# Accurate ab Initio Quantum Chemical Determination of the Relative Energetics of Peptide Conformations and Assessment of Empirical Force Fields

**Michael D. Beachy,[†] David Chasman,[†] Robert B. Murphy,[‡] Thomas A. Halgren,[§] and Richard A. Friesner*,[†]**

*Contribution from the Department of Chemistry and Center for Biomolecular Simulation, Columbia University, New York, New York 10027, Schrödinger Inc., 121 SW Morrison, Suite 1212, Portland, Oregon 97204, and Department of Molecular Design and Diversity, Merck Research Laboratories, Rahway, New Jersey 07065*

*Received July 8, 1996. Revised Manuscript Received April 14, 1997*[⊗]

**Abstract:** Correlated *ab initio* calculations have been carried out with a parallel version of the PSGVB electronic structure code to obtain relative energetics of a number of conformations of the alanine tetrapeptide. The highest level of theory utilized, local MP2 with the cc-pVTZ(−f) correlation-consistent basis set, has previously been shown to provide accurate conformational energies in comparison with experiment for a data set of small molecules. Comparisons with published and new canonical MP2 calculations on the alanine dipeptide are made. Results for ten gas-phase tetrapeptide conformations and a $\beta$-sheet dipeptide dimer are compared with 20 different molecular mechanics force field parametrizations, providing the first assessment of the reliability of these models for systems larger than a dipeptide. Comparisons are made with the LMP2/cc-pVTZ(−f) results, which are taken as a benchmark for the tetrapeptides. Statistical summaries with regard to energetics and structure are produced for each force field, and a discussion of qualitative successes and failures is provided. The results display both the successes and limitations of the force fields studied and can be used as benchmark data in the development of new and improved force fields. In particular, comparisons of hydrogen-bonding energetics as a function of geometry suggest that future force fields will need to employ a representation for electrostatics that goes beyond the use of atom-centered partial charges.

## 1. Introduction

The development of an accurate model for calculating the relative energetics of peptide conformations in the gas phase and in solution is a crucial objective for successful molecular modeling of biological systems. Peptides play an important role in many biological processes, and knowledge of their lowest energy conformations would be a major step forward in elucidating the modality of peptide−receptor binding and providing template structures for rational drug design. At present, however, reliable determination of peptide structure by purely computational means is beyond the capabilities of molecular modeling methodology. Peptide structures that have been determined have come from analysis of NMR or other experimental data, in many cases supplemented with theoretical modeling.

Currently, the standard approach to calculation of gas-phase peptide energetics is the use of molecular mechanics force fields, such as AMBER,[1,2] OPLS,[3] CHARMM,[4−6] MMFF,[7−11] and MM3.[12−14] These force fields are parametrized to fit small-molecule experimental data and, in some cases, quantum chemical calculations. However, it has been extremely difficult to demonstrate that such procedures create analytic potentials that are transferable to proteins and larger peptides. More generally, it has been difficult to elucidate the magnitude of the errors incurred when using these potentials to compute the energetics of larger and more complex structures.

Over the past 5 years, we have been developing novel quantum chemical methods that are capable of treating large systems at a high level of electron correlation. When implemented on a scalable parallel supercomputer such as the IBM SP2, a large number of peptide conformations of substantial size can be studied via correlated *ab initio* quantum chemical calculations. This new capability provides a way of investigating the accuracy and reliability of molecular modeling force fields. With these benchmark quantum chemical computations in hand, we are able to determine the accuracy of force fields in ranking compact, low-energy peptide structures. Our goal in the present paper is to demonstrate the feasibility of such investigations and to present preliminary data assessing the accuracy of a number of widely used force fields.

The calculations in this paper demonstrate a qualitative advance in the ability to carry out quantum chemical calculations for large biological systems while utilizing both a high level of electron correlation and an adequately large one-particle basis set. This capability will be essential in designing a new and

[†] Columbia University.
[‡] Schrödinger Inc.
[§] Merck Research Laboratories.
[⊗] Abstract published in *Advance ACS Abstracts,* June 1, 1997.

(1) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230−252.
(2) McDonald, D. Q.; Still, W. C. *Tetrahedron Lett.* **1992**, *33*, 7743.
(3) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(4) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, J.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(5) Mackerell, A. D., Jr.; Wiórkiewicz-Kuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.
(6) Momany, F. M.; Rone, R. J. *J. Comput. Chem.* **1992**, *13*, 888−900.
(7) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490−519.
(8) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 520−552.
(9) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 553−586.
(10) Halgren, T. A.; Nachbar, R. B. *J. Comput. Chem.* **1996**, *17*, 587−615.
(11) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 616−641.
(12) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551.
(13) Lii, J.-H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8566.
(14) Lii, J.-H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8576.

S0002-7863(96)02310-4 CCC: $14.00 

more accurate generation of force fields. For, while the performance of the existing force fields is in many respects encouraging, all of the force fields examined exhibit some serious discrepancies. A discussion of one way in which this objective might be accomplished is presented in the Conclusion. In any case, the benchmark results obtained here will be available as an evaluative criterion for all force field developers. Future calculations will incorporate amino acids other than alanine and will also be made available as benchmarks.

## 2. Computational Methods

The quantum mechanical methodology we have employed is straightforward. We have studied a total of ten conformations of the "alanine tetrapeptide", which consists of three central alanyl residues blocked on the N-terminal end by an acetyl group and on the C-terminal end by an *N*-methylamine group. Initial geometries were generated by carrying out a limited conformational search in the MacroModel[15] suite of programs using the AMBER* force field with the BatchMin 5.0 default dielectric of $\epsilon = 1.0r$. The five lowest energy conformations found on the AMBER* surface were chosen, along with five additional conformations of higher energy. These initial geometries were optimized via analytical gradient calculations at the Hartree−Fock level while using the 6-31G** basis. Geometry optimizations were carried out to the PS-GVB standard convergence criteria, comprised of the following: a maximum element of the gradient of less than $4.5 \times 10^{-4}$ hartree/bohr (0.53 kcal mol$^{-1}$ Å$^{-1}$), an RMS of gradient elements of less than $3.0 \times 10^{-4}$ hartree/bohr (0.36 kcal mol$^{-1}$ Å$^{-1}$), a change in energy of less than $5.0 \times 10^{-5}$ hartrees between final iterations, a maximum nuclear displacement of less than $1.8 \times 10^{-3}$ bohr between final iterations, and an RMS nuclear displacement of less than $1.2 \times 10^{-3}$ bohr between final iterations. LMP2 calculations were performed at these geometries with the cc-pVTZ(−f) basis. (Here and elsewhere, cc-pVTZ(−f) refers to the triple-$\zeta$ correlation-consistent basis of Dunning,[16] excluding f functions on second row elements and d functions on hydrogen.) The HF calculations also used the HF/6-31G** geometries. Calculations to confirm that these geometries are true minima on the HF/6-31G** surface are on-going.

Alanine dipeptide calculations were carried out at the LMP2/cc-pVTZ(−f) level while using HF/6-31G** minimized geometries. The dipeptide geometries were minimized via analytic gradient calculations to PS-GVB standard convergence.

The comparison of gas-phase energetics of molecular mechanical and quantum mechanical results requires care with regard to the geometries that are used. The distance at which strongly repulsive forces become activated varies slightly from method to method. Consequently, the molecular mechanics energy evaluated at a quantum chemical minimized geometry can be substantially more positive than is reasonable, due to "atom clashes". A fair comparison of relative energies can be made by optimizing the quantum chemical geometries in the force field being assessed. The relative energies of each force field at its own minima are therefore compared with the LMP2/cc-pVTZ(−f)//HF/6-31G** energies. For each force field, geometry optimization was carried out for the ten conformers starting from the quantum chemical geometry. To precisely locate the minima for the MacroModel, MMFF, and CHARMM force fields, optimization was carried out until the norm of the gradient fell below 0.001 kcal mol$^{-1}$ Å$^{-1}$. Similar protocols were cited by colleagues who contributed results for other force fields (see the Acknowledgment). Because the geometries differ from the HF/6-31G** geometries, we also compare the RMS deviation of the force field local minima from the HF/6-31G** geometries. Both values are measures of the quality of the force field.

A second comparison of force field relative energetics has been made by performing force field optimizations in which the $\phi$, $\psi$ torsional angles are restrained to the values of the HF/6-31G** minima. The torsional angles for the optimized force field geometries typically varied from the HF/6-31G** torsional angles by less than 0.10°. In cases where the location of the force field minimum is very different from the quantum chemical one, this procedure can lead to problems such as high sensitivity of the energy to the precise form of the coordinate restraints. Despite this caveat, the procedure has some qualitative value in assessing the nature of the errors in a force field when minima do not correspond to the quantum chemical ones. We must note that all MacroModel geometries were restrained to $\phi$, $\psi$ angles rounded to the nearest degree, due to the inability of BatchMin 5.0 to restrain torsional angles to non-integer values.

The present study also investigates the abilities of the force fields to reproduce hydrogen bonding energies in $\beta$-sheets. For this purpose, two alanine dipeptides were brought together in a representative antiparallel $\beta$-sheet geometry[17] with $\phi$ and $\psi$ torsional angles of −139° and 135°, respectively. This region of the HF/6-31G** potential dipeptide $\phi$, $\psi$ surface is relatively flat, with torsional gradients of around 0.25 kcal/mol per 10°. However, to minimize any nontorsional gradients, constrained HF/6-31G** geometry optimization was carried out with frozen torsional angles, and force field geometry optimizations were carried out with the angles tightly restrained to their *ab initio* values.[18]

**1. Quantum Chemical Methods.** We have used the PSGVB electronic structure code for all quantum chemical calculations reported here unless otherwise noted. Previous publications document the accuracy and efficiency of PSGVB. For geometry optimization of large structures, factors of 5−10 in speed as compared to Gaussian 92 are observed.[19,20] For local MP2 (LMP2) calculations with the cc-pVTZ(−f) basis set (which we use in all calculations reported here), a scaling of $N^{2.5}$ with basis set size $N$ is obtained, whereas Gaussian 92 achieves an $N^5$ scaling.[21] Based on timings for the MP2/cc-pVTZ alanine dipeptide calculations included herein and assuming the $N^5$ experimental scaling found in ref 21, we estimate that the computational cost of a single canonical MP2/cc-pVTZ(−f) alanine tetrapeptide calculation (658 basis functions) would be on the order of 24 cpu days on an IBM 580 workstation, assuming sufficient disk storage and memory. It is therefore unlikely that the suite of calculations reported here could have been performed with any other *ab initio* electronic structure code without extraordinary computational expenditures.

We have recently developed a parallel version[22,23] of PSGVB that has been optimized for the IBM SP2 parallel supercomputer. Our LMP2 parallelization effort had two major goals: (1) distribution of the two electron integrals in the molecular orbital basis over the local disks on the SP2, thus allowing truly large systems to be studied, and (2) maintenance of a ∼90% efficiency level as compared to single node performance for a calculation in which the number of nodes employed is sufficient to reduce the wall clock run time to ∼12 h. These objectives have been accomplished with our current parallel version. For example, the LMP2 alanine tetrapeptide calculations reported here require ∼6 h of time on 4 nodes of the SP2, and retain an efficiency of better than 95% in comparison with single-node calculations. Calculations have been performed on the SP2 at the Cornell Theory Center and on our own SP2 at Columbia University.

**1.a. Accuracy of the Local MP2 Method for Conformational Energetics.** In previous papers,[10,21,24] we have carried out extensive investigations of the accuracy of Hartree−Fock, local and canonical MP2, and DFT methods for the determination of experimentally known energy differences. When a large basis set such as cc-pVTZ(−f) is used, the local MP2 method provides the highest level of accuracy and reliability.[21] The use of any MP2 method with smaller basis sets leads to qualitative errors for some test cases. The RMS error for

(15) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440−467.

(16) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

(17) Dickerson, R. E. In *The Proteins*; Neurath, H., Ed.; Academic Press: New York, 1964; Vol. 2.

(18) The optimized force-field angles typically differed by less than 0.10° from these nominal values.

