

Force Field Modeling of Conformational Energies: Importance of Multipole Moments and Intramolecular Polarization

THOMAS D. RASMUSSEN,¹ PENGYU REN,² JAY W. PONDER,³
FRANK JENSEN¹

¹*Department of Physics and Chemistry, University of Southern Denmark, DK-5230 Odense, Denmark*

²*Department of Biomedical Engineering, University of Texas at Austin, 1 University Station, C0800, ENS 618, 2501 Speedway, Austin, Texas 78712-0238, USA*

³*Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110, USA*

Received 4 October 2006; accepted 31 October 2006

Published online 5 December 2006 in Wiley InterScience (www.interscience.wiley.com).

DOI 10.1002/qua.21278

ABSTRACT: The ability of force fields to reproduce relative conformational energies for seven molecules is probed. The correlation against LMP2/cc-pVTZ results for standard force field employing fixed partial charges deteriorates as the molecules become more polar. Inclusion of multipole moments and intramolecular polarization can improve the correlation, and both contributions are of similar importance. © 2006 Wiley Periodicals, Inc. *Int J Quantum Chem* 107: 1390–1395, 2007

Key words: force field; polarization; conformations

Introduction

Force field methods parameterize the internal energy of a molecule by an additive energy function containing constants that are fitted to re-

produce experimental or quantum mechanical results [1, 2]. All force fields contain the terms shown in eq. (1), and additional terms may be added to improve the performance:

$$E_{\text{FF}} = E_{\text{str}} + E_{\text{bend}} + E_{\text{tors}} + E_{\text{vdw}} + E_{\text{el}} \quad (1)$$

The stretch and bending terms are relatively easy to parameterize, and they have little influence on conformational degrees of freedom. The nonbonded

Correspondence to: F. Jensen; e-mail: frj@ifk.sdu.dk

Contract grant sponsors: Danish Natural Science Research Council; Danish Center for Scientific Computing.

van der Waals and electrostatic terms determine intermolecular interactions, and are thus very important for calculating, e.g., binding affinities of drug molecules to enzymes. The nonbonded terms together with the torsional energy determine the internal (conformational) degrees of freedom. The existence and relative stabilities of conformations for a given molecule is thus determined by a delicate balance among these three energy terms. Being able to predict the energetic ordering of conformations correctly is an essential quantity for bringing force field methods into the predictive regime of theoretical methods.

The van der Waals and electrostatic parameters are normally assigned to be atomic quantities, and are primarily aimed at generating good intermolecular interaction energies. The torsional energy is subsequently parameterized as a residual correction to reproduce rotational barriers and energetic preferences. The electrostatic energy is unimportant for nonpolar systems like hydrocarbons, and is often completely neglected for such cases. The conformational space is consequently determined by E_{vdw} and E_{tors} . Since the van der Waals interaction is short-ranged, this means that the transferability assumption inherent in force field methods is valid, and torsional parameters determined for small model systems can be used for predicting conformations also for large systems.

For polar systems, however, the long-ranged electrostatic interaction is often the dominating energy term, and the use of torsional parameters determined for small model systems becomes more problematic. In the majority of commonly used force fields, the electrostatic interaction is parameterized in terms of a Coulombic interaction between partial atomic charges, with the latter typically determined by fitting to the electrostatic potential sampled in the near vicinity of the molecule:

$$E_{\text{el}} = \sum_{i>j}^N \frac{Q_i Q_j}{\epsilon R_{ij}} \quad (2)$$

The dielectric constant ϵ is sometimes given a value larger than one or made distant dependent to model the implicit screening by solvent or other molecules. It is well known that the partial atomic charges depend on the molecular geometry [3], but this issue is ignored in fixed charge force fields. The procedure for determining atomic charges based on intermolecular interactions, the neglect of the ge-

ometry dependence, and the transfer of torsional parameters from small model systems, suggest that the conformational space is less well reproduced for polar systems. In a previous publication, we have shown that the correlation coefficient for relative conformational energies against accurate reference data deteriorates as the systems become more polar [4]. In the present case we examine whether inclusion of higher-order multipoles and polarization for representing the electrostatic interaction can improve the performance for predicting relative conformational energies for a test set of seven molecules.

Computational Details

All force field calculations have been done using the MacroModel [5] program for fixed charge force fields and with the Tinker [6] program for the Amoeba polarizable force field. The MM2, MM3, AMBER, and MMFF results all refer to the versions implemented in the MacroModel package, where the AMBER results refer to the AMBER94 parameter set. Electronic structure calculations have been performed with the Gaussian [7], Jaguar [8], and GAMESS [9] program packages. All calculations have been carried out in vacuum with a dielectric constant of one. Relative energies have been calculated at the LMP2/cc-pVTZ level on MP2/cc-pVDZ optimized geometries, generated from partial scanning of potential energy surfaces (PESs). Although we do not claim to have performed an exhaustive scan of the whole surface, the number of conformations located should represent a substantial fraction of the minima on the PES. The quality of the force field energies is evaluated based on single point energy calculation on the MP2/cc-pVDZ optimized geometries.

