Available online at www.sciencedirect.com
SCIENCE DIRECT®
Polymer 45 (2004) 669–676
polymer
www.elsevier.com/locate/polymer

Characterization of non-alpha helical conformations in Ala peptides

Angel E. Garcia

Theoretical Biology and Biophysics Group, T10 MS K710, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

Received 28 April 2003; received in revised form 17 September 2003; accepted 17 September 2003

Abstract

The folded (alpha helical) and unfolded (non-alpha helical) ensembles of the 21 amino acid peptide Ace-A21-Nme are characterized structurally. The replica exchange molecular dynamics approach is used to generate these ensembles at 46 different temperatures ranging from 278 to 487 K. Each replica system is simulated in explicit solvent for a period of 10 ns/replica, for a total of 460 ns. In addition to alpha helices, poly proline II (PPII) structures were identified to occur significantly. At low T the alpha helical content is larger than the PPII content, but near 300 K the PPII population is larger. Below 300 K, the PPII population increases with T, but it decreases above 300 K. The alpha helical content decreases with temperature. At temperatures below 300 K, there is a PPII propagation free energy that enhances the formation of long segments of PPII structure. This propagation term is smaller than for alpha helices. PPII segments of length 8 or less are more likely to form than alpha helices of the same length. The obtained low propensity for the formation of PPII segments of length shorter than five suggest that the interactions are responsible for the formation of PPII structures require a PPII segment of at least four amino acids. Stretches of four consecutive amino acids in the PPII conformations are needed for the formation of a groove around the peptide backbone that is strongly hydrated. Water in this groove is delocalized along a channel formed by the peptide in the PPII conformation.

Keywords: Alpha helices; poly proline II; amino acid

Alpha helices and beta sheets are the most common secondary structural motifs found in folded proteins. The propensity of helix formation is largely a local property of the peptide sequence, while the formation of beta hairpins and beta sheets is determined by tertiary interactions in the folded state. Detailed time-resolved studies of the helix coil transition in peptides [1,2] have shown that helix formation occurs in the 200 ns time scale, while the formation of beta hairpins occurs in the microsecond time scale [3]. The fast rate of protein folding has been attributed to a funneled energy landscape where the large configurational entropy associated with the chain is balanced by intramolecular and solvent interactions along a multitude of folding paths [4–8]. Although the native state of proteins and peptides is well known, little is known about the unfolded state.

The unfolded state of a peptide may adopt a wide range of conformations, with some preference for particular regions of the $\Phi-\Psi$ map. Tiffany and Krimm, based on CD measurements, suggested that a significant fraction of the unordered structure has short segments (4–7 amino acids) of poly prolyne II (PPII) conformations, interspersed with sharp bends [9]. PPII is a structure adopted by poly proline peptides, and collagen. An ideal PPII structure has backbone dihedrals $(\Phi, \Psi) = (-75, 145)$, leading to an extended, left handed helix, with three amino acids per turn. Host guest experiments show the propensity for PPII in PXP to follow P > Q > A > (G, L, M) > N > (I, V) [10]. In globular proteins, however, the propensity of PPII conformation is large for A and V amino acids, but not for G [11,12]. The lack of correlation between the host guest experiments and the observation in globular proteins have been suggested to be due to the backbone hydration present in peptides, and not in globular proteins [13]. Monte Carlo calculations on a simple model of the Ala_7 showed that the PPII, and a structure close to α' occupy the lowest energy on the $\Phi-\Psi$ map [14]. Ab initio calculations in of Ala dipeptide with four explicit water molecules and a continuum treatment for other water molecules show that the PPII conformation is favored [15]. Experiments of alanine tripeptide show a large propensity for the PPII conformation [16–19]. The population of PPII over beta conformation depends on the solvent, with PPII being favored in water and the beta conformation favored in DMSO [18]. Combined experimental and theoretical work

E-mail address: angel@atlas.lanl.gov (A.E. Garcia).

0032-3861/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.polymer.2003.10.070

show that the PPII in an Ala tripeptide is 80% PPII, and 20% alpha [20]. Shi et al. have shown that unfolded heptapeptides adopt a PPII structure, and have suggested that the helix coil transition is an order-to-order transition, where the unfolded state is mostly PPII [21]. This work has raised many questions: is the formation of PPII higher than expected from random configurations? Will long-range ordered PPII structures form? Is there a PPII nucleation energy? Is there a PPII propagation energy?

In this manuscript, we study the folded (helical) and unfolded states of Ac-A21-methyl amide (Nme) in water, and over a wide range of temperatures. The ensemble of states were generated in a previous work where we applied the replica exchange molecular dynamics method (REMD) [22,23] to study the equilibrium folding/unfolding thermodynamics of Ac-A21-Nme and the Fs peptide (Fs stands for folded short peptide with sequence Ac-AAAAA(AAARA)$_3$ A-Nme designed by Lockhard and Kim [24]) over a broad range of temperatures [25]. We use the ensemble of configurations, over a broad range of T, previously generated for Ac-A21-Nme to characterize the folded and unfolded states in terms of their structures, $\Phi-\Psi$ maps, end-to-end distance distributions, and hydration. In our previous calculations, we found that a small modification in the dihedral angle energy Cornell et al. [26] force field was needed to obtain a better description of the helix coil thermodynamics, when compared to experiments. The modified Amber force field eliminates the backbone dihedral angle energy terms for rotations around the Φ and Ψ angles (i.e. the energy term is set to zero). With the modified Amber force field (parm_mod) the transition temperature for Ala based peptides, the Zimm–Bragg helix nucleation, and propagation parameters were in good agreement with measurements [27]. In this manuscript, we show that the propensity of PPII configurations observed is significantly larger than expected from random distributions. We show that segments of at least four consecutive amino acids in the PPII conformation are stable. We also show that segments of PPII of length 8 are common, and that segments as long as 17 amino acids are formed at low temperature. We will also show that water forms a delocalized channel around the PPII backbone. This water channel may play a role in stabilizing the PPII conformation.

