A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules

Wendy D. Cornell,[†] Piotr Cieplak,[‡] Christopher I. Bayly,[§] Ian R. Gould,[⊥] Kenneth M. Merz, Jr.,[∥] David M. Ferguson,[&] David C. Spellmeyer,[#] Thomas Fox, James W. Caldwell, and Peter A. Kollman*

Contribution from the Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143

Received November 10, 1994[⊗]

Abstract: We present the derivation of a new molecular mechanical force field for simulating the structures, conformational energies, and interaction energies of proteins, nucleic acids, and many related organic molecules in condensed phases. This effective two-body force field is the successor to the Weiner *et al.* force field and was developed with some of the same philosophies, such as the use of a simple diagonal potential function and electrostatic potential fit atom centered charges. The need for a 10−12 function for representing hydrogen bonds is no longer necessary due to the improved performance of the new charge model and new van der Waals parameters. These new charges are determined using a 6-31G* basis set and restrained electrostatic potential (RESP) fitting and have been shown to reproduce interaction energies, free energies of solvation, and conformational energies of simple small molecules to a good degree of accuracy. Furthermore, the new RESP charges exhibit less variability as a function of the molecular conformation used in the charge determination. The new van der Waals parameters have been derived from liquid simulations and include hydrogen parameters which take into account the effects of any geminal electronegative atoms. The bonded parameters developed by Weiner *et al.* were modified as necessary to reproduce experimental vibrational frequencies and structures. Most of the simple dihedral parameters have been retained from Weiner *et al.*, but a complex set of ϕ and ψ parameters which do a good job of reproducing the energies of the low-energy conformations of glycyl and alanyl dipeptides has been developed for the peptide backbone.

Downloaded by UNIVERSITAT WIEN BIBLIOTHEKS on July 15, 2009
Published on May 1, 2002 on http://pubs.acs.org | doi: 10.1021/ja00124a002

Introduction

The application of computer-based models using analytical potential energy functions within the framework of classical mechanics has proven to be an increasingly powerful tool for studying molecules of biochemical and organic chemical interest. These applications of molecular mechanics have employed energy minimization, molecular dynamics, and Monte Carlo methods to move on the analytical potential energy surfaces. Such methods have been used to study a wide variety of phenomena, including intrinsic strain of organic molecules, structure and dynamics of simple and complex liquids, thermodynamics of ligand binding to proteins, and conformational transitions in nucleic acids. In principle, they are capable of giving insight into the entire spectrum of non-covalent interactions between molecules, and, when combined with quantum mechanical electronic structure calculations, modeling covalent bonding changes, essentially all molecular reactions and interactions. Given their importance, much effort has gone into consideration of both the functional form and the parameters that must be established in order to apply such analytical potential energy functions (or "force fields").

In the area of organic molecules, the book by Allinger and Burkert[1] provides a thorough review pre-1982 and the subsequent further development of the MM2[2] and MM3[3] force fields by Allinger and co-workers has dominated the landscape in this area. The number of force fields developed for application to biologically interesting molecules is considerably greater, probably because of the greater complexity of the interactions which involve ionic and polar groups in aqueous solution and the difficulty of finding an unequivocal test set to evaluate such force fields. Many of these force fields developed prior to 1987 are described briefly by McCammon and Harvey.[4]

Given the complexities and subjective decisions inherent in such biological force fields, we have attempted to put the development of the force field parameters on a more *explicitly stated* algorithmic basis than done previously, so that the force field could be extended by ourselves and others to molecules and functional groups not considered in the initial development. This is important, because, if the assumptions, approximations, and inevitable imperfections in a force field are at least known, one can strive for some cancellation of errors.

Approximately a decade ago, Weiner *et al.*[5,6] developed a force field for proteins and nucleic acids which has been widely

[†] Graduate Group in Biophysics, University of California, San Francisco.
[‡] Permanent address: Department of Chemistry, University of Warsaw, Pasteura 1, 02-093, Warsaw, Poland.
[§] Current address: Merck Frosst Canada, Inc., C.P. 1005 Pointe Claire-Dorval, Quebec H9R 4P8, Canada.
[⊥] Current address: Department of Chemistry, University of Manchester, Lancs M13 9PL, U.K.
[∥] Current address: Department of Chemistry, The Pennsylvania State University, State College, PA 16802.
[&] Current address: Department of Medicinal Chemistry, University of Minnesota, Minneapolis, MN 55455.
[#] Current address: Chiron Corporation, Emeryville, CA 94608.

