Thermodynamic Mechanism and Consequences of the Polyproline II (P_{II}) Structural Bias in the Denatured States of Proteins†

James B. Hamburger,‡ Josephine C. Ferreon,‡ Steven T. Whitten, and Vincent J. Hilser*

Department of Human Biological Chemistry and Genetics, and Sealy Center for Structural Biology, University of Texas Medical Branch at Galveston, Galveston, Texas 77555

Received April 2, 2004; Revised Manuscript Received May 14, 2004

ABSTRACT: A quantitative characterization of the structure and energy of the denatured states of proteins represents the cornerstone to a molecular-level understanding of both protein stability and fold specificity. Recent studies have revealed a significant bias in unstructured peptides toward the polyproline II (P_{II}) conformation, even when no prolines are present in the sequence. This indicates that the P_{II} conformation is a dominant component of the denatured states of proteins, although a quantitative description of the component enthalpy and entropy functions associated with this conformation (i.e., the thermodynamic mechanism) has thus far proven elusive. An experimental system has been designed that, when analyzed with high-precision isothermal titration calorimetry, provides direct access to the residue-specific thermodynamics of the P_{II} structure formation in disordered proteins and peptides. Here, it is shown that the P_{II} bias is driven by a favorable and significant enthalpy (Δh) of -1.7 kcal mol^{-1} residue^{-1}, which is partially offset by an unfavorable entropy ($T\Delta s$) of -0.7 kcal mol^{-1} residue^{-1}, relative to the ensemble of disordered conformations of the molecule. In addition to impacting dramatically the interpretation of thermal denaturation experiments, these experimental values form the framework of a quantitative energetic description of the denatured states of proteins.

Protein folding describes the process whereby a nascent polypeptide chain achieves a unique and functional tertiary structure through a search of conformational space. Several decades ago, the observation was made that, for a polypeptide chain to achieve its unique structure in a biologically relevant time frame, a random search of all accessible conformations is not possible and that the polypeptide chain must therefore search a significantly smaller space than would be predicted from steric considerations alone (*1*). Indeed, numerous efforts have focused on the effects of residual native structure, excluded volume, and non-native conformational propensities on the stability as well as the folding kinetics of proteins (*2*−*10*). Notwithstanding these efforts, a quantitative thermodynamic description of the myriad of conformations in the denatured state has proven problematic, primarily because of the experimental difficulty in deciphering the relationship between the structural and thermodynamic properties of the individual conformations, which can vary significantly, and the average properties of the ensemble as a whole.

The polyproline II (P_{II})[1] structure has emerged as an important component in the conformational manifold of disordered peptides (*11*−*22*). Indeed, recent work has suggested that the P_{II} structure may in fact dominate the ensemble of conformations for polypeptides (*23*) even at nonproline positions (*24*−*26*). The P_{II} conformation is characterized by a left-handed helical turn with the amide hydrogen and the carboxyl oxygen of each peptide backbone projecting into solution, presumably making favorable interactions with the solvent (*27*−*30*). In addition, the P_{II} conformation appears to facilitate favorable intrachain $n \rightarrow \pi^*$ interactions, which should also be stabilizing (*31*). Despite clear experimental evidence for the existence of this structural bias and a number of plausible explanations for its origin, little is known of the detailed energetics of the P_{II} conformation. For instance, is the P_{II} state favored enthalpically, entropically, or both? A detailed energetic description of the P_{II} bias, relative to the ensemble of disordered states, is therefore of significant value because it provides direct access to the energetic variations within the denatured state. The significance of such a description is that it provides a residue-specific description of how the denatured state will change throughout the unfolding transition, providing a quantitative estimate of the error associated with the assumption of a single thermodynamic denatured state.

MATERIALS AND METHODS

Isothermal Titration Calorimetry (ITC). The binding energetics used here were those reported previously (*24*) and are listed in parts A and B of Table 1, The previously reported mean and standard deviation in $\Delta H_{binding}$ (Table 2 in ref *24*) correspond to the results obtained by averaging the enthalpy columns in Table 1B.

† Supported by grants from the NSF (MCB-9875689) and the Welch Foundation (H-1461).

* To whom correspondence should be addressed. Fax: 409-747-6816. Telephone: 409-747-6813. E-mail: vince@hbcg.utmb.edu.

‡ Sealy Center for Structural Biology Predoctoral Fellow.

[1] Abbreviations: SH3, src-homology domain 3; C-SH3, C-terminal SH3 domain; Sos, Son of Sevenless; ITC, isothermal titration calorimetry; PDB, Protein Data Bank; H-bond, hydrogen bond; P_{II}, polyproline II.

10.1021/bi049352z CCC: $27.50
Published on Web 07/07/2004

Table 1

(A) Binding Constants for the Interaction of SEM5 Analogues with the P3 and P3A−Sos Peptides as well as the Calculated Microscopic Free Energy Difference between P_{II} and Non-P_{II} Conformations[a]

protein	P3−Sos peptide [K_{obs} (M^{-1})]	P3A−Sos peptide [K_{obs} (M^{-1})]	$\Delta\Delta G$ binding [$-RT \ln(K^{P3}/K^{P3A})$ (cal/mol)]	microscopic P_{II} equilibrium (κ_{mic})	microscopic Δg_{mic} ($g_{nonPII} - g_{PII}$) (cal $mole^{-1}$ $residue^{-1}$)
wild type	$2.81 \times 10^4 \pm 493$[b]	$1.12 \times 10^4 \pm 335$	-545 ± 16	0.191 ± 0.011	981 ± 33
S170G	$2.18 \times 10^4 \pm 366$	$0.86 \times 10^4 \pm 108$	-551 ± 17	0.194 ± 0.005	972 ± 17
S170A	$2.58 \times 10^4 \pm 329$	$1.01 \times 10^4 \pm 150$	-556 ± 16	0.197 ± 0.006	962 ± 17
					average 972 ± 24

(B) Enthalpy of Binding of SEM5 Analogues to the P3 and P3A−Sos Peptides as well as the Calculated Microscopic Enthalpy Difference between P_{II} and Non-P_{II} Conformations[c]

protein	P3−Sos peptide [ΔH_{bind} (cal/mole)]	number of binding sites (N)	P3A−Sos peptide [ΔH_{bind} (cal/mole)]	number of binding sites (N)	microscopic Δh_{mic} ($h_{nonPII} - h_{PII}$) (cal $mole^{-1}$ $residue^{-1}$)
wild type	-7917 ± 39[d]	0.902	-8579 ± 86	1.168	1103
S170G	-7614 ± 35	0.970	-8900 ± 50	0.961	2143
S170A	-7771 ± 23	1.050	-8866 ± 47	1.094	1825
					average 1690 ± 530[e]

[a] κ_{mic} is determined from eq 5 with P_{II} as the reference state, and the microscopic free energy is calculated as $\Delta g_{mic} = -RT \ln \kappa_{mic}$. Positive values for Δg_{mic} indicate that the P_{II} conformation is favored. A description of the materials and methods can be found elsewhere (*24*). Experiments were conducted at 298 K. [b] Standard deviation of the fitted parameters to a two-state model. [c] Values are reported using P_{II} as the reference state; thus, positive values indicate that P_{II} is favored. [d] Standard deviation of the fitted parameters to a two-state model. [e] Standard deviation of the mean for the binding to the three SEM5 analogues.

RESULTS AND DISCUSSION

Binding Equilibrium Linked to Folding. The P_{II} bias at each position in a peptide is locally driven (*32*). As shown below, the result is in an expression wherein the residue-specific P_{II} propensity of the peptide backbone can be determined directly from ITC experiments, which monitor the binding of the C-terminal Sem-5 SH3 domain (SEM5) to analogues of the polyproline peptide sequence ($P_1P_2P_3$-$V_4P_5P_6R_7R_8R_9$) found in the Son of Sevenless (Sos) protein (Figure 1). For a protein binding to a disordered peptide (Figure 1), the overall binding reaction can be divided into two separate equilibrium processes. First, the unbound peptide exists as an ensemble of interconverting conformational states, which must shift to the binding-competent conformation to bind. In the case of the Sos peptide, the binding-competent state is the P_{II} conformation. Second, an equilibrium exists between the unbound state of the binding-competent conformation and the state that is bound to the SEM5 protein. Thus, the overall binding energy ($\Delta G_{bind} = -RT \ln K_{bind}$) is the difference between the free energy required to fold the peptide into the P_{II} conformation (ΔG_{conf}) and the free energy of interaction between the P_{II} conformation of the peptide and the SEM5 protein (ΔG_{int}). If the peptide ensemble is completely biased toward P_{II} in the unbound state, ΔG_{conf} ($=-RT \ln Q_{conf}$) will be $\sim$0 and the measured binding energy will simply correspond to the intrinsic free energy of interaction (i.e., $\Delta G_{bind} = \Delta G_{int}$). If, on the other hand, the unbound peptide molecules are populating a structurally heterogeneous ensemble, the conformational free energy will correspond to the energy difference between the heterogeneous state and the P_{II} conformation. In this study, mutations are made that impact the conformational free energy in a known way, as described below. By monitoring the impact of these mutations on the observed binding, it is possible to elucidate the degree and thermodynamic mechanism of P_{II} conformational bias in the unbound ensemble at the level of the individual amino acid.

