Molecular Dynamics and Monte Carlo Simulations Favor the α-Helical Form for Alanine-Based Peptides in Water

Julian Tirado-Rives,* David S. Maxwell, and William L. Jorgensen*

Contribution from the Department of Chemistry, Yale University, New Haven, Connecticut 06511

*Received June 3, 1993**

Abstract: Molecular dynamics and Monte Carlo simulations of a capped undecaalanine peptide were conducted to determine the preferred helical conformation in water. The results clearly favor the α-helical form over the 3_{10} alternative, in contrast to recent ESR findings. A molecular dynamics simulation started at the 3_{10} conformation converges rapidly to an α-helix, and the free energy profile calculated *via* Monte Carlo simulations shows the α-helix to be more stable by *ca.* 1.0 kcal·mol^{-1} per residue.

Introduction

Though small peptides typically do not exhibit secondary structure in aqueous solution, several helix-forming exceptions have recently been reported. These structured peptides include natural sequences like the S-peptide from ribonuclease A[1] and Pα5 from BTPI[2] and several designed polyalanine-based peptides with regular Lys and Glu insertions.[3,4] By analogy to known protein structures,[5] it has been assumed that the isolated helices exist predominantly in the α rather that the 3_{10} form.[6] However, the spectroscopic techniques most commonly used to study conformation in aqueous solution, circular dichroism and nuclear magnetic resonance, have not yet been able to discriminate between these alternatives. Remarkably, a recent electron spin resonance study with doubly spin-labeled analogues of the trilysine-substituted polyalanine peptide 3K[4] has been interpreted as evidence that these 16-residue peptides are mostly in the 3_{10} conformation.[7] Molecular dynamics (MD) and Monte Carlo (MC) simulations can also probe the structures and energetics for peptides in solution and have been used successfully to study unfolding transitions of an S-peptide analogue.[8] Here we report the results of similar calculations on undecaalanine in water which clearly demonstrate predominance of the α-helical form.

Force Field and Computational Methods

The peptide chosen for this study, undecaalanine with neutral acetyl and methylamide groups at the N- and C-termini, was represented by the OPLS/AMBER force field.[9,10] As in any simulation, the reliability of the results depends critically on the force field. The OPLS nonbonded parameters have been developed to accurately reporduce experimentally measured properties of liquids.[9] The parameters for alanine are based on simple amides such as acetamide, whose free energy of hydration was calculated *via* Monte Carlo simulations[11] to be −9.5 ± 0.4 kcal·mol^{-1}, in excellent agreement with the experimental value of −9.68 kcal·mol^{-1}.[12]

The conformational energetics of alanine are also well described by the force field. Though the α and 3_{10} forms are not minima in the gas-phase conformational surfaces of the *N*-formylalaninamide dipeptide with either *ab initio* HF/3-21G[13] or OPLS/AMBER calculations, the energy difference between equivalent points near these conformations is approximately the same, *ca.* 1.0 kcal·mol^{-1} in both cases. Higher level *ab initio* calculations including correlation effects (MP2/6-31 + G*//HF/6-31G*) give an energy difference of 1.1 kcal·mol^{-1} between the two most stable conformations, C_5 and C_7^{eq}, of *N*-formylalaninamide, the latter being the most stable.[13] More recent calculations at the MP2/6-311G** level yield a 1.7 kcal·mol^{-1} difference for the same molecule[14] and 1.5 kcal·mol^{-1} at the MP2/TZVP//HF/6-31G** level for *N*-acetyl-*N'*-methylalaninamide.[15] Energy minimizations with the OPLS/AMBER force field give a similar energy difference (1.5 kcal·mol^{-1}) between these conformations for both dipeptides.

