Molecular Mechanics Parameters and Conformational Free Energies of Proline-Containing Peptides

D. Quentin McDonald and W. Clark Still*

Department of Chemistry, Columbia University, New York, New York 10027

Received October 2, 1995[⊗]

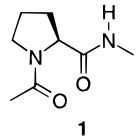
A new set of AMBER* parameters for proline-containing peptides is reported based on the results of high level *ab initio* calculations for a model proline dipeptide, **1**. These new parameters correct a number of deficiencies in the previous parameters for proline, and Monte Carlo/stochastic dynamics (MC/SD) free energy simulations for **1** with the new parameter set give a good agreement with experiment for both the degree of internal hydrogen bonding and the population of the *cis* isomer. The new parameter set is also applied to a conformational study of the sequence Ac-(L)Pro-X-NHMe in $CHCl_3$, where X = Gly, (D/L)Ala, (D/L)Val, (D/L)Ser, and (D/L)Asn(*N*-Me). β-Turn formation is favored by alternating chirality in our simulations so the D-residues in the "X" position always show higher β-turn populations. The enthalpic preference for β-turns in $CHCl_3$ is also reproduced in our simulations. Calculated absolute enthalpies are lower than those found experimentally, but the simulations do correctly reproduce the trends in enthalpic preference for β-turn formation with X = Gly > Ala > Val. Finally the new parameter set is applied to a conformational search of cyclo((D)Pro-(L)Pro-(D)Pro-(L)Pro) and is shown to reproduce the experimentally observed structure having alternating *cis* and *trans* amide bonds.

Introduction

Proline plays an important role in directing the secondary structure of polypeptides. While the high frequency of occurrence of proline in the $i + 1$ position of reverse-turn motifs in proteins has been recognized for some time,[1] it has only recently been reported that proline-containing sequences of short linear peptides can also have significant secondary structure in water.[2] Recent studies of the binding preferences of SH3 domains of proteins important in cellular signal transduction show a high preference for proline-containing sequences.[3] Biological properties aside, the proline Φ-angle is highly constrained by the five-membered pyrrolidine ring and contributes a conformational rigidifying effect to proline-containing peptides. Such a contribution makes proline unique among the genetically encoded amino acids.

We recently described a reparameterization of the original AMBER force field for peptides,[4] a set of parameters we designate as the AMBER* force field. That parameterization was based on high level *ab initio* calculations using the glycine and alanine dipeptides as models for all natural α-amino acids. Subsequently, we decided to specifically reparameterize AMBER* for proline for the following reasons. First, proline is the only encoded amino acid having a secondary amine. Secondly, proline has a constrained Φ angle that makes its conformational preferences very different from other amino acids. Thirdly, because the ring of proline is five-membered, cyclic torsion angles are constrained to values around 0°, angles which in acyclics and larger rings are commonly associated with energy maxima, not minima. Thus, we felt it was likely that a specific set of parameters for proline would be necessary to accurately model its derivatives.

In this paper we report a new set of molecular mechanics parameters for proline derivatives that are based upon high level *ab initio* calculations on *N*-acetylproline methylamide (**1**). Using these parameters, we carry out free energy simulations of **1** both in solution



1

and *in vacuo*. We also apply the new parameter set to a series of dipeptides of the form Pro-X where the thermodynamics of β-turn formation have been measured in organic solutions. Finally we show that the new parameters are able to reproduce the experimentally observed conformation of a cyclic tetrapeptide cyclo((D)Pro-(L)Pro-(D)Pro-(L)Pro). In all cases, the new quantum mechanically derived parameter set shows a significant improvement over our previous alanine-based parameter set in their ability to reproduce experimental observations on derivatives of proline.

Methods

All molecular mechanics calculations were performed using the MacroModel/BatchMin 5.0 program with its associated united-atom and all-atom AMBER* force fields. Except for the new force field parameters described in the text, all proline parameters (*e.g.* atomic partial charges, stretch and bend parameters) were the

[⊗] Abstract published in *Advance ACS Abstracts,* February 1, 1996.

(1) Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Protein Chem.* **1985**, *37*, 1. Chou, P. Y.; Fasman, G. D. *Adv. Enzymol.* **1978**, *47*, 45.

(2) Montelione, G. T.; Arnold, E.; Meinwald, Y. C.; Stimson, E. R.; Denton, J. B., Huang, S.-G.; Clardy, J.; Scheraga, H. A. *J. Am. Chem. Soc.* **1984**, *106*, 7946. Dyson, H. J.; Rance, M.; Houghten, R. A.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1988**, *201*, 161. Imperiali, B.; Fisher, S. L.; Moats, R. A.; Prins, T. J. *J. Am. Chem. Soc.* **1992**, *114*, 3182. Yao, J.; Brüschweiler, R.; Dyson, H. J.; Wright, P. E. *J. Am. Chem. Soc.* **1994**, *116*, 12051.

(3) Chen, J. K.; Schreiber, S. L. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 953. Simon, J. A.; Schreiber, S. L. *Chem. Biol.* **1995**, *2*, 53. Feng, S.; Chen, J. K.; Yu, H.; Simon, J.; Schreiber, S. *Science* **1994**, *266*, 1241.