The strategy used here is to target a surface-exposed proline residue at position 3 (P3) on the Sos peptide for mutation to alanine (P3A) (Figure 1). Because the side chain of residue 3 does not make contact with the SEM5 protein, as determined from solvent-accessible surface area calculations on the crystal structure of the complex (*33*) (data not shown), the binding interfaces of the P3−SEM5 and P3A−SEM5 complexes are expected to be identical. For such a case, the difference in binding free energy to each Sos variant should arise solely from the difference in the conformational free energies of folding each Sos variant into the P_{II} conformation.

$$\Delta\Delta G_{bind}(P3 \rightarrow A) = -RT \ln\left(\frac{^{P3}K_{bind}}{^{P3A}K_{bind}}\right) \approx -RT \ln\left(\frac{^{P3A}Q_{conf}}{^{P3}Q_{conf}}\right) \quad (1)$$

Indeed, as demonstrated previously (*24*), the NMR {^{1}H-^{15}N} HSQC spectra of SEM5 in complex with either the P3−Sos or P3A−Sos peptides are virtually super-imposable, indicating that the bound states are essentially identical and suggesting that eq 1 is a reasonable approximation.

Microscopic Equilibrium Constant for P_{II} Formation. The conformational entropy, S_{conf}, of disordered peptides and proteins describes the number of conformations accessible to the system (i.e., the degeneracy, Ω) and, assuming equal a priori probability for each conformation, can be related by the following expression:

$$S_{conf} = R \ln \Omega \quad (2)$$

where R is the molar gas constant. It is well-known (*34*, *35*) that the free energy of unfolding a protein can be affected through point mutation of surface-exposed residues, even if the side chain for the substituted amino acid projects into solution and makes no interactions in the native state. The

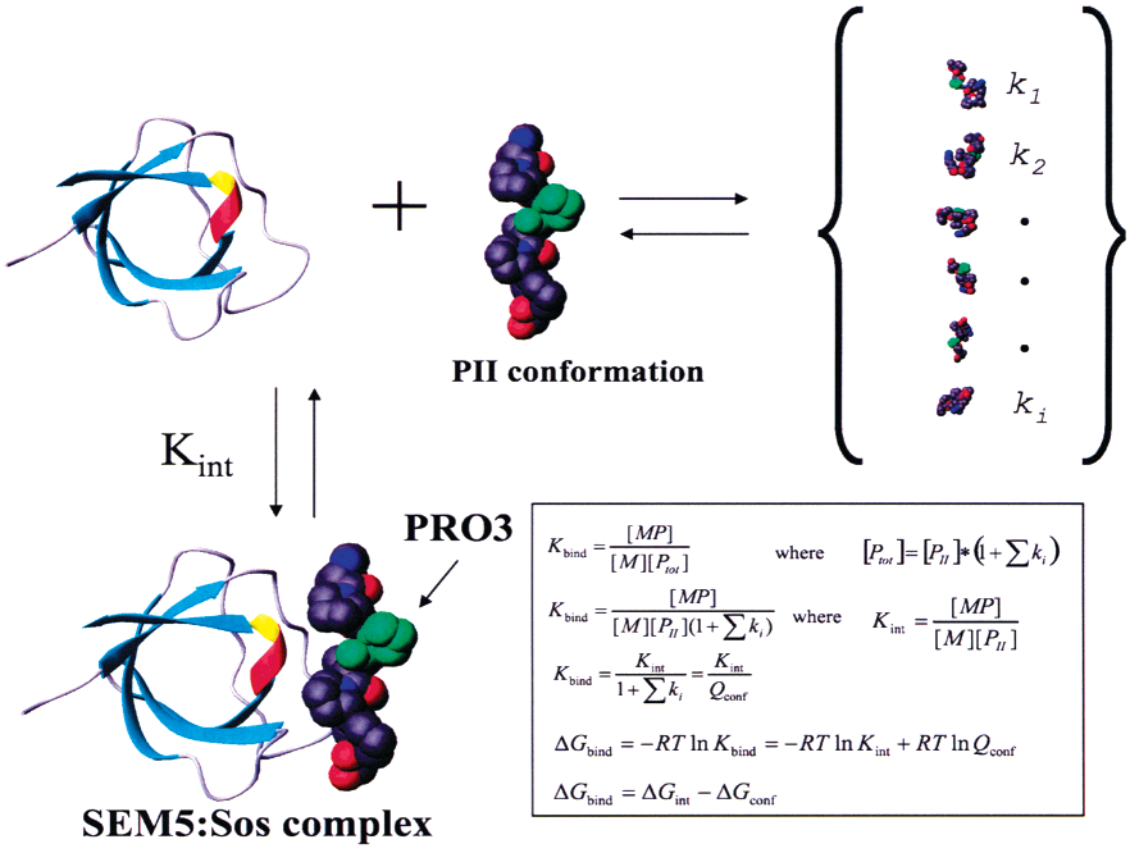

FIGURE 1: Experimental model system. Sem-5 C-SH3 domain (ribbon) binds to the polyproline Sos peptide in the P_{II} conformation (Ac-VPPPVPPRRRY-amide, CPK model), (PDB 1SEM). PRO3 of the Sos peptide (green) is targeted for a mutation to Ala. Shown in the brackets are alternative peptide conformations in the unbound state. The equilibrium binding constant can be expanded to account for the conformational equilibrium in the peptide, where K_{bind} is the equilibrium binding constant, K_{int} is the intrinsic binding constant, [*M*] is the total protein concentration, [*MP*] is the concentration of the SEM5:Sos complex, [P_{II}] is the concentration of peptide in the P_{II} conformation, k_i is the statistical weight of the *i*th peptide species, Q_{conf} is the conformational partition function of the peptide, ΔG_{int} is the change in the intrinsic Gibbs free energy upon binding, ΔG_{conf} is the change in conformational Gibbs free energy upon binding ($-RT$ ln Q_{conf}), and ΔG_{bind} is the total change in Gibbs free energy upon binding.

physical basis of the stability difference stems from the fact that the particular chemical structure of each amino acid determines the number of accessible conformations that it can adopt at each position when it is not in the folded structure. Likewise, when an unstructured peptide binds to a protein in a unique conformation, the degeneracy of the unbound state of the peptide will influence the binding affinity, independent of the interactions that are made at the binding interface. Indeed, substituting Ala or Gly at the solvent-exposed P3 position of Sos results in a significant difference in binding affinity for SEM5 (*24*).

Previous results from this laboratory (*24*) have demonstrated that although Ala, Pro, and Gly residues show a difference in the degree of P_{II} bias, with the rank order: Pro > Ala > Gly, the observed differences between each amino acid can be accounted for by the differences in the degeneracy of each residue when not in the P_{II} conformation. A direct consequence of this result is that the per residue equilibrium constant (K_i) between the P_{II} state and the ensemble of non-P_{II} conformations can be expressed as a product of two terms, the degeneracy (Ω_i) of all non-P_{II} conformations, which is residue-specific (*34*) and the microscopic P_{II} bias, κ_{mic}, which is common to Pro, Ala, and Gly.

$$K_i = \Omega_i \kappa_{mic} \quad (3)$$

Chart 1

State No.	Conformation	Stat. Wt. (SW)	
0	$P_1P_2P_3P_4P_5P_6$	1	1
1	$\boldsymbol{U}_1P_2P_3P_4P_5P_6$	K_1	$\Omega_1\kappa_{mic}$
2	$\boldsymbol{U}_1\boldsymbol{U}_2P_3P_4P_5P_6$	K_1K_2	$\Omega_1\Omega_2\kappa_{mic}^2$
3	$\boldsymbol{U}_1P_2\boldsymbol{U}_3P_4P_5P_6$	K_1K_3	$\Omega_1\Omega_3\kappa_{mic}^2$
•	•	•	•
2^{Nres}-1	$\boldsymbol{U}_1\boldsymbol{U}_2\boldsymbol{U}_3\boldsymbol{U}_4\boldsymbol{U}_5\boldsymbol{U}_6$	$K_1K_2K_3K_4K_5K_6$	$\prod_i^{N_i} \Omega_i \cdot \kappa_{mic}^{N_i}$
			$Q_{conf} = \Sigma$ SW

In other words, the results indicate that a baseline energetic bias exists that is common to Ala, Gly, and Pro residues, which can be interpreted as a per residue free energy difference ($\Delta g_{mic} = -RT$ ln κ_{mic}) between the ensemble of non-P_{II} conformations and the P_{II} state. The importance of eq 3 is that it provides straightforward means of simplifying the conformational free energy term (eq 1) for the Sos peptide. This is shown in Chart 1, where the ensemble of conformational states of the peptide (Q_{conf}) is enumerated and consists of all possible combinations for each amino acid to be in either the P_{II} conformation (P) or the disordered/non-P_{II} conformation (U).

