JCTC Journal of Chemical Theory and Computation

Force Field Modeling of Amino Acid Conformational Energies

Jakub Kaminský[†,‡] and Frank Jensen*[,†]

Department of Physics and Chemistry, University of Southern Denmark, DK-5230 Odense, Denmark, and Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 166 10 Prague, Czech Republic

Received April 2, 2007

Abstract: The conformational degrees of freedom for four amino acids in a model peptide environment have been sampled with density functional and second-order Møller-Plesset methods. Geometries have been optimized with an augmented double-ζ basis set and relative energies estimated by extrapolation of results using double, triple, and quadruple-ζ basis sets and including higher order correlation effects. In addition, the effects of vibrational zero point energies and solvation have been considered. The density functional method is unable to locate all the minima found at the MP2 level, which most likely is due to the inability for describing dispersion interactions. The use of basis sets smaller than augmented polarized double-ζ with the MP2 method may also in some cases lead to artifacts. The effects on relative energies by enlarging the basis set beyond an augmented triple-ζ and including higher order correlation beyond MP2 is small. The MP2/aug-cc-pVTZ level is recommended as a level of theory capable of an accuracy of ∼1 kJ/mol for relative conformational energies. Eight different force fields are tested for reproducing the electronic structure reference data. Force fields that represent the electrostatic energy by fixed partial charges typically only account for half of the conformations, while the AMOEBA force field, which includes multipole moments and polarizability, can reproduce ∼80% of the conformations in terms of geometry. This not only suggests that multipole moments and polarizability are important factors in designing new force fields but also indicates that there is still room for improvements.

Introduction

Investigating the three-dimensional structure of proteins is of prime importance for understanding their biological function. X-ray diffraction methods have in recent years provided a wealth of structural information,[1] but these methods only provide a static picture corresponding to a solid-state structure, and certain classes of proteins are difficult to investigate due to crystallization problems. NMR methods are capable of providing structural information under conditions closer resembling the natural biological environment (solution), but these methods are limited to relatively small systems.[2] Experimental information regarding the time-dependent changes in structural features can be obtained from time-resolved spectroscopy, but this can rarely provide details at the atomic level.

Theoretical simulations can in principle provide (time-dependent) atomic resolution information of proteins and other macromolecules. Predicting the three-dimensional structure of a protein from its primary amino acid sequence is still an unsolved problem due to the huge size of the phase space. Simulations of macromolecules therefore usually start from geometries derived from experiments, either directly from an X-ray structure of the actual system or by homology modeling of a closely related structure.

An essential part of a simulation is the energy function. Car−Parrinello methods take the electrons explicitly into

* Corresponding author e-mail: frj@ifk.sdu.dk.
† University of Southern Denmark.
‡ Czech Academy of Sciences.

10.1021/ct700082f CCC: $37.00 © 2007 American Chemical Society
Published on Web 08/18/2007

account by propagating both the nuclear and wave function parameters in time, but the computational requirements of these methods are sufficiently high that only relatively small systems can be simulated for short time spans (picoseconds).[3] Performing atomistic simulations for systems containing thousands of atoms in the nano- or microsecond time regime is only possible using parametrized energy functions.[4] Such force field methods rely on using atoms as the fundamental building block and providing the bonding information explicitly. It is evident that the quality of the energy function determines the ultimate accuracy that can be extracted from simulations.

The internal energy of a molecule is parametrized in a force field approach by a function containing constants that are fitted to reproduce experimental or quantum mechanical results.[5,6] All force fields contain the terms shown in eq 1, and additional terms may be added to improve the performance.

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

The stretch and bending terms are relatively easy to parametrize, and they have little influence on conformational degrees of freedom. The nonbonded van der Waals and electrostatic terms determine intermolecular interactions and are thus very important for calculating, e.g., binding affinities of drug molecules to enzymes. The nonbonded terms together with the torsional energy determine the internal (conformational) degrees of freedom. The existence and relative stabilities of conformations for a given molecule is thus determined by a delicate balance between these three energy terms. Being able to correctly predict structural and energetic features of conformations is an essential quantity for bringing force field methods into the predictive region of theoretical methods. Despite this, there have only been a few studies where the performance of force field for predicting structural and energetic information of conformations has been addressed.[7] Part of this calibration problem is, of course, due to the lack of reference data of sufficient accuracy.

One of the main limitations of current force field methods is the use of fixed partial charges for parametrizing the electrostatic interaction, as this neglects the known conformational dependence of the atomic charges and polarization effects.[8,9] This becomes especially problematic for polar systems, like proteins, where the electrostatic component often accounts for a large fraction of the total energy.[10] Currently research therefore focuses on improving the representation of the electrostatic energy by including multipole moments and atomic polarization.[11] These quantities are usually derived from electronic structure calculations on small model systems, but there is no clear consensus regarding which electronic structure method to use or the procedure for extracting the parameters. Lacking an objective criterion, a distinction between the different procedures will have to rely on the performance for reproducing structural and energetic feature of the target molecular systems. We here report a systematic study of the conformational space of four amino acids in a model peptide environment and evaluate the performance of different force fields for reproducing the reference data. The aim is 2-fold, to establish the necessary theoretical level for reliably determining structures and energies of amino acid conformations for use on a larger variety of systems and to subsequently use this data for developing more accurate force fields.

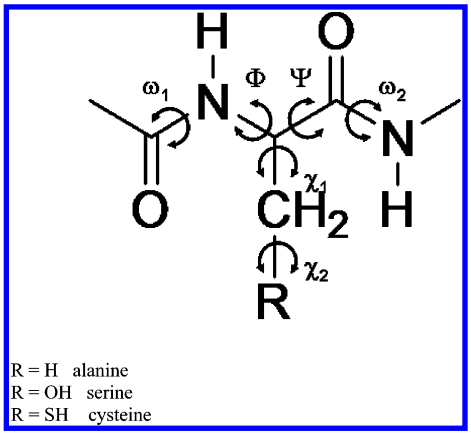

Figure 1. Illustration of the torsional angles in the model systems.

Computational Details

All electronic structure calculations have been carried out using the Gaussian 03 program package.[12] Force field calculations have been performed with the MacroModel[13] and Tinker programs.[14] The MacroModel program allows calculations with the AMBER94, MM2*, MM3*, MMFFs, and OPLS_05 force fields which are slightly modified versions of the parent AMBER,[15] MM2,[16] MM3,[17] MMFF,[18] and OPLS[19] force fields. The Tinker program has been used for the AMBER99,[20] CHARMM27,[21] and AMOEBA[22] results. All calculations have been done using a constant dielectric constant of 1.0 for comparisons with the calculated gas-phase conformational energies. The AMOEBA force field includes distributed multipoles and polarizabilities for representing the electrostatic energy, while the others use a fixed partial charge model.

We have selected the glycine, alanine, serine, and cysteine amino acids as our trial set of systems, with side chains corresponding to H, CH_3, CH_2OH, and CH_2SH, respectively. N-Acetyl and N-methylamide functional groups were added to the C- and N-terminus to mimic the environment in longer peptide chains. The notation for the torsion angles is depicted in Figure 1. Trial structures were generated by varying the Φ and Ψ torsion angles in steps of 30°, while the χ_1 torsional angle was varied in steps of 120°. The torsion angle χ_2 was preset at a value of 180°, while the torsional angles ω_1 and ω_2 associated with the terminal amide bonds were given initial values of 180°, corresponding to trans junctions. This produces a total of 144 starting geometries for glycine and alanine and 432 for serine and cysteine. Some of these conformations are, of course, related by symmetry. We generated additional trial conformations for the serine and cysteine systems by varying the χ_2 torsional angle. All trial structures were optimized without any restraints at the MP2/6-31G(d,p) level and characterized as true minima by frequency calculations, and the unique structures were reoptimized with the aug-cc-pVDZ basis set. Although we do not claim that this procedure exhausts the conformational

space, it certainly will locate the large majority of the important conformations.

Improved estimates of the relative conformational energies were obtained by single point MP2 calculations using the cc-pVDZ, cc-pVTZ, cc-pVQZ, and aug-cc-pVTZ basis sets, and method dependency was tested using CCSD(T)/cc-pVDZ calculations. Our best estimate of relative energies was obtained by separately extrapolating the Hartree−Fock[23] and MP2[24] energies to the basis set limit and adding the difference between the CCSD(T) and MP2 results with the cc-pVDZ basis set to the extrapolated MP2 results.

The same initial pool of trial structures were also optimized at the B3LYP/6-31G(d,p) level, and improved estimates of relative energies were obtained using single point calculations at the B3LYP/aug-cc-pVTZ level in order to investigate the performance of a popular density functional model. Solvent effects were estimated by single point B3LYP/6-31G(d,p) and MP2/6-31G(d,p) calculations using the PCM continuum solvent model for water.[25] Zero point energy differences were evaluated using harmonic frequencies. Force field conformations were determined by a Monte Carlo type random variation of the torsional angles followed by energy minimization, where the number of trial steps was set sufficiently high to ensure that all possible conformations are located.

In order to establish the correspondence between the force field and ab initio conformations, we have calculated root-mean-square (rms) deviations in the torsional angle space. In most cases this allows a unique connection to be made. For large rms deviations we have performed MP2/6-31G-(d,p) optimizations starting from the force field geometries in order to establish the nearest conformation.

