

See discussions, stats, and author profiles for this publication at:
<https://www.researchgate.net/publication/229901510>

Derivation of Class II force fields. 5. Quantum force field for amides, peptides, and related compounds

ARTICLE in JOURNAL OF COMPUTATIONAL CHEMISTRY · MARCH 1998

Impact Factor: 3.59 · DOI: 10.1002/(SICI)1096-987X(199803)19:4<430::AID-JCC5>3.0.CO;2-T · Source: DBLP

CITATIONS

78

READS

88

5 AUTHORS, INCLUDING:



[T. P. Stockfisch](#)

Stockfisch Consulting

17 PUBLICATIONS 809 CITATIONS

SEE PROFILE



[Arnie Hagler](#)

University of Massachusetts Amherst

134 PUBLICATIONS 8,122 CITATIONS

SEE PROFILE

Derivation of Class II Force Fields: V. Quantum Force Field for Amides, Peptides, and Related Compounds

J. R. MAPLE, M.-J. HWANG,* K. J. JALKANEN,[†] T. P. STOCKFISCH, A. T. HAGLER[‡]

Molecular Simulations Inc., 9685 Scranton Road, San Diego, California 92121

Received 23 December 1996; accepted 6 October 1997

ABSTRACT: As the field of biomolecular structure advances, there is an ever-growing need for accurate modeling of molecular energy surfaces to simulate and predict the properties of these important systems. To address this need, a second generation amide force field for use in simulations of small organics as well as proteins and peptides has been derived. The critical question of what accuracy can be expected from calculations in general, and with this class II force field in particular, is addressed for structural, dynamic, and energetic properties. The force field is derived from a recent methodology we have developed that involves the systematic use of quantum mechanical observables. Systematic *ab initio* calculations were carried out for numerous configurations of 17 amide and related compounds. Relative energies and first and second derivatives of the energy of 638 structures of these compounds resulted in 140,970 *ab initio* quantum mechanical observables. The class II peptide quantum mechanical force field (QMFF), containing 732 force constants and reference values, was parameterized against these observables. A major objective of this work is to help establish the role of anharmonicity and coupling in improving the accuracy of molecular force fields, as these terms have not yet become an agreed upon standard in the ever more extensive simulations being used to probe biomolecular properties. This has been addressed by deriving a class I harmonic diagonal force field (HDFF), which was fit to the same energy surface as the QMFF, thus providing an opportunity to quantify the effects of

* *Present address:* Institute of Biomedical Sciences, Academia Sinica, Taipei 11529, Taiwan, ROC

[†] *Present address:* DKFZ, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

[‡] *Present address:* ScienceMedia Inc., 6540 Lusk Boulevard, Suite C144, San Diego, CA 92121

Correspondence to: A. T. Hagler (at Molecular Simulations Inc.)

Contract/grant sponsor: Potential Energy Functions Consortium

This article contains Supplementary Material available from the authors upon request or via the Internet at <ftp.wiley.com/public/journals/jcc/suppmat/19/430> or <http://journals.wiley.com/jcc/>

these coupling and anharmonic contributions. Both force field representations are assessed in terms of their ability to fit the observables. They have also been tested by calculating the properties of 11 stationary states of these amide molecules. Optimized structures, vibrational frequencies, and conformational energies obtained from the quantum calculations and from both the QMFF and the HDFF are compared. Several strained and derivatized compounds including urea, formylformamide, and butyrolactam are included in these tests to assess the range of applicability (transferability) of the force fields. It was found that the class II coupled anharmonic force field reproduced the structures, energies, and vibrational frequencies significantly more faithfully than the class I harmonic diagonal force field. An important measure, rms energy deviation, was found to be 1.06 kcal/mol with the class II force field, and 2.30 kcal/mol with the harmonic diagonal force field. These deviations represent the error in relative configurational energy differences for strained and distorted structures calculated with the force fields compared with quantum mechanics. This provides a measure of the accuracy that might be expected in applications where strain may be important such as calculating the energy of a system as it approaches a (rotational) barrier, in ligand binding to a protein, or effects of introducing substituents into a molecule that may induce strain. Similar results were found for structural properties. Protein dynamics is becoming of ever-increasing interest, and, to simulate dynamic properties accurately, the dynamic behavior of model compounds needs to be well accounted for. To this end, the ability of the class I and class II force fields to reproduce the vibrational frequencies obtained from the quantum energy surface was assessed. An rms deviation of 43 cm^{-1} was achieved with the coupled anharmonic force field, as compared to 105 cm^{-1} with the harmonic diagonal force field. Thus, the analysis presented here of the class II force field for the amide functional group demonstrates that the incorporation of anharmonicity and coupling terms in the force field significantly improves the accuracy and transferability with regard to the simulation of structural, energetic, and dynamic properties of amides. © 1998 John Wiley & Sons, Inc. J Comput Chem 19: 430–458, 1998

Keywords: force fields; amides; conformations; *ab initio*; quantum chemistry; molecular mechanics

Introduction

The rapid pace of experimental research in complex molecular systems, particularly in biochemistry and biophysics, has led to a greatly increased need for accurate molecular simulation techniques. The most powerful techniques at present are molecular mechanics and dynamics, which presuppose an analytic energy function of the atomic coordinates (i.e., a force field) that governs all the relevant molecular energetic, structural, and dynamic properties. Although in many applications a qualitative picture derived from the simulation is sufficient, in an ever-increasing number of cases quantitative accuracy is critical for obtaining meaningful results.

Although a number of factors affect the accuracy of simulations, including the treatment of

solvent and counterions,¹ perhaps the most basic factor is the force field itself. Formulation of the force field depends on the functional form and the associated parameters that relate molecular energies to inter- and intramolecular distances, bond lengths, bond angles, and torsion angles. Research into this fundamental problem is the topic of the present series of studies.^{2–5}

Closely connected to the form assumed for the force field is the type, amount, and quality of the data used in its derivation. Early force fields relied primarily on analysis of experimental data, particularly gas phase vibrational spectra or x-ray crystallographic data. Examples of such experimentally derived force fields are AMBER,⁶ CVFF,⁷ CHARMM,⁸ GROMOS,⁹ MM3,¹⁰ MM4,¹¹ and OPLS.¹² However, in many cases, the amount of available experimental data is insufficient for obtaining well-determined and transferable force field parameters for even the simplest energy func-

tions for many, if not most, classes of molecules. Furthermore, inadequate availability of experimental data has inhibited research on improved energy functions for more accurate simulations of energetic, structural, and spectroscopic properties. This has led to the more recent development of force fields such as the class II CFF^{2-4,13-15} and MMFF,¹⁶ which are based primarily on quantum mechanical calculations of the energy surface.

The method for deriving class II CFF force fields has been previously described.^{2-4,13-15} It is based on using both existing experimental data and the vast amount of information that can be obtained from *ab initio* quantum mechanical calculations.^{2-4,13-15} With this method a quantum mechanical force field (QMFF) giving the molecular energy as a function of atomic coordinates is first derived from least-squares fits of *ab initio* data consisting of the energy and the first and second derivatives of the energy with respect to the Cartesian coordinates of the nuclei in a set of molecules. To sample the *ab initio* potential energy surface, the training set incorporates distorted molecular structures that are generated by displacing the nuclei in the directions of the normal modes of vibrations.² Systematic errors inherent in the quantum data are then removed by scaling the resulting QMFF to fit experimental data that includes vibrational frequencies plus gas phase and crystal structural data.³ Applying this methodology to the alkane functional group, 78 adjustable force constants and reference values in the QMFF force field were derived from a total of 128,376 independent quantum observables.² The resulting CFF alkane force field was then derived by adjusting only seven scale factor and reference value parameters to reproduce 150 experimental observables consisting of gas phase vibrational frequencies and equilibrium bond lengths.³ The resulting force field was then shown to account for observed conformational energies and rotational barriers as well. Large ratios of observables to parameters were maintained in this manner by fitting both the quantum and experimental data, thus enabling the derivation of a well-determined, transferable force field for a relatively complex energy function with many anharmonic and interaction terms. One may view the methodology as providing an amplification of the experimental data through the use of the quantum energy surface.

In addition to providing a very large and almost inexhaustible reservoir of observables, the use of quantum data in force field development is invaluable because the detailed types of molecular data

that can be provided are currently not available or even accessible from experiment. For example, the gas phase conformational energies of simple dipeptides or the rotational barriers in most small amide molecules are not known. Another example can be found in conformationally dependent structural trends,¹⁷ which are also difficult if not impossible to obtain from experimental measurements. This information, which can be readily determined from *ab initio* calculations, is highly useful for the derivation and testing of force fields. The resulting force fields, in turn, are needed to accurately simulate important biological processes such as protein dynamics or ligand binding, or to assist in the development of new therapeutic agents.

The functional form of the force field is a crucial factor in determining how accurately the force field will be able to reproduce quantum or experimental information on conformational energies, the energetic barriers to transitions between conformers, conformationally dependent structural trends, and vibrational frequencies. The quantum mechanical data allow parameterization of more accurate functional forms, including anharmonicity and coupling between internal coordinates. Recently, in studies of the alkane functional group, it has been demonstrated that the anharmonic and coupling interaction terms in the MM3, MMFF, and CFF functional forms substantially improve the accuracy of force field calculations of many of the aforementioned properties. The designation "class II force field," which is defined more rigorously elsewhere,³ was used to distinguish the functional form of such force fields as MM3 and CFF, which contain these anharmonic and coupling interactions, relative to those of older class I force fields, which are for the most part diagonal quadratic force fields; that is, based only on quadratic functions of the bond lengths and bond angles plus simple cosine forms for torsion potentials. (We note, however, that for the amide functional group, MM3 omitted the use of coupling interactions from the functional form, although anharmonic terms for bonds and angles were retained.¹⁰ MMFF includes only bond length-bond angle coupling.¹⁶)

In the present work, the fifth in this series²⁻⁵ a quantum force field for amide molecules is derived with the CFF functional form. The ultimate objective is to use the resulting QMFF together with experimental data as the basis for the derivation of the protein force field as just described. In this study we first derive the quantum force field and then test the adequacy of the functional form for the amide functional group by comparing force

field and quantum calculations of the energies, structures, and frequencies of model compounds. The results of the examination of the energy surface of amides complement previous tests of the CFF functional form on quantum^{2,15} and experimental³ data for alkanes. In particular, because amides contain partially conjugated bonds and undergo pyramidal inversion processes, the results of these calculations provide new information on the importance of anharmonicity and coupling interactions to describe the energetics of deformation of these structural features. In this article, the derivation of an amide quantum force field and the fit of the force field to the quantum observables (energies and derivatives) is presented. The force field is then tested by assessing the accuracy with which it reproduces *ab initio* conformational energies, structures, and frequencies of stable conformers of 11 simple amides and related compounds in the training set. (Analysis of stationary states of the three-membered rings and peptides in the training set is deferred to a subsequent report.) As part of the assessment of the anharmonic and coupling interactions in the energy surface, parallel results are determined for a quadratic diagonal force field, in which the former terms are not included. A more stringent test of the force field for amides will be presented in a subsequent article, which will include tests on rotational barriers, nitrogen pyramidal inversion barriers, and the effects of rotation about conjugated bonds and of pyramidal inversion on structures and vibrational frequencies.¹⁸

Potential Energy Surface Sampling

As previously described,^{2,14,15} a three-step procedure is used to sample the potential energy surfaces of a family of molecules. First, a training set of molecular structures must be selected for parameterization. Then, different distorted configurations of the model compounds comprising the training set are generated by methods such as displacing the atoms along the directions of linear combinations of the normal modes of vibration. Finally, the energy, forces, and second derivatives of the distorted molecules are calculated at the HF/6-31G* level of theory. Because the resulting force field will be scaled to fit experimental data, here we only require the determination of the overall shape of the potential energy surface. HF/6-31G* has pragmatically been shown to be

adequate for this purpose, and, because literally thousands of configurations of fairly large molecules (in the context of *ab initio* calculations) are being considered, it has been the method of choice.

Molecules chosen for the training set are depicted in Figure 1, organized in four groups (Fig. 1A–D). Simple amides are represented by formamide, acetamide, *N*-methylformamide, *N*-methylacetamide, *N,N*-dimethylformamide, and *N,N*-dimethylacetamide. These compounds sample all the needed internal coordinates (bond lengths, angles, torsions, and out-of-plane motions) except for torsions involving ethyl or propyl groups. The latter were sampled in *N*-ethylformamide and propionamide to complete group A. The three dipeptides in group B, *N*-formylglycineamide, *N*-formylalanineamide, and *N*-formyl-*N*-methylglycineamide, were also selected for the training set. Urea, *N*-formylformamide, and butyrolactam (group C) were included to extend both the range of molecular topology and of deformations to which the force field may be applied. The three-membered ring molecules in group D, *N*-formylaziridine, azacyclopropanone, and methylazacyclopropanone, were selected for similar reasons.