The statistical weight (SW) of each conformational state is therefore the product of all κ_{mic} and degeneracy terms (Ω_i) at each position. If no P_{II} bias exists (i.e., $\kappa_{mic} = 1$), the statistical weight is determined only by the degeneracy and

the ensemble corresponds to a disordered state (*34*). If a P_{II} bias exists, however (i.e., $\kappa_{mic} < 1$), the statistical weight for adopting a non-P_{II} conformation will be decreased and the conformational free energy of folding the peptide into the binding competent conformation will likewise be decreased.

The ability to represent the P_{II} bias in terms of a baseline κ_{mic} (for Pro, Ala, and Gly) allows the conformational partition function, Q_{conf}, to be reduced to a binomial expansion of the energetic bias at each position, and eq 1 takes the form:

$$\Delta\Delta G_{bind}(P3 \rightarrow A) \approx \Delta\Delta G_{conf}(P3 \rightarrow A) = -RT \ln\left(\frac{\prod_{i=1}^{N_{res}}(1 + {}^{P3A}\Omega_i\kappa_{mic})}{\prod_{i=1}^{N_{res}}(1 + {}^{P3}\Omega_i\kappa_{mic})}\right) \quad (4)$$

where the products are over all residues in the P3A and P3−Sos peptides that occupy the P_{II} conformation in the bound state (i.e., residues 1−7) (*33*).

For the experimental protocol described here, a single amino acid substitution is made at the 3rd position of the peptide, leaving all other positions unchanged

P3−Sos: $Ac-V_0P_1P_2\mathbf{P}_3V_4P_5P_6R_7R_8R_9Y_{10}-NH_3$

P3A−Sos: $Ac-V_0P_1P_2\mathbf{A}_3V_4P_5P_6R_7R_8R_9Y_{10}-NH_3$

Because the degeneracy of the Pro residues are equivalent at position 2 of the P3A−Sos peptide and position 3 of the P3−Sos peptide as described previously (*24*), the change in binding free energy between SEM5 and the two Sos peptides (eq 4) simplifies to

$$\Delta\Delta G_{bind}(P3 \rightarrow A) \approx -RT \ln(1 + {}^{P3A}\Omega_3\kappa_{mic}) \quad (5)$$

where ${}^{P3A}\Omega_3$ is the degeneracy, relative to the P_{II} conformation, of the substituted Ala at position 3 in the P3A−Sos peptide. Of note is the fact that the only remaining term in eq 5 involves the description of the equilibrium at position 3, and this description in terms of κ_{mic} and ${}^{P3A}\Omega_3$ follows directly from experimental substitution data at that position (Table 1A).

The importance of eq 5 is that it reveals a direct connection between the experimentally observed difference in binding free energy and the microscopic equilibrium constant for P_{II} unfolding. Here, the previously determined backbone entropy contribution for Ala, $\Delta S = 4.1$ eu (entropy units = cal mol^{-1} K^{-1}) (34), is used, which according to eq 2, corresponds to a degeneracy of ${}^{P3A}\Omega_3 = 7.9$. Table 1A shows the values of κ_{mic} determined from the analysis of the binding of the P3 and P3A−Sos peptides to three different SEM5 analogues. The good agreement between the $\Delta\Delta G_{bind}$ values (and hence the κ_{mic}) obtained with different SEM5 variants indicates that the specific protein used in the study does not impact the analysis and that the effects can be entirely attributed to the changes in the peptide conformational ensemble as suggested by eq 5.

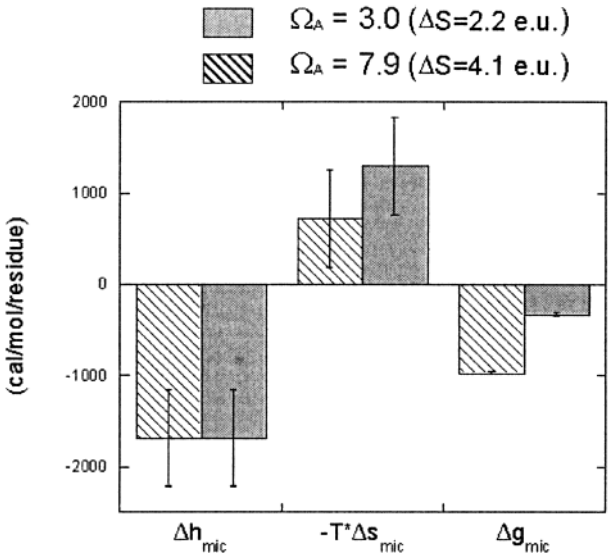

FIGURE 2: Microscopic thermodynamic parameters for P_{II} formation are shown, where the microscopic entropy is Δs_{mic}, the microscopic enthalpy is Δh_{mic}, and the microscopic free energy is Δg_{mic}, as shown in parts A and B of Table 1. The error bars are derived from the standard deviations of these terms based on the binding of the peptides to three different C-SH3 variants. The values are plotted with the disordered non-P_{II} conformation as the reference state (i.e., negative values indicate that the P_{II} conformation is favored). The solid bars correspond to the values determined using the lower limit for the degeneracy $\Omega_A = 3.0$ ($\Delta S = 2.2$ eu), and the hashed bars correspond to the values determined using $\Omega_A = 7.9$ ($\Delta S = 4.1$ eu).

The values of κ_{mic} shown in Table 1A indicate that the residue-specific energy of adopting a non-P_{II} conformation is unfavorable (i.e., positive), demonstrating that the P_{II} conformation is favored. The thermodynamic mechanism of the microscopic P_{II} bias (i.e., the component enthalpy, Δh_{mic}, and entropy functions, Δs_{mic}) can be obtained from the appropriate temperature derivative of eq 5

$$\Delta\Delta H_{bind}(P3 \rightarrow A) = -R\frac{\partial \ln(1 + {}^{P3A}\Omega_3\kappa_{mic})}{\partial (1/T)} = \frac{{}^{P3A}\Omega_3\kappa_{mic}}{1 + {}^{P3A}\Omega_3\kappa_{mic}}\Delta h_{mic} \quad (6)$$

Because $\Delta\Delta H_{bind}(P \rightarrow A)$ can be ascertained directly from ITC experiments (Table 1B), it is straightforward to obtain an estimate of the microscopic enthalpy Δh_{mic} and entropy Δs_{mic} [$=(\Delta h_{mic} - \Delta g_{mic})/T$] for the process. The average values are shown in Figure 2 and reveal that the energetically favorable P_{II} bias of -972 ± 24 cal mol^{-1} residue^{-1}, relative to the unfolded or random-coil state (i.e., all other states), is driven by a decrease in enthalpy (-1690 ± 530 cal mol^{-1} residue^{-1}), which is partially offset by an increase in entropy ($-T\Delta S = 718 \pm 530$ cal mol^{-1} residue^{-1}). We note from inspection of the data in Table 1B that the Δh_{mic} obtained from the analysis of the wild-type protein is significantly less than the values obtained from the S170A and S170G proteins. Although we have no unambiguous explanation for this difference, we do note that the values of N (number of binding sites) obtained from the fit differ considerably for the binding to the P3 and P3A peptides. Nonetheless, the results that are obtained from the analysis of each protein

are of the same sign and order of magnitude, and the corresponding uncertainty does not impact the interpretation of the results.

The microscopic P_{II} thermodynamic values (Figure 2) are important because they provide a per residue description of the energetic cost of P_{II} formation, which is common to Pro and Ala residues. The fact that the reported P_{II} bias in Gly is consistent with the difference in degeneracy between Pro, Ala, and Gly, as determined previously (*24*), suggests that indeed these values represent a baseline energetic cost for all amino acids to adopt P_{II}. Although the studies presented here do not address the molecular origins of observed energetics, the apparent similarity in the thermodynamic mechanism for the different amino acids would appear to implicate a phenomenon related to backbone−backbone or backbone−solvent interactions (*20*, *21*, *23*, *27*−*31*) or, as recently suggested, a combination of multiple subtle effects (*36*). It should be noted that, regardless of the intrinsic backbone contributions to the P_{II} formation identified here, it is also plausible that additional energetic factors related to specific side-chain−side-chain and side-chain−backbone interactions may need to be considered for other amino acids.

