Helix Unfolding and Intramolecular Hydrogen Bond Dynamics in Small α-Helices in Explicit Solvent

C. J. Margulis, H. A. Stern, and B. J. Berne*

Department of Chemistry, Columbia University, 3000 Broadway, New York City, New York 10027

Received: February 25, 2002; In Final Form: July 3, 2002

The dynamics of folding and unfolding of secondary structural motifs found in proteins is crucial to understanding the protein folding problem. In this paper we study the short-time dynamics of loop unfolding and hydrogen bond formation and breaking in the alanine pentapeptide at room temperature, using several constant-energy molecular dynamics simulations of about 4 ns of duration. We analyze the results in terms of time histories of "core" α-carbon dihedral angles. We also perform a principal component analysis of the data. We find that the time scale for a considerable deformation of the dihedral angles formed by the α-carbons is on the order of 1 ns whereas hydrogen bonds seem to break and form on a shorter time scale.

1. Introduction

The thermodynamics of the helix−coil transition in polypeptides has been extensively studied in the past using simple statistical models such as those developed by Lifson and Roig[1] and Zimm and Bragg,[2] but the folding dynamics itself has only been experimentally and theoretically[3−7] probed during the past few years. Several ultrafast techniques have been used to study helix to coil dynamics. In some, the rate process is first initiated by a temperature jump[8−14] and the subsequent folding or unfolding is studied by monitoring transient IR absorption, UV−Raman or the fluorescence of a chromophore. A different technique is to initiate the transition with a pH jump, and monitor the time-dependent circular dicroism.[15] The results from these different experiments seem to be contradictory. For example, the time-dependent CD experiments[15] give a slow helix nucleation rate of 15 s^{-1}, whereas temperature-jump experiments[8,10] give a rate between 10^7 and 10^8 s^{-1} for peptides of similar size and amino acid sequence.

Recent constant-temperature molecular dynamics simulations of an alanine pentapeptide appear to indicate that helix nucleation occurs rapidly, on a time scale of 0.1 ns. This estimate is based on the analysis of the time dependence of the root-mean-square distance from a reference structure computed from principal component analysis (PCA).[3] To better understand the above disagreements, we performed a molecular dynamics study of the short-time dynamics of loop and hydrogen bond formation and breaking in the alanine pentapeptide in explicit solvent at room temperature. We focused on the formation and disappearance of a helical core[16] by monitoring two dihedral angles of backbone α-carbon atoms, and we monitored the formation and disappearance of the intramolecular hydrogen bond in the pentapeptide. Our analysis shows that core melting is much slower than 0.1 ns whereas hydrogen bond breaking and formation occurs on a time scale on the order of 100 ps. We find that even when the core is intact the intramolecular hydrogen bond can be made or broken many times. We also perform a PCA analysis of our simulations and find that the free-energy landscape as a function of the first two principal components closely resembles that computed from backbone α-carbon torsional angles. It is therefore reasonable to assume that the time history of these two torsional angles conveys most of the important dynamical features of the system.

2. Methods

In this study constant-energy molecular dynamics (MD) simulations of the peptide Ac−Ala_5−NHMe (Ala_5) using the program SIM developed in the Berne group[17] were performed. The system simulated contained one peptide molecule and 470 SPC water molecules in a cubic box with periodic boundary conditions. We used the particle−particle particle-mesh Ewald (P3ME) method[18,19] to treat the long-range electrostatic interactions. The reference system propagator algorithm (r-Respa) with two nested Trotter levels[18,20,21] was employed to minimize the number of force evaluations computed for slower degrees of freedom. The force field employed was the all-atom OPLS-AA potential of Jorgensen et al.[22] A time step of 2 fs was used for the slow degrees of freedom and 0.5 fs for the fast ones.