Parameterization Procedure

The representation of the electrostatic interaction for a given geometry can be improved by including higher-order electric moments, and the geometry dependence can be included by allowing polarization of these moments. Recent developments in this direction have focused on adding static dipole and quadrupole moments and dipole polarizabilities at atomic sites [10]; this is also the approach taken in the present study [11]. A nontrivial issue is how the static moments and the polarization effects are sep-

arated, and which reference data the force field is parameterized against. The parameterization procedure employed has been described earlier [12], and only a short overview will be given.

The static moments are assigned based on a procedure using the original version of the distributed multipole analysis (DMA) developed by Stone [13]. The results from this analysis are in the form of multipole expansions at atomic sites, expressed as spherical tensors. The expansion is truncated at $l = 2$ (quadrupoles) and converted to local frames with the atomic sites as origins:

$$\mathbf{M}_i = \lfloor q_i \quad \mu_{i,\alpha} \quad \Theta_{i,\alpha\beta} \rfloor; \quad \alpha, \beta = \{x, y, z\}. \quad (3)$$

These multipole moments can be regarded as composed of a permanent and an induced contribution:

$$\mathbf{M}_i = \mathbf{M}_i^{\text{perm}} + \mathbf{M}_i^{\text{ind}}. \quad (4)$$

$\mathbf{M}_i^{\text{ind}}$ is produced by direct and mutual induction from all sites in the absence of an external field:

$$\mathbf{M}_{i,\alpha}^{\text{ind}} = \alpha_i \left(\sum_j T_{\alpha}^{ij} \mathbf{M}_j^{\text{perm}} + \sum_{j'} T_{\alpha}^{ij'} \mathbf{M}_{j'}^{\text{ind}} \right), \quad (5)$$

where T represents the interaction matrix elements. The first term in the above equation corresponds to induction due to the static multipoles, while the second term arises from induced multipoles at other sites. Here α_i represents atomic isotropic polarizabilities generated from a Thole model, where a smeared charge distribution $\rho(\mathbf{r})$ is used to prevent the polarization catastrophe [14]:

$$\rho(u) = \frac{3a}{4\pi} e^{-au^3}. \quad (6)$$

The quantity $u = R_{ij}/(\alpha_i\alpha_j)^{1/6}$ is the effective distance as a function of the atomic polarizabilities at site i and j , and α is a dimensionless constant controlling the damping. Inserting eq. (4) in eq. (5), and assuming that the same scaling factors are employed for mutual and direct induction using identical groups j and j' , produces

$$\mathbf{M}_{i,\alpha}^{\text{ind}} = \alpha_i \left(\sum_{j \neq i} T_{\alpha}^{ij} \mathbf{M}_j \right). \quad (7)$$

In the case in which j and j' are not the same, or when different fractional scaling factors are desired

for short-range direct and mutual interactions, a more complex formulation requiring iterative solution is necessary. The permanent moments can now be calculated from eq. (4) and will return the truncated quantum mechanical DMA result when combined with the force field polarization model defined by Eq. (5).

It is well known that the original DMA version generates distributed multipole moments that do not converge as the basis set is improved. Especially the use of diffuse functions is troublesome and the problem is further aggravated as the size of the system increases. However, it is also known that the present of diffuse functions substantially improves the prediction of total molecular moments and polarizabilities. A recent modification of the DMA procedure shows improved stability for deriving atomic parameters [15], but the present data was generated before the release of the modified procedure. For developing force field parameters, a necessary assumption is that atomic parameters are transferable, i.e. parameters derived from a (small) model system can be used for large systems without modifications. The systems employed in the present case are small enough that this assumption can be tested explicitly. We will compare the results from deriving multipoles for the entire system with the results from using smaller model systems. The DMA parameters are calculated at the MP2 level with the 6-311G(d,p) basis set for the full and partial systems, while the 6-311++G(2d,2p) basis set is used only for the partial systems.

Results

We have chosen the seven molecules shown in Table I as our test systems, representing a progression from nonpolar to zwitterionic structures. The molecules are large enough to have a nontrivial set of conformations, but small enough that reasonably high level electronic structure calculations can be employed to establish a suitable set of reference data. We have performed a partial scan of the PESs for each molecule at the MP2/cc-pVDZ level with the resulting number of unique minima given in Table I. For compound 5 we restrained the OH distance in the carboxyl group in order to prevent a migration of the acid hydrogen to form compound 7. Improved estimates of relative energies have been obtained from single-point LMP2/cc-pVTZ calculations, and these results are used as our reference energies.