Several caveats should be noted, which are relevant to the interpretation of the values calculated with eqs 5 and 6. First, the value of κ_{mic} determined from eq 5 is inversely correlated to the value for the degeneracy that is used for Ala ($^{P3A}\Omega_3$) because they appear as a product in eq 5. As such, an overestimate of $^{P3A}\Omega_3$ will translate into an underestimate of κ_{mic}. The importance of this observation stems from the large variability in the reported conformational entropy of Ala, with some estimates being as low as 2.2 eu (*37*) (corresponding to an apparent degeneracy of $^{P3A}\Omega_3 = 3.0$), in contrast to the value of 4.1 eu used here. Although these estimates vary significantly, Figure 2 indicates that only the microscopic entropy Δs_{mic} and free energy Δg_{mic} and not the enthalpy Δh_{mic} are affected by the precise value of $^{P3A}\Omega_3$ used. As is evident, using realistic values for the high and low extremes of $^{P3A}\Omega_3$ does not affect the enthalpy of P_{II} formation, which is obtained directly from the calorimetric analysis (eq 5). In either case, the interpretation is the same; the P_{II} conformation is driven by a large favorable enthalpy.

Second, we note that the degeneracy term calculated from eq 2 represents an upper limit, because it is predicated on equal a priori probability for all of the non-P_{II} conformations. Use of this expression also carries with it the assumption that the apparent conformational entropy of Ala is temperature-independent, an assumption that is supported by the observation that Ala to Gly mutations result in similar free energy changes regardless of the experimental temperature (*24*, *34*, *35*). In any case, it is important to note that the specific form of the relationship between the degeneracy and the entropy does not qualitatively impact the interpretation of the results as shown in Figure 2.

The determination of the component microscopic enthalpy and entropy of the P_{II} formation represents a significant advance in our understanding of the observed conformational bias, and this level of understanding is important for two interrelated reasons. First, it provides unprecedented insight into the thermodynamic mechanism, which, as described below, is the basis for understanding how the bias will be affected by changes in the environmental conditions. Second, from a detailed knowledge of how the denatured ensemble responds to environmental changes, a quantitative impact of the P_{II} bias on the biophysical and functional properties can be assessed directly.

Consequences of the P_{II} Bias on Equilibrium Thermodynamic Measurements. The picture of the denatured state that emerges from this study reveals several consequences that must be considered in future attempts to model the unfolded states of proteins, as well as to interpret experimental results. Specifically, as the P_{II} bias is driven by a favorable and significant enthalpy, the effect of increased temperature will be to populate the ensemble of non-P_{II} states at the expense of P_{II}. Although a precise quantitative description of the impact of the temperature-induced redistribution within the denatured state ensemble will ultimately require explicit consideration of the energetic contributions of each residue in the sequence, the values determined here can be used to estimate the effect on a hypothetical polyalanine protein.

Because the P_{II} conformation is locally driven (*32*, *38*) and therefore highly noncooperative, the transition from P_{II} to the ensemble of non-P_{II} states will be characterized by two features. First, the overall probability of finding a molecule with a specific number of residues in P_{II} will be describable in terms of the individual probabilities at each position (see below). Second, the transition will occur over a broad temperature range, making it difficult to evaluate the impact of this transition on the thermodynamic parameters associated with unfolding of proteins. This can be demonstrated by considering the enthalpy fluctuations as a function of temperature. For the denatured state of a hypothetical polyalanine protein with *N* amino acids, wherein each residue can be folded in the P_{II} conformation in all combinations with the other residues in the sequence, the average excess enthalpy (relative to the P_{II} state) can be expressed as

$$\langle \Delta H_{P_{II}} \rangle = \sum_{i=1}^{N_{states}} P_i \Delta H_i = \sum_{m=0}^{N_{aa}} \frac{m\Delta h_{mic} \dfrac{^{P3A}\Omega_3{}^m \kappa_{mic}{}^m N!}{m!(N-m)!}}{\sum_{m=0}^{N_{aa}} \dfrac{^{P3A}\Omega_3{}^m \kappa_{mic}{}^m N!}{m!(N-m)!}} \quad (7)$$

where the summation on the right describes the contribution of all states that have *m* number of residues in the non-P_{II} conformation. As indicated by Figure 3A, the polyalanine denatured state ensemble experiences a gradual yet significant heat effect throughout the entire temperature range modeled (i.e., 50−423 K), as the ensemble transitions from the fully P_{II} to the fully non-P_{II} state. The corresponding excess heat capacity function (i.e., $\langle \Delta C_{p_{PII}} \rangle = \mathrm{d}\langle \Delta H_{P_{II}} \rangle / \mathrm{d}T$), which is the quantity that is obtained directly from differential scanning calorimetry (DSC) experiments, is likewise significantly affected by the shift in the denatured state ensemble (Figure 3B).

The practical manifestation of these effects can be observed when the magnitude of the expected heat effect for a protein of a specific size is compared to the heat capacity of unfolding (ΔC_{p_u}) for that same protein. Shown in Figure 3C are the experimental ΔC_{p_u} values for a database of proteins (Table 2) plotted as a function of the protein length (*39*). The linear regression provides an estimate of

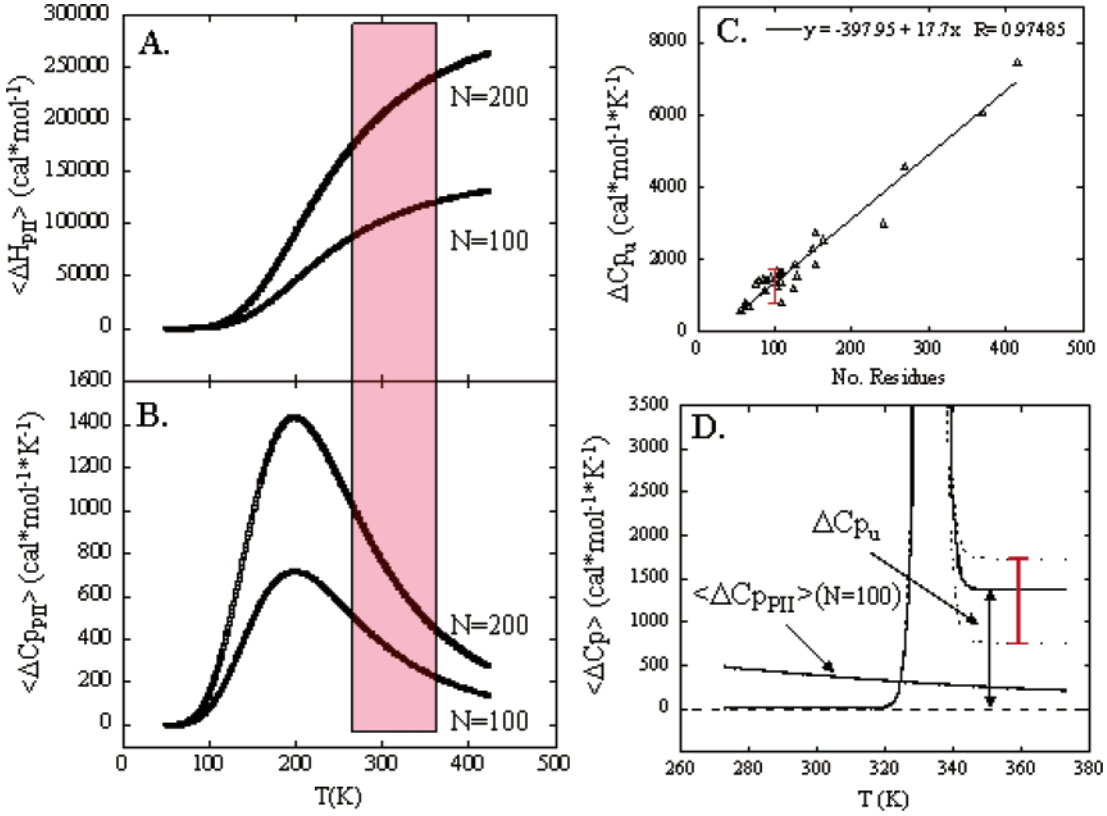

FIGURE 3: (A) Average excess enthalpy and (B) excess heat capacity for the transition from P_{II} to unfolded for a hypothetical 100 (N = 100) and a 200 amino acid polyalanine protein (N = 200). The highlighted area is the experimentally observable region for protein unfolding (~273−373 K). Although the midpoint of the transition is far below 273 K, the heat effect associated with the transition is nonetheless considerable at temperatures where proteins typically unfold. (C) Heat capacity of unfolding (ΔC_{p_u}) for the database of proteins shown in Table 2 plotted as a function of the number of residues. Linear regression identifies the mean ΔC_{p_u} for a given protein length, and the error bar at N = 100 (red) provides the approximate experimental range. (D) Simulated excess heat capacity function for a protein undergoing a two-state temperature induced transition with ΔC_{p_u} values corresponding to the mean and the approximate range (dashed curves) for a 100-residue protein (from Figure 3C). For comparison, the excess heat capacity function for the P_{II} to unfolded transition in the denatured state of a 100 amino acid protein is shown (from Figure 3B). Results show that the P_{II}-related denatured state fluctuations constitute a sizable fraction of the observed ΔC_{p_u}. Simulation parameters for the unfolding transition are ΔH_u = 150 kcal/mol, Tm = 60 °C (333.15 K), and ΔC_{p_u} = 750, 1360, and 1725 cal/(mol K).