1. Methods

The replica exchange method has been implemented with M replicas distributed over M processors, where each replica system is a peptide in explicit solvent. Each system is simulated at different T distributed over a broad range (275–487 K). Replicas are sorted according to T and couple to each other via a temperature exchange Monte Carlo (MC) procedure, described next. At fixed time intervals, neighboring systems, i and j, with temperatures T_i and T_j, respectively, can exchange configurations, such that system i changes to temperature T_j, and system j to temperature T_i. The probability that this exchange occurs satisfies detailed balance:

$$W(X)w(X,X') = W(X')w(X',X),$$

where $W(X)$ is the weighting factor for the state X, and $w(X,X')$ is the transition probability of exchanging system X to system X'. $W(X)$ is given by the product of the Boltzmann factors for each of the M replicas. In the exchange process, we scale the velocities of all particles by $\sqrt{(T_i/T_j)}$ in system j, and $\sqrt{(T_j/T_i)}$ in system i, such that there are no net changes in the total kinetic energy. This gives

$$w(X,X')/w(X',X) = \exp(-\Delta),$$

where $\Delta = (\beta_i - \beta_j)(U_j - U_i)$, β_i is $(k_B T)^{-1}$, k_B is the Boltzmann constant, and U_i is the potential energy of system i, before the exchange. These transition probabilities are implemented using the Metropolis criterion [22,23]. Only replicas having neighboring temperatures are allowed to exchange. The direction of exchange of neighboring replicas is chosen at random.

The Ac-A21-Nme peptide consists of 21 amino acid and 222 atoms. This system is contained in a cubic box containing 2640 TIP3P water molecules [28]. The initial configurations for the replicas are generated by a 1.0 ns simulation of the extended peptide, in vacuo, at 1000 K. Configurations sampled during the last 0.9 ns are clustered based on their pairwise root mean square distance (*rmsd*). We generate 100 clusters, and select 46 structures from different clusters. Each one of the structures is at least 2.5 A in *rmsd* from any structure belonging to other clusters. No biases are imposed regarding secondary structure or energy. The selected configurations are solvated and assigned as initial configurations for the 46 replicas. The assignment of configurations to each of the chosen temperatures is done at random. The solvated system is immersed in a cubic box is 43.5 Å, which is 1.5 times the linear dimension of the folded alpha helix.

The solvated systems are subjected to 500 steps of steepest descent energy minimization and a 100 ps MD simulation at constant NPT, with $P = 1$ at. and $T = 300$ K. We used the modified Force field of Cornell et al. [25,26] and the suite of programs in Amber [29], modified to include the generalized reaction field (GRF) treatment of electrostatic interactions [30,31] and the replica exchange algorithm. We used the GRF with a cut-off of 8.0 A. Non-bonded pairs lists were updated every 10 integration steps. The integration step in all simulations is 0.002 ps. The system is coupled to an external heat bath with relaxation time of 0.1 ps [32]. All bonds involving hydrogen atoms were constrained using SHAKE with a tolerance of 0.0005 A. The final configuration of the MD simulation was then used to execute 100 ps MD simulation at NTV at $T = 275$, 300, 325, 350, 400, 450, 500 and 550 K. The volume of the system was maintained to be the final volume obtained from the initial MD NPT simulation

at 300 K. These runs were used to determine the average potential energy of the system as a function of T, which is then used to determine the set of temperatures to be employed in the replica exchange method. The exchange rate among replicas ranges from 8 to 20%, with an overall exchange rate of 14%. Exchanges are attempted every 125 integration steps (0.25 ps). We simulated 46 replicas with $T = 278–487$ K, spanning a range of T in which the peptide samples folded and unfolded states. We simulated this system for 10 ns/replica, with 46 replicas, for a total of 460 ns of MD sampling.

The REMD methods has been applied to study the energy landscape of peptides in explicit solvent models [33,34]. Side-by-side comparisons of REMD with molecular dynamics (MD) simulations have shown that the REMD provides a much broader sampling of configurational space than regular MD simulations that cover the same time scales [35].

2. Analysis

Equilibrium quantities, based on ensemble averages at each temperature, are evaluated over the second half of the 10 ns/replica simulations. Block averages over 0.25 ns were performed to estimate errors in the ensemble averages. The hydration of the peptide structure around the PPII configurations is analyzed over the last 2 ns/replica of the simulation.

We analyze the configurations generated by the replica exchange simulations in terms of the backbone dihedral angles $\Phi–\Psi$ maps and the sequence helical content. To analyze the alpha helical content of the peptide, the sampled configurations are labeled by a sequence of 21 h or c's, depending on the Φ and Ψ angles of each amino acid. We label (Φ, Ψ) pairs as h's if the angles are $(-60 \pm 30, -60 \pm 35)°$, and as c's, otherwise. Calculations of the helical content follow the Lifson–Roig theory. An N- and C-capped 21 amino acid sequence can have a maximum helical length of 19, and 21 amino acids are characterized as h's or c's [36–38]. Configurations are labeled by their h–c patterns.

