist et al.[48]

[b] For the sources and measure types of data, see Table I. For nos. 40–44, no experimental data were available, so *ab initio* energies at the MP4SDQ/TZP level were used.

pairs, the 34-molecule set and 55-molecule set, for comparison purposes. We compared our results with MMFF, MM3, and CHARMm, which have been used widely and have shown good performance in characterizing the geometries and relative conformational energies. If the force field value differs from the experiment or *ab initio* by >1.0 kcal/mol, it is regarded as an outlier. Both the average absolute error and RMSD are used to evaluate the performance of different force fields.

TABLE V.
Summary of AMBER Conformational Energies (kcal/mol) versus Experiment by Compound Classes.

	Number	Average of absolute error	RMS deviation
Alkane	14	0.32	0.47
Alkene	3	0.04	0.04
Conjugated compound	2	0.36	0.37
Fluoride	5	0.24	0.27
Chloride	7	0.30	0.43
Bromide	4	0.41	0.46
Amide and amino compound	10	0.20	0.23
Alcohol	9	0.48	0.60
Ether	17	0.34	0.51
Compound with carbonyl	11	0.40	0.59
Total	82	0.33	0.47

TABLE VI.
Summary Comparisons of Difference Force Fields for 34-Molecule Set.[a]

	Experiment		MP4/TZP		GVB-LMP2[b]	
	Average of absolute error	RMS dev.	Average of absolute error	RMS dev.	Average of absolute error	RMS dev.
AMBER	0.28	0.43	0.33	0.43	0.27	0.37
MMFF	0.31	0.38	0.18	0.33	0.22	0.29
MM3	0.53	0.74	0.53	0.81	0.47	0.73
CHARMm	0.58	0.77	0.51	0.70	0.51	0.71

[a] Conformational energies in kcal/mol.
[b] With the cc-pVTZ (-f) basis set.

Table VI lists the results of comparisons for the 34-molecule set. Compared with the experimental data, AMBER shows the smallest AAE (0.28 kcal/mol), slightly better than MMFF, which shows an AAE of 0.31 kcal/mol. MM3 and CHARMm should be regarded at the same level, having AAEs of 0.53 and 0.58 kcal/mol, respectively. When compared with MP4/TZP and GVB/LMP2, MMFF shows the smallest AAE and RMS deviation, both marginally better than AMBER. It is understandable that MMFF gives the best performance for the comparisons to the *ab initio* data, given that these data were used extensively in the MMFF parameterization.

Table VII gives a summary of comparison results for the 55-molecule set. Both the AAE and RMS deviation for each compound class are listed for the four force fields. The AAEs are 0.28, 0.43, 0.52, and 0.57 kcal/mol for AMBER, MMFF, MM3, and CHARMm, respectively. As for RMS, the same trend can be seen. It is of interest to point out that MMFF, MM3, and CHARMm perform better than AMBER if only alcohols are considered. For compounds in other classes, AMBER always has the smallest AAE and RMS deviations. The performance of each force field is illustrated in Figures 1–3. From these, we can quickly find the number of outliers and the number of wrong conformers favored (wrong sign). For the

TABLE VII.
Summary Comparisons of Difference Force Fields for 55-Molecule Set.[a]

		AMBER		MMFF		MM3		CHARMm	
	Number	Average of absolute error	RMS dev.	Average of absolute error	RMS dev.	Average of absolute error	RMS dev.	Average of absolute error	RMS dev.
Alkane	9	0.31	0.42	0.35	0.40	0.46	0.82	0.49	0.69
Alkene	2	0.04	0.04	0.10	0.11	0.32	0.35	0.32	0.34
Conjugated compound	2	0.36	0.37	0.48	0.48	1.10	1.10	0.60	0.61
Fluoride	5	0.24	0.27	0.63	0.79	0.35	0.38	0.29	0.33
Chloride	7	0.30	0.43	0.70	0.81	0.65	0.76	0.69	0.91
Bromide	4	0.41	0.46	0.50	0.63	0.72	0.89	0.72	0.81
Amide and amino compound	10	0.20	0.23	0.33	0.38	0.21	0.31	0.56	0.78
Alcohol	3	0.28	0.39	0.14	0.17	0.28	0.30	0.18	0.19
Ether	6	0.10	0.17	0.37	0.44	0.40	0.54	0.65	0.87
Compound with carbonyl	7	0.48	0.66	0.45	0.54	1.02	1.27	0.82	0.99
Total	55	0.28	0.40	0.43	0.54	0.52	0.75	0.57	0.76

[a] Conformational energies in kcal/mol.

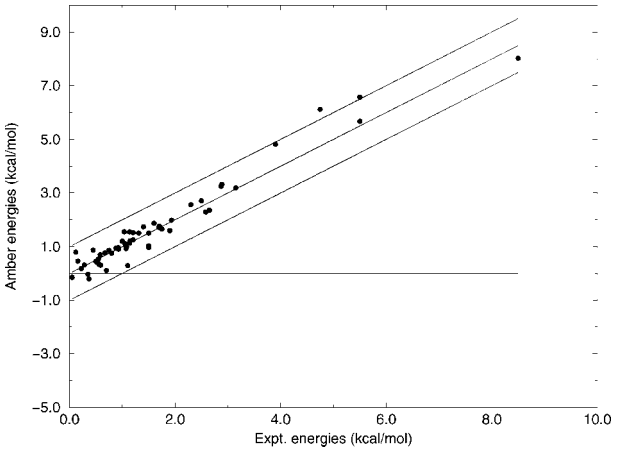

FIGURE 1. The performance of AMBER (Parm99) force field for reproducing the relative experimental conformational energies. Three parallel lines are defined as $y = x + 1.0$, $y = x$, and $y = x - 1.0$ from top to bottom. The greater density of dots close to the middle line indicates better performance. The dots above $y = x + 1.0$ or below $y = x - 1.0$ are defined as outliers. A dot below the line that is parallel with the *X* axis predicts the wrong sign (favoring the wrong conformation). For Parm99, there are two outliers and three wrong signs.

55-molecule set, AMBER has two outliers, whereas MMFF, MM3, and CHARMm have four, five, and nine outliers, respectively. As for wrong sign, AMBER has three, whereas MMFF, MM3, and CHARMm have five, five, and seven, respectively.