(1) Burkert, U.; Allinger, N. J. *Molecular Mechanics;* American Chemical Society: Washington, DC, 1982.
(2) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127−8134 and subsequent versions, e.g. MM2−87, MM2−89, MM2−91.
(3) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551−8566, 8566−8576, 8576−8582.
(4) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, 1987.
(5) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765−

DOCKET ALARM
Find authenticated court documents without watermarks at docketalarm.com.

used. Important independent tests of this force field were performed by Pavitt and Hall for peptides[7] and Nilsson and Karplus[8] for nucleic acids and it was found to be quite effective. Nonetheless, it was developed in the era before one could routinely study complex molecules in explicit solvent. Weiner *et al.* attempted to deal with this issue by showing that the same force field parameters could be effectively used both without explicit solvent (using a distance-dependent dielectric constant ($\epsilon = R_{ij}$)) and with explicit solvent ($\epsilon = 1$) on model systems. Further support for this approach was provided by molecular dynamics simulations of proteins[9-11] and DNA[12,13] which compared the implicit and explicit solvent representations.

As computer power has grown, it has become possible to carry out more realistic simulations which employ explicit solvent representations. It is therefore appropriate that any new force field for biomolecules focus on systems modeled in the presence of an explicit solvent representation. This approach has been pioneered by Jorgensen and co-workers in their OPLS (Optimized Potentials for Liquid Simulations) model.[14] In particular, the development of parameters which reproduce the enthalpy and density of neat organic liquids as an essential element ensures the appropriate condensed phase behavior. The OPLS non-bonded parameters have been combined with the Weiner *et al.* bond, angle, and dihedral parameters to create the OPLS/Amber force field for peptides and proteins,[15] which has also been effectively used in many systems.[16]

We have been influenced by the OPLS philosophy of balanced solvent−solvent and solute−solvent interactions in our thoughts about a second-generation force field to follow that of Weiner *et al.*[5,6] The Weiner *et al.* force field used quantum mechanical calculations to derive electrostatic potential (ESP) fit atomic centered charges, whereas the OPLS charges were derived empirically, using mainly the liquid properties as a guide. For computational expediency, Weiner *et al.* relied principally on the STO-3G basis set for their charge derivation. This basis set leads to dipole moments that are approximately equal to or smaller than the gas-phase moment but tends to underestimate quadrupole moments. Thus, it is not well *balanced* with the commonly used water models (SPC/E,[17] TIP3P,[18] TIP4P[18]) which have dipole moments that are about 20% higher than the gas-phase value for water. These water models, which have empirically derived charges, include condensed-phase electronic polarization implicitly. Kuyper *et al.*[19] suggested that the logical choice of a basis set for deriving ESP-fit partial charges for use in condensed phases is the 6-31G* basis set, which uniformly overestimates molecular polarity. Standard ESP charges derived with that basis set were shown to lead to excellent relative free energies of solvation for benzene, anisole, and trimethoxyanisole.[19]

A 6-31G* based ESP-fit charge model, like the OPLS model, is capable of giving an excellent reproduction of condensed-phase *inter* molecular properties such as liquid enthalpies and densities and free energies of solvation.[20] A major difference between such a model and most others is the magnitude of the charges on hydrocarbons. For example, 6-31G* standard ESP charges derived from the trans conformation of butane have values of −0.344 for the methyl carbon and 0.078 for the methyl hydrogen. In both cases, however, the carbon and hydrogen charges offset each other, resulting in small net charges on the methyl groups of −0.110 and −0.059 for the trans and gauche charges, respectively. Furthermore, free energy perturbation calculations involving the perturbation of methane with standard ESP charges ($q_C = -0.464$ and $q_H = 0.116$) to methane with charges of 0.0 in solution yield essentially no change in free energy.[21] The standard ESP charges also result in conformational energies for butane which are in reasonable agreement with experiment, when used with a 1−4 electrostatic scale factor of 1/1.2.[20]