Several methods were used to generate varied configurations of the molecules in the training set to sample the peptide energy surface. To ensure that all deformations were sampled, most of the distorted configurations were obtained by displacing the atoms along the directions of the normal modes of vibration, while randomly selecting the magnitude of the atomic displacements, as discussed in detail elsewhere.^{1,2} In addition, a number of configurations were obtained from *ab initio* structure optimizations. Optimized structures for stable conformers, as well as other stationary states (e.g., transition states), were generated to sample conformational energies, rotational barriers, and barriers to amino inversion. These optimized structures are also useful for testing the force field and are described later in this work. In a few cases (e.g., for *N*-formylglycineamide, *N*-formylalanineamide, and *N*-formyl-*N*-methylglycineamide), additional points on the *ab initio* potential energy surface were sampled by extracting the coordinates and forces generated during the (iterative) *ab initio* optimizations (these, in general, sample lower distortion energies). Finally, to improve the sampling of the torsion angle energy dependence, additional configurations were obtained from rigid rotations about the amide bond in optimized structures of formamide, acetamide, *N*-methylfor-

(A) Amides

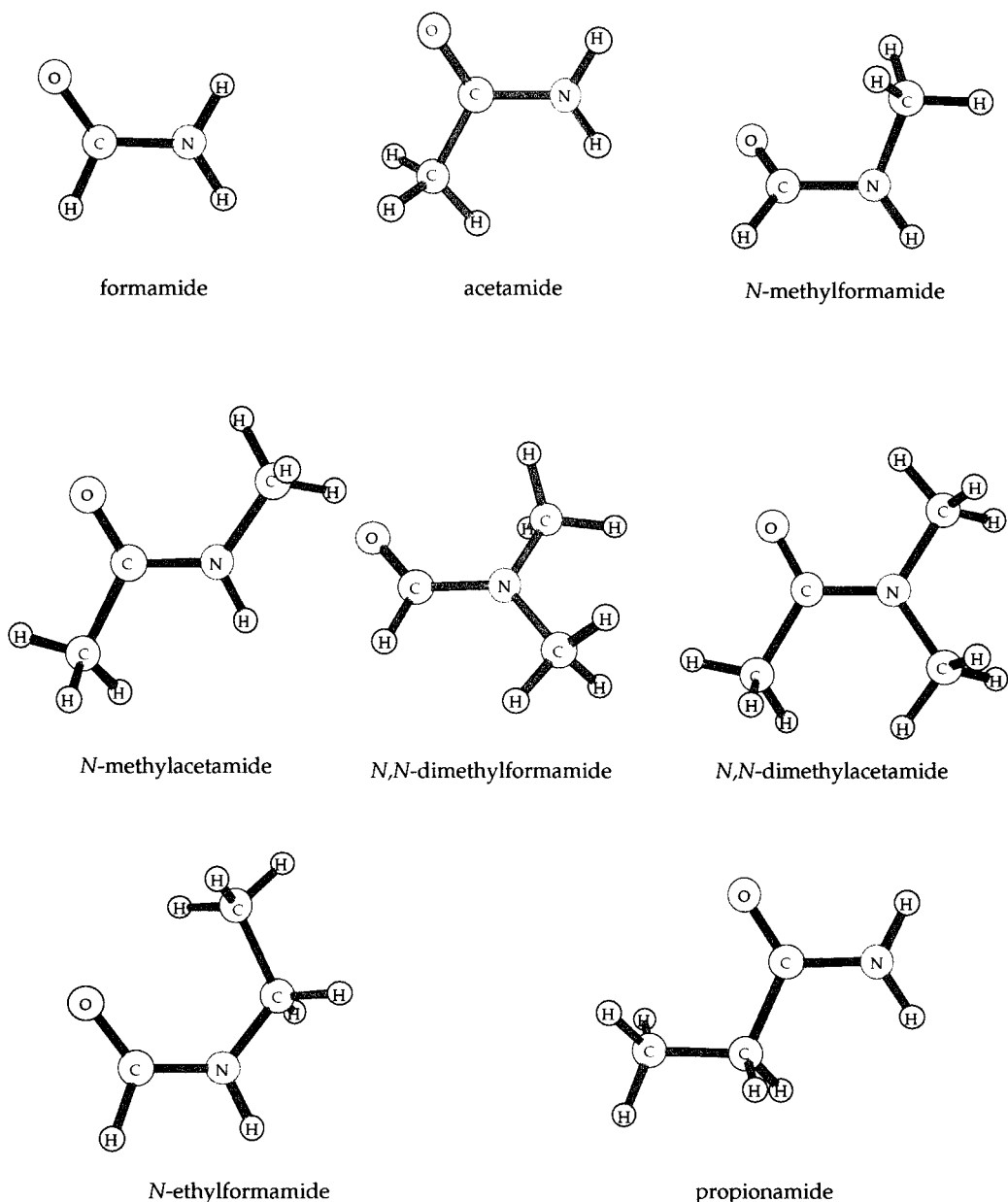


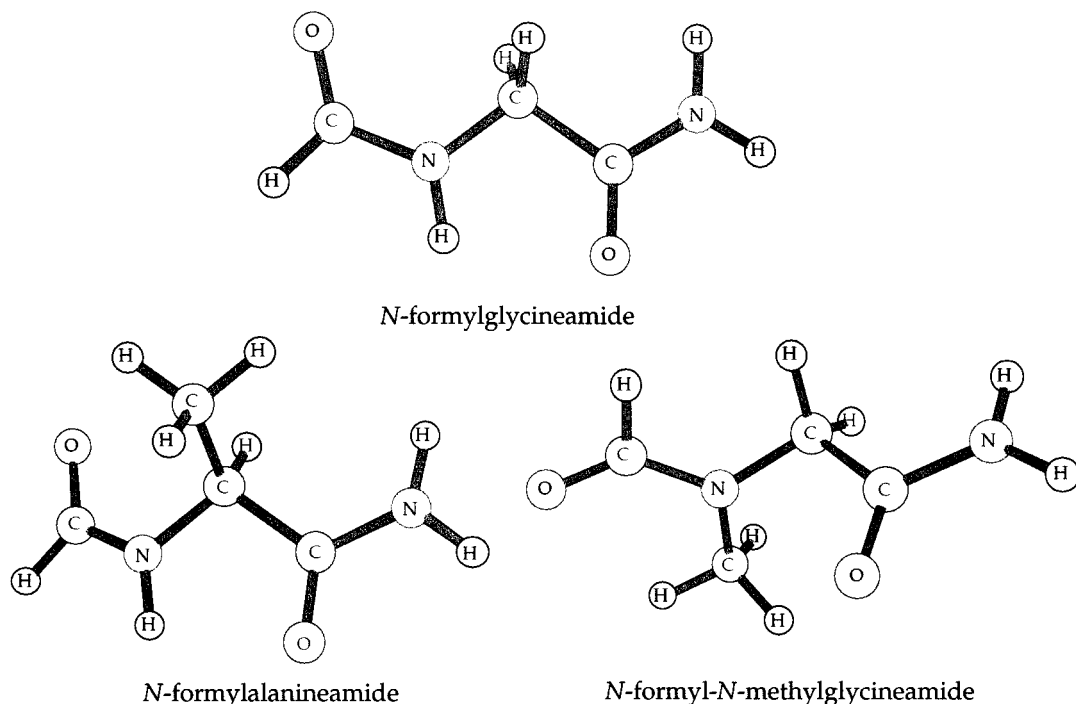
FIGURE 1. Model compounds used in the derivation of the quantum mechanical force field (QMFF) for the peptide group. A diverse set of molecules is chosen to constitute the training set. These molecules include (A) amides; (B) peptides; (C) lactams, and other amide-related compounds; and (D) three-membered rings.

mamide, and *N*-methylacetamide. Use of these differing methods to sample the energy surface results in a wide range of deformation energies to be accounted for, as seen in Table I.

The extent of the potential energy surface sampling is characterized in Figure 2, which illustrates the range of exploration of the energy surface for three typical internal coordinates: $C'-N$ bonds,

$O'-C'-N$ angles, and $O'-C'-N-C$ torsion angles. (We used O' and C' to denote the carbonyl oxygen and carbon atoms.) For example, a total of 869 $C'-N$ bond lengths ranging from 1.22–1.50 Å were sampled. Of these, it can be seen from Figure 2a that the training set included a maximum of 320 bonds in the 1.35–1.37-Å range with the remainder spread over a range of ~ 0.28 Å. Similarly, Figure

(B) Dipeptides



(C) Related Compounds

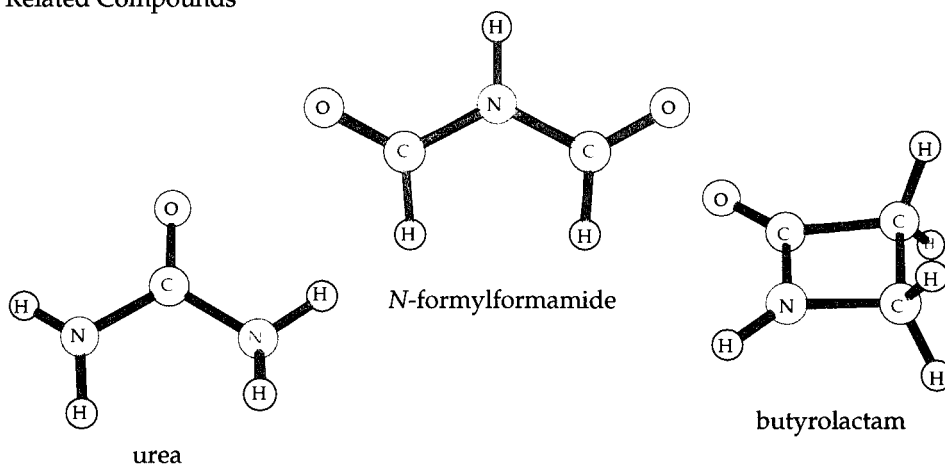


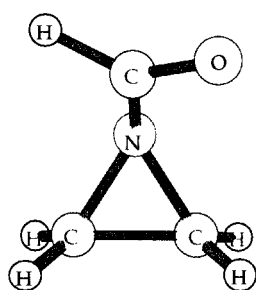
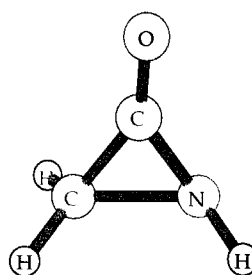
FIGURE 1. (Continued)

2b and c illustrate that the $O'-C'-N$ angle and the $O'-C'-N-C$ torsion angle were sampled through ranges of $90-150^\circ$ and $0-360^\circ$, respectively. These internal coordinate sampling ranges are typical of the ranges used for the other internal coordinates of the training set molecules in Figure 1 and of the previously reported ranges utilized for the parameterization of the hydrocarbon functional group.² These obviously span high and low energy regions of the potential surface and test the

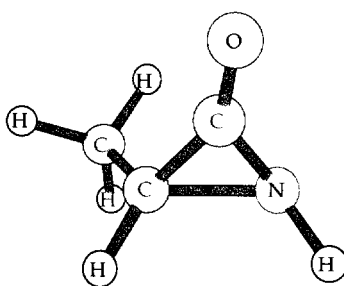
capability of the functional form to account for the distortions likely to be encountered in practical applications.

The potential energy surface sampling was completed by calculating the *ab initio* energy and the first and second derivatives of the energy with respect to the atomic coordinates. All calculations were performed at the HF/6-31G* level of theory with the GRADSCF,¹⁹ TURBOMOLE,²⁰ or GAUSS-
IAN 90²¹ programs. The maximum energy of the

(D) Three-membered rings

*N*-formylaziridine

azacyclopropanone



methylazacyclopropanone

FIGURE 1. (Continued)

distorted structures of each molecule in the training set is given in Table I, along with the number of independent quantum mechanical observables generated for the training set. In general, for M configurations of a molecule containing N atoms, there will be $M - 1$ relative energies, $M(3N - 6)$ independent first derivatives, and $M(0.5)(3N - 6)(3N - 5)$ independent second derivatives of the energy.² For example, the number of independent quantum observables for 36 distorted formamide structures comprises 35 relative energies, 432 first derivatives, and 2808 second derivatives. The final training set contained a total of 140,970 quantum observables consisting of 621 relative energies, 18,039 first derivatives, and 122,310 second derivatives. The latter two quantities represent the slope and curvature of the potential surface at the particular configuration.

Parameterization

The quantum force field (QMFF) was derived by fitting the parameters in the CFF functional form^{2,4,15} to these quantum mechanical observables, using the least-squares fitting procedure previously discussed.² The CFF functional form is given by:

$$E = \sum_b \left[{}^2K_b(b - b_0)^2 + {}^3K_b(b - b_0)^3 + {}^4K_b(b - b_0)^4 \right] + \sum_\theta \left[{}^2K_\theta(\theta - \theta_0)^2 + {}^3K_\theta(\theta - \theta_0)^3 + {}^4K_\theta(\theta - \theta_0)^4 \right]$$

TABLE I.

Summary of Energy Data Used for Derivation of Quantum Force Field for Amides, Including the Energy (E_{\max}) of Maximally Distorted Structure and the Number of Quantum Mechanical Observables for each Molecule.

Molecule	E_{\max}^a	Number of observables		
		Energies	First derivatives	Second derivatives
Formamide	26.7	35	432	2808
Acetamide	68.6	44	945	10,395
<i>N</i> -methylformamide	63.4	78	1659	18,249
<i>N</i> -methylacetamide	115.0	66	2010	31,155
<i>N,N</i> -dimethylformamide	55.1	18	570	4650
<i>N,N</i> -dimethylacetamide	17.8	6	273	5460
<i>N</i> -ethylformamide	54.1	34	1050	10,695
Propionamide	48.8	86	3393	7800
<i>N</i> -formylglycineamide	128.9	102	3090	12,555
<i>N</i> -formylalanineamide	18.4	47	2016	0
<i>N</i> -formyl- <i>N</i> -methylglycineamide	31.1	19	840	0
Urea	89.4	23	432	4104
<i>N</i> -formylformamide	96.3	28	522	4959
Butyrolactam	51.9	7	192	2400
<i>N</i> -formylaziridine	50.1	11	288	3600
Azacyclopropanone	27.2	8	135	1080
Methylazacyclopropanone	36.7	9	192	2400
Total		621	18,039	122,310

^a The maximum energy among all distorted molecular structures of each species relative to its most stable equilibrium structure (kcal/mol).

$$\begin{aligned}
& + \sum_{\phi} \left[{}^1K_{\phi}(1 - \cos \phi) + {}^2K_{\phi}(1 - \cos 2\phi) \right. \\
& \left. + {}^3K_{\phi}(1 - \cos 3\phi) \right] + \sum_{\chi} K_{\chi} \chi^2 \\
& + \sum_b \sum_{b'} K_{bb'}(b - b_0)(b' - b'_0) \\
& + \sum_{\theta} \sum_{\theta'} K_{\theta\theta'}(\theta - \theta_0)(\theta' - \theta'_0) \\
& + \sum_b \sum_{\theta} K_{b\theta}(b - b_0)(\theta - \theta_0) \\
& + \sum_{\phi} \sum_b (b - b_0) \\
& \times \left[{}^1K_{\phi b} \cos \phi + {}^2K_{\phi b} \cos 2\phi + {}^3K_{\phi b} \cos 3\phi \right] \\
& + \sum_{\phi} \sum_{b'} (b' - b'_0) \\
& \times \left[{}^1K_{\phi b'} \cos \phi + {}^2K_{\phi b'} \cos 2\phi + {}^3K_{\phi b'} \cos 3\phi \right] \\
& + \sum_{\phi} \sum_{\theta} (\theta - \theta_0) \\
& \times \left[{}^1K_{\phi\theta} \cos \phi + {}^2K_{\phi\theta} \cos 2\phi + {}^3K_{\phi\theta} \cos 3\phi \right] \\
& + \sum_{\phi} \sum_{\theta} \sum_{\theta'} K_{\phi\theta\theta'}(\theta - \theta_0)(\theta' - \theta'_0) \cos \phi \\
& + \sum_{i>j} \varepsilon \left[2 \left(\frac{r^*}{r_{ij}} \right)^9 - 3 \left(\frac{r^*}{r_{ij}} \right)^6 \right] \\
& + \sum_{i>j} \frac{q_i q_j}{r_{ij}} \quad (1)
\end{aligned}$$

where b , θ , ϕ , χ , and r denote bond lengths, bond angles, torsion angles, out-of-plane coordinates, and distances between nonbonded atoms, respectively. The parameters K , q , ε , and r^* represent force constants, atomic charges, van der Waals well depths, and atomic diameters, respectively. The first four summations describe the quartic bond stretching, quartic angle bending, three-term Fourier expansion of the torsion energy, and the harmonic out-of-plane deformation energy, respectively. The next set of terms represent coupling interactions between internal coordinates. In order, these couplings are described by the following cross-terms: bond/bond, angle/angle, bond/angle, bond/torsion 1 (describing the coupling between a torsion angle and the central bond in the