Previous Work

Previous work on calculating conformations for peptide model systems has employed various methods based on wave function or density functional theory (DFT). A relatively large number of calculations of peptide systems have been published during the last two decades, but most of them focused on alanine or glycine systems[26−33] and only a few have considered other amino acids (e.g., serine,[34−37] cysteine,[38] asparagine,[39] proline[40]). Several of these studies have used di- or tetrapeptides as model systems, but the size and flexibility of these systems compared to simple amino acids necessitated simplification in the exploration of the conformational space. A typical approach has been to optimize a number of Φ, Ψ restrained conformers at the Hartree−Fock (HF) level of theory, followed by single point MP2 calculations. Low-energy conformations identified by this procedure is subsequently fully relaxed. Beachy et al. reported that for their test suite consisting of simple flexibile molecules such as methyl vinyl ether, the use of MP2/6-31G(d,p) geometries instead of HF/6-31G(d,p) ones changes the conformational energies by only a few tenths of a kcal/mol, and they therefore considered HF geometries as adequate.[41] Gould et al., however, showed that some minima on the potential energy surface could not be located at the HF level, while the MP2 level performed much better.[42] Beachy et al. showed for the alanine dipeptide model that canonical MP2 energies are contaminated by basis set superposition errors, which artificially lower the energies of the compact structures relative to extended ones.[43] Császár showed that inclusion of higher order (MP3, MP4, CCSD, and CCSD(T)) electron correlation in most cases makes only small contributions to the energy differences of glycine conformers and concluded that large basis set MP2 calculations should result in highly accurate energies.[44]

DFT methods have in recent years been accepted by the chemistry community as a cost-effective approach for the computing of molecular structures, vibrational frequencies, and energies of chemical reactions.[45−49] Unfortunately, most current DFT methods do not account for dispersion, and for large systems it is anticipated that intramolecular dispersion type interactions will play a significant role. DFT results for describing the conformational space are consequently likely to be of lower accuracy than MP2 results.[50]

For systems as large as tetrapeptides the number of possible starting structures for performing a systematic conformational search is prohibitive for use with electronic structure methods. Most previous work on such systems has therefore employed molecular mechanics methods for an initial evaluation of the conformational space. Beachy et al. selected 10 conformers of the alanine tetrapeptides based on AMBER* results and recomputed them with several different molecular mechanics force fields and compared the results to LMP2/cc-pVTZ//HF/6-31G(d,p) data.[53] A number of higher energy structures were deliberately chosen among the 10 structures in order to sample the conformational space representing low-lying minima. Higher energy conformers could potentially be populated due to solvation effects or by interaction with other parts of a particular protein. Gresh et al. treated Beachy's 10 alanine tetrapeptide conformations with their SIBFA force field, which includes multipoles and polarization, and evaluated relative energies by single point HF, DFT, and MP2 calculations.[51] The relative energies calculated at the LMP2/6-311G(d,p) level were reproduced by the SIBFA force field with an rms value of ∼5 kJ/mol. It was also found that the MP2 method tended to exaggerate the stabilization of folded structures, presumably due to BSSE effects.

The conformations of small linear peptides containing the Arg-Gly-Asp sequence have been studied by NMR and/or molecular dynamics simulations. These studies showed, that although linear peptides are expected to be very flexibile, the linear Arg-Gly-Asp-containing peptides could adopt mainly β-turns for Arg-Gly and Gly-Asp sequences.[52−54] The unfolding of the Ac-Ala-Ala-NHMe and Ac-Pro-Ala-NHMe sequences in water was monitored by Tobias et al.[55,56] using MD simulations with the CHARMM force field. Recently, ab intio molecular dynamic simulation techniques employing a Gaussian implementation of Kohn−Sham density functional theory was used to study the gas-phase conformational dynamics of an alanine dipeptide analog,[57] although the full range of Φ, Ψ values have yet to be sampled using this approach.

The conformational energies of dipeptides have a direct relationship with protein modeling by force fields, which has motivated several molecular mechanics studies of alanine and glycine dipeptides. Despite the fact that many force fields

have been applied to a wide range of peptides, the question of their accuracy still remains. Molecular dynamics simulations of polypeptides in solution showed the existence of π-helical conformations,[58,59] but this has been reported by Feig et al. to be associated with force fields bias.[60] For systems like dipeptides or tetrapeptides, most force fields have problems reproducing quantum mechanical results.[43,61−64] We therefore feel that there is a need for improving the treatment of peptide conformations by empirical force fields. To date, optimization of empirical force fields for the protein backbone has typically been based on reproducing the relative energies of selected conformers of the alanine and glycine dipeptides.[55,65−69] This approach has the limitation that it ignores high-energy regions of the Φ, Ψ conformational space, although MacKerell et al. recently have reported an optimized CHARMM force field, where cross terms were added in order to treat the entire range of Φ, Ψ values.[70]

As indicated by the above references, a large majority of previous work has relied on force field, HF or DFT methods for selecting and optimizing conformations, and only rarely has the influence of improving the theoretical procedure by enlarging the basis set or treatment of electron correlation been addressed. The results presented in the next section provide a systematic study of the effects of different computational procedures.

Results

While HF and DFT methods have been popular for searching the conformational space due to their favorably computational requirements, it should be recognized that they lack a description of dispersion interactions. Even for a medium sized system, however, it is expected that intramolecular dispersion interactions will be important for the topology of the potential energy surface. The MP2 method is the lowest level of theory that provides a qualitative correct description of dispersion, and it is at present also the only computationally feasible method for performing a large number of structure optimization. The overall topology of the energy surface is expected to be correctly reproduced with a double-ζ type basis set, and we have selected the 6-31G(d,p) basis set for computational efficiency for the initial screening of the conformational space. The resulting conformations have been refined with the more flexible aug-cc-pVDZ basis set, which in some cases leads to different conformations collapsing to the same structure, i.e., the 6-31G(d,p) basis set may in some cases produce artificial minima.

Relative energies are sensitive to the quality of the basis set, and our reference energies are based on MP2 energies extrapolated to the basis set limit from calculations with the cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets and with an additive correction for higher order correlation based on CCSD(T)/cc-pVDZ results. From the observed changes upon improving the basis set or treatment of electron correlation, we estimate that relative energies calculated by this procedure are converged to ~1 kJ/mol. In general we do not expect such an elaborate treatment to be necessary and also report the results obtained at the MP2/aug-cc-pVTZ level, which we feel represent a suitable compromise between accuracy and computational cost.

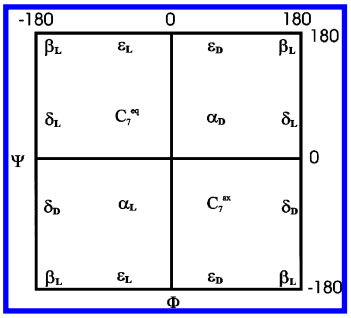

Figure 2. Connection between backbone torsional angles and conformational labeling.

Conformations for peptides are commonly discussed in terms of a Ramachandran map with Φ and Ψ as the variables, where a positive sign indicates a clockwise rotation. Most peptide residues exhibit nine unique conformations in accordance with values of their Φ and Ψ torsion angles, traditionally labeled as α_D, ϵ_D, C_7^{ax}, δ_L, β_L, δ_D, C_7^{eq}, ϵ_L, and α_L, as illustrated in Figure 2.[71] Since we in the present work only consider conformations with a trans peptide bond, the full conformational space includes two (glycine, alanine) or four (serine, cysteine) torsion angles, respectively, and the potential energy surface can therefore be described as a function of two or four variables.

Glycine. Optimization at the MP2/aug-cc-pVDZ level produced two structures, one in the C_7^{ax} and one in the β_L region, as summarized in Table 1. Both conformations were confirmed to be separate minima by frequency calculations and have been reported previously.[52,72,73] Note that C_7^{ax} and C_7^{eq} are symmetrically equivalent for glycine. Bohm et al. reported also a minimum in the α_L region on the AMBER potential energy surface, but this conformation is not a local minimum at the HF/DZP level and is not found in the present work either.[72] Our best estimate for the energy difference between the two conformations is 4 kJ/mol, but both zero point energies and solvent corrections suggest that the β_L conformer may be essentially isoenergetic under solution-phase conditions.

The B3LYP method reproduces the geometries of both minima with a root-mean-square deviation of torsion angles of only 1° relative to the MP2 structures, but the relative energy of the β_L conformer is significantly smaller. Single point MP2 calculations at the B3LYP geometries give relative energies almost identical to those at the MP2 optimized geometries, indicating that this is not a geometry effect. This is slightly at variance with work by Barone et al. where it was concluded that the B3LYP method can reproduce MP2 relative energies for free amino acids with accuracy about 1 kJ/mol,[74] but it confirms the conclusions of Beachy[43] that MP2 prefers folded structures. The zero point energy and solvent corrections determined at the B3LYP level resemble closely those at the MP2 level. Although the PCM model is expected to provide a qualitative indication of the solvent effect, it is unlikely to be able to quantitatively predict the effect of surrounding these polar systems with water, as for example hydrogen bonding is not accounted for. We therefore primarily use the PCM results

Table 1. Glycine Conformations[a]

				MP2				B3LYP			
conf	region	Φ	Ψ	A	B	ΔZPE	ΔPCM	A	ΔZPE	ΔPCM	rms
1	C_7^{ax}	82	−71	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1
2	β_L	180	180	5.7	4.1	−2.8	−3.2	1.3	−2.9	−4.6	0
MAD				1.6				2.8			1

[a] Relative energies are in kJ/mol. Glycine conformations optimized at the MP2/aug-cc-pVDZ (MP2) and B3LYP/6-31G** (B3LYP) levels. Relative energies from single point calculations with the aug-cc-pVTZ basis set are labeled with A, while the best estimates obtained by basis set extrapolation and CCSD(T) corrections are labeled B. Changes in relative energies due to zero point corrections (ΔZPE) and PCM (ΔPCM) solvation corrections have been calculated at the MP2/6-31G** and B3LYP/6-31G** levels. Root-mean-square deviation (rms) of the B3LYP torsional angles relative to the MP2 values in deg. MAD indicates the mean absolute deviation over the conformations.

Table 2. Force Field Predictions of Glycine Conformations[a]

			AMBER94		MM2*		MM3*		MMFFs		OPLS_2005		AMBER99		CHARMM27		AMOEBA	
conf	region	ref	rms	Δ*E*	rms	Δ*E*	rms	Δ*E*	rms	Δ*E*	rms	Δ*E*	rms	Δ*E*	rms	Δ*E*	rms	Δ*E*
1	C_7^{ax}	0.0	9	0.0	11	0.0	8	0.0	2	0.0	21	0.0	17	0.0	1	0.0	4	0.0
									7	*0.3*	*77*	*29.9*	*54*	*12.5*			*44*	*9.0*
2	β_L	4.1	0	7.9	0	23.6	27	5.2	0	5.9	46	−3.8	0	−1.0	0	3.9	7	4.8
					61	*26.4*												
MAD			5	3.8	6	19.5	18	1.1	1	1.8	34	7.9	9	5.1	1	0.2	5	0.7

[a] Relative are energies in kJ/mol. The best estimates of the relative energies from Table 1 are labeled as ref. Root-mean-square deviation (rms) of the force field torsional angles relative to the MP2 values in deg. MAD indicates the mean absolute deviation over the conformations not marked in italics.

to estimate solvation effects, but clearly better methods are needed for reliably calculating the effect of placing these systems in an aqueous environment. The difference in the PCM solvent corrections is sufficiently large that it is likely that the β_L conformation is the global minimum in solution.