The molecular dynamics simulations were conducted using the AMBER 3.0a software[16] and the protocol developed in our previous study of an S-peptide analogue,[8] including the use of SHAKE[17] to fix the bond lengths. The initial coordinates for two different MD runs were obtained from structures generated at the ideal values for both the α- ($\phi = -60°, \psi = -50°$) and the 3_{10}-helices ($\phi = -50°, \psi = -28°$), which were then minimized *in vacuo*. The resulting structures were centered in rectangular cells of TIP3P water[18] to produce periodic boxes of dimensions 34.8 × 30.5 × 30.5 Å^3, containing 1039 water molecules for the α-helix, and 38.8 × 28.9 × 28.9 Å^3, containing 1031 molecules, for the 3_{10}-helix. The solvent in these initial boxes was relaxed *via* 100 steps of steepest descent minimization followed by 3 ps of NVT (constant volume and temperature)-MD, during which the temperature was raised from 100 to 278 K, and by 4 ps of NVT-MD at 278 K with the solute fixed. This was followed by an additional 2 ps of NVT-MD, during which the solute was allowed to move freely and the temperature was again brought to 278 K. The systems thus obtained were used as the starting points for the NPT (constant pressure and temperature) simulations of 300 (α) and 100 ps (3_{10}) at 278 K and 1 atm with a 2 fs time step. Residue-based cutoffs were applied at 9 Å, i.e., if two residues or a residue and a water molecule have any atoms within 9 Å, the interaction between the entire pair is evaluated.

The Monte Carlo calculations were run with the BOSS 3.4 program[19] to obtain a free energy profile for interconversion of the α- and 3_{10}-helices. The nonbonded interactions were truncated at 9 Å with quadratic smoothing to zero over the last 0.5 Å. The initial system was created by centering the peptide in a 3_{10}-helical conformation ($\phi = -60°, \psi = -30°$)

* Abstract published in *Advance ACS Abstracts,* November 15, 1993.

(1) Klee, W. A. *Biochemistry* **1968**, *7*, 2731–2736.
(2) Goodman, E. M.; Kim, P. S. *Biochemistry* **1989**, *28*, 4343–4347.
(3) Marqusee, S.; Baldwin, R. L. *Proc. Nat. Acad. Sci. U.S.A.* **1987**, *84*, 8898–8902.
(4) Marqusee, S.; Robbins, V. H.; Baldwin, R. L. *Proc. Nat. Acad. Sci. U.S.A.* **1989**, *86*, 5286–5290.
(5) Barlow, D. J.; Thornton, J. M. *J. Mol. Biol.* **1988**, *201*, 601–619.
(6) Scholtz, J. M.; Baldwin, R. L. *Annu. Rev. Biomol. Struct.* **1992**, *21*, 95–118.
(7) Miick, S. M.; Martinez, G. V.; Fiori, W. R.; Todd, A. P.; Millhauser, G. L. *Nature* **1992**, *359*, 653–655.
(8) Tirado-Rives, J.; Jorgensen, W. L. *Biochemistry* **1991**, *30*, 3864–3871.
(9) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
(10) Weimer, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. J. *J. Am. Chem. Soc.* **1984**, *106*, 765–784.
(11) Duffy, E. M.; Severance, D. L.; Jorgensen, W. L. *Isr. J. Chem.*, in press.
(12) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. *Biochemistry* **1981**, *20*, 849–855.
(13) Head-Gordon, T.; Head-Gordon, M.; Frisch, M. J.; Brooks, C. L., III; Pople, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5989–5997.
(14) Frey, R. F.; Coffin, J.; Newton, S. Q.; Ramek, M.; Cheng, V. K. W.; Momany, F. A.; Schäfer, L. *J. Am. Chem. Soc.* **1992**, *114*, 5369–5377.
(15) Gould, I. R.; Kollman, P. A. *J. Phys. Chem.* **1992**, *96*, 9255–9258.
(16) Singh, U. C.; Weiner, P. K.; Caldwell, J.; Kollman, P. A. *AMBER 3.0a*; University of California: San Francisco, CA, 1989.
(17) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
(18) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(19) Jorgensen, W. L. *BOSS 3.4*; Yale University: New Haven, CT, 1993.