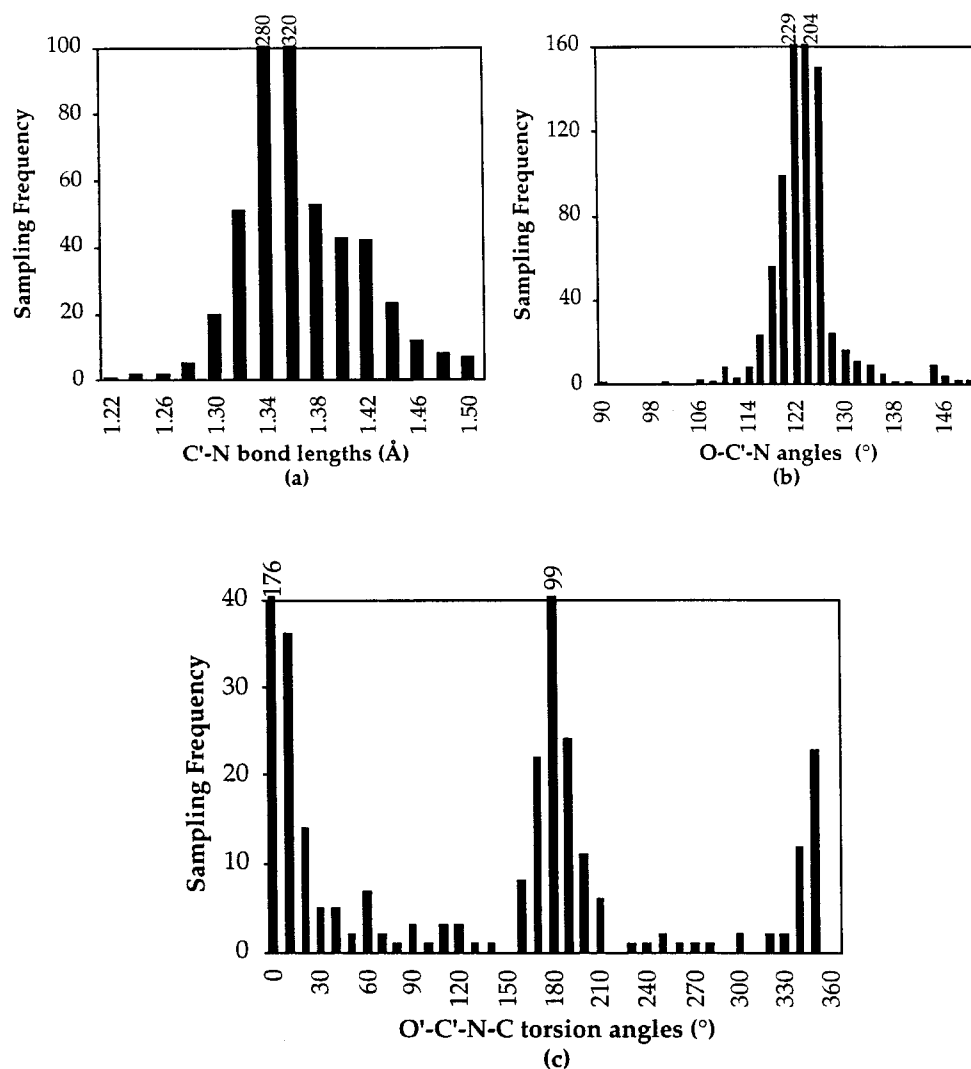


FIGURE 2. Distribution of (a) C'—N bond lengths, (b) O'—C'—N bond angles, and (c) O'—C'—N—C torsion angles in structures used to sample the *ab initio* potential energy surface.

torsion), bond/torsion 2 (describing the coupling between a torsion angle and a peripheral bond in the torsion), angle/torsion, and angle/angle/torsion. The final two terms describe the van der Waals interaction, which is represented by an inverse ninth power exchange repulsion plus an inverse sixth power dispersion interaction, and the Coulombic electrostatic interaction. This is the same functional form as previously reported,^{2,4,15} except for the addition of the harmonic out-of-plane term.

OUT-OF-PLANE COORDINATE DEFINITION

The out-of-plane coordinates, χ , which describe the planarity or, conversely, the degree of pyramidalization at nitrogen and carbonyl carbon atoms

in amides, are defined as the angle between a plane formed by three (of the four) atoms constituting the out-of-plane center and the bond formed between the fourth atom and the out-of-plane center²² (e.g., the C—N bond and the plane formed by the carbon, oxygen, and hydrogen atoms in formamide). To obtain a unique definition for the out-of-plane coordinate, the average value of the three possible bond/plane angles was used for the out-of-plane coordinate in the CFF functional form. This out-of-plane coordinate definition, which we have previously denoted as a “symmetrized Wilson angle,”²² was found to give a much better description of the potential energy surface than the “improper torsion” coordinate previously used in most common force fields.

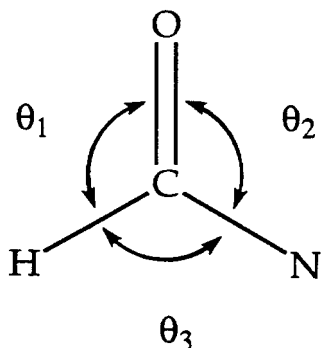


FIGURE 3. Schematic diagram of the amide group.

TRANSFERABILITY OF PARAMETERS

The amide nitrogen (N), nonpolar hydrogen (H), carbonyl carbon (C'), and carbonyl oxygen (O') atoms, as depicted in Figure 3, were assumed to be chemically identical in all of the molecules in the training set. In other words, parameters for the carbonyl (C'—O') bond were assumed to be transferable between each of the amide, peptide, and related (e.g., urea, *N*-formylformamide, and butyrolactam) molecules, and similarly, transferability was assumed for each of the other bond, angle, torsion, out-of-plane, nonbond, and coupling interactions. Alkyl carbon atoms (C) with sp^3 hybridization were distinguished from the sp^2 -hybridized carbonyl carbon, and the parameters for methyl, methylene, and methine groups were transferred from the QMFF for hydrocarbons.² The justification for these assumptions is based on the quality of the fit to the *ab initio* potential energy surface and the ability of the QMFF to reproduce derived quantities such as energies, structures, and frequencies, as described in later sections of this work.

The electrostatic and van der Waals parameters for the potential energy function in eq. (1) were derived from fits of experimental crystal structures and sublimation energies of a variety of molecular crystals containing carbonyl groups (e.g., aldehydes, ketones, carboxylic acids, esters, and amides) and amino groups (e.g., amides and amines).²³ As was done for the derivation of the alkane QMFF,² a pragmatic approach was used wherein nonbond parameters were fixed at these values in the determination of the force constants and reference values by the least-squares fit of the quantum mechanical energy surfaces. This assured maximum transferability of the nonbond amide parameters to

intramolecular and intermolecular atomic interactions in chemically similar functional groups.

CONSTRAINTS AND RESTRAINTS

The derivation of molecular mechanics force fields is complicated by the need to use a redundant set of internal coordinates. For a molecule with N atoms, $3N - 6$ internal coordinates are needed to uniquely describe the structure. However, the sum of the total number of bond lengths, bond angles, torsion angles, and out-of-plane coordinates required for the force field is almost always greater than $3N - 6$, thereby introducing an excess number of force constants that result in correlations between parameters for the angle, torsion, and out-of-plane energy functions [i.e., terms 2–4 in eq. (1)] and for all cross-term interactions involving one of these three angular deformation coordinates.

Parameter redundancies can severely impede the determination of force fields, because, in principle, there may be an infinite number of solutions (i.e., parameter sets) to the least-squares fit of the observables that result in equivalent mathematical descriptions of the energy surfaces. However, in practice, most of these solutions consist of parameters without a plausible physical basis (e.g., negative quadratic force constants for energy function terms for bond angles or out-of-plane deformations) and are consequently not useful for molecular modeling.

As a specific example of a common parameter correlation problem in molecules with sp^2 -hybridized atoms, consider the planar amide group in formamide (see Fig. 3). The in-plane angle deformations are uniquely described by only two bond angles (θ_1 and θ_2), because the value of the third bond angle (θ_3) is equal to 360° minus the sum of the other two angles. By assuming the harmonic approximation, which is valid for small deformations, the energy changes caused (solely) by in-plane angle deformations can be characterized by:

$$E = k_1(\theta_1 - \theta_1^0)^2 + k_2(\theta_2 - \theta_2^0)^2 + k_{12}(\theta_1 - \theta_1^0)(\theta_2 - \theta_2^0) \quad (2)$$

Thus, three force constants (k_1 , k_2 , and k_{12}) describe the energetics of in-plane deformations. However, in force fields used for molecular me-

chanics and dynamics, three bond angle coordinates are normally used to describe the in-plane deformations located at atoms with sp^2 hybridization. If the energy equation used by the force field includes all possible quadratic deformation terms, as in eq. (3):

$$E = k_1(\theta_1 - \theta_1^0)^2 + k_2(\theta_2 - \theta_2^0)^2 + k_3(\theta_3 - \theta_3^0)^2 \\ \times k_{12}(\theta_1 - \theta_1^0)(\theta_2 - \theta_2^0) \\ + k_{13}(\theta_1 - \theta_1^0)(\theta_3 - \theta_3^0) \\ + k_{23}(\theta_2 - \theta_2^0)(\theta_3 - \theta_3^0) \quad (3)$$

then it can be seen that force constants for three diagonal terms and three cross-terms, or a total of six force constants, are used, even though only three are needed to completely describe a harmonic potential energy surface. As a result the force constants in eq. (3) are redundant, and there are, in principle, an infinite number of equivalent parameter sets that would mathematically describe the energetics of bond angle deformations in the amide group used for this example.

The method chosen for treating the redundancy problem for angular deformations in formyl, acetyl, and amino groups was to constrain the force constants for angle/angle interactions to a value of 0, thereby reducing eq. (3) to:

$$E = k_1(\theta_1 - \theta_1^0)^2 + k_2(\theta_2 - \theta_2^0)^2 + k_3(\theta_3 - \theta_3^0)^2 \quad (4)$$

In other words, angle/angle cross-terms were not used for angles centered at atoms with three bonds.

Although this procedure removed the parameter redundancies just described, many other sources of redundancies exist. The majority of these are caused by the use of dihedral angles for torsion coordinates. For example, in formamide, a well-defined coordinate set should contain a single torsion coordinate, whereas four dihedral angles are customarily used in molecular mechanics force fields for this molecule. The problem is compounded by the use of torsion-angle-dependent cross-terms, such as the bond/torsion interactions used by CFF^{2,15} and MM3.²⁴ Using the formamide example again, the force constants for the coupling interaction between the amide bond and the four dihedral angles are redundant, because only a single set of Fourier coefficients is needed to describe this interaction with a well-defined coordinate set.

Parameter redundancy problems can be reduced (but not eliminated) by using force fields

without cross-terms. As discussed earlier and noted elsewhere,^{2,25} angle/angle cross-terms are not needed to describe the angle bending potential for angles located at trigonal centers (i.e., apex atoms with three bonds). However, there is ample evidence that many other interaction cross-terms (e.g., bond/bond, bond/angle, and bond/torsion coupling interactions) are necessary for accurate predictions of the energetics, structures, and vibrational spectra of molecules.^{2-4,24-29} The cross-term interactions listed in eq. (1) are examples of specific types that are needed for accurate simulations.^{2,29}

PENALTY FUNCTIONS

One of the practical consequences of parameter redundancy is that there is a tendency for some force constants to be substantially larger than "reasonable," particularly in the case of force constants for terms dependent on torsion coordinates (i.e., the internal coordinate, if it were isolated from correlated internals, would have a nonphysical energy surface). To restrain these parameters to the "physical domain" restraints in the form of penalty functions were introduced for terms containing torsions, as in eq. (5):

$$F_p = \sum_i S_i^2 \left[1 + \sum_j W_j P_j^2 \right] \quad (5)$$

Here, S_i are the residuals (i.e., differences between energies and energy derivatives calculated by quantum mechanics and by the force field), P_j are parameters, and W_j are penalty coefficients. This modified function of the residuals, F_p , was then minimized in the least squares process. The penalty coefficients were chosen so as to damp the torsion constants, while having a minimal effect on the sum of squares. We note that this restraint can only reduce the ability of the parameters to fit the data, because an additional restraint cannot improve the fit. Thus, it will not lead to an artificially good fit.

PARAMETERS

The amide QMFF parameters obtained from the least-squares fit of the quantum mechanical energies and energy derivatives as just described are given in Table I of the Supplementary Material. Parameters specific to alkanes (e.g., C—H bond stretching and C—H/H—C—H bond/angle cross-term force constants) were derived previous-

ly,² but are given there for completeness. The 732 force constants and reference values for amides were derived from a total of 140,970 quantum observables (see Table I). Thus, an observable-to-parameter ratio of 193:1 was achieved in the fit of the data.

Quality of Fit to *Ab Initio* Potential Energy Surfaces

AMIDES IN TRAINING SET

The quality of fit of the quantum force field (QMFF) to the *ab initio* potential energy surfaces of the core set of eight amide molecules (group A in Fig. 1) in the training set is indicated in Table II, which gives the rms deviations from *ab initio* results for energies, first derivatives, and second derivatives calculated by the QMFF for the distorted structures of the molecules.