(4) McDonald, D. Q.; Still, W. C. *Tetrahedron Lett.* **1992**, *33*, 7743.

(5) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230. 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.

0022-3263/96/1961-1385$12.00/0

Table 1. Amide Geometry and *Cis*−*Trans* Relative Energy for 1 Calculated with HF/6-31+G*, Original All Atom AMBER*, and New All Atom AMBER*

	ω-angle, deg	nitrogen improper torsion angle	relative energy (kcal/mol)
trans ($\Psi = 74.3$)			
original AMBER*	−178.1	−177.1	0.00
HF/6-31+G*//HF/6-31+G*	−172.8	−167.2	0.00
new AMBER*	−171.2	−165.9	0.00
cis ($\Psi = -9.1$)			
original AMBER*	−1.5	−177.7	5.03
HF/6-31+G*//HF/6-31+G*	9.3	−164.2	2.38, 2.28[b]
new AMBER*	10.2	-161.3	2.26

[a] The nitrogen improper torsion is defined as: $C_{carbonyl}-N-C_{\delta}-C_{\alpha}$. [b] Calculated at the MP2/6-31+G*//HF/6-31+G* level of theory.

same as those described by Kollman et al.[5] AMBER* force field equations are also identical to those described by Kollman et al. for AMBER. For solution phase calculations, the GB/SA continuum models for chloroform or water were used.[6] When comparing molecular mechanics and *ab initio* energies, we used a constant dielectric electrostatic treatment with $\epsilon = 1.0$. All simulations were performed using the recently described Monte Carlo/stochastic dynamics (MC/SD) hybrid simulation algorithm.[7] A timestep of 1.5 fs and a frictional coefficient (γ) of 2.5 ps^{-1} were used for the stochastic dynamics (SD) part of the algorithm. The MC part of the algorithm used random torsional rotations between ±60° and ±180° that were applied to all rotatable bonds except the proline amide C−N bond where the random rotations were between ±120° and ±180°. No torsion rotations were applied to bonds in the pyrrolidine ring of proline as the barriers are low enough to permit adequate sampling from the SD part of the simulation. In our free energy simulations, the total simulation time was 5 ns, and the first 50 ps of each simulation was taken as an equilibration period. The ratio of SD to MC steps was 1:1, and the acceptance rate for the MC part of the algorithm was found to be 0.5−5%. Nonbonded interactions were not truncated. Hydrogen bonding populations were estimated by monitoring geometries around acceptor (O) and donor (H) atoms during each simulation: a hydrogen bond was counted as present if the N−H···O(=C) distance was <2.5 Å, the N−H···O(=C) angle was >120° and the N−H···O=C angle was >90°.[8] Conformational searching was performed using the MCMM method.[9] The conformational searches were carried out in blocks of 5000 MC steps and repeated until subsequent blocks failed to locate new low energy minima (<3 kcal/mol of global minimum).

Ab initio molecular orbital calculations for **1** were performed with the Gaussian 92 package.[10] Initial geometry optimization was performed at the HF/3-21G level of theory and those geometries were reoptimized at the HF/6-31+G* level of theory. Single point energy calculations were performed at the MP2/6-31+G* level of theory on the HF/6-31+G* geometries.[11]

Parameterization was limited to adjusting the torsion and improper torsion parameters. The original AMBER* partial charges were used in order to retain compatibility with the parameters for all other amino acids. Where a torsion around a single bond had multiple substituents we chose to use only one or two of the component torsional arrays to adjust the overall energy profile; the other components were arbitrarily set to zero.

Results

New AMBER* Parameters for Proline. The first step in our parameterization was examination of the geometry of the proline amide bond of **1**. Optimized geometries were obtained for the minimum energy conformations of *cis*- and *trans*-**1** at the HF/6-31+G* level of theory. For the *cis* form, the minimum energy conformation has a Ψ angle of −9.1° while the *trans* minimum has $\Psi = 74.3°$. The geometry around the amide bonds for each minimum and the *cis*−*trans* energy difference are summarized in Table 1.

The *ab initio* calculations show that the ring nitrogen of the proline acetamide is significantly pyramidalized (nearly 20° out of plane in the *cis* form). The pyramidalization of amide nitrogens has been identified previously in *ab initio* calculations of acetamide[12] were it was attributed to a favorable interaction of the nitrogen lone pair and a hydrogen on an adjacent methyl group. In a recent *ab initio* study of **1**, Karplus et al.[13] performed geometry optimizations at the HF/3-21G level of theory and reported the ring nitrogen to be distorted out of the best plane by ∼14−16° for the *cis* and *trans* forms, and considerably more in the transition states for amide rotation. The proline ω-angle is calculated to be nearly 10° from planarity for the *cis* form and 7° for the *trans* form of **1**. For both the nitrogen out-of-plane angle and the amide ω-angle, the original AMBER proline parameters[5] underestimated the deviations from planarity predicted by the *ab initio* calculations. We therefore reoptimized the AMBER* parameters for the proline amide bond and the ring nitrogen improper torsion to give a closer agreement with the *ab initio* calculations. The results are given in Table 1.

