Host−Guest Study of Left-Handed Polyproline II Helix Formation†

Melissa A. Kelly,‡ Brian W. Chellgren,‡ Adam L. Rucker,‡ Jerry M. Troutman,‡ Michael G. Fried,§ Anne-Frances Miller,∥,⊥ and Trevor P. Creamer*,‡,⊥

Department of Molecular and Cellular Biochemistry, Department of Chemistry, and Center for Structural Biology, University of Kentucky, 800 Rose Street, Lexington, Kentucky 40536, and Department of Biochemistry and Molecular Biology, The Pennsylvania State University College of Medicine, 500 University Drive, Hershey, Pennsylvania 17033

Received May 21, 2001; Revised Manuscript Received September 10, 2001

ABSTRACT: The importance of the left-handed polyproline II (PPII) helical conformation has recently become apparent. This conformation generally is involved in two important functions: protein−protein interactions and structural integrity. PPII helices play vital roles in a variety of processes including signal transduction, transcription, and cell motility. Proline-rich regions of sequence are often assumed to adopt this structure. Remarkably, little is known about the physical determinants of this secondary structure type. In this study, we have explored the formation of PPII helices by a short poly(proline) peptide. In addition, the results from experiments used to determine the propensities for apolar residues, plus glycine, asparagine, and glutamine, to adopt this structure in a poly(proline)-based host peptide are reported here. Proline possesses the highest intrinsic propensity, with glutamine, alanine, and glycine having surprisingly high propensities. β-Branched residues possess the lowest propensities of the residues examined. It is postulated that propensities possessed by apolar residues are due in part to peptide−solvent interactions, and that the remarkably high propensity possessed by glutamine may be due to a side chain to backbone hydrogen bond. These data are the first step toward a molecular understanding of the formation of this important, and yet little studied, secondary structure.

In recent years, the left-handed polyproline II (PPII)[1] helical conformation has been elevated from the status of a relatively rare and seemingly uninteresting secondary structure to one that is surprisingly common and of the utmost importance. This structure plays a central role in numerous vital processes including signal transduction, transcription, cell motility, and the immune response. Proline-rich ligands of the cytoskeletal protein profilin (*1*), as well as those of the SH3, WW, and EVH1 protein interaction domains, are bound in this conformation (*2*). The peptide ligands of class II MHC molecules are also bound in the PPII conformation (*3*). PPII helices are major features of collagens (*4*) and plant cell wall proteins (*5*). The PPII helix is believed to be the dominant conformation for many proline-rich regions of sequence (PRRs) (*6*). Sequences not rich in proline also adopt this structure. For example, poly(lysine), poly(glutamate), and poly(aspartate) peptides form PPII helices (*7*). Around 2% of all residues in known protein structures are found in PPII helices at least four residues long (*8*, *9*). As many as 10% of all residues are found in the PPII conformation, although not necessarily as part of PPII helices (*10*). PPII helices have also been hypothesized to be a major component of protein denatured states (*11*−*14*), giving them a role in a most fundamental process. Recently, Blanch et al. (*15*) have suggested that the PPII helix might be the precursor conformation in amyloid formation. Given the preceding, it is truly remarkable how little is known about the physical determinants of the PPII helical conformation. Here we are concentrating on the formation of PPII helices by poly(proline)-based peptides. These data will be most applicable to studies of the conformational properties of PRR's, but will also shed light on the ability of sequences poor in proline content to adopt this structure.

An ideal PPII helix has backbone dihedrals of $(\phi,\psi) = (-75°,+145°)$, leading to a very extended structure, with a highly solvated backbone, and three residues per turn (*8*). This conformation is known to be adopted by poly(proline) peptides and polymers (*16*). A cartoon of a poly(proline) peptide in the PPII conformation is shown in Figure 1. PPII helical structure arises in this case as a result of steric interactions between adjacent prolyl rings (*17*). Poly(proline) peptides can also adopt the compact right-handed polyproline I helical conformation, which has similar backbone dihedrals

† Acknowledgment is made to the donors of the Petroleum Research Fund, administered by the American Chemical Society, for support of this research.