Nevertheless, the 6-31G* standard ESP charges are less than ideal for two reasons. First, when charges generated using different conformations of a molecule are compared, there is often considerable variation seen. This was demonstrated by Williams, who studied the conformational variation of ESP-fit charges in alanyl dipeptide for 12 different conformations.[22] Butane is another example, where charges from the gauche conformation have values of −0.197 and 0.046 for the methyl carbon and hydrogen, respectively. Another example is propylamine, which was studied at length by Cornell *et al.*[20] Five low-energy conformations can be identified for propylamine, and the 6-31G* standard ESP charges calculated for each conformation show significant variation. The average and standard deviation for the charge on a given atom over the five conformations are as follows: α-carbon $q_{av} = 0.339$ and $\sigma = 0.059$, β-carbon $q_{av} = 0.033$ and $\sigma = 0.060$, and γ-carbon $q_{av} = -0.205$ and $\sigma = 0.146$. This inconsistency is potentially problematic in terms of deriving other force field parameters which may be sensitive to the variation. Furthermore, it reduces the reproducibility of a particular calculation, which is not a problem in other force fields where the charges are assigned empirically.

The second reason that the 6-31G* standard ESP charges are less than ideal is that the charges on "buried" atoms (such as the sp^3 carbons described above for butane and propylamine) are statistically underdetermined and often assume unexpectedly large values for nonpolar atoms. Bayly *et al.*[23] found that the electrostatic potential of methanol could be fit almost equally well using either the standard ESP charges determined by the linear least-squares fit or an alternative set of charges derived with the methyl carbon constrained to have a much smaller value.

Considering the problems associated with the standard ESP charge model, it might seem tempting to adopt the OPLS approach of empirically derived charges. However, any empirically derived charge model cannot easily describe transition states and excited states, as can an electrostatic potential fit

Downloaded by UNIVERSITAT WIEN BIBLIOTHEKS on July 15, 2009
Published on May 1, 2002 on http://pubs.acs.org | doi: 10.1021/ja00124a002

(7) Pavitt, N.; Hall, D. *J. Comput. Chem.* **1984**, *5*, 441−450
(8) Nilsson, L.; Karplus, M. *J. Comput. Chem.* **1986**, *7*, 591−616.
(9) Tilton, R. F.; Singh, U. C.; Weiner, S. J.; Connolly, M. L.; Kuntz, I. D., Jr.; Kollman, P. A.; Max, N.; Case, D. *J. Mol. Biol.* **1986**, *192*, 443−456.
(10) Guenot, J. M.; Kollman, P. A. *Protein Sci.* **1992**, *1*, 1185−1205.
(11) York, D. M.; Wlodawer, A.; Redersen, L.; Darden, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8715−8718.
(12) Singh, U. C.; Weiner, S. J.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 755−759.
(13) Seibel, G. L.; Singh, U. C.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 6537−6540.
(14) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008−2010.
(15) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(16) (a) Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1990**, *112*, 2773−2781. (b) Orozco, M.; Tirado-Rives, J.; Jorgensen, W. L. *Biochemistry* **1993**, *32*, 12864−12874.
(17) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269−6271.
(20) Cornell, W.; Cieplak, P.; Bayly, C.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.
(21) (a) Sun, Y. X.; Spellmeyer, D.; Pearlman, D. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1992**, *114*, 6798−6801. (b) Sun, Y. X.; Kollman, P. A. Hydrophobic Solvation of Methane and Nonbond Parameters of the TIP3P Water Model. *J. Comput. Chem.*, in press. Pang, Y. P.; Kollman, P. A.

DOCKET ALARM
Find authenticated court documents without watermarks at docketalarm.com.

model. Furthermore, the conformational dependence of *N*-methylacetamide (NMA) is better represented with an ESP-fit model.[24] Finally, the requirement of Monte Carlo calculations on requisite liquids including appropriate fragments makes it more problematic to make an empirical charge model that will cover most or all of chemical/biochemical functionality.