The results of conformational searches with several force fields are shown in Table 2. For each force field is given the rms deviation of the torsional angles relative to the MP2 structure, and the relative energies calculated by the force field, which can be compared with our reference data (labeled ref in Table 2). Except for the OPLS and AMBER99 force fields, the global minimum in the C_7^{ax} region is reproduced with an acceptable accuracy, and all force fields except MM3 reproduced the β_L minimum accurately in terms of the Φ and Ψ angles. The MM2 force field predicts one additional minimum in the ϵ_D region which collapses to the β_L conformation upon MP2/6-31G(d,p) optimization, and it is therefore labeled as conformation **2** in Table 2. The MMFFs, OPLS, AMBER99, and AMOEBA force fields similarly display a third stable conformation, which upon MP2 optimization converges to the C_7^{ax} minimum, and they are listed under conformation **1** in Table 2. These minima should be considered as "artificial" minima on their respective surface. For characterization purposes, we have arranged them in the tables according to the minima to which they collapse upon MP2 optimization, but the large rms deviations in the torsional energy parameters indicate that they have substantial different geometries.

For systems with many conformations it is useful to characterize the performance in terms of a mean absolute deviation (MAD) relative to the reference data. The lack of a clear one-to-one correspondence between the force field and reference data, however, requires some decisions to be made regarding the selection of data. We have decided to classify the force field conformations into four groups: (1) "Good" minima have torsional rms values less than 40° relative to the MP2 structure. (2) "Poor" minima have torsional rms values larger than 40° but collapse to the indicated minimum upon MP2 optimization. (3) "Missing" minima indicate that there is no force field minimum that is connected to the corresponding MP2 conformation. (4) "Artificial" minima indicate that the force field minimum collapses to the indicated conformation upon MP2 optimization, but there is another force field minimum which provides a better representation of the same conformation. These artificial minima are marked with italics in the tables.

Clearly the rms value of 40° for distinguishing between "good" and "poor" minima is somewhat arbitrary. The MAD values over the conformations have only been calculated for the "good" and "poor" minima, i.e., only the best representation has been included for the conformations with multiple associated force field structures.

The ability to correctly predict geometries of local minima is only one parameter for characterizing an accurate potential energy surface. Of similar importance is the ability to predict the energetic ordering of these minima. All of the present force field can reproduce the C_7^{ax} and β_L minima, and our reference calculations place the latter 4 kJ/mol above the global minimum. This value is reasonably reproduced by the AMBER94, MM3, MMFFs, CHARMM27, and AMOEBA force fields with deviations of a few kJ/mol, while the MM2 provides a significantly higher value. The OPLS and AMBER99 force fields reverse the energetic preference of the two structures (Table 2).

Alanine. The presence of an additional methyl group in the alanine system leads to seven different minima at the MP2/aug-cc-pVDZ level, which are labeled C_7^{eq}, β_L, C_7^{ax}, δ_L, α_D, ϵ_D, and δ_D in agreement with the results obtained by Rodriguez et al.[75] The Φ and Ψ torsion angles and relative energies are presented in Table 3. The minimum in the ϵ_D region (conformation **6**) deserves a special comment. Rodriguez et al. located this structure at the HF/3-21G level as

Table 3. Alanine Conformations[a]

				MP2				B3LYP			
conf	region	Φ	Ψ	A	B	ΔZPE	ΔPCM	A	ΔZPE	ΔPCM	rms
1	C_7^{eq}	−82	76	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3
2	β_L	−161	157	6.3	6.0	−1.4	−6.0	3.4	−1.4	−4.7	5
3	C_7^{ax}	74	−54	9.5	10.1	0.7	−2.8	10.7	0.6	−3.0	3
4	α_L	−83	−10	12.8	13.5	−1.0	−9.9	12.1	−1.1	−9.9	41
5	α_D	64	30	19.8	19.3	−0.4	−13.6	23.3	−0.8	−12.0	4
6	ϵ_D	52	−130	19.7	20.5	−0.1	−2.4				
7	δ_D	−165	−38	26.7	27.1	−0.7	−12.3	27.3	−1.0	−10.0	2
MAD				0.6				1.9			10

[a] Relative energies are in kJ/mol. Alanine conformations optimized at the MP2/aug-cc-pVDZ (MP2) and B3LYP/6-31G** (B3LYP) levels. Relative energies from single point calculations with the aug-cc-pVTZ basis set are labeled with A, while the best estimates obtained by basis set extrapolation and CCSD(T) corrections are labeled B. Changes in relative energies due to zero point corrections (ΔZPE) and PCM (ΔPCM) solvation corrections have been calculated at the MP2/6-31G** and B3LYP/6-31G** levels. Root-mean-square deviation (rms) of the B3LYP torsional angles relative to the MP2 values in deg. MAD indicates the mean absolute deviation over the conformations.

Table 4. Force Field Predictions of Alanine Conformations[a]

			AMBER94		MM2*		MM3*		MMFFs		OPLS_2005		AMBER99		CHARMM27		AMOEBA	
conf	region	ref	rms	Δ*E*	rms	Δ*E*	rms	Δ*E*	rms	Δ*E*	rms	Δ*E*	rms	Δ*E*	rms	Δ*E*	rms	Δ*E*
1	C_7^{eq}	0.0	10	0.0	15	0.0	9	0.0	4	0.0	11	0.0	27	0.0	4	0.0	1	0.0
2	β_L	6.0	14	6.3	18	25.3	14	2.4	4	5.9	7	4.1	9	3.8	12	3.8	5	5.1
3	C_7^{ax}	10.1	9	6.2	3	6.4	3	5.6	1	7.7	4	10.4	19	9.3	10	8.6	1	10.5
4	α_L	13.5									51	12.9	42	10.8			28	11.8
5	α_D	19.3							8	19.8							1	18.7
6	ϵ_D	20.5																
7	δ_D	27.1	20	25.8			14	17.8	6	23.7	9	25.6					3	23.5
MAD			13	1.8	12	11.5	10	5.8	5	1.6	8	1.1	24	1.5	9	1.9	7	1.4

[a] Relative energies are in kJ/mol. The best estimates of the relative energies from Table 3 are labeled as ref. Root-mean-square deviation (rms) of the force field torsional angles relative to the MP2 values in deg. MAD indicates the mean absolute deviation over the conformations not marked in italics.

well as with several force fields (AMBER, MM+, BIO+, OPLS, and ECEPP/2) but failed to find it at the semiempirical AM1 level. This conformation was not reported in the work of Yu[28] (MP2/6-311G(d,p)), Bohm[72] (HF/DZP), and MacKerell[70] (MP2/6-31G(d) and MP2/6-311++G(d,p)). We were able to optimize the structure at the HF/6-311++G(d,p) and MP2/6-311++G(d,p) levels, in addition to MP2/aug-cc-pVDZ, but it does not exist at the B3LYP level. A B3LYP/6-31G(d,p) optimization starting from the MP2 structure converged to structure **3** in the C_7^{ax} region.

The MP2/6-31G(d,p) optimized geometry for conformer **4** belongs to the δ_L region (−138°, 23°), but a reoptimization with the aug-cc-pVDZ basis set changes the value of the Ψ angle to −10°, which places the conformation in the α_L region. This value is significantly different from previously published values of ∼40°.[27,70,75] The OPLS value is 41°, while AMOEBA, AMBER99, and B3LYP/6-31G(d,p) give intermediate values (11°−23°), and the remaining force fields fail to locate this structure. This conformation corresponds to a right-handed helix in polypeptides, where the Ψ angle has a value ∼ −40°. The computational results suggest that the Ψ torsional energy surface is very flat in this region, and the exact geometry is therefore sensitive to the level of theory employed.

The B3LYP method gives a good agreement with the MP2 geometries, except for the missing conformation **6** and conformation **4**, which has an rms deviation of 41°. Compared to the reference values for the relative energies, the MP2/aug-cc-pVTZ method gives a MAD value of only 0.6 kJ/mol, while the corresponding B3LYP value is 1.9 kJ/mol. Zero point energy corrections give only minor changes in the relative conformational energies, while the PCM solvent model again preferentially stabilized the high-energy conformations and makes the C_7^{eq} and β_L conformations essentially isoenergetic.

All force fields are able to locate the three lowest energy minima in the C_7^{eq}, β_L, and C_7^{ax} regions (Table 4) with good representations of the geometries, and all except MM2, AMBER99, and CHARMM27 can also account for the highest energy conformation **7** in the δ_D region. Moreover, MMFFs locates conformation **5**, OPLS and AMBER99 locate conformation **4**, but none of the force fields finds conformation **6**. The polarizable AMOEBA force field performs well and locates six out of the seven MP2 minima, of which five have very low rms values for the torsional angles. The relative stabilities of the conformations found by the different force field reproduce fairly well the MP2 results, except that conformation **2** is predicted to be significantly destabilized by the MM2 force field.

Serine. The CH_2OH side chain in serine provides additional conformational flexibility compared to alanine, and the polarity and possibilities for hydrogen-bonding result in a more complicated potential energy surface. A conformational search at the MP2/aug-cc-pVDZ level along the lines described in the Computational Details section resulted in 39 different conformations spanning a 70 kJ/mol energy range (Table 5). The global minimum is a C_7^{eq} type conformation, and it is 13 kJ/mol below the second lowest

Table 5. Serine Conformations[a]