* To whom correspondence should be addressed. Phone: (859) 323-6037. E-mail: trevor@euripides.gws.uky.edu.

‡ Department of Molecular and Cellular Biochemistry, University of Kentucky.

§ Department of Biochemistry and Molecular Biology, The Pennsylvania State University.

∥ Department of Chemistry, University of Kentucky.

⊥ Center for Structural Biology, University of Kentucky.

[1] Abbreviations: PPII, left-handed polyproline II helix; PRR, proline-rich region of sequence; SH3, sarc homology 3 domain; CD, circular dichroism; NMR, nuclear magnetic resonance; HPLC, high-performance liquid chromatography; UV, ultraviolet; Ac, acetyl; Gdn, guanidine; PPP, Ac−$(Pro)_7$-Gly-Tyr-NH_2; PAP, Ac-$(Pro)_3$-Ala-$(Pro)_3$-Gly-Tyr-NH_2; PGP, Ac−$(Pro)_3$-Gly-$(Pro)_3$-Gly-Tyr-NH_2; PLP, Ac−$(Pro)_3$-Leu-$(Pro)_3$-Gly-Tyr-NH_2; PMP, Ac-$(Pro)_3$-Met-$(Pro)_3$-Gly-Tyr-NH_2; PVP, Ac-$(Pro)_3$-Val-$(Pro)_3$-Gly-Tyr-NH_2; PIP, Ac−$(Pro)_3$−Ile-$(Pro)_3$-Gly-Tyr-NH_2; PNP, Ac-$(Pro)_3$-Asn-$(Pro)_3$-Gly-Tyr-NH_2; PQP, Ac-$(Pro)_3$-Gln-$(Pro)_3$-Gly-Tyr-NH_2; PAAP, Ac-$(Pro)_3$-$(Ala)_2$-$(Pro)_3$-Gly-Tyr-NH_2; TOCSY, total correlation spectroscopy.

10.1021/bi011043a CCC: $20.00
Published on Web 11/03/2001

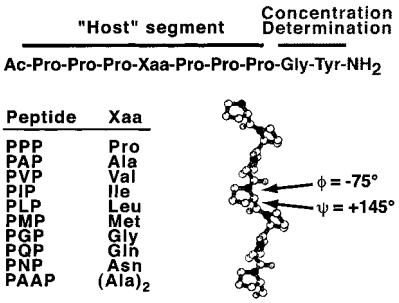

FIGURE 1: Host-peptide design and cartoon of an ideal PPII helix formed by poly(proline).

of $(\phi,\psi) = (-75°,+160°)$, all peptide bonds in the cis conformation and 3.3 residues/turn. However, this conformation is disfavored in aqueous solution (*18*).

The aim of this work was to develop a host−guest system suitable for deriving an intrinsic propensity scale for residues in PRRs to adopt the PPII helical conformation. We present here the host system, a short poly(proline) peptide, and a model for estimating peptide PPII helical content. The results of our initial propensity measurements for a number of residues are given. We have concentrated on aliphatic residues, plus glycine, asparagine, and glutamine, to test our host−guest system. We also present data on the propagation of the PPII helix through two adjacent alanines in our poly(proline)-based host system.

MATERIALS AND METHODS

Host−Guest Experiment Design. Host−guest experiments have proven to be an extremely useful vehicle for determining intrinsic propensities of residues to adopt particular secondary structures such as α-helices (*19*, *20*), β-sheets (*21*, *22*), and even collagen triple-helices (*23*). In such experiments, guest residues are inserted into a host peptide known to adopt the structure of interest. The effect of the guest residue upon the host structure is then measured. The host peptide used in this work is shown in Figure 1. A poly(proline)-based host was chosen since such peptides are known to form stable PPII helices in aqueous solution (*7*, *24*), and such a model is appropriate for studying the effects of nonproline residues upon the conformational properties of PRRs. Only three prolines are required to form and detect a PPII helix (*24*). The C-terminal proline of such a triplet is not necessarily in the PPII conformation since it is not followed by a restricting proline (*17*). The host peptide in Figure 1 will then always possess two short PPII helices, one on either side of the guest position. The C-terminal -Gly-Tyr pair was included for concentration determination purposes (*25*), and the N- and C-termini were acetylated and amidated respectively in order to remove charge interactions. Nine peptides with proline (peptide PPP; Figure 1), alanine (PAP), valine (PVP), isoleucine (PIP), leucine (PLP), methionine (PMP), asparagine (PNP), glutamine (PQP), and glycine (PGP) single guest residues have been examined. We have also examined the conformational properties of our host peptide with two adjacent alanines in the center (peptide PAAP).

