SHORT COMMUNICATION

Unfolded State Of Polyalanine Is a Segmented Polyproline II Helix

Alex Kentsis, Mihaly Mezei, Tatyana Gindin, and Roman Osman*  
Department of Physiology and Biophysics, Mount Sinai School of Medicine, New York University, New York, New York

ABSTRACT  Definition of the unfolded state of proteins is essential for understanding their stability and folding on biological timescales. Here, we find that under near physiological conditions the configurational ensemble of the unfolded state of the simplest protein structure, polyalanine α-helix, cannot be described by the commonly used Flory random coil model, in which configurational probabilities are derived from conformational preferences of individual residues. We utilize novel effectively ergodic sampling algorithms in the presence of explicit aqueous solvation, and observe water-mediated formation of polyproline II helical (PII) structure in the natively unfolded state of polyalanine, and its facilitation of α-helix formation in longer peptides. The segmented PII helical coil preorganizes the unfolded state ensemble for folding pathway entry by reducing the conformational space available to the diffusive search. Thus, as much as half of the folding search in polyalanine is accomplished by preorganization of the unfolded state. Proteins 2004;55:493–501. © 2004 Wiley-Liss, Inc.

Key words: protein folding; hierarchical organization; search nucleation; unfolded state preorganization

INTRODUCTION

The importance of accurately defining the unfolded state of proteins was recognized early by Levinthal, who concluded that folding of a random coil by way of a diffusive search of its combinatorially vast conformational space is incompatible with biological energies and timescales of protein folding.1 Consequently, either the conformational space of unfolded polypeptides deviates significantly from that of a random coil, or the conformational search is not entirely diffusive, being guided along folding pathway(s). The latter is a foundation of the framework folding models,2 including the experimentally supported search-nucleation mechanism,3 while the former is an implication of the hierarchical organization of proteins that are composed of nested structural motifs,4 the most basic of which, such as helices and hairpin turns, can form on timescales several orders of magnitudes faster than those of protein folding.5

In support of the nonrandom nature of the unfolded state, original studies by Tanford and colleagues suggested that unfolded polypeptides can possess residual structure under nondenaturing conditions,6 and using CD spectroscopy Tiffany and Krimm7 observed PII conformations in natively unfolded polyglutamyl peptides and hypothesized their origin in PII helices 4–7 residues in length. Recently, advanced spectroscopic studies indicated that a substantial fraction of residues in unfolded polyalanine peptides exists in PII conformation.8–11 Polypeptides in canonical PII configuration are characterized by backbone dihedral angles of \((\phi, \psi) = (-75°, +145°)\), which result in \(3_{10}\)-helices that do not contain intramolecular hydrogen bonds and are more extended and solvated as compared to α-helices. It is unknown whether the macroscopic experimentally observed PII structural features correspond to conformational preferences of individual residues, or to an average of microscopic configurational states that are composed of contiguous stretches of multiple residues in PII conformation. The former is consistent with Flory’s isolated pair hypothesis and the random, albeit conformationally biased (statistical), coil model of the unfolded state.12 The latter cannot be obtained from Flory’s model and implies that the unfolded state configurational space of proteins is preorganized in the form of PII structural segments. A quantitative demonstration of such preorganization may provide insight into its contribution to the folding search, and is thus of major importance for understanding, predicting, and designing protein structures.

Numerous simulations of short peptides have been carried out in the past, including those of polyalanine in...
aqueous environment, and a few have focused on the role of unfolded state structure. Recent work explored a blocked A21 peptide in explicit water using replica exchange molecular dynamics (MD). The authors observed a substantial fraction of α-helix, but the rest of the conformations, labeled as coil, remained undifferentiated. Recently, Mu et al. and Zaman et al. examined conformational preferences of alanine tripeptides and noted that the P2H conformation has a major population in the unfolded state.