						MP2				B3LYP			
conf	region	Φ	Ψ	χ_1	χ_2	A	B	ΔZPE	ΔPCM	A	ΔZPE	ΔPCM	rms
1	C_7^{eq}	−82	73	55	65	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1
2	C_7^{eq}	−82	68	179	−66	13.5	12.5	−1.5	−6.4	11.3	−1.8	−6.9	3
3	β_L	−157	−173	−167	79	15.2	14.8	0.1	−9.0	12.2	0.4	−10.4	2
4	β_L	−176	165	−92	51	18.2	17.0	−2.5	−5.7	11.9	−2.6	−6.7	4
5	ϵ_D	63	−138	84	−47	18.3	18.0	−2.5	−2.3	14.0	−2.6	−1.8	20
6	β_L	−157	−174	−169	172	19.1	18.4	−3.8	−8.3	15.6	−3.9	−7.6	3
7	C_7^{eq}	−85	69	−54	−179	24.0	23.0	−3.1	−7.6	22.6	−3.3	−7.0	5
8	C_7^{eq}	−85	72	−51	−68	24.2	23.5	−3.0	−11.8	21.2	−3.0	−10.4	4
9	C_7^{ax}	74	−37	84	−57	26.6	24.1	−0.6	−12.0	26.5	−1.1	−12.6	6
10	β_L	−155	177	66	−58	25.5	25.0	−3.9	−15.6	23.3	−4.4	−11.9	2
11	C_7^{eq}	−82	79	−65	51	28.3	26.2	−3.4	−14.9	26.3	−3.4	−13.0	3
12	β_L	−158	162	−175	−76	27.1	27.5	−3.7	−16.0				
13	C_7^{eq}	−104	8	53	170	27.3	27.5	−4.8	−19.2	24.5	−5.2	−17.5	8
14	ϵ_D	60	34	−166	−62	29.3	27.9	−1.8	−21.3	29.3	−2.3	−18.4	3
15	C_7^{ax}	76	−53	−52	82	31.0	29.4	−1.4	−8.0	31.0	−1.5	−7.4	6
16	α_L	−75	−20	−53	−72	31.2	31.2	−2.0	−8.0	28.3	−2.5	−7.3	4
17	C_7^{ax}	77	−50	−58	175	31.4	31.3	−3.2	−12.2	30.4	−3.3	−10.6	3
18	β_L	−161	174	71	−160	32.8	31.9	−5.6	−21.0	29.1	−6.0	−16.6	2
19	C_7^{ax}	70	−29	−161	−43	31.2	32.1	−3.6	−21.1				
20	C_7^{ax}	76	−48	−58	−76	33.4	33.3	−3.0	−16.9	31.5	−3.3	−15.6	2
21	α_L	−75	−18	−54	−178	33.0	33.9	−4.0	−17.1	33.4	−4.2	−14.3	5
22	δ_L	−150	19	−48	−52	34.1	34.1	−2.8	−21.9	31.6	−3.5	−18.5	4
23	C_7^{eq}	−81	89	−173	70	35.2	34.5	−3.7	−24.8	31.9	−4.2	−21.0	4
24	C_7^{ax}	69	−79	−178	−72	35.6	35.4	−1.6	−12.2	35.9	−2.2	−10.5	2
25	ϵ_D	59	−162	−159	−178	35.9	35.4	−2.4	−19.4	36.8	−3.0	−17.1	7
26	α_D	66	31	−55	−178	36.5	36.5	−3.7	−20.9	39.1	−4.4	−17.2	2
27	ϵ_D	57	−159	−165	70	37.4	37.0	−2.2	−23.3	39.0	−2.9	−21.6	7
28	α_D	47	52	55	60	37.3	37.9	−2.0	−23.2	42.6	−3.0	−21.1	4
29	ϵ_L	−68	161	59	−176	40.0	39.4	−4.7	−31.1				
30	δ_D	171	−37	−88	61	40.1	39.9	−2.2	−19.4	36.5	−2.3	−20.0	11
31	δ_D	−159	−69	57	−173	39.7	40.0	−4.4	−21.6	35.7	−4.8	−15.8	9
32	α_D	65	30	−56	−90	40.8	40.6	−3.8	−25.5				
33	α_D	66	26	−55	86	40.9	41.2	−5.1	−23.9	42.4	−5.6	−21.4	3
34	ϵ_D	49	−134	−58	56	42.0	41.3	−2.9	−15.2	44.5	−3.7	−11.9	4
35	δ_D	−167	−26	−172	−41	49.5	49.7	−2.9	−18.0	49.9	−3.3	−18.1	7
36	C_7^{ax}	58	−29	67	173	52.7	53.4	−3.9	−21.3	49.8	−4.3	−19.9	1
37	ϵ_D	37	−120	65	−175	53.2	53.8	−3.9	−27.1	55.2	−4.5	−24.3	4
38	α_L	−65	−37	−175	−172	60.4	61.3	−6.1	−39.7				
39	δ_D	−126	−69	−63	−176	69.9	69.1	−7.9	−39.3				
MAD						**0.6**				**2.2**			**5**

[a] Relative energies are in kJ/mol. Serine conformations optimized at the MP2/aug-cc-pVDZ (MP2) and B3LYP/6-31G** (B3LYP) levels. Relative energies from single point calculations with the aug-cc-pVTZ basis set are labeled with A, while the best estimates obtained by basis set extrapolation and CCSD(T) corrections are labeled B. Changes in relative energies due to zero point corrections (ΔZPE) and PCM (ΔPCM) solvation corrections have been calculated at the MP2/6-31G** and B3LYP/6-31G** levels. Root-mean-square deviation (rms) of the B3LYP torsional angles relative to the MP2 values in deg. MAD indicates the mean absolute deviation over the conformations.

conformation. There are six conformations within 20 kJ/mol of the global minimum and 15 conformations within a 30 kJ/mol window. The differences in zero point energies are small, but the estimates of solvent effects indicate that many of these higher energy conformations may be energetically accessible in solution. The MP2/aug-cc-pVTZ method again provides a good correlation with the accurate reference energies, with a MAD value of 0.6 kJ/mol.

A corresponding conformational search at the B3LYP level produced 33 conformations with the large majority being in good agreement with the MP2 geometries. It is significant, however, that six conformations located at the MP2 level do not exist on the B3LYP energy surface.[76] Most of these missing conformations are relatively high in energy, although solvation potentially could bring some of these down in energy. The MAD for the relative energies of the 33 conformations located at the B3LYP level is 2.2 kJ/mol.

Table 6 shows that the various force fields perform erratically for locating conformations, producing 18−41 different conformation compared to the 39 found at the MP2 level. Not only are there several conformations which do not exist on the force field energy surfaces, but there are also many artificial minima. The MMFFs, for example, not only have a good description of the global minimum, with an rms deviation of only 2° compared to the MP2 geometry, but also have another minimum 10.4 kJ/mol higher in energy

Table 6. Force Field Predictions of Serine Conformations[a]

			AMBER94		MM2*		MM3*		MMFFs		OPLS_2005		AMBER99		CHARMM27		AMOEBA	
conf	region	ref	rms	ΔE	rms	ΔE	rms	ΔE	rms	ΔE	rms	ΔE	rms	ΔE	rms	ΔE	rms	ΔE
1	C_7^{eq}	0.0	5	0.0	10	0.0	5	0.0	2	0.0	6	0.0	13	0.0	3	0.0	4	0.0
							6	*1.5*	*50*	*10.4*			*45*	*19.5*			*45*	*26.1*
2	C_7^{eq}	12.5	8	13.3	14	5.7	5	7.5	8	24.8	5	17.3	17	10.9	5	9.9	7	18.4
									36	*35.7*	*32*	*27.2*	*36*	*22.0*			*36*	*31.3*
									55	*32.4*							*58*	*37.8*
3	β_L	14.8	7	14.4	58	50.8			9	17.0	1	10.7	9	10.8	9	3.0	5	11.8
					72	*50.8*							*66*	*40.5*			*70*	*53.3*
4	β_L	17.0	12	4.6	40	39.0	11	7.0	3	5.6	8	14.2	11	−0.6	7	10.7	5	22.7
					45	*41.2*												
5	ϵ_D	18.0	46	19.5					21	15.7	21	16.2	37	16.4	17	17.2	6	30.8
													64	*23.7*	*43*	*19.6*	*34*	*30.7*
																	65	32.0
6	β_L	18.4	6	13.7	8	27.9	12	−3.9	7	16.4	5	13.3	7	9.8	4	−0.1	9	16.6
									64	*61.1*			*62*	*48.1*				
7	C_7^{eq}	23.0	28	5.9	9	13.1	6	9.7	9	11.4	13	20.0	23	4.5	9	9.8	5	19.2
													46	*33.9*			*68*	*29.1*
8	C_7^{eq}	23.5	29	7.3	13	19.8			9	13.4	13	14.6	23	6.1	6	4.0	5	17.2
9	C_7^{ax}	24.1	13	12.3	11	15.8	6	17.6	10	22.0	8	21.2	14	13.8			10	40.9
																	62	51.8
10	β_L	25.0	2	24.8	11	43.7	16	20.9	9	22.5	5	18.8	3	20.9	44	14.8	22	29.8
			55	*27.8*			*51*	*5.0*	*49*	*22.4*	*34*	*22.1*	*71*	*25.6*			*54*	*34.5*
11	C_7^{eq}	26.2			16	27.2			5	16.2	5	23.5					16	24.9
12	β_L	27.5									6	20.4					10	21.7
13	C_7^{eq}	27.5							39	30.5	39	30.7						
14	ϵ_D	27.9			7	20.1			10	30.3					11	18.5	7	26.7
15	C_7^{ax}	29.4	9	17.2	7	22.9	5	14.0	7	19.7	8	28.0	14	18.5			5	31.3
							14	*17.4*										
16	α_L	31.2																
17	C_7^{ax}	31.3	19	13.4	4	20.3	7	14.8	4	18.9	8	33.7	19	14.5	62	31.6	5	32.7
18	β_L	31.9	6	30.1	20	41.7	17	22.5	9	30.9	7	32.3	4	26.3				
19	C_7^{ax}	32.1	11	16.3	60	30.7	9	13.6	13	29.2	7	25.7	11	17.5	12	22.7	58	40.8
											53	*34.7*						
20	C_7^{ax}	33.3	19	19.5					5	26.5	8	33.8	19	20.7	11	23.5	3	32.1
															52	*23.3*		
21	α_L	33.9															4	33.9
22	δ_L	34.1	17	20.2					50	19.4			21	11.6	7	21.4		
23	C_7^{eq}	34.5	7	37.5			10	27.3									7	37.3
24	C_7^{ax}	35.4	52	30.1	53	22.7	50	18.7			53	38.1	53	34.4			54	43.6
25	ϵ_D	35.4	7	33.8	22	35.0			15	30.1	21	29.1	7	39.6	7	18.7	6	29.9
26	α_D	36.5							4	22.3							8	36.2
									36	35.7								
27	ϵ_D	37.0							51	26.4	21	30.9			11	13.3	5	29.4
28		37.9			18	31.4			21	32.9					11	28.9		
29	ϵ_L	39.4																
30	δ_D	39.9							5	26.1	13	41.7	31	24.9	35	44.3	45	65.1
																	80	42.9
																	63	53.3
31	δ_D	40.0	5	26.0			17	4.8	7	25.0	5	29.7	15	21.2			7	37.6
32	α_D	40.6															4	38.9
33	α_D	41.2															7	36.3
34	ϵ_D	41.3															13	40.7
35	δ_D	49.7					14	33.7	8	55.8								
36	C_7^{ax}	53.4	7	39.1	19	33.2			11	50.3			7	40.3				
37	ϵ_D	53.8													25	44.4	6	59.8
38	α_L	61.3															9	65.3
39	δ_D	69.1																
MAD			15	8.9	21	10.7	13	14.0	13	7.4	13	4.1	17	10.9	16	11.1	11	4.2

[a] Relative energies are in kJ/mol. The best estimates of the relative energies from Table 5 are labeled as ref. Root-mean-square deviation (rms) of the force field torsional angles relative to the MP2 values in deg. MAD indicates the mean absolute deviation over the conformations not marked in italics.