The energy difference between the *cis* and *trans* forms of **1** was calculated by quantum mechanics to be 2.38 kcal/mol favoring the *trans* form. Introduction of electron correlation (MP2) lowers this value slightly to 2.28 kcal/mol. The original AMBER* parameters greatly overestimated the *cis*/*trans* energy difference, making it 5.03 kcal/mol *in vacuo*. The torsional parameters for the proline amide bond were therefore reoptimized to reproduce the *ab initio* energy difference and give an energy difference of 2.26 kcal/mol. No attempt was made to recalculate the height of the barrier for *cis*/*trans* isomerism. Karplus *et al.*[13] studied this interconversion in detail and report values of 17.91 and 20.65 kcal/mol for the *syn* and *anti* transition states at the HF/6-31G*//HF/

(6) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127.
(7) Guarnieri, F.; Still, W. C. *J. Comput. Chem.* **1994**, *15*, 1302.
(8) Kurz, M.; Mierke, D. F.; Kessler, H., *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 210.
(9) Chang, G.; Guida, W. C.; Still, W. C. *J. Am. Chem. Soc.* **1989**, *111*, 4379.
(10) M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. W. Wong, J. B. Foresman, M. A. Robb, M. Head-Gordon, E. S. Replogle, R. Gomperts, J. L. Andres, K. Raghavachari, J. S. Binkley, C. Gonzalez, R. L. Martin, D. J. Fox, D. J. Defrees, J. Baker, J. J. P. Stewart, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1993.
(11) A recent study has concluded that calculations at the MP2/6-31G*//HF/6-31G* level give relative energies for conformations of amino acids which are in good accord with those obtained at the MP2/6-31+G*//MP2/6-31+G* level: Gronert, S; O'Hair, R. A. J. *J. Am. Chem. Soc.* **1995**, *117*, 2071.
(12) Wong, M. W.; Wiberg, K. B. *J. Phys. Chem.* **1992**, *96*, 668.
(13) Fischer, S.; Dunbrack, R. L.; Karplus, M. *J. Am. Chem. Soc.* **1994**, *116*, 11931.

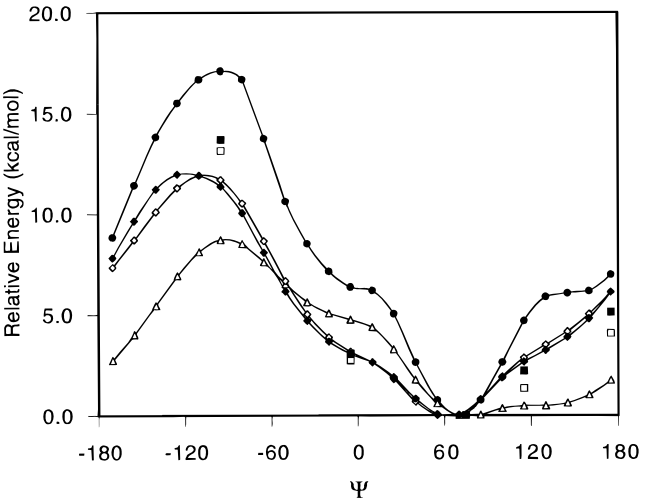


Figure 1. Relative energy as a function of Ψ-angle for **1** calculated at a number of levels of *ab initio* theory and with AMBER* parameters: (filled circles) HF/3-21G//HF/3-21G; (open squares) HF/6-31+G*//HF/6-31+G*'; (filled squares) MP2/6-31+G*//HF/6-31+G*; (open triangles) original united atom AMBER*; (filled diamonds) modified united atom AMBER*; (open diamonds) modified all atom AMBER*.

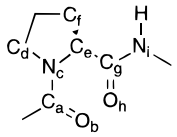
6-31G* level of theory. We have calculated Ψ/ξ[14] plots for **1** with our new AMBER* parameter set (data not shown) and find that they are in good agreement with data by Karplus *et al.* From these plots, the AMBER* barrier for *cis/trans* isomerization is estimated to be 15−16 kcal/mol. These values are in good agreement with the barriers for *N,N*-dimethylacetamide of 14.1 and 17.0 kcal/mol ($\Delta H^{\ddagger}_{0\,K}$) for *syn* and *anti* transition states respectively reported at the G2(MP2) level of theory by Wiberg *et al.*[15] and with the experimental[16] gas phase value of 15.8 ± 1.1 kcal/mol ($\Delta H^{\ddagger}_{298\,K}$).

With a set of parameters which accurately reproduces the amide geometry and the *cis/trans* energy difference, we then turned to the energy profile of the proline Ψ-angle. A complete rotational energy profile of the Ψ-angle for the *trans* form of **1** was calculated at the HF/3-21G level of theory to define the major features of the potential energy surface (Figure 1). This profile shows a minimum at *ca.* 70° which corresponds to an internally hydrogen bonded conformation and a maximum near −100° where the amide (*N*-methyl) nitrogen is close to pyrrolidine ring hydrogens and near-eclipsing of the α-hydrogen with the proline carbonyl oxygen and of the exocyclic nitrogen with the proline *β*-carbon are present. The minimum and maximum points on the HF/3-21G profile were then used as starting geometries for higher level optimizations at HF/6-31+G*. The maximum optimized to a Ψ angle of −94.8° while the minimum was located at a Ψ = 74.3°. In addition to these points, structures with the Ψ-angle constrained to −5°, 115°, and 175° were also optimized at the HF/6-31+G* level of theory. For all the HF/6-31+G* optimized geometries, a single point energy calculation with electron correlation (MP2/6-31+G*) was performed. Total energies and Cartesian coordinates for all optimized geometries are given as supporting information. The relative energies obtained from these calculations are shown graphically in Figure 1. The HF/3-21G profile shown in Figure 1 indicates that there may be additional minima at Ψ-angles near 5° and 145°. Reoptimizations of all degrees of freedom starting from these values of Ψ at the HF/3-21G level of theory resulted in new minima with Ψ = 7.8° and 147.0°. However, further optimizations of these at the HF/6-31+G* level of theory consistently collapsed to the *γ*-turnlike, internally hydrogen bound structure (Ψ = 74.3°). Hence we conclude that these additional minima are either not present or insignificant on the gas phase potential energy surface of **1**.