Peptides and CD spectroscopy. Poly(proline)-based peptides were purchased from PeptidoGenic Research, and Co. (Livermore, CA) and purified via reversed-phase HPLC. The identities of the peptides were confirmed using mass spectrometry. Peptides were dissolved in buffer containing 5 mM potassium phosphate, 5 mM sodium fluoride, and 0.02% sodium azide, with the pH adjusted to 7.0. Peptide concentration was determined using the method of Brandts and Kaplan (*25*). Absorbance was measured using a 1.0 cm path-length cell in a Genesys 5 spectrophotometer. CD spectra were measured using a Jasco J-710 spectropolarimeter employing a water-jacketed 1 mm path-length quartz cuvette. Measurements were carried out at 0.5 nm resolution and a scan rate of 100 nm min.$^{-1}$ Reported spectra are averages of 30 scans with no smoothing. Spectra were measured with peptide concentrations of ∼100−200 μM and are corrected for solvent/buffer contributions. Errors in molar ellipticities are estimated to be around ±3%.

Equilibrium Sedimentation. Peptide samples in 100 mM potassium phosphate buffer were brought to sedimentation equilibrium in a Beckman XL-A analytical ultracentrifuge fitted with an AN-60Ti rotor, operating at 5.0 ± 0.1 °C. Data sets were collected at four rotor speeds (21 500, 30 000, 42 000, and 48 000 rpm). Absorbance values were measured at 275 nm as functions of radial position. Five scans were averaged for each sample at each rotor speed. Approach to equilibrium was considered to be complete when replicate scans separated by ≥6 h were indistinguishable.

At sedimentation equilibrium, the absorbance at a specified wavelength and position in the solution column is given by eq 1 (*26*, *27*).

$$A(r) = \sum_n \alpha_{n,0} \exp[\sigma_n(r^2 - r_0^2)] + \zeta \qquad (1)$$

Here, $A(r)$ is the absorbance at radial position r, the summation is over all species, n; $\alpha_{n,0}$ is the absorbance of the nth species at the reference position r_0, $\sigma_n = M_n(1 - \overline{\upsilon_n}\rho)\omega^2/2RT$ with M_n the molecular weight of the nth species, $\overline{\upsilon_n}$ its partial specific volume, ρ the solution density, ω the rotor angular velocity, R the gas constant, and T the absolute temperature. The baseline offset term ζ compensates for slight position-independent differences in the optical properties of different cell assemblies. The solvent density was calculated from the buffer composition using data tabulated in McRorie and Voelker (*26*). The value of $\overline{\upsilon_n}$ was estimated from the amino acid composition of the polypeptides by the method of Cohn and Edsall (*28*) and corrected for temperature according to McRorie and Voelker (*26*).

Weight-average molecular weights were estimated by simultaneous least-squares fitting of eq 1 (with $n = 1$) to multiple data sets ("global analysis") using the NONLIN 2 software running on a Macintosh computer (*29*).[2] Typical analyses combined three data sets, obtained at three rotor speeds.

NMR Spectroscopy. ^{1}H one-dimensional and two-dimensional ^{1}H total correlation spectroscopy [TOCSY (*30*)] spectra were collected for the Ace-(Pro)7-Gly-Tyr-NH2 peptide (PPP), along with a ^{1}H one-dimensional spectrum for monomeric proline. Samples were dissolved in 10% D_2O/

[2] NONLIN for the Macintosh was obtained from the website http://www.cauma.edu/software.