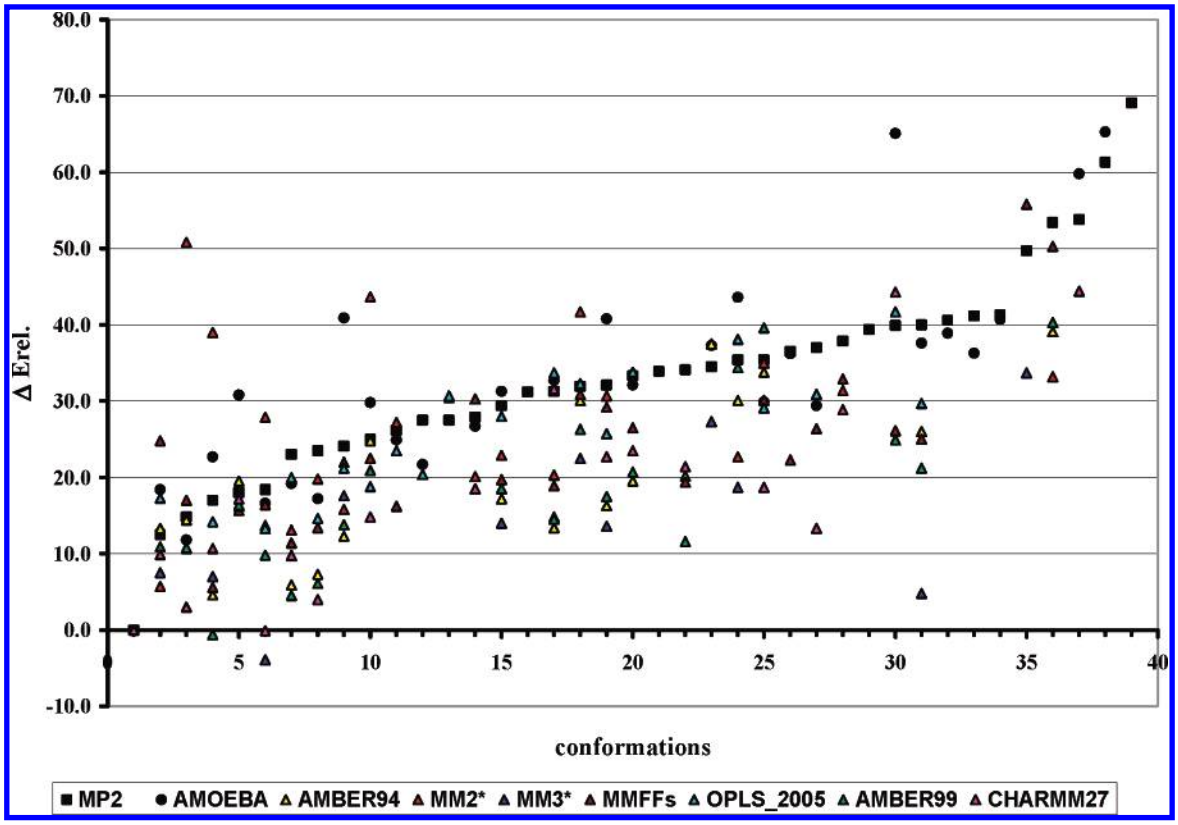

Figure 3. Grapical representation of the relative energies for the serine conformations in Table 6.

with an rms deviation of 50° which converges to the global minimum upon MP2 minimization. Using an rms deviation of 40° as a criterion of a satisfactory representation of the geometries, the MM2, MM3, and CHARMM27 force fields have a similar performance and account satisfactorily for ~15 conformations, the AMBER94, AMBER99, and OPLS reproduce ~20 conformations, while the MMFFs force field gives reasonable representations for 25 of the 39 conformations. The AMOEBA again performs better than the other force fields, being able to locate 27 of the 39 MP2 minima with rms deviation lower than 40°, but also finds 11 artificial minima. The difference in performance is also indicated by the MAD rms values, where the MMFFs, OPLS, and AMOEBA have a value of ~12°, while the other force fields have values of ~17°. The energetic ordering of the force field conformations is somewhat more erratic. The AMOEBA and OPLS force fields perform best and reproduce the reference energies with MAD values of ~4.0 kJ/mol, while the other force fields have MAD values of ~10 kJ/mol. Figure 3 shows a graphical representation of the relative energies of the different force fields as a function of the conformation numbering in Table 6.

Cysteine. The CH_2SH side chain in cysteine has the same torsional degrees of freedom as serine, but the SH group is more polarizable and has less hydrogen bonding ability than the OH group in serine. An MP2 conformational search located 47 conformations of which 14 are within 20 kJ/mol of the global minimum and 28 are within a 30 kJ/mol energy window (Table 7). Estimates of the solvation effect again suggest that many of these will be substantially stabilized in solution, perhaps to the point where the global minimum will change. The computationally less demanding MP2/aug-cc-pVTZ level has a MAD value of 0.7 kJ/mol compared to the reference energies.

The B3LYP method fails to account for six of the 47 conformations, again mostly in the (gas phase) higher energy region. With a couple of exceptions, the geometries of the 41 conformations are in good agreement with the MP2 values. Most of the differences between the B3LYP and MP2 results are in the L-regions, i.e., β_L, α_L, δ_L, and C_7^{eq}. The B3LYP relative energies are in fair agreement (MAD = 2.5 kJ/mol) with the reference values. We note that Zamora and Bombaraso[77,78] reported 47 local minima located by the HF/3-21G model and 42 minima at the B3LYP/6-31G level for this system.

The force field results in Table 8 again show erratic performance. The best force field by far is the AMOEBA, which provides a fair description of 39 of the 47 conformations but also has five artificial minima. The MMFFs is again the best of the fixed charge force fields and accounts for 25 conformations, while the other force fields reproduce between 15 and 24 of the 47 minima. The MAD over the rms in torsion angles is ~20° for the AMBER94, AMBER99, MM2, and MM3 force fields, while the other force fields give values in the range 9−14°. The MAD for the energetic ordering of the conformations is largest for the MM2 and MM3 force fields with values of 10−14 kJ/mol, while the remaining force fields provide values of ~5 kJ/mol. Figure 4 shows a graphical representation of the relative energies of the different force fields as a function of the conformation numbering in Table 8.

Table 7. Cysteine Conformations[a]

						MP2				B3LYP			
conf	region	Φ	Ψ	χ_1	χ_2	A	B	ΔZPE	ΔPCM	A	ΔZPE	ΔPCM	rms
1	C_7^{eq}	−82	64	51	65	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2
2	β_L	−161	173	−161	69	8.8	7.6	−1.1	−9.9	7.9	−2.4	−6.7	4
3	C_7^{eq}	−82	80	−174	−71	9.8	8.9	−0.1	−16.7	6.4	−0.8	−15.6	2
4	C_7^{eq}	−81	64	55	−119	11.7	10.8	0.1	−13.1	10.4	−0.7	−10.7	2
5	δ_L	−128	22	59	71	11.6	11.8	−0.5	−15.9	9.9	−1.1	−15.1	3
6	C_7^{eq}	−86	76	−52	−56	13.5	12.1	−0.8	−15.6	9.9	−1.9	−14.3	2
7	β_L	−164	149	−174	−80	13.8	12.7	−1.6	−23.1	10.8	−2.3	−16.2	5
8	C_7^{eq}	−86	73	−57	−179	16.1	14.8	−1.0	−18.1	12.5	−1.9	−17.1	2
9	C_7^{eq}	−83	77	−68	52	16.8	15.6	−1.0	−19.9	13.7	−2.0	−18.4	4
10	C_7^{eq}	−82	85	−174	78	17.4	15.8	−1.0	−25.1	17.0	−2.0	−21.5	5
11	β_L	−158	179	68	−59	16.5	16.2	−2.2	−24.9	12.9	−3.1	−20.9	3
12	C_7^{ax}	78	−52	−54	81	17.7	17.7	−0.4	−19.5	16.6	−1.0	−13.8	3
13	β_L	−171	157	54	54	18.1	17.7	−1.7	−25.4	17.4	−2.3	−21.7	5
14	C_7^{ax}	79	−51	−56	−70	18.8	18.3	0.0	−21.8	16.6	−1.1	−17.5	4
15	δ_L	−144	23	58	−95	20.8	20.4	−1.9	−27.3	17.2	−2.4	−24.1	6
16	β_L	−140	148	−54	−32	22.7	21.9	−1.4	−24.8	22.0	−1.9	−22.4	5
17	α_L	−74	−22	−55	−60	22.1	22.5	−1.3	−22.4	21.4	−2.1	−18.0	32
18	C_7^{ax}	78	−50	−52	−162	23.0	23.0	−1.0	−22.4	24.5	−1.0	−18.3	19
19	C_7^{ax}	72	−74	−174	−68	23.8	23.5	0.4	−19.7				
20	β_L	−167	165	66	−166	24.0	23.6	−0.8	−31.6	27.2	−2.6	−27.8	2
21	C_7^{ax}	73	−68	−179	59	25.8	25.0	−1.2	−27.3	28.7	−1.3	−22.7	3
22	α_D	67	29	−57	−75	26.1	25.5	−0.6	−27.3	25.1	−2.4	−21.9	2
23	α_D	68	26	−57	79	25.9	25.6	−1.7	−27.6	22.4	−2.4	−20.1	3
24	α_L	−78	−17	−61	165	27.5	27.4	−0.7	−27.1				
25	β_L	−138	146	−63	173	30.0	28.4	−0.5	−34.2	33.2	−3.2	−31.1	3
26	δ_D	−176	−44	47	−77	27.8	28.8	−0.8	−17.1				
27	α_D	70	26	−51	−162	29.5	29.5	−2.2	−28.6	31.5	−2.0	−22.7	6
28	α_D	60	39	−157	−67	30.1	29.9	−0.9	−26.8	25.3	−2.0	−26.3	4
29	C_7^{ax}	74	−64	−173	−178	31.4	30.9	−0.5	−28.0	34.8	−1.4	−23.8	5
30	ϵ_D	58	−161	−158	56	32.4	31.0	−1.2	−20.2	32.1	−2.2	−12.5	5
31	δ_D	−164	−34	60	87	31.7	31.9	−1.4	−23.4	30.3	−2.4	−16.6	2
32	C_7^{ax}	71	−34	98	−65	33.6	32.9	−2.7	−20.2	34.5	−0.6	−15.6	9
33	α_L	−74	−23	−165	−42	32.7	33.0	−1.3	−27.0				
34	ϵ_D	58	−159	−151	−166	35.1	33.8	0.0	−19.1	37.5	−2.3	−10.9	4
35	ϵ_D	40	−134	72	−19	34.6	35.3	−2.1	−14.6				
36	α_D	44	52	51	59	35.4	35.6	0.2	−26.4	42.6	−1.8	−18.6	3
37	C_7^{ax}	52	−26	58	55	39.2	39.5	−2.1	−18.7	38.3	−1.5	−15.9	3
38	δ_D	−154	−69	175	60	40.9	40.1	0.1	−35.9	43.5	−3.9	−30.4	5
39	α_D	58	42	−164	108	41.4	40.3	−0.6	−40.1	38.2	−2.5	−34.8	2
40	δ_D	−163	−44	−173	−48	40.8	40.6	−0.1	−29.4	47.2	−2.4	−26.4	4
41	ϵ_D	42	−129	62	176	41.6	41.0	−1.7	−25.9	41.3	−1.3	−12.9	5
42	α_L	−70	−30	−180	165	44.7	44.6	−0.7	−42.3				
43	δ_D	−158	−58	175	171	47.3	46.8	−2.8	−36.7	46.1	−3.0	−32.2	7
44	α_D	51	43	51	−161	47.6	46.9	0.0	−30.7	54.0	−3.7	−26.7	6
45	δ_D	−134	−68	−55	−30	47.6	48.8	−1.1	−36.2	46.4	−2.1	−28.1	27
46	δ_D	−173	−41	−118	64	50.0	49.0	−0.5	−27.9				
47	δ_D	−135	−69	−64	175	55.8	54.3	−0.8	−39.6	51.0	−3.8	−34.2	4
MAD						0.7				2.5			6