In this article, we present results of statistically converged and apparently accurate calculations of the unfolded state ensemble of 7 (A7) and 14 (A14) residue polyalanine peptides flanked by pairs of diaminobutyrate and ornithine residues (see Methods section), recently examined by Kallenbach and colleagues using NMR and circular dichroism (CD) spectroscopies. These peptides are of particular interest because A7 is natively unfolded, allowing direct characterization of the unfolded state, without the use of chemical or thermal denaturation, and A14 is about 30% helical, enabling us to evaluate the contribution of unfolded state structure to α-helical (protein) stability. Moreover, the physics of polyalanine reflects inherent and universal properties of the polypeptide backbone, which are preserved in natural proteins and modified in a sequence specific manner.

METHODS

Molecular Systems

Simulations were performed with all-atom CHARMM27 and AMBER98 (http://amber.scripps.edu) force fields and the TIP3P water model. CHARMM27 and AMBER98 force fields were produced by setting force constants for φ and ψ torsion potentials to zero. CHARMM27a98 and AMBER98c27 were produced by replacing the φ and ψ torsion potentials in CHARMM27 with those of AMBER98, and those of AMBER98 with those of CHARMM27, respectively.

Calculations were performed on the blocked polyalanine peptides Ac-X2A7O2-NH2 (A7) and Ac-X2A14O2-NH2 (A14), where X and O represent diaminobutyrate and ornithine, respectively, as studied experimentally by Kallenbach and colleagues. Force field parameters for both diaminobutyrate and ornithine were derived from lysine, since their amino acid preferences of alanine tripeptides and noted that the PII conformation has a major population in the unfolded state.

Recent work characterizes its performance using larger solutes and extends its capabilities as described.

MC simulations were performed using a force-biased Metropolis procedure as implemented in the program MMC (http://inka.mssm.edu/~mezei/mmc). Systems were thermalized for 10,000 sweeps, as judged from energy equilibration, and evolved for 50 million sweeps, saving configurations every 100th sweep, where one sweep represents one step of all the degrees of freedom, including those of the solute and the solvent. Bond lengths and angles were kept constant. We utilized the shuffled cyclic procedure for single-solute torsional moves, and the reverse proximity criterion for local torsional moves of several torsional angles, which were performed half of the time. Appendix 1 in the Supplementary Material describes the iterative algorithm used to obtain the proximal solution to the loop-closing problem for a peptide backbone. Although not necessary for convergence of simulations of pentapeptides, usage of local torsional moves was required for efficient sampling of the larger solutes studied here (Supplementary Tables I–III). Both solute and solvent step sizes were tuned to yield mean acceptance rates of 20–40%. The radius of the PHS was updated every third sweep, using the normalized reference shell energy of 0.15 kcal/mol/molecule and the restraining force constant of 3.0 kcal/mol/Å2, as described previously. All nonbonded interactions were included.

Convergence and Accuracy

To evaluate sampling efficiency, we calculate the evolution of the apparent self-diffusion coefficient, $\Omega_a \rightarrow b (x) = (f_{a}(x) - f_{b}(x))^2$, as a function of simulation length x, where $f_a$ and $f_b$ are phase space variables, such as protein backbone dihedral angles, of two independent simulations starting from different initial conditions a and b. If the sampling of phase space is ergodic, $\Omega_{a \rightarrow b} (x)/\Omega_{a \rightarrow b} (0)$...
decays to zero at long x. This is a necessary but insufficient condition of ergodicity, since it depends on the choice of initial conditions \(a\) and \(b\). To evaluate convergence and ergodicity directly, we carried out 4 independent simulations using different random number seeds for the initial MC moves and their sizes, and 4 different initial solute configurations, chosen to span the configurations of interest: 

\[ P_{II}(\phi, \psi) = (-75^\circ, +145^\circ), \]

\[ P_{II}(\phi, \psi) = (+145^\circ, -75^\circ), \]

\[ \alpha\text{-helix}(\phi, \psi) = (-57^\circ, -47^\circ), \]

and \( \beta\text{-strand}(\phi, \psi) = (-139^\circ, +135^\circ). \)

**Analysis**

Although canonical secondary structure elements have well-defined regular geometries, conformations and configurations in solution at ambient temperature exhibit considerable plasticity, which in computational studies are additionally dependent on force field parameters. Moreover, torsional geometries and their associated basins may depend on the aggregate nature of polypeptide interactions.