Considering that all the molecules for comparisons are selected from our 82-molecule training set, we should note that the comparison with MM3 and

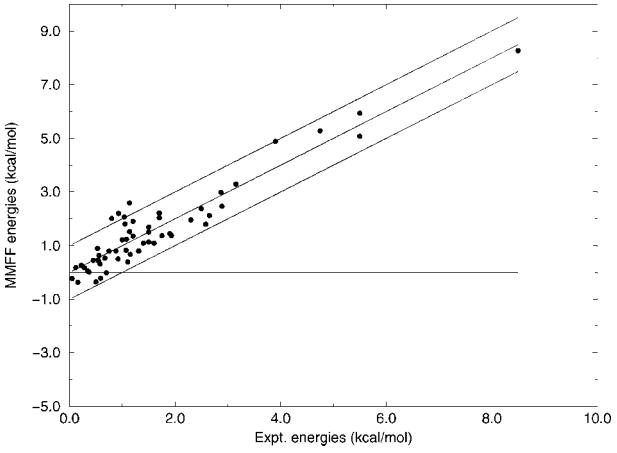

FIGURE 2. The performance of the MMFF force field for reproducing the relative experimental conformational energies. Notation is the same as that in Figure 1. For MMFF, there are four outliers and five wrong signs.

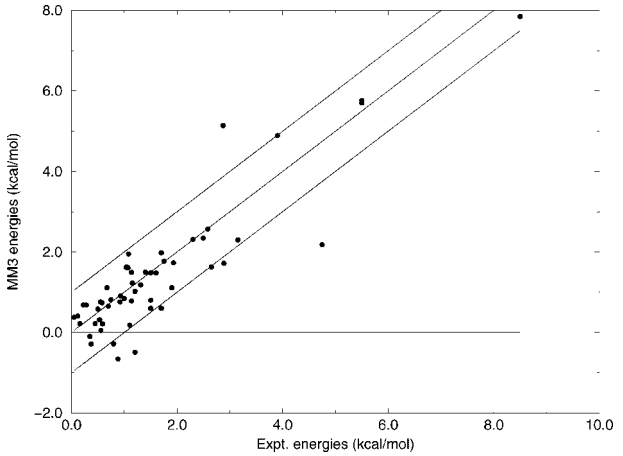

FIGURE 3. The performance of the MM3 force field for reproducing the relative experimental conformational energies. Notation is the same as that in Figure 1. For MM3, there are six outliers and five wrong signs.

CHARMm might be biased considering that their force fields may have not included all of the molecules for parameterization. We thus investigated how well the parameters transfer from training set to test set by considering two examples for further study. For the first example, 11 hydrocarbons that involve the CT—CT—CT—CT torsional parameter are grouped into the training set (nos. 2, 4, 5, 7, 9, and 11 in Table VIII) and the test set (nos. 1, 3, 6, 8, and 10 in Table VIII). For the second example, seven chlorides, three (nos. 12, 14, and 17 in Table VIII) in the training set and four (nos. 13, 15, 16, and 18 in Table VIII) in the test set, involve the torsional parameter Cl—CT—CT—HC. Parameter set I

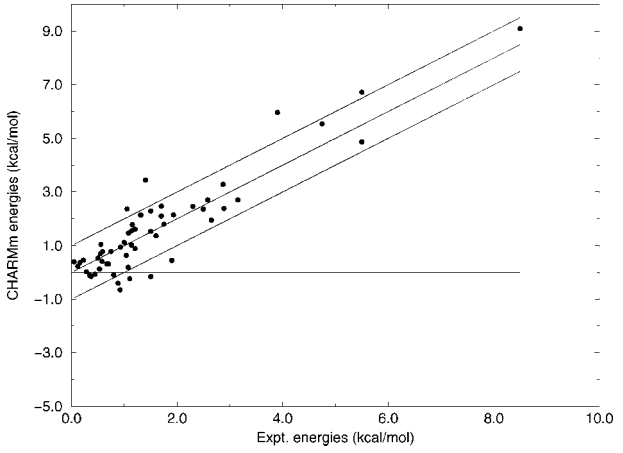

FIGURE 4. The performance of the CHARMm force field for reproducing the relative experimental conformational energies. Notation is the same as in Figure 1. For CHARMm, there are nine outliers and eight wrong signs.

TABLE VIII.
Conformational Energies of 11 Hydrocarbons and 7 Chlorides with Parameters Derived Using Training Set and Full Molecular Set (Parm99).

Compound no.	Compound name and conformational comparison[a]	Expt.	Parm99	Parameter derived with training set
1	Butane, g–a	0.75	0.86	0.86
2	Cyclohexane, twist boat–chair	5.50	6.58	6.41
3	Methylcyclohexane, ax–eq	1.75	1.66	1.68
4	2,3-Dimethylbutane, H—C2—C3—H, g–a	0.05	−0.15	−0.13
5	Cyclooctane, D4d–Cs boat–chair	1.90	1.59	1.82
6	Cyclononane, [255]C2–[333]D3	1.00	1.20	1.39
7	Phenylcyclohexane, ax–eq	2.87	3.25	3.25
8	*Trans*-1,2-dimethylcyclohexane, ax,ax–eq,eq	2.58	2.29	2.35
9	*Cis*-1,3-dimethylcyclohexane, ax,ax–eq,eq	5.50	5.67	5.73
10	Cyclopenta, plane-pucker	5.20	4.03	3.92
11	Butane, cis–trans	4.89	5.09	5.00
12	1,2-Dichloroethane, g–a	1.08	0.99	1.04
13	1-Chloropropane, a–g	0.37	−0.21	−0.24
14	1,3-Dichloropropane, ga–gg	1.10	0.29	0.26
15	1,3-Dichloropropane, aa–gg	1.50	0.97	0.91
16	Chlorocyclohexane, ax–eq	0.50	0.46	0.52
17	*Trans*-1,2-dichlorocyclohexane, eq,eq–ax,ax	0.93	0.91	0.86
18	*Trans*-1,4-dichlorocyclohexane, eq,eq–ax,ax	0.80	0.75	0.64

[a] For the sources and measure types of data, see Table I.

is defined as the torsional parameter derived using all the molecules (Parm99); set II is the same parameter optimized using only the training set. Table VIII lists the relative energies of these two torsional parameter sets. For CT—CT—CT—CT, set I (V_1 [0.18, 0.0], V_2 [0.25, 180.0], V_3 [0.20, 180.0]) and set II (V_1 [0.18, 0.0], V_2 [0.29, 180.0], V_3 [0.24, 180.0]) are very similar, and the AAEs of the 11 hydrocarbons are almost the same for the two sets (0.355 vs. 0.359 kcal/mol). It is understandable that set II gives the smaller AAE for the training set (0.35) than that of set I (0.39); on the contrary, for the test set, the AAE of set II (0.37) is a little larger than that of set I (0.31). For the second example, parameter set II has a marginally smaller force constant (V_1 [0.23, 0.0]) than that of set I (V_1 [0.25, 0.0]). The AAEs of the seven chlorides are 0.30 and 0.33 kcal/mol for sets I and II, respectively. For the training set, set II also gives smaller AAE (0.297) than that of set I (0.313), whereas, for the test set, set I gives a smaller AAE (0.30) than that of set II (0.36).