We perform a similar analysis for PPII conformations, where we label (Φ, Ψ) pairs as p's (representing PPII configurations) if the angles are $-120 \leq \Phi \leq -30°$, and $60 \leq \Psi \leq 180°$, and 0's otherwise. We label conformations as beta if the angles are $-180 \leq \Phi < -120°$, and $60 \leq \Psi \leq 180°$. The PPII content is calculated assuming that an ordered PPII segment consists of three or more consecutive amino acids in the PPII conformation. This choice is arbitrary, but consistent with the criteria used for alpha helical segments. In analyzing the conformations of the peptide, we adopt a binary code for describing the conformations as alpha or non-alpha, PPII or non-PPII, and beta or non-beta. In this code, there are 2^{21} different conformations. We use the patterns observed in this set as a reference state in the statistical analysis of the propensities of the alpha and PPII conformations.

We calculate the ${}^3J_{\mathrm{HN}\alpha}$ coupling coefficients as a function of T for all amino acids. We use the Karplus relation [39] between the ${}^3J_{\mathrm{HN}\alpha}$ coupling coefficients and the Φ dihedral angle using ${}^3J_{\mathrm{HN}\alpha} = 6.51\cos^2\theta - 1.76\cos\theta + 1.6$, with $\theta = |\Phi - 60°|$ [40].

We calculate the density of water around PPII structures of length 5. We use the $\Phi–\Psi$ angles to identify configurations that have PPII segments of length 5 of more. Each segment of five is aligned to an ideal PPII structure by performing rotations and translation of the whole system. The local density of water is calculated in a 1 Å grid in cube covering ± 20 Å from the segment center of mass. These calculations are done for all 46 temperatures. The triplet correlation function of central two consecutive carbonyl oxygen atoms in the backbone of the PPII segment and water are calculated. The calculation of the triplet correlation functions for water have been described before [41,42].

3. Results

We first describe the helix coil transition for the A21 peptide, as described by the REMD simulations. In Fig. 1, we show the alpha helical and PPII content profile for A21 as a function of temperature. This alpha helix profile have been reported earlier [25]. We define the midpoint of the helix coil transition as the temperature at which the helical content is 0.5. For A21 this T is below 273 K. The helical content of A21 at 278 K is 34%. Spek et al.'s data on A13 flanked by charged amino acid side chains is near 40% at 273 K in 1 M NaCl solution [27]. The Zimm–Bragg helix propagation and nucleation parameters for A21 are $w = 1.37$ and $v = 0.076$ at 273 K, which correspond to $s = 1.3$, and $\sigma = 0.004$. Fig. 1 also shows the PPII content profile as a function of T. This profile shows that the PPII content first increases slightly with T, up to $T \sim 300$ K, and then decreases. The non-monotonous shape of this curve

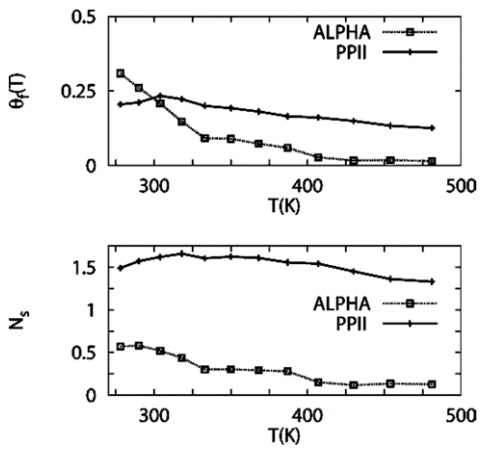

Fig. 1. Alpha helical and PPII content profiles (top), and average number of continuous alpha helical and PPII segments (bottom), as a function of T for the A21 (Ac-A21-methyl amide).

suggests that two competing effects affect the stability of PPII. At $T \sim 294$ K the PPII content equals the alpha helical content. At $T \sim 350$ K the PPII + alpha helical content add to 28%, and other, disordered, conformations add to 72%, showing that the helix-coil transition is an order to disorder transition, but with the caveat that the ordered state consists of two populations, alpha helices and PPII.

The bottom curve in Fig. 1 shows the average number of continuous alpha helical and PPII segments, as a function of T. For alpha helices, we obtain an average number less than one, indicating that the average number of helical segments is dominated by chains with zero or one continuous segments. This result is consistent with a high alpha helical nucleation energy. For PPII segments, we obtain an average number of 1.5, suggesting chains with one or two PPII segments dominate the ensemble. This result is consistent with the existence of a PPII nucleation energy that is smaller than the corresponding energy for alpha helices.

Next, we will characterize the sampled ensembles in terms of structural parameters. Fig. 2 shows the Φ–Ψ map of the N-terminal amino acid in A21 at low (278 K) and high (487 K) temperatures. The N- and C-terminal amino acids are the most disordered at all temperatures. From these maps, we can clearly identify six basins, identified as: (i) alpha, centered near $(\Phi, \Psi) \sim (-60, -60)$; (ii) PPII centered near $(\Phi, \Psi) \sim (-60, 120)$; (iii) beta, centered near $(\Phi, \Psi) \sim (-150, 120)$; (iv) alpha′, centered near $(\Phi, \Psi) \sim (-150, -90)$; (v) alpha-L, centered near $(\Phi, \Psi) \sim (60, 60)$, and (vi) centered near $(\Phi, \Psi) \sim (60, -90)$. The alpha, beta, and PPII configurations are the most populated. At low temperature the population is dominated by PPII and alpha configurations. At high temperature the population is still dominated by PPII, followed by beta and alpha configurations. Fig. 3 shows the population of beta, PPII, and alpha configurations by the central amino acid in A21, and the N- and C-termini, as a function of temperature. From these plots we see that the