To emulate the conditions of the *T*-jump experiment, in which a short nanosecond near-infrared pump pulse increases the temperature of the region surrounding the helix, we started with Ala_5 in an α-helical conformation and performed a gas-phase energy minimization. The peptide was then inserted into a box of water that had been previously equilibrated at 300 K. Overlapping water molecules were removed, and the system was equilibrated for an additional 100 ps at 300 K using constant-temperature and pressure Andersen−Hoover molecular dynamics. In general, 100 ps is a sufficiently long time to equilibrate a box of water molecules in an MD simulation; however, it is still much shorter than the duration of the short-infrared pump pulse. It is therefore reasonable to assume that the configuration after this relatively short equilibrium represents the state of the system immediately after the pulse. Our production runs consisted of four independent constant-energy MD simulations, each of them 4 ns long; they are presumed to reflect the short-time behavior of the system after the external perturbation.

3. Results

The Ala_5 peptide studied is capable of forming one and a half loops when in a complete α-helical conformation. To study

10.1021/jp0205158 CCC: $22.00
Published on Web 09/25/2002

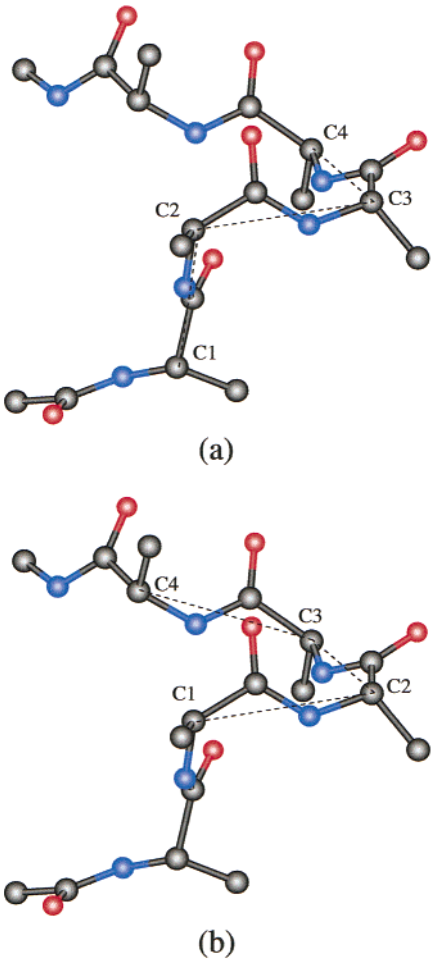

Figure 1. Snapshot of a configuration of the peptide studied showing the α-carbons involved in the dihedral angle ϕ_N closest to the N terminus (a). Snapshot of a configuration of the peptide studied showing the α-carbons involved in the dihedral angle ϕ_C closest to the C terminus (b).

the backbone dynamics, we monitor two dihedral angles, one closest to the N terminus, ϕ_N, and the other closest to the C terminus ϕ_C where each of them is defined by four contiguous backbone α-carbon atoms. We regard these dihedral angles as order parameters or reaction coordinates for the formation and destruction of the "helical core." These dihedral angles between core α-carbons are shown in Figure 1a,b, respectively. In terms of these two angles the system can be found in either of the following configurations: cis−cis, cis−trans, trans−cis, or trans−trans. The cis−cis configuration corresponds to a completely folded structure, whereas trans−trans corresponds to the completely extended or unfolded structure, cis−trans corresponds to an unfolded C terminus α-carbon dihedral angle, and trans−cis corresponds to an unfolded N terminus angle.

Figure 2 displays a contour plot of the free energy at 300 K computed as $-k_BT \log(P(\phi_N,\phi_C))$, where $P(\phi_N,\phi_C)$ is the joint distribution function of the dihedral angles and k_B is Boltzmann's constant. This free energy surface is based on the combined data from all of our constant-energy MD trajectories and corresponds to a total time of about 16 ns. Four basins corresponding to the aforementioned cis−cis, cis−trans, trans−cis, and trans−trans configurations can be distinguished. The size of the basin corresponding to trans−trans configurations is larger, although the cis−cis basin is slightly deeper. As will be made clearer in the next paragraph, the lifetime of these folded and unfolded states is on the order of nanoseconds; therefore, 16 ns might not be enough run time to achieve an exact statistical description of the basin's relative depths or