0002-7863/93/1515-11590$04.00/0

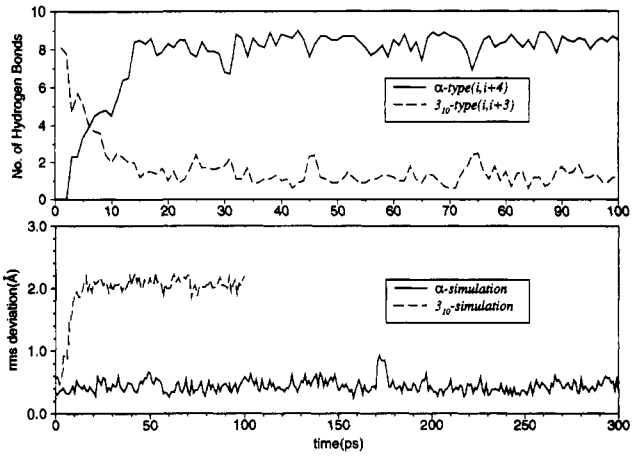

Figure 1. (Top) Backbone–backbone hydrogen bond distribution during the 3_{10} simulation, averaged over 1-ps intervals. A hydrogen bond is defined by a H–O distance less than 2.5 Å and an N–H–O angle between 120 and 180°. (Bottom) Plots of the rms deviations of the main-chain heavy atoms during the MD simulations from the initial 3_{10}- and α-helical structures.

in a large box of TIP4P water.[18] It is our practice to use the somewhat preferable TIP4P model rather than TIP3P water in Monte Carlo simulations since the placement of the negative charge at a massless site in the TIP4P model is computationally awkward in MD simulations.[18] All water molecules farther than 12.0 Å in any one Cartesian direction or closer than 2.5 Å to the peptide were then removed to produce a rectangular box of dimensions 30.9 × 37.8 × 51.7 Å^3, containing the peptide plus 1719 water molecules. Free energy changes were obtained using statistical perturbation theory[20] with the backbone angles ψ altered in tandem as the reaction coordinate and the angles ϕ fixed at –60°; the computational procedure was similar to that described for computing solvent effects on the barrier to rotation about the central C–N bond in dimethylacetamide.[21] All other bond angles and dihedral angles were sampled during the simulations, with bond lengths again fixed. The present system was then subjected to an equilibration regime comprised of 5 × 10^5 configurations of NVT-MC allowing only solvent to move, 5 × 10^5 configurations of NVT-MC in which both solvent and solute could move, and 4.9 × 10^7 configurations of NPT-MC. The averaging period of 2 × 10^7 configurations of NPT-MC started from this structure. The initial configurations for subsequent runs were generated from the structure at the end of the previous equilibration period by modifying the ψ angles to their new value while maintianing maximum overlap with the original structure. The system was reequilibrated for 4 × 10^6 configurations prior to the new averaging period of at least 2 × 10^7 configurations. In this fashion, 32 distinct values of $\Delta G/\Delta\psi$ were computed from separate NPT (298 K, 1 atm) simulations at 16 values of ψ using double-wide sampling with $\Delta\psi = \pm 0.5°$. These values were then integrated over the full range of ψ to obtain the free energy profile.[22] The Monte Carlo calculations have several advantages over molecular dynamics for these free energy calculations. In particular, the use of internal coordinates in the Monte Carlo simulations facilitates the imposition of constraints on internal coordinates, such as the backbone ϕ angles, and the utilization of internal coordinates as the reaction coordinate, *e.g.*, the ψ angles.