For comparison, Table II also lists the rms deviations from the quantum results for a harmonic diagonal force field (HDFF), which has a class I functional form. The harmonic diagonal functional form is widely used for simulations of large biological systems in force fields such as AMBER,⁶ CHARMM,⁸ and GROMOS.⁹ The HDFF was derived in a separate fit to the observables in the same manner as described for the QMFF [i.e., by

eliminating all anharmonic and coupling interactions in eq. (1) and then fitting the harmonic diagonal force constants to the same data]. However, the penalty functions were not applied in deriving the HDFF. The primary reason for using penalty functions is to make the parameters more physically reasonable and hence transferable. Because the HDFF parameters were not intended to be transferred to other systems, parameter transferability was not an issue. Thus, the HDFF was allowed to fit the data as well as possible given its functional form, and pragmatic considerations to restrict the parameter ranges to physically reasonable values were not required. As such, the HDFF was, in effect, given the benefit of the doubt, to make it compare as favorably as possible with the QMFF. Nevertheless, we will see that, with a few exceptions, it still performs significantly less well than the coupled anharmonic QMFF force field, although not restrained by penalty functions, as just noted.

Because the HDFF is more restricted in the range of molecular deformation energies it can accurately describe, the force field parameters in the HDFF derived using quantum mechanical data depend on the choice of deformations in the training set of molecules more strongly than those in the QMFF. In the present case, we employed both equilibrium and moderately distorted structures

TABLE II.
Ability of Diagonal Harmonic (Class I) and QMFF (Class II) Functional Forms to Account for Quantum Mechanical Energy Surface of Amides: rms Deviations from *Ab Initio* Energies and First and Second Derivatives of Distorted Structures.

Molecule	$\Delta E_{\text{rms}}^{\text{a}}$ (kcal / mol)		% $\Delta E'_{\text{rms}}^{\text{b}}$		% $\Delta E''_{\text{rms}}^{\text{c}}$	
	HDFF (class I)	QMFF (class II)	HDFF (class I)	QMFF (class II)	HDFF (class I)	QMFF (class II)
Formamide	1.58	0.70	38.2	10.0	21.5	4.4
Acetamide	2.35	1.31	39.3	15.5	18.3	4.1
<i>N</i> -methylformamide	1.36	0.71	43.0	15.8	17.3	4.1
<i>N</i> -methylacetamide	5.97	1.92	38.0	12.9	27.0	5.7
<i>N,N</i> -dimethylformamide	2.17	1.15	46.7	12.6	27.2	5.5
<i>N,N</i> -dimethylacetamide	0.59	1.42	^d	^d	12.6	4.6
<i>N</i> -ethylformamide	2.07	0.72	39.9	9.8	29.0	4.9
Propionamide	2.28	0.53	51.3	27.3	28.1	5.8
Average deviation	2.30	1.06	42.3	14.8	21.4	4.9

^a rms deviation of relative energies calculated with the force field from the *ab initio* values.

^b Percent rms deviation of first derivatives of the energy.

^c Percent rms deviation of second derivatives of the energy.

^d All configurations are near an energy minimum, wherein the forces are small and the relative deviations are therefore not meaningful.

typical of those found in macromolecules. For example it is known experimentally that bond angles about the C $^{\alpha}$ atoms in polypeptides can vary by up to 8 $^{\circ}$,^{30,31} with even larger ranges found in molecular dynamics simulations. These variations are reflected by the ranges of internal coordinates as shown by Table I and Figure 2 for both force fields. Alternatively, the HDFF could have been fit to a relatively small range of structures, in which case the values of equilibrium energies and other computed equilibrium properties would be better reproduced. However, when applied to real systems containing the range of angles (and other internals) alluded to above, or strained systems such as calculation of rotational barriers, molecules with substituents that introduce strain, ligands with strain induced on binding to proteins, dynamic fluctuations, etc., the errors would be large because the system was outside the "domain" for which the force field was derived. Thus, the present approach compares the two representations derived using exactly the same data, focusing on how well the force fields reproduce *trends* in computed properties among different molecules and conformations.

As can be seen in Table II, from the improvement achieved by the full force field over the harmonic diagonal force field, the quantum energy surface of amides contains significant coupling and anharmonicity. This is similar to the conclusions drawn for the formate energy surface¹⁴ and for the intramolecular interactions in hydrocarbons.² The deviations in the energies of the distorted amide configurations observed with the QMFF range from 0.53 to 1.92 kcal/mol, whereas the class I diagonal quadratic force field yields rms errors in energy that range from 0.59 to 5.97 kcal/mol. The average rms energy deviations of 1.06 and 2.30 kcal/mol for the QMFF and HDFF, respectively, demonstrate that the neglect of anharmonicity and coupling interactions results in energies that are less accurate by an average of 1.2 kcal/mol over this set of structures.

The errors in first derivatives (forces) obtained with the QMFF range from 9.8% for *N*-ethylformamide to 27.3% for propionamide, whereas the corresponding error range for the harmonic diagonal force field is from 38.2% for formamide to 51.3% for propionamide. The average deviations in first derivatives are 14.8% for the QMFF and 42.3% for the HDFF, which demonstrates the extreme importance of anharmonicity and coupling interactions in determining the atomic forces in distorted

structures. As is well known, accurate forces are necessary for obtaining accurate trajectories in molecular dynamics simulations.

The importance of anharmonicity and coupling interactions are also clearly indicated in the fit of the second derivatives (curvature of the potential surface). For the QMFF, the relative deviations range from 4.1% for acetamide and *N*-methylformamide to 5.8% for propionamide, in contrast to the much larger deviations in the harmonic diagonal force field, which range from 12.6% for *N,N*-dimethylacetamide to 29.0% for *N*-ethylformamide. The average relative deviations obtained with the QMFF and HDFF are 4.9% and 21.4%, respectively, thus indicating that the neglect of anharmonicity and coupling interactions results in an increase in errors for calculated second derivatives of the energy by a factor of four.

A comparison of the deviations in the energies and energy derivatives given in Table II with the corresponding deviations obtained for alkanes² indicates that the potential energy surfaces of amides are not fit quite as well. For example, the deviations in energies and second derivatives obtained with the QMFF functional form for alkanes are 0.87 kcal/mol and 4.6%, respectively. These deviations are somewhat smaller than the average amide deviations (in Table II) of 1.06 kcal/mol and 4.9%, respectively. The larger errors in the amide potential energy surfaces seem to indicate that there may be coupling interactions not accounted for by the QMFF functional form in eq. (1) that are significant for amides but not for alkanes. Obvious examples of such possible coupling interactions are out-of-plane/torsion or bond/out-of-plane couplings, because alkanes do not contain out-of-plane coordinates. (There are, of course, other interaction terms that are not accounted for and might be more important in amides, such as more accurate expressions for the electrostatic terms, whose absence can also lead to increased deviations.) In any case, it should be kept in mind that, although the current functional form provides a significantly better description of the energy surface than previous diagonal forms, there are still significant residual deficiencies indicating that there remain interactions that are inadequately accounted for.

BIFUNCTIONAL AMIDES

In Table III we compare the HDFF and the QMFF by examining their ability to fit derivatized amide groups (groups C and D). These molecules include highly strained and functionalized amides

TABLE III.

Ability of Diagonal Harmonic (Class I) and QMFF (Class II) Functional Form to Account for Quantum Mechanical Energy Surface of Amide-Related Molecules: rms Deviations from *Ab Initio* Energies and First and Second Derivatives of Distorted Structures.

Molecule	$\Delta E_{\text{rms}}^{\text{a}}$ (kcal / mol)		% $\Delta E'_{\text{rms}}^{\text{b}}$		% $\Delta E''_{\text{rms}}^{\text{c}}$	
	HDFP (class I)	QMFF (class II)	HDFP (class I)	QMFF (class II)	HDFP (class I)	QMFF (class II)
<i>N</i> -formylformamide	2.94	1.34	34.7	20.2	27.4	6.1
Urea	8.55	1.88	41.1	13.1	33.6	6.0
Butyrolactam	5.57	0.39	40.8	7.2	29.6	7.2
<i>N</i> -formylaziridine	5.40	4.88	48.2	20.9	25.3	9.3
Azacyclopropanone	4.39	1.58	47.1	20.0	33.6	8.5
Methylazacyclopropanone	3.74	1.67	57.2	23.1	23.7	9.5
Average deviation	5.10	1.96	44.8	17.4	28.9	7.8

^a rms deviation of relative energies calculated with the force field from those calculated *ab initio*.

^b Percent rms deviation of first derivatives of the energy.

^c Percent rms deviation of second derivatives of the energy.

and provide a stringent challenge for the force field. As with the corresponding results for the amides in Table II, the QMFF shows a substantially better agreement with the quantum mechanical energy surface. For example, the average rms deviations of the energies are 2.0 and 5.1 kcal/mol for the QMFF and the HDFP, respectively. Similarly, the improved accuracy that results from including anharmonicity and coupling interactions in the functional form is demonstrated by the 17.4% and 44.8% relative rms deviations from the quantum first derivatives that result, on average, from the QMFF and the HDFP, respectively. The average relative rms deviation for the second derivatives is also smaller by over a factor of three (i.e., 7.8% for the QMFF, as compared to 28.9% for the HDFP) when anharmonicity and coupling interactions are taken into account.

The average rms deviations in the energies and energy derivatives are significantly larger for the functionally substituted amides than for the amides themselves. However, the rms deviations in the energies and the first derivatives of most of these molecules fall within the ranges discussed for the amides. For example, with the QMFF, the 4.88 kcal/mol rms deviation in the energies calculated for the substituted three-membered ring *N*-formylaziridine are the only energy deviations that do not fall within the 0.53–1.92 kcal/mol range (see Table II) for the amides. Similarly, all relative rms deviations in the first derivatives calculated by the QMFF for the amide-related molecules are

within the 9.8–27.3% range found for the amides in the training set. On the other hand, the relative rms deviations in the second derivatives range from 6.0% to 9.5% for these molecules, which are larger than those (4.1–5.8%) for the amides themselves. On the whole, these results indicate that the assumption of parameter transferability among amides and their derivatives (listed in Table III) is not unreasonable, but does result in some degradation in the accuracy with which the force field can account for the quantum mechanical energy surface.

Comparison of Relative Energies, Internal Coordinates, and Frequencies from *Ab Initio* and Force Field Calculations of Specific Compounds

AMIDES

As a further test of the quantum force field, and to assess the importance of anharmonicity and coupling interactions, the relative conformational energies, structures, and vibrational frequencies obtained quantum mechanically were compared with those calculated by the derived force fields, HDFP and QMFF, for a variety of stable conformers of the amide molecules shown in Table IV. The most significant differences between the QMFF and *ab initio* results for the optimized structures, frequencies, and conformational energy differences are discussed next.

TABLE IV.
Comparison of Accuracy of Bond Lengths and Bond Angles in Planar Structures Optimized by Harmonic Diagonal Force Field (HDFF) and QMFF: rms Deviations of Eight *Ab Initio* Structures of Amides.

Molecule / structure	Bonds (Å)		Angles (°)	
	HDFF	QMFF	HDFF	QMFF
Formamide	0.013	0.001	1.9	1.3
Acetamide	0.007	0.004	0.9	0.9
<i>N</i> -methylformamide, trans	0.020	0.002	1.4	0.8
<i>N</i> -methylformamide, cis	0.015	0.002	1.5	0.5
<i>N</i> -methylacetamide, trans	0.015	0.002	1.0	0.8
<i>N</i> -methylacetamide, cis	0.011	0.002	1.4	0.7
<i>N,N</i> -dimethylformamide	0.021	0.002	1.4	0.7
<i>N,N</i> -dimethylacetamide	0.020	0.003	1.4	0.8
Average rms deviation	0.015	0.002	1.4	0.8

Because most of the amides studied here are either planar or very nearly planar, except for the methyl groups, the values and deviations of the computed torsion angles are not a very useful probe of the force fields and will not be discussed in detail. However, in some cases, the minimum energy structure obtained by the force fields disagreed with the *ab initio* structure with regard to the rotation of one or more methyl groups. Specifically, the HDFF structure for acetamide differs from the *ab initio* structure with regard to the methyl rotation. Similarly the QMFF structures differ from the *ab initio* structures with regard to methyl orientations for *trans-N*-methylacetamide, *N,N*-dimethylformamide, and *N,N*-dimethylacetamide. In these cases, the stationary state that most closely approximates the *ab initio* structure is a transition state (i.e., the methyl group is at a rotation barrier) with the force field. These stationary states were used in comparing vibrational frequencies, as they represent structures with the closest corresponding geometries. The structures with differing methyl torsions are noted as such in the Supplementary Material.

Detailed comparisons of results are given in Tables II and III of the Supplementary Material, which give all the internal coordinates and frequencies for each of the structures optimized by *ab initio* calculation and by the HDFF and QMFF force fields.

Formamide

As indicated in Table IV, the *ab initio* and QMFF structures are nearly identical. The rms deviations in the bond lengths and angles are only 0.001 Å

and 1.3°, respectively. The corresponding deviations (0.013 Å and 1.9°) resulting with the harmonic diagonal force field are significantly larger. Both the quantum and force field structures are planar, in agreement with microwave spectra^{32–34} and recent high-level quantum calculations.³⁵

As seen from Table V, the frequencies for all normal modes except the amino puckering mode agree quite closely, with an rms deviation of 28 cm⁻¹ using the QMFF. However, the deviation in the puckering mode is 294 cm⁻¹, the QMFF frequency being calculated as 403 cm⁻¹ as compared to the 109 cm⁻¹ obtained for the quantum result. In fact, as seen from Table V, both the HDFF and the QMFF consistently exhibit large deviations in frequency for the amino puckering mode relative to the frequency obtained from the *ab initio* calculations, not only for formamide but also for most of the other primary amides.

Amino Puckering Mode in Primary Amides

The very large error in the frequency of the amino puckering mode results from the fact that the force constant for the nitrogen out-of-plane coordinate has been constrained to a value of zero. This constraint was imposed to avoid the use of an out-of-plane force constant with a negative value. The root of the anomaly, wherein a physically unrealistic force constant value must be used to reproduce *ab initio* results for the frequency of the amino puckering mode, lies in the nonorthogonality of the out-of-plane and torsion (i.e., dihedral angle) coordinates used in the QMFF, as well as in all commonly used molecular mechanics/dynamics force fields. The torsion potential for rota-

TABLE V. Comparison of Accuracy of Conformational Energy Differences and Vibrational Frequencies Calculated by HUFF and QMFF for Eight Amide Structures.