[a] Relative energies are in kJ/mol. Cysteine conformations optimized at the MP2/aug-cc-pVDZ (MP2) and B3LYP/6-31G** (B3LYP) levels. Relative energies from single point calculations with the aug-cc-pVTZ basis set are labeled with A, while the best estimates obtained by basis set extrapolation and CCSD(T) corrections are labeled B. Changes in relative energies due to zero point corrections (ΔZPE) and PCM (ΔPCM) solvation corrections have been calculated at the MP2/6-31G** and B3LYP/6-31G** levels. Root-mean-square deviation (rms) of the B3LYP torsional angles relative to the MP2 values in deg. MAD indicates the mean absolute deviation over the conformations.

Discussion

The present results for four amino acid systems show that the B3LYP method cannot locate all minima found at the MP2 level. One would expect similar problems for density functional methods in general, as they lack a proper description of dispersion interactions. The missing conformations tend to be among the higher energy ones in the gas phase, but whether this is a general trend will have to await results for a larger selection of systems. For the minima that actually exist on the B3LYP surface, the geometries and

Table 8. Force Field Predictions of Cysteine Conformations[a]

			AMBER94		MM2*		MM3*		MMFFs		OPLS_2005		AMBER99		CHARMM27		AMOEBA	
conf	region	ref	rms	ΔE	rms	ΔE	rms	ΔE	rms	ΔE	rms	ΔE	rms	ΔE	rms	ΔE	rms	ΔE
1	C_7^{eq}	0.0	7	0.0	11	0.0	10	0.0	4	0.0	5	0.0	14	0.0	2	0.0	4	0.0
2	β_L	7.6	53	11.2	57	26.3	59	−3.0	4	14.6	5	7.1	4	10.0	4	8.1	3	10.6
													53	*9.2*	*53*	*7.8*		
															72	*28.2*		
3	C_7^{eq}	8.9	8	8.4	11	4.6	4	1.5	3	15.9	7	12.1	18	10.1	5	5.3	4	7.5
			51	*18.1*	*51*	*10.0*	*53*	*5.2*					*52*	*23.0*	*52*	*12.9*		
4	C_7^{eq}	10.8			15	13.0	18	13.2	19	17.3	21	17.6	42	8.0	17	20.5	8	17.1
					37	*5.4*			*37*	*9.8*	*36*	*9.8*			*36*	*10.5*		
5	δ_L	11.8							53	16.8	6	10.9	3	9.7				
											53	16.0						
6	C_7^{eq}	12.1	10	19.1	13	0.2	9	−2.0	7	3.6	11	12.8	23	2.8	7	2.8	4	9.5
7	β_L	12.7	10	19.1					3	21.3	8	11.2	14	16.0	10	11.8	4	13.0
									52	*28.2*			*61*	*21.1*			*40*	*22.0*
8	C_7^{eq}	14.8	27	1.8	10	0.8	5	−2.7			10	16.1	21	2.0	4	2.2	4	13.1
9	C_7^{eq}	15.6	6	9.4	15	4.8	36	5.5					16	11.0	4	6.9	2	16.4
																	61	21.7
10	C_7^{eq}	15.8	10	17.4			10	7.5	9	23.6			16	27.1	8	15.5	5	19.6
																	41	22.0
11	β_L	16.2	4	18.7	7	28.2			8	15.3	9	11.4	3	17.2	6	12.4		
											61	*37.3*	*87*	*22.1*				
12	C_7^{ax}	17.7	9	13.1	3	6.3	5	1.1	6	10.0	9	23.1	15	16.4	10	12.9	3	23.9
					64	*39.9*	*85*	*31.4*	*68*	*38.2*								
13	β_L	17.7	15	18.6	30	29.2	11	4.5	6	17.0	12	11.9	13	15.9			11	17.0
14	C_7^{ax}	18.3					86	42.3					18	16.4	9	16.1	5	22.5
15	δ_L	20.4	61	6.6			13	10.4					46	12.2			11	25.6
																	47	23.9
16	β_L	21.9	41	12.9					44	16.6			42	10.9	35	9.7	4	21.8
17	α_L	22.5											34	15.1			8	22.7
18	C_7^{ax}	23.0	20	9.3			11	0.8	70	46.8	69	55.0	54	12.4	14	13.2	13	26.5
19	C_7^{ax}	23.5	26	14.6	10	7.6	12	2.5	16	20.4	19	16.9	27	17.8	7	14.8	13	17.4
20	β_L	23.6							6	22.1	10	17.8	10	19.4	11	15.0	9	25.3
21	C_7^{ax}	25.0	85	33.1			8	8.1	9	26.4	12	22.4	13	28.5	6	19.5	82	42.5
									82	*46.0*	*78*	*45.3*	*85*	*36.4*	*82*	*35.2*		
22	α_D	25.5															6	26.6
23	α_D	25.6							2	17.0							4	26.7
24	α_L	27.4															7	28.2
25	β_L	28.4	9	28.2									5	27.4				
26	δ_D	28.8	20	26.8	5	31.6			2	23.9	11	22.7	10	24.5			7	28.0
									21	*27.2*								
27	α_D	29.5															14	30.1
28	α_D	29.9							10	27.4							4	24.5
									52	*39.1*								
29	C_7^{ax}	30.9	6	20.9			6	5.6					55	27.8	6	18.2	5	26.1
30	ϵ_D	31.0	9	38.2							18	28.9	9	45.7			6	26.2
																	74	36.3
31	δ_D	31.9					8	19.5	6	29.4					15	30.3	5	28.6
							48	*14.2*	*50*	*26.0*								
32	C_7^{ax}	32.9	17	19.1	12	21.5			14	31.2	12	27.2	18	22.4	20	23.7	6	35.5
33	α_L	33.0															6	28.1
34	ϵ_D	33.8	9	33.8	49	43.9			16	34.9	51	37.4	9	41.3			7	28.4
35	ϵ_D	35.3															6	30.2
36	α_D	35.6							18	41.2							6	33.8
37	C_7^{ax}	39.5			15	20.4	15	21.9			12	38.5			16	31.4	7	32.9
38	δ_D	40.1	8	35.0	7	36.2	11	18.1	13	43.7	7	37.4	10	34.6				
39	α_D	40.3															18	40.4
40	δ_D	40.6			11	40.5			6	43.0	2	39.2						
41	ϵ_D	41.0															3	41.3
42	α_L	44.6															7	40.6
43	δ_D	46.8							9	50.1	5	44.3	4	39.6			3	44.6
44	α_D	46.9															5	54.7
45	δ_D	48.8					15	30.1	13	42.4							8	42.1
							48	32.3										
46	δ_D	49.0							11	39.7	11	46.7						
47	δ_D	54.3																
MAD			20	6.3	17	10.0	18	13.9	14	5.4	14	4.6	19	5.4	10	6.6	9	3.1

[a] Relative energies are in kJ/mol. The best estimates of the relative energies from Table 7 are labeled as ref. Root-mean-square deviation (rms) of the force field torsional angles relative to the MP2 values in deg. MAD indicates the mean absolute deviation over the conformations not marked in italics.