We use the term **configuration** to refer to backbone geometries of individual peptide residues, and **configuration** to refer to molecular geometries of groups of residues. Consequently, we utilize a self-consistent method for defining configurational basins using a stepwise optimal clustering algorithm based on a self-organizing neural net, as implemented in ART-2 by Brooks and coworkers. Briefly, the cluster assignment of the dihedral angles extracted from simulation trajectories are optimized subject to a constraint on the cluster radius, such that no member of a cluster is farther than a specified distance from the cluster center. Because the convergence of such minimizations is sensitive to initial conditions, we test the robustness of assignments to configurational basins by recalculating the cluster assignments using reshuffled simulation trajectories (data not shown). Probabilities of sampling of the configurational basins based on configurational assignments as defined in this manner were calculated using a set of home-built programs, available upon request. Additionally, robustness of self-consistent clustering was evaluated by assigning secondary structure elements using the definitions implemented in PROSS (http://roselab.jhu.edu/utils/pross.html). In this manner, we use \( P_{II} \) to refer to backbone geometries in the \( P_{II} \) conformational basin, \( \alpha \) to refer to helical geometries, including those of \( \alpha\)-helical residues, and \( S \) to refer to residues having strand geometries, including residues in \( \beta\)-strands [Fig. 1(a)].

Scalar \( J_{NNN} \), coupling constants were calculated using the Karplus relation, as calibrated by Vuister and Bax. Interaction energies, radial and orientational distribution functions, and coordination numbers were calculated according to standard methods, referenced to the center of mass of the solute, as implemented in MMC using proximity analysis. Sensitivity analysis of the decomposed interaction energies was done according to the method of Zhang et al.

**RESULTS AND DISCUSSION**

In order to ensure computational accuracy, we have employed two different solute potential energy functions, an effectively ergodic MC algorithm to sample configurational space, and an explicit MC-PHS water to represent molecular aqueous solvation. MC-PHS maintains principal solvation effects such as density, solvation energy, and fine water structure, while efficiently coupling structural rearrangements of the solvent to those of the solute. As a result, simulation lengths of \( 5 \times 10^7 \) sweeps exceed the apparent time for self-diffusion of the backbone dihedral angles by more than 3 orders of magnitude (Supplementary Figs. 1 and 2, Supplementary Tables 1–III), a degree of sampling currently unattainable by MD and MC simulations of canonical ensembles in condensed phase. The simulations appear to be converged, as judged from the results obtained from 4 independent simulations using initial configurations maximally distributed in torsional phase space (Fig. 1).

Unfolded state ensembles calculated using CHARMM27 and AMBER98 force fields exhibit several differences (Supplementary Fig. 3), most notably in the geometry of calculated \( P_{II} \) conformations [Fig. 1(a and b)], but are statistically indistinguishable in the mean probabilities of individual alanine residues sampling the respective \( P_{II} \) conformational basins, in spite of significant differences in their parameterisations of the polypeptide backbone [Fig. 1(c and d)]. More importantly, mean \( P_{II} \) conformational probabilities of \( 0.42 \pm 0.089 \) and \( 0.47 \pm 0.076 \) calculated using CHARMM27 and AMBER98, respectively, are in excellent agreement with the experimentally observed value of \( 0.4 \pm 0.08 \). Additionally, calculated \( J_{NNN} \) scalar coupling constants (Supplementary Table IV) are in agreement with those observed experimentally.