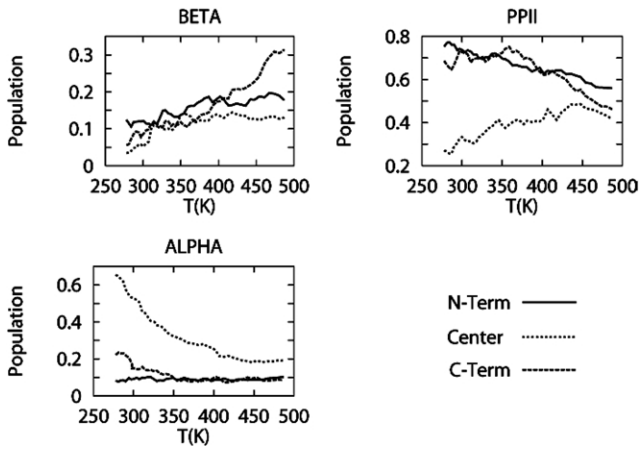

Fig. 3. Populations of configurations sampling the alpha, beta, and PPII regions of the Φ–Ψ map as a function of temperature for the N-terminal, C-terminal and center amino acids in Ac-A21-Nme.

population of beta configurations increases with T, the population of the PPII configurations for the N- and C-termini decreases with T, while for the central amino acid it increases. The population of the alpha configurations for the central amino acid decreases with T, while it decreases for the C-terminal amino acid, and remains constant for the N-terminal amino acid. Notice that for the N- and C-termini amino acids the populations are 70–75% PPII, 8–20% alpha, and 5–10% beta, at 278 K. This is not far from the 80% PPII, 20% alpha (or not PPII) populations measured by pump probe IR spectroscopy on tripeptides [17].

Fig. 4 shows the $^3J_{\mathrm{HN}\alpha}$ coupling coefficients for amino acids Ala 2, 3, 10, 11, 15, 16, 19 and 20, as a function of temperature. We find that the $^3J_{\mathrm{HN}\alpha}$ near the N- and C-termini change from 5.2 to 5.8, while for the central amino acids it changes from 4.6 (corresponding to alpha helices) to 5.8, over the same T range. The $^3J_{\mathrm{HN}\alpha}$ at the termini, and at high T are in reasonable agreement with

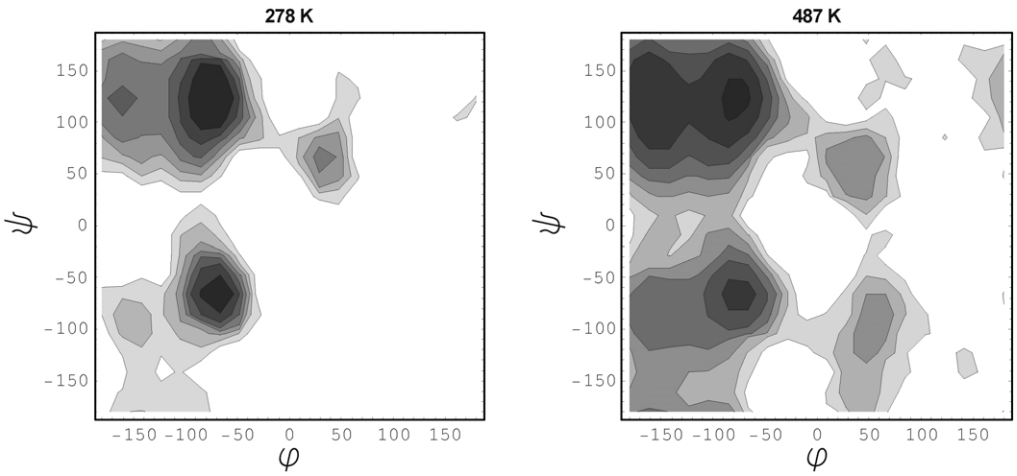

Fig. 2. Contour maps of the population of Φ–Ψ angles for the N-terminal amino acid at two temperatures, 278 and 487 K. The contours are plotted in a logarithmic scale, in $k_B T$ units. We identify six basins near $(\Phi, \Psi) \sim (-60, -60)$ (alpha), $(-60, 120)$ (PPII), $(-150, 120)$ (beta), $(-150, -90)$ (alpha′), (60, 60) (alpha-L), and $(60, -90)$.

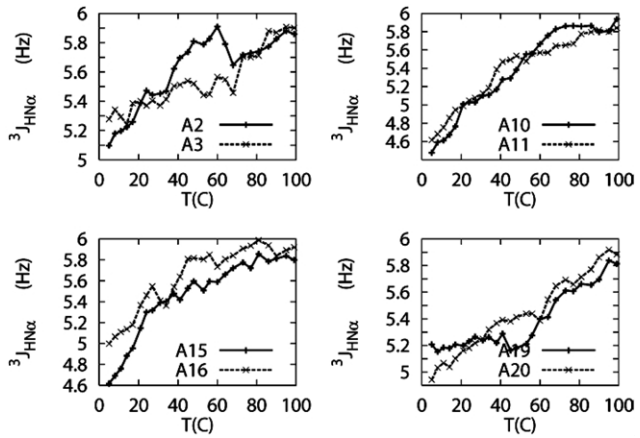

Fig. 4. $^3J_{HN\alpha}$ coupling coefficients as a function of temperature for selected amino acids in A21.

the measurements by Shi et al. in shorter peptides [21]. The $^3J_{HN\alpha}$ data by Shi et al. offered some evidence for the formation of ordered PPII structures. However, the $^3J_{HN\alpha}$ couplings measure the average coefficient for an amino acid and do not provide information about the conformation of the neighboring amino acids. Given the presence of various basins in the Φ–Ψ maps, the large propensity of individual amino acids for PPII and alpha configurations does not imply that long-range ordered structures will be formed.