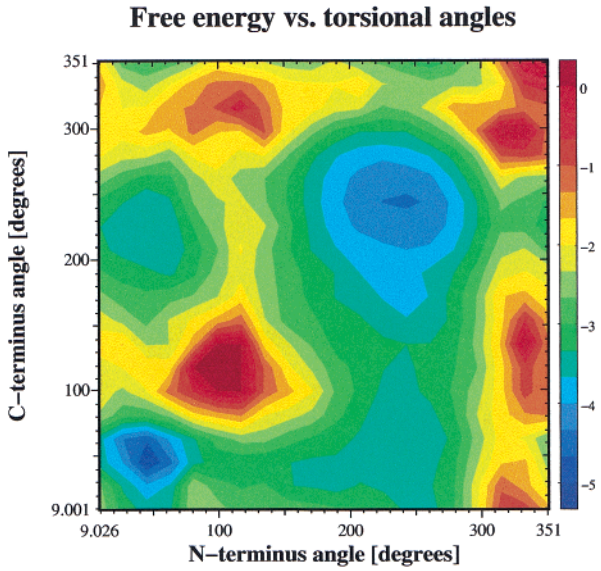

Figure 2. Free energy computed combining our four MD trajectories as a function of the two α-carbon torsional angles $\{\phi_N,\phi_C\}$. On the *X* axis is the torsional angle ϕ_N closest to the N terminus; on the *Y* axis is the torsion angle ϕ_C closest to the C terminus. The lower left basin corresponds to the cis−cis configuration, the upper right basin to the trans−trans configuration, and the upper left and lower right to the cis−trans and trans−cis configurations, respectively.

barrier heights. Improved statistics can be generated using the replica exchange method as described in ref 23 and references therein. In any case, we observe a barrier of about 1 or 2 kcal/mol to cross from cis−cis to trans−trans in the space of these reaction coordinates. The sizes of these barriers seem to be consistent with the PCA analysis presented later in this paper and with experimental results from refs 8−11 and 14.

Inspection of the free energy surface in Figure 2 is also helpful when trying to analyze the time history of the dihedral α-carbon angles. In most cases, during the course of our simulations, we observe that the system passes through the cis−trans or trans−cis configurations, shown as shallower minima in Figure 2, before unfolding into the trans−trans extended state. This can be established by inspecting Figure 3a, where during the first 1.4 ns of constant-energy MD, the system exists in the cis−cis conformation; immediately after this, the system can be found for 0.6 ns in the cis−trans conformation before finally unfolding to the totally stretched trans−trans unfolded state. In Figure 3d, however, it is the trans−cis configuration that is visited for about 0.5 ns as an intermediary between the cis−cis and trans−trans states. Figure 3c is somehow different in that both intermediate conformers, the cis−trans and the trans−cis, are visited intermitently, before finally unfolding into the trans−trans configuration at about 3.2 ns. We observe from these time histories of the torsional α-carbon angles that some transitions are sharp (they happen in 50 ps), whereas some others are slow and take 200 ps or more to occur. Note that these transient periods between conformers are long, even longer perhaps, than the helix or coil lifetimes reported in ref 3.

To understand the relationship between the *T*-jump experiments and the CD measurements, one should address the fundamental question of whether the hydrogen bond holds the loop together or if a stable loop structure can exist without the formation of this intramolecular hydrogen bond. Clearly, the transient infrared probe used in the *T*-jump experiments[8,10] can discriminate between the presence or absence of the intramolecular hydrogen bond, whereas CD measurements[15] sense the presence or absence of the helical core.