To characterize the **configurational** preferences of the unfolded state ensemble, we enumerate the occurrences of varying lengths of contiguous stretches of alanines in \( P_{II} \) conformation, and compare the calculated probability distributions with those expected for a Flory polymer with identical but independent configurational preferences [Fig. 2(a and b)]. Probabilities of sampling \( P_{II} \) configurations are significantly greater for both CHARMM27 and AMBER98 simulations as compared to the rapidly decaying power law distributions of their respective Flory polymers [Fig. 2(a and b)], indicating that the unfolded state ensemble of polyalanine deviates significantly from that of a statistical coil. Specifically, \( P_{II} \) segments 4–6 residues in length, corresponding to about 1–2 helical turns, appear to be roughly isoenergetic with stretches of residues in the strand conformation, as indicated by their sampling probability of 0.4–0.5 [Fig. 2(a)]. Biphase nature of the apparent configurational \( P_{II} \) probabilities relative to those of a Flory-like statistical coil with a breakpoint at 4–6 residues is consistent with such a segmental structure [Fig. 2(b)]. Snapshots of the calculated ensemble are shown in Figure 2(c). Consequently, the unfolded state ensemble of polyalanine is a segmented \( P_{II} \) helical coil, whose \( P_{II} \) segments are roughly isoenergetic with intervening strand configurations [Fig. 2(a)].
Fig. 1.

Fig. 2.
The corresponding difference in conformational entropy between the observed $P_{II}$ segmentally helical and expected Flory statistical coils is about 1 cal/mol/K per residue and 3–4 cal/mol/K per $P_{II}$ helical turn (Table I). The apparent linear additivity of conformational entropies of $P_{II}$ helix formation in the unfolded state of $\alpha$-polyalanine indicates that $P_{II}$ formation is largely noncooperative (Table I), consistent with predominantly diffusive nature of unfolded state configurational sampling. Considering that the calorimetrically determined loss of conformational entropy for $\alpha$-helix formation in polyalanine is about 2 cal/mol/K/residue, $^{11,34}$ formation of $P_{II}$ helical structures in unfolded polypeptides significantly reduces the contribution of conformational entropy to protein stability, conventionally thought to be the major opposing force for folding, as dictated by random and statistical coil models of the unfolded state.

To examine the contribution of solvating water to the configurational preferences of polyalanine, the accompanying article in this issue by Mezei et al. $^{35}$ presents results of thermodynamic integration calculations of conformational free energy of polyalanine. Remarkably, they observe that regular $P_{II}$ helices produce a less disruptive effect on surrounding water organization as compared to $\beta$-strands, the solvation of the latter requiring the formation of entropically unfavorable peptide:water:peptide bridges and general reorganization of the surrounding bulk, much akin to a hydrophobic moiety. $^{36}$ We take a complementary approach and decompose the interaction energies of the conformationally segregated $\alpha$-unfolded state ensembles using proximity analysis (Supplementary Fig. 4). In agreement with Mezei et al., our analysis suggests that the stability of $P_{II}$ conformations relative to strand is entropic in origin, being a property of the surrounding water solvent (Supplementary Fig. 4). Such entropic, water-mediated origin of stability of $P_{II}$ helical structure in the unfolded state of polyalanine, coupled with the noncooperative nature of $P_{II}$ itself constitute a fundamental revision in our understanding of protein folding. In this manner, the hydrophobic effect appears to be not only a principal determinant of protein stability and folding pathways, $^{36}$ but also to act early in the folding process to preorganize the unfolded state structure.