TABLE I
Systems used in the present work.*

Label	Structure	Conformations
Mol 1		12
Mol 2		10
Mol 3		8
Mol 4		23
Mol 5		58
Mol 6		43
Mol 7		20

* Conformations refer to number of different conformations optimized at the MP2/cc-pVDZ level of theory.

The performance of various force fields with respect to relative energies of the unique minima in Table I is quantized based on the correlation coefficient (R^2) and slope (α) of the least squares fitted line against the reference data and is shown in Table II. All the force fields provides satisfactory performance for molecules 1 and 2, but as the polarity of the systems increases, the correlation with the LMP2 reference data becomes erratic, and all the standard force fields have problematic perfor-

mance for at least two of the seven systems. We interpret this as an inadequacy of using fixed partial charges to model the electrostatic interaction, i.e., neglect of multipole moments and intramolecular polarization.

To improve the treatment of the electrostatic interaction, we have tested three different versions of the Amoeba force field, which include polarization as described above. Version (1) uses multipole parameters derived from the whole

TABLE II
Correlation of force field relative energies against LMP2/cc-pVTZ results.*

Force field	Mol 1		Mol 2		Mol 3		Mol 4		Mol 5		Mol 6		Mol 7	
	α	R^2	α	R^2	α	R^2	α	R^2	α	R^2	α	R^2	α	R^2
AMBER	1.01	0.99	1.17	0.99	0.20	0.16	1.68	0.65	0.53	0.28	1.45	0.49	0.93	0.73
MM2	1.05	0.99	1.05	0.92	1.54	0.91	1.21	0.62	-1.00	0.04	0.91	0.36	-0.04	0.00
MM3	1.12	0.95	1.14	0.95	1.06	0.74	1.27	0.63	-0.17	0.00	0.56	0.13	-0.14	0.01
MMFF	1.42	1.00	1.26	0.98	0.53	0.35	1.23	0.47	1.40	0.76	1.43	0.55	0.92	0.94
OPLS	1.06	0.96	1.09	0.96	0.79	0.98	1.12	0.17	1.22	0.57	1.14	0.20	0.93	0.87
Amoeba(1)	0.77	0.97	0.97	0.95	0.91	0.93	0.90	0.57	0.81	0.39	0.77	0.20	0.46	0.83
Amoeba(2)	0.78	0.95	0.90	0.94	0.87	0.93	0.95	0.60	0.75	0.35	1.11	0.54	0.69	0.91
Amoeba(3)	0.79	0.94	0.92	0.86	0.94	0.92	1.02	0.63	1.11	0.74	1.27	0.70	0.70	0.90

* Amoeba(1) is parameterized based on MP2/6-311G(*d,p*) results for the entire molecule.

Amoeba(2) is parameterized based on MP2/6-311G(*d,p*) results for smaller molecular systems.

Amoeba(3) is parameterized based on MP2/6-311++G(2*d,2p*) results for smaller molecular systems.

TABLE III
Fragment systems used for parameterization.

Group	Model systems	Target systems
Alkane	CH ₃ CH ₂ CH ₃ , CH ₃ (CH ₂) ₂ CH ₃	1, 2, 3, 5, 6, 7
Ether	CH ₃ OCH ₃ , CH ₃ CH ₂ OCH ₃	2, 3, 4
Amine	CH ₃ NHCH ₃ , CH ₃ CH ₂ NHCH ₃	4, 5
Amide	CH ₃ C(=O)NH ₂ , CH ₃ CH ₂ C(=O)NH ₂	6
Anion	CH ₃ COO ⁻ , CH ₃ CH ₂ COO ⁻	7
Acid	CH ₃ C(=O)OH, CH ₃ CH ₂ C(=O)OH	5
Cation	CH ₃ NH ₂ CH ₃ ⁺ , CH ₃ CH ₂ NH ₂ CH ₃ ⁺	7

system at the MP2/6-311G(*d,p*) level, while the parameters in versions (2) and (3) are derived from smaller model systems (Table III), using the 6-311G(*d,p*) and 6-311++G(2*d,2p*) basis sets, respectively. The torsional parameters have in most cases been reparameterized after assignment of the polarization and multipole moments. It should be noted that deriving parameters with the 6-311++G(2*d,2p*) basis set for the full system produces significantly poorer results due to instabilities in the employed DMA procedure. The good agreement between the Amoeba(1) and Amoeba(2) models shown in Table II indicates that the assumption of transferability is valid for these systems. Table II also shows that the Amoeba models perform significantly better than the regular force fields, and all seven molecules are treated qualitatively correct. The best model in term of correlation with the reference data is Amoeba(3), although the correlation coefficients are only ~0.70 for systems 4, 5, and 6.