(19) Greeley, B. H.; Russo, T. V.; Mainz, D. T.; Friesner, R. A.; Langlois, J.-M.; Goddard, W. A., III; Donnelly, R. E., Jr.; Ringnalda, M. N. *J. Chem. Phys.* **1994**, *101*, 4028.

(20) Won, Y.; Lee, J.-G.; Ringnalda, M. N.; Friesner, R. A. *J. Chem. Phys.* **191**, *94*, 8152.

(21) Murphy, R. B.; Beachy, M. D.; Friesner, R. A.; Ringnalda, M. N. *J. Chem. Phys.* **1995**, *103*, 1481−1490.

(22) Chasman, D.; Beachy, M. D.; Wang, L.; Friesner, R. A. *J. Comput. Chem.* Submitted for publication.

(23) Beachy, M. D.; Chasman, D.; Friesner, R. A.; Murphy, R. B. *J. Comput. Chem.* Submitted for publication.

(24) St.-Amant, A.; Cornell, W. D.; Kollman, P. A.; Halgren, T. A. *J. Comput. Chem.* **1995**, *16*, 1483−1506.

LMP2/cc-pVTZ(−f) in comparison with experiment for the entire test suite is 0.42 kcal/mol with HF/6-31G** geometries and 0.36 kcal/mol with 6-31G* MP2 geometries. Only one case, methyl vinyl ether, exhibits a deviation from experiment greater than 1 kcal/mol. More recent work, using a multireference local MP2 method,[25] demonstrates that this can be attributed primarily to the low-lying $\pi$ state associated with the C=C double bond. When this case (which is not relevant to alanine polymers) is eliminated, the LMP2 results have a maximum deviation of 0.73 kcal/mol from experiment. It should be noted that many of the molecules in our test suite are of respectable size, the largest being only a few atoms smaller than the alanine dipeptide. Also, the test suite contains several molecules that possess the amide functional group, and hence are directly relevant to the present investigation.

While the best results for our test suite are obtained from the use of MP2 geometries, substitution of 6-31G** Hartree−Fock geometries increases the errors by no more than a few tenths of a kilocalorie per mole. Our view is therefore that HF geometry optimization is adequate for this preliminary study in which accuracy is being assessed at the level of 1.0 kcal/mol rather than 0.1 kcal/mol. This point is further addressed in our discussions of the alanine dipeptide (Section 3.1).

In this paper, we carry out only HF/6-31G** optimizations for the tetrapeptides. Single-point energies are then computed at the LMP2/cc-pVTZ(−f) level. For the time being, we shall take these LMP2 results as a benchmark against which other quantum chemical methods and force field results are to be compared. When we have parallelized our LMP2 gradient algorithm, we will perform tests with LMP2 geometries for the tetrapeptides. Convergence of the correlation energy will be checked in future work via a GVB-LMP2 approach that we are currently in the process of testing, and the effects of diffuse and f functions on relative energetics will be examined via the aug-cc-pVTZ basis set.

As noted above, the overall accuracy of this protocol has been demonstrated for a large database (36 molecules) of small-molecule conformational energy differences. The question of how well the observed errors translate to performance for larger, more flexible molecules such as peptides remains open. In Section 3.1, we attempt to provide a preliminary answer by studying the alanine dipeptide. However, a few general points should be noted. First, the LMP2 method in great part eliminates basis set superposition error (BSSE) in the correlation energy, as has been shown by Saebøand Pulay for the water dimer.[26] Thus, we expect that for larger systems (where one must compare compact and extended conformations), LMP2 results will be preferable to those obtained from canonical MP2 calculations. Second, we expect LMP2 calculations to provide a good estimation of long-range dispersion interactions, as these are typically dominated by the second-order energy. Our belief is that an analysis of the errors in the dipeptide will provide a reasonable estimation for those in the tetrapeptide. However, it should be noted that truly rigorous error estimation for either system will require using higher levels of electron correlation and larger basis sets, as has been suggested above.

**2. Molecular Mechanics Calculations.** We have undertaken a comprehensive study of protein molecular mechanics force fields, including most of those that are widely distributed in the molecular modeling community. In many cases we investigated a number of variants of each force field, in part to determine whether there has been improvement as a function of time. The presentation of these data in a single location should be of great value for those who wish to assess the relative performance of the different methodologies.

The following methods have been surveyed:

**(1) MacroModel Force Fields.** These methods include the force fields designated MM2*, MM3*, AMBER*, AMBER94, and OPLS*. These force fields are modified versions of the parent MM2,[27,28] MM3,[12−14,29] AMBER,[1,30] AMBER4.1,[31] and OPLS[3] force fields. The modifications in AMBER94, however, are minimal. The recent paper by Gundertofte et al.[32] further characterizes each of these force fields except for AMBER94, which has just been implemented in the BatchMin 5.5 module of the MacroModel program suite. In assessing MM2*, MM3*, and AMBER*, these authors employed the default BatchMin distance-dependent dielectric of $\epsilon = 1.0r$ in seeking to reproduce experimental values for small-molecule conformational energies, most of which were measured in nonpolar solution or in the gas phase. The developers of the BatchMin force fields recommend, however, that a constant dielectric ($\epsilon = 1.0$) be used in comparisons to calculated gas-phase conformational energies.[33] Indeed, this dielectric model has been made the default choice as of BatchMin 5.5. We report comparisons for both $\epsilon = 1.0$ and $1.0r$ for most of these force fields. For MM2* and MM3* we also include calculations using a constant dielectric of 1.5, the value used in the parent MM2 and MM3 force fields. All calculations were performed with BatchMin 5.0 except the AMBER94 calculations, which were performed with BatchMin 5.5.

**(2) CHARMM Force Fields.** We report results for three variants. The CHARMM 19 results refer to parametrization 19 of the polar-hydrogen force field,[4] which we use in the united atom approximation, and those labeled CHARMM 22 utilize the recently published all-atom force field.[5] Finally, the MSI CHARMm results employ the Momany and Rone force field[6] as subsequently updated and distributed with version 4.1 of QUANTA.[34] All calculations with CHARMM force fields were performed with a constant dielectric of 1.

**(3) AMBER Force Fields.** AMBER 3 refers to the original all-atom force field of Kollman and co-workers[1,30] as implemented in the IMPACT program suite.[35] Cross checks of some of the IMPACT results with the authentic AMBER program have confirmed the accuracy of the AMBER 3/IMPACT calculations.[36] AMBER 4.1 refers to the recently described reparametrization of Cornell et al.[31] as employed in AMBER 4.1. All AMBER calculations used a constant dielectric of 1 and were carried out by collaborators (see the Acknowledgment).

**(4) Merck Force Fields.** The calculations labeled MMFF used the recently described MMFF94 force field,[7−11] while those labeled MMFFs used the MMFF94s variant[37] in which modifications to out-of-plane force constants at trigonal delocalized nitrogen (with compensating adjustments to torsional parameters) are used to produce planar, or nearly planar, "quasi-experimental" energy-optimized geometries at amide and other resonance-delocalized nitrogens. MMFF94 typically yields puckered geometries that reflect the nonplanarity of the reference MP2/6-31G* *ab initio* structures. MMFF93 denotes an earlier version[10,32] of MMFF94, and MM2X[38] refers to a predecessor of MMFF that has been widely used in modeling applications at Merck.

**(5) OPLS Force Fields.** Three varieties of OPLS were considered. OPLS/A-UA(2,8) is the united atom force field first introduced some years ago,[3,39] while OPLS-AA(2,2) refers to the newer all-atom force field.[40] Finally, OPLS-UA(2,2) refers to a newer united atom force field.[41] All OPLS calculations used a constant dielectric of 1 and were carried out by one of the developers of OPLS.

**(6) Discover Force Fields.** We considered both CVFF, the original Discover force field,[42] and CFF95, a recent version of the force field being developed by the Biosym Consortium on potential energy

(25) Murphy, R. B.; Pollard, W. T.; Friesner, R. A. *J. Chem. Phys.* **1997**, *106*, 5073−5084.
(26) Saebø, S.; Tong, W.; Pulay, P. *J. Chem. Phys.* **1993**, *98*, 2170.
(27) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.
(28) Bukert, U.; Allinger, N. L. *Molecular Mechanics*; American Chemical Society: Washington, DC, 1982.
(29) Allinger, N. L.; Yan, L. *J. Am. Chem. Soc.* **1993**, *115*, 11918−11925.
(30) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765−784.
(31) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(32) Gundertofte, K.; Liljefors, T.; Norrby, P.-O.; Pettersson, I. *J. Comput. Chem.* **1996**, *17*, 429−449.
(33) W. C. Still, personal communication with T.A.H.
(34) QUANTA 4.1 is available from Molecular Simulations, Inc.: San Diego, CA.
(35) Kitchen, D. B.; Hirata, F.; Westbrook, J. D.; Levy, R. M.; Kofke, D.; Yarmush, M. *J. Comput. Chem.* **1990**, *11*, 1169-1180.
(36) P. A. Kollman, personal communication with T.A.H.
(37) T. A. Halgren, to be submitted for publication. The rationale behind the development of MMFF94s is discussed in refs 7 and 9.
(38) Holloway, M. K. *et al. J. Med. Chem.* **1995**, *38*, 3006−317.
(39) Jorgensen, W. L.; Laird, E. R.; Nguyen, T. B.; Tirado-Rives, J. *J. Comput. Chem.* **1993**, *14*, 206−215.
(40) This force field, which is still under development, uses torsion parameters fit to the *ab initio* data described in: Maxwell, D. S.; Tirado-Rives, J.; Jorgensen, W. L.; *J. Comput. Chem.* **1995**, *16*, 984−1010.
(41) J. Tirado-Rives, personal communication.

**Table 1.** Alanine Dipeptide Torsional Angles (deg)

| | HF/6-31G** present work | | HF/6-31G** Gould et al. | | MP2/6-31G* | |
|---|---|---|---|---|---|---|
| conf | $\phi$ | $\psi$ | $\phi$ | $\psi$ | $\phi$ | $\psi$ |
| $C7_{eq}$ | −85.8 | 78.5 | −85.8 | −79.0 | −83.1 | 77.8 |
| C5 | −157.9 | 160.3 | −157.2 | 159.8 | −158.4 | 161.3 |
| $\beta_2$ | −128.6 | 23.2 | −130.9 | 22.3 | −137.9 | 22.9 |
| $C7_{ax}$ | 75.8 | −56.5 | 76.0 | −55.4 | 74.4 | −64.2 |
| $\alpha_L$ | 66.9 | 29.7 | 67.0 | 30.2 | 63.5 | 34.8 |
| $\alpha'$ | −166.4 | −40.1 | | | −166.1 | −37.2 |

functions.[43,44] These calculations used a constant dielectric of 1 and were performed by the developers of CFF95.

**(7) GROMOS.** These calculations used the 1987 GROMOS force field,[45] in conjunction with a suggested "hydrophilic" modification for C···O interactions;[46] this modification has been applied to all nonbonded interactions between carbon and oxygen atoms. These calculations, contributed by a collaborator, used a constant dielectric of 1.

## 3. Results

**1. Alanine Dipeptide Gas-Phase Quantum Chemical Calculations.** Before investigating the tetrapeptide energetics, it is useful to examine the results obtained from cc-pVTZ(−f) LMP2 calculations for the alanine dipeptide (acetyl−Ala−$NHCH_3$ ). There are several reasons for presenting these calculations. First, we can compare our results with previous work in the literature to see whether they are altered by the use of a larger basis set. Since many of the molecular mechanics force fields have been parametrized against earlier dipeptide results by using smaller basis sets, such discrepancies may contribute to errors in the tetrapeptide force field energetics. We can also examine the effects of using an even larger basis than the cc-pVTZ(−f). Second, we can explore the effects of using localized MP2, as opposed to canonical MP2. For the small-molecule cases considered in ref 21, the differences in relative conformational energetics were typically small, less than 0.3 kcal/mol. Here, however, a larger, more flexible molecule offers greater possibilities for disagreement, especially when considering that intramolecular hydrogen bonds can be formed. We have therefore carried out canonical MP2 calculations of the dipeptide in the identical cc-pVTZ(−f) basis with Gaussian 94 for comparison. Third, we can examine the effects of using Hartree−Fock geometries instead of MP2 geometries. Fourth, we can examine the effects of using a higher level of correlation than second-order perturbation theory.