To directly examine the contribution of unfolded state structure to $\alpha$-helical (protein) stability, we calculated the configurational partition functions of $A_{14}$ polyalanine, which due to its increased length is able to sample $\alpha$-helical configurations at sufficient probabilities for analysis. These calculations also appear to be converged as judged from the equivalency of the results of 4 independent simulations using CHARMM27, and accurate, as estimated from the mean $\alpha$-helical content of 0.25 ± 0.035, as compared with the experimental value of about 30%. $^{32}$ Furthermore, calculated length-dependent configurational probabilities of $\alpha$-helix formation are in general agreement with experimental values, $^{37,38}$ whereby formation of short $\alpha$-helices (nucleation) is unfavorable, with an apparent mean $\alpha = 0.00084 \pm 0.00027$, and formation of long $\alpha$-helices is favorable (propagation), with an apparent mean $s = 1.1 \pm 0.053$ [Fig. 3(a)]. Configurational $P_{II}$

### Table I. Differences in Conformational Entropies for the $A_2$ Peptide Between Observed Segmentally Helical $P_{II}$ coil and Statistical Coil With Identical Conformational Preferences

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\Delta S_{seg}^a$</th>
<th>$\Delta S_{res}^b$</th>
<th>$\Delta S_{seg}^a$</th>
<th>$\Delta S_{res}^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.59</td>
<td>0.62</td>
<td>0.63</td>
<td>0.65</td>
</tr>
<tr>
<td>2</td>
<td>0.75</td>
<td>0.77</td>
<td>0.78</td>
<td>0.80</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>1.05</td>
<td>1.05</td>
<td>1.07</td>
</tr>
<tr>
<td>4</td>
<td>1.6</td>
<td>1.75</td>
<td>1.72</td>
<td>1.84</td>
</tr>
<tr>
<td>5</td>
<td>1.8</td>
<td>1.95</td>
<td>1.82</td>
<td>1.94</td>
</tr>
</tbody>
</table>

$^a$Calculated using $S = R \ln P$, where $R$ is the gas constant, and expressed in cal/mol/K $\cdot 10^{-1}$.

$^b$Calculated by dividing $\Delta S$ by the number of contained residues. Note that for segments greater than 3 residues in length, the calculated $\Delta S$ values report on the conformational entropy of forming a $P_{II}$ helical turn. Since the differences in conformational entropies per residue calculated using different segment lengths are equal to each other, this linear additivity indicates that $P_{II}$ helix formation is noncooperative, consistent with the largely diffusive nature of the conformational search in the unfolded state.
stability appears to have a maximum at length of 4–5 residues [Fig. 3(a)], consistent with the segmentally P II helical nature of the A7 unfolded state [Fig. 2(a)]. As for the A7 peptide, A14 also exhibits significant differences from Flory behavior for lengths up to 7–8 residues (Table II). Most importantly, length-dependent stabilities of P II and
α-helical configurations are inversely related, whereby melting of PII (in the unfolded state) is concomitant with formation of α-helices [Fig. 3(a)], suggesting that PII formation in the unfolded state may facilitate length-dependent α-helix formation.

In order to directly evaluate the contribution of unfolded state PII structure to the kinetics of α-helix formation, we calculated the potential of mean force (PMF) for the formation of an α-helical turn in A14 by enumerating the backbone torsions of alanine residues when they are in the field of 3 α-helical residues. Such a PMF represents the Gibbs free energy surface forming a 4-residue α-helical turn in the mean field of 3 α-helical residues, as referenced to the most stable conformation of the fourth residue [Fig. 3(b)]. The conformational basins in the PMF surface identified from the analysis correspond to distorted PII, β, and α, with population-weighted average (ϕ, ψ) = (−81°, +137°), (−157°, +173°), and (−98°, −69°), respectively [Fig. 3(b)]. Most importantly, the barrier for α-helix formation from the PII conformation is about 0.8 kcal/mol lower than that from β [Fig. 3(b)], suggesting that PII structure in the unfolded state kinetically facilitates α-helix formation.