Given the above-mentioned deficiencies in the standard ESP model, along with the desire to retain the general strategy of fitting charges to the electrostatic potential, Bayly *et al.*[23] were motivated to develop the RESP (restrained ESP-fit) charge model. The RESP model still involves a least-squares fit of the charges to the electrostatic potential, but with the addition of hyperbolic restraints on charges on non-hydrogen atoms. These restraints serve to reduce the charges on atoms which *can* be reduced without impacting the fit, such as buried carbons. The final RESP model requires a two-stage fit, with the second stage needed to fit methyl groups which require equivalent charges on hydrogen atoms which are not equivalent by molecular symmetry. The new charge model has been shown to perform well at reproducing interaction energies and free energies of solvation. When used with a 1−4 electrostatic scale factor of 1/1.2 (as opposed to the scale factor of 1/2 employed by Weiner *et al.*), both the RESP (and standard ESP) charges also result in good conformational energies for many of the small molecules studied to date without the necessity for an elaborate dihedral potential.[20]

In addition to the new charges which have been tailored for condensed phase simulations, new van der Waals (VDW) parameters have also been adopted and developed which are optimized for reproducing liquid properties. The VDW parameters in the Weiner *et al.*[5,6] force field are primarily a modification of a set originally proposed by Hagler−Euler−Lifson,[25] which were fit to lattice energies and crystal structures of amides. The new VDW parameters for aliphatic and aromatic hydrogens take into account the effects of any vicinal electronegative atoms.[26,27]

High-level quantum mechanical data are now available on the conformational energies of the glycyl and alanyl dipeptides[28] and these data are critical for developing ϕ and ψ dihedral parameters for the peptide backbone. Because such high-level data were unavailable at the time the Weiner *et al.* force field was developed, torsional parameters for the ϕ and ψ angles were left as 0.0 kcal/mol since the resulting molecular mechanical energies seemed to be in reasonable agreement with the best theoretical data available at that time. That force field led to conformational energies for glycyl dipeptide where the C5 extended conformation was about 1 kcal/mol too high in energy and for alanyl dipeptide where the C5 conformation was nearly 2 kcal/mol too high in energy but the $C7_{ax}$ conformation was about 1 kcal/mol too low in energy. The error in the alanyl dipeptide $C7_{ax}$ energy is not critical since it is rarely found in proteins[29] (only in γ-turns), but the errors in the energies of the C5 conformations are more important since that is the conformation found in β-sheets. Any errors in the energies of the C5 conformations are multiplied by the length of the secondary structure. The new force field includes V_1, V_2, V_3, and V_4 dihedral parameters for ϕ and ψ which result in good agreement between the molecular mechanical and quantum mechanical energies of the dipeptides.

Finally, the benzene molecule as modeled by the Weiner *et al.* all-atom force field has been shown to possess excessive flexibility for out-of-plane distortions.[30] This was caused by the use of the V_2 potential derived for the united atom model. This underestimate of the benzene V_2 parameter is noteworthy, because it affects not only the flexibility of benzene and benzene-like moieties but also the interpolation scheme used for determining the V_2 barriers for X−C−N−X and X−C−C−X dihedrals in conjugated rings. These V_2 parameters are determined by interpolating according to the bond length either between a pure single bond and a partial double bond (benzene) or between a partial double bond and a pure double bond. The excessive out-of-plane motion of benzene has been easily fixed by adjusting the V_2 parameter from 5.5 to 14.5 kcal/mol to match the experimental normal mode frequencies.

General Description of the Model

The model presented here (eq 1) can be described as "minimalist" in its functional form, with the bond and angles represented by a simple diagonal harmonic expression, the VDW interaction represented by a 6−12 potential, electrostatic interactions modeled by a Coulombic interaction of atom-centered point charges, and dihedral energies represented (in most cases) with a simple set of parameters, often only specified by the two central atoms. Electrostatic and van der Waals interactions are only calculated between atoms in different molecules or for atoms in the same molecule separated by at least three bonds. Those non-bonded interactions separated by exactly three bonds ("1−4 interactions") are reduced by the application of a scale factor.