Results and Discussion

The progress of the MD simulations can be followed by the root mean square (rms) deviations of the peptide main-chain atoms (N, C$^\alpha$, C, and O) from their respective starting structures, as plotted in Figure 1. The results reflect the overall motion of the peptide and show deviations of 0.34 and 2.19 Å for the final instantaneous structures from the α and 3_{10} runs, respectively. Notably, the final structure from the 3_{10} run has a rms deviation of only 0.48 Å with the α-helical initial structure, well within the range observed during the α-helix simulation. Though the α-helical structure is stable, the initial 3_{10} structure converts completely to the α-helical conformation within 15 ps. The initial and final structures from this simulation are shown in Figure 2. This transformation is also obvious in the number and nature of the backbone–backbone hydrogen bonds in Figure 1; the eight initial 3_{10} hydrogen bonds convert to eight of α-type and 1 of 3_{10}-type. The hydrogen bonds for residues 2 ← 5 and 9 ← 12 in the initial structure of the 3_{10}-helix were broken during the 2 ps of NVT-MD prior to the NPT simulation.

Table I. Results of the MD Simulations

property	simulation α	simulation 3_{10}
averaging period (ps)	0–300	20–100
density (g·cm^{-3})	1.024 ± 0.008	1.021 ± 0.007
solvent-accessible surface area (Å^2)	860 ± 8	858 ± 8
no. of α-helical hydrogen bonds	8.42 ± 0.49	8.31 ± 0.48
no. of 3_{10}-helical hydrogen bonds	1.09 ± 0.40	1.23 ± 0.45
$C_i^\beta - C_{i+2}^\beta$ distance (Å)	7.21 ± 0.22	7.20 ± 0.20
$C_i^\beta - C_{i+3}^\beta$ distance (Å)	5.62 ± 0.38	5.60 ± 0.35
$C_i^\beta - C_{i+4}^\beta$ distance (Å)	6.31 ± 0.29	6.33 ± 0.30
ϕ (deg)	–70 ± 11	–70 ± 11
ψ (deg)	–38 ± 11	–38 ± 11
ω (deg)	178 ± 10	178 ± 7
rms deviation (Å) from starting structure	0.44 ± 0.10	2.07 ± 0.09
end-to-end distance[a] (Å)	17.89	17.21

[a] Distance between the methyl carbons in the terminal acetyl and methyl amide groups, measured only in the last, instantaneous structure. The corresponding distances in the starting structures after minimization and NVT dynamics were 17.76 and 23.03 Å for the α- and 3_{10}-helices, respectively.

A sample of properties calculated during the MD simulations is given in Table I. Due to the change in conformation, the averages for the 3_{10} simulation were accumulated in the period from 20 to 100 ps. The similarity of the average values and their fluctuations indicates that, although started from significantly different structure, both simulations are sampling the same region of conformation space after the initial transition. The average values for the backbone dihedral angles ϕ and ψ and the marked predominance of $i, i+4$ hydrogen bonds identify this region as α-helical.

The free energy profile from the MC simulations for interconverting the all-3_{10}- and the all-α-helical forms of the same capped undecaalanine in water is shown in Figure 3. The convergence of the calculated free energies was ascertained by examination of the curve at periodic intervals over the last 12 × 10^6 configurations of averaging, and while there were only small changes after 4 × 10^6 steps, any differences were negligible beyond 8 × 10^6 configurations. Advantage has been taken of the similar ϕ values for α- and 3_{10}-helices[5] to permit convenient use of ψ as the reaction coordinate. Both the α and 3_{10} forms are found to be minima, with the α conformation lower in free energy by 10.6 ± 0.7 kcal·mol^{-1}, *ca.* 1.0 kcal·mol^{-1} per residue. The small barrier, 2.8 ± 0.3 kcal·mol^{-1}, for the 3_{10}- to α-helical transition is consistent with the behavior observed in the MD calculations. These quantitative results could change somewhat with inclusion of sampling over the ϕ and ψ angles, *i.e.*, in perturbing precisely between the α and 3_{10} minima.