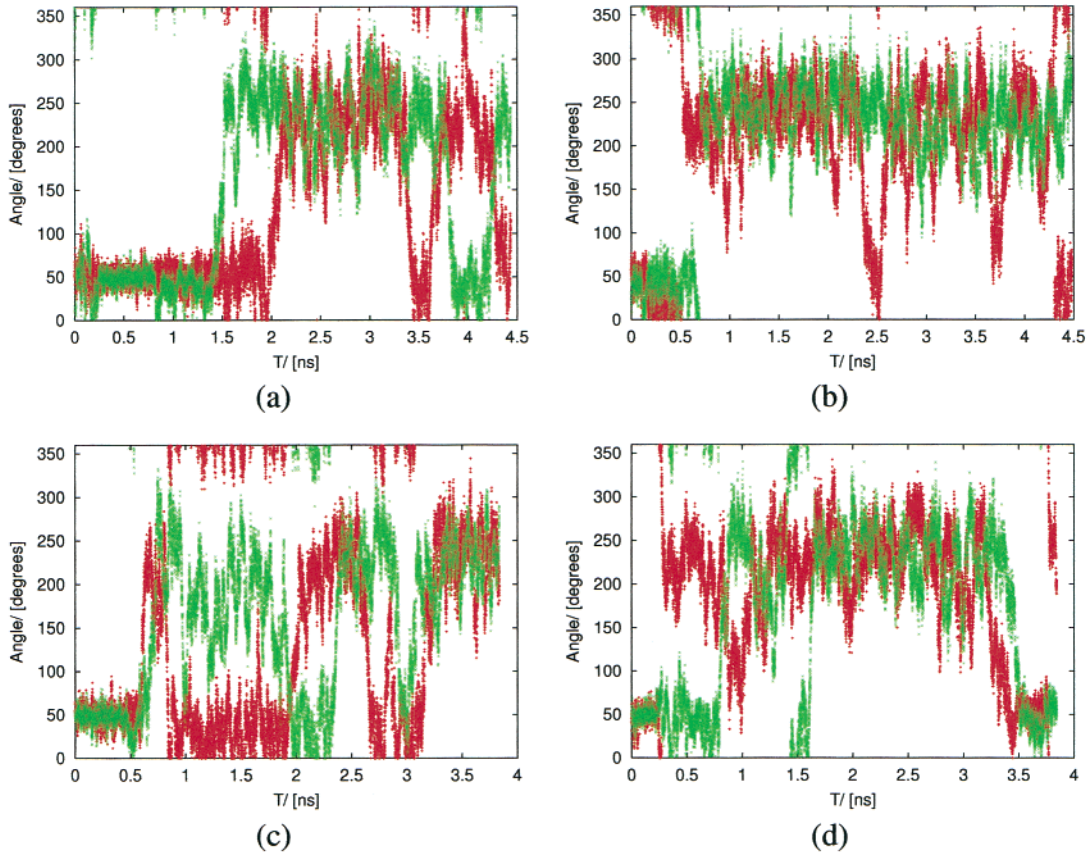

Figure 3. Time evolution of the two torsional α-carbon angles for each of the constant-energy MD trajectories: (red) the angle ϕ_N closest to the N terminus; (green) the angle ϕ_C closest to the C terminus.

Following Kabsch and Sander,[24] we monitor the formation and destruction of $(i, i + 4)$ intramolecular hydrogen bonds by computing the function

$$\epsilon = q_1 q_2 [1/r(\mathrm{ON}) + 1/r(\mathrm{CH}) - 1/r(\mathrm{OH}) - 1/r(\mathrm{CN})]f \quad (1)$$

ϵ is an order parameter used to define the intramolecular H bond and does not correspond to any term in the force field used to generate the molecular dynamics. Here $q_1 = 0.42$ and $q_2 = 0.2$ are in units of the electron charge, r(AB) is the interatomic distances from A to B in Å, $f = 332$ is a dimensional factor, and ϵ is in kcal/mol. ϵ is sensitive to the N−H−O angle and N−O distance. Values of ϵ smaller than −0.5 kcal/mol correspond to angles below 63° when the N−O distance is 2.5 Å, or to distances below 5.2 Å for the perfect 0° N−H−O alignment. An ideal intramolecular hydrogen bond has an N−O distance of approximately 2.9 Å, an N−H−O angle of 0°; and an energy of −3.0 kcal/mol. The intramolecular H bond is said to exist when $\epsilon \leq -0.5$ kcal/mol.