Molecule / structure	Energies (kcal / mol) ^a			rms deviations in frequencies (cm ⁻¹) ^b	
	HF / 6-31G*	HUFF	QMFF	HUFF	QMFF
Formamide	0.0	0.0	0.0	108 (245)	28 (293)
Acetamide	0.0	0.0	0.0	79 (320)	44 (525)
N-methylformamide, trans	0.0	0.0	0.0	119 (95)	45 (100)
N-methylformamide, cis	1.1	2.1 (1.0)	1.6 (0.5)	121 (–117)	38 (–141)
N-methylacetamide, trans	0.0	0.0	0.0	100 (39)	50 (22)
N-methylacetamide, cis	2.5	3.2 (0.7)	2.4 (0.1)	101 (8)	39 (77)
N,N-dimethylformamide	0.0	0.0	0.0	112 (16)	48 (0)
N,N-dimethylacetamide	0.0	0.0	0.0	101 (48)	48 (10)
Average rms deviation		0.9	0.4	105 (154)	43 (223)

^a Deviations from *ab initio* relative energies are in parentheses. For N-methylformamide and N-methylacetamide, the energies are relative to the trans planar conformation.

^b rms deviations exclude frequency deviations from the amino puckering mode in molecules with NH₂ groups; the numbers in parentheses are the deviations in the frequency of this mode.

tion about the C'—N bond is quite steep, due to the large amount of energy required for breaking the partial π bond during the rotation, and this steep potential contributes to the vibrational frequency of out-of-plane deformation modes, because the nonorthogonal dihedral angles must change during an out-of-plane deformation. The torsion potential favors the planar configuration, resists out-of-plane deformations, and increases the frequency of the amino puckering mode. The out-of-plane force constant must then be reduced to mathematically counter the contribution of the torsion potential to the frequency of the amino puckering mode.

In the case of formamide, the intrinsic force constant for the out-of-plane coordinate in the amino group is quite small and the use of nonorthogonal torsion coordinates contributes more than the intrinsic out of plane force constants, thus requiring the force constant to be reduced to a smaller, and therefore negative, value. More generally, coordinate nonorthogonality is problematic for all commonly used force fields for molecular mechanics whenever the molecule of interest contains sp²-hybridized orbitals. One potential solution to this problem is the use of local symmetry torsion coordinates,³⁶ which consist of linear combinations of dihedral angles and which are orthogonal to out-of-plane coordinates. For example, Ermer and Lifson successfully used local symmetry

coordinates to distinguish twisting and puckering normal modes calculated from an alkene force field.^{37,38}

For force field calculations of the frequency of the amino puckering mode in formamide and acetamide, a (nitrogen out-of-plane) force constant value of about $-3 \text{ kcal mol}^{-1} \text{ rad}^{-2}$ would entirely eliminate the systematic error in the frequency. It should be noted, however, that, for the QMFF result with formamide, the 403 cm^{-1} frequency for the amino puckering mode is actually more consistent with the 289 cm^{-1} experimental³⁹ frequency than the 109 cm^{-1} *ab initio* result. The anomalously low frequency resulting from the quantum calculation reflects a large anharmonicity in the nitrogen out-of-plane deformation mode,³² and a consequent breakdown in the harmonic approximation for calculating frequencies.⁴⁰

To our knowledge there have been only two previous force field publications^{10,16b} that showed results for formamide frequencies. The MMFF parameter derivation resulted in a negative force constant for the nitrogen out-of-plane force constant for amides.^{16b} It is of interest to note that, in the MM3 study,¹⁰ a problem with the amino puckering frequency was not noted, because the force field was fit to a 507 cm^{-1} frequency⁴¹ that was obtained from an *ab initio* calculation and then scaled to fit an apparently misassigned experimental frequency. As a result, a negative force constant

was not required. The aforementioned 289 cm^{-1} experimental frequency³⁹ used above is now commonly assigned to the amino puckering mode^{32, 40, 42} and would presumably require a "negative force constant" in that earlier work as well. Unfortunately, however, because formamide is planar, a negative out-of-plane force constant is counterintuitive. The negative value erroneously implies an intrinsic instability of the sp^2 orbitals in amide nitrogen atoms and a consequent tendency toward nonplanar minimum energy structures with pyramidal nitrogen, as in nonconjugated functional groups (e.g., amines). In amides, the partial π bond in the C'—N bond stabilizes the sp^2 nitrogen orbitals, thereby causing unstrained amides to be planar or nearly planar, and a positive force constant for the nitrogen out-of-plane coordinate would therefore be expected. For these reasons, we chose to avoid the use of a negative, and therefore physically counterintuitive, value for the nitrogen out-of-plane force constant, but the drawback in this choice was that the force field cannot reproduce the frequency of the amino puckering mode in primary amides including formamide and acetamide.

Planarity of Amides

The planarities of the amides studied here, as determined by either experiment or quantum calculations, have not been determined conclusively. For formamide, analyses of microwave spectra have given conflicting conclusions.³⁵ In the case of acetamide, *N*-methylformamide (NMF), and *N*-methylacetamide (NMA), planarity was assumed in deriving the electron diffraction structures.^{43–45} For *N,N*-dimethylacetamide, recent electron diffraction results are ambiguous with respect to planarity.⁴⁶ Furthermore, the differences in energy between planar and nonplanar structures of amides as obtained by *ab initio* calculation are very small and depend on the basis set and the level of electron correlation used.^{47–52} Consequently, there does not appear to be any justification for attempting to obtain better agreement between the planarity of *ab initio* and QMFF structures, because the differences involve tenths of a kilocalorie per mole or less, as seen in what follows. Therefore, in this study, for all the species except urea, planarity was imposed in optimizing the molecular structures for consistency in comparing the two types of calculations.

Acetamide

The global minimum energy structure calculated by the force field is planar, whereas in the *ab initio* structure the carbonyl group is planar, but the amino group is somewhat puckered, as manifested by the 15.5° nitrogen out-of-plane coordinate. However, as previously discussed, the *ab initio* energy of the planar structure is only 0.02 kcal/mol higher in energy than the energy of the nonplanar global minimum. The force field also displays an extremely shallow energy surface along the out-of-plane coordinate, and the structure optimized with a 15.5° nitrogen out-of-plane coordinate is only 0.23 kcal/mol higher in energy than the planar structure. Thus, variations of up to 15.5° in the nitrogen out-of-plane coordinate result in relatively negligible energy changes in both the force field and quantum structures.

For the most part the internal coordinates and frequencies obtained for the planar structures with the quantum and force field (QMFF) calculations agree quite well. As indicated in Table IV, the rms bond length and bond angle deviations are only 0.004 \AA and 0.9° . As in formamide, the large rms frequency deviation including all modes (122 cm^{-1}) is mostly attributable to the frequency for the amino puckering mode. The *ab initio* planar structure is a transition state, as indicated by the negative eigenvalue for the amino puckering eigenvector, whereas the corresponding eigenvalue in the (planar) global minimum energy structure calculated with the force field is positive (381 cm^{-1}). The only other large frequency deviation is for the methyl torsion mode. The force field and quantum result for this frequency is 106 and 12 cm^{-1} , respectively. Excluding the amino inversion, the force field and quantum frequencies differ by 44 cm^{-1} (rms) as shown in Table V.

N-Methylformamide

The bond lengths calculated by the force field agree extremely well with *ab initio* calculations, and the rms and maximum deviations for both conformers are only 0.002 \AA and 0.003 \AA , respectively. Table IV shows that the rms deviations in the bond angles are 0.8° and 0.5° for the trans and cis structures respectively, which indicates that the bond angles also agree closely with the quantum structures. The maximum bond angle deviations in these two structures are 2.0° for the C'—N—H* angle in the former and 1.0° in the C—N—H* angle in the latter.

The frequencies calculated by the QMFF for both the trans and cis structures exhibit good agreement with the quantum mechanical results, with rms deviations of 45 and 38 cm^{-1} , respectively. The largest errors in frequency for both conformers are in the amino puckering mode. For example, in the trans conformer, the frequencies obtained by *ab initio* calculation and by the force field are 228 and 328 cm^{-1} , respectively. In the cis conformer, the N—CH₃ torsion and the amino puckering modes are heavily mixed. The deviations in frequencies calculated with the force field for the nitrogen out-of-plane deformation mode in the trans and cis conformers are small compared with the errors obtained in formamide and acetamide discussed previously. The addition of a methyl group to the amide nitrogen atom causes an increase in the intrinsic nitrogen out-of-plane force constant and, consequently, the absence of the need for a negative value for this force constant, unlike the results discussed previously for formamide. Thus, the QMFF gives much better agreement with the *ab initio* results for the frequency of the amino puckering mode in molecules with at least one carbon bonded to the nitrogen atom (as is the case for the peptide bond in peptides and proteins).

As indicated in Table V, the cis–trans conformational energy differences calculated by the QMFF and the *ab initio* method are 1.6 and 1.1 kcal/mol, respectively. Both results are in good agreement with the value of 1.6 kcal/mol determined from experiment.^{49,53}

N-Methylacetamide

The minimum energy structures for the trans and cis conformers are nonplanar according to both the *ab initio* and QMFF calculations. As in the case for acetamide and NMF, the energy difference between the planar and nonplanar structures is small. For example, the cis and trans planar structures of NMA are both only 0.02 kcal/mol higher in energy, according to calculations at the HF/6-31G* level of theory. Thus, the differences in planarity are not significant energetically and, as noted earlier, the exceptionally shallow energy surface is the reason that the extent of planarity is extremely sensitive to basis set, correlation, and other details of the quantum calculation.

A comparison of the QMFF force field and quantum results for structures and frequencies for the planar cis and trans configurations reveals only small differences. For example, Table IV shows

that the rms bond length and bond angle deviations are 0.002 Å and 0.8°, respectively, for the trans structure and are about the same for the cis conformer. Similarly, as seen from Table V, the rms frequency deviations are 50 and 39 cm^{-1} for the trans and cis conformers, respectively. In the trans structure, the maximum frequency deviation is 141 cm^{-1} for the C'—N torsion mode (Supplementary Material, Table III.E). All other deviations are less than 130 cm^{-1} . For the cis conformer the maximum deviation is 77 cm^{-1} for both the C'—CH₃ asymmetric deformation and the N—CH₃ torsion modes (Supplementary Material, Table III.F).

Importance of Coupling Terms to Account for Conformational Effects on Internal Geometry

Close inspection of differences between the QMFF and the harmonic diagonal force field (HDF) for the cis and trans conformers of NMA reveals the inability of the HDF to fit specific trends that occur for various structural parameters. For example, the *ab initio* bond angles for C₂—C₁—H₈, C₂—N₃—C₄, and C₂—N₃—H₆* change from 113.7°, 121.4°, and 119.6° for the trans structure to 108.3°, 127.4°, and 114.0° for the cis structure, respectively (Supplementary Material, Table II.D). Thus, these three angles change by −5.4°, 6.0°, and −5.6° from the trans to cis conformer. The changes in the angles with conformation are well accounted for by the QMFF: −5.5°, 6.0°, and −6.3°, respectively. However, the changes in these angles calculated with the HDF are only −1.4°, 3.3°, and −3.7° (Supplementary Material, Table II.D). Thus, the HDF is unable to reproduce significant changes in bond angles that occur between the trans and cis conformers. Instead, it tends to produce bond angles that average out the differences between the two conformers. The presence of cross-terms in the QMFF such as angle/torsion reflects the coupling in the energy surface and in turn accounts for these structural trends very accurately.

According to experimental measurements⁵⁴ the cis conformer is 2.8 kcal/mol higher in energy than the trans conformer. Table V demonstrates that the corresponding conformational energy differences calculated by HF/6-31G* and QMFF are 2.5 and 2.4 kcal/mol. Thus, both quantum and force field methods give good agreement with experiment for this important energy difference.

N,N-Dimethylformamide

The minimum energy structures obtained from the QMFF force field (with the planarity constraint) and *ab initio* calculation are very similar, as evidenced by the very small rms deviations in the bond lengths (0.002 Å), and bond angles (0.7°). Excluding the asymmetric methyl torsion, the rms frequency deviation is only 29 cm⁻¹, which is relatively small compared with those of many of the other molecules that have been tested. The force field calculation gives a negative eigenvalue for this problematic mode. This negative eigenvalue results from the imposed planarity constraint on the nitrogen. Due to coupling between the amino out-of-plane coordinate and *N*-methyl torsion coordinates in this mode, this manifests itself as a negative eigenvalue for the *N*-methyl torsion mode (rather than the expected amino inversion mode). The QMFF reproduces bond angles (0.7° rms) significantly better than the HDFF (1.4° rms) for this molecule. In addition, there are several bond angles that are significantly worse for the HDFF than the QMFF in comparison with the *ab initio* result. These angles are O₁—C₂—H₆, N₃—C₂—H₆, and N₃—C₅—H₁₀. The QMFF deviations for these angles are -0.6°, 0.3°, and 0.9°, respectively, and the HDFF deviations are -2.7°, 2.5°, and 2.8° (Supplementary Material, Table II.E).

N,N-Dimethylacetamide

The agreement between the QMFF and *ab initio* structures is good for *N,N*-dimethylacetamide, as indicated by the respective 0.003 Å and 0.8° rms deviations for bond lengths and angles. The 48 cm⁻¹ rms deviation between frequencies obtained by the force field and *ab initio* calculation is comparable to deviations found in other amides that were tested. The frequency calculated by the force field for the C'—N bond stretching mode is 109 cm⁻¹ larger than the *ab initio* result.

Comparison of the QMFF and HDFF results shows the coupled anharmonic force field to be able to account for the structural parameters of this molecule, particularly bond angles, significantly better than the diagonal quadratic form (HDFF). The overall rms deviation for bond angles improves from 1.4° for the HDFF to 0.8° for QMFF as shown in Table IV. In addition, there are four bond angles in which the deviation between the HDFF and *ab initio* result ranges from 2.1° to 4.2°, whereas the maximum deviation for the QMFF is only 1.4° (Supplementary Material, Table II.F).