The different components of the solute *intramolecular* energy calculated during the evaluation of the free energy profile are shown in Figure 4. As intuitively expected, the angle bending energy is higher in the 3_{10} form by 5.1 kcal·mol^{-1}, but this is to a large extent compensated by the torsional energy, which is lower by 3.8 kcal·mol^{-1}. The nonbonded energy, on the other hand, is lower in the α form by 20.6 kcal·mol^{-1}. The latter energy difference completely dominates the conformational preference of the peptide. The nonbonded energy has a profile very similar to that of the free energy in solution, as can be seen in Figure 4.

(20) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420–1426.
(21) Duffy, E.; Severance, D. L.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1992**, *114*, 7535–7542.
(22) van Gunsteren, W. F.; Mark, A. E. *Eur. J. Biochem.* **1992**, *204*, 947–961.

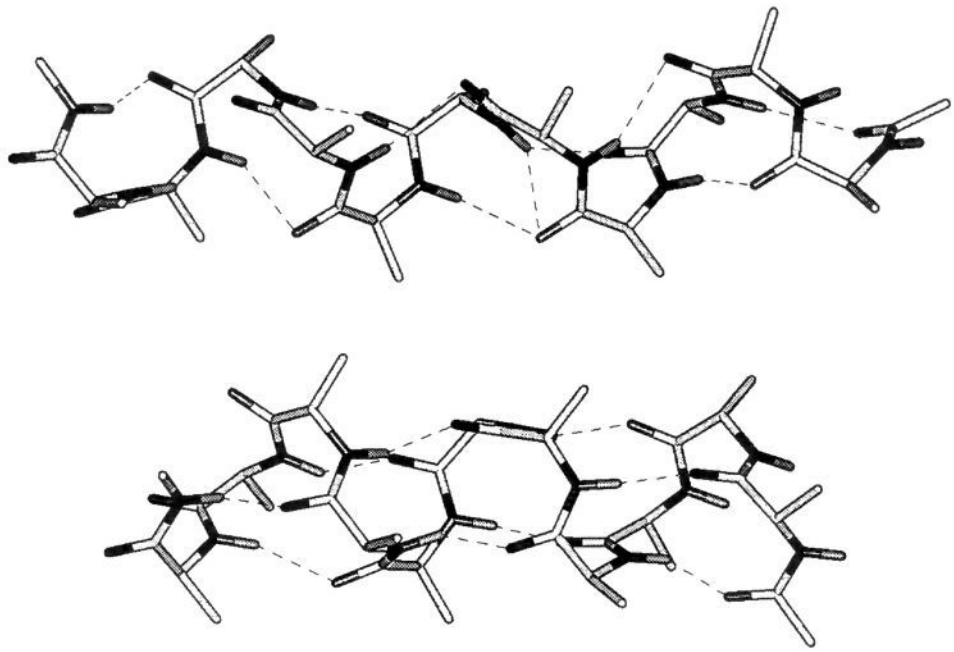

Figure 2. Initial (top) and final (bottom) structures from the 3_{10} simulation. The initial 3_{10} structure converts to an α-helix in 15 ps.

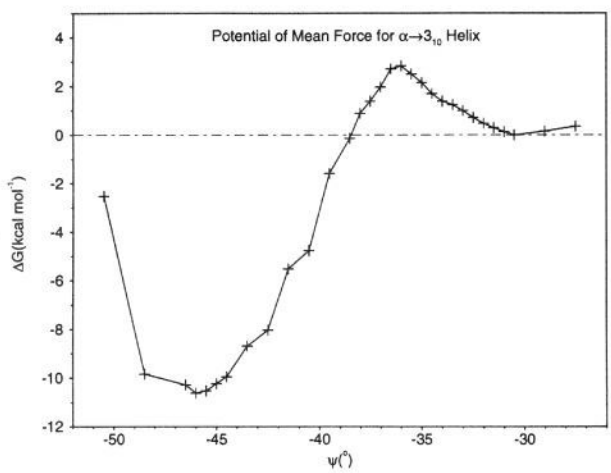

Figure 3. Free energy profile from the MC simulations for the 3_{10}- to α-helix interconversion.