Three $(i, i + 4)$ intramolecular hydrogen bond pairs are consistent with an α-helix in the pentapeptide studied here. These hydrogen bonds are depicted in Figure 4. A typical time history of the evolution of the hydrogen bond function is displayed in Figure 5. The first point to notice is the absence of hydrogen bonding after 1.5 ns. This correlates with the begining of the unfolding process. It is obvious from this picture that no intramolecular $(i, i + 4)$ hydrogen bonds are observed in an unfolded helix. It is important to notice, however, that the absence of hydrogen bonding is not necessarily a sign of unfolding. If we concentrate on the period from 0 to 250 ps in Figure 5 we clearly observe that two of the three hydrogen bonds are broken. If we look at the same period of time in Figure 3a we find that although the torsional angles seem to have larger

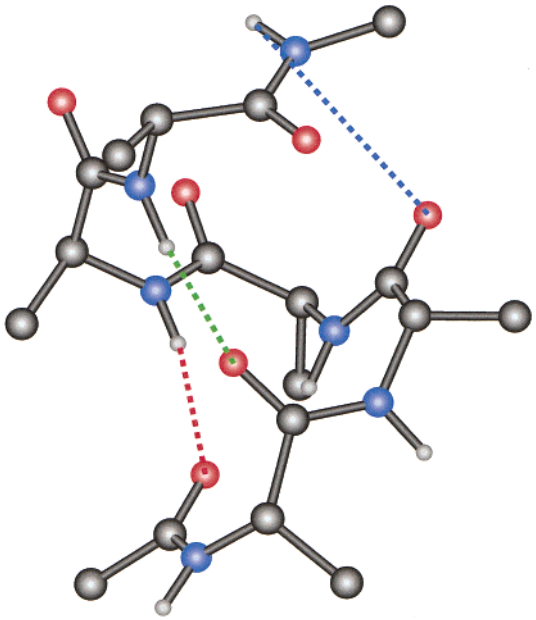

Figure 4. Snapshot of the gas-phase energy-minimized peptide studied showing the three $(i,i+4)$ hydrogen bonds, including the one closest to the N terminus (red), the one closest to the C terminus (blue), and the one halfway between the two termini (green).

fluctuations around their equilibrium structure than when all hydrogen bonds are formed, they still correspond to helical configurations. The system in fact remains a helix until about 1.5 ns. Some of these nonintramolecularly hydrogen bonded periods correspond to fast exchanges of intra- and intermolecular hydrogen bond acceptors. Proof of this is given by the radial distribution functions in Figure 6 between the hydrogen atom

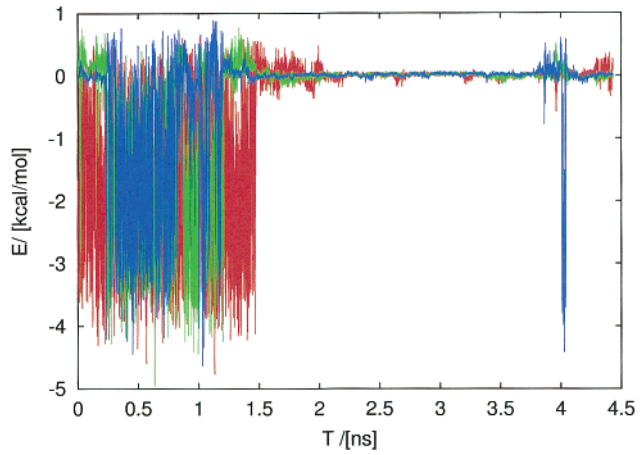

Figure 5. Time evolution of all possible (*i*,*i*+4) intramolecular hydrogen bond functions from trajectory one. In red is the hydrogen bond closest to the N terminus depicted in Figure 4, in green is the central hydrogen bond, and in blue is the hydrogen bond closest to the C terminus.

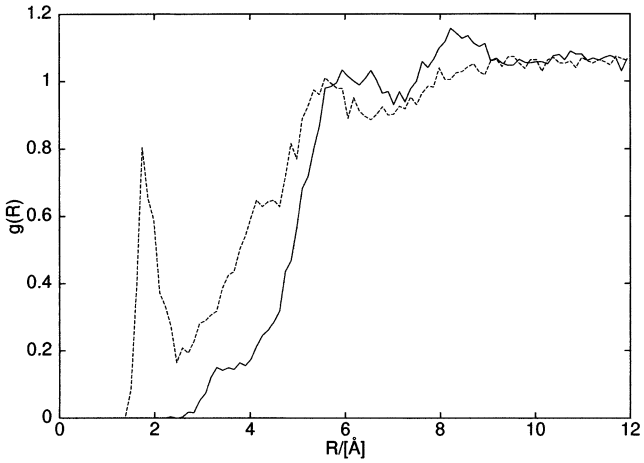

Figure 6. Water oxygen radial distribution function centered on the central amide hydrogen shown in Figure 4. The solid line corresponds to an intramolecularly hydrogen bonded period of about 200 ps, and the dashed line corresponds to the same duration of time but spans the time interval immediately previous in which the intramolecular hydrogen bond was absent.