From the two examples just presented, we believe that the torsional parameters of Parm99 can be reliably transferred from the training set to the test set. It is our estimate that, for the organic compounds outside the training set, in cases of no parameter missing, Parm99 can reproduce the relative energies of experiment with an AAE of <0.40 kcal/mol. We conclude that a well-parameterized, simple additive model with RESP charge can describe the structure and the intramolecular energies for organic systems very well.

NONADDITIVE FORCE FIELD (NAFF) MODEL FOR ORGANIC MOLECULES

For the nonadditive model described by eqs. (2)–(7), we worked out three schemes. The first is a simple polarizable model without lone pairs [NAFF (I)], whereas the other two models include lone pairs for oxygen and nitrogen. The distances from the lone pairs to the center atoms were set to 0.2 Å [NAFF (II)] and 0.35 Å [NAFF (III)] for the latter two models, respectively. Table IV lists the conformational energies of the three models. Sixty conformational pairs were studied. For the first 26

TABLE IX.
Summary Comparisons of Additive Force Field (AFF) and Nonadditive Force Fields (NAFF) for 60 Compounds.[a]

		AFF		NAFF (project I)		NAFF (project II)		NAFF (project III)	
	Number	Average of absolute error	RMS dev.	Average of absolute error	RMS dev.	Average of absolute error	RMS dev.	Average of absolute error	RMS dev.
Alkane	7	0.31	0.45	0.36	0.49	—	—	—	—
Alkene	2	0.04	0.04	0.02	0.02	—	—	—	—
Conjugated compound	2	0.36	0.37	0.28	0.38	—	—	—	—
Fluoride	5	0.24	0.27	0.24	0.25	—	—	—	—
Chloride	7	0.30	0.43	0.66	0.73	—	—	—	—
Bromide	3	0.37	0.44	0.42	0.44	—	—	—	—
Amide and amino compound	10	0.20	0.23	0.48	0.62	0.51	0.57	0.52	0.61
Alcohol	8	0.52	0.63	0.56	0.67	0.51	0.63	0.52	0.64
Ether	8	0.19	0.31	0.28	0.35	0.33	0.39	0.38	0.62
Compound with carbonyl	8	0.44	0.62	0.52	0.64	0.46	0.72	0.53	0.76
Total	60 (34)[b]	0.31	0.44	0.43 (0.46)[c]	0.55 (0.50)[c]	0.46	0.56	0.48	0.66

[a] Three projects of NAFF are studied: I—polarizable model without lone pairs; II—polarizable model with lone pairs having distances 0.2 Å from heavy atoms; III—polarizable model with lone pairs having distances 0.35 Å from heavy atoms (conformational energies in kcal/mol).
[b] Nos. 27–60 in Table IV.
[c] For 34-molecule subset in Table IV (nos. 27–60).

conformational pairs that are isolated hydrocarbons and halides, only the standard nonadditive model (NAFF (I)] was used, because the electrostatic and polarization effects are negligible for hydrocarbons and no lone pairs are used for halogens in our model. For the remaining 34 conformational pairs, all of the three NAFF models were investigated.

All of the NAFF models applied the same torsional parameters as those developed for the AFF model. Table IX lists the results of comparison of each model with the experimental data or *ab initio* energy in cases unavailable experimental values. For the 60 conformational pairs, the AFF model achieves better performance (AAE and RMS deviation 0.31 and 0.44 kcal/mol, respectively) than all of the three polarizable models. As for NAFF (I), the AAE is 0.44 and the RMS deviation is 0.55 kcal/mol. The AAEs are almost the same for NAFF (II) and NAFF (III), which are 0.46 and 0.48 kcal/mol, respectively. The RMS deviations of NAFF (II) and NAFF (III) are also similar, 0.56 kcal/mol for NAFF (II) and 0.62 kcal/mol for NAFF (III). It is encouraging that our polarizable models perform reasonably well when applying the same parameters developed for the additive model. All three nonadditive models have performance levels comparable to MMFF, which has an AAE of 0.43 kcal/mol (55-molecule set).

It is likely that our nonadditive model can be improved significantly with the logical adjustment of some torsional parameters. In this case, "a logical adjustment" means that both the additive and nonadditive models use different force constants, but they should use the same phase angles. Table X presents some results of comparing a polarizable model with and without logical adjustments of the torsional parameters. From Table X we conclude that, in most cases, the polarizable models do achieve comparable or better performance than the additive model, after making a logical adjustment of torsional parameters.

PEPTIDES

The development of an accurate model to reproduce precisely the relative energies of peptide conformations is crucial in successfully modeling biological systems. With the great progress made in computer performance, one can undertake high-level *ab initio* calculations for large systems such as peptides. The reconsideration of organic mole-

TABLE X.

Comparison of Performances of Polarizable Models with and without Further Torsional Adjustments.

Compound class	Fluoride[a]	Chloride[b]	Bromide[c]	Acid[d]	Conjugated aldehyde[e]
Nonadditive model type[f]	NAFF (I)	NAFF (I)	NAFF (I)	NAFF (I)	NAFF (II) and (III)
Number of conformational pairs	5	7	4	4	1
Torsional parameter[g] before adjustment	F—C_{sp3}—C_{sp3}—F (1.20, 180.0, 1) F—C_{sp3}—C_{sp3}—H (0.19, 0.0, 1)	Cl—C_{sp3}—C_{sp3}—Cl (0.45, 180.0, 1) Cl—C_{sp3}—C_{sp3}—H (0.25, 0.0, 1)	Br—C_{sp3}—C_{sp3}—Br (0.00, 180.0, 1) Br—C_{sp3}—C_{sp3}—H (0.55, 0.0, 1)	O=C_{sp2}—O—H (1.50, 0.0, 1)	C_{sp2}—C_{sp2}—C_{sp2}=O (0.30, 0.0, 3)
Torsional parameter[g] after adjustment	F—C_{sp3}—C_{sp3}—F (1.00, 180.0, 1) F—C_{sp3}—C_{sp3}—H (0.24, 0.0, 1)	Cl—C_{sp3}—C_{sp3}—Cl (0.25, 180.0, 1) Cl—C_{sp3}—C_{sp3}—H (0.35, 0.0, 1)	Br—C_{sp3}—C_{sp3}—Br (0.0, 180.0, 1) Br—C_{sp3}—C_{sp3}—H (0.65, 0.0, 1)	O—C_{sp2}—O—H (1.90, 0.0, 1)	C_{sp2}—C_{sp2}—C_{sp2}=O (0.55, 0.0, 3)
AAE[h] without adjustment	0.24	0.66	0.42	0.66	0.50 and 0.43
AAE[h] with adjustment	0.12	0.47	0.32	0.27	0.063 and 0.01

[a] Nos. 12–16 in Table IV.
[b] Nos. 17–23 in Table IV.
[c] Nos. 24–26 in Table IV.
[d] Nos. 81–82 in Table I and 53–54 in Table IV.
[e] Nos. 60 in Table IV.
[f] The definitions of the three nonadditive models are given in Table IX.
[g] The three numbers in the parentheses correspond to the force constant, phase angle, and which fold of Fourier component.
[h] Average absolute error (in kcal/mol).