the expected mean and range of ΔC_{p_u} values for each protein length. Despite the high correlation coefficient, there is considerable variability in ΔC_{p_u} for a given protein size (almost 100% at N = 100). The importance of this correlation, as well as the range in the data, becomes clear upon inspection of Figure 3D, which shows the simulated excess heat capacity function for a two-state unfolding transition of a protein, wherein the mean and the range of ΔC_{p_u} values from Figure 3C are modeled. As is well-known (*40*), the difference between the pre- and post-transitional baselines obtained from a DSC scan corresponds to the ΔC_{p_u} for the unfolding transition. Thus, from the range of ΔC_{p_u} values presented in Figure 3C for the unfolding of a 100 amino acid protein, Figure 3D shows the range of expected post-transitional baselines. Also shown in the figure is the heat effect associated with the fluctuations between P_{II} and non-P_{II} conformations within the denatured state ensemble (from Figure 3B). A comparison of this heat effect with the expected values for ΔC_{p_u} reveals at least two noteworthy features. First, the heat effect associated with the P_{II} fluctuations represents a sizable fraction of the apparent ΔC_{p_u}. In fact, from the range of values, the heat capacity change associated with these fluctuations represents between 10 and 30% of the total ΔC_{p_u} of unfolding.

Second, the apparent $\langle \Delta C_{p_{PII}} \rangle$ associated with fluctuations is temperature-dependent, decreasing at higher temperatures (Figure 3D). This result is a direct consequence of the fact that the midpoint of the P_{II} to non-P_{II} transition (where the heat effect is maximal) is centered far below 0 °C (273 K) (parts A and B of Figure 3). The excess heat capacity function associated with these fluctuations therefore decreases at temperatures above the midpoint of the transition (Figure 3B). Although slight differences in the P_{II} thermodynamics are likely to be found when other amino acids are investigated, the phenomenological implications are nonetheless not likely to be qualitatively different from what is shown in Figure 3D. Namely, the temperature-dependent redistribution of the denatured state ensemble will produce a substantial heat effect because of the magnitude of Δh_{mic}, and the magnitude of that heat effect will decrease with increased temperature, so long as the midpoint of the P_{II} transition shown in parts A and B of Figure 3 does not change dramatically (i.e., it occurs at a lower temperature than the Tm for the protein). The result will be an apparent temperature dependence in ΔC_{p_u}. It is noteworthy that the apparent temperature dependence in ΔC_{p_u} has been observed previously (*41*), although to date, the physical basis for this dependence has not been experimentally established. The

Table 2: Amino Acid Length and Heat Capacity for 31 Proteins

protein name	number of residues	ΔC_p
ovomucoid third domain (turkey)[a]	56	590
IgG-binding domain of protein G[b]	56	620
SH3 domain of α spectrin[c]	62	813
chymotrypsin inhibitor 2[d]	62	720
Btk[e]	67	740
ubiquitin[e]	76	1360
λ repressor 6−85[f]	80	1440
EcHPr[g]	85	1490
BsHPr[h]	88	1160
barstar[i]	89	1460
ribonuclease Sa[j]	96	1520
cytochrome *c* (horse heart)[k]	104	1730
ribonuclease T1[l]	104	1270
arc repressor[m]	106	1600
iso-1-cytochrome *c* (yeast)[n]	108	1370
thioredoxin (*E. coli*)[o]	108	1660
binase[p]	109	850
barnase[q]	110	1650
ribonuclease A[k]	124	1230
ROP[r]	126	1890
lysozyme (hen egg white)[k]	129	1540
lysozyme (human)[s]	130	1580
staphylococcal nuclease[t]	149	2320
interleukin 1-β[u]	153	1890
metmyoglobin (horse)[v]	153	1870
metmyoglobin (sperm whale)[k]	153	2770
T4 lysozyme (T54, A97)[w]	163	2570
α chymotrypsin[k]	241	3020
tryptophan synthase, α subunit[x]	268	4600
pepsinogen[k]	370	6090
phosphoglycerate kinase (yeast)[y]	415	7500

[a] Swint and Robertson (1993) (*51*). [b] O'Neil et al. (1995) (*52*). [c] Viguerra et al. (1994) (*53*). [d] Jackson et al. (1993) (*54*). [e] Knapp et al. (1998) (*55*). [f] Huang and Oas (1996) (*56*). [g] Nicholson et al. (1996) (*57*). [h] Scholtz, J. M. (1995) (*58*). [i] Agashe and Udgaonkar (1995) (*59*). [j] Pace et al. (1998) (*60*). [k] Privalov and Gill (1988) (*61*). [l] Yu et al. (1994) (*62*). [m] Bowie and Sauer (1989) (*63*). [n] Cohen and Pielak (1994) (*64*). [o] Santoro and Bolen (1992) (*65*). [p] Protasevich et al. (1987) (*66*). [q] Griko et al. (1994) (*67*). [r] Munson et al. (1994) (*68*). [s] Herning et al. (1992) (*69*). [t] Carra et al. (1994) (*70*). [u] Makhatadze et al. (1994) (*71*). [v] Kelly and Holladay (1990) (*72*). [w] Hu et al. (1992) (*73*). [x] Stackhouse et al. (1988) (*74*). [y] Murphy and Freire (1992) (*44*).

thermodynamic values for P_{II} unfolding reported here suggest that the P_{II}-related structural fluctuations in the denatured state are at least partially responsible for this behavior.

The significance of the results presented in Figure 3 is that it impacts the molecular interpretations of the origin of ΔC_{p_u} for proteins. As is well-known, the excess heat capacity function for a protein can be expressed as

$$\langle \Delta C_p \rangle = \frac{\mathrm{d}\sum_{i=1}^{N} P_i \Delta H_i}{\mathrm{d}\,T} = \sum_{i=1}^{N} \Delta H_i \frac{\partial P_i}{\partial T} + \sum_{i=1}^{N} P_i \Delta C_{p_i}$$

$$\langle \Delta C_p \rangle = \langle \Delta C_{p_{\mathrm{tran}}} \rangle + \langle \Delta C_{p_{\mathrm{baseline}}} \rangle \qquad (8)$$

where P_i, ΔH_i, and ΔC_{p_i} are the probability, enthalpy difference, and heat capacity difference (relative to the native state) for state *i* and the summation is over all *N* states accessible to the protein. The two terms in eq 8 represent contributions from distinctly different origins (*42*, *43*). The first term is the transition excess heat capacity, $\langle \Delta C_{p_{\mathrm{tran}}} \rangle$, and corresponds to the heat effect associated with fluctuations between the different conformational states of the protein. It is the $\langle \Delta C_{p_{\mathrm{tran}}} \rangle$ that is responsible for the bell-shaped curve seen in the DSC scans at temperatures close to the transition temperature (Figure 3D). The second term, which is referred to as the baseline excess heat capacity, $\langle \Delta C_{p_{\mathrm{baseline}}} \rangle$, corresponds to the population weighted difference in heat capacity between the beginning and the end state. For a two-state transition, the difference between the pretransitional heat capacity value, where the native state dominates, and the post-transitional signal, where the unfolded state is populated, provides access to the ΔC_{p_u} of unfolding, as shown in Figure 3D. Because the value of $\langle \Delta C_{p_{\mathrm{tran}}} \rangle$ presumably takes into account all of the heat effects associated with conformational fluctuations on the part of the protein, $\langle \Delta C_{p_{\mathrm{baseline}}} \rangle$ reflects the difference between how the native and denatured states of the molecule interact with the solvent (*42*, *43*). Indeed, numerous research efforts have focused on the correlations between the measured ΔC_{p_u} and changes in solvent-accessible surface area (*39*, *44*−*46*). A direct consequence of the temperature-dependent redistribution of the denatured state ensemble from the P_{II} to non-P_{II} states is the fact that the post-transitional baseline will not simply reflect the difference between how the native and denatured states interact with the solvent. On the contrary, the denatured state itself will be changing, and a sizable fraction of the ΔC_{p_u} will be attributable to the changing character of the denatured state throughout the unfolding transition. This contribution, which can vary in an unknown way, does not need to be proportional to the surface area in the denatured state of the molecule. However, even in cases where proportionality between surface area and ΔC_{p_u} is observed, the results presented here indicate that the proportionality constant cannot be compared in a physically meaningful way to values obtained from non-peptide-based model compound studies.