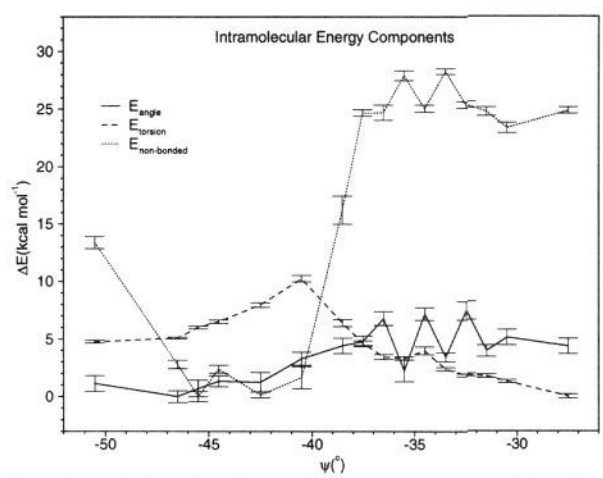

Figure 4. Relative values for the different components of the solute intramolecular energy during the 3_{10}- to α-helix interconversion.

According to these results, the 3_{10}-helix is intrinsically less stable due to increased unfavorable intramolecular electrostatic and steric interactions, in spite of containing an additional intramolecular hydrogen bond.

Calculation of the free energy profile for interconversion of the helical forms in the absence of solvent yields qualitatively similar results. The minimum for the 3_{10}-helix is shifted toward more positive values of ψ, the energy barrier is higher, 4.75 ± 0.10 kcal·mol^{-1}, and the difference in free energy between α and 3_{10} is also larger, 13.5 ± 0.43 kcal·mol^{-1}. Combination of the latter number with the results in water yields a $\Delta\Delta G$ of hydration that favors the 3_{10} form by 2.9 kcal·mol^{-1}, in spite of its having one less hydrogen-bonding site accessible to the solvent at each terminus. It should be noted that in order to obtain a correct thermodynamic cycle that accounts properly for any intramolecular effects, the free energy profiles in solution and in the gas phase were calculated using identical constraints (all $\phi = 60°$) and the same definition of the reaction coordinate. The use of the constraint on ϕ provides some uncertainty to the difference in free energies of hydration for the α and 3_{10} forms at their minima. Display of structures from both the MC and MD simulations (Figure 2) shows that the carbonyl groups splay away from the helix axis toward the solvent more in the 3_{10} form than in the α-helix. Quantitatively, from the MD simulations, the average solid angles between the C=O vectors and the helix axis are 33° and 19° in the 3_{10} and α regions, respectively. The increased solvent exposure helps compensate for increased strain in the 3_{10} hydrogen bonds. Overall, the present results indicate that the α-helical form for the capped undecaalanine is intrinsically much more stable than the 3_{10} alternative, and whether or not there is compensation by the solvent, it is not nearly enough to invert the relative stabilities. Interestingly, 3_{10}-helices are often thought to become less competitive with α-helices in more polar solvents;[23,24] however, this conclusion has been based on comparisons of a few short Aib-rich peptides in chloroform versus dimethyl sulfoxide. Though this is reasonable from consideration of the helix end interactions, these solvents do not provide hydrogen bond donation as found in water, which is critical for the interactions with the carbonyl groups.

(23) Marshall, G. R.; Hodgkin, E. E.; Langs, D. A.; Smith, G. D.; Zabrocki, J.; Leplawy, M. T. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 487–491.
(24) Balaram, H.; Sukumar, M.; Balaram, P. *Biopolymers* **1986**, *25*, 2209–2223.