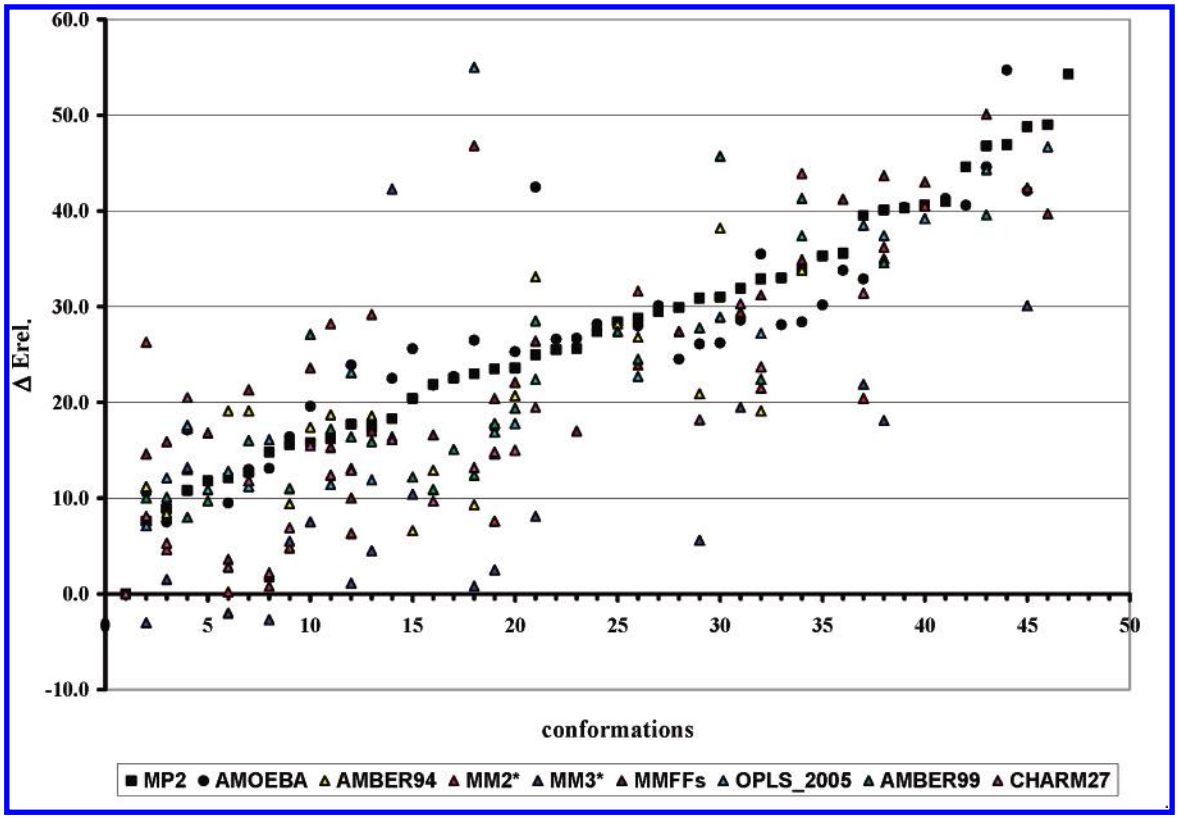

Figure 4. Grapical representation of the relative energies for the cysteine conformations in Table 8.

relative energies are mostly in good agreement with the MP2 results, although there certainly are exceptions. Inclusion of zero point energies only makes small changes in relative energies, typically a few kJ/mol. Estimates of the solvation effect with the PCM model indicate that the energetic ordering in solution may be substantially different than in the gas phase. For these polar systems with hydrogen-bonding capabilities, however, it is questionable whether a continuum solvation model can satisfactorily account for solvation.

Figure 5 shows a graphical representation of the performance of eight different force fields and the B3LYP model against the reference data consisting of all 95 conformations for the four systems. The MM2, MM3, CHARMM27, AMBER94, AMBER99, and OPLS force fields perform almost at par but are only able to satisfactorily reproduce the geometries of approximately half of the conformations. The MMFFs performs somewhat better and is the best of the traditional fixed charge force fields. The more recent AMOEBA force field, which includes multipole moments and polarizabilities, represents a significant improvement and performs almost as well as the B3LYP method. Nevertheless, the AMOEBA force field only accounts for ~80% of the conformations and, in addition, has ~20% artificial minima which are not present on the MP2 energy surface.

Our reference energies are derived from basis set extrapolated MP2 results combined with an additive correction for higher order correlation effects. Compared to these results, the MP2/aug-cc-pVTZ level provides a MAD for relative energies of only 0.7 kJ/mol, indicating that this level of theory is a good compromise between accuracy and computational efficiency and capable of providing relative conformational energies accurate to ~1 kJ/mol.

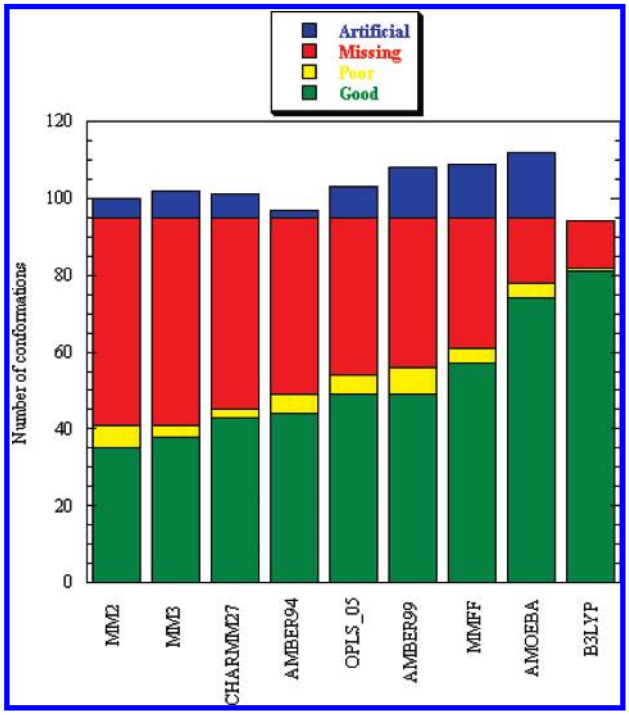

Figure 5. Performance of different force fields and the B3LYP method for reproducing MP2 conformations for all four systems. See the text for the definition of Good, Poor, Missing, and Artificial conformations.

Figure 6 shows the MAD values for the rms torsional angles and relative energies for all four systems for the "good" and "poor" conformations in Figure 5. The MM2

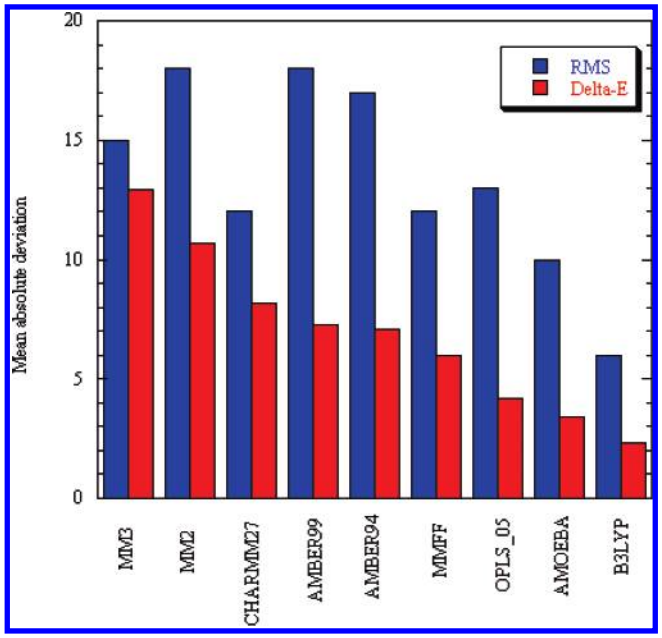

Figure 6. Mean absolute deviations for root-mean-square (rms) differences in the torsional angles and relative energies (Delta-E) of different force fields and the B3LYP method for all conformations over all four systems. rms values in degrees and Delta-E values in kJ/mol.

and MM3 force fields provide the poorest performance, with the CHARMM27, AMBER99, and AMBER94 being slightly better for reproducing relative energies. The MMFFs and OPLS force fields are the best of the fixed partial charge methods, with the former performing best for geometries and the latter performing slightly better for relative energies. The polarizable AMOEBA force field provides the best results with a typical rms error in the torsional angles of 10° and a deviation of 3.4 kJ/mol for relative energies. Nevertheless, the inability of this force field to describe ~20% of the conformations, the presence of artificial minima, and a mean rms deviation for the torsional angles of 10° for those conformations that can be described indicate further room for improvement.

The longer term goal is to use data of the present type for developing new force fields with improved capabilities for reproducing the potential energy surface of, e.g., proteins. For this purpose a decision must be made regarding which reference data to use. Including both solvation and zero point energies in the reference data produces a force field where these effects are absorbed in the parameters. Alternatively, the force field can be parametrized against data without solvent and/or zero point energies, and these effects can then be calculated explicitly within the force field model, for example using the GB/SA solvent model[79] or by explicit solvation. The use of continuum solvent models, however, is only expected to produce qualitative solvent effects, as for example hydrogen bonding is neglected in these models, and using explicit solvation with accurate electronic structure methods is computationally expensive. Zero point energies, on the other hand, are cumbersome to calculate within a force field environment. A viable strategy could be to parametrize the force field for reproducing gas-phase results including zero point energies and subsequently parametrize the interaction with explicit solvent. Since zero point energies are a relatively minor correction, the parametrization could also be done using just electronic energies.

The present work has focused on locating stable conformations, i.e., minima on the potential energy surface for a few simple systems. For parametrization of force field it is necessary to extend the present work to more diverse systems, i.e., other amino acids and larger peptide models. For modeling the dynamics it is also of importance to be able to describe the energetics of interconversion between conformations, i.e., reproduce geometries and stabilities of transition structures connecting minima. Such extensions will be considered at a later date.

Acknowledgment. This work was supported by grants from the Danish Natural Science Research Council, the Danish Center for Scientific Computing, an EU Marie Curie training grant, and the Grant Agency of the Academy of Sciences of the Czech Republic (IAA400550702).

Supporting Information Available: MP2/aug-cc-pVDZ optimized geometries and associated energies for all 95 conformations. The material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) http://www.rcsb.org (accessed April 1, 2007).

(2) Fernandez, C.; Wider, G. *Mod. Mag. Res.* **2006**, *1*, 483. Candel, A. M.; Conejero-Lara, F.; Martinez, J. C.; van Nuland, N. A. J.; Bruix, M. *FEBS Lett.* **2007**, *581*, 687. Richter, B.; Gsponer, J.; Varnai, P.; Salvatella, X.; Vendruscolo, M. J. *Biomol. NMR* **2007**, *37*, 117.

(3) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471. Tse, J. S. *Ann. Rev. Phys. Chem.* **2002**, *53*, 249.

(4) Schlick, T. *Molecular modeling and simulation*; Springer: 2002.

(5) Burkert, U.; Allinger, N. L. *Molecular Mechanics*; Caserio, M. S., Ed.; ACS Monograph 177; American Chemical Society: Washington, DC, 1982.

(6) Jensen, F. In *Introduction to Computational Chemistry*, 2nd ed.; Wiley: Chichester, England, 2006; Chapter 2, pp 22−77.

(7) Gundertofte, K.; Liljefors, T.; Norrby, P. O.; Pettersson, I. *J. Comput. Chem.* **1996**, *17*, 429. Mackerell, A. D., Jr.; Feig, M.; Brooks, C. L., III *J. Am. Chem. Soc.* **2004**, *126*, 698. Buck, M.; Bouguet-Bonnet, S.; Pastor, R. W.; Mackerell, A. D., Jr. *Biophys. J.* **2006**, *90*, L36. Wang, Z.-X.; Zhang, W.; Wu, C.; Lei, H.; Cieplak, P.; Duan, Y. *J. Comput. Chem.* **2006**, *27*, 781.