P_{II} Conformation and the Size of the Denatured State Ensemble. As recognized by Levinthal, an unguided search of conformational space from the unfolded random coil to the folded native state is not possible in a biologically reasonable time (*1*). This problem is often graphically depicted using the funnel representation of the energy landscape (Figure 4A), wherein the top of the funnel describes the unfolded conformations (*47*). The folding problem arises because the energy landscape of unstructured states is believed to be comparatively featureless (i.e., few low energy states exist). In the absence of distinct minima in the energy surface, protein molecules will be randomly distributed over a large number of conformations [e.g., a 100 amino acid protein with a degeneracy of 7.9 per residue gives an overall degeneracy of $\Omega = 7.9^{100}$ ($\sim 10^{89}$)]. Exploring such a search space is clearly intractable, and this point highlights the nature of the Levinthal paradox. Namely, although it is physically possible for a native state to be in equilibrium with a large ensemble of denatured conformations, the size of the ensemble will determine the duration of time it takes the system to reach equilibrium. As such, for a protein to achieve equilibrium in a biologically reasonable time, a conformational bias must exist in the denatured state that serves as a preorganization step, effectively reducing the search space accessible to the polypeptide (*20*, *23*).

The microscopic P_{II} parameters determined here provide a straightforward estimate of the quantitative impact of P_{II} on the accessible conformations in the denatured state. In the case of a single Ala residue, for example, the apparent

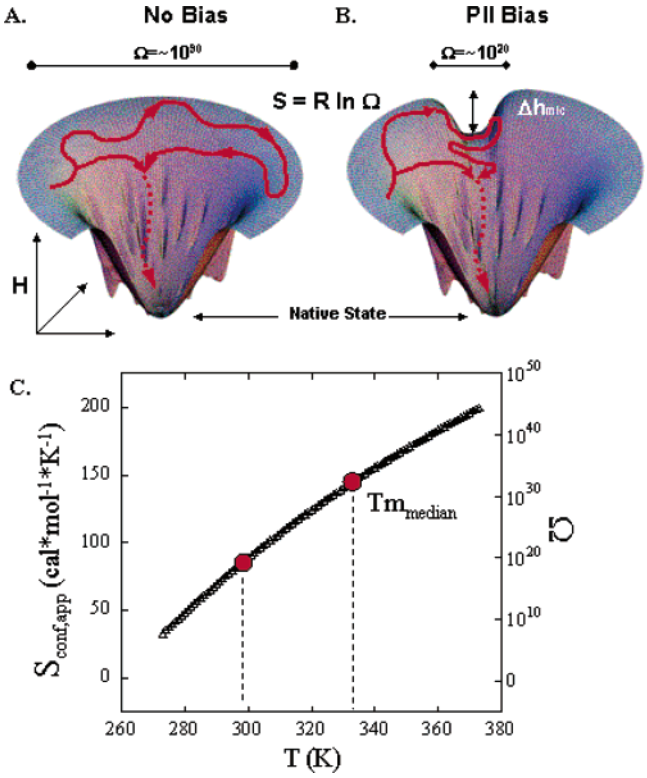

FIGURE 4: Folding funnel representation of the energy landscape of proteins (A) with no bias and (B) with a P_{II} bias. According to this representation, the native state is at the bottom of the funnel and the unfolded conformations are at the top. The width schematically depicts the conformational entropy of the denatured state ($S = R \ln \Omega$). When the P_{II} bias obtained here is applied to a 100 amino acid protein (schematically depicted in B as a groove with a depth $\Delta h_{mic} \sim -1.7$ kcal mol^{-1} residue^{-1}), the apparent degeneracy of the denatured state is decreased by a factor of $\sim 10^{70}$. The P_{II} bias requires that the conformational search of each amino acid will be significantly restricted relative to a completely random search. The solid red arrow represents a random search through unstructured conformations, and the dashed arrow represents the process involving the accretion of the native structure (i.e., the folding pathway). Note: the direction and length of the arrows are for illustrative purposes only and are not intended to imply the order or mechanism of folding. The fact that proteins fold rapidly implies that the classically defined folding pathway can be accessed from the P_{II} conformation. (C) Temperature dependence of the apparent conformational entropy ($S_{conf,app} = R \ln \Omega^{100}$) for a hypothetical 100 amino acid polyalanine protein. Shown are the values at 298 and 333 K (60 °C), which is the approximate median unfolding temperature for the database of proteins used to generate Figure 3C.

degeneracy $\Omega = 7.9$ in the absence of a P_{II} bias is reduced by a factor of ~ 5 (i.e., $1/\kappa_{mic} = \exp[-\Delta g_{mic}/RT]$) to $\Omega = 1.6$ when the P_{II} bias is considered. In the context of a 100 amino acid protein with a P_{II} bias at each position, the apparent degeneracy is reduced to $\sim 10^{19}$, a factor of more than 10^{70} over the unbiased ensemble! Consequently, the intrinsic P_{II} bias at each residue significantly *reduces and restricts* the conformational search space. In terms of the funnel representation, an energetic minimum of -1.7 kcal/mol is present at the P_{II} conformation for each residue (Figure 4B) and each molecule will spend a significant fraction of time in a conformation wherein some region of the molecule occupies the P_{II} space. Remarkably, from the κ_{mic} values determined here (Table 1), the probability of a 100 amino acid protein to adopt a conformation that does not have at least one residue in the P_{II} conformation is less than 1 part in 10^{20}. Interestingly, the fact that the P_{II} bias shows significant temperature dependence indicates that the apparent conformational entropy of the denatured state also increases with temperature. From the Δh_{mic} and Δs_{mic} values determined here, the impact of the shift in the P_{II} bias at high temperatures provides a quantitative estimate of the apparent change in the conformational entropy of unfolding (Figure 4C). For example, the denatured state of a 100 amino acid protein at 298 K is a factor of 10^{13} smaller than the denatured state at the approximate median unfolding temperature of the proteins listed in Table 2 (i.e., 333 K) (*48*) and a factor of 10^{24} smaller than the denatured state at 373 K.

Finally, although the current study does not address the kinetic aspects of P_{II} formation, one of the most important implications of this work is the fact that the folding process of proteins will be largely restricted to a specific region of conformational space, in effect providing a bias in the ensemble of starting conformations for the folding reaction. This finding by itself does not imply that folding should be accelerated relative to a more disordered or random-coil ensemble. The fact that proteins do indeed fold rapidly (*49*) indicates that the series of steps involving the accretion of higher order structure must be accessible from the P_{II} state (although such a scenario does not imply that P_{II} is an obligatory conformation). This is shown schematically in Figure 4, where the P_{II} bias in the denatured state is shown to restrict the conformational search to a space wherein the folding reaction can begin. If it is true that the P_{II} conformation is indeed a major reason that the folding reaction proceeds quickly, it would represent a somewhat paradoxical result. It is difficult to envision how an energetically favored conformation, which is not overwhelmingly prevalent in the final native folds of proteins (~6%) (*50*), could facilitate the folding reaction. It would seem at first consideration to significantly *slow* the folding process because it traps the residues in a low energy well, thus creating an energy gap between P_{II} and the native φ/ψ angles for each residue. Whether the folding process involves exhaustive exploration of an intractably large conformational space or whether it depends on multiple residues becoming excited from a low-energy state simultaneously, the problem would appear similarly challenging. However, the P_{II} conformation or any conformation for that matter could facilitate rapid folding if the energy barrier between the P_{II} conformation and native-state conformations was small or if the P_{II} state prevented nonproductive folding events. Along the lines of the latter, it is noteworthy that the P_{II} conformation is considerably expanded (i.e., it extends linearly and translates 3.12 Å/residue) and would disfavor (at least locally) the collapse of the polypeptide chain into small, nonspecific structural units. In effect, the P_{II} conformation could be viewed as a local conformational "place holder", preventing the misfolding of small units of the sequence. Whether this is indeed the case awaits further study, where the relationship between folding rates and P_{II} propensities can be quantitatively evaluated.