Table 1: Estimated Weight-Average Molecular Weights from Equilibrium Sedimentation Experiments[a]

peptide	actual MW	estimated weight-average MW	95% confidence interval in estimated MW
PPP	959	1069	223
PAP	933	1176	364
PGP	919	1009	245
PVP	961	1146	174
PAAP	1004	1192	233

[a] Three data sets were included in each fit. Actual molecular weights were calculated from the composition of each peptide.

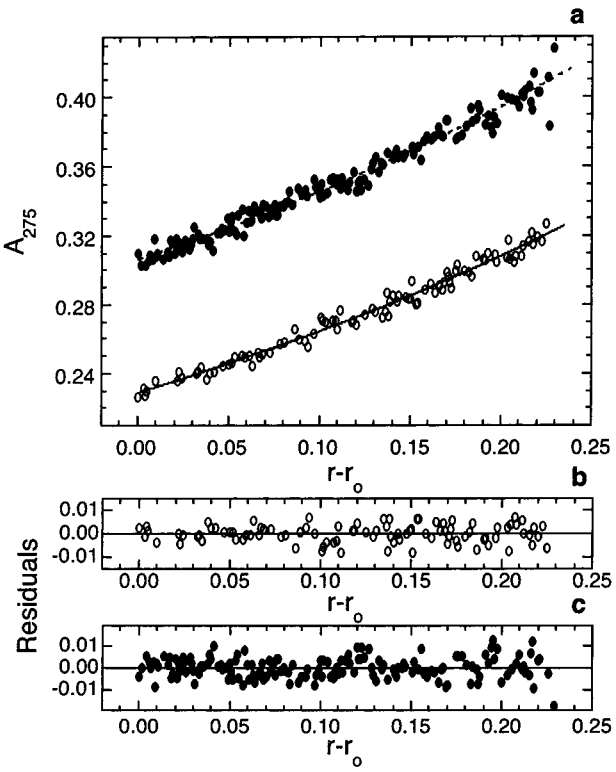

FIGURE 2: Results of equilibrium sedimentation experiments for peptides PPP (open symbols) and PGP (closed symbols). (a) Absorbance at 275 nm against radial position. Solid line is the fit for PPP, and dashed line is the fit for PGP. Data and fit for PGP have 0.1 added in order to off-set them from the PPP data and fit. Residuals from (b) PPP fit, and (c) PGP fit.

90% H_2O 5 mM potassium phosphate buffer at pH 7 with 0.02% azide as a preservative. All spectra were collected at 37 °C at 600 MHz with presaturation to suppress the resonance of solvent water. The TOCSY was collected with 512 complex points in t_1, 1 s acquisition time in t_2, 5400 Hz sweep widths in both dimensions, and a 100 ms spin lock at a field of 3.5 kHz.

RESULTS AND DISCUSSION

Oligomeric State of the Peptides. Sedimentation equilibrium experiments were performed for the PPP, PAP, PGP, PVP, and PAAP peptides. Estimated weight-average molecular weights are given in Table 1. As shown by the symmetrical residuals (Figure 2; data for peptides PPP and PGP shown) and the small range of 95% confidence intervals, the single species model (Equation 1 with $n = 1$) is fully compatible with these peptides being monomeric. Fits were not significantly improved by inclusion of additional species