TABLE XI.
MM Results for Alanine Dipeptides and Tetrapeptides Using Standard Charges.[a]

		Parm 99		Parm 94	
Conformers	*Ab initio*[b]	Additive model	Polarizable model	Additive model	Polarizable model
C7eq	0.00	0.00	0.00	0.00	0.00
C7ax	2.10	2.23	2.70	1.48	2.00
C5	0.89	0.90	1.25	1.50	1.70
AlphaR	3.91	4.23	3.83	3.66	3.15
AlphaL	4.28	5.14	5.40	4.50	4.66
AlphaP	5.45	5.32	4.87	6.71	6.16
Beta 2	2.53	2.58	2.55	3.64	3.50
Conf1	2.71	2.00	0.93	6.29	4.87
Conf2	2.84	3.64	2.94	7.14	5.99
Conf3	0.00	0.00	0.00	0.00	0.00
Conf4	4.13	5.35	4.49	8.71	7.58
Conf5	3.88	1.95	0.72	5.54	4.00
Conf6	2.20	2.80	2.71	6.16	5.84
Conf7	5.77	5.34	5.41	7.42	7.23
Conf8	4.16	7.68	7.82	7.61	7.50
Conf9	6.92	8.42	9.03	7.18	7.52
Conf10	6.99	7.35	6.82	9.02	8.26
AlphaR	8.40	6.16	3.80	5.54	3.45

[a] Conformational energies in kcal/mol. Torsional restraints were applied to all conformers except C7eq, C7ax, C5 of alanine dipeptide, and Conf1, Conf3 of alanine tetrapeptide. The mean RMS of distances between the minimized and the *ab initio* structures are 0.20, 0.21, 0.20, and 0.22 Å for the four models from left to right, respectively.

[b] *Ab initio* energies of C7ax, C5 and AlphaR of alanine dipeptide are from ref. 28 and other conformational data of the dipeptide are from ref. 27. All conformational data of alanine tetrapeptide are from ref. 27 except the AlphaR conformation, which is obtained from private correspondence.

cules in this study provides an opportunity to revisit the parameterizations of the Ψ, Φ torsional angles retaining the same bond, angle, and nonbonded parameters as those of the Cornell et al. model. For example, the parameter H—C_{sp3}—C=O for carbonyl compounds will change peptide energies unless one uses different atom types for ketone C=O other than amide C=O, which we choose not to do. In addition, the availability of more accurate *ab initio* energies for di- and tetrapeptides than when the Cornell et al. model was derived provides further motivation to rederive the Ψ, Φ torsional parameters. We thus attempted to improve our model for peptides and proteins by optimizing the Ψ, Φ torsional parameters to reproduce the relative *ab initio* energies. Once again, the parameter optimization engine *parmscan* was used to optimize the torsional parameters of peptides, which are N—C—C—N (V_1 [1.70, 180.0], V_2 [2.00, 180.0]) and C—N—C—C (V_1 [0.80, 0.0], V_2 [0.85, 180.0]). Our main aim was to reproduce to the extent possible the relative energies of alanine dipeptides and tetrapeptides: In the present case we could not able to give a clear "physical picture" of every Fourier component of the torsional angles.

The six alanine dipeptide conformational energies were selected as a training set to optimize the torsional parameters, and 11 alanine tetrapeptide conformational energies were used to test the parameters. Table XI lists the relative conformational energies of our new model and that of the Cornell et al. model. Both the additive model and the nonadditive model without lone pairs were investigated. Table XII summarizes the results of Table XI. For the alanine dipeptides (the training set), the average absolute errors are 0.21 and 0.58 kcal/mol for this model and the Cornell model, respectively. For the alanine tetrapeptides (the test set), the average absolute errors are 1.21 and 2.58 kcal/mol, respectively. When the polarizable energy is included, the performance of the Cornell et al. model is improved slightly, which now has an average absolute error of 2.20 kcal/mol. The new model with polarization is slightly worse than the additive one, but we have not reoptimized the torsional parameters for this case. It must be emphasized that a torsional

TABLE XII.
Comparison of Performance of Parm99 and Parm94 for Alanine Peptides.[a]

		Parm 99 (additive model)		Parm 99 (nonadditive model)		Parm 94 (additive model)		Parm 94 (nonadditive model)	
	Number	Average of absolute error	RMS dev.	Average of absolute error	RMS dev.	Average of absolute error	RMS dev.	Average of absolute error	RMS dev.
Alanine dipeptide	7	0.21	0.35	0.39	0.55	0.58	0.73	0.53	0.64
Alanine tetrapeptide	11	1.21	1.56	1.53	2.20	2.58	3.00	2.20	2.69
Total	18	0.82	1.24	1.09	1.74	1.80	2.37	1.55	2.14

[a] Conformational energies in kcal/mol.

constraint was added if the final structure deviated greatly from the optimized structure found in the *ab initio* calculations. This makes sense because many of the *ab initio* structures were optimized only at HF/6-31G*, which does not include any dispersion energies. Table XI also shows which conformation was minimized with the torsional constraint. Least-square fittings were performed for the minimized and the *ab initio*-optimized structures and the average RMSDs were found to be approximately 0.2 Å for the aforementioned models. One should note that, for those conformations minimized without adding torsional constraints, our new model has a somewhat larger RMSD than that of the Cornell et al. model. For example, for the C7eq conformation of the dipeptide, the RMSD in our model is 0.34 Å compared with 0.29 Å in the Cornell et al. model.