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

Our assumption is that such a simple representation of bond and angle energies is adequate for modeling most unstrained systems. The goal of this force field is to accurately model conformational energies and intermolecular interactions involving proteins, nucleic acids, and other molecules with related functional groups which are of interest in organic and biological chemistry.

A. Atom Types. The atom types employed are similar to those defined previously and are given in Table 1. The one significant departure is the definition of new atom types for hydrogens bonded to carbons which are themselves bonded to one or more electronegative atoms. This is similar in spirit to the electronegativity based bond length correction used in MM2 and MM3.

B. Bond and Angle Parameters. The r_{eq}, θ_{eq}, K_r, and K_θ values[5,6] were used as starting values and adjusted as necessary to reproduce experimental normal mode frequencies. These values were initially derived by fitting to structural and vibrational frequency data on small molecular fragments that make up proteins and nucleic acids. For example, in complex fragments such as the nucleic acid bases, the r_{eq} and θ_{eq} values

(24) Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **1991**, *12*, 1232−1236.
(25) Hagler, A.; Euler, E.; Lifson, S. *J. Am. Chem. Soc.* **1974**, *96*, 5319−5327.
(26) Gough, C.; DeBolt, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 963−970.
(27) Veenstra, D.; Ferguson, D.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 971−978.
(28) (a) Gould, I. R.; Kollman, P. A. *J. Phys. Chem.* **1992**, *96*, 9255−

Downloaded by UNIVERSITAT WIEN BIBLIOTHEKS on July 15, 2009
Published on May 1, 2002 on http://pubs.acs.org | doi: 10.1021/ja00124a002

DOCKET ALARM Find authenticated court documents without watermarks at docketalarm.com.

Table 1. List of Atom Types[a]

atom	type	description
carbon	CT	any sp^3 carbon
	C	any carbonyl sp^2 carbon
	CA	any aromatic sp^2 carbon and (Cϵ of Arg)
	CM	any sp^2 carbon, double bonded
	CC	sp^2 aromatic in 5-membered ring with one substituent + next to nitrogen (Cγ in His)
	CV	sp^2 aromatic in 5-membered ring next to carbon and lone pair nitrogen (e.g. Cδ in His (δ))
	CW	sp^2 aromatic in 5-membered ring next to carbon and NH (e.g. Cδ in His (ϵ) and in Trp)
	CR	sp^2 aromatic in 5-membered ring next to two nitrogens (Cγ and Cϵ in His)
	CB	sp^2 aromatic at junction of 5- and 6-membered rings (Cδ in Trp) and both junction atoms in Ade and Gua
	C*	sp^2 aromatic in 5-membered ring next to two carbons (e.g. Cγ in Trp)
	CN	sp^2 junction between 5- and 6-membered rings and bonded to CH and NH (Cϵ in Trp)
	CK	sp^2 carbon in 5-membered aromatic between N and N-R (C8 in purines)
	CQ	sp^2 carbon in 6-membered ring between lone pair nitrogens (e.g. C2 in purines)
nitrogen	N	sp^2 nitrogen in amides
	NA	sp^2 nitrogen in aromatic rings with hydrogen attached (e.g. protonated His, Gua, Trp)
	NB	sp^2 nitrogen in 5-membered ring with lone pair (e.g. N7 in purines)
	NC	sp^2 nitrogen in 6-membered ring with lone pair (e.g. N3 in purines)
	N*	sp^2 nitrogen in 5-membered ring with carbon substituent (in purine nucleosides)
	N2	sp^2 nitrogen of aromatic amines and guanidinium ions
	N3	sp^3 nitrogen
oxygen	OW	sp^3 oxygen in TIP3P water
	OH	sp^3 oxygen in alcohols, tyrosine, and protonated carboxylic acids
	OS	sp^3 oxygen in ethers
	O	sp^2 oxygen in amides
	O2	sp^2 oxygen in anionic acids
sulfur	S	sulfur in methionine and cysteine
	SH	sulfur in cysteine
phosphorus	P	phosphorus in phosphates
hydrogen	H	H attached to N
	HW	H in TIP3P water
	HO	H in alcohols and acids
	HS	H attached to sulfur
	HA	H attached to aromatic carbon
	HC	H attached to aliphatic carbon with no electron-withdrawing substituents
	H1	H attached to aliphatic carbon with one electron-withdrawing substituent
	H2	H attached to aliphatic carbon with two electron-withdrawing substituents
	H3	H attached to aliphatic carbon with three electron-withdrawing substituents
	HP	H attached to carbon directly bonded to formally positive atoms (e.g. C next to NH_3^+ of lysine)
	H4	H attached to aromatic carbon with one electronegative neighbor (e.g. hydrogen on C5 of Trp, C6 of Thy)
	H5	H attached to aromatic carbon with two electronegative neighbors (e.g. H8 of Ade and Gua and H2 of Ade)