CONCLUSION

The results presented here provide the first calorimetric characterization of the thermodynamic origins (i.e., enthalpy and entropy) of the P_{II} bias in the disordered states of

peptides, which can be applied to the denatured states of proteins. We have used this bias to extract residue-specific energetics, which correspond to the backbone contribution for Ala. The component enthalpy and entropy functions of P_{II} formation indicate that the energetics are driven by a large favorable enthalpy change and are partially offset by an unfavorable entropy change. Furthermore, the thermodynamic parameters were used to provide an estimate of the quantitative impact of denatured state fluctuations on equilibrium thermodynamic measurements. In effect, the results described here represent an experimentally based quantitative description of the energetic consequences of the structural variations within the denatured state ensemble. Last, the values determined here reveal a denatured ensemble that is effectively reduced by a factor of 10^{70} at 298 K from that predicted for a random-coil model of an average size unfolded protein. However, because of the significant and favorable heat associated with adopting the P_{II} conformation, the P_{II} bias decreases with increased temperature, resulting in a sizable temperature dependence of the apparent conformational entropy of the denatured state.

ACKNOWLEDGMENT

The authors thank Drs. Bill DeGrado, Trevor Creamer, Rohit Pappu, Wayne Bolen, and Kip Murphy for critical review of this paper.

REFERENCES

1. Levinthal, C. (1969) How to fold graciously. Mossbauer spectroscopy in biological systems, *Proc. Univ. Illinois Bull. 41*, 22−24.
2. Kataoka, M., Nishii, I., Fujisawa, T., Ueki, T., Tokunaga, F., and Goto, Y. (1995) Structural characterization of the molten globule and native states of apomyoglobin by solution X-ray scattering, *J. Mol. Biol. 249*, 215−228.
3. Wong, K. B., Clarke, J., Bond, C. J., Neira, J. L., Freund, S. M., Fersht, A. R., and Daggett, V. (2000) Towards a complete description of the structural and dynamic properties of the denatured state of barnase and the role of residual structure in folding, *J. Mol. Biol. 296*, 1257−1282.
4. Wong, K. B., Freund, S. M., and Fersht, A. R. (1996) Cold denaturation of barstar: ^{1}H, ^{15}N, and ^{13}C NMR assignment and characterisation of residual structure, *J. Mol. Biol. 259*, 805−818.
5. Schwalbe, H., Fiebig, K. M., Buck, M., Jones, J. A., Grimshaw, S. B., Spencer, A., Glaser, S. J., Smith, L. J., and Dobson, C. M. (1997) Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea, *Biochemistry 36*, 8977−8991.
6. Gillespie, J. R., and Shortle, D. (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels, *J. Mol. Biol. 268*, 158−169.
7. Penkett, C. J., Redfield, C., Jones, J. A., Dodd, I., Hubbard, J., Smith, R. A., Smith, L. J., and Dobson, C. M. (1998) Structural and dynamical characterization of a biologically active unfolded fibronectin-binding protein from *Staphylococcus aureus*, *Biochemistry 37*, 17054−17067.
8. Bond, C. J., Wong, K. B., Clarke, J., Fersht, A. R., and Daggett, V. (1997) Characterization of residual structure in the thermally denatured state of barnase by simulation and experiment: Description of the folding pathway, *Proc. Natl. Acad. Sci. U.S.A. 94*, 13409−13413.
9. Zhang, O., and Forman-Kay, J. D. (1997) NMR studies of unfolded states of an SH3 domain in aqueous solution and denaturing conditions, *Biochemistry 36*, 3959−3970.
10. Zhang, O., Forman-Kay, J. D., Shortle, D., and Kay, L. E. (1997) Triple-resonance NOESY-based experiments with improved spectral resolution: Applications to structural characterization of unfolded, partially folded, and folded proteins, *J. Biomol. NMR 9*, 181−200.
11. Sreerama, N., and Woody, R. W. (1999) Molecular dynamics simulations of polypeptide conformations in water: A comparison of α, *β*, and poly(pro)II conformations, *Proteins 36*, 400−406.
12. Krimm, S., and Tiffany, M. L. (1974) The circular dichroism spectrum and structure of unordered polypeptides and proteins, *Isr. J. Chem. 40*, 14376−14383.
13. Woody, R. W. (1992) Circular dichroism and conformation of unordered polypeptides, *Adv. Biophys. Chem. 2*, 37−79.
14. Tiffany, M. L., and Krimm, S. (1972) Effect of temperature on the circular dichroism spectra of polypeptides in the extended state, *Biopolymers 11*, 2309−2316.
15. Tiffany, M. L., and Krimm, S. (1968) New chain conformations of poly(glutamic acid) and polylysine, *Biopolymers 6*, 1379−1382.
16. Drake, A. F., Siligardi, G., and Gibbons, W. A. (1988) Reassessment of the electronic circular dichroism criteria for random coil conformations of poly(L-lysine) and the implications for protein folding and denaturation studies, *Biophys. Chem. 31*, 143−146.
17. Dukor, R. K., and Keiderling, T. A. (1991) Reassessment of the random coil conformation: Vibrational CD study of proline oligopeptides and related polypeptides, *Biopolymers 31*, 1747−1761.
18. Park, S. H., Shalongo, W., and Stellwagen, E. (1997) The role of P_{II} conformations in the calculation of peptide fractional helix content, *Protein Sci. 6*, 1694−1700.
19. Ding, L., Chen, K., Santini, P. A., Shi, Z., and Kallenbach, N. R. (2003) The Pentapeptide GGAGG has P_{II} conformation, *J. Am. Chem. Soc 125*, 8092−8093.
20. Pappu, R. V., and Rose, G. D. (2002) A simple model for polyproline II structure in unfolded states of alanine-based peptides, *Protein Sci. 11*, 2437−2455.
21. Wilson, G., and Hecht, L. A. B. L. D. (1996) Residual structure in unfolded proteins revealed by Raman optical activity, *Biochemistry 35*, 12518−12525.
22. Vila, J. A., Baldoni, H. A., Ripoll, D. R., Ghosh, A., and Scheraga, H. A. (2004) Polyproline II helix conformation in a proline-rich environment: A theoretical study, *Biophys. J. 86*, 731−742.
23. Shi, Z., Olson, C. A., Rose, G. D., Baldwin, R. L., and Kallenbach, N. R. (2002) Polyproline II structure in a sequence of seven alanine residues, *Proc. Natl. Acad. Sci. U.S.A. 99*, 9190−9195.
24. Ferreon, J. C., and Hilser, V. J. (2003) The effect of the polyproline II (PPII) conformation on the denatured state entropy, *Protein Sci. 12*, 447−457.
25. Rucker, A. L., and Creamer, T. P. (2002) Polyproline II helical structure in protein unfolded states: Lysine peptides revisited, *Protein Sci. 11*, 980−985.
26. Chellgren, B. W., and Creamer, T. P. (2004) Short sequences of non-proline residues can adopt the polyproline II helical conformation, *Biochemistry 43*, 5864−5869.
27. Han, W.-G., Jalkanen, K. J., Elstner, M., and Suhai, S. (1998) Theoretical study of aqueous *N*-acetyl-L-alanine *N*′-methylamide: Structures and Raman, VCD, and ROA spectra, *J. Phys. Chem. B 102*, 2587−2602.
28. Eker, F., Griebenow, K., and Schweitzer-Stenner, R. (2003) Stable conformations of tripeptides in aqueous solution studied by UV circular dichroism spectroscopy, *J. Am. Chem. Soc. 125*, 8178−8185.
29. Eker, F., Cao, X., Nafie, L., and Schweitzer-Stenner, R. (2002) Tripeptides adopt stable structures in water. A combined polarized visible Raman, FTIR, and VCD spectroscopy study, *J. Am. Chem. Soc.* 124, 14330−14341.
30. Weise, C. F., and Weisshaar, J. C. (2003) Conformational analysis of alanine dipeptide from dipolar couplings in a water-based liquid crystal, *J. Phys. Chem. B 107*, 3265−3277.
31. Hinderaker, M. P., and Raines, R. T. (2003) An electronic effect on protein structure, *Protein Sci. 12*, 1188−1194.
32. Creamer, T. P. (1998) Left-handed polyproline II helix formation is (very) locally driven, *Proteins 33*, 218−226.
33. Lim, W. A., Richards, F. M., and Fox, R. O. (1994) Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains, *Nature 372*, 375−379.
34. D'Aquino, J. A., Gomez, J., Hilser, V. J., Lee, K. H., Amzel, L. M., and Freire, E. (1996) The magnitude of the backbone conformational entropy change in protein folding, *Proteins 25*, 143−156.
35. Smith, C. K., Bu, Z., Anderson, K. S., Sturtevant, J. M., Engelman, D. M., and Regan, L. (1996) Surface point mutations that significantly alter the structure and stability of a protein's denatured state, *Protein Sci. 5*, 2009−2019.