The energetic preference for such a pathway of α-helix nucleation from PII may be due to the requirement of moving only the ψ angle, ϕ already having an approximate α-helical geometry in PII. On the other hand, formation of α from β requires the motion of both ϕ and ψ angles [Fig. 3(b)]. Moreover, there is an absence of peptide:water:peptide bridges in PII,35 structures that may need to be disrupted in order to complete an α-helical turn from β conformation. In a sense that the transition state for the conversion of PII to α resembles the configurational basin from which it originates, PII structures in the unfolded state of polyalanine act to preorganize as well as guide the ensemble for folding pathway entry. In this way, the segmentally helical PII coil acts to preorganize the unfolded state ensemble of polyalanine and to kinetically guide its α-helix formation.

Cooperativity that defines α-helix formation is characteristic of protein folding in general, with both phenomena exhibiting a nonlinear, sigmoidal dependence on temperature. Both the largely all-or-none behavior and the associated two-state transition mean that the microscopic ensemble is either largely folded or unfolded. However, statistical mechanical models of the helix–coil transition that assume full cooperativity among residues39 overestimate helical content, as well as its dependence on temperature. Consequently, Zimm and Bragg,40 and later, Lifson and Roig41 reduced the overall cooperativity of the transition by employing an Ising-like model, limiting it to nearest neighbor cooperativity. Thus, helix formation is partitioned into fully cooperative nucleation and largely noncooperative extensions steps, thereby reducing the overall cooperativity of the reaction, and more accurately reproducing the observed helical content. The Ising model and its variants are the only way to reduce two-state reaction cooperativity in the presence of a random reference state.42 However, nearest neighbor cooperativity is not required if the reference state is organized, as is the case for the coil state of polyalanine (Fig. 2).

Because the loss of conformational entropy in the formation of a PII helix is about half of α-helix formation, the entropic penalty associated with nucleating the first α-helical turn in the Zimm–Bragg model is significantly reduced in the segmentally helical PII coil (Table I). The Ising-like reduction of overall cooperativity in the helix–(random) coil model of Zimm and Bragg is thus effectively equivalent to a model based on the (preorganized) segmentally helical PII coil state, in which α-helix formation itself is fully cooperative, but the overall cooperativity is reduced by the noncooperative PII formation. Such a model is supported by the calculated configurational thermodynamics and kinetics of A14, in which depletion of PII is concomitant with appearance of α-helical structure (Fig. 3). Modification of the Zimm–Bragg formalism to include a third state leads to a reduction of overall cooperativity, as was demonstrated by inclusion of the 310-helix.43 A more pronounced effect is expected from the introduction of a PII-helix, since its formation is less cooperative as compared to that of a 310-helix (Table I). Finally, the Zimm–Bragg random coil model is incompatible with the observed kinetics of α-helix formation, whereby in addition to the exponential phase corresponding to cooperative barrier crossing, several stretched and nonexponential relaxations are observed,44,45 indicative of noncooperative processes; these could be attributed to processes such as PII formation. A mechanism of α-helix formation that emerges from these considerations involves diffusive, noncooperative PII formation in the segmentally helical PII coil state that facilitates nucleation and propagation of α-helices.

Although recently Flory’s independent pair hypothesis was observed to be sterically invalid in the α-helical region of the Ramachandran map, no steric interference was observed for extended conformations such as β and PII.46 Here, we demonstrate that Flory’s hypothesis and the associated assumption of randomness of the unfolded state fail for these conformations as well, albeit as a result of an

### Table II: Differences in Conformational Entropies for the A14 Peptide Between Observed Segmentally Helical PII Coil and Statistical Coil With Identical Conformational Preferences

<table>
<thead>
<tr>
<th>N</th>
<th>ΔS/seg*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>1.2 ± 0.21</td>
</tr>
<tr>
<td>3</td>
<td>2.3 ± 0.35</td>
</tr>
<tr>
<td>4</td>
<td>3.7 ± 0.52</td>
</tr>
<tr>
<td>5</td>
<td>4.6 ± 0.96</td>
</tr>
<tr>
<td>6</td>
<td>5.1 ± 1.3</td>
</tr>
<tr>
<td>7</td>
<td>4.97 ± 2.0</td>
</tr>
<tr>
<td>8</td>
<td>0.13 ± 1.3</td>
</tr>
<tr>
<td>9</td>
<td>0.044 ± 0.52</td>
</tr>
<tr>
<td>10</td>
<td>0.206 ± 0.42</td>
</tr>
<tr>
<td>11</td>
<td>0.015 ± 0.58</td>
</tr>
<tr>
<td>12</td>
<td>0.010 ± 0.72</td>
</tr>
<tr>
<td>13</td>
<td>0.0070 ± 0.41</td>
</tr>
<tr>
<td>14</td>
<td>0.0039 ± 0.24</td>
</tr>
</tbody>
</table>