AMIDE-RELATED COMPOUNDS

Urea

For urea, the QMFF force field and *ab initio* calculations give optimized structures in which the nitrogen out-of-plane coordinates are 37.4° and 40.8°, respectively. Thus, both calculations are in agreement that the nitrogen atoms are sp³ hybridized. Analysis of the force field energies shows that the nonplanarity of urea is accounted for by the electrostatic terms in the nonbonded potential. The Coulombic attraction between each nitrogen atom and a hydrogen attached to the other nitrogen favors puckering because it shortens the N...H distance. The puckering is also favored by a Coulombic repulsion between the two closest nonbonded hydrogen atoms. These interactions are sufficiently strong to overcome the tendency of the sp² hybridization of the nitrogen to remain planar. The 2.3° rms deviation in bond angles (Table VI) is larger than the deviations found for the amides (Table IV). The maximum bond angle deviation is 4.1°, found in one of the C'—N—H* bond angles in both amino groups (Supplementary Material, Table II.G). Finally, as indicated by the small rms deviation in bond lengths (0.003 Å), the force field accurately reproduces the quantum results for bond lengths in urea.

Because urea contains the NH₂ group, systematic errors in the frequencies for the amino puckering modes might be expected, as discussed previously for formamide and acetamide. In urea, the QMFF force field underestimates the quantum frequencies for the symmetric and asymmetric NH₂ bending modes by 51 and 85 cm⁻¹, respectively

TABLE VI.
Comparison of Accuracy of Bond Lengths and Bond Angles in Structures Optimized by HDFF and QMFF: rms Deviations from *Ab Initio* Structures of Three Amide-Related Molecules.

Molecule / structure	Bonds (Å)		Angles (°)	
	HDFF	QMFF	HDFF	QMFF
Urea	0.035	0.003	1.9	2.3
<i>N</i> -formylformamide	0.005	0.003	1.9	0.3
Butyrolactam	0.013	0.004	3.2	0.4
Average rms deviation	0.018	0.003	2.3	1.0

(Supplementary Material, Table III.I). Contrary to the results for formamide and acetamide, however, the amino puckering mode frequencies obtained by the force field and *ab initio* calculation differ by only 9 and 72 cm^{-1} , even though the deviations are quite large for primary amides that are planar. Overall, the 61 cm^{-1} rms deviation in the frequencies (Table VII) is larger than the 28–50 cm^{-1} range (Table V) found for amides that have been tested. The largest frequency deviations (-102 and -135 cm^{-1}) are due to the NH_2 rocking modes.

Overall, the quality of fit by the QMFF force field for urea is somewhat lower than for the simple amides, possibly due to the presence of two amino groups bonded to the single carbonyl. This degradation in fit to the quantum energy surface may indicate some lack of transferability of the parameters to urea. This is the subject of further investigation.

N-Formylformamide

Tables VI and VII demonstrate that the QMFF force field gives very good agreement with *ab initio* results for both the structure and the frequencies of N-formylformamide. Both the force field and the quantum calculations indicate that N-formylformamide is planar. The rms deviations in the bond lengths, angles, and frequencies are 0.003 Å, 0.3°, and 33 cm^{-1} , respectively, whereas the largest frequency deviation is 69 cm^{-1} for the NH bond stretching mode (Supplementary Material, Table III.J). The agreement with the quantum results indicates that the amide force field is also transferable to bifunctional molecules in which the

amide has an extra carbonyl group attached to the nitrogen.

Comparison of the HDFF and QMFF for this system reveals, once again, the significant improvement produced by the QMFF in the quality of the fit, particularly for bond angles. The rms deviation in bond angles from the *ab initio* structure changes from 1.9° for the HDFF to 0.3° for the QMFF, as seen from Table VI. Most of this deviation is caused by the poor fits to the angles involving the two C—H bonds. These deviations are 2.7° and 2.9° for the HDFF and 0.0° and -0.4° for the QMFF (Supplementary Material, Table II.H).

Butyrolactam

Butyrolactam provides a stringent test of the range of applicability of the amide QMFF force field, because here the amide group is incorporated into a strained four-membered ring. The force field and *ab initio* calculations show good agreement on the structure of butyrolactam, and both results agree that the amino group is planar. In addition, the 0.004 Å and 0.4° rms deviations in bond lengths and bond angles, respectively, are small (Table VI). In comparison, the HDFF does substantially worse for this molecule, once again especially for bond angles. The rms bond angle deviation is 3.2° as compared with the 0.4° deviation in the QMFF. Five bond angles (all involving the nitrogen atom) have deviations ranging from 4.6° to 6.5° whereas the maximum deviation for the QMFF for this system is only 0.9° (Supplementary Material, Table III.I). Obviously, the HDFF functional form is not able to account for the structure of a four-membered ring compound using the same parameters needed to fit acyclic structures. On the other hand, the QMFF fits this small ring compound quite accurately *with no additional parameters or atom types* beyond those used for the acyclic amides.

As seen in Table VII, the overall rms deviation between frequencies obtained by the QMFF force field and *ab initio* calculation is 70 cm^{-1} . Five deviations (Supplementary Material, Table III.K) exceed 100 cm^{-1} : -166 , 150, -142 , 126, and -100 cm^{-1} for the CH_2 wag, ring puckering, in-plane ring deformation, amino puckering, and NH rocking modes, respectively. The large frequency deviation for the CH_2 wag mode is comparable to the -150 cm^{-1} deviation for the D_{4h} transition state structure of cyclobutane, and thus is inherent in the alkane parameters.

TABLE VII.
Comparison of Accuracy of Vibrational Frequencies Calculated by HDFF and QMFF for Three Amide-Related Molecules.

Molecule / structure	rms deviations in frequencies (cm^{-1})	
	HDFF	QMFF
Urea	136	61
N-formylformamide	116	33
Butyrolactam	225 ^a	70

^a This rms deviation excludes the frequency for the ring puckering mode. According to *ab initio* calculation, the frequency of this mode is 408 cm^{-1} . A negative eigenvalue was calculated for this mode with the HDFF.

Comparison of Specific Structural Features among Differing Compounds

BOND LENGTHS

From the summary including all bond lengths, in Table IV, it can be seen that in general bond lengths in amides are calculated much more accurately with the QMFF than with the harmonic diagonal force field. A more detailed assessment of the errors is given in what follows, where the focus is on the ability of the force field to reproduce trends in bond lengths in equilibrium conformations. The results of applying the QMFF to transition states, such as those involved in *cis*–*trans* isomerization or in the rotation of a methyl group, are discussed elsewhere.¹⁸ The bond lengths calculated by the two force fields are used as a probe to demonstrate the extent to which coupling interactions are reflecting properties inherent in the quantum mechanical energy surface. Although these effects are small, they demonstrate fundamental properties of the molecular energy surface, which can become significant when addressing such properties as valence angles about the C^α atom in peptides.^{30, 31}

C'—N BONDS

From Table VIII it can be seen that the average of the amide bond lengths for these equilibrium

configurations is 1.352 Å according to the *ab initio* calculations. The 0.016 Å spread in C'—N bond lengths is mainly due to somewhat longer bonds in the acetamides relative to the formamides. With the harmonic diagonal force field, the average bond length is overestimated by 0.044 Å relative to the *ab initio* values. In contrast, with the QMFF the average bond length of 1.353 Å deviates by only 0.001 Å from the average quantum result.

The accuracy of the bond lengths calculated by the QMFF is primarily due to the bond/torsion coupling interaction term in eq. (1). This term accounts for changes in forces acting along a bond as a rotation about that bond takes place.⁴ In the case of the amide C'—N bond, the forces are different for structures that are approximately planar than for structures in which the amino group is rotated about the amide bond, because the partial π bond in the nearly planar structures becomes weaker in twisted structures. The coupling interaction allows the QMFF to accurately reproduce the variations in the forces and the resulting configuration dependence of the bond lengths throughout all the structures in the training set, which include both planar and nonplanar geometries. This coupling is also responsible in part for the ability of the QMFF to reproduce trends in the *ab initio* results, such as the longer bond lengths for acetamides (relative to formamides) and for the longer bond length in *cis*-*N*-methylacetamide relative to the *trans* conformer.

TABLE VIII. Comparison of Amide and Carbonyl Bond Lengths in Structures Optimized by *Ab Initio* Calculation, H DFF, and QMFF.^a

	C' — N bond lengths (Å)			C' — O' bond lengths (Å)		
	HF / 6-31G*	H DFF	QMFF	HF / 6-31G*	H DFF	QMFF
Formamide	1.348	1.377 (0.029)	1.347 (–0.002)	1.193	1.187 (–0.006)	1.193 (0.000)
Acetamide	1.356	1.369 (0.013)	1.352 (–0.004)	1.198	1.187 (–0.011)	1.196 (–0.001)
<i>N</i> -methylformamide, <i>trans</i>	1.345	1.401 (0.056)	1.348 (0.003)	1.196	1.190 (–0.006)	1.196 (0.001)
<i>N</i> -methylformamide, <i>cis</i>	1.347	1.389 (0.041)	1.349 (0.002)	1.196	1.189 (–0.006)	1.194 (–0.001)
<i>N</i> -methylacetamide, <i>trans</i>	1.350	1.398 (0.048)	1.351 (0.001)	1.201	1.190 (–0.011)	1.202 (0.001)
<i>N</i> -methylacetamide, <i>cis</i>	1.357	1.389 (0.032)	1.358 (0.001)	1.200	1.189 (–0.012)	1.197 (–0.003)
<i>N,N</i> -dimethylformamide	1.349	1.415 (0.067)	1.350 (0.002)	1.196	1.192 (–0.004)	1.197 (0.001)
<i>N,N</i> -dimethylacetamide	1.361	1.430 (0.069)	1.366 (0.006)	1.203	1.192 (–0.010)	1.201 (–0.001)
Average ^b	1.352	1.396 (0.044)	1.353 (0.001)	1.197	1.190 (–0.008)	1.197 (–0.001)

^a Deviations between bond lengths calculated from the *ab initio* method and from the force fields are given in parentheses.
^b Average bond length. The numbers in parentheses are the average deviations between bond lengths calculated *ab initio* and by the H DFF or QMFF force fields.
^c rms deviation from the average bond length. This is a measure of the spread or range of bond lengths calculated by a given method. Similarly, the numbers in parentheses are the rms deviations in the differences between bond lengths calculated *ab initio* and by the force field.

Because coupling interactions are not included in the functional form of harmonic diagonal force fields, these effects cannot be accounted for by the HDFF. As a result, the HDFF gives bond lengths that are relatively independent of the conformation, contrary to observed behavior, and the fit to the complete training set of molecular structures, which contains both planar and rotated geometries, is degraded. The HDFF is not able to reproduce the trends in the C'—N bond lengths such as the difference between *cis*- and *trans*-*N*-methylacetamide, or the longer C'—N bond lengths in the acetamides than in the formamides (Table VIII). All three of these deficiencies in the HDFF are directly attributable to its neglect of coupling interactions that are present in the QMFF.

OTHER BONDS IN AMIDES

As previously discussed, the bond/torsion interaction is the most important coupling term for determining the conformational dependence of C'—N and C'—O' bond lengths in amides. The bond/bond and bond/angle interactions also are significant but have a smaller effect, as ascertained by removing these terms from the QMFF and observing the resulting changes in the optimized bond lengths. In general, this conclusion is also valid for other bonds found in the amide functional group. Table IX summarizes the average values of amide bond lengths determined from quantum and force field structure optimizations of planar amide conformations, in addition to the rms deviations between bond lengths determined from

quantum and force field methods. From a comparison of the average bond lengths and the rms deviations in Table IX, it can be seen that, in all cases, except for the C'—C bond, the coupled anharmonic QMFF reproduces the features of the quantum energy surface much more accurately than the HDFF.

BOND ANGLES

The use of coupling interactions by force fields also results in small but significant improvements in the accuracy with which bond angles can be calculated, as has been pointed out earlier in several examples. From Table IV it can be seen that the rms deviations for bond angles are almost a factor of two lower for the QMFF than for the HDFF. This improvement in many cases reflects that trends observed from *ab initio* calculations are better reproduced by the coupled, anharmonic potential. As an example, from both the quantum and the QMFF results in Table X, it can be seen that the O'—C'—N angle is an average of 3.1° smaller in acetamides than in formamides. With the HDFF, this trend is less pronounced, being on average only 1.8° smaller. Additional examples demonstrating the extent of coupling implicit in the quantum energy surface are given in Table XI, wherein the rms deviation of each bond angle type is compared for the two force fields. The largest rms deviation from the *ab initio* result is 1.6° (for the C'—N—H* angle) with the QMFF, but is 3.0° (for the H—C'—N angle) with the harmonic diagonal representation (HDFF).

TABLE IX. Comparison of Average Values of Bond Lengths (Angstroms) and Their rms Deviations for Eight Amide Structures^a Optimized by *Ab Initio* Calculation, HDFF, and QMFF.

Bond	Bond lengths (Å)				
	HF/6-31G*	HDFF		QMFF	
	Average	Average	rms dev.	Average	rms dev.
C'—N	1.352	1.396	0.048	1.353	0.003
C'—O'	1.197	1.190	0.009	1.197	0.001
C'—C	1.515	1.514	0.002	1.512	0.005
C'—H	1.091	1.095	0.004	1.090	0.002
N—C	1.446	1.454	0.010	1.446	0.004
N—H*	0.994	0.999	0.005	0.995	0.002

^a The eight structures are listed in Table IV. Individual C'—N and C'—O' bond lengths are given in Table VIII.

TABLE X. Comparison of O' — C' — N Bond Angles in Planar Structures Optimized by *Ab Initio* Calculation, HDFF, and QMFF.

	O' — C' — N bond angles (°)		
	HF / 6-31G*	HDFF	QMFF
Formamide	124.9	123.2 (– 1.8)	124.2 (– 0.7)
Acetamide	122.2	121.3 (– 0.9)	121.4 (– 0.9)
<i>N</i> -methylformamide, trans	124.8	125.8 (1.0)	124.6 (– 0.2)
<i>N</i> -methylformamide, cis	124.9	123.4 (– 1.5)	125.1 (0.2)
<i>N</i> -methylacetamide, trans	122.2	123.7 (1.4)	122.1 (– 0.1)
<i>N</i> -methylacetamide, cis	121.3	120.0 (– 1.4)	121.9 (0.6)
<i>N,N</i> -dimethylformamide	125.9	126.1 (0.2)	126.2 (0.3)
<i>N,N</i> -dimethylacetamide	122.1	122.7 (0.6)	122.2 (0.1)
Average	123.5	123.3 (– 0.3)	123.5 (– 0.1)

TABLE XI. rms Deviations between Bond Angles Obtained by *Ab Initio* Calculation and by Force Fields for Eight Amide Structures.^a

Angles	rms deviations (°)	
	HDFF	QMFF
O' — C' — N	1.2	0.5
O' — C' — C	0.7	0.8
O' — C' — H	2.4	0.4
C — C' — N	1.4	1.1
H — C' — N	3.0	0.6
C' — N — C	1.4	1.1
C' — N — H*	1.0	1.6
C — N — C	1.3	0.8
C — N — H*	1.3	0.8
H* — N — H*	1.0	1.1

^a The eight amide structures are listed in Table IV. Individual O' — C' — N angles are given in Table X.