A complete rotational profile for the Ψ-angle of **1** calculated with the original AMBER* parameters is also shown as open triangles in Figure 1. While these parameters correctly predict a single minimum corresponding to an internally hydrogen bound form, the minimum they create is much shallower than that found quantum mechanically, especially in the region of Ψ = 80−180°. Thus, the original AMBER* parameters underestimate the stability of hydrogen bound conformations relative to non-hydrogen bound conformations. We therefore reoptimized the appropriate AMBER* torsional parameters for proline to better reproduce the Ψ-angle profile calculated by *ab initio* methods. The Ψ-angle profile calculated with these new AMBER* parameters is also shown in Figure 1 as diamonds. These profiles more closely reproduces the relative conformational energies calculated at the MP2/6-31+G*//HF/6-31+G* level of theory. The largest difference between the new AMBER* and the *ab initio* energies is at the energy maximum (Ψ = −95°) where the difference between the new AMBER* and MP2/6-31+G* is ∼2 kcal/mol. This difference is, however, acceptable for most purposes because it is found at a relatively high energy (and therefore sparsely populated) region of the potential energy surface.

The new proper and improper torsional parameters developed in this study are listed in Table 2.

Table 2. New AMBER* Torsional Parameters for Proline



	united atom			all atom		
	$V_1/2$	$V_2/2$	$V_3/2$	$V_1/2$	$V_2/2$	$V_3/2$
Proper Torsion						
$C_a-N_c-C_g-O_h$[a]	0.00	0.00	0.00	0.00	0.00	0.00
$N_c-C_e-C_g-O_h$	0.00	0.00	0.00	0.00	0.00	0.00
$C_f-C_e-C_g-N_i$	1.19	-0.16	0.00	1.21	-0.08	0.00
$C_f-C_e-C_g-O_h$	0.00	0.00	0.00	0.00	0.00	0.00
$N_c-C_e-C_g-N_i$	-0.74	0.52	0.02	0.40	0.42	0.00
$O_b-C_a-N_c-C_e$	0.50	2.50	0.00	0.60	2.50	0.00
Improper Torsion						
$C_a-N_c-C_e-C_d$	0.00	0.00	1.30	0.00	0.00	1.50

[a] "Remote torsion" as defined in reference 4.

The new proper and improper torsional parameters developed in this study are listed in Table 2.

Free Energy Simulations for Proline-Containing Peptides. While a good agreement with high level *ab initio* calculations is necessary for a new set of molecular mechanics parameters, the ultimate validation of parameters rests on their ability to reproduce experimental observations. For this reason, we carried out a series of free energy simulations of proline-containing peptides where solution phase experimental data is available.

The first system we examined was **1**. The conformational preferences of this compound have been examined

(14) For the definition of the virtual torsion ζ see reference 13.
(15) Wiberg, K. B.; Rablen, P. R.; Rush, D. J.; Keith, T. A. *J. Am. Chem. Soc.* **1995**, *117*, 4261.
(16) Ross, B. D.; True, N. S.; Matson, G. B. *J. Phys. Chem.* **1984**, *88*, 2675.

Table 3. Calculated and Experimental % Hydrogen Bonding and % *Cis* Form for 1 from 5 ns MC/SD Simulations Using New All-Atom AMBER* Parameters

	% hydrogen bonding	% *cis* form
in vacuo		
new AMBER*	78	5
$CHCl_3$		
experiment[a]	68	14
new AMBER*	72	10
original AMBER*	37	<1
H_2O		
experiment[b]	not detected	30
new AMBER*	6	40

[a] Experimental value measured in CH_2Cl_2 from ref 18. [b] From reference 17.

in both water[17] and organic solvents.[18] Two quantities which can be measured experimentally are the *cis/trans* ratio and the fraction of the population having an internal hydrogen bond. In an attempt to reproduce this data, we used the Monte Carlo/stochastic dynamics (MC/SD) simulation method to generate converged ensembles of Boltzmann-weighted, three-dimensional states near room temperature (300 K). This method has recently been demonstrated to generate converged ensembles that are good agreement with the experimentally observed conformational free energies a series of diamides in organic solution.[19] Convergence of our simulations was ensured by showing that final populations did not vary significantly with starting geometry or simulation time. The results of our 5 ns MC/SD simulations for **1** *in vacuo* and in GB/SA $CHCl_3$ and H_2O are summarized in Table 3.