NUCLEIC ACID BASES

Given that we have added some parameters to our molecular mechanical models (e.g., H—C_{sp3}—C_{sp3}—O) and changed others (e.g., C_{sp3}—C_{sp3}—C_{sp3}—C_{sp3}) that will affect the molecular mechanical energies for nucleic acids, it was also necessary to test how this model reproduces the properties of nucleosides. Recently, we modified our torsional parameters for χ (OS—CT—N*—CK, OS—CT—N*—CM) and CT—OS—CT—N*, and also OS—CT—CT—OS, in order to better reproduce the sugar-puckering properties and χ value of nucleosides to derive Parm98.[52] Despite this, one of the less desirable properties of Parm98 when compared with Parm94 was the increased barrier between C2′ endo and C3′ endo sugars. Thus, this study enabled us to revisit this parameterization and to achieve nucleoside torsional energies better than or comparable to Parm98. We achieved this with a significantly smaller V_2 (1.15, 0.0) instead of (1.5, 0.0) for OCCO and V_2 (0.65, 0.0) instead of (1.0, 0.0) for CT—OS—CT—N* (Table XIII). Furthermore, the new parameters lead to a significantly lower barrier ($\Delta E_{O1'endo-C2'endo}$ = 2.15 kcal/mol for Parm99), comparable to that of Parm94 (2.61 kcal/mol) and significantly less than that of Parm98 (3.27 kcal/mol). Table XIV compares the structural parameters of the minimized structures of our new model and those of MP2/6-31G*. It can be seen that MP2/6-31G* and Parm99 obtain quite similar structures.

TABLE XIII.
Conformational Energies Difference for DNA and RNA Bases.[a]

Base name and puckers	*Ab initio* (HF/6-31G*)	*Ab initio* (MP2/6-31G*)	Parm99	Parm98	Parm94
Deoxyadenosine, C3 endo_anti–C2 endo_anti	0.35	0.37	0.36	0.27	0.48
Riboadenosine, C3 endo_anti–C2 endo_anti	−1.00	−1.85	−1.76	−2.07	−1.59
Deoxycytidine, C3 endo_anti–C2 endo_anti	−0.46	−0.39	0.67	0.71	0.84
Deoxyguanosine, C3 endo_anti–C2 endo_anti	0.75	0.78	0.59	0.53	0.73
Deoxythymidine, C3 endo_anti–C2 endo_anti	0.65	0.90	1.05	1.10	1.00

[a] Conformational energies in kcal/mol.

TABLE XIV.
Structural Parameters for DNA and RNA Bases.[a]

Base name and puckers		*Ab initio* (MP2/6-31G*)			Parm99		
		q	*W*	χ	*Q*	*W*	χ
Deoxyadenosine	C2 endo_anti	0.34	163.90	228.00	0.32	163.00	221.30
	C3 endo_anti	0.36	12.40	198.50	0.34	7.48	205.30
Riboadenosine	C2 endo_anti	0.37	163.20	233.70	0.36	169.26	237.80
	C3 endo_anti	0.40	10.30	193.70	0.40	3.66	189.70
Deoxycytidine	C2 endo_anti	0.34	160.60	205.00	0.34	155.46	216.90
	C3 endo_anti	0.36	15.00	195.20	0.33	15.68	208.40
Deoxyguanosine	C2 endo_anti	0.34	164.40	231.90	0.33	162.95	226.20
	C3 endo_anti	0.35	14.20	203.20	0.33	7.97	211.20
Deoxythymidine	C2 endo_anti	0.35	159.60	226.70	0.34	158.42	220.50
	C3 endo_anti	0.35	16.60	199.00	0.33	16.02	211.30

[a] *q* and *W* are defined in ref. 50; χ is defined in ref. 51; angles in degrees; conformational energies in kcal/mol.

MOLECULAR DYNAMICS SIMULATION ON UBIQUITIN AND d(CGCGAATTCGCG)$_2$

Molecular dynamics simulations were performed for the protein ubiquitin and DNA dodecamer deoxynucleic acid duplex using both our new model and the Cornell et al. model (and in the case of the DNA, Parm98). Figure 5a shows that the RMS deviation from the crystal structure of ubiquitin is comparable for the two models, consistent with a previous analysis by Fox and Kollman.[53] For DNA, Parm98 and Parm99 give comparable RMSDs for the crystal structure, both of which are lower than in Parm94 (Fig. 5b). For DNA, Parm99 shows an A DNA to B DNA transition at the same time scale as for Parm94[54] (Fig. 5c).

Discussion

Molecular mechanics methods have now become a standard tool for chemists to study the structure and conformational energy as well as noncovalent molecular interactions. To calculate these properties as accurately as possible, one should choose a good force field model. Which factors are responsible for a successful force field model? First, accurate treatment of electrostatic interactions is crucial for correctly describing the intermolecular energies, which is also an important factor for differentiating the force fields that have similar functional forms. Second, a successful force field should have accurate and enough specific parameters—especially torsional parameters, which are important for describing the geometry and conformational energy. Good parameter transferability is another crucial element for a successful molecular mechanical model. In our opinion, the minimal functional form in eq. (1) should be adequate to study geometry, intermolecular interactions, and conformational energies accurately for all but the most restrained molecules. Among the several widely used force fields, MM2/MM3 and MMFF apply more complex functional forms, including anharmonic stretch and bend and cross terms, whereas AMBER, OPLS, DREIDING, and other harmonic force fields use the simpler functional forms of eq. (1). MM3 and MMFF usually represent conformational energies better than the simpler force fields [eq. (1)], but one should remember that MM3 and MMFF are also two of the most extensively parameterized force fields. Many torsional parameters are specified and, of these, many have three Fourier terms. The results presented suggest that it is not the functional form that is crucial in representing conformational energies but rather the parameterization. We have shown that a well-parameterized model, although simple, is capable of studying tasks for both organic *and* bioorganic systems.

Different force fields have different methods for handling electrostatic interactions. Some force fields calculate atomic charges with heuristic algorithms that are based on electronegativities (e.g., MMFF). Other force fields, such as the OPLS,[6] derive charge empirically, based on fitting to liquid properties (structure, vaporization, and sublimation enthalpy). Both the Weiner et al.[2] and the Cornell et al.[1] versions of AMBER force fields use electrostatic potential charges based on quantum-mechanical calculation of fragments (RESP in the case of the