Torsional angles for HF/6-31G** alanine dipeptide geometries obtained with PS-GVB, MP2/6-31G* geometries obtained with Gaussian 94, and the HF/6-31G** geometries obtained by Gould et al.[47] are listed in Table 1. It can be seen that the change in geometry when going from HF/6-31G** to MP2/6-31G* is small; the maximum change in any torsional angle is less than 10°. For most conformations, there is excellent quantitative agreement between the HF/6-31G** results of this paper and those of Gould et al. and direct comparisons of the energetics are possible. However, we found that the $\alpha_R$ and $\beta$ geometries presented by Gould et al. are not minima on the PSGVB HF/6-31G** potential surface. We checked these results by carrying out geometry optimization using Gaussian 92 and obtained results identical with those of PSGVB. The discrepancy cannot be resolved without direct access to the data of Gould et al., so we do not pursue the issue further in this paper. In addition to the minima studied by Gould et al., we also include the $\alpha'$ minimum found by Head-Gordon et al. in their studies of the alanine dipeptide analog.[48]

**Table 2.** Alanine Dipeptide Relative Energies at HF/6-31G** Geometries (kcal/mol)

| | LMP2 | | MP2 | | |
|---|---|---|---|---|---|
| conf | cc-pVTZ(−f) | cc-pVTZ | cc-pVTZ(−f) | cc-pVTZ | TZP |
| $C7_{eq}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| C5 | 0.95 | 0.84 | 1.33 | 1.19 | 1.47 |
| $C7_{ax}$ | 2.67 | 2.63 | 2.64 | | 2.05 |
| $\beta_2$ | 2.75 | 2.53 | 3.14 | | 3.25 |
| $\alpha_L$ | 4.31 | 4.28 | 4.40 | | 4.42 |
| $\alpha'$ | 5.51 | 5.45 | 5.84 | | |

Table 2 presents the relative energies of the alanine dipeptide HF/6-31G** geometries for LMP2/cc-pVTZ(−f) and LMP2/cc-pVTZ calculations using PSGVB and canonical MP2/cc-pVTZ(−f) and MP2/cc-pVTZ calculations using Gaussian 94. Also presented are the canonical MP2 results of Gould et al. using the older TZP basis of Dunning.[49]

The easiest configurations to consider are the $\alpha_L$ and $C7_{ax}$. For the former, all methods are in near quantitative agreement. For the latter, the LMP2/cc-pVTZ(−f), LMP2/cc-pVTZ, and MP2/cc-pVTZ(−f) calculations agree quite well and differ significantly (0.6 kcal/mol) from the TZP results, so the difference in conformational energy relative to $C7_{eq}$ can be attributed to basis set differences. A number of discrepancies of this magnitude between these basis sets were discovered in ref 21 for small-molecule conformational energetics, and in all cases the cc-pVTZ(−f) results were nearer to experiment.

The C5 and $\beta_2$ relative energies display systematic disagreement between the MP2 and LMP2 methods, on the order of 0.5 kcal/mol, with basis set playing an apparently smaller role. We note that the C5 and $\beta_2$ conformations are considerably more extended than the $C7_{eq}$ and $C7_{ax}$ conformations, each of which has an internal hydrogen bond. This suggests that the canonical MP2 values are contaminated by basis set superposition error, which would artificially lower the energies of the compact structures compared to the more extended ones. In an intermolecular complex, the binding energy is enhanced due to increased monomeric correlation energy from basis functions on the other monomer, and the effect here is similar. The LMP2 results are expected to have the correlation part of the BSSE eliminated by construction. The value of 0.5 kcal/mol is typical of the size of the effect for a basis set of the quality of cc-pVTZ(−f) for hydrogen bonding calculations.[26] In support of this hypothesis, we note that the canonical MP2 energies move toward the LMP2 results as one goes from TZP to the better cc-pVTZ(−f) basis by 0.1 kcal/mol for both C5 and $\beta_2$. Internal

(42) Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T. *Proteins* **1988**, *4*, 31−47.
(43) Maple, J. R.; Hwang, M.-J.; Stockfish, T. P.; Dinur, U.; Waldman, M.; Ewig, C. S.; Hagler, A. T. *J. Comput. Chem.* **1994**, *15*, 161−182.
(44) Hwang, M.-J.; Stockfish, T. P.; Hagler, A. T. *J. Am. Chem. Soc.* **1994**, *116*, 2515−2525.
(45) van Gunsteren, W.; Berendsen, H. *Groningen Molecular Simulation Package and Manual (GROMOS);* Biomos: Groningen, The Netherlands, 1987.
(46) Van Buuren, A. R.; Marrink, S.-J.; Berendsen, H. J. C. *J. Phys. Chem.* **1993**, *97*, 9206−9212.
(47) Gould, I. R.; Cornell, W. D.; Hillier, I. H. *J. Am. Chem. Soc.* **1994**, *116*, 9250−9256.
(48) Head-Gordon, T.; Head-Gordon, M.; Frisch, M. J.; Brooks, C. L., III; Pople, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5989−5997.
(49) Dunning, T. H. *J. Chem. Phys.* **1971**, *55*, 716.

**Table 3.** Alanine Dipeptide Relative Energies at MP2/6-31G* Geometries (kcal/mol)

| | LMP2 | MP2 | | "MP4" | "MP4−BSSE" |
|---|---|---|---|---|---|
| conf | cc-pVTZ(−f) | cc-pVTZ(−f) | TZP | cc-pVTZ(−f) | cc-pVTZ(−f) |
| $C7_{eq}$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| C5 | 1.11 | 1.61 | 1.86 | 1.39 | 0.89 |
| $C7_{ax}$ | 2.48 | 2.41 | 2.13 | 2.48 | 2.55 |
| $\beta_2$ | 2.78 | 3.27 | 3.42 | 3.05 | 2.56 |
| $\alpha_L$ | 4.36 | 4.38 | 4.40 | 4.23 | 4.21 |
| $\alpha'$ | 5.49 | 5.94 | 5.67 | 5.62 | 5.17 |

BSSE diminishes as the basis set size is increased, though the convergence of BSSE with basis set size is notoriously slow. One further check of our hypothesis was made by carrying out canonical MP2 calculations with Gaussian 92 of the C5 and $C7_{eq}$ conformers in the even larger standard cc-pVTZ basis (i.e. f functions are included on first-row elements and d functions on hydrogen). The energy difference between C5 and $C7_{eq}$ was lowered to 1.19 kcal/mol, a value that is an additional 0.14 kcal/mol closer to the LMP2/cc-pVTZ(−f) results. The two MP2/cc-pVTZ calculations each used 468 basis functions and required a total of 186 cpu hours on an IBM 580 RS/6000 workstation. Due to the expense of these calculations, no other conformers were examined at this level. We have completed all LMP2/cc-pVTZ calculations, however, and note that the largest change in relative energy from the LMP2/cc-pVTZ(−f) results is 0.22 kcal/mol, suggesting that the cc-pVTZ(−f) basis is adequate for our desired accuracy of 1 kcal/mol.

Turning to results for the MP2/6-31G* geometries, we see that the trends found for the HF/6-31G** geometries hold. Table 3 presents LMP2 and MP2 results for the cc-pVTZ(−f) basis, along with MP2/TZP results. The $\alpha_L$ results are again in excellent agreement across the board, and the differences in $C7_{ax}$ results again seem to be mainly affected by an inadequacy of the TZP basis. The C5 and $\beta_2$ minima seem to be contaminated by internal BSSE again. The $\alpha'$ energies do not fall into the standard pattern where MP2/cc-pVTZ(−f) decreases the gap between MP2/TZP and LMP2/cc-pVTZ(−f), though it could be that the MP2/TZP energy for this conformation is also affected greatly by the inadequacy of the TZP basis.

Table 3 also presents "MP4/cc-pVTZ(−f)" results obtained by adding perturbative MP3 and MP4SDQ contributions to the MP2/cc-pVTZ(−f) energies. The MP3 and MP4SDQ results were calculated with a "6−31G#"[10] basis set that differs from 6-31G* in using different (more "correlation-consistent") polarization exponents (0.80 for H, 0.60 for C, 0.85 for N, and 1.20 for O). Another composite model, "MP4-BSSE/cc-pVTZ(−f)", presents the "MP4/cc-pVTZ(−f)" results with an additive BSSE correction obtained from the difference between the LMP2/cc-pVTZ(−f) and MP2/cc-pVTZ(−f) energies at the MP2/6-31G* geometries. We regard this composite model as our best estimation of the alanine dipeptide minima relative energies.

What is important to note is that for no pair of conformers are the relative energies at the LMP2/cc-pVTZ(−f)//HF/6-31G** level more than 0.35 kcal/mol in error when compared with the LMP2/cc-pVTZ(−f) energies at MP2/6-31G* geometries or more than 0.28 kcal/mol in error when compared with the "MP4-BSSE"/cc-pVTZ(−f) energies at MP2/6-31G* geometries. The excellent agreement between LMP2/cc-pVTZ(−f) energies at HF/6-31G** geometries and "MP4-BSSE"/cc-pVTZ(−f) relative energies at MP2/6-31G* geometries must be regarded as somewhat fortuitous, as the LMP2/cc-pVTZ(−f) energies at the better MP2/6-31G* geometries are in slightly worse agreement with the "MP4-BSSE"/cc-pVTZ(−f)//MP2/6-31G* results. This does not detract from the overall very good agreement between the three methods, however, and supports our claim that the HF geometries are not greatly inferior to MP2 geometries for conformational energetics and that the absolute accuracy of the LMP2/cc-pVTZ(−f)//HF/6-31G** relative energies is within 1 kcal/mol. This observation is confirmed by the work of Frey et al.,[50] in which the MP2/6-311G** energy difference between the $C7_{eq}$ and C5 conformers of the alanine dipeptide analog (*N*-formylalanylamide) changed by only 0.33 kcal/mol when going from HF/6-311G** geometries to MP2/6-311G** geometries.

**Table 4.** HF/6-31G** Alanine Tetrapeptide Torsional Angles (deg)[a]

| conf | $\phi_1$ | $\psi_1$ | $\phi_2$ | $\psi_2$ | $\phi_3$ | $\psi_3$ |
|---|---|---|---|---|---|---|
| 1 | −158.5 | 163.5 | −157.8 | 163.4 | −156.2 | 160.8 |
| 2 | −158.6 | 163.9 | −154.9 | 158.1 | −86.0 | 79.2 |
| 3 | −81.7 | 93.4 | 76.3 | −53.4 | −80.5 | 85.1 |
| 4 | −156.9 | 161.3 | −88.8 | 83.5 | −156.0 | 152.8 |
| 5 | −157.2 | 170.0 | −76.2 | −19.6 | −153.8 | 160.8 |
| 6 | −89.0 | 67.3 | 63.0 | 24.3 | −165.0 | 149.8 |
| 7 | 56.0 | −158.5 | −93.0 | 63.8 | −163.3 | −50.0 |
| 8 | 72.8 | −70.5 | −58.1 | 134.7 | 62.0 | 25.7 |
| 9 | 75.7 | −59.5 | 76.1 | −55.3 | 75.5 | −53.0 |
| 10 | 62.5 | 29.0 | 65.1 | 20.6 | 73.8 | −51.5 |

[a] Numbering begins at the N terminus.