The results in Table 3 show that the new AMBER* parameters reproduce the experimentally observed trends for **1** both in the population of the *cis* form and in the extent of hydrogen bonding as a function of solvent. In the organic solvent, the new parameters give a much better agreement with the experimentally observed percentage of hydrogen bonding than do the original parameters. This observation is consistent with fact that the new parameters define a narrower minimum for the hydrogen bound conformation *in vacuo* (see Figure 1). The *cis* population is also better reproduced by simulations with the new AMBER* parameters. Again this finding is consistent with the fact that the original parameters significantly overestimated *cis/trans* energy differences. $CHCl_3$ solvation slightly reduces the amount of internal hydrogen binding relative to the gas phase value. The *cis*−*trans* free energy difference for **1** is also slightly smaller in the organic solvent than in vacuum. In water, however, the internal hydrogen bond could not be detected experimentally and the simulations in GB/SA water reflect this with only a small (6%) population for hydrogen bound forms. Water is known to substantially reduce the *cis*−*trans* energy difference in **1**, and our simulations in GB/SA water reflect this.

Further insight into the conformational preferences of **1** in various solvents can be obtained by examining the distribution of Ψ-angles populated during the simulation (Figure 2). In vacuum and $CHCl_3$, the Ψ-angle most highly populated corresponds to the minima around 75°

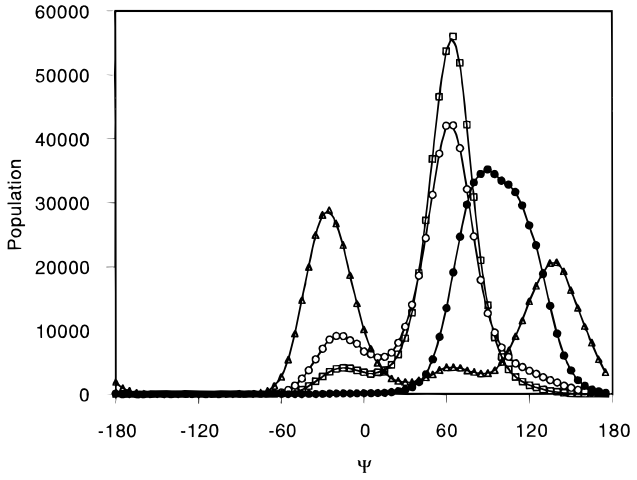


Figure 2. Distribution of Ψ-angles from a 5 ns MCSD simulation of **1** using all atom AMBER* parameters, *in vacuo*, with GB/SA $CHCl_3$ and GB/SA H_2O: (open squares) modified AMBER* *in vacuo*; (open circles) modified AMBER* with GB/SA $CHCl_3$; (filled circles) original AMBER* with GB/SA $CHCl_3$; (open triangles) modified AMBER* with GB/SA H_2O.

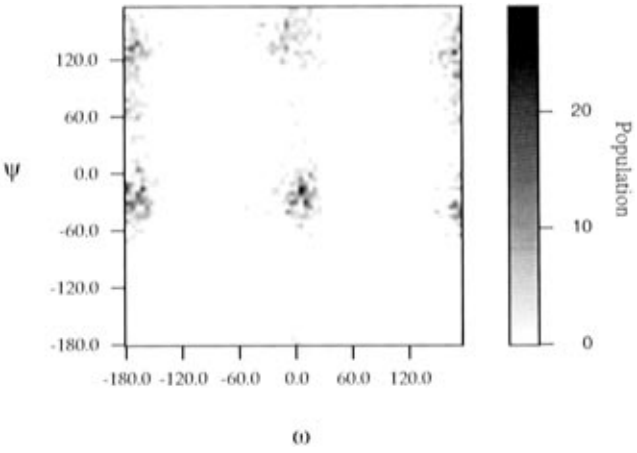


Figure 3. Distribution of Ψ- and ω-angles from a 5 ns MCSD simulation of **1** using the modified all atom AMBER* parameters and GB/SA H_2O.

with a smaller distribution centered around Ψ = −20° which arises from the *cis* isomer that is sampled occasionally during the simulation. The original AMBER* parameters in $CHCl_3$ result in a broader distribution of angles centered around Ψ = 90°, and this is the reason for the lower populations of hydrogen bonded conformations predicted with the old parameter set. In GB/SA H_2O, Ψ indicates a small population corresponding to the hydrogen bonded minimum around Ψ = 60°, and two larger and approximately equal distributions at Ψ = −30° (corresponding to a right-handed α helix (α_R) conformation) and Ψ = 140° (corresponding to a polyproline II (P_{II}) conformation). Experimental observations[17] suggest that the γ-turn, internally hydrogen bonded conformation was largely absent, but the populations of other conformations were not established. Our Figure 2 data represents, however, the Ψ-angle distribution for both *trans* and *cis* forms of **1**, with the latter having a significant (40%) population in GB/SA H_2O. To pursue this issue, we plotted the distribution of both the Ψ- and ω-angles from the aqueous MC/SD simulation. Figure 3 shows the result: the α_R and P_{II} conformations are approximately equally populated in both the *cis* and *trans* forms of **1**.