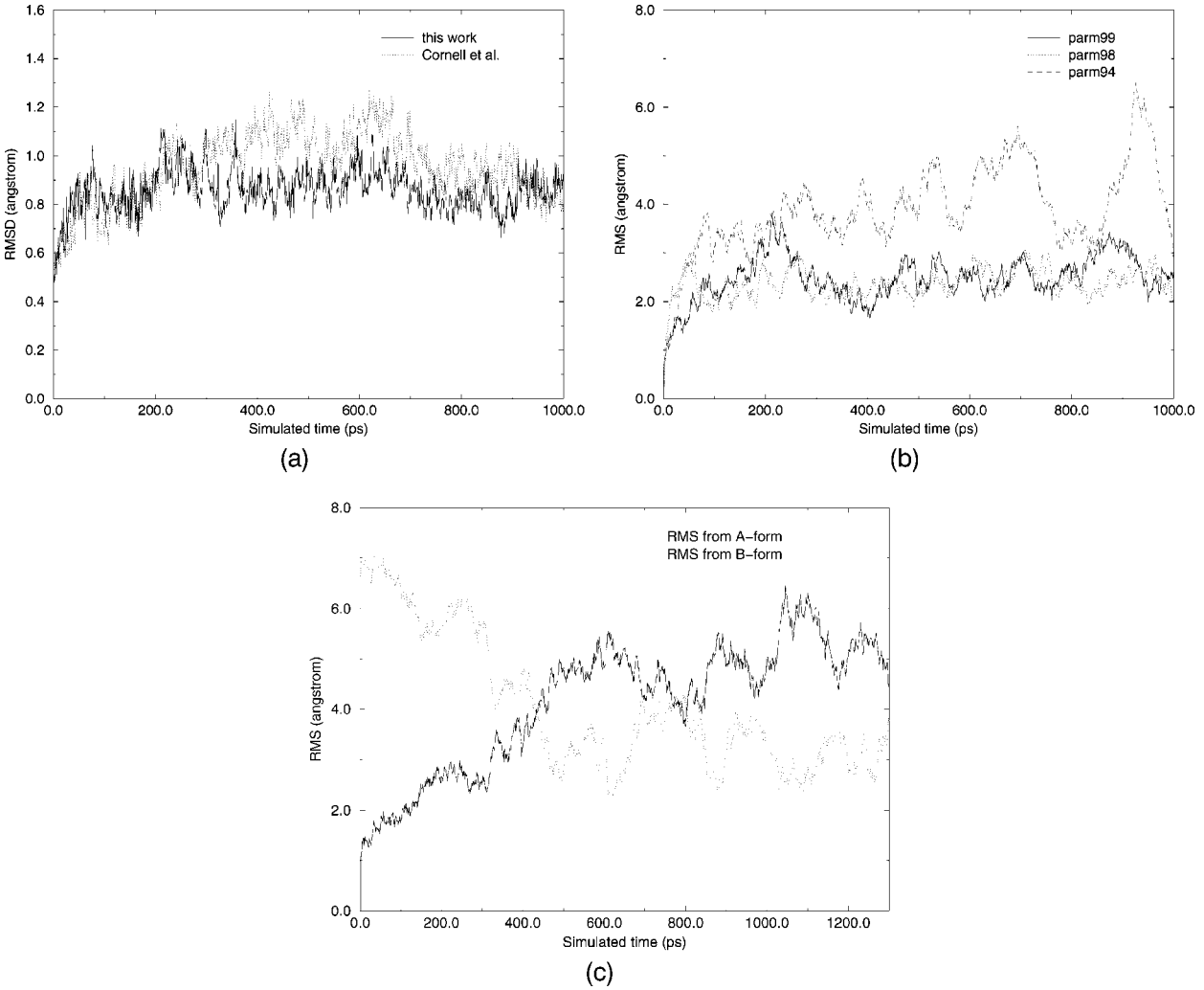

FIGURE 5. (a) RMS deviation (Å) of backbone atoms (only amino acids 1–72 are taken into account as noted in ref. 50) between the crystal structure of ubiquitin and the structure along the MD trajectory as given by the Cornell et al. model (dotted line) and our new model (solid line). Average RMS deviations are 0.95 and 0.86 Å for the Cornell et al. model and the new model, respectively. (b) All-atom RMS deviation (Å) as a function of time with respect to initial structure for the Dickerson dodecamer d(CGCGAATTCGCG)$_2$ MD simulation in aqueous solution. Average RMSs over the 1-ns trajectory are 2.46, 2.33, and 3.92 Å for Parm99, Parm98, and Parm94, respectively. (c) All-atom RMS deviation (Å) as a function of time for the Dickerson dodecamer d(CGCGAATTCGCG)$_2$ MD simulation initiated from the A-form in aqueous solution. Solid and dotted lines are RMSs with respect to initial A-form structure and minimized crystallographic B-form structure, respectively.

Cornell et al.[1]) to describe the electrostatic interactions. Although the previous electrostatic potential models (Weiner et al.[2] and Cornell et al.[1]) have not been widely parameterized to reproduce the relative conformational energy of a large number of organic systems, they have done a reasonable job in describing intermolecular interactions. Gundertofie suggested that it is the treatment of charges that made AMBER significantly better than other harmonic force fields with regard to such a test (see ref. 24). Hobza et al. also showed that the Cornell et al. model represented nucleic acid–base interactions better than any other MM and many quantum-mechanical models relative to high-level *ab initio* calculations.[55]

As to which *ab initio* model is appropriate for the derivation of RESP charges, it makes sense to choose the 6-31G* basis set for an effective two-body model [eq. (1)] for simulations in polar media. It has been suggested that the 6-31G*-based ESP-fit charge

model is capable of providing an excellent reproduction of condensed-phase intermolecular properties such as liquid enthalpies and densities as well as free energies of solvation.[21] Moreover, Kuyper et al. suggested that the logical choice of basis set for ESP-fit charges for use in condensed phases is the 6-31G* basis set,[56] which uniformly overestimates molecular polarity; that is, it is well balanced with the commonly available water models (SPC/E,[57] TIP3P,[34] TIP4P[34]), which have dipole moments that are about 20% higher than the gas phase value for water. Unfortunately, the electrostatic potential (EP)-based charge methods are less than ideal for two reasons. First, one must carry out quantum-mechanical calculations to obtain the electrostatic potentials before performing the least-squares fit. Thus, it cannot be applied for deriving charges as efficiently as the empirical methods, which may be a bottleneck for database screening studies. Second, charges generated using different conformations of a molecule are usually not identical. The variation can be $\geq$10% for those charges that have large absolute values (>0.1). We should point out that this deficiency is much more serious for standard ESP charges. Although it is partially mitigated by the RESP charge-fitting scheme, it may remain unclear as to how to determine which conformations should be chosen for deriving the RESP charges.

To investigate how this drawback can affect the molecular mechanical energies, we designed three charge schemes to calculate the relative energies for five fluorides in Table I. Charge sets I and II were RESP charges derived from the electrostatic potentials of the lower energy conformation and those of the higher energy conformation, respectively. Charge set III was generated by multiple molecular fitting procedures using the electrostatic potentials of both conformers. The torsional parameters used for all three schemes were derived based on charge set I, and it is not surprising that the best result was obtained for this charge set. The AAE compared with the experiment is 0.24 kcal/mol. For charge set II, the AAE was 0.58 kcal/mol, and for charge set III it was 0.35 kcal/mol, which is slightly worse than charge set I. We conclude that the multiple-molecule RESP charge scheme performs adequately and can be used effectively for complicated molecules in which the global minimum conformation is not obvious. A similar tendency was found in the Cornell et al.[1] model (Parm94). We suggest one may use the RESP charges derived from the electrostatic potentials of the more stable conformers, or use the multiple-molecule RESP charges.