making the second intramolecular hydrogen bond (green) in Figure 4 and the solvent's oxygen. Figure 6 plots $g(r)$ for the period in which no intramolecular hydrogen bond is present, and the period inmediately after in which the internal hydrogen bond is re-formed. It is clear from the plot that the peak at around 2 Å occurs during the period when there is no intramolecular H bond whereas it is absent in the period during which there is an intramolecular H bond. This behavior is exactly what one would expect when there is a switch to intermolecular hydrogen bonding to a water molecule. We conclude that there are sharp transitions between intramolecular hydrogen bonding and intermolecular hydrogen bonding. Although these events might be correlated with small changes of the core angles, they do not seem to alter the stability of the backbone appreciably. This can be established by observing that dihedral angle values consistent with the existence of an α-helical core exist even during periods where the intramolecular hydrogen bond is absent.

Monitoring the α-carbon dihedral angles shows that the helix−coil transition in alanine pentapeptide occurs on a time scale on the order of 1 ns. Studying the constant-temperature dynamics of this system with the AMBER force field at 300 K

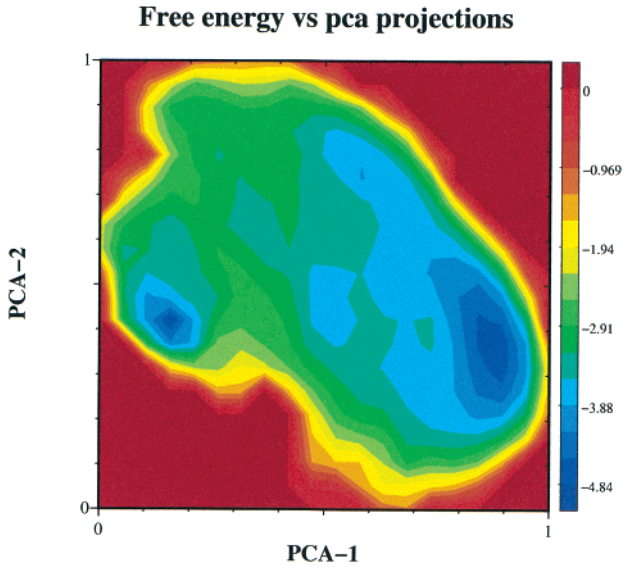

Figure 7. Free energy computed from PCA analysis using projections into the first and second principal component eigenvectors as the *X* and *Y* axes, respectively.

and other temperatures and analyzing the root-mean-square distance from a PCA structure allowed some investigators[3] to report that helix melting and formation occurs rapidly on a time scale of 0.1 ns. We believe that this difference can be explained by their criterion for unfolding. Their definition of melting is based on whether an aligned structure deviates by more than 0.6 Å from an average reference structure. If one looks carefully at their reported data, no stretched configurations (trans−trans) occur at 300 K even for times as long as 8 ns. Even more intriguing is the fact that stretched configurations do not occur even at higher temperatures. Thus their criterion for an unfolded structure does not correspond to a trans−trans structure and rather than the time for unfolding being 0.1 ns as claimed, it is much longer than the 1 ns we find for the OPLSAA potential. The fact that their dynamics is at constant temperature and ours is at constant energy could also influence the results, but we believe that the main reason is the force field. The AMBER force field appears to stabilize the α-helix. This raises again the fundamental question of how force fields are engineered and whether polarizability needs to be included in the parametrization.

For the principal component analysis of the data collected from the four trajectories, we proceeded in the following fashion: only backbone heavy atoms were considered; that is, hydrogen atoms and methyl groups were left out of the calculation. We took conformations every 0.2 ps, making a total of about 90 000 snapshots of the peptide. Each of these configurations was aligned to the energy-minimized α-helical initial structure, with minimal RMSD using the algorithm in ref 25.