[a] See refs 5 and 6.

determined by linear interpolation between pure single and double bond values using the observed bond distances and the K_r value taken from vibrational analysis of a simple sp^2 atom and suggested by the critical analysis of Halgren of the diagonal force constants used in different force fields.[31]

One "difficulty" arose in the development of this new force field compared to that of Weiner *et al.* which was related to the switch to the 6-31G* basis set for charge derivation. With 6-31G* standard ESP charges and a 1−4 electrostatic scale factor of 1/1.2 rather than 1/2.0 (see below), we found that the exocyclic $-NH_2$ groups of the bases moved considerably away from their r_{eq} and θ_{eq} values upon energy minimization. This problem was considerably reduced with RESP charges and a 1−4 electrostatic scale factor of 1/1.2, so we chose not to selectively increase the K_θ values around the $-NH_2$ group to force it to more "canonical" geometries.

In general, however, one might have resorted to a more complex optimization of r_{eq}, θ_{eq}, K_r, and K_θ to ensure that the geometries of simple fragments were as close as possible to experiment *after energy minimization*, rather than taking r_{eq} and θ_{eq} from experiment and assuming little distortion would occur (which is generally the case, with the slight exception of the case of the $-NH_2$ groups noted above). We chose not to undertake a more time-consuming iterative self-consistent derivation of geometrical parameters, because of our assumption that any such errors which we were making were of much smaller consequence for accurately representing conformations and intermolecular interactions than the inaccuracies remaining in the dihedral and non-bonded (charge and VDW) parameters.

C. Dihedral Parameters. Weiner *et al.*[5,6] developed a limited set of general and specific dihedral parameters which were appropriate for the functionalities found in proteins and DNA and calibrated to adjust the energies of small model compounds. In this strategy, a dihedral parameter is optimized on the simplest molecule possible and then applied to larger and more complex molecules. This approach is in contrast to one employed by many other force field developers where the parameters are optimized to best reproduce the conformational energies of a large number of molecules. An advantage of our approach is the lack of dependence of the resulting parameters on the particular molecules chosen for the test set.

For the most part, a minimalist approach has been retained with regards to dihedral parameters. For example, we have only a 3-fold Fourier component (V_3) for dihedrals around −C−C− bonds, with the exception of cases such as E−C−C−E′ where E and E′ are electronegative atoms like O or F. In these cases, there is a "gauche" effect which stabilizes the gauche conformation over the trans and this can be modeled with a 2-fold Fourier component (V_2). The rotation around phosphorus−ester bonds (CT−OS−P−OS) also requires a 2-fold component. In these cases, we have been able to go beyond the Weiner *et al.* force field by making use of reasonably high level *ab initio* models (MP2/6-31G*) to fit the values of such V_2 Fourier components.

Two exceptions were made to the principle of adding extra Fourier terms to the dihedral energies only in the presence of a compelling physical basis. These exceptions are the dipeptide ψ and ϕ and the nucleoside χ dihedrals. Here we used additional Fourier components to try to reproduce as well as possible the relative energies of the alanyl and glycyl dipeptides and a model nucleoside fragment calculated at a high level of theory without the requirement of "a physical picture". An alternative approach would be to empirically adjust the atomic partial charges to achieve the same aim. Given the power of the RESP methodology for deriving atomic partial charges which lead to good representations of intermolecular interactions and the importance of maintaining an accurate balance between intra- and inter-

Downloaded by UNIVERSITAT WIEN BIBLIOTHEKS on July 15, 2009
Published on May 1, 2002 on http://pubs.acs.org | doi: 10.1021/ja00124a002

DOCKET ALARM
Find authenticated court documents without watermarks at docketalarm.com.