Another key characteristic of a successful force field is parameter transferability. One can consider "transferability" on two different levels. First, each parameter should not only work well for the model molecules in the training set, but also for the other related molecules. Second, all the parameters of the force field should be consistent. For instance, as to the torsional parameters of HC—CT—CT—X (X can be F, Cl, or Br), not only should the Fourier series and phase angles be the same, but the force constants should also follow the periodic table.

To improve transferability one can use *parmscan* or another heuristic search program[37–39] to derive the parameters for all representative test cases available. The more high-quality the experimental or high-level the *ab initio* data, the more reliably the parameters can be transferred to the test sets. As to how to collect high-quality reference data, there are several possibilities. First, the molecules must be chosen to reflect the chemistry area for which the force fields are developed. Second, one should understand that the relative energies of molecular mechanics are steric energies (ΔE), which can usually be substituted for the enthalpy difference (ΔH). However, these data cannot be used to predict the free energy differences (ΔG) unless there is good reason to suspect that the entropy difference (ΔS) is near zero. Although the reference data should include the experimentally determined enthalpy differences and the energy differences from the potential energy curves in the gas phase, such values are relatively scarce, and, to obtain a reasonable scope for the test set, a few instances of ΔG may need to be included.

It is the parameter developer that is responsible for the second level of transferability—the parameter consistency of the whole force field. There are some encouraging examples of parameter consistency with our new model. As an example of torsional parameters of halides, both H—C_{sp3}—C_{sp3}—X and X—C_{sp3}—C_{sp3}—X (X can be F, Cl, or Br) apply the same Fourier components with the same phase angles for all three kinds of halides. Moreover, the torsional force constants follow the periodic table. For H—C_{sp3}—C_{sp3}—X, the force constant changes from 0.19, 0.25, to 0.55 kcal/mol for fluorides, chlorides, and bromides, respectively; for X—C_{sp3}—C_{sp3}—X, the force constant changes from 1.20, to 0.45, to 0.0 kcal/mol for fluorides, chlorides, and bromides, respectively. Furthermore, the torsional parameter of H—C_{sp3}—C_{sp3}—O also applies the onefold Fourier component with the phase angle of 0.0°. The force constant (0.25 kcal/mol) is the same as that of chloride. Another example

is that, for both C_{sp3}—C_{sp3}—C_{sp3}—C_{sp3} and C_{sp3}—C_{sp3}—N_{sp3}—C_{sp3}, a onefold as well as a threefold term can improve the performance of reproduction of the relative energies. The third example is that all the following parameters apply a twofold Fourier component with a phase angle of 180.0° to represent planar systems: X—CA—CA—X (CA is the carbon in pure benzene); X—C_{sp2}—N_{sp2}—X, X—CD—CD—X (CD is the middle carbon in butadiene); and X—C_{sp2}—CT—X. The force constants change from 14.5 to 10.0 to 4.0 to 0.0 kcal/mol. The tendency responses the strength order of the double bonds formed between the two middle atoms.

The more complex intramolecular force field involves anharmonic terms and coupling terms such as "stretch–bend" in addition to the harmonic terms of eq. (1). This kind of model may be necessary to study other molecular properties (e.g., vibrational frequencies) besides structure and inter- and intramolecular interactions. Including polarization and three-body exchange repulsion can also extend a simple additive force field.[58] Polarization energy can be described using eqs. (2)–(7) and it can be solved for self-consistently. In most cases, it is very difficult to extend a simple additive model to a nonadditive one while keeping most of the parameters unchanged. We are encouraged that our new model performs reasonably well in reproducing the relative energies for the three polarizable models with the same torsional parameters developed for the additive model. Furthermore, we find that performance can be improved significantly after slight adjustment of some of the torsional parameters (Table X). Of course, using the high-level (B3LYP-pVTZ) *ab initio* RESP-based charges with polarization should enable more accurate intermolecular interaction in various different solvent environments. In the future, we plan to develop a complete nonadditive force field for proteins, nucleic acids, and organic molecules.

Conclusion

A new version of the AMBER force field has been developed using 82 conformational pairs with RESP charges. For the pairwise additive model, very encouraging results have been obtained upon adding a limited set of torsional parameters. The new model gives the smallest AAE and RMS deviation compared with experimental data for both 34- and 55-molecule sets of MMFF, MM3, and CHARMm. As for comparison to high-level *ab initio* methods, the present model also provides excellent results, and is comparable in performance to MMFF. Considering that all the molecules chosen for comparisons are selected from our "training set" and MM3 and CHARMm may have not included all molecules for parameterization, the comparison with the two force fields might be biased. To make comparisons more valid and to investigate the transferability of the torsional parameters, we took 11 hydrocarbons and 7 chlorides for further study by dividing them into a training set and a test set. The absolute errors of molecules in the test set are only slightly larger than in the training set (differences <0.1 kcal/mol). In summary, we conclude that a well-parameterized harmonic force field with a reliable charge method can describe the structure and intramolecular energies for organic systems very well.

It is very encouraging to point out that nonadditive models also give reasonable results using the same parameters derived for the additive model, although additional torsional parameterization is required to achieve the same high level of accuracy as that found using the additive model.

The torsional angle parameters involved in peptides have been optimized by the new program *parmscan* to reproduce the relative energies of alanine dipeptide and tetrapeptide conformers calculated by high-level *ab initio* methods. Although, for those conformers, which are fully flexible during minimization, the new model predicts geometry slightly less accurately than the Cornell et al. model, it reproduces the relative energies much better than the Cornell et al. force field. The average absolute errors for alanine dipeptide and tetrapeptide are 0.21 and 1.21 kcal/mol, respectively. This is noteworthy because the absolute accuracy of the LMP2/cc-pVTZ (-f) relative energies for the tetrapeptide has been estimated to be 1.0 kcal/mol. Moreover, two conformations, Conf8 and Conf9, of alanine tetrapeptide are of rather high energy and are not often found; if we neglect them, the AAE of the test set is reduced to 0.92 kcal/mol, which is within the absolute accuracy of the *ab initio* method.

Parmscan has also been applied to optimize some torsional parameters involved in nucleic acids and the new model achieves DNA structures closer to the crystal structures than Parm94 and comparable to those found with Parm98. In the process of modifying Parm98 to "reach" Parm99, we have been able to lower the sugar-pucker barrier of Parm98 to a level comparable to that of Parm94, which should enable it to sample conformational space as efficiently as Parm94.