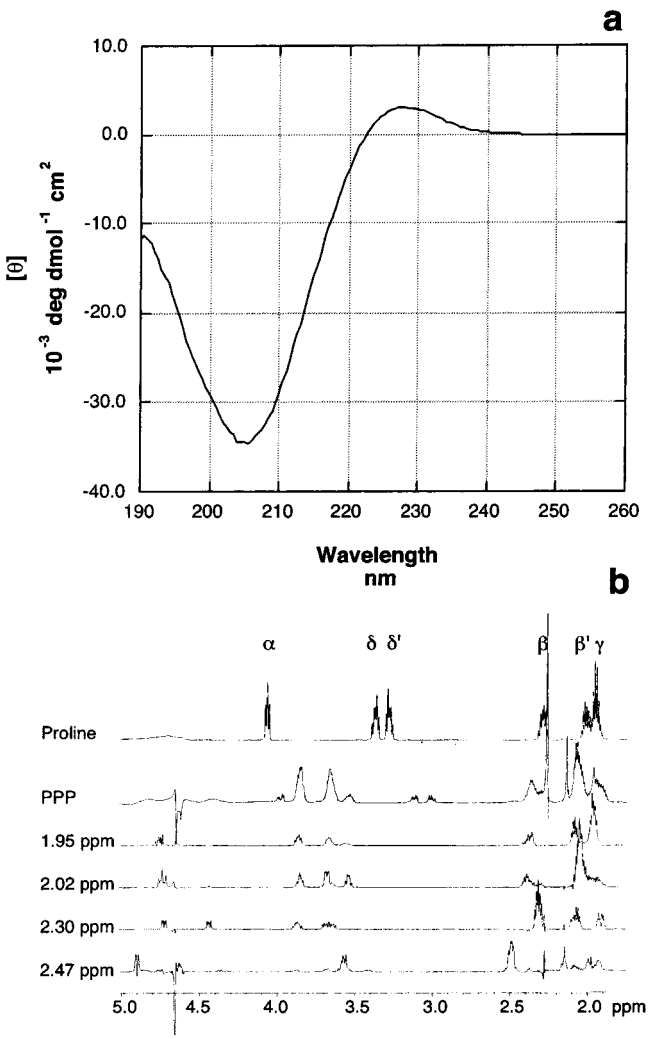

FIGURE 3: PPII character of the PPP peptide. (a) CD spectrum taken at 5 °C. (b) One-dimensional ^{1}H spectra for proline monomers and PPP, and slices from a two-dimensional ^{1}H TOCSY spectrum of PPP.

in the model. In particular, the data are incompatible with models in which a second species of molecular weight ≥2000 (expected for multimers) is present in significant concentrations.

The five peptides examined in the sedimentation equilibrium experiments were chosen as either a control that would not be expected to oligomerize (PPP) or as representative of those most likely to oligomerize. PGP is of particular interest since it is the most collagen-like in sequence. Clearly PGP is monomeric (Table 1 and Figure 2). Feng et al. (*31*) have demonstrated that at least five repeats of the collagen Gly-Pro-Hyp triplet are required for significant trimer formation. The PVP peptide might be expected to oligomerize or aggregate as a result of the bulky apolar valine side chain, and yet we see no evidence for this. The data in Table 1, which indicate that all five peptides examined are monomeric, are supported by concentration dependence experiments (data not shown). The CD spectra of all peptides listed in Figure 1 have been measured for peptide concentrations ranging from ~5 up to ~500 μM. We find that the molar ellipticities do not change with peptide concentration.

Poly(proline) Host−Guest Peptide Structure. The CD spectrum of the PPP peptide at 5 °C is shown in Figure 3a. This possesses the maximum at 228 nm and minimum at

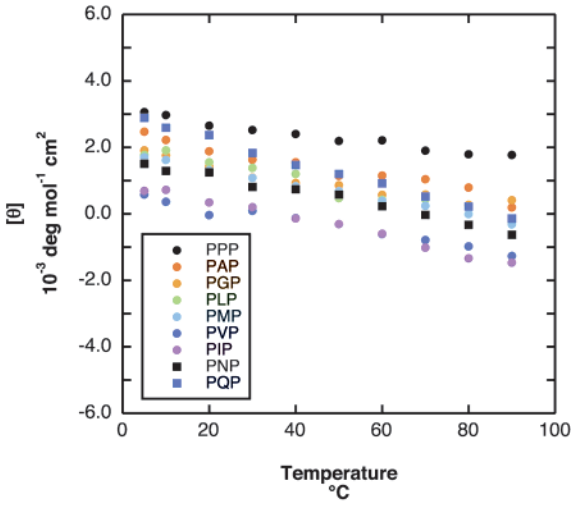

FIGURE 4: Temperature dependencies of the molar ellipticities at 228 nm for PPP, PAP, PGP, PLP, PMP, PIP, PVP, PNP, and PQP.