**2. Alanine Tetrapeptide Structures.** The HF/6-31G** $\phi$, $\psi$ angles for the ten tetrapeptide conformers are listed in Table 4. In most cases, the $\phi$, $\psi$ pairs are in regions commonly found in protein structures.[51] We view these conformers as being representative of structures that might be found in proteins outside of $\beta$-sheet or α-helical regions. They do not, however, necessarily represent the lowest energy conformers on the *ab initio* surface, or even on the AMBER* surface used to generate initial structures. For example, optimization of an extensive set of 10 000 trial conformations generated via the distance-geometry technique *jg*[52] yielded 3 distinct conformers on the MMFF94 surface that were up to 2.3 kcal/mol lower in energy than Conformer 3, which was the lowest of the ten examined conformers on the LMP2/cc-pVTZ(−f) and MMFF94 surfaces. A number of higher energy structures were deliberately chosen, on the theory that it was desirable to test more regions of conformational space than those representing low-lying minima only. Higher energy conformers could be populated due to effects of solvation or being packed against another part of the protein. An accurate protein force field should robustly treat all of the structures considered here, and others as well. We do, however, intend to examine additional low-lying conformers in a subsequent study.

Figure 1 depicts the ten conformations, which range from relatively extended to quite compact, include several conformations with one to three internal hydrogen bonds, and populate various regions of the Ramachandran torsional angle map. Comparison of Tables 1 and 4 shows that seven conformers (5, 7, and 8 are exceptions) contain only recognizably dipeptide-minimum $\phi$, $\psi$ pairs and that all dipeptide minima except $\beta_2$ are represented. Several conformations contain $C7_{ax}$ fragments, however, a local geometry seldom found in protein structures. In addition, none of the tetrapeptides have a $\phi$, $\psi$ pair in the α-helical region, which is commonly found in proteins. We expect, in future work, to base model conformations more closely on patterns found in actual protein structures. Nevertheless, these conformations are sufficient to test whether key regions of the dipeptide map are accurately represented by the force field models.

(50) Frey, R. F.; Coffin, J.; Newton, S. Q.; Ramek, M.; Cheng, V. K. W.; Momany, F. A.; Schäfer, L. *J. Am. Chem. Soc.* **1992**, *114*, 5369.
(51) Moult, J.; James, M. N. G. *Proteins* **1986**, *1*, 146−163.
(52) *jg* (jiggle): a distance-geometry based program written by S. K. Kearsley at the Merck Research Laboratories.

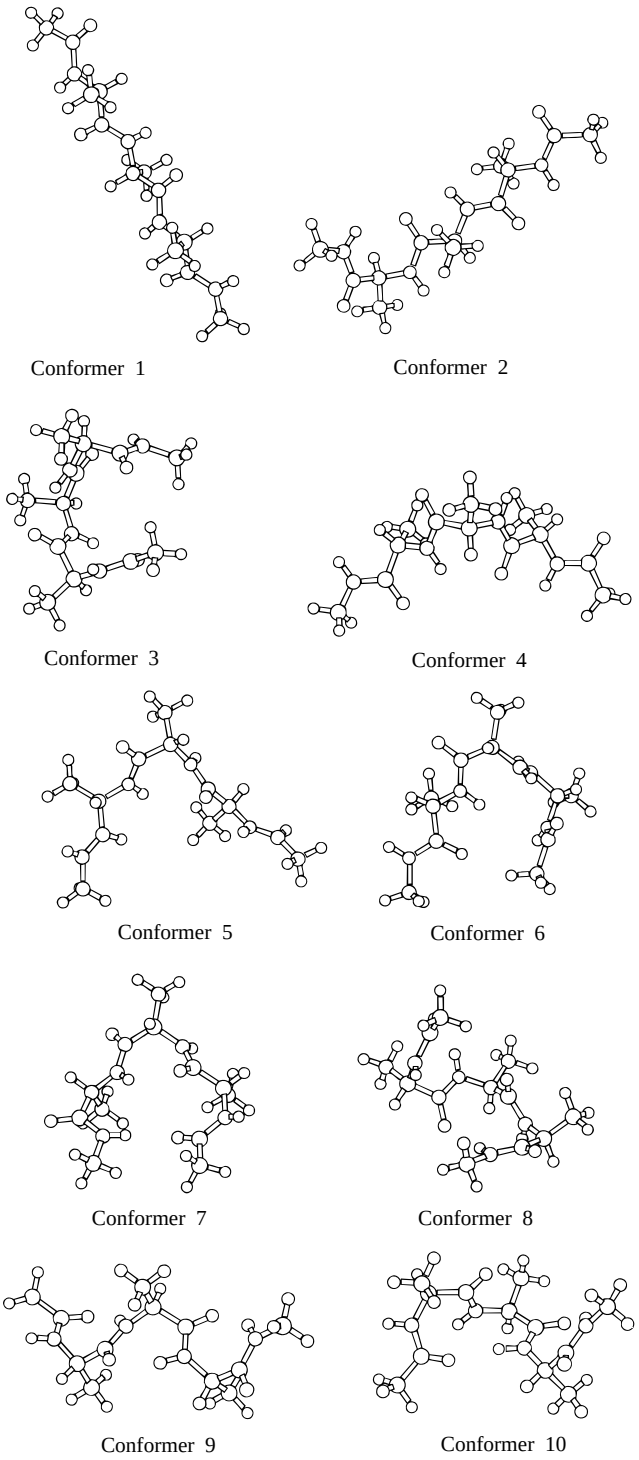



**Figure 1.** Alanine tetrapeptide Conformers 1−10.

Before turning to a statistical analysis of the performance of the force fields, we would like to briefly discuss some limitations of our tetrapeptide data set that potentially affect the conclusions that can be drawn about overall force field quality.

First, our data set is intended to examine regions of phase space relevant to small peptides and to the interior loop regions in proteins. This has led us to deemphasize the α-helical region, as α-helices are typically not stable for small peptides. Such structures have been automatically removed by the geometry optimization step in our conformational search protocol. Our goal is to test the ability of force fields to rank putative peptide and/or loop structures against each other, and not an investigation of the relative stability of the α-helix. The distribution of angles that have been used here is not unreasonable for these purposes. Future studies will take up the question of the performance of force fields in the α-helical region of phase space. It is quite possible that the rankings of the force fields by their ability to predict α-helical stability would be very different than those presented here. It is also possible that some force field developers have focused considerable effort on the optimization of α-helical stability and in the process sacrificed accuracy in prediction of other arguably less important relative energetics. Our results should therefore be taken to apply only to the restricted domain for which they have been designed.

Second, we would like to address the fact that some of our $\phi$, $\psi$ pairs are not commonly found in proteins. More generally, while most of our $\phi$, $\psi$ pairs do fall within $\beta$, $\lambda$, and α regions commonly observed in protein loops,[53] the overall distribution of torsion angles in our data set does not closely match that found in nature. It is not clear, however, that an accurate reproduction of the observed distribution would lead to a better evaluation of force field performance than the actual distribution used here. For example, suppose that a molecular mechanics force field exhibited a very low relative energy for the $C7_{ax}$ region, in contradiction to the quantum chemical data and the observed infrequent occurrence of this conformation in protein structures. Such an occurrence would represent a serious flaw in the force field, but would not be found if the $C7_{ax}$ region were to be excluded from our data set. In addition, we consider the overall structures of the tetrapeptide conformations to be of equal importance to the local torsional angle values in our studies, as variation in spatial extent examines the general performance of the nonbonded part of the potential, and not just the specific dipeptide torsional parametrization. The structures shown in the figures clearly indicate that a wide range of conformations, varying between extended and compact, have been studied.

The third question is the most serious one: whether our sample is large enough to draw conclusions concerning the agreement of the force fields and quantum chemical data. Our view is that ten tetrapeptide conformations, representing 30 $\phi$, $\psi$ pairs, does represent a data set that is large enough for gross differences in performance to be meaningful. For example, it seems unlikely that the consistent ability of variants of the OPLS and MMFF force fields to reproduce all ten quantum chemical structural minima to better than 0.5 Å RMS deviation is accidental. However, small differences between force field RMS energy errors are probably not particularly significant. There is no question that a larger data set would improve upon the current one, however, and that the results reported here should therefore be regarded as tentative.

The summary in Table 5 includes average RMS errors (the full table is included in the Supporting Information) of the force-field minimized geometries from the HF/6-31G** minima, for consistency across the force fields based on heavy atoms and polar hydrogens only. The best performance is obtained by three OPLS force fields and by MMFF and MMFFs. The accuracy of the contributed OPLS-AA(2,2) and OPLS/A-UA-(2,8) structures, in particular, is quite remarkable. AMBER 3 also performs well in a RMS sense, but one conformer slightly exceeds the somewhat arbitrary RMS cutoff of 0.6 Å we have used in Table 6 to highlight poor comparison to the reference ab initio geometry. MSI CHARMm, GROMOS, CFF95, and MM3* display a small number of large errors, but perform reasonably well on average. On the other hand, AMBER*($\epsilon = 1.0r$), AMBER94, CHARMM 19 and 22, CVFF, OPLS*, and MM2* all have at least four conformations whose optimized geometries differ significantly from the HF/6-31G** geometries used to initiate the optimizations. In many cases in which large

(53) Kwasigroch, J.-M.; Chomilier, J.; Mornon, J.-P. *J. Mol. Biol.* **1996**, *259*, 855.

**Table 5.** Summary Table for Force Field Comparisons to LMP2/cc-pVTZ(−f)//HF/6-31G**[a]

| | av geometry RMS | geometry RMS >0.60 | energy RMS unrestrained minima | energy RMS restrained minima | pairwise errors >3 | max pairwise error |
|---|---|---|---|---|---|---|
| LMP2/cc-pVTZ(−f) | | | 0.00 | 0.00 | 0 | 0.00 |
| HF/6-31G** | 0.00 | 0 | 1.10 | 1.10 | 1 | 3.08 |
| MMFF93 | | | 1.20 | | 3 | 3.39 |
| MMFF | 0.32 | 0 | 1.24 | 1.21 | 5 | 3.35 |
| OPLS-AA(2,2) | 0.16 | 0 | 1.31 | | 6 | 4.72 |
| MMFFs | 0.24 | 0 | 1.40 | 1.59 | 8 | 4.31 |
| OPLS/A-UA(2,8) | 0.18 | 0 | 1.43 | | 8 | 4.81 |
| MM2X($\epsilon = 1.5$) | 0.51 | 3 | 1.49 | 1.38 | 9 | 5.44 |
| MM3* | 0.48 | 2 | 1.53 | 1.78 | 9 | 5.54 |
| OPLS* | 0.49 | 4 | 1.55 | 1.82 | 7 | 5.47 |
| MM3*($\epsilon = 1.5$) | 0.45 | 3 | 1.58 | 1.44 | 9 | 5.96 |
| GROMOS | 0.39 | 1 | 1.60 | 2.42 | 10 | 5.67 |
| HF/cc-pVTZ(−f) | | | 1.69 | 1.69 | 19 | 4.63 |
| MMFF($\epsilon = 1.5r$) | 0.27 | 1 | 1.75 | 2.00 | 13 | 5.54 |
| CFF95 | 0.41 | 3 | 1.86 | 1.91 | 14 | 6.48 |
| MM3*($\epsilon = 1.0r$) | 0.48 | 2 | 2.00 | 1.85 | 12 | 7.26 |
| Null Hypothesis | | | 2.07 | 2.07 | 16 | 6.99 |
| OPLS-UA(2,2) | 0.26 | 0 | 2.19 | | 18 | 6.55 |
| AMBER* | 0.48 | 4 | 2.39 | 2.55 | 21 | 7.35 |
| MSI CHARMm | 0.40 | 2 | 2.54 | 2.12 | 21 | 7.48 |
| MMFF($\epsilon = 2.0r$) | 0.29 | 1 | 2.56 | 2.77 | 23 | 7.49 |
| CHARMM 22 | 0.86 | 5 | 2.56 | 3.78 | 13 | 10.20 |
| CHARMM 19 | 0.76 | 5 | 2.73 | 2.05 | 24 | 8.63 |
| AMBER*($\epsilon = 1.0r$) | 0.47 | 2 | 2.75 | 2.88 | 23 | 7.97 |
| AMBER 4.1 | 0.55 | 3 | 3.35 | | 19 | 12.24 |
| AMBER94 | 0.58 | 4 | 3.42 | 1.42 | 20 | 12.53 |
| CVFF | 0.98 | 6 | 3.91 | | 27 | 13.99 |
| AMBER 3 | 0.35 | 1 | 4.17 | | 28 | 12.17 |
| MM2*($\epsilon = 1.5$) | 0.96 | 8 | 4.91 | 4.48 | 34 | 15.47 |
| MM2* | 0.80 | 7 | 6.09 | 5.64 | 37 | 19.54 |
| MM2*($\epsilon = 1.0r$) | 0.83 | 7 | 6.14 | 6.03 | 35 | 19.51 |

[a] The null hypothesis is the case for which all conformers are equal in energy. All quantum mechanical results used HF/6-31G** geometries. All energy units are kcal/mol, all length units are Å. Unless otherwise specified, $\epsilon = 1.0$.

departures are found, the force-field optimization arrived at a common equilibrium geometry from two (and for CVFF, from three) different initial HF/6-31G** geometries. This is signaled in Table 6 by those instances in which two or more "conformers" display identical conformational energies, and occurs especially frequently for Conformers 4 and 5. In order to clarify the physical meaning of the RMS deviations, Figure 2 presents examples of a low RMSD (0.28 Å) and high RMSD (1.62 Å) structure obtained with the MM2* force field for Conformers 6 and 5, respectively. It can be seen that the high RMSD structure is qualitatively in error for this conformation.