Acknowledgment

P.A.K. is grateful to acknowledge research supported from NIH GM-29072 and NIH CA-25644 (P. Kollman, P. I.) and NIH GM-56609 (E. Arnold, P. I.)

Supplementary Material

The parameter file (Parm99) and topology files for the organic molecules in our test sets are available as Supplementary Material. All the atom types of Parm99 are the same as those in the Cornell et al. model (Parm94), with the exception of the sp^3 neutral nitrogen (NT) and the inner sp^2 carbon of dienes (CD). Detailed information on the nonbond, bond, and angle parameters of the new atom types are available in the "Results" section of this work.

References

1. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. J Am Chem Soc 1995, 117, 5179–5197.
2. 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, 765–784.
3. Burkert, U.; Allinger, N. L. Molecular Mechanics; American Chemical Society: Washington, DC, 1982.
4. Allinger, N. L. J Am Chem Soc 1977, 99, 8127.
5. Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. J Am Chem Soc 1989, 111, 8551–8566.
6. Jorgensen, W. L.; Tirado-Rives, J. J Am Chem Soc 1988, 110, 1657–1666.
7. Clark, M.; Cramer, R. D., III; van Opdenbosch, N. J Comput Chem 1989, 10, 982.
8. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. J Comput Chem 1983, 4, 187–217.
9. Lifson, S.; Hagler, A. T.; Dauber, P. J Am Chem Soc 1979, 101, 5111–5121.
10. Hwang, M. J.; Stockfish, T. P.; Halgler, A. T. J Am Chem Soc 1994, 116, 2515–2525.
11. Momany, F. M.; Rone, R. J Comput Chem 1992, 13, 888–900.
12. Mayo, S. L.; Olafson, B. D.; Goddard, W. A., III J Phys Chem 1990, 94, 8879.
13. Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A., III; Skiff, W. M. J Am Chem Soc 1992, 114, 10024–10035.
14. Halgren, T. A. J Comput Chem 1996, 17, 490–519.
15. Halgren, T. A. J Comput Chem 1996, 17, 520–552.
16. Halgren, T. A. J Comput Chem 1996, 17, 553–586.
17. Halgren, T. A. J Comput Chem 1996, 17, 587–615.
18. Halgren, T. A. J Comput Chem 1996, 17, 615–641.
19. Halgren, T. A. J Comput Chem 1999, 20, 730–748.
20. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. J Phys Chem 1993, 97, 10269–10280.
21. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. J Am Chem Soc 1993, 9620–9631.
22. Fox, T.; Kollman, P. A. J Phys Chem 1998, 102, 8070–8079.
23. Cieplak, P.; Cornell, W. D.; Bayly, C. I.; Kollman, P. A. J Comput Chem 1995, 16, 1357–1377.
24. Gundertofie, K.; Liljefors, T.; Norrby, P.-O.; Pettersson, I. J Comput Chem 1996, 17, 429–449.
25. Howard, A. E.; Cieplak, P.; Kollman, P. A. J Comput Chem 1995, 16, 243–261.
26. Rychnovsky, S. D.; Yang, G.; Powers, J. P. J Org Chem 1993, 58, 5251–5255.
27. Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. J Am Chem Soc 1997, 119, 5908–5920.
28. Gould, I. R.; Kollman, P. A. J Comput Chem 1992, 13, 971–978.
29. Cornell, W. D.; Gould, I. R.; Kollman, P. A. J Mol Struct Theochem 1997, 392, 101–109.
30. Halgren, T. A. J Comput Chem 1999, 20, 720–729.
31. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. 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-94; Gaussian, Inc., Pittsburgh, PA, 1995.
32. Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Ferguson, D. M.; Radmer, R. J.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. AMBER-5; University of California at San Francisco, San Francisco, CA, 1997.
33. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. J Chem Phys 1995, 103, 8577–8593.
34. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. J Chem Phys 1983, 79, 926–935.
35. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. J Comput Phys 1977, 23, 327–341.
36. Drew, H. R.; Dickerson, R. E. J Mol Biol 1981, 151, 535–556.
37. Liang, G.; Fox, P. C.; Bowen, J. P. J Comput Chem 1996, 17, 940–953.
38. Norrby, P. O.; Liljefors, T. J Comput Chem 1996, 19, 1146–1166.
39. Faller, R.; Schmitz, H.; Biermann, O.; Müller-Plathe, F. J Comput Chem 1999, 20, 1009–1017.
40. Radom, L.; Hehre, W. J.; Pople, J. A. J Am Chem Soc 1972, 94, 2371–2381.
41. Brunck, T. K.; Weinhold, F. J Am Chem Soc 1979, 101, 1700–1709.
42. Heenan, R. K.; Bartell, L. S. J Chem Phys 1983, 78, 1270–1274.
43. Cornell, W. D.; Ha, M. P.; Sun, Y.; Kollman, P. A. J Comput Chem 1996, 17, 1541–1548.
44. Carreira, L. A.; Jiang, G. L.; Person, W. B.; Willis, J. N. J Chem Phys 1972, 56, 1440.

45. Allinger, N. L.; Grev, R. S.; Yates, B. F.; Schaefer, H. F., III J Am Chem Soc 1990, 112, 114–118.
46. Hirota, E.; Emdo, Y.; Saito, S.; Duncan, J. J Mol Spectrosc 1981, 89, 285–295.
47. Payne, P.; Allen, L. C. Modern Theoretical Chemistry: Applications of Electronic Structure Theory; Plenum: New York, 1987; Chapter 2.
48. Lowe, J. P. Prog Phys Org Chem 1968, 6, 1.
49. Applequist, J.; Carl, J. R.; Fung, K.-K. J Am Chem Soc 1972, 94, 2952.
50. Cremer, D.; Pople, J. J Am Chem Soc 1975, 97, 1354–1358.
51. Saenger, W. Principles of Nucleic Acid Structure; Springer: Tokyo, 1984.
52. Cheatham, T. E., III; Cieplak, P.; Kollman, P. A. J Biomol Struct Dynam 1999, 16, 845–862.
53. Fox, T; Kollman, P. A. Prot Struct Funct Genet 1996, 25, 315–334.
54. Cheatham, T. E., III; Kollman, P. A. J Mol Biol 1996, 259, 434–444.
55. Hobza, P.; Kabelác, M.; Šponer, J.; Mejzlík, P.; Vondrášek, J. J Comput Chem 1997, 18, 1136–1150.
56. Kuyper, L.; Ashton, D.; Merz, K. M., Jr.; Kollman, P. A. J Phys Chem 1991, 95, 6661–6666.
57. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. J Phys Chem 1987, 91, 6269–6271.
58. Dixon, R. W.; Kollman, P. A. J Comput Chem 1997, 18, 1632–1646.