Range of Applicability of Force Field

Urea, *N*-formylformamide, and butyrolactam provide a challenging test of the range of applicability of the amide force field. Of particular interest are the structural and spectroscopic features that are strikingly outside the range of values found for the eight structures spanning the six simple amides previously discussed.

C'—N BOND LENGTHS

The C'—N bond lengths obtained by *ab initio* calculation and by the force fields are listed in Table XII. The bond length from the *ab initio* result is 1.373 Å for both urea and *N*-formylformamide. This value is significantly longer than the bond

TABLE XII. Comparison of C' — N and C' — O' Bond Lengths in Amide-Related Molecules Optimized by *Ab Initio* Calculation, HDFF and QMFF.

	C' — N bond lengths (Å)			C' — O' bond lengths (Å)		
	HF / 6-31G*	HDFF	QMFF	HF / 6-31G*	HDFF	QMFF
Urea	1.373	1.308 (– 0.065)	1.372 (– 0.001)	1.197	1.180 (– 0.017)	1.192 (– 0.005)
<i>N</i> -formylformamide	1.373	1.379 (0.006)	1.373 (0.000)	1.185	1.188 (0.003)	1.189 (0.003)
Butyrolactam	1.358	1.384 (0.027)	1.354 (– 0.004)	1.186	1.189 (0.002)	1.186 (– 0.001)
rms deviation ^a		0.041	0.002		0.010	0.004

^a rms deviation between bond lengths calculated *ab initio* and by the force field. For consistency with the results listed in Tables VI and XIII, the rms deviations for the C' — N bond lengths were counted twice in urea and *N*-formylformamide, because there are two C' — N bonds in these molecules. Similarly, the rms deviations for the carbonyl bond were counted twice in *N*-formylformamide.

lengths of the simple amides seen in Table VIII (average value 1.352 ± 0.005 Å). For urea and *N*-formylformamide the QMFF force field result is 1.372 and 1.373 Å, respectively, which is in exceptionally good agreement with the *ab initio* result. The coupling interactions that are most responsible for lengthening the bond are the bond/angle interactions in urea and the bond/torsion interactions in *N*-formylformamide. Because coupling interactions are not accounted for by the harmonic diagonal force field, significant deviations from the quantum result are expected, and the error of -0.065 that results in urea with the HUFF is indicative of the importance of cross-terms for determining the C'—N bond length in this molecule. The use of cross-terms clearly improves the applicability of the amide force field to these related chemical functionalities.

The four-membered ring compound, butyrolactam, on the other hand, has the shortest C'—N bond of the three compounds, a relatively standard amide bond length (1.358 Å). This again is well reproduced by the QMFF force field, while simultaneously accounting for the stretching in urea and *N*-formylformamide. The harmonic diagonal force field, however, breaks down when applied to the four-membered ring, predicting a C'—N bond length of 1.384 Å, longer than either of the previous two compounds and almost 0.03 Å too long.

Table XIII gives a summary of the rms and maximum deviations between bond lengths obtained by *ab initio* calculation and by the force fields for all three derivatized amides. In all cases except for the N—H* bond, wherein both force fields have an rms deviation of 0.003 Å, the rms deviations resulting from the HUFF are substantially larger than those resulting from the QMFF.

As discussed earlier, the inclusion of coupling interactions is the primary reason for the improved accuracy and transferability of the QMFF.

Comparison of Specific Vibrational Frequencies of Amides

In addition to the structural properties, it is important to test the force field's ability to account for dynamic properties. Vibrational spectra are a sensitive probe of this property, and thus we have explored the ability of the two force fields' functional forms to account for amide dynamics.

Both anharmonicity and coupling interactions are well-known to be important for accurate calculations of vibrational frequencies.^{4, 24, 27–29} Consequently, vibrational frequencies calculated with the QMFF would be expected to be more accurate than those determined with harmonic diagonal force fields. This turns out to be the case as seen from the overview given in Table V, wherein the accuracy of vibrational frequencies calculated with the QMFF and HUFF force fields are compared for the six simple amide molecules. In general, the rms deviations from the *ab initio* results tend to be at least a factor of two lower with the QMFF. To delineate the intramolecular coupling and anharmonic interactions responsible for these frequency shifts, we analyze their effects on several important modes.

C'—N BOND STRETCHING MODE

In Table XIV we compare the rms deviations for 20 characteristic amide normal modes. An extreme example of the combined importance of anharmonicity and coupling is the stretch of the amide

TABLE XIII.
rms and Maximum Deviations between Bond Lengths (Angstroms) Obtained by *Ab Initio* Calculation and by Force Fields for Three Amide-Related Molecules.^a

Bonds	HUFF		QMFF	
	rms dev.	Maximum dev. ^b	rms dev.	Maximum dev. ^b
C'—N	0.043	−0.065 (urea)	0.002	−0.004 (but.)
C'—O'	0.009	−0.017 (urea)	0.003	0.005 (urea)
C'—C	0.006	−0.006 (but.)	0.005	−0.005 (but.)
C'—H	0.006	0.006 (NFF)	0.001	0.001 (NFF)
N—C	0.026	−0.026 (but.)	0.006	−0.006 (but.)
N—H*	0.003	−0.004 (NFF)	0.003	−0.007 (NFF)

^a Molecules: urea, *N*-formylformamide (NFF), and butyrolactam (but.).

^b The molecule in which the maximum deviation occurred is listed in parentheses.

TABLE XIV.
Comparison of rms Deviations between Vibrational Frequencies Obtained from *Ab Initio* Calculations and from Force Fields for Eight Amide Structures.^a

Mode	rms deviations (cm ⁻¹)	
	HDFF	QMFF
N—H stretch	46.0	28.0
C'—H stretch	59.9	16.9
C'—O' stretch	21.5	15.8
C'—N stretch	231.6	29.2
C'—C stretch	171.6	24.0
N—C stretch	160.1	41.4
N—H bend	100.5	20.9
NH ₂ scissors	260.9	12.7
NH ₂ rock	192.3	65.9
C'—H bend	52.8	25.0
C'—O' bend	107.6	14.8
C'—C bend	81.3	34.5
N—C bend	23.6	19.7
CNC scissors	5.6	17.1
CNC rock	36.3	9.7
C'—N torsion	52.3	65.5
C'—CH ₃ torsion	39.1	88.2
N—CH ₃ torsion	33.7	113.2
C'—O' oop def	102.1	35.8
Amino inversion ^b	66.8	77.8

^a The eight structures are listed in Table IV. Specific C'—N stretch, C'—O' bend, and C'—N torsion frequencies are given in Tables XV—XVII.

^b Excludes frequencies in formamide and acetamide, for reasons discussed in the text.

bond (i.e., C'—N bond). It can be seen that the rms deviation for this mode with the harmonic diagonal force field (HDFF) is 232 cm⁻¹, almost an order of magnitude larger than the corresponding deviation of 29 cm⁻¹ resulting from the use of the

QMFF. As previously discussed, the C'—N bond length substantially increases as the amino group is rotated about the amide bond. Because of anharmonicity both the effective force constant and the frequency for stretching of the C'—N bond would be expected to decrease. More importantly, however, as the amino group is rotated away from the planar structure, the partial π bond is broken, which means that the C'—N bond length is coupled to the torsion angles about the amide bond. This coupling interaction will also reduce the effective force constant and therefore the stretching frequency for the C'—N bond.

From Table XV it can be seen that the HDFF systematically underestimates the frequency of the C'—N stretching mode by an average of 218 cm⁻¹ for these eight amide conformations. This indicates that, during the HDFF parameter fitting, the force constant for the C'—N bond was strongly influenced by the energy derivative data for transition state configurations in which the effective C'—N force constant was substantially lower, due to breaking the partial π bond. Although, in principle, the data for the nearly planar configurations could be weighted more strongly, it would nevertheless be impossible for the HDFF to be sufficiently flexible to correctly determine the effective force constant and hence the stretching frequency for the C'—N bond for all sampled amide configurations with the range of distortions corresponding to the energies in Table I. With the QMFF this problem does not occur, as can be seen from the small rms deviation for this frequency, because the energy function terms for coupling interactions, such as the bond/torsion coupling, as well as anharmonicity, account for the bond lengthening and the reduced stretching frequency for the C'—N

TABLE XV.
Comparison of Vibrational Frequencies for C'—N Stretch Mode in Structures Optimized by *Ab Initio* Calculation, and by HDFF and QMFF.

	Frequencies (cm ⁻¹) for C'—N stretch		
	HF / 6-31G*	HDFF	QMFF
Formamide	1378.6	1168.6 (–210.0)	1333.7 (–44.9)
Acetamide	1463.6	1324.0 (–139.6)	1485.5 (21.9)
N-methylformamide, trans	1348.6	1113.1 (–235.5)	1304.3 (–44.3)
N-methylformamide, cis	1412.8	1049.9 (–362.9)	1386.1 (–26.7)
N-methylacetamide, trans	1408.2	1289.9 (–118.3)	1416.8 (8.6)
N-methylacetamide, cis	1469.5	1314.4 (–155.1)	1492.6 (23.1)
N,N-dimethylformamide	1586.8	1284.7 (–302.1)	1601.8 (15.0)
N,N-dimethylacetamide	1593.0	1371.7 (–221.3)	1621.9 (28.9)
Average deviation		–218.1	–2.3

bond in configurations in which the partial π bond is weakened or broken. We note that this range of twisting or distortion of the amide occurs in equilibrium structures of proteins, due to environment, as well as being visited due to thermal motion at room temperature, even for planar amides. Thus, for an accurate representation of these properties, the energy surface must account for the coupling over this range of distortion.

C'—O' ROCKING MODE

Another example of the difference in accuracy with which the QMFF and the harmonic diagonal force field are able to reproduce the frequencies obtained by *ab initio* calculation is given in Table XVI for the in-plane rocking mode of the carbonyl bond. In general, it can be seen that, according to the *ab initio* results, the frequency of this mode tends to be higher for formamides than for acetamides. For example, the frequencies in the trans conformers of NMF and NMA are 838 and 674 cm^{-1} , respectively, a difference of 164 cm^{-1} . Both the HDFF and QMFF force fields give higher frequencies for NMF, consistent with this result. However, the shift in the harmonic diagonal force field is only half that obtained from quantum mechanics (difference of 86 cm^{-1}), whereas the QMFF gives essentially the same shift (159 cm^{-1}).

The *ab initio* calculations also show that this frequency is larger for the trans conformers of NMF and NMA than for the cis, although the conformational dependence of this frequency is significantly greater for NMF, wherein the frequencies of the trans and cis conformers are 838

and 659 cm^{-1} , as opposed to the corresponding frequencies of 674 and 612 cm^{-1} in NMA. One reason for these differences is that the effective force constants are conformationally dependent, as in the case of the force constant for the C'—N bond stretching, whereas the other reason is that there is a qualitative difference in the mechanical coupling in the two conformations. For example, in *cis*-NMA the rocking motion of the C'—H bond is significantly more coupled to the C'—O' rocking mode than in *trans*-NMA. To the extent that mechanical coupling contributes to shifts in the frequency of the C'—O' rocking mode, the harmonic diagonal force field would be expected to at least qualitatively reproduce the conformational dependence observed for the frequencies in the quantum calculation, as is found. However, as in the case of the C'—N stretching mode, to the extent that the effective force constant for the N—C'—O' bond angle is conformationally dependent, the HDFF would not be able to quantitatively reproduce the variation in the frequencies obtained from *ab initio* calculation. The ability of the QMFF (and the inability of the HDFF) to quantitatively reproduce the *ab initio* frequency differences between the trans and cis conformers of NMF and NMA suggests that the effective force constants are indeed significantly different in the two sets of conformers. In NMA, in particular, the 654 and 598 cm^{-1} frequencies calculated by the QMFF for the trans and cis conformers deviate from the quantum result by only 20 and 14 cm^{-1} , respectively. In contrast, the HDFF gives frequencies of 606 and 458 cm^{-1} for these two conformers, and these differ from the *ab initio* values by 68 and 154 cm^{-1} . The

TABLE XVI.
Comparison of Vibrational Frequencies for C' — O' Rocking Mode in Structures Optimized by *Ab Initio* Calculation, and by HDFF and QMFF.

	Frequencies (cm^{-1})		
	HF / 6-31G*	HDFF	QMFF
Formamide	617.6	578.3 (–39.3)	627.6 (9.7)
Acetamide	598.1	558.2 (–39.9)	605.7 (7.6)
N-methylformamide, trans	837.8	692.0 (–145.8)	812.8 (–25.0)
N-methylformamide, cis	659.3	522.5 (–136.8)	648.9 (–10.4)
N-methylacetamide, trans	673.6	605.8 (–67.8)	653.6 (–20.0)
N-methylacetamide, cis	611.9	457.7 (–154.2)	598.0 (–13.9)
N,N-dimethylformamide	711.9	586.9 (–125.0)	695.7 (–16.2)
N,N-dimethylacetamide	639.1	564.8 (–74.3)	635.7 (–3.4)
Average		–97.9	–8.9

large difference in these two deviations is further evidence of a conformational dependence of the effective force constants.

From Table XVI it is clear that the HDFF systematically underestimates the frequencies of the C'—O' rocking mode for all of the amides. As in the case of the C'—N stretching mode, this is an indication that, during the parameterization of the force fields, energy derivative data for configurations with relatively low effective force constants had an important effect in determining the force constants (and, as noted previously, the harmonic diagonal form cannot account for the conformational coupling). The effective force constants can be calculated by a direct transformation of the energy derivatives from the Cartesian to the internal coordinate representation.³⁶ This calculation has been done, and it was found that configurations of NMA for which the partial π bond character of the amide bond is reduced have particularly small effective force constants for C'—O' rocking. Thus, as noted earlier, the HDFF is insufficiently flexible to simultaneously accommodate the energy surface of equilibrium and moderately distorted structures typical of those found in macromolecules, strained molecules, or molecular dynamics simulations.