(8) Koch, U.; Popelier, P. L. A.; Stone, A. J. *Chem. Phys. Lett.* **1995**, *238*, 253. Koch, U.; Stone, A. J. *J. Chem. Soc., Faraday Trans.* **1996**, *92*, 1701.

(9) Halgren, T. A.; Damm, W. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236.

(10) Rasmussen, T. D.; Jensen, F. *Mol. Simul.* **2004**, *30*, 801. Mackerell, A. D., Jr. *J. Comput. Chem.* **2004**, *25*, 1584. Mackerell, A. D., Jr. *Ann. Rep. Comp. Chem.* **2005**, *1*, 91.

(11) Patel, S.; Brooks, C. L., III *J. Comput. Chem.* **2006**, *25*, 1. Patel, S.; MacKerell, A. D., Jr.; Brooks, C. L., III *J. Comput. Chem.* **2006**, *25*, 1504. Maple, J. R.; Cao, Y.; Damm, W.; Halgren, T. A.; Kaminski, G. A.; Zhang, L. Y.; Friesner, R. A. *J. Chem. Theor. Comput.* **2005**, *1*, 694.

(12) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 03*; Gaussian, Inc.: Pittsburgh, PA, 2003.

(13) 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.

(14) Ponder, J. W. *Tinker*, *Version 4.2*; Biochemistry & Molecular Biophysics, Washington University School of Medicine: St. Louis, MO, 2004.

(15) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.

(16) Alinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.

(17) Allinger, N. L.; Yan, L. *J. Am. Chem. Soc.* **1993**, *115*, 11918.

(18) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490.

(19) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.

(20) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.

(21) Foloppe, N.; MacKerell, A. D., Jr. *J. Comput. Chem.* **2000**, *21*, 86.

(22) Ponder, J. W.; Case, D. A. *Adv. Prot. Chem.* **2003**, *66*, 27.

(23) Karton, A.; Martin, J. M. L. *Theor. Chem. Acc.* **2006**, *115*, 330.

(24) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639.

(25) Amovilli, C.; Barone, V.; Cammi, R.; Cances, E.; Cossi, M.; Mennuci, B.; Pomelli, C. S.; Tomasi, J. *Adv. Quantum Chem.* **1999**, *32*, 227.

(26) Moon, S.; Case, D. A. *J. Comput. Chem.* **2006**, *27*, 825.

(27) Hudáky, I.; Hudáky, P.; Perczel, A. *J. Comput. Chem* **2004**, *25*, 1522.

(28) Yu, C.-H.; Norman, M. A.; Schäfer, L.; Ramek, M.; Peeters, A.; van Alsenoy, C. *J. Mol. Struct.* **2001**, *567*, 361.

(29) Černohorský, M.; Vaultier, M.; Koča, J. *J. Mol. Struct.* **1999**, *489*, 213.

(30) Brijbassi, S. U.; Sahai, M. A.; Setiadi, D. H.; Chass, G. A.; Penke, B.; Csizmadia, I. G. *J. Mol. Struct.* **2003**, *666*, 291.

(31) Adamo, C.; Dillet, V.; Barone, V. *Chem. Phys. Lett.* **1996**, *263*, 113.

(32) Chakraborty, D.; Manogaran, S. *J. Phys. Chem. A* **1997**, *101*, 6964.

(33) Kaschner, R.; Hohl, D. *J. Phys. Chem. A* **1998**, *102*, 5111.

(34) Gronert, S.; O'Hair, R. A. J. *J. Am. Chem. Soc.* **1995**, *117*, 2071.

(35) Jarmelo, S.; Lapinski, L.; Nowak, M. J.; Carey, P. R.; Fausto, R. *J. Phys. Chem. A* **2005**, *109*, 5689.

(36) Gong, X.; Zhou, Z.; Du, D.; Dong, X.; Liu, S. *Int. J. Quantum Chem.* **2005**, *103*, 105.

(37) Bogar, F.; Szekeres, Z.; Bartha, F.; Ladik, J. *Phys. Chem. Chem. Phys.* **2005**, *7*, 2965.

(38) Fernandez-Ramos, A.; Cabaleiro-Lago, E.; Hermida-Ramón, J. M.; Martínez-Núnez, E.; Pena-Gallego, A. *J. Mol. Struct.* **2000**, *498*, 191.

(39) Aleman, C.; Puiggali, J. *J. Phys. Chem. B* **1997**, *101*, 3441.

(40) Antohi, O.; Naider, F.; Sapse, A.-M. *J. Mol. Struct.* **1996**, *360*, 99.

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

(42) Gould, I. R.; Cornell, W. D.; Hillier, I. H. *J. Am. Chem. Soc.* **1994**, *116*, 9250.

(43) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 5908.

(44) Czaszar, A. G. *J. Phys. Chem.* **1996**, *100*, 3541.

(45) Gang, Z.; Chunsheng, D.; Zhengyu, Z.; Qunyan, W.; Jinfeng, L. *J. Mol. Struct.* **2005**, *765*, 143.

(46) Korter, T. M.; Balu, R.; Campbell, M. B.; Beard, M. C.; Gregurick, S. K.; Heilweil, E. J. *Chem. Phys. Lett.* **2006**, *418*, 65.

(47) Broda, M. A.; Siodłak, D.; Rzeszotarska, B. *J. Pept. Sci.* **2005**, *11*, 546.

(48) Linder, R.; Nispel, M.; Häber, T.; Kleinermanns, K. *Chem. Phys. Lett.* **2005**, *409*, 260.

(49) Gerlach, A.; Unterberg, C.; Fricke, H.; Gerhards, M. *Mol. Phys.* **2005**, *103*, 1521.

(50) Toroz, D.; van Mourik, T. *Mol. Phys.* **2006**, *104*, 559. van Mourik, T.; Karamertzanis, P. G.; Price, S. L. *J. Phys. Chem. A* **2006**, *110*, 8. Holroyd, L. F.; van Mourik, T. *Chem. Phys. Lett*. **2007**, *442*, 42.

(51) Gresh, N.; Kafafi, S. A.; Truchon, J.-F.; Salahub, D. R. *J. Comput. Chem.* **2004**, *25*, 823.

(52) Pieshoff, C. D.; Ali, F. E.; Bean, J. W.; Calvo, R.; D'Ambrosio, C. A.; Eggleston, D. S.; Hwang, S. M.; Kline, T. P.; Koster, P. F.; Nichols, A.; Powers, D.; Romoff, T.; Samanen, J. M.; Stadel, J.; Vasko, J. A.; Kopple, K. D. *J. Med. Chem.* **1992**, *35*, 3962.

(53) Reed, J.; Hull, W. E.; von der Lieth, C.-W.; Kübler, D.; Suhai, S.; Kinyel, V. *Eur. J. Biochem.* **1988**, *178*, 141.

(54) Johnson, W. C., Jr.; Pagano, T. G.; Basson, T. C.; Madri, J. A.; Gooley, P.; Armitage, I. M. *Biochemistry* **1993**, *32*, 268.

(55) Tobias, D. J.; Brooks, C. L., III *J. Phys. Chem.* **1992**, *96*, 3864.

(56) Tobias, D. J.; Sneddon, S. F.; Brooks, C. L., III *J. Mol. Biol.* **1990**, *216*, 783.

(57) Wei, D.; Guo, H.; Salahub, D. R. *Phys. Rev. E* **2001**, *64*, 011907.

(58) Shirley, W. A.; Brooks, C. L., III *Proteins* **1997**, *28*, 59.

(59) Sajot, N.; Garnier, N.; Genest, M. *Theor. Chem. Acc.* **1999**, *101*, 67.

(60) Feig, M.; MacKerell, A. D., Jr.; Brooks, C. L., III *J. Phys. Chem. B* **2002**, *107*, 2831.

(61) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999.

(62) Park, H. S.; Kim, C.; Kang, Y. K. *Biopolymers* **2002**, *63*, 298.

(63) Ono, S.; Nakajima, N.; Higo, J.; Nakamura, H. *J. Comput. Chem.* **2000**, *21*, 748.

(64) Bashford, D.; Case, D. A.; Choi, C.; Gippert, G. P. *J. Am. Chem. Soc.* **1997**, *119*, 4964.

(65) Gresh, N.; Tiraboschi, G.; Salahub, D. R. *Biopolymers* **1998**, *45*, 405.

(66) Gould, I. R.; Kollman, P. A. *J. Phys. Chem.* **1992**, *96*, 9255.

(67) Möhle, K.; Hofmann, H.-J. *J. Mol. Struct.* **1995**, *339*, 57.

(68) Hermida-Ramón, J. M.; Brdarski, S.; Karlström, G.; Berg, U. *J. Comput. Chem.* **2003**, *24*, 161.

(69) Yang, Z. Z.; Zhang, Q. *J. Comput. Chem.* **2006**, *27*, 1.

(70) Mackerell, A. D., Jr.; Feig, M.; Brooks, C. L., III *J. Comput. Chem.* **2004**, *25*, 1400.

(71) Some authors chose a different labeling where β_L is equivalent to C5, α_D is equivalent to α_L, and α_L is equivalent to α_R.

(72) Bohm, H. J.; Brode, S. *J. Am. Chem. Soc.* **1991**, *113*, 7129.

(73) Iwaoka, M.; Okada, M.; Tomoda, S. *Theochem* **2002**, *586*, 111.

(74) Barone, V.; Adamo, C.; Lelj, F. *J. Chem. Phys.* **1995**, *102*, 364.

(75) Rodriguez, M. A.; Baldoni, H. A.; Suvire, F.; Vásquez, R. N.; Zamarbide, G.; Enriz, R. D.; Farkas, O.; Perczel, A.; McAllister, M. A.; Torday, L.; Papp, J. G.; Csizmadia, I. G. *J. Mol. Struct.* **1998**, *455*, 275.

(76) Perczel reported the presence of 44 local minima at the HF/3-21G level for the HCO-Ser-NH_2 system: Perczel, A.; Farkas, Ö.; Csizmadia, I. G. *J. Am. Chem. Soc.* **1996**, *118*, 7809.

(77) Zamora, M. A.; Baldoni, H. A.; Bombaraso, J. A.; Mak, M. L.; Perczel, A.; Farkas, O.; Enriz, R. D. *J. Mol. Struct.* **2001**, *540*, 271.

(78) Bombaraso, J. A.; Zamora, M. A.; Baldoni, H. A.; Enriz, R. D. *J. Phys. Chem. A* **2005**, *109*, 874.

(79) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.

CT700082F