To characterize the structural ensemble sampled by A21, we calculate the histograms of the number of PPII and alpha conformations of lengths 1–19, without having any amino acid in between in another conformation. The negative of the logarithm of the probability distributions obtained from these histograms define the potential of mean force (pmf), $W(N)/k_BT$, for forming PPII or alpha structures of length N. The pmfs at four temperatures (278, 300, 326 and 350 K) are shown in Fig. 5. The pmf profiles show the free energy for forming long-range ordered structures, relative to each other. At temperatures 278–326 K the free energy to form helices of length 12–16 is lower than to form PPII structures. However, the free energy of PPII structures of

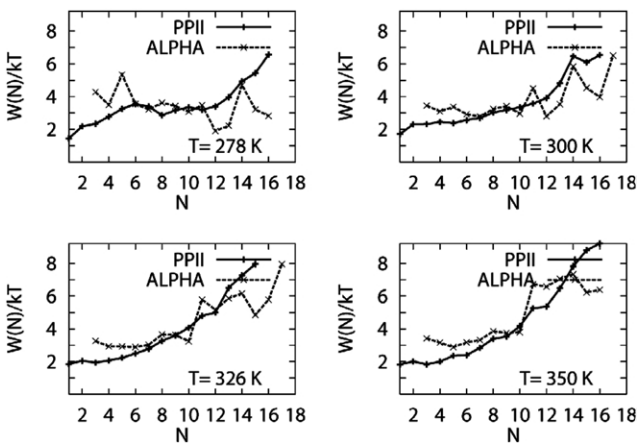

Fig. 5. Potential of mean force, $W(N)$, for forming alpha and PPII segments of various length, N, as a function of segment length.

length 3–9 is lower or equal than for alpha helices of the same length, at all temperatures shown. At 278K the PPII pmf has minima at lengths 8 and 11. This figure shows that long-range ordered PPII segments are formed, although they are not the dominant structure in A21. Alpha helices, on the other h and, are the dominant, long-range, ordered structures at all temperatures shown. The cooperativity imparted by the $i, i+4$ hydrogen bond pattern is responsible for this long-range order [37].

To determine the statistical significance of forming long-range ordered structures, we define an unbiased ensemble that consists of all possible configurations that can be adopted by a chain of length 21, considering only two possible states for each member of the chain. There are 2^{21} possible microstates in this ensemble. This unbiased binary ensemble has the highest Shannon entropy for a chain of length 21, where each member of the chain has probability 0.5 of being in either 1 or 0. We also calculate binary ensembles for PPII (PPII = 1, other = 0), and alpha (alpha = 1, other = 0) structures. These two ensembles are compared to the unbiased ensemble. We define $N_L(X)$ as the number of islands of 1's of length L in X, where X = PPII, alpha, or the reference ensemble. We define $P_L(X) = N_L(X)/N_{configuration}$, where $N_{configuration}$ is the number of configurations in the ensemble (20,000 for the PPII and alpha ensembles, and 2^{21} for the reference ensemble. Fig. 6 shows a semilog plot of the ratio of $P_L(X)/P_L(REF)$, for the PPII and alpha ensembles, as a function of the segment length. This ratio will be large if the number of configurations of X of length L is much larger than in the reference state.

First, we will describe the results for the alpha helical ensemble. For alpha helices at low temperature the number of conformations with helices of length 18 is 10,000 times larger at the lower temperatures than for the unbiased ensemble. However, for helices of length 1–6 fewer conformations are observed than for the unbiased ensemble, with the lowest count for two consecutive segments in an

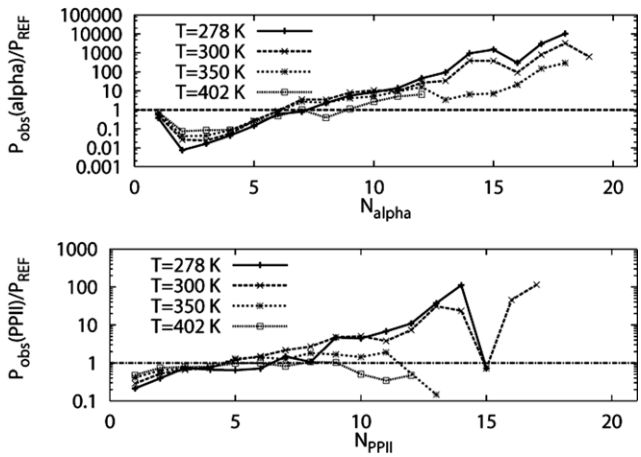

Fig. 6. Ratio of the observed number of counts for alpha and PPII configurations of various lengths over the expected number of counts in the reference, unbiased, state.

alpha helical configuration, reflecting the nucleation energy cost for helix formation. For segments of length larger than three, there is an increase in the number of observed configurations with the helix length, up to length 17. As T increases the number of longer helical configurations decreases. The curves corresponding to the alpha helices can be understood in terms of the nucleation and propagation parameters in the Zimm–Bragg theory. The lower than expected propensity of helical segments of length 2 is due to the helix nucleation energy. Helical segments of length 2 do not form hydrogen bonds but pay an entropic price for restricting their conformations to the alpha region. At length 3, one hydrogen bond is formed, the energy is lowered, but over all this is still a high free energy state. For every hydrogen bond formed there is a gain in free energy and propensity for forming longer segments is enhanced. This construction does not tell us anything new, but will help understand the energetics of PPII formation in the next plot.