(17) Madison, V.; Kopple, K. D. *J. Am. Chem. Soc.* **1980**, *102*, 4855.
(18) Liang, G.-B.; Rito, C. J.; Gellman, S. H. *Biopolymers* **1992**, *32*, 293.
(19) McDonald, D. Q.; Still, W. C. *J. Am. Chem. Soc.* **1994**, *116*, 11550.

Table 4. Conformational Populations for the Ac-L-Pro-X-NHMe Dipeptides from 5 ns, 300 K MC/SD Simulations Using United Atom AMBER* and GB/SA $CHCl_3$

X	% β-turn (**2a**)	% Pro γ-turn (**2b**)	% X γ-turn (**2c**)
Gly	12	28	5
(L)Ala	5	35	8
(D)Ala	11	43	7
(L)Val	4	38	12
(D)Val	6	46	10
(L)Ser	6	45	20
(D)Ser	14	45	14
(L)Asn(*N*-Me)	<1	33	5
(D)Asn(*N*-Me	<1	39	5

We next studied dipeptides **2** having the sequence Ac-Pro-X-NHMe where X = Gly, Ala, Val, Ser, and Asn(*N*-Me) to examine their ability to form a β-turn. On the basis of our results with the proline peptide **1**, we would not expect much internal hydrogen bonding in water; however, significant intramolecular hydrogen bonding might be expected in organic solvents, and some experimental data is available in this regard. For dipeptides **2**, there are three hydrogen bonding possibilities: a β-turn conformation involving a ten-membered hydrogen bonded ring (**2a**), a γ-turn across the proline residue (**2b**), or a γ-turn across the "X" residue (**2c**). To investigate the hydrogen bonding of **2**, we performed 5 ns MC/SD simulations at 300 K and monitored the populations of each of the three possible hydrogen bonds. The results of these simulations are summarized in Table 4, and the populations shown there are converged to within 1%.

The data in Table 4 shows a number of interesting trends. First, the calculations predict that the β-turn hydrogen bonding conformation **2a** never has a population above 14% for any of the sequences studied. The β-turn appears to be less favored as the X-residue becomes more bulky, so that the β-turn population diminishes as X goes from Gly → Ala → Val. Heterochiral sequences (*e.g.* (L)Pro(D)Ala) always increase the amount of β-turn present relative to the corresponding homochiral sequences. Alternating chirality is a common feature of linear peptides that form β-turns in solution or in the solid state[1] and has recently been reported to promote reverse turn formation in a computational study of reverse turn mimetics.[20] While the β-turn hydrogen bonding pattern is present in small amounts in most of the dipeptides **2**, the calculations predict the most populated hydrogen bonding pattern in $CHCl_3$ to be a γ-turn across the proline (**2b**). γ-Turns across the "X" residue (**2c**) in the Ac-Pro-X-NHMe sequence become increasingly favored as X becomes more bulky. A notable exception to this trend is **2** (X = *N*-Me Asn) where a suitably placed hydrogen bond acceptor on the Asn side chain appears to disrupt γ-turn formation across the main chain. In most of the sequences examined, however, the population of the γ-turn across the "X" residue is comparable to or larger than that of the β-turn.

While direct a comparison of the populations shown in Table 4 with experimental observations is not possible, Gellman *et al.*[21] did report a thermodynamic analysis of **2** that determines the enthalpic and entropic factors associated with β-turn formation. This was done by

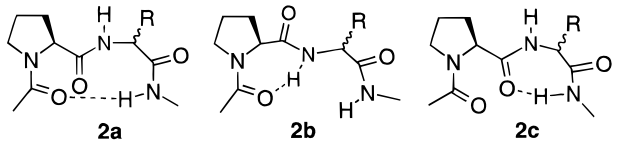


Figure 4. Hydrogen bonding patterns in the NAc-Pro-X-NMe dipeptide **2**.

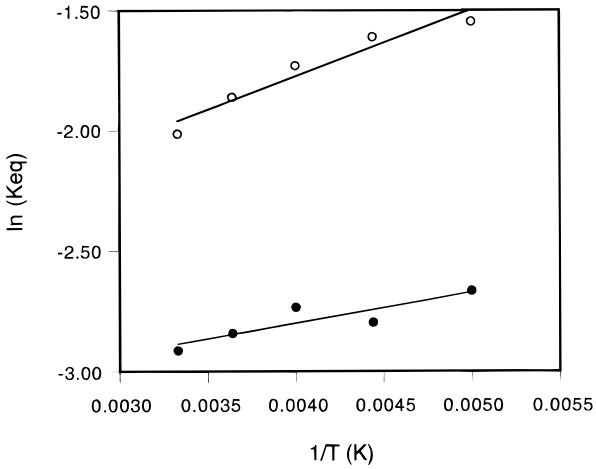


Figure 5. Van't Hoff plot constructed from β-turn hydrogen bond populations of the Ac-Pro-X-NHMe sequence from 5 ns MCSD simulations using the modified AMBER* parameters and GB/SA $CHCl_3$: (open circles) X = Gly; (filled circles) X = Ala.