Computational results from other workers using different peptides and force fields also lend indirect support to the predominance of the α-helix. The MD simulations of deca- and trisdecaalanine in water by DiCapua *et al.*[25,26] at 300 K indicated stable α-helical conformations with some terminal fraying. The MD simulation of capped tridecaalanine in water by Dagget and Levitt[27] at 278 K found *ca.* six or nine backbone hydrogen bonds to be α-helical. Furthermore, the potential of mean force for folding a capped alanine tripeptide in TIP3P water calculated at Tobias and Brooks[28] at 300 K shows local minima corresponding to 3_{10}- and α-helices with the latter lower by *ca.* 1.4 kcal·mol^{-1}, in good agreement with our results in spite of the large difference in the size of the system studied and the differences in the force field and simulation techniques.

The inconsistency of these results with the ESR data for the 3K peptides[7] is striking. In their study, the observed trend in the reciprocal line widths (δ^{-1}) of doubly spin-labeled peptides at different positions ($\delta^{-1}_{i,i+3} < \delta^{-1}_{i,i+2} < \delta^{-1}_{i,i+4}$) was compared to the trend in the distances between the corresponding β carbons C^{β}–C^{β} in idealized 3_{10}- ($R^{\beta\beta}_{i,i+3} < R^{\beta\beta}_{i,i+2} < R^{\beta\beta}_{i,i+4}$) and α-helices ($R^{\beta\beta}_{i,i+3} \approx R^{\beta\beta}_{i,i+4} < R^{\beta\beta}_{i,i+2}$). The C^{β}–C^{β} distances in Table I are somewhat different than those in the ESR study, though qualitatively the $i,i+2$ distance remains greater than $i,i+4$ for the α form, opposite to the 3_{10} case. Effects from, for example, the lower helicity of the 3K-(4,6) peptide (evident from the circular dichroism curves and not surprising since there are three consecutive non-Ala residues), the substitution of side chains, and the conformational mobility for the spin labels themselves warrant further investigation. It may be noted that even synthetic peptides of lengths over 10 residues containing as much as 50% Aib, a known 3_{10}-helix former, have been shown to favor α-helical conformations in organic solvents.[29]

Conclusion

The present results have addressed both the kinetic and the thermodynamic stabilities of helical forms of a capped undecaalanine in water. The molecular dynamics results found that, even after relaxation of the peptide and the solvent, the 3_{10}-helical form has marginal kinetic stability and converts in 15 ps to the α-helix. Similar simulations for the α-helix revealed no tendency for decay during 300 ps at 278 K. Monte Carlo simulations were then used to obtain the free energy profile for interconversion of the two helical forms. Both conformations appear to be local minima in water, with the α-helix lower in free energy by *ca.* 1 kcal·mol^{-1} per residue. The small computed barrier, 2.8 ± 0.3 kcal·mol^{-1}, for the 3_{10}- to α-helical conversion is consistent with the instability of the 3_{10}-helix found in the MD simulation. It seems likely that the qualitative preference for the α-helical form in water should extend to other alanine-based peptides, except perhaps at the shortest lengths. We hope that the present results will encourage further efforts to clarify this fundamental issue concerning helix structure in aqueous solution.

Acknowledgment. Gratitude is expressed to the National Institutes of Health (GM 32136) for support of this work and to the U.S. Department of Education for a graduate student fellowship (D.S.M.).

(25) DiCapua, F. M.; Swaminathan, S.; Beveridge D. L. *J. Am. Chem. Soc.* **1990**, *112*, 6768–6771.
(26) DiCapua, F. M.; Swaminathan, S.; Beveridge D. L. *J. Am. Chem. Soc.* **1991**, *113*, 6145–6155.
(27) Dagget, V.; Levitt, M. *J. Mol. Biol.* **1992**, *223*, 1121–1138.
(28) Tobias, D. J.; Brooks, C. L., III. *Biochemistry* **1991**, *30*, 6059–6070.
(29) Basu, G.; Kuki, A. *Biopolymers* **1992**, *32*, 61–71.