Once the alignment procedure is acomplished, we build and diagonalize the covariance matrix, obtaining eigenvalues and eigenvectors. There seems to be a most important eigenvalue that is about 10 times larger than any other. Other eigenvalues, except from the second and third perhaps, are much smaller, being only about 1% or less of the first one. In Figure 7 we show the calculated free energy at 300 K resulting from the projection of all structures into the first two principal component axis. As one can establish from examining Figure 7, the two deepest minima lie at oposite ends of the first principal component axis. The cis−cis conformers are on the left, and the trans−trans conformers are on the right. The other conform-

ers, cis−trans and trans−cis appear as shallower minima, on the upper center and center of the graph. To visualize the multidimensional direction of the most important PCA component, we perform small displacements on the gas-phase energy-minimized structure along this first PCA eigenvector. We observe that this corresponds to a "screw-like" unfolding motion of the molecule.

The resemblance of the PCA free-energy landscape with the one computed from the torsional angles is remarkable. Barriers from cis−cis to trans−trans are on the order of 1 or 2 kcal/mol, the same value obtained from the free energy as a function of α-carbon torsional angles. Basin depths for cis−cis and trans−trans are also similar to the ones obtained from torsional angles. This leads to the conclusion that the system can indeed be thought of as a combination of two simple "most important" variables, the two α-carbon dihedral angles, each fluctuating according to the dynamics imposed by the interaction with the surrounding solvent and able to be found in any of two discrete configurations (cis or trans).

4. Conclusions

This paper focuses on the time scales for intra- and intermolecular hydrogen bond formation and destruction and for core melting in a small alanine polypeptide. We also studied the energetics for unfolding from a totally folded, cis−cis, structure, to a stretched, trans−trans, configuration. Starting from a perfect helical configuration, each system was equilibrated at 300 K using constant-temperature molecular dynamics to mimic a temperature-jump pump pulse, the subsequent short-time evolution was monitored by four constant-energy MD simulations, each of them 4 ns long. Time-resolved infrared measurements after a *T*-jump[10] predict that unfolding of a mainly alanine 21-peptide occurs on a 10−100 ns scale. We find in our pentapeptide that although the intramolecular hydrogen bond breaks and forms on a time scale faster than nanoseconds, the dihedral angles defining the core helix in our small system can be preserved up to a period of about 2 ns. This may be consistent with the *T*-jump observation. Time-dependent infrared measurements after a temperature-jump probe the local structure of a molecule, specifically the intramolecular hydrogen bonds between amides and carbonyls. Our simulation shows that these bonds are weak compared to the energy required to disassemble a single turn helix or that, at least, the time scale in which these hydrogen bonds form and break is probably much faster than core melting. Hence the question arises whether it is possible that in the aforementioned experiments a propagation barrier (i.e. fraying of helix termini and cleavage of intramolecular hydrogen bonds) and not the nucleation barrier is crossed.[15] To address this point, simulations at least 1 or 2 orders of magnitude longer on larger systems than studied here could be performed.

In this paper we use α-carbon dihedral angles to monitor the helix−coil transition in a small alanine pentapeptide and report a time scale on the order of 1 ns. Others have instead run the AMBER force field for this system using constant-temperature dynamics at 300 K and other temperatures and used the RMSD from a PCA structure[3] to monitor the time scale for helix melting and formation. They report that the transitions occur rapidly on a time scale of 0.1 ns. The reason for this difference appears to spring from their criterion for helix melting. Their definition of melting is based on whether an aligned structure has an RMSD greater than 0.6 Å from an average reference structure. We find that under this criteria their "unfolded structure" is not necessarily a trans−trans structure. From their reported data, no configurations resembling a trans−trans structure are observed at 300 K for 8 ns, and, even more intriguing, no trans−trans conformation occurs even at higher temperatures. The unfolding time scale for AMBER is thus not 0.1 ns, as reported, but much longer than the 1 ns that we observed for the OPLSAA system. It appears that the AMBER force field predicts much more stable and long-lived α-helices than OPLSAA. Given the large differences predicted by different force fields, it is essential that more work be done comparing force fields with each other and with experiment. This raises again the fundamental question of how force fields are engineered and whether polarizability needs to be included in the parametrization.