To investigate further the importance of the various terms, we have performed additional cor-

relations with the Amoeba(3) model, where either the multipole moments, the polarization, or all the electrostatic terms have been set to zero, as shown in Table IV. In agreement with expectations, the electrostatic and polarization terms are unimportant for molecules 1 and 2. For the systems 3–6, the neglect of either multipole or polarization effects decreases the performance, and both contributions are of similar importance. The conformational space of the zwitterionic system 7 is completely dominated by the localized charges and neglect of multipole and polarization terms have only minor influence. Neglecting the electrostatic contribution entirely in the Amoeba(3.3) model significantly degrades the performance for molecules 3 and 4, produces models with no correlation against the reference data for molecules 5 and 6, and produces a strong anti-correlation for molecule 7. This substantiates the claim that the electrostatic energy plays a dominating role for predicting conformational energies of polar molecules.

TABLE IV
Correlation of force field relative energies against LMP2/cc-pVTZ results.*

Force field	Mol 1		Mol 2		Mol 3		Mol 4		Mol 5		Mol 6		Mol 7	
	α	R^2	α	R^2	α	R^2	α	R^2	α	R^2	α	R^2	α	R^2
Amoeba(3)	0.79	0.94	0.92	0.98	0.94	0.92	1.02	0.63	1.11	0.74	1.27	0.70	0.70	0.90
Amoeba(3.1)	0.86	0.95	0.99	0.88	0.87	0.84	1.01	0.60	0.58	0.11	1.25	0.76	0.57	0.83
Amoeba(3.2)	0.90	0.98	0.91	0.96	1.07	0.83	1.06	0.59	0.72	0.31	1.03	0.56	0.53	0.80
Amoeba(3.3)	0.87	0.96	0.73	0.97	0.97	0.65	0.82	0.38	-0.30	0.01	-0.20	0.04	-0.85	0.93

* Amoeba(3) is parameterized based on MP2/6-311++G(2*d,2p*) results for smaller molecular systems.

Amoeba(3.1) corresponds to neglecting the quadrupole contributions.

Amoeba(3.2) corresponds to neglecting the polarization contributions.

Amoeba(3.3) corresponds to neglecting all of the electrostatic contributions.

Summary

The present results show that inclusion of higher-order electric moments and polarization can improve the performance of force field for reproducing relative conformational energies. Permanent multipole moments and polarization appear to be of similar importance, and should therefore be included at the same stage. The results show that it is possible to derive parameters from carefully chosen smaller model systems, and that the resulting force field display a more consistent performance for all the systems. The correlation analysis indicates that there is still room for improving the optimum way of deriving transferable multipole and polarization parameter, and this will require further analysis.

References

- Burkert, U.; Allinger, N. L. *Molecular Mechanics*; ACS Monograph 177; American Chemical Society: Washington, DC, 1982.
- Jensen, F. *Introduction to Computational Chemistry*; Wiley: Chichester, UK, 1999.
- (a) Koch, U.; Popelier, P. L. A.; Stone, A. J. *J Chem Phys Lett* 1995, 238, 253; (b) Koch, U.; Stone, A. J. *J Chem Soc Faraday Trans* 1996, 92, 1701; (c) Halgren, T. A.; Damm, W. *Curr Opin Struct Biol* 2001, 11, 236.
- Rasmussen, T. D.; Jensen, F. *Mol Simul* 2004, 30, 801.
- Macromodel 8.1, Schrodinger: Portland, OR, 2003.
- Tinker 4.1, Ponder, J. W. Department of Biochemistry and Molecular Biophysics Washington University School of Medicine, St. Louis, MO 63110, 2003.
- Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K. J.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B. Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M. P.; Gill, M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian: Pittsburgh, PA, 1998.
- Jaguar 4.2; Schrodinger: Portland, OR, 2000.
- Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L. *GAMESS*; Iowa State University, 2002; with Dupuis, M.; Montgomery, J. A., Jr. *J Comput Chem* 1993 14, 1347.
- Rick, S. W.; Stuart, S. J. *Rev Comp Chem* 2002, 18, 89.
- (a) Ponder, J. W.; Case, D. A. *Adv Protein Chem* 2003, 66, 27; (b) Ren, P.; Ponder, J. W. *J Phys Chem* 2002, 107, 5933.
- Ren, P.; Ponder, J. W. *J Comput Chem* 2002, 23, 1497.
- (a) Stone, A. J. *J Chem Phys Lett* 1981, 83, 233; (b) Stone, A. J.; Alderton, M. *Mol Phys* 2002, 100, 221.
- (a) Van Duijnen, P. T.; Swart, M. *J Phys Chem A* 1998, 102, 2399; (b) Stone, A. *The Theory of Intermolecular Forces*; Clarendon: Oxford, UK, 1996.
- Stone, A. J. *J. Chem. Theory Comput.* 2005, 1, 1128.