**3. Alanine Tetrapeptide Energetics.** In evaluating the quality of force field energetics, there is a question as to whether structures with high RMS deviation from a given quantum chemical structure should be included. On one hand, it could be argued that agreement with the location of quantum chemical minima is a different issue than the accuracy of the energies for the cases where the minima agree to a given tolerance. On the other hand, the removal from the data set of numerous cases with large RMS deviations could be viewed as artificially biasing the results in favor of force fields with a large number of erroneous geometries. Thus, we have examined two cases. The first comparison is made in Table 6, between Hartree−Fock and LMP2 energies at HF/6-31G** minima and force field energies at force field minima. Bold quantities denote that the force field geometry for that conformer differs from the HF/6-31G** geometry by an RMS deviation of 0.6 Å or more. (We have highlighted all cases with an RMS deviation of greater than 0.6 Å in order to indicate where energy deviations may be primarily due to differences in the locations of force field and quantum mechanical minima, as opposed to serious errors in the energy function. Differentiation between these two effects was of concern to a number of force field developers.[54]) The second comparison is made in Table 7 for force field structures that, as previously described, have been restrained to have nearly the same $\phi$, $\psi$ torsional angles as the HF/6-31G** minima listed in Table 4. In order to obtain an unbiased measure of average performance for both unrestrained and restrained force field minima, we have added a constant to the zero of energy for each method to minimize its RMS deviation from the LMP2 relative energies. Columns three and four of Table 5 present the minimized RMS deviations of the quantum chemical and force field results from the LMP2 relative energies, which are taken as a benchmark. Note that the quantum mechanical results are the same in the unrestrained and restrained geometry cases because they use the HF/6-31G** geometries.

The use of restrained geometries leads to significant improvement of the results for two force fields, AMBER94 and CHARMM 19, suggesting a specific problem for these force fields in their prediction of geometries. Note, however, that poor correlation of the unrestrained force field geometry with the corresponding HF minimum may mean that the restrained force field geometry is in a region of the potential surface with high gradients. The energy of structures in such regions can be sensitive to small displacements in geometry.

Interestingly, two force fields—GROMOS and CHARMM 22—reproduce the LMP2 relative energies significantly more poorly at the restrained minima. This finding indicates that these energy models are not as robust as the comparisons based on unrestrained minima might have suggested.

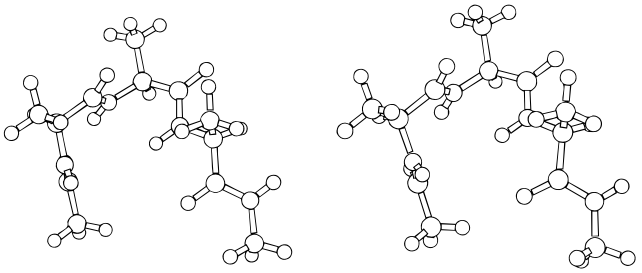
To give a reference point for calibration of the energy RMS error, we also include in Table 5 a model in which all

(54) Personal communication with Carl Ewig (CFF95), Peter Kollman (AMBER), and Martin Karplus (CHARMM).

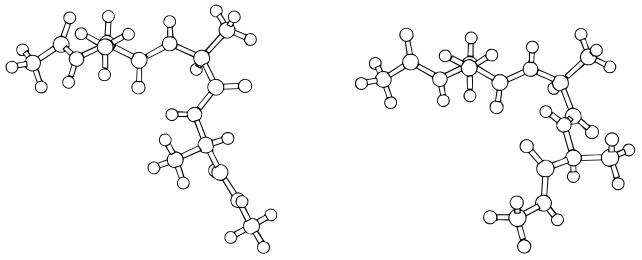
**Table 6.** Relative Conformational Energies for Unrestrained Alanine Tetrapeptide Conformations 1−10[a]

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| LMP2/cc-pVTZ(−f) | 2.71 | 2.84 | 0.00 | 4.13 | 3.88 | 2.20 | 5.77 | 4.16 | 6.92 | 6.99 |
| AMBER* | −0.67 | 0.79 | 2.33 | **2.04** | **2.04** | 1.69 | **4.73** | 8.13 | 8.76 | **9.77** |
| AMBER*($\epsilon = 1.0r$) | −1.41 | 0.54 | 3.84 | 1.26 | **1.26** | **2.15** | 6.11 | 7.93 | 8.14 | 9.77 |
| AMBER 3 | 9.56 | 6.12 | −2.94 | 7.51 | 8.09 | 4.19 | 1.83 | 1.54 | 1.60 | **2.14** |
| AMBER 4.1 | 6.36 | 5.91 | 0.07 | 6.90 | 4.91 | **1.29** | **−2.83** | 5.75 | 5.61 | **5.66** |
| AMBER94 | 6.40 | 5.95 | 0.08 | 6.93 | 4.95 | **1.30** | **−3.08** | **5.75** | 5.63 | **5.69** |
| CFF95 | 5.70 | 4.72 | −0.98 | 5.04 | **5.04** | **−0.11** | 5.49 | 5.05 | 3.43 | **6.22** |
| CVFF | **−0.74** | **−0.74** | 1.20 | **−0.74** | **1.15** | **1.20** | 7.23 | **5.27** | 9.72 | 16.10 |
| CHARMM 19 | 6.33 | **0.58** | −1.12 | **0.58** | **3.63** | **−1.10** | 6.78 | 6.08 | **5.80** | 12.06 |
| CHARMM 22 | 3.83 | 3.80 | 1.65 | **3.83** | 4.89 | **1.93** | **−1.26** | **7.34** | 6.79 | **6.79** |
| MSI CHARMm | 6.97 | 4.91 | −3.08 | 6.06 | 6.65 | 2.35 | 2.55 | 3.13 | **3.91** | **6.14** |
| GROMOS | 3.41 | 2.74 | −0.85 | 2.77 | 4.72 | **−0.33** | 6.98 | 4.94 | 5.08 | 10.13 |
| HF/6-31G** | 0.94 | 1.53 | 0.71 | 3.09 | 3.12 | 1.95 | 7.07 | 5.05 | 7.94 | 8.20 |
| HF/cc-pVTZ(−f) | 0.09 | 0.91 | 1.46 | 2.37 | 2.24 | 2.42 | 7.29 | 5.41 | 8.40 | 9.00 |
| MM2* | **13.16** | **8.00** | **−3.32** | **8.90** | **8.90** | 4.20 | **−3.33** | **2.26** | 0.37 | 0.45 |
| MM2*($\epsilon = 1.0r$) | **13.53** | **8.21** | **−2.50** | **8.44** | **8.44** | 4.85 | **−2.93** | **1.38** | 0.64 | −0.46 |
| MM2*($\epsilon = 1.5$) | **10.75** | **7.07** | **−0.61** | **7.65** | **7.65** | 4.29 | **−1.66** | **1.03** | 3.56 | **−0.14** |
| MM3* | 3.18 | 3.14 | 1.40 | 3.58 | **3.58** | 2.18 | 8.09 | 5.50 | 5.17 | **3.77** |
| MM3*($\epsilon = 1.0r$) | 4.08 | 3.75 | 2.18 | 3.84 | **3.84** | 2.88 | 7.19 | 4.89 | 5.07 | **1.90** |
| MM3*($\epsilon = 1.5$) | 2.35 | 2.74 | 2.82 | 2.86 | **2.89** | 2.33 | 7.49 | 5.41 | **6.87** | **3.84** |
| MM2X($\epsilon = 1.5$) | 5.24 | 3.82 | −0.01 | 4.34 | **4.34** | 2.93 | 3.69 | **5.03** | 6.13 | **4.08** |
| MMFF | 4.05 | 3.58 | −1.14 | 5.12 | 5.93 | 1.09 | 4.47 | 4.98 | 5.63 | 5.88 |
| MMFF($\epsilon = 1.5r$) | 0.91 | 1.89 | 3.61 | 2.20 | 2.95 | 2.54 | 4.11 | **5.69** | 8.81 | 6.87 |
| MMFF($\epsilon = 2.0r$) | −0.20 | 1.22 | 4.57 | 1.24 | 1.89 | 2.38 | 4.09 | **6.09** | 10.65 | 7.68 |
| MMFFs | 4.28 | 3.82 | −0.62 | 5.35 | 5.66 | 0.89 | 5.86 | 4.68 | 4.40 | 5.28 |
| MMFF93 | 4.08 | 3.44 | −1.42 | 5.11 | 5.41 | 1.45 | 3.91 | 5.35 | 5.95 | 6.28 |
| OPLS* | 1.55 | 2.87 | **−0.13** | **3.89** | **3.19** | 1.70 | 2.60 | 6.23 | 9.22 | **8.49** |
| OPLS-AA(2,2) | 2.78 | 2.50 | −1.34 | 3.48 | 4.46 | 3.31 | 3.82 | 6.93 | 5.90 | 7.78 |
| OPLS/A-UA(2,8) | 2.43 | 2.56 | −2.08 | 3.79 | 4.54 | 3.10 | 4.43 | 5.82 | 5.30 | 9.71 |
| OPLS-UA(2,2) | 5.35 | 4.20 | −2.54 | 7.39 | 5.85 | 2.04 | 2.48 | 4.56 | 4.54 | 5.69 |

[a] Zero of energy set at the value which minimizes the RMS deviation from LMP2/cc-pVTZ(−f) energies for all conformations. Bold quantities indicate the structures for these conformations have an RMSD of 0.6 Å or more from the HF/6-31G** geometries (except for MMFF93, where structural information is unavailable). All quantum mechanical results used HF/6-31G** geometries. All energies in kcal/mol. Unless otherwise specified, $\epsilon = 1.0$.



MM2* Conformer 6 6-31G** Conformation 6



MM2* Conformation 5 6-31G** Conformation 5

**Figure 2.** Alanine tetrapeptide MM2* structures. Conformation 6 has an 0.28 Å RMS deviation from the 6-31G** structure, and Conformation 5 has a 1.62 Å RMS deviation from the 6-31G** structure.

conformations have the same energy, the value of which is again adjusted to minimize the RMS deviation. We refer to this reference point (with an energy RMS error of 2.07 kcal/mol) as the null hypothesis. Arguably, an RMS error significantly above this value suggests that the predictive capability of the force field in question is not better than random with regard to quantitative energetic accuracy. Approximately half the force field models examined fail to better the null hypothesis value.

The results of comparisons made with LMP2/cc-pVTZ(−f) energies on a pairwise basis are displayed in columns five and six of Table 5. For each method, we list the largest absolute deviation from the LMP2 conformational energy difference for any pair of tetrapeptide conformations and tabulate the number of such errors that exceed 3 kcal/mol, in each case using conformational energies for unrestrained minima. These data allow an assessment of how likely a method is to make large errors, a somewhat different issue than that addressed by the RMS deviation.