36. Drozdov, A. N., Grossfield, A., and Pappu, R. V. (2004) Role of solvent in determining conformational preferences of alanine dipeptide in water, *J. Am. Chem. Soc. 126*, 2574−2581.
37. Alexandrescu, A. T., Rathgeb-Szabo, K., Rumpel, K., Jahnke, W., Schulthess, T., and Kammerer, R. A. (1998) ^{15}N backbone dynamics of the S-peptide from ribonuclease A in its free and S-protein bound forms: Toward a site-specific analysis of entropy changes upon folding, *Protein Sci. 7*, 389−402.
38. Kelly, M. A., Chellgren, B. W., Rucker, A. L., Troutman, J. M., Fried, M. G., Miller, A. F., and Creamer, T. P. (2001) Host−guest study of left-handed polyproline II helix formation, *Biochemistry 40*, 14376−14383.
39. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant *m* values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding, *Protein Sci. 4*, 2138−2148.
40. Freire, E. (1995) Differential scanning calorimetry, *Methods Mol. Biol. 40*, 191−218.
41. Gomez, J., Hilser, V. J., Xie, D., and Freire, E. (1995) The heat capacity of proteins, *Proteins 22*, 404−412.
42. Freire, E. (1995) Thermal denaturation methods in the study of protein folding, *Methods Enzymol. 259*, 144−168.
43. Biltonen, R. L., and Freire, E. (1978) Thermodynamic characterization of conformational states of biological macromolecules using differential scanning calorimetry, *CRC Crit. Rev. Biochem. 5*, 85−124.
44. Murphy, K. P., and Freire, E. (1992) Thermodynamics of structural stability and cooperative folding behavior in proteins, *Adv. Protein Chem. 43*, 313−361.
45. Makhatadze, G. I. (1998) Heat capacities of amino acids, peptides, and proteins, *Biophys. Chem. 71*, 133−156.
46. Spolar, R. S., and Record, M. T., Jr. (1994) Coupling of local folding to site-specific binding of proteins to DNA, *Science 263*, 777−784.
47. Dill, K. A., and Chan, H. S. (1997) From Levinthal to pathways to funnels, *Nat. Struct. Biol. 4*, 10−19.
48. Freire, E. (1995) Thermodynamics of partly folded intermediates in proteins, *Annu. Rev. Biophys. Biomol. Struct. 24*, 141−165.
49. Ferguson, N., and Fersht, A. R. (2003) Early events in protein folding, *Curr. Opin. Struct. Biol. 13*, 75−81.
50. Sreerama, N., and Woody, R. W. (1994) Poly(pro)II helices in globular proteins: Identification and circular dichroic analysis, *Biochemistry 33*, 10022−10025.
51. Swint, L., and Robertson, A. D. (1993) Thermodynamics of unfolding for turkey ovomucoid third domain: Thermal and chemical denaturation, *Protein Sci. 2*, 2037−2049.
52. O'Neil, K. T., Hoess, R. H., Raleigh, D. P., and DeGrado, W. F. (1995) Thermodynamic genetics of the folding of the B1 immunoglobulin-binding domain from streptococcal protein G, *Proteins 21*, 11−21.
53. Viguera, A. R., Martinez, J. C., Filimonov, V. V., Mateo, P. L., and Serrano, L. (1994) Thermodynamic and kinetic analysis of the SH3 domain of spectrin shows a two-state folding transition, *Biochemistry 33*, 2142−2150.
54. Jackson, S. E., Moracci, M., elMasry, N., Johnson, C. M., and Fersht, A. R. (1993) Effect of cavity-creating mutations in the hydrophobic core of chymotrypsin inhibitor 2, *Biochemistry 32*, 11259−11269.
55. Knapp, S., Mattson, P. T., Christova, P., Berndt, K. D., Karshikoff, A., Vihinen, M., Smith, C. I., and Ladenstein, R. (1998) Thermal unfolding of small proteins with SH3 domain folding pattern, *Proteins 31*, 309−319.
56. Huang, G. S., and Oas, T. G. (1996) Heat and cold denatured states of monomeric λ repressor are thermodynamically and conformationally equivalent, *Biochemistry 35*, 6173−6180.
57. Nicholson, E. M., and Scholtz, J. M. (1996) Conformational stability of the *Escherichia coli* HPr protein: Test of the linear extrapolation method and a thermodynamic characterization of cold denaturation, *Biochemistry 35*, 11369−11378.
58. Scholtz, J. M. (1995) Conformational stability of HPr: The histidine-containing phosphocarrier protein from *Bacillus subtilis*, *Protein Sci. 4*, 35−43.
59. Agashe, V. R., and Udgaonkar, J. B. (1995) Thermodynamics of denaturation of barstar: Evidence for cold denaturation and evaluation of the interaction with guanidine hydrochloride, *Biochemistry 34*, 3286−3299.
60. Pace, C. N., Hebert, E. J., Shaw, K. L., Schell, D., Both, V., Krajcikova, D., Sevcik, J., Wilson, K. S., Dauter, Z., Hartley, R. W., and Grimsley, G. R. (1998) Conformational stability and thermodynamics of folding of ribonucleases Sa, Sa2, and Sa3, *J. Mol. Biol. 279*, 271−286.
61. Privalov, P. L., and Gill, S. J. (1988) Stability of protein structure and hydrophobic interaction, *Adv. Protein Chem. 39*, 191−234.
62. Yu, Y., Makhatadze, G. I., Pace, C. N., and Privalov, P. L. (1994) Energetics of ribonuclease T1 structure, *Biochemistry 33*, 3312−3319.
63. Bowie, J. U., and Sauer, R. T. (1989) Equilibrium dissociation and unfolding of the Arc repressor dimer, *Biochemistry 28*, 7139−7143.
64. Cohen, D. S., and Pielak, G. J. (1994) Stability of yeast iso-1-ferricytochrome *c* as a function of pH and temperature, *Protein Sci. 3*, 1253−1260.
65. Santoro, M. M., and Bolen, D. W. (1992) A test of the linear extrapolation of unfolding free energy changes over an extended denaturant concentration range, *Biochemistry 31*, 4901−4907.
66. Protasevich, I. I., Platonov, A. L., Pavlovsky, A. G., and Esipova, N. G. (1987) Distribution of charges in *Bacillus intermedius* 7P ribonuclease determines the number of cooperatively melting regions of the globule, *J. Biomol. Struct. Dyn. 4*, 885−893.
67. Griko, Y. V., Makhatadze, G. I., Privalov, P. L., and Hartley, R. W. (1994) Thermodynamics of barnase unfolding, *Protein Sci. 3*, 669−676.
68. Munson, M., O'Brien, R., Sturtevant, J. M., and Regan, L. (1994) Redesigning the hydrophobic core of a four-helix-bundle protein, *Protein Sci. 3*, 2015−2022.
69. Herning, T., Yutani, K., Inaka, K., Kuroki, R., Matsushima, M., and Kikuchi, M. (1992) Role of proline residues in human lysozyme stability: A scanning calorimetric study combined with X-ray structure analysis of proline mutants, *Biochemistry 31*, 7077−7085.
70. Carra, J. H., Anderson, E. A., and Privalov, P. L. (1994) Thermodynamics of staphylococcal nuclease denaturation. I. The acid-denatured state, *Protein Sci. 3*, 944−951.
71. Makhatadze, G. I., Clore, G. M., Gronenborn, A. M., and Privalov, P. L. (1994) Thermodynamics of unfolding of the all β-sheet protein interleukin-1 β, *Biochemistry 33*, 9327−9332.
72. Kelly, L., and Holladay, L. A. (1990) A comparative study of the unfolding thermodynamics of vertebrate metmyoglobins, *Biochemistry 29*, 5062−5069.
73. Hu, C. Q., Kitamura, S., Tanaka, A., and Sturtevant, J. M. (1992) Differential scanning calorimetric study of the thermal unfolding of mutant forms of phage T4 lysozyme, *Biochemistry 31*, 1643−1647.
74. Stackhouse, T. M., Onuffer, J. J., Matthews, C. R., Ahmed, S. A., and Miles, E. W. (1988) Folding of homologous proteins: Conservation of the folding mechanism of the α subunit of tryptophan synthase from *Escherichia coli*, *Salmonella typhimurium*, and five interspecies hybrids, *Biochemistry 27*, 824−832.

BI049352Z