For PPII conformations, we observe that segments of length 14 are 100 times more likely to occur at the lower temperatures than in the unbiased ensemble. Notice, however, that alpha helical segment of the same length are 1000 times more likely to occur than in the reference ensemble, reflecting a higher cooperativity for the formation of alpha helices. At 300 K we observed PPII as long as 17 amino acids, with a count 100 times larger at lower temperatures than the unbiased ensemble. At the lowest two temperatures, we see an exponential increase in the propensity of forming PPII segments if increasing length (up to length 14). In analogy with the alpha helix, this implies that there is a PPII propagation free energy that enhances PPII formation. The propensity to form long PPII structures disappears at higher temperatures. PPII segments of length shorter than five are less likely to occur than in the reference ensemble. This suggests that the energy driving the formation of PPII requires PPII segments of length 5 or more to act favorably. Interestingly, the ratio for PPII segments increase with temperature for lengths 5–6, and decrease for longer and shorter segments. A comparison of the profiles for alpha and PPII structures show that short lengths, up to length 8, the chain will prefer PPII over alpha helices. If we assume that the segment length preferences apply to any peptide of length 21 or less, sequences of length 8 or less will show a predominant PPII conformation. This is consistent with the observation of PPII structures by Shi et al. [21]

A property that can be measured for peptides and that plays a crucial role in the modeling of polymers is the end-to-end distance distribution. For a random coil, we expect the end-to-end distance distribution to be a Gaussian distribution, with mean proportional to the square root of the number of amino acids. For a 21 amino acid alpha helix we expect an end-to-end distance around 30 A, and for helix-turn helix it could be as short as an atomic contact (3.2 A). For a PPII segment of length 16, we should observe an end-to-end distance of 40 A. For the ensembles studied here, we expect a mixture of all of the above. We use the methyl groups in the acetyl and methyl amide termini of the chain to calculate the end-to-end distance. Fig. 7 shows the end-to-end distance distributions over a broad temperature range. At high temperatures, we see a distribution that could be well approximated to a Gaussian distribution with average distance near 15 A. At the lowest temperature, we seen the most complex distribution, with multiple peaks at ~13, 35, 50–60 A, corresponding to distributions dominated by disordered and bent helical conformations, helices, and PPII extended conformations. Configurations with end-to-end distances larger than the simulation box length (43.5 A) may be affected by the periodic boundary conditions. We can neglect extended beta conformations in our analysis, since we did not find segments of more than four consecutive beta angle pairs. As the temperature increases, the peak corresponding to alpha helices gets smaller, and the peaks corresponding to PPII extended conformations disappear. Also, the peak near 15–20 A broadens.

Given the simplicity of the amino acid being studied (Ala), and the lack of significant intermolecular packing interactions, it has been speculated that water stabilizes PPII conformations [10,11,21,43]. We characterize the density of water around PPII segments of length 5 obtained during the last 2 ns of the simulation. All configurations identified to have PPII segments of length 5 or more are aligned to an ideal PPII segment. The density was mapped on 41 A cubic grid, using a 1 A mesh. The grid is centered at the five amino acid fragment center of mass. Fig. 8 shows the position of large water density around these PPII conformations, at three different temperatures. The high density grid points are labeled with spheres in yellow for water densities in the range $3\rho_0 < \rho < 2\rho_0$, and orange for higher water densities, $\rho \geq 3\rho_0$. ρ_0 is the bulk density of water. At low temperature the sites of high water density form a continuous chain of sites, 1 A apart, surrounding the PPII backbone, and bounded on the sides by side chain CB atoms three amino acid apart. This high water density chain suggests that water can occupy multiple positions near the carbonyl atoms. That is, water molecules are delocalized in this channel, and are not pinned around polar atoms. Interestingly, this groove requires four consecutive amino acids to form—an interesting number in light of the previous observation regarding the formation of PPII segments of length 5 or more. Water molecules coordinated to the backbone amino nitrogen atom have much higher density ($3–5\rho_0$), and are more tightly

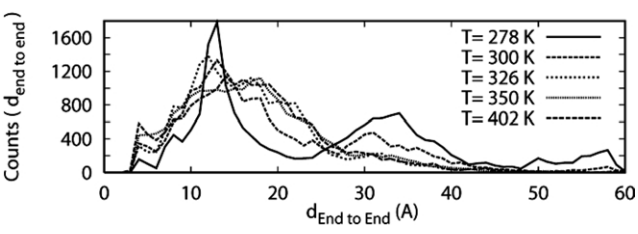

Fig. 7. End-to-end distance distribution for A21 at various temperatures.