Table 2. Standardized Parameters for Scaling Algorithms

bond	r_{eq}[a]	K_r[b]
pure C−C	1.507[c]	317[d]
pure C=C	1.336[e]	570[f]
pure C−N	1.449[g]	337[h]
pure C=N	1.273[i]	570[j]
torsion	r_{eq}[a]	V_2[k]
pure X−C−C−X	1.507[c]	0.0[l]
partial X−C=C−X	1.397[m]	14.5[n]
pure X−C=C−X	1.336[e]	30.0[o]
pure X−C−N−X	1.449[g]	0.0[p]
partial X−C=N−X	1.335[q]	10.0[r]
pure X−C=N−X	1.273[i]	30.0[s]

[a] In Å. [b] In kcal/(mol Å^2). [c] Microwave data from acetone (ref 32). [d] Value taken from MM2, ref 2. [e] Microwave data from propene (ref 32). [f] Default from NMA normal mode analysis for carbonyl force constant. [g] Benedetti structural data (ref 33). [h] Value derived from normal mode analysis on NMA. [i] Microwave data from methylenimine (ref 32). [j] Default value, see footnote f. [k] In kcal/mol. [l] Assumed free rotation about pure C−C single bond. [m] Structural data from benzene (ref 32). [n] From normal modes analysis of benzene. [o] Approximate rotational barrier of ethylene is ~60 kcal/mol (see ref 34). [p] Assumed free rotation about a pure single C−N bond. [q] Benedetti structural data (ref 33). [r] Reference 35. [s] Calculated rotational barrier in methylenimine is 57.5 kcal/mol (see ref 36).

In our previous force field, the bond length and V_2 parameters for X−C−N−X and X−C−C−X fragments involving sp^2 hybridized atoms were determined by a linear interpolation approach (according to the experimental bond length) between the known barriers of pure single, pure double, and partial double bonded systems (benzene for X−C−C−X and NMA for X−C−N−X). We have used the same approach here, but have adjusted the V_2 term of benzene to more accurately describe its out-of-plane frequencies (Weiner *et al.*[5,6] had used the V_2 derived for a united atom model of benzene, which was significantly different). Table 2 presents the parameters used. For example, given a C(sp^2)−C(sp^2) bond length, its bond stretching force constant is linearly interpolated between the values for pure single bond and double bond given in Table 2. Its V_2 torsional potential is interpolated between the values for pure double and partial double or between partial double and single, depending on whether the bond length is greater or less than the 1.397 Å of benzene. This is exactly the procedure used by Weiner *et al.*[5,6]

D. VDW Parameters. Given the success of the OPLS approach in modeling liquids, we have developed all-atom sp^3 carbon and aliphatic hydrogen VDW parameters by carrying out Monte Carlo simulations on CH_4, C_2H_6, C_3H_8, and C_4H_{10} liquids and empirically adjusting R^* and ϵ for the C and H to reproduce the densities and enthalpies of vaporization of these liquids.[37] Such parameters have also been employed in calculations of relative free energies of solvation of CH_4, C_2H_6, and C_3H_8.[21,38] We also derived VDW parameters for sp^2 C and aromatic H employing Monte Carlo simulations on benzene liquid and adjusting the R^* and ϵ of these atoms to reproduce the density and enthalpy of liquid benzene.[37] At the time these parameters were developed, such all-atom parameters were unavailable for the OPLS force field. These Monte Carlo simulations were the first calculations carried out as part of the development of this new force field, and as such employed 6-31G* standard ESP charges. The electrostatic contribution for the *n*-alkanes was very small regardless of the charge model—at most a few tenths of a kcal/mol. We note that the standard ESP charges for benzene (q_C= −0.145 and q_H= 0.145) accurately reproduce the quadrupole moment of that molecule.