**3.a. Quantum Chemical Energetics.** The HF/6-31G** energies display an RMS deviation of 1.10 kcal/mol from the LMP2 results, a number smaller than any force field examined. Furthermore, the maximum pairwise error for the HF/6-31G** method is 3.08 kcal/mol, a number better than the best displayed by a force field—MMFF, for which the maximum error is 3.34 kcal/mol. The results with the larger cc-pVTZ(−f) basis are not as good, with a maximum pairwise error of 4.63 kcal/mol for the Hartree−Fock results and an RMS error of 1.69 kcal/mol. However, even when considering the larger maximum error of the cc-pVTZ(−f) basis, only the MMFF force fields do as well as Hartree−Fock in not committing any grave errors in pairwise relative energies.

Despite the generally good quantum chemical results, a careful study reveals systematic discrepancies in the HF computations. Examination of Conformers 1−10 reveals that Conformers 1 and 2 are considerably more extended than the remaining structures, and that Conformers 4 and 5 are intermediate between the extended and compact structures. It is striking that the HF calculations predict systematically lower energies for the extended and partially extended structures (with the effect being significantly larger for the extended structures) as compared to the LMP2 results (see Table 6). A reasonable explanation of these results (although one that we are not able

**Table 7.** Relative Conformational Energies for Restrained $\phi$, $\psi$ Alanine Tetrapeptide Conformations 1−10[a]

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| LMP2/cc-pVTZ(−f) | 2.71 | 2.84 | 0.00 | 4.13 | 3.88 | 2.20 | 5.77 | 4.16 | 6.92 | 6.99 |
| AMBER* | −1.38 | −0.02 | 1.43 | 1.36 | 3.64 | 2.12 | 6.27 | 8.48 | 7.34 | 10.36 |
| AMBER*($\epsilon = 1.0r$) | −2.08 | 0.01 | 3.45 | 0.81 | 3.24 | 2.73 | 6.27 | 8.88 | 6.86 | 9.43 |
| AMBER94 | 4.07 | 3.86 | −1.10 | 5.44 | 2.31 | 3.47 | 4.26 | 5.56 | 4.48 | 7.26 |
| CFF95 | 4.47 | 3.91 | −1.10 | 4.10 | 6.30 | 3.68 | 4.96 | 4.60 | 2.26 | 6.46 |
| CHARMM 19 | 3.27 | 2.21 | −3.62 | 3.07 | 4.62 | 2.74 | 4.41 | 4.97 | 6.14 | 11.78 |
| CHARMM 22 | 0.69 | 0.51 | −1.74 | 1.81 | 1.32 | 5.29 | 2.78 | 9.34 | 4.46 | 15.15 |
| MSI CHARMm | 6.04 | 4.28 | −3.14 | 5.79 | 5.93 | 2.41 | 2.61 | 2.88 | 5.17 | 7.63 |
| GROMOS | 3.45 | 3.05 | −3.80 | 2.96 | 4.02 | 3.68 | 4.02 | 5.39 | 4.47 | 12.40 |
| MM2* | 11.54 | 7.55 | −1.99 | 8.94 | 8.95 | 2.53 | 2.90 | 2.86 | −2.41 | −1.27 |
| MM2*($\epsilon = 1.0r$) | 12.02 | 8.05 | −0.44 | 9.20 | 8.54 | 3.02 | 2.17 | 2.36 | −2.45 | −2.87 |
| MM2*($\epsilon = 1.5$) | 10.09 | 6.84 | −0.29 | 7.77 | 6.97 | 2.20 | 2.71 | 3.32 | 0.43 | −0.43 |
| MM3* | 2.72 | 2.40 | 0.94 | 2.77 | 5.02 | 1.82 | 7.08 | 7.67 | 3.52 | 5.65 |
| MM3*($\epsilon = 1.0r$) | 3.59 | 3.24 | 2.06 | 3.37 | 4.76 | 2.34 | 6.59 | 6.39 | 3.68 | 3.58 |
| MM3*($\epsilon = 1.5$) | 2.42 | 2.29 | 2.14 | 2.24 | 3.77 | 1.68 | 6.70 | 6.95 | 5.77 | 5.63 |
| MM2X($\epsilon = 1.5$) | 4.72 | 3.28 | 0.30 | 3.99 | 5.58 | 2.63 | 3.20 | 4.90 | 5.10 | 5.90 |
| MMFF | 3.65 | 3.14 | −0.98 | 5.03 | 5.65 | 1.68 | 4.96 | 5.77 | 5.09 | 5.61 |
| MMFF($\epsilon = 1.5r$) | 0.44 | 1.37 | 3.43 | 2.13 | 2.81 | 2.60 | 4.54 | 7.70 | 8.06 | 6.53 |
| MMFF($\epsilon = 2.0r$) | −0.77 | 0.61 | 4.26 | 1.01 | 1.75 | 2.45 | 4.62 | 8.22 | 9.93 | 7.52 |
| MMFFs | 3.57 | 3.21 | −0.51 | 4.80 | 5.33 | 0.99 | 5.86 | 7.11 | 4.05 | 5.19 |
| OPLS* | 0.32 | 1.71 | 0.01 | 2.84 | 4.19 | 1.56 | 3.80 | 7.96 | 8.49 | 8.72 |

[a] Zero of energy set at the value which minimizes the RMS deviation from LMP2/cc-pVTZ(−f) energies. All energies in kcal/mol. Unless otherwise specified, $\epsilon = 1.0$.

to prove rigorously here) is that more compact structures have greater dispersion energy. Therefore, methods such as HF that do not include dispersion energy will systematically predict relative energies that are too low for extended structures (or, equivalently, relative energies that are not sufficiently low for compact structures). Verification of this hypothesis will require a much larger and more structurally varied data set than we have examined here.

**3.b. Force Field Energetics.** The most successful force fields in structure prediction are the OPLS, MMFF, and AMBER 3 force fields, as indicated above. In relative energy prediction, MMFF93, MMFF, OPLS-AA(2,2), MMFFs, and OPLS/A-UA-(2,8) are the top five force fields in RMS energy error for the unrestrained minima (Table 5). In addition, MMFF, MMFF93, and MMFFs display the lowest maximum errors for the force field models tested and are comparable with the HF results. OPLS-AA(2,2) and OPLS/A-UA(2,8) do not do as well in minimizing the maximum error.

A sizable group of force fields are near this group of five in overall performance, including MM2X($\epsilon = 1.5$), MM3*, OPLS*, MM3*($\epsilon = 1.5$), GROMOS, MMFF($\epsilon = 1.5r$), CFF95, and MM3*($\epsilon = 1.0r$). All of these force fields better the null hypothesis for both unrestrained and restrained minimizations. None, however, have quite the accuracy in geometry prediction of the first five, and all are worse off in maximum error and the number of errors greater than 3 kcal/mol.

Most of the remaining force fields are clearly inferior in geometric and/or energetic performance, with RMS errors for the unrestrained optimization data set that are significantly larger than that of the null hypothesis. For example, AMBER* inverts the relative energies of the first few conformations for both the constant and distance-dependent dielectric models. AMBER 3, while being a good performer in structure prediction, is one of the worst performers in minimizing the RMS deviation from the LMP2/cc-pVTZ(−f) energies. MSI CHARMm also predicts structures well, but performs in the middle of the pack matching relative energies with LMP2/cc-pVTZ(−f). In contrast, CHARMM 19 and 22 show relatively high RMS deviations both for geometries and for energies at unrestrained minima, as does CVFF. Finally, AMBER94 (AMBER 4.1 is similar) has some problems with geometry, but reproduces the relative conformational energies fairly well at the restrained minima.

The influence of the dielectric model is also evident in Table 5. For example, the RMS energy error for MMFF increases substantially on going from a constant dielectric of $\epsilon = 1$ (for which the force field was parametrized) to distance-dependent dielectrics of $\epsilon = 1.5r$ and $2.0r$. The RMS errors for MM3* and MM2* also show a dependence on the dielectric used. We note that in each case, the restrained minima energy RMS is the lowest when using the dielectric for which the force field was parametrized. Interestingly, for none of these force fields are the RMS geometry comparisons greatly affected. As intermolecular interaction energies also strongly depend on the dielectric model (see the following section), changes to the dielectric model clearly should be used with caution in applications that rely on energetic comparisons.

**4. $\beta$-Sheet Interaction Energetics.** The $\beta$-sheet dimer (Figure 3) has two intermolecular carbonyl−amide hydrogen bonds. While $\beta$-sheet structures in actual proteins may not conform precisely to the geometry we have used, the accurate prediction of interaction energetics at this geometry seems a necessary condition for accurate energetic calculation of closely related structures. A quantitative understanding of the energetics of such structures will require the examination of a hypersurface of considerable size, which is beyond the scope of what we hope to accomplish. Our more modest goal is to compare quantum chemical and force field energetics for one point on the hypersurface, which is at the very least a fair representation of the ensemble of relevant structures.

Energetic comparisons are made in Table 8, which reports the negative interaction energy of the complex, defined as the complex energy minus the sum of the energies of the separated fragments. When calculating the energy of the fragments, they were kept frozen at the geometry that they exhibited in the complex. For the HF/cc-pVTZ(−f) and LMP2/cc-pVTZ(−f) calculations, the HF results are corrected for BSSE by the counterpoise method. The LMP2 correlation energy is *not* counterpoise corrected because most of the BSSE is removed through the use of local virtual spaces. Counterpoise and non-counterpoise corrected HF/6-31G**, BLYP/6-31G**, B3LYP/6-31G**, and B3LYP/cc-pVTZ(−f) results are also presented.

The $\beta$-sheet test examines a different aspect of the force fields—namely, their accuracy for intermolecular interactions, which are critical in both protein structure and ligand binding applications. In contrast to the conformational comparisons, however, it is less obvious that the force field results should reproduce gas-phase quantum chemical binding energies. If a polarizable functional form for the force field was being used,

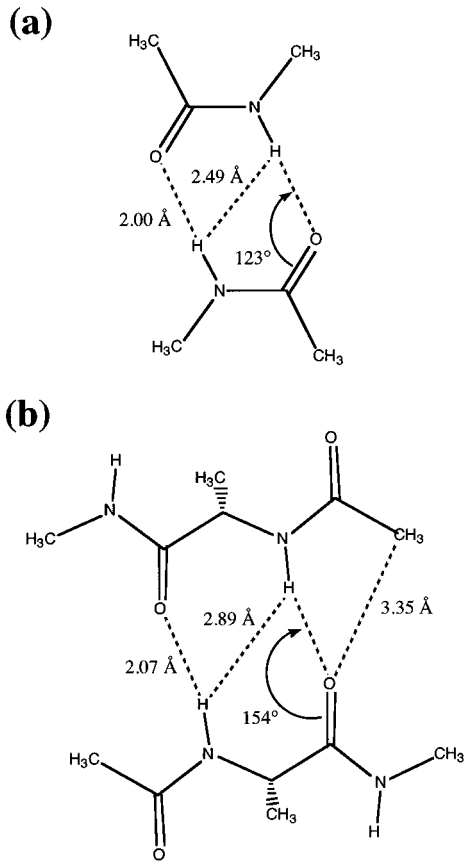



**Figure 3.** HF/6-31G* optimized geometries for (a) the *cis-N*-methylacetamide dimer and (b) the alanine dipeptide dimer (with the $\phi$ and $\psi$ angles fixed at $-139°$ and 135°, respectively). Key interatomic distances are shown.