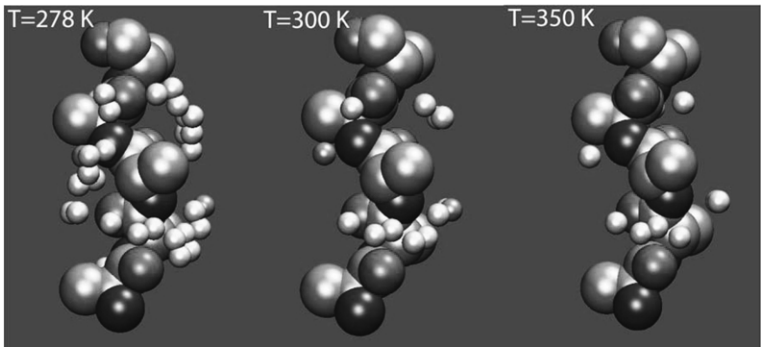

Fig. 8. Sites of high water density around PPII conformations at three different temperatures (278, 300 and 350 K). The density is calculated in a 41 A grid, at 1 A resolution. The grid points with high water density are labeled with spheres in yellow ($3\rho_0 < \rho > 2\rho_0$) and orange ($\rho \geq 3\rho_0$). The yellow and orange spheres indicate positions of high water density, not positions simultaneously occupied by individual water molecules. The PPII forms a left-handed helix with three amino acids per turn. The color codes are C (cyan), O (red), and N (blue). We used the program VMD to create this image.

localized. At higher temperatures the continuous water chain breaks, and water forms clusters around the nitrogen and oxygen backbone atoms.

Fig. 9 shows the triplet correlation function of water molecules around two consecutive carbonyl oxygen atoms (O_{l+2} and O_{l+3}), in a five amino acid segment consisting of amino acids l to l + 4, in conformations that adopt the PPII structure, $g3_{OOW}$, and the water triplet correlation function, $g3_{WWW}$. The triple correlation functions depend on three variables, (r, s, t), describing the sides of a triangle. We only show the triplets with $r = d(O_{l+2}, O_{l+3}) = 3.6$ Å. For PPII structures, the O–O distance are sharply distributed around 3.6 Å. In the $g3_{OOW}$ triplet correlation function we observe three larger peaks at $(s, t) = (3.0$ Å, 5.5 Å), (5.5 Å, 3.0 Å) and (5.5 Å, 5.5 Å). The first two peaks correspond water molecules hydrogen bonded to one O. The third peak corresponds to water molecules hydrogen bonded to the backbone nitrogen atom between the two carbonyls. These water molecules are strongly bound. The water density around the nitrogen atoms is $3–5\rho_0$. These sites are shown with orange spheres in Fig. 8. This distribution is significantly different from the $g3_{WWW}$ obtained for three water molecules, showing that the PPII structure changes the water structure near its surface.

4. Conclusions

The structural characterization of the folded and unfolded ensembles of A21 shows a system where alpha helices and PPII configurations coexist at low temperatures. At low temperature the PPII content increases with temperature, up to 300 K, when is shows a maximum. At higher temperatures the PPII content decreases with temperature. As temperature increases, the propensity to form alpha helices decrease. At all temperatures a disordered state dominates the ensemble. At higher temperatures the end-to-end distance distribution is similar to a Gaussian chain, and the system does not hold long-range ordered structures, except for alpha helices. PPII structures of length 5–8 are more likely to form than in an unbiased ensemble, but only by a factor of two. At low T

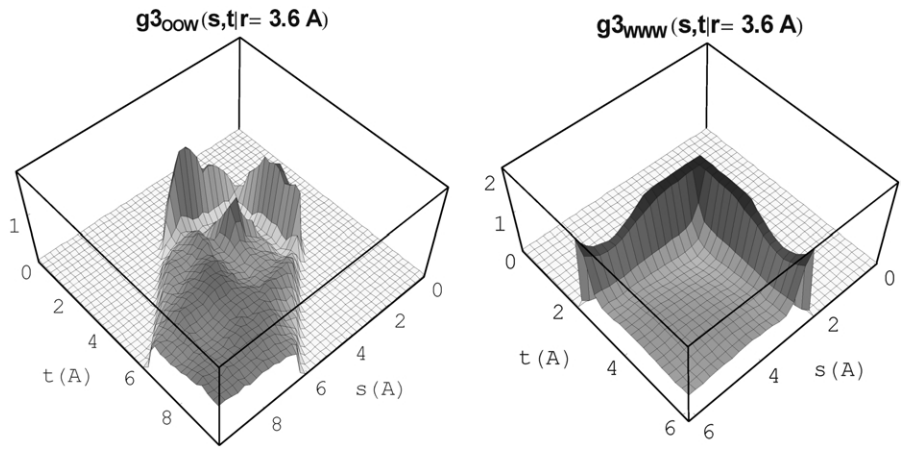

Fig. 9. Triplet correlation function of water molecules around consecutive two carbonyl oxygen (O) atoms in conformations that adopt the PPII structure, $g3_{OOW}$, and the water triplet correlation function, $g3_{WWW}$. For PPII structures the O–O distance is sharply distributed around 3.6 A. We show the $g3(r, s, t)$, where (r, s, t) are the sides of a triangle defined by O(l = 2), O(l + 3) and OW, for $r = 3.6$ A, and all values of s, t between 0 and 10 A.

fragments of length 14 are 100 times more likely to occur than in the unbiased ensemble. In our calculations, we do not see the helix coil transition occurring as an order to order, but rather as an order to disorder, as it has been commonly described, with the caveat that there are two ordered structures at low temperature, one with a large population (alpha helices), and the other with a minor population (PPII). For peptides of length 8 or less the PPII structure will dominate over the alpha conformation. We found that at low temperatures the propensity of forming longer PPII segments is enhanced, indicating the existence of a significant propagation free energy. We also found that PPII segments of length less than four are significantly less likely to form than in an unbiased ensemble. This suggests that segments of length 4 are needed to stabilize PPII structures. Calculations of the hydration around PPII segments of length 5 or more suggested that four amino acids in the PPII conformation are needed to form a channel occupied by water molecules. These water molecules are delocalized along the channel at low temperature. These observations suggest that the formation of the water channel may provide a mechanism for the cooperativity observed in PPII.