Table 5. Calculated and Experimental Enthalpic Preference for β-Turn Formation with Ac-(L)Pro-X-NHMe (2) in Chlorohydrocarbon Solution

X	ΔH_{calc} (kcal/mol)[a]	ΔH_{exp} (kcal/mol)[b]
Gly	−0.6 (±0.3)	−1.7 (±0.5)
(L)Ala	−0.3 (±0.3)	−1.4 (±0.4)
(L)Val	0.0 (±0.3)	−0.6 (±0.2)

[a] From the data in Figure 5, solvent $CHCl_3$. [b] From reference 21, solvent CH_2Cl_2. Error bars are shown based on the error estimate of 30% (see reference 21).

treating the peptide as a two-state system equilibrating between a β-turn state and a "non-β-turn state" that includes both the alternative hydrogen bonding patterns (**2b** and **2c**) plus any non-hydrogen bonded conformations. By measuring the population of the β-turn state as a function of temperature, a van't Hoff plot can be constructed and the thermodynamic parameters associated with β-turn formation extracted. We can perform a similar analysis computationally. Thus we carry out MC/SD simulations over the temperature range 200−300 K, determine the equilibrium population of the β-turn population at each temperature (T), convert that to a equilibrium constant (K_{eq}) for the equilibrium involving the two states (**2a** and everything else), and plot $\ln(K_{eq})$ against $1/T$. The results for **2** (X = (L)Ala) and **2** (X = Gly) are shown in Figure 5. For **2** (X = (L)Val), the population of the β-turn was essentially invariant with temperature. From the data shown, the enthalpic preference (ΔH) for formation of the β-turn in the Ac-Pro-X-NHMe sequence can be calculated and this data together with the experimental values are shown in Table 5.

The most obvious feature of the data in Table 5 is that absolute magnitudes of the calculated enthalpic preference for β-turn formation in **2** are consistently lower than those measured experimentally. However, the experimental observation that β-turn formation is more en-

(20) Chalmers, D. K.; Marshall, G. R. *J. Am. Chem. Soc.* **1995**, *117*, 5927.

(21) Liang, G.-B.; Rito, C. J.; Gellman, S. H. *J. Am. Chem. Soc.* **1992**, *114*, 4440.

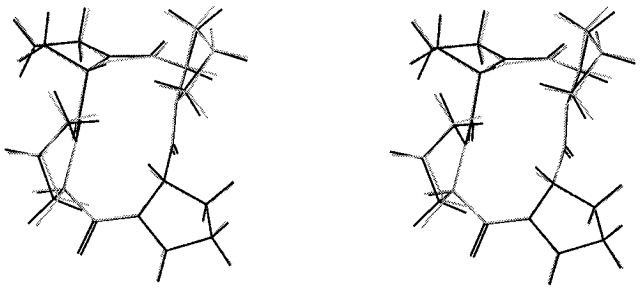
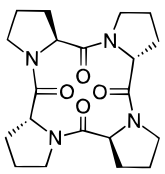


Figure 6. Superimposition of calculated (all atom AMBER*) global minimum and X-ray structures of cyclo((D)Pro-(L)Pro-(D)Pro-(L)Pro) (**3**).

thalpically favored by Gly > Ala > Val is reproduced by the simulations. There are a number of possible explanations for the differences between these calculated and experimental values of ΔH. Firstly the experimental values were measured in CH_2Cl_2, whereas our simulations were performed using a $CHCl_3$ solvent model. Previous studies of hydrogen bonding with similar compounds in CH_2Cl_2 and $CHCl_3$ have concluded, however, that solvent effects with these two solvents are unlikely to be large. The experimental values also depend on a limiting value for $\Delta\delta$ NH/ΔT, and this varies somewhat with the choice of a suitable model compound. It was estimated that this uncertainty could result in errors in ΔH_{exp} as large as 30%.[21] Even taking such factors into account, it still appears that the new AMBER* parameters, while able to reproduce trends in β-turn formation for simple peptides, are not able to accurately reproduce the magnitude of the experimental enthalpic preference for β-turn formation. Thus it is likely that the population of β-turns relative to other hydrogen bonding patterns will be somewhat underestimated in simulations with our new parameters. This may be a result of the absence of a directional hydrogen bonding term in the force field which could overestimate the strength of nonplanar hydrogen bonding arrays such as those found in γ-turn conformations (*e.g.* **2b**, **2c**).

As a final test of the new parameter set, we studied a proline-containing peptide where a recent study with our original AMBER parameters found a poor correlation with experiment. In their recent study of reverse turn mimetics, Marshall and Chalmers[20] compared the relative energies, as calculated with a number of force fields and the GB/SA water, of low energy conformations of cyclo((D)Pro-(L)Pro-(D)Pro-(L)Pro) (**3**).



3

This cyclic tetrapeptide is well characterized in water solution and in the solid state[22] and was found to exist exclusively in a conformation having alternating *cis* and *trans* amide bonds: the *ctct* conformation. With our original parameters, however, the authors found that conformations with a *cis*−*cis*−*trans*−*trans* (*cctt*) conformation were at least 4.9 kcal/mol lower in energy than the experimentally observed *ctct* conformations. We therefore reevaluated the conformational preferences of this compound using our new parameters. To this end we performed an extensive conformational search of **3** with our new AMBER* all-atom parameters in GB/SA water. From this search we found that all conformations within 3 kcal/mol of the global minimum have the experimentally observed *ctct* structure. The global minimum is 1.2 kcal/mol lower in energy than the next lowest energy conformation and hence is predicted to be highly populated in water. This same structure is very close to that of the X-ray structure of **3**. They are shown superimposed in Figure 6 and have an rms superimposition deviation of 0.11 Å.