**Table 8.** Interaction Energies (kcal/mol) of Two $\beta$-Sheet Conformation Alanine Dipeptides[a]

| | |
|---|---|
| MM3*($\epsilon = 1.5$) | −7.23 |
| HF/cc-pVTZ(−f) (CP corrected) | −8.15 |
| MMFF($\epsilon = 2.0r$) | −8.24 |
| HF/6-31G** (CP corrected) | −9.25 |
| HF/cc-pVTZ(−f) (non-CP corrected) | −9.42 |
| MM2*($\epsilon = 1.5$) | −9.69 |
| MM3* | −9.78 |
| AMBER 3 | −9.84 |
| MM3*($\epsilon = 1.0r$) | −10.23 |
| LMP2/cc-pVTZ(−f) (HF CP corrected) | −10.73 |
| CVFF | −10.77 |
| MMFF($\epsilon = 1.5r$) | −11.01 |
| AMBER* | −11.07 |
| HF/6-31G** (non-CP corrected) | −11.68 |
| LMP2/cc-pVTZ(−f) (HF non-CP corrected) | −12.00 |
| CFF95 | −12.14 |
| AMBER*($\epsilon = 1.0r$) | −12.98 |
| MSI CHARMm | −12.99 |
| MM2* | −13.02 |
| OPLS-AA(2,2) | −13.21 |
| MM2*($\epsilon = 1.0r$) | −13.47 |
| CHARMM 22 | −14.10 |
| MMFFs | −14.97 |
| CHARMM 19 | −15.21 |
| MMFF | −15.38 |
| AMBER94 | −16.01 |
| MM2X | −16.11 |
| AMBER94($\epsilon = 1.0r$) | −16.50 |
| OPLS/A-UA(2,8) | −16.70 |
| OPLS-UA(2,2) | −16.91 |
| OPLS* | −17.63 |

[a] Unless otherwise specified, $\epsilon = 1.0$.

this would not be an issue: such a potential function should accurately reproduce both gas-phase and condensed-phase intermolecular interactions. However, all of the force fields we have tested use fixed charges, which are assumed to have the effects of the environment built in implicitly in the parametrization process.

The traditional point of view in developing an empirical force field that does not explicitly incorporate polarizability is that one should employ enhanced atomic charges to simulate the effects of polarization in a high-dielectric medium. A plausible case can be made that fixed charge models like all of those tested here should reproduce binding energies that reflect, in an averaged manner, the increase in polarization of the monomer units that results from immersion in a dielectric medium. This philosophy, for example, underlies the parametrization of MMFF94[8] and of CHARMM 22[5] against scaled HF/6-31G* intermolecular interaction energies and geometries, in which *ab initio* well depths are systematically increased by 10−15% and heteroatom distances are shortened by 0.2−0.3 Å. These adjustments are of the order of those required to convert the (non-counterpoise corrected) HF/6-31G* binding energy of −5.6 kcal/mol and O···O distance of 2.97 Å for the linear water dimer to the values of approximately −6.5 kcal/mol and 2.75 Å required to reproduce the properties of liquid water when a fixed-charge model such as TIP3P[55] or SPC[56] is used. For comparison, we find that the LMP2/cc-pVTZ(−f) model for the water dimer yields an interaction energy of −4.8 kcal/mol, a result that is encouragingly close to the gas-phase value of $-5.4 \pm 0.7$ kcal/mol[57] and nearly identical to a very high level quantum chemical estimate[58] but that diverges even further from the value of −6.5 kcal/mol used by the empirical water models. Given that we find the non-counterpoise corrected HF/6-31G** interaction energy for the $\beta$ dimer to be −11.68 kcal/mol, this argument would suggest that the macromolecular force fields tested in this paper should yield an interaction energy of −12.8 to −13.4 kcal/mol.

A major problem with this argument arises from the special nature of cooperative hydrogen bonding in liquid water. Water was the first polar polyatomic molecule to be parametrized for molecular dynamics simulation, and has served throughout the past two decades as a testing ground for potential function development. Thus, scaling of interactions based upon a procedure that yields reliable experimental results for water is in many ways a historical legacy.

A close examination of the underlying physics of the enhanced average hydrogen bonding strength in water reveals that the large shift from the gas-phase value in solution is due to the uniquely cooperative nature of the hydrogen bonding network. That is, each hydrogen bond in liquid water enhances the strength of every other hydrogen bond. This will not always be the case, however. Consider, as a trivial example, a molecule with two nearby carbonyl groups. It is obvious that hydrogen bonding to one carbonyl will not enhance hydrogen bonding to the second. In fact, it will have just the opposite effect, as the first hydrogen bond will tend to withdraw electrons from the second carbonyl. More generally, one would expect a substantial diminishment of the huge effect in liquid water due to the heterogeneous nature of the typical "condensed-phase environment". Thus, it is difficult to justify the use of uniform scaling based upon liquid water for every hydrogen bonding interaction in a protein or other complex biological system.

(55) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926-935.
(56) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, Holland, 1981.
(57) Curtiss, L. A.; Frurip, D. J.; Blander, M. *J. Chem. Phys.* **1979**, *71*, 2703.
(58) Szalewicz, K.; Cole, S. J.; Kołos, W.; Bartlett, R. J. *J. Chem. Phys.* **1988**, *89*, 3662−3673.

**Table 9.** Hydrogen-Bonded Dimer Binding Energies (kcal/mol)

| method | cis-NMA | alanine dipeptide | $\Delta E$ |
|---|---|---|---|
| HF/6-31G* | −13.15 | −11.80 | 1.35 |
| AMBER* | −11.16 | −10.65 | 0.51 |
| MM3* | −8.72 | −9.62 | −0.90 |
| OPLS* | −14.58 | −16.44 | −1.86 |
| MMFF94 | −11.66 | −14.44 | −2.78 |
| CHARMM 22 | −11.34 | −13.59 | −2.25 |
| MM2X | −13.34 | −15.93 | −2.59 |
| MM2* | −9.65 | −12.88 | −3.23 |
| AMBER94 | −11.07 | −14.43 | −3.36 |
| MSI CHARMm | −9.31 | −12.74 | −3.43 |
| CHARMM 19 | −11.14 | −14.96 | −3.82 |

On balance, it is still likely that calibrating intermolecular interactions in a fixed-charge model to values in excess of those observed in the gas phase will serve to improve the accuracy with which the model can describe such quantities as ligand−enzyme binding energies. The effects of medium polarization will typically lead to increased partial atomic charges. However, the cautionary arguments presented above with regard to using pure liquid water as a benchmark for this calibration suggest that the best value (i.e. that leading to the smallest RMS deviation for the interaction when averaged over an ensemble of environments relevant to protein structure or ligand binding predictions) may well be smaller than the estimates made above. It is at present not possible to validate objectively either this point of view or the more traditional one (the authors themselves are not in complete agreement on this point), and the reader will have to form his or her own judgment concerning which perspective to adopt (and hence how to interpret the results in Table 8).

Despite this uncertainty, it appears that MM2X, AMBER94, OPLS*, OPLS/A-UA(2,8), and OPLS-UA(2,2) substantially overestimate the interaction energy. To a lesser extent, CHARMM 22 and 19, MMFF, and MMFFs also overestimate the interaction energy. Even if one accepts the argument for enhanced intermolecular interactions, the upper end of the target range is −12.8 to −13.4 kcal/mol, as stated above. Within this target range are AMBER*($\epsilon = 1.0r$), MSI CHARMm, and MM2*($\epsilon = 1.0$). However, if the gas-phase dimerization energy is taken as the proper benchmark, it should be noted that MM3*-($\epsilon = 1.0r$), CVFF, MMFF($\epsilon = 1.5r$), and AMBER* come the closest to the LMP2 value of −10.73 kcal/mol (with the HF portion counterpoise corrected), each having deviations of 0.5 kcal/mol or less. Two force fields are well below even the LMP2/cc-pVTZ(−f) gas-phase value in magnitude and almost certainly too low—MM3*($\epsilon = 1.5$) and MMFF($\epsilon = 2.0r$).

It is especially disconcerting that substantial overestimates are given by MMFF and AMBER94, two of the most recently introduced force fields, and by OPLS-UA(2,2) and OPLS/A-UA(2,8), which are widely considered to describe intermolecular interactions well. In investigating the origin of these discrepancies, we turned to the cyclic *cis*-*N*-methylacetamide (*cis*-NMA) dimer. We felt that this system might prove interesting because the similar cyclic formamide dimer was the only hydrogen-bonded dimer in the parametrization of MMFF94 in which the calculated binding energy was *less negative* than the *ab initio* HF/6-31G* value.

Binding energies ($\epsilon = 1.0$ ) for the two dimers (obtained with the respective monomers optimized, not frozen as in Table 8), are listed in Table 9; the dimer geometries are illustrated in Figure 3. As the table shows, the HF/6-31G* binding energy (non-counterpoise corrected) is 1.35 kcal/mol larger (more negative) for the *cis*-NMA dimer than for the $\beta$-dipeptide dimer. All but one of the force-field models, however, reverse this pattern, finding the $\beta$ dimer to be more stable by 2.37 kcal/mol on average. The typical force field error in the relative binding energies for the two dimers is thus a sizable 3.7 kcal/mol. Moreover, none of the force field binding energies for *cis*-NMA exceeds the *ab initio* value by more than 15%, and most are actually smaller in magnitude. Clearly, errors in the force field energetics are just as serious for the *cis*-NMA dimer as for the $\beta$ dimer—but occur in the opposite sense!

We suggest that these discrepancies arise from an insufficiently appreciated shortcoming of current-generation force fields. Note, as Figure 3 shows, that the HF/6-31G** optimized C═O···H angle rises from 123° in the *cis*-NMA dimer to 154° in the $\beta$ dimer. We posit that the anisotropy of the electron distribution around oxygen favors a coordination angle more like that seen in the *cis*-NMA dimer, in which the polar hydrogen interacts with the carbonyl oxygen along a "rabbit-ear" lone-pair direction. In the $\beta$ dimer, this angle (which would require translating the lower dipeptide monomer in Figure 3 to the right) cannot be attained, because such a geometry would intensify the intermolecular steric clash between the carbonyl oxygen and the terminal methyl carbon and its attached hydrogens (cf. Figure 3). As a result, the HF/6-31G*-optimized H···O distance in the $\beta$ dimer lengthens, despite the fact that the smaller electrostatic clash between the polar hydrogens would facilitate a closer approach.

For the force field models, in contrast, the reliance on atom-centered charges appears to yield a preference for the more nearly linear C═O···H angle found in the $\beta$ dimer. For the simpler formaldehyde···HOH dimer, for example, we find that the C═O···H angle increases from 101° at the HF/6-31G* geometry to 149° upon MMFF optimization. For the $\beta$ dimer, the altered geometric preference enables the C═O···H angle to increase further, e.g., to 166° at the MMFF-optimized geometry. This variation (which corresponds to moving the lower monomer in Figure 3 to the *left*) is beneficial because it reduces the steric clash between the carbonyl oxygens and terminal methyl groups and thus allows the two monomers to approach more closely. For nine of the ten force-field models, these factors combine to produce a more negative binding energy for the $\beta$ dimer; the exception, AMBER*, is not yet understood.

If this interpretation is correct, it follows that force fields that represent electrostatic interactions solely in terms of atom-centered partial charges will not describe the whole range of plausible hydrogen-bonding geometries accurately. Significantly, OPLS*, the only force-field model in Table 9 that exceeds the HF/6-31G* interaction energy for the *cis*-NMA dimer by the amount (10−15%) discussed above, is also the force field that performs most poorly for the $\beta$ dimer when judged by the same standard. As was proposed in the development of MMFF94,[7,8] future force fields will need to include a more complex model for electrostatic interactions than is provided by atom-centered charges. Such a model, for example, might employ charges offset from the nuclear center-(s), as does the TIP4P model for water.[55] For TIP4P, this modification leads to an angle of 46° between the O···O axis and the acceptor H−O−H bisector in the dimer. This angle agrees much more closely with the HF/6-31G* angle of 62.6° and the experimental angle[59] of ∼60° than do the much smaller angles of 26−27° found for three-point (atom-centered charge) models such as TIP3P,[55] SPC,[56] and MMFF94.[8] Alternatively, an improved model might employ atom-centered dipoles as well as charges, as does Dinur's model for water, which yields an angle of 66° between the O···O axis and the acceptor H−O−H plane,[60] or might employ still another mechanism for enforcing proper hydrogen bond directionality.

(59) Dyke, T. R.; Mack, K. M.; Muenter, J. S. *J. Chem. Phys.* **1977**, *66*, 498.

(60) Dinur, U. *J. Phys. Chem.* **1990**, *94*, 5669−5671.