C'—N TORSION MODE

In Table XVII we compare the frequencies for twisting about the C'—N bond in the six amides. Neither the HDFF nor the QMFF fits the *ab initio* frequencies with an accuracy comparable to that of the other modes. For example, for the trans conformers of NMF and NMA, the frequencies calculated by the QMFF are too large by 88 and 141

cm⁻¹, respectively. A detailed analysis, to be presented elsewhere,¹⁸ shows that the *ab initio* frequencies are strongly influenced by the presence of out-of-plane/torsion coupling, which is missing in both the HDFF and the QMFF functional forms. The absence of this term is responsible for the relatively poor fit of the QMFF to these frequencies. In addition, this is one of the few cases where the HDFF fits the quantum mechanical energy surface as well as, if not slightly better than, the QMFF. Analysis reveals that larger ¹K_φ torsion constants for the amide bond in the HDFF compared with the QMFF are apparently responsible for the HDFF's ability to reproduce these frequencies. These parameters were damped in deriving the QMFF through the use of penalty functions, which is why it reproduces the *ab initio* frequencies slightly less accurately. However, it is our experience that the relatively large ¹K_φ parameters, such as those that result in this case from fitting the HDFF without penalty functions, exhibit relatively poor transferability when applied to other systems. A more correct (transferable) way to model the amide torsion frequencies involves the use of orthogonal torsion and out-of-plane coordinates and out-of-plane/torsion cross-terms, as will be discussed in more detail elsewhere.

Comparison of Frequencies Calculated for Amide-Related Molecules by HDFF and QMFF

Table XVIII summarizes the accuracy by which the HDFF and the QMFF reproduce frequencies obtained by *ab initio* calculations for a variety of

TABLE XVII.
Comparison of Vibrational Frequencies for C' — N Torsion Mode in Structures Optimized by *Ab Initio* Calculation, and by HDFF and QMFF.

	Frequencies (cm ⁻¹) for C' — N torsion		
	HF / 6-31G*	HDFF	QMFF
Formamide	673.4	679.5 (6.1)	697.2 (23.8)
Acetamide	562.6	641.6 (79.0)	614.1 (51.5)
N-methylformamide, trans	486.9	560.8 (73.9)	574.4 (87.5)
N-methylformamide, cis	648.6	631.0 (− 17.6)	698.3 (49.7)
N-methylacetamide, trans	365.9	373.4 (7.5)	506.5 (140.6)
N-methylacetamide, cis	508.3	594.3 (86.0)	477.9 (− 30.4)
N,N-dimethylformamide	346.1	394.9 (48.8)	352.0 (5.9)
N,N-dimethylacetamide	143.1	145.8 (2.7)	128.7 (− 14.4)
rms deviation		52.3	65.5

TABLE XVIII.
Comparison of rms Deviations Between Vibrational
Frequencies Obtained from *Ab Initio* Calculations
and from Force Fields for Three
Amide-Related Molecules.^a

Mode	rms Deviations (cm ⁻¹)	
	HDFF	QMFF
N—H stretch	75.9	29.2
C'—H stretch	94.6	20.4
C'—O' stretch	39.1	20.3
C'—N stretch	298.6	210.7
N—H bend	345.5	100.2
NH ₂ bend	133.2	70.2
NH ₂ rock	180.6	119.8
C'—H bend	57.3	7.3
C'—O' bend	88.9	26.3
C'—N torsion	52.3	32.5
C'—O' oop def	74.4	49.8
Amino inversion	162.3	72.6
All modes	157.8	89.2

^a Urea, *N*-formylformamide, butyrolactam.

normal modes in urea, *N*-formylformamide, and butyrolactam. The rms deviations are lower for the QMFF for all normal modes, including the modes due to torsions or out-of-plane deformations, which in some cases were problematic for the amides (see Table XIV). This suggests that the inclusion of anharmonicity and coupling interactions in the potential energy function substantially extends the range of applicability of the amide force field to related functional groups.

Summary

A force field for amides and functionally substituted compounds was derived by fitting 732 force constants and reference values to 140,970 quantum observables consisting of the relative energies and first and second derivatives of the energy of 638 structures of 17 different molecules. The importance of anharmonicity and cross-term interactions was demonstrated by comparing force fields derived with and without these terms. The harmonic diagonal force field fit the energies, and the first and second energy derivatives of amides in the training set to an accuracy of 2.3 kcal/mol, 42%, and 21%, respectively, whereas the force field with terms representing anharmonicity and cross-terms fit the same data to an accuracy of 1.1 kcal/mol, 15%, and 5%. Thus, these interaction terms, which

are for the most part not present in class I protein force fields (e.g., AMBER,⁶ CHARMM,⁸ GROMOS,⁹ or in the MM3 amide force field¹⁰), improve the accuracy of calculated energies by 1.2 kcal/mol, while also resulting in factors of three and four improvements in the accuracies of calculated first and second derivatives of the energy, respectively. Similar improvements in the fit of *ab initio* data by the QMFF (relative to HDFF) were obtained for six amide-related molecules (i.e., urea, *N*-formylformamide, butyrolactam, *N*-formylaziridine, azacyclopropanone, and methylazacyclopropanone).

It is generally recognized that there are many other terms not considered here, such as improved electrostatic representations⁵⁵ and charge polarization.⁵⁶ These effects must be included in developing increasingly accurate molecular force fields, and work is in progress to address them. Also, because in the present work the same functional form was used for both the alkane and amide calculations, the larger deviations for amides suggests that interaction terms not included in eq. (1) may be important for amides but not for alkanes. This possibility has been investigated and confirmed in a detailed analysis of rotational barriers, inversion barriers, structures, and frequencies calculated by the force field and by quantum mechanics.¹⁸ **Nevertheless, the comparison of results obtained with the HDFF and QMFF force fields demonstrates that the anharmonic and coupling interaction terms included in the current functional form [i.e., eq. (1)] dramatically improve not only the ability of the QMFF to account for the amide energy surface, but also the molecular structures and vibrational frequencies calculated by *ab initio* methods.**

Acknowledgment

The authors thank Drs. Marvin Waldman and Carl Ewig for carefully reviewing this manuscript and making a number of useful suggestions.

References

1. D. H. Kitson, F. Avbelj, J. Moulton, D. T. Nguyen, J. E. Mertz, D. Hadzi, and A. T. Hagler, *Proc. Natl. Acad. Sci. USA*, **90**, 8920 (1993).
2. J. R. Maple, M.-J. Hwang, T. P. Stockfisch, U. Dinur, M. Waldman, C. S. Ewig, and A. T. Hagler, *J. Comput. Chem.*, **15**, 162 (1994).
3. M.-J. Hwang, T. P. Stockfisch, and A. T. Hagler, *J. Am. Chem. Soc.*, **116**, 2515 (1994).

4. J. R. Maple, M.-J. Hwang, T. P. Stockfisch, and A. T. Hagler, *Israel J. Chem.*, **34**, 195 (1994).
5. Z. Peng, C. S. Ewig, M.-J. Hwang, M. Waldman, and A. T. Hagler, *J. Phys. Chem. A*, **101**, 7243 (1997).
6. (a) S. J. Weiner, P. A. Kollman, D. T. Nguyen, and D. A. Case, *J. Comput. Chem.*, **7**, 230 (1986); (b) W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.*, **117**, 5179 (1995).
7. P. Dauber-Osguthorpe, V. A. Roberts, D. J. Osguthorpe, J. Wolff, M. Genest, and A. T. Hagler, *Proteins: Structure, Function and Genetics*, **4**, 31 (1988).
8. (a) L. Nilsson and M. Karplus, *J. Comput. Chem.*, **7**, 591 (1986); (b) S. Ha, J. Gao, B. Tidor, J. W. Brady, and M. Karplus, *J. Am. Chem. Soc.*, **113**, 1553 (1991).
9. B. P. Van Eijck, R. W. W. Hooft, and J. Kroon, *J. Phys. Chem.*, **97**, 12093 (1993).
10. J.-H. Lii and N. L. Allinger, *J. Comput. Chem.*, **12**, 186 (1991).
11. N. L. Allinger, K. Chen, and J.-H. Lii, *J. Comput. Chem.*, **17**, 642 (1996).
12. (a) W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, **110**, 1657 (1988); (b) W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, *J. Am. Chem. Soc.*, **118**, 11225 (1996).
13. A. T. Hagler, J. R. Maple, T. S. Thacher, G. B. Fitzgerald, and U. Dinur, *Theoretical and Experimental Applications*, ESCOM, Leiden, 1989, p. 149.
14. J. R. Maple, U. Dinur, and A. T. Hagler, *Proc. Natl. Acad. Sci. USA*, **85**, 5350 (1988).
15. A. T. Hagler and C. S. Ewig, *Comput. Phys. Commun.*, **84**, 131 (1994).
16. (a) T. A. Halgren, *J. Comput. Chem.*, **17**, 490 (1996); (b) T. A. Halgren, *J. Comput. Chem.*, **17**, 553 (1996).
17. F. A. Momany, V. J. Klimkowski, and L. J. Schäfer, *J. Comput. Chem.*, **11**, 654 (1990).
18. J. R. Maple and A. T. Hagler, manuscript in preparation.
19. A. Komornicki, GRADSCF, Polyatomics, Inc.
20. *TURBOMOLE User's Guide, Version 2.3*, Molecular Simulations Inc., San Diego, CA, 1996.
21. M. J. Frisch, M. Head-Gordon, G. W. Trucks, J. B. Foresman, H. B. Schlegel, K. Raghavachari, M. A. Robb, J. S. Binkley, C. Gonzalez, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, S. Topiol, and J. A. Pople, *Gaussian 90*; Gaussian, Inc., Pittsburgh, PA, 1990.
22. E. B. Wilson, J. C. Decius, and P. C. Cross, *Molecular Vibrations*, Dover, New York, 1980.
23. C. S. Ewig, T. S. Thacher, and A. T. Hagler, manuscript in preparation.
24. N. L. Allinger, Y. H. Yuh, and J.-H. Lii, *J. Am. Chem. Soc.*, **111**, 8551 (1989).
25. T. A. Halgren, *J. Am. Chem. Soc.*, **112**, 4710 (1990).
26. A. T. Hagler, P. S. Stern, S. Lifson, and S. Ariel, *J. Am. Chem. Soc.*, **101**, 813 (1979).
27. A. Warshel and S. Lifson, *Chem. Phys. Lett.*, **4**, 255 (1969).
28. S. Lifson and P. S. Stern, *J. Chem. Phys.*, **77**, 4542 (1982).
29. *Molecular Mechanics*, U. Burkert and N. L. Allinger, Eds., American Chemical Society, Washington, DC, 1982.
30. P. A. Karplus, *Prot. Sci.*, **5**, 1406 (1996).
31. X. Jiang, C.-H. Yu, M. Cao, S. Q. Newton, E. F. Paulus, and L. Schäfer, *J. Mol. Struct.*, **403**, 83 (1997).
32. E. Hirota, R. Sugisaki, C. J. Nielsen, and G. O. Sørensen, *J. Mol. Spectrosc.*, **49**, 251 (1974).
33. C. C. Costain and J. M. Dowling, *J. Chem. Phys.*, **32**, 150 (1960).
34. R. D. Brown, P. D. Godfrey, and B. Kleibomer, *J. Mol. Spectrosc.*, **124**, 34 (1987).
35. G. Fogarasi and P. G. Szalay, *J. Phys. Chem. A*, **101**, 1400 (1997).
36. P. Pulay, G. Fogarasi, F. Pang, and J. E. Boggs, *J. Am. Chem. Soc.*, **101**, 2550 (1979).
37. O. Ermer and S. Lifson, *J. Am. Chem. Soc.*, **95**, 4121 (1973).
38. O. Ermer, In *Structure and Bonding*, Vol. 27, J. D. Dunitz, Ed., Springer, Berlin, 1976, p. 161.
39. S. T. King, *J. Phys. Chem.*, **75**, 405 (1971).
40. K. B. Wiberg and C. M. Breneman, *J. Am. Chem. Soc.*, **114**, 831 (1992).
41. G. Fogarasi and A. Balazs, *J. Mol. Struct. (Theochem)*, **133**, 105 (1985).
42. C. L. Brummel, M. Shen, K. B. Hewett, and L. A. Philips, *J. Opt. Soc. Am. B*, **11**, 176 (1994).
43. M. Kitano and K. Kuchitsu, *Bull. Chem. Soc. Jpn.*, **47**, 631 (1974).
44. M. Kitano and K. Kuchitsu, *Bull. Chem. Soc. Jpn.*, **46**, 384 (1973).
45. M. Kitano and K. Kuchitsu, *Bull. Chem. Soc. Jpn.*, **46**, 3048 (1973).
46. H.-G. Mack and H. Oberhammer, *J. Am. Chem. Soc.*, **119**, 3567 (1997).
47. N. R. Carlsen, L. Radom, N. V. Riggs, and W. R. Rodwell, *J. Am. Chem. Soc.*, **101**, 2233 (1979).
48. P. L. Polavarapu, Z. Deng, and C. S. Ewig, *J. Phys. Chem.*, **98**, 9919 (1994).
49. Y. Sugawara, A. Y. Hirakawa, M. Tsuboi, S. Kato, and K. Morokuma, *Chem. Phys.*, **62**, 339 (1981).
50. N. G. Mirkin and S. Krimm, *J. Mol. Struct.*, **242**, 143 (1991).
51. Y. Sugawara, A. Y. Hirakawa, M. Tsuboi, S. Kato, and K. Morokuma, *J. Mol. Spectrosc.*, **115**, 21 (1986).
52. P. G. Jasien, W. J. Stevens, and M. Krauss, *J. Mol. Struct. (Theochem)*, **139**, 197 (1986).
53. F. A. L. Anet and M. Squillacote, *J. Am. Chem. Soc.*, **97**, 3243 (1975).
54. T. Drakenberg and S. Forsen, *Chem. Commun.*, **1404** (1971).
55. U. Dinur and A. T. Hagler, *J. Chem. Phys.*, **91**, 2949, 2959 (1989); U. Dinur, *J. Comput. Chem.*, **12**, 469 (1991).
56. J. Bajorath, J. Kraut, Z. Li, D. H. Kitson, and A. T. Hagler, *Proc. Natl. Acad. Sci. USA*, **88**, 6423 (1991).