Conclusion

As noted above, calculations on the simple proline peptide **1** with our original AMBER* parameters show significant differences from the results of *ab initio* molecular orbital calculations. In particular, the original parameters result in a proline amide bond that is too planar, overestimate the *cis*/*trans* energy difference for the tertiary amide bond, and overestimate the flexibility of internally hydrogen bonded conformations with respect to the Ψ-angle profile. These differences justify our original concern that proline, with its constrained Φ-angle, would not be adequately described by a parameter set that was developed for the alanine peptides as models for all α-substituted amino acids. Thus we developed a new set of AMBER* parameters specifically for proline residues that gives considerably better agreement with *ab initio* results in the areas of amide geometry, *cis*/*trans* energy difference, and Ψ-angle energy profile.

While the ability to reproduce high level molecular orbital calculations on suitable model compounds is an important starting point in the validation of a new parameter set, it is also important to demonstrate that the parameters can be used to model known conformational preferences of real molecules under typical laboratory conditions, *i.e.* in solution at room temperature. This requires that we add terms to the gas phase potential energy surface that is defined by the molecular mechanics parameters to allow for the effects of solvent. In this work, we used the GB/SA continuum model for solvent because it has been demonstrated to reproduce both absolute and relative solvation free energies for small molecules.[6,19,23] Hence we expect that, once the gas-phase potential energy surface is adequately defined using quantum mechanics, the GB/SA model will correctly add the effects of solvation. The use of a continuum solvation model is also important here because it allows us to use the highly convergent MC/SD simulation procedure[7] to sample conformational space. Because an experimental observation at non-zero temperatures represents an average over the entire ensemble of populated conformations and vibrational states, a calculation needs to include representative samples of all these states if it is to reproduce the observation. Thus once the problems of solvation and gas phase parameters are adequately addressed, some type of free energy simulation methodology is required to generate the Boltzmann-weighted ensemble of molecular geometries from which the ensemble-averaged properties may be calculated. Confor-

(22) Mästle, W.; Weber, T.; Thewalt, U.; Rother, M. *Biopolymers* **1989**, *28*, 161.

(23) Varnek, A. A.; Wipff, G.; Glebov, A. S.; Feil, D. *J. Comput. Chem.* **1995**, *16*, 1.

mational searching and energy minimization has often been used as a simple approach to ensemble generation. However, it can introduce significant errors (especially with highly flexible molecules) because accurate conformational entropy differences are difficult to include.[7] Furthermore, properties such as the degree of internal hydrogen bonding cannot be meaningfully analyzed in terms of minimum energy conformers alone, because all the structures associated with one energy well may not have the same property. For example, while a given minimum energy conformer may have a strong hydrogen bond, other populated geometries within the same energy well may not.[24] In some systems it is possible to use a conventional molecular simulation methods (MD or MC) to generate the required ensemble averages.[24] In general, however, convergence will be difficult or impossible to obtain when energy barriers between conformations are significant. Even barriers as small as 3 kcal/mol can make it difficult to converge a dynamics simulation to high precision.[7] With *cis/trans* interconversion barriers of ~16 kcal/mol, proline derivatives obviously require some special approach for ensemble generation. For problems such as this, we find the MC/SD free energy simulation methodology[7] to be quite effective. Its use of large Monte Carlo moves in torsion angle space allows it to jump over energy barriers easily and hence to generate converged ensembles of significantly populated states efficiently with multiconformational molecules like **1** and **2**. Using these approaches to solvation and ensemble generation, we can apply our new parameter set to experimental observations.

When MC/SD simulations were performed with our quantum mechanically-derived parameter set in the appropriate solvent for proline derivatives **1**-**3**, calculated observables such as the population of given conformations and proline *cis/trans* ratios are found to be generally in good agreement with available experimental data measured in solution.

The results described here clearly demonstrate the importance of high quality molecular mechanics parameters and the correct choice of representative model compounds for parameter development. While we believe that our original AMBER* parameters give acceptable performance for most α-amino acids, proline is sufficiently different to warrant a specific set of molecular mechanics parameters. Considering that there are still many functional groups (and combinations of structurally proximate functional groups) that are inadequately parameterized in all existing force fields, this study may be useful as a guide to future parameterizations. Force field parameterization is a necessary prelude to successful molecular modeling, and *ab initio* calculations provide an exceptionally powerful tool for testing existing parameters and generating new ones as we have done here. This study shows that once accurate parameterization of the gas phase energy surface is achieved, then use of a reliable model for solvation and an efficient free energy simulation procedure for generating of Boltzmann-weighted ensembles can give results that are in good agreement with experiment.

Acknowledgment. This work was support by NSF CHE92 08254.

Supporting Information Available: HF/3-21G and HF/6-31+G* optimized geometries and total energies for **1** (7 pages). This material is contained in libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.

JO951788K

(24) McDonald, D. Q.; Still, W. C. *Tetrahedron Lett.* **1992**, *33*, 7747.