## 4. Discussion and Conclusions

Though the RMS deviations in conformational energy for the best force fields examined here are in the range of 1.2−1.4 kcal/mol for unrestrained optimizations of the alanine tetrapeptide, it is clear from the results that a truly quantitative prediction of peptide and protein energetics via molecular mechanics is not yet available. Previous attempts at validation of protein force fields have relied upon measures such as RMS deviation of the model protein in a simulation from the X-ray crystal structure after a given length of simulation time. This procedure is not a reliable test of the ability to rank widely different compact structures in the gas phase or in solution, as it simply examines whether or not the X-ray structure (or a close approximation of it) is a local minimum on the force field surface. And, because of packing considerations and the local stability of secondary structure elements in the gas phase due to hydrogen bonding, it almost always is. The energy of an α-helix or $\beta$-sheet conformation does not have to be quantitatively (or even qualitatively) accurate in order to be a local minimum that survives short simulations. However, if one wants to predict the native structure of a protein (or of a small piece, such as a loop) without knowing the answer beforehand, the fact that the native structure is a local minimum will be insufficient. There will be countless other local minima, many of which could easily be lower in energy than the native on the force field surface. In reinforcement of the conclusions of the present study, we do not believe that there have yet been convincing demonstrations that robust predictions of loop or peptide geometries are possible with current force fields.

An argument is often made that one does not want a force field to reproduce accurate results in the gas phase (the comparison being made here) because the actual system is in a condensed phase, and the parametrization may have been designed to take this into account. As discussed earlier, this argument has some merit. Nevertheless, it is ultimately flawed because there is no single condensed-phase environment. The environment inside a protein is surely quite different from that of bulk water. If the effect of the environment is large, a pair potential will be inadequate because it has no possibility of responding to variations in the environment. To demonstrate otherwise, one would have to show that the polarization in all physically important condensed-phase environments was uniform, a point of view that appears highly implausible. Thus, even if one grants that the use of effective pair potentials is the best approach available within the confines of a fixed-charge model, it is unequivocally clear that this approach must be limited in accuracy.

The conclusion of these arguments dovetails with results we have recently obtained concerning solvation free energies of a series of methylated amines and amides.[61] In this study, we showed that the large errors in hydration free energies made across a given series in continuum-dielectric calculations could be readily rationalized by the inaccuracy of the potential functions in reproducing quantum chemical pair hydrogen bonding energies. Though scaled quantum-chemical calculations have been used as benchmarks in developing CHARMM 22 and MMFF, as previously noted, no force field has yet been developed in which rigorous agreement with unscaled, gas-phase quantum chemical pair energies for all relevant chemical groups has either been sought or achieved. Moreover, even if this were to be done, approximation of such interactions by fixed, atom-centered charges would not allow for physically appropriate changes in the force-field description as a function of the dielectric environment, and hence would not yield a proper representation of binding energetics. Furthermore, the comparison of the calculated binding energies for the $\beta$-dipeptide and *cis*-NMA dimers suggests that force fields that use fixed atom-centered charges will not reliably reproduce energetic trends. Even for gas-phase modeling, a more complex electrostatic representation appears to be needed.

What needs to be done is to incorporate high-level correlated quantum chemical results for intermolecular interactions directly into molecular mechanics force fields, much as was done for torsional interactions in the development of MMFF94.[10] The force field model would then need to be able both to describe gas-phase conformational and intermolecular interactions correctly and to adjust self-consistently for the effects of immersion in a medium of varying dielectric constant. Indeed, just this approach was cited in the development of MMFF94 as being the proper long-term objective.[7,8] The essential problem in implementing this approach lies in the difficulty of including condensed-phase polarization effects in the parametrization. Without these effects, the results would likely be every bit as erroneous as the gas-phase energetics examined here, if not more so.

Recent work has suggested that condensed-phase effects can be accommodated by the use of polarizable force fields,[62−67] and furthermore that such force fields can be used in simulations with reasonable computational effort; Berne and co-workers, in particular, report only a 10% increase in computation time for a liquid water simulation utilizing a fluctuating charge model with multiple time scale simulation methods.[68] Efforts along these lines are in progress in our laboratories and those of others. Hopefully, the benchmark structures and energies presented here, which are available upon request,[69] will facilitate such developments by providing an objective test as to whether a given implementation of a polarizable model represents real improvement over existing pair potentials. A uniform accuracy of better than 1 kcal/mol compared to LMP2-level quantum chemistry should be achievable and would represent a significant advance over any of the force fields tested in this paper. Moreover, we expect future work using higher levels of electron correlation (e.g., GVB-LMP2 wave functions) in conjunction with still larger basis sets to yield substantially more accurate quantum-chemical benchmarks against which empirical force fields can be even more reliably developed and tested. A goal of peptide benchmarks accurate to 0.5 kcal/mol with these improved quantum chemical methods would appear to be in reach in the coming year.

Despite the evident simplifications in current force field models, the performance of the top-rated force fields in many ways is encouraging. Much progress has been made in peptide and protein molecular mechanics. The ability to reproduce the relative energies for the tetrapeptide conformers to an RMS precision of 1.5 kcal/mol or less clearly is highly nontrivial and is achieved by several of the force fields. The fact that the force fields vary widely in performance and suffer episodic

(61) Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. *J. Phys. Chem.* **1996**, *100*, 11775−11788.

(62) Corongiu, G.; Migliore, M.; Clementi, E. *J. Chem. Phys.* **1989**, *90*, 4629.

(63) Dang, L. X.; Rice, J. E.; Caldwell, J.; Kollman, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 2481−2486.

(64) Sprik, M. *J. Phys. Chem.* **1991**, *95*, 6762−6769.

(65) Zhu, S.-B.; Yao, S.; Zhu, J.-B.; Singh, S.; Robinson, G. W. *J. Phys. Chem.* **1991**, *95*, 6211−6217.

(66) Bernardo, D. N.; Ding, Y.; Krogh-Jespersen, K.; Levy, R. M. *J. Phys. Chem.* **1994**, *98*, 4180−4187.

(67) Ding, Y.; Bernardo, D. N.; Krogh-Jespersen, K.; Levy, R. M. *J. Phys. Chem.* **1995**, *99*, 11575−11583.

(68) Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141.

(69) Structures are also available via the World Wide Web at http://www.chem.columbia.edu/∼beachy/structures.html.

failures for quite different conformers undoubtedly reflects the fact that radically different data and approaches have been used in their derivation. As noted above, we have already inferred from observations on hydrogen bond energetics as a function of angular geometry that electrostatic interactions need to be described by more than a fixed, atom-centered charge model. Similarly, a careful analysis of the successes and failures of each method for conformational energies should provide important new insights into the necessary and sufficient conditions for improvement of force-field development.

While a comprehensive analysis is beyond the scope of this paper, we can offer a preliminary assessment. First, we note that how the individual force fields generate their spread of tetrapeptide conformational energies varies widely. In particular, some heavily parametrize the torsional terms while others do not. For example, for the restraint-optimized tetrapeptide conformers the torsional energies range from 5.7 to 16.8 kcal/mol for MMFF94, from 7.9 to 17.1 kcal/mol for AMBER94, and from 10.0 to 19.0 kcal/mol for CHARMM 22. In contrast, these energies vary just from 0.5 to 1.1 kcal/mol for CHARMM 19 and from 2.3 to 3.2 kcal/mol for MSI CHARMm. Clearly, torsional terms contribute strongly to the relative conformational energies only in the first three cases. However, no correlation is found between the range of the torsional energies and the range of the total conformational energies themselves, which Table 7 shows to be 6.7, 8.4, 16.9, 15.4, and 10.8 kcal/mol, respectively. Neither is a large torsional energy necessarily associated with a high total conformational energy. For MMFF, for example, the lowest energy structure, conformation 3, has the second highest torsional energy, but also the second most stabilizing electrostatic energy. Nor is there a strong correlation between the magnitude of the torsional contribution and the magnitude of the RMS error in conformational energy, which Table 5 gives as 1.21, 1.42, 3.78, 2.05, and 2.12 kcal/mol, respectively. While the best performance among these five methods is achieved by the strong torsion MMFF94 and AMBER94 force fields, CHARMM 22 (see below) is an exception. Though further study is needed, a possible conclusion for force-field development is that strong torsional parametrization is a necessary, but not a sufficient, condition for accurate peptide conformational energies.

Lastly, for several of the force fields considered in this paper, we have determined the degree to which errors in the tetrapeptide conformational energetics can be traced to the conformational energy errors the force field makes for the dipeptide. Such an analysis is possible for seven of the ten conformations (all those except for 5, 7, and 8) because each $\phi$, $\psi$ pair (cf. Table 4) corresponds reasonably closely to one of the dipeptide minima (Table 1]). Purely for purposes of analysis, the tetrapeptide energies obtained in the restrained optimizations can be additively corrected for the associated dipeptide errors.[70] We find that these corrections almost always improve the fit to the LMP2/cc-pVTZ(−f) conformational energies, sometimes dramatically so. For MMFF, in particular, the RMS energy error for conformations 1, 2, 3, 4, 6, 7, and 10 falls from 1.02 to a scant 0.44 kcal/mol. Equally dramatically, the RMS error for CHARMM 22 falls from 3.79 to 0.95 kcal/mol, largely because the dipeptide correction compensates for the large error CHARMM 22 makes for the α′ conformation, which is placed ∼6 kcal/mol too high. Other force fields we have examined to date are affected less strongly. Nevertheless, these preliminary results provide grounds for hope that rigorous parametrization against *ab initio* data of sufficiently high quality will yield a force-field model that accurately describes systems that are larger and more complex than those used in its parametrization.

As for the present study, the reader who wishes to know which of the current-generation force fields examined here is "best" will have to use his or her own judgment. The decision will also have to take into account the application. One issue for prospective applications is the range of chemistry for which the force field is parametrized. In this regard, MMFF94 and CFF95 are among the most widely parametrized, while OPLS, AMBER, and the academic CHARMM force fields focus mainly on functional groups that appear in biological macromolecules. As we have seen, there are significant differences in the ability to locate the *ab initio* minima on the tetrapeptide surface, to predict relative conformational energies accurately, and to produce what (on one basis or another) appears to be an appropriate interaction energy in the $\beta$-dipeptide and *cis*-NMA dimer calculations. Clearly, several force fields are deficient in one or more respects. But, while none is best in all, it can be seen that MMFF (1994 version), MMFF93, and OPLS-AA-(2,2) earn top scores in nearly all the categories in Table 5. OPLS/A-UA(2,8) and MMFFs also tend to do consistently well. But, again, the requirements for success in a given application will vary and the final judgment must be the reader's.

**Acknowledgment.** R.A.F. acknowledges support from the National Center for Research Resources, Biomedical Technology Research Program (P41-RR06892). Work done at Schrödinger Inc. was partially supported by a Phase I Small Business Innovative Research Grant (GM-53301-01). We thank the following professional colleagues for performing conformational and interaction energy calculations: Julian Tirado-Rives (authentic OPLS methods), Jay Banks (AMBER 3), Christopher Bayly and Peter Kollman (AMBER 4.1), Carol Parish (AMBER94 and AMBER*), Pieter Stouten (GROMOS), and Carl Ewig and Xiangshan Ni (CVFF, MMFF93, and CFF95). We also thank Alex Mackerell for assisting in the validation of the CHARMM 22 calculations reported herein.

**Supporting Information Available:** Table 1S containing the geometry RMS deviations for individual force field minima in comparison with the corresponding HF/6-31G** minima, Tables 2S and 3S containing the force field and quantum mechanical relative energies of the tetrapeptides relative to Conformation 3 (Table 2S for the unrestrained optimizations and Table 3S for the restrained ones), Table 4S containing energy comparisons made on a pairwise basis and binned in 1-kcal/mol increments (9 pages). See any current masthead page for ordering and Internet access instructions.

JA962310G

(70) T. A. Halgren, unpublished results. The force field calculations minimized the dipeptide conformers while restraining their $\phi$, $\psi$ angles tightly to the HF/6-31G** present work values listed in Table 1. The relative dipeptide energies were then compared to the LMP2/cc-pVTZ(−f) results listed in Table 2 to ascertain the force-field errors in these quantities.