Acknowledgements

I thank S. Gnanakaran, N.R. Kallenbach, J.J. Portman, R. Pappu, M.L. Paulaitis, L.R. Pratt and K.S. Sanbonmatsu for insightful comments and collaborations. This work was supported by the US DOE under contract W-740-ENG-36 and the Laboratory Directed Research and Development Program at Los Alamos National Laboratory.

References

[1] Williams S, Causgrove TP, Gilmanshin R, Fang KS, Callender RH, Woodruff WH, Dyer RB. Biochemistry 1996;35:691–7.
[2] Thompson PA, Eaton WA, Hofrichter J. Biochemistry 1997;36: 9200–10.
[3] Munoz V, Thompson PA, Hofrichter J, Eaton WA. Nature 1997;390: 196–9.
[4] Bryngelson J, Wolynes P. Proc Natl Acad Sci USA 1987;84:7524–8.
[5] Leopold PE, Montal M, Onuchic JN. Proc Natl Acad Sci USA 1992; 89:8721–5.
[6] Nymeyer H, Garcia AE, Onuchic JN. Proc Natl Acad Sci USA 1998; 95:5921–8.
[7] Onuchic JN, Luthey-Schulten Z, Wolynes PG. Annu Rev Phys Chem 1997;48:545–600.
[8] Onuchic JN, Wolynes PG, Luthey-Schulten Z, Socci ND. Proc Natl Acad Sci USA 1995;92:3626–30.
[9] Tiffany M, Krimm S. Biopolymers 1968;6:1379–82.
[10] Kelly M, Chellgren B, Rucker A, Troutman J, Fried M, Miller A, Creamer T. Biochemistry 2001;40:14376–83.
[11] Stapley B, Creamer T. Protein Sci 1999;8:587–95.
[12] Adzhubei A, Sternberg M. J Mol Bio 1993;229:472–93.
[13] Rucker A, Creamer T. Protein Sci 2002;11:980–5.
[14] Pappu R, Rose G. Protein Sci 2002;11:2437–55.
[15] Han W, Jalkanen K, Elstner M, Suhai S. J Phys Chem B 1998;102: 2587–602.
[16] Woutersen S, Hamm P. J Phys Chem B 2000;104:11316–20.
[17] Woutersen S, Hamm P. J Chem Phys 2001;114:2727–37.
[18] Eker F, Cao X, Nafie L, Huang Q, Schweitzer-Stenner R. J Phys Chem B 2003;107:358–65.
[19] Eker F, Cao X, Nafie L, Schweitzer-Stenner R. J Am Chem Soc 2002; 124:14330–41.
[20] Woutersen S, Pfister R, Hamm P, Mu Y, Kosov D, Stock G. J Chem Phys 2002;117:6833–40.
[21] Shi Z, Olson C, Rose G, Baldwin R, Kallenbach N. Proc Natl Acad Sci USA 2002;99:9190–5.
[22] Hansmann U, Okamoto Y. Curr Opin Struct Bio 1999;9:177–83.
[23] Sugita Y, Okamoto Y. Chem Phys Lett 1999;314:141–51.
[24] Lockhart D, Kim P. Science 1992;257:947–51.
[25] Garcia AE, Sanbonmatsu KY. Proc Natl Acad Sci USA 2002;99: 2782–7.
[26] Cornell WD, Cieplak P, Bayley CI, Gould R, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell J, Kollman PA. J Am Chem Soc 1995;117:5179–97.
[27] Spek E, Olson C, Shi Z, Kallenbach N. J Am Chem Soc 1999;121: 5571–2.
[28] Jorgensen W, Chandrasekhar J, Madura J, Impey R, Klein M. J Chem Phys 1983;79:926–35.
[29] Pearlman D, Case D, Caldwell J, Ross W, Cheatham T, Debolt S, Ferguson D, Seibel G, Kollman P. Comput Phys Commun 1995;91: 1–41.
[30] Pratt LR, Hummer G, Garcia AE. Biophys Chem 1994;51:147–65.
[31] Hummer G, Soumpasis D, Neumann M. J Phys: Condens Matter 1994;6:A141–4.
[32] Berendsen H, Postma J, van Gunsteren W, Dinola A, Haak J. J Chem Phys 1984;81:3684–90.
[33] Gnanakaran S, Nymeyer H, Portman JJ, Sanbonmatsu KY, Garcia AE. Curr Opin Struct Bio 2003;13:168–75.
[34] Garcia AE, Sanbonmatsu KY. Proteins 2001;42:345–54.
[35] Sanbonmatsu KY, Garcia AE. Proteins 2002;46:225–34.
[36] Lifson S, Roig A. J Chem Phys 1961;34:1963–74.
[37] Zimm B, Bragg J. J Chem Phys 1959;31:526–35.
[38] Hong Q, Schellman J. J Phys Chem 1992;96:3987–94.
[39] Karplus M. J Chem Phys 1959;30:11–15.
[40] Vuister G, Bax A. J Am Chem Soc 1993;115:7772–7.
[41] Hummer G, Garcia AE, Soumpasis DM. Faraday Discuss 1996;103: 175–89.
[42] Garcia AE, Hummer G, Soumpasis DM. Proteins 1997;27:471–80.
[43] Sreerama N, Woody R. Proteins 1999;36:400–6.