We have taken most of the remaining VDW parameters from the OPLS model[15]—sp^2 and sp^3 N; sp^2 O, ether ester (OS), hydroxyl (OH) and TIP3P water (OW) sp^3 oxygens; and sulfur (SH and S)—since it has been optimized for reproducing liquid properties. The Weiner *et al.*[5,6] phosphorus (P) parameters were not re-optimized since that atom is most frequently found buried inside of four other heavy atoms.

The VDW model is minimalist as well, with some exceptions. A standard VDW parameter is used for a given atom and hybridization, e.g. all sp^2 carbons have the same VDW parameters. The only heavy atom exceptions are sp^3 O, where oxygens in water (OW), alcohol (OH), and ether (OS) have slightly different parameters, as found in OPLS. We suspect that this is due to the use of a zero VDW radius on hydrogens bound to oxygen, so that an effectively larger R^* is required for a water oxygen than alcohol than ether.

A significant departure has been made from the previous model in the treatment of hydrogens. The current model does not employ 10−12 hydrogen bonding H···X parameters, although these are still supported within the AMBER software. The original Hagler *et al.*[25] and OPLS approach[14,15] suggested a zero R^* and ϵ for hydrogen binding hydrogens. Thus the TIP3P water model has R^* and ϵ equal to 0.0 for its hydrogen (HW). We opted not to develop a new water model, but to use the TIP3P one.

Hydrogen and helium are unique in the periodic table in not having an inner shell of electrons. Consequently, it makes physical sense for the hydrogen VDW radius, unlike other atoms, to be very sensitive to its bonding environment. This has been extensively analyzed for the hydrogen R^* in X−C−*H* systems by Gough *et al.* and Veenstra *et al.*,[26,27] who demonstrated the sensitivity of R^* to the electron-withdrawing properties of X. For example, a "normal" C−*H* has VDW R^* = 1.487 Å; whereas in CF_3−*H* it is ~0.3 Å shorter and in $CH_3NH_3^+$ it is ~0.4 Å shorter still.

We have employed the following approach here. A C−*H* has R^* = 1.487 Å and, based on nucleic and base pairing energy minimization, an N−*H* has R^* = 0.6 Å. This qualitative dependence on electronegativity makes physical sense. Based on the Veenstra *et al.*[27] studies we have chosen to reduce the R^* on sp^3 C−*H* atoms by 0.1 Å for each electronegative (O, N, F, S) substituent. The hydrogen atom types are then defined as H1, H2, and H3 for 1, 2, and 3 electronegative groups, respectively. The hydrogen R^* is reduced by 0.4 Å for each neighboring positively charged group (atom type HP). For sp^2 C−*H*, R^* has been reduced by 0.05 Å for each electronegative neighbor (atom types H4 and H5).

Given our retention of the simplicity of a 6−12 rather than a 6-exponential VDW representation, we have continued to reduce 1−4 VDW interactions since the 6−12 approximation and the lack of polarization in the model both will lead to exaggerated short-range repulsion. It is difficult to determine the scale factor unambiguously so we have retained the value of 1/2.0 used by Weiner *et al.*[5,6]

E. Electrostatic Energies. In Cornell *et al.*[20] and Cieplak *et al.*[39] we have extensively analyzed the development of our

(32) Harmony, M.; Laurie, V.; Kuezkowski, R.; Schwendeman, R.; Ramsay, D.; Lovas, F.; Lafferty, W.; Maki, A. *J. Phys. Chem. Ref. Data* **1979**, *8*, 619−721.
(33) Benedetti, E. In *Peptides-Proceedings of the 5th American Peptide Symposium*; Goodman, M., Meienhofer, J., Eds.; J. Wiley and Co.: New York, 1977; pp 257−273.
(34) Douglas, J.; Rabinovich, B. S.; Looney, F. *J. Chem. Phys.* **1955**, *23*, 315−323.
(35) Momany, F.; McGuire, R.; Burgess, A.; Scheraga, H. *J. Phys. Chem.*

Downloaded by UNIVERSITAT WIEN BIBLIOTHEKS on July 15, 2009
Published on May 1, 2002 on http://pubs.acs.org | doi: 10.1021/ja00124a002

DOCKET ALARM
Find authenticated court documents without watermarks at docketalarm.com.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.