To address whether temperature-jump experiments do actually measure core melting and not just hydrogen bond dynamics, shifts in the time-dependent vibrational frequencies of carbonyl bonds can be computed during a molecular dynamics simulation. To calculate the actual spectra, an ensemble of constant-energy molecular dynamics simulations must be run with initial conditions obtained from a Boltzmann distribution generated, for example, from a constant-temperature MD run. To extract data that can be directly compared with experiments, the ensemble of trajectories would need to run for microseconds. This is still computationally very expensive but faster computers, better algorithms for treating the electrostatics, and parallelization should make this task feasible soon as long as force fields are reliable enough to deliver the correct results.

Acknowledgment. This work was supported by Grant GM 43320 from the National Institute of Health.

References and Notes

(1) Lifson, S.; Roig, A. *J. Chem. Phys.* **1961**, *34*, 1963−1974.
(2) Zimm, B. H.; Bragg, J. K. *J. Chem. Phys.* **1959**, *31*, 526−535.
(3) Hummer, G.; Garcıa, A. E.; Garde, S. *Proteins: Struct., Funct., Genet.* **2001**, *42*, 77−84.
(4) Brooks, C. L. *J. Phys. Chem.* **1996**, *100*, 2546−2549.
(5) Elmer, S.; Pande, V. S. *J. Phys. Chem. B* **2001**, *105*, 482−485.
(6) Buchete, N. V.; Straub, J. E. *J. Phys. Chem. B* **2001**, *105*, 6684−6697.
(7) Shea, J. E.; Brooks, C. L. *Annu. Rev. Phys. Chem.* **2001**, *52*, 499−535.
(8) Thompson, P. A.; Eaton, W. A.; Hofrichter, J. *Biochemistry* **1997**, *36*, 9200−9210.
(9) Thompson, P. A.; Muñoz, V.; Jas, G. S.; Henry, E. R.; Eaton, W. A.; Hofrichter, J. *J. Phys. Chem. B* **2000**, *104*, 378−389.
(10) Williams, S.; Causgrove, T. P.; Gilmanshin, R.; Fang, K. S.; Callender, R. H.; Woodruff, W. H.; Dyer, R. B. *Biochemistry* **1996**, *35*, 691−697.
(11) Lednev, I. K.; Karnoup, A. S.; Sparrow, M. C.; Asher, S. A. *J. Am. Chem. Soc.* **2001**, *123*, 2388−2392.
(12) Werner, J. H.; Dyer, R. B.; Fesinmeyer, R. M.; Andersen, N. H. *J. Phys. Chem. B* **2002**, *106*, 487−494.
(13) Jas, G. S.; Eaton, W. A.; Hofrichter, J. *J. Phys. Chem. B* **2001**, *105*, 261−272.
(14) Huang, C.; Klemke, J. W.; Getahum, Z.; DeGrado, W. F.; Gai, F. *J. Am. Chem. Soc.* **2001**, *123*, 9235−9238.
(15) Clarke, D. T.; Doig, A. J.; Stapley, B. J.; Jones, G. R. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 7232−7237.
(16) Levitt, M.; Greer, J. *J. Mol. Biol.* **1977**, *114*, 181−239.
(17) Sim molecular dynamics simulation program. Stern, H.; Xu, H.; Harder, E.; Rittner, F.; Pavesse, M.; Berne, B. J.
(18) Zhou, R.; Harder, E.; Xu, H.; Berne, B. J. *J. Chem. Phys.* **2001**, *115*, 2348−2358.
(19) Luty, B. A.; Tironi, I. G.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *103*, 3014.
(20) Tuckerman, M.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990−2001.
(21) Stuart, S. J.; Zhou, R.; Berne, B. J. *J. Chem. Phys.* **1996**, *105*, 1426−1436.
(22) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
(23) Zhou, R.; Berne, B. J.; Germain, R. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14931−14936.
(24) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577−2637.
(25) Arun, K. S.; Huang, T. S.; Blostein, S. D. *IEEE Transactions on pattern analysis and machine intelligence* **1987**, *PAMI-9 No. 5*, 698−700.