*Calculated using S = R P ln P, where R is the gas constant, and expressed in cal/mol K ± 1σ.
entropic effect of the surrounding water, whereby polyalanine segments up to 7 residues in length form PII helices [Figs. 2(a) and 3(a)]. Thus, the conformational space actually sampled by unfolded polypeptides is considerably smaller than the one assumed by random and statistical coil models of the unfolded state. The extent of cooperation (κ) in achieving observed biological efficiencies of protein folding between such (thermodynamic) preorganization of the unfolded state and (kinetic) facilitation by folding pathways, regardless of their nature, can be quantified by 𝜅 = Δ𝑆_Δ𝑆_F, where Δ𝑆_F is the loss of conformational entropy in the unfolded state, and Δ𝑆_F is the change in conformational entropy upon folding. In the case of α-helix formation by polyalanine, calorimetrically determined Δ𝑆_F ∼ 2 cal/mol/K/residue,11 and Δ𝑆_F, calculated as the difference in Shannon entropies of the observed segmentally helical PII coil and the expected Flory statistical coil with identical conformational biases, is about 1 cal/mol/K/residue (Table I), yielding 𝜅 ~ 0.5. Thus, as much as half of the folding search of α-helix formation in polyalanine is accomplished by water-mediated, entropic preorganization of the unfolded state into segmentally helical PII coil. Values of 𝜅 and the underlying unfolded state structure propensities are expected to depend both on polymer nature and sequence (unpublished observations), and promise to be useful both for polymer structure prediction and design.

Although the cooperative nature of protein stability and folding requires the existence of folding pathways, diffusive search in the unfolded state prior to the crossing of the rate-limiting barrier can nevertheless be productive, particularly if the configurational space available to the search is reduced relative to conventional random coil models.4 Such a process is in fact a corollary of the statistical mechanics of polymer folding of Wolynes and colleagues47 and the associated funnel-like topologies of several experimentally and theoretically determined folding free energy landscapes.48,49 The segmentally helical PII coil observed in the unfolded state of polyalanine in water, reflecting the inherent physics of the polypeptide backbone, accomplishes just that, capturing more than half of the unfavorable conformational entropy of α-helix formation. Such reduction of the otherwise biologically intractable conformational space acts to preorganize the unfolded state ensemble for folding pathway(s) entry. Indeed, although local structural motifs such as helices and turns can form on ultrafast nanosecond timescales in isolation, several orders of magnitude faster than those of protein folding,5 their formation in proteins is often,50 but not always,51 coupled to the (cooperative) traversal of the rate-limiting barrier, suggesting that preorganization of the unfolded state both locally in the vicinity of PII helices, and globally as a result of chain compaction, is directly related to the nucleation of folding pathways. In this way, the requirements for efficient folding are satisfied both by the thermodynamic organization of the conformational search in the unfolded state and its kinetic guidance by the resulting folding pathways.

ACKNOWLEDGMENTS

We are grateful to Neville Kallenbach, George Rose, William Eaton, Tobin Sosnick, Angel Garcia, and Alex Proekt for helpful and critical discussions, to Allan Capili and Kathy Borden for reading the manuscript, and to Benjamin Goldsteen for technical support. We thank George Rose, Tobin Sosnick, and Angel Garcia for sharing unpublished manuscripts.

REFERENCES