205 nm that are indicative of a PPII helix formed by a poly(proline) peptide in aqueous solution (*7*). Previous workers (*24*, *32*, *33*) have shown that the intensities of these positive and negative bands are peptide length dependent. Isemura et al. (*33*, *34*) have shown that these bands will increase in intensity, on a per residue basis, up to a peptide length of about 15 residues. Petrella et al. (*32*) suggest that this length dependence is due to a combination of electronic effects and flexibility at the ends of proline peptides. Calculations do not support the idea of flexibility at the termini, with the exception of the C-terminal proline (*17*), leaving electronic effects as the remaining potential explanation for the observed length dependence. The preceding aside, the CD spectrum in Figure 3a clearly demonstrates that our seven proline peptide possesses considerable PPII helical character.

Slices from a two-dimensional ^{1}H total correlation spectroscopy (TOCSY) spectrum (*30*) of PPP are shown in Figure 3b along with one-dimensional ^{1}H spectra for proline monomers and PPP. The PPP sample displays sharp contaminant signals, most likely from residual DMSO and acetone, near 2.3 and 2.2 ppm (Figure 3*b*). In the monomeric proline sample the α protons are visible near 4 ppm, β and β' near 2.3 and 2.1 ppm, γ protons all near 2.0 ppm, and δ and δ' near 3.2 and 3.3 ppm. In the one-dimensional PPP spectrum, proline resonances are visibly shifted and α resonances are obscured by residual water. Resonances near 3.0 and 3.1 ppm are assigned to the β protons of tyrosine and the quadruplet at 4.0 ppm is one of the α protons of glycine, the other being superimposed on the proline δ protons at 3.9 ppm. The PPP peptide proline resonances are heavily overlapped, indicating that each type of proton experiences predominantly the same environment, regardless of the particular proline. This high spectral overlap could be interpreted as complete absence of structure. However, the peptide proline resonances are shifted from those for monomeric proline in the same buffer at the same temperature (37 °C). In addition, the lines are significantly broadened. Combined with the CD spectrum (Figure 3a), these data are taken to indicate that PPP is highly structured in solution and that this structure is PPII helix.

The temperature dependence of the CD signals around the PPII maxima (228 nm) are shown in Figure 4 for all nine peptides containing a single guest residue. Although there appears to be a steady decrease in structure with increasing temperature, it is slight, and PPII structure is still present in all peptides at 90 °C. Clearly there are no large transitions. These peptides do not undergo a clear unfolding transition in this temperature range. This is not surprising since the host peptide will possess two short stretches of PPII helix on either side of the guest site. These arise as a result of strong steric interactions between prolyl rings (*17*), making the peptides difficult to completely denature.

Limits of PPII Structure. To quantify the effects of each guest residue upon PPII structure, it is necessary to define CD signals for both 100 and 0% PPII helix. We assume that the former corresponds to a rigid PPII structure whereas the latter refers to completely disordered states. Previous methods for estimating PPII content employed by Park et al. (*14*) and Bienkiewicz et al. (*35*) are not suitable for the peptides studied here. The poly(proline)-based peptides used in this work have CD spectra with the maxima and minima shifted to significantly higher wavelengths than those employed by Park et al. (*14*) and Bienkiewicz et al. (*35*). This wavelength shift is a result of the differences in absorbances of tertiary versus secondary versus primary amides (*7*). Consequently, the wavelengths used by these workers (222 and 200 nm, respectively) to estimate PPII helix content correspond to steep portions of CD spectra for poly(proline)-based peptides, introducing the possibility of large errors in PPII content estimates (see Figure 3a). New estimates of the upper and lower bounds for the characteristic maximum, which will be used here to estimate PPII helix content, need to be established.

Obtaining a rigid PPII helix using the PPP peptide is relatively straightforward. Poly(proline) peptides become stronger, more rigid PPII helices in the presence of the chaotropic agent guanidine (Gdn) (*7*). This is presumably the result of the denaturing agent binding to the backbone, forcing the polypeptide to become more extended. A poly(proline) peptide however is sterically limited to the PPII region of (ϕ,ψ) space (*17*), so addition of Gdn will result in a more rigid PPII helical structure. A Gdn solution with pH balanced to 7 was titrated into a PPP solution at 5 °C, up to a Gdn concentration of 8.4 M (Figure 5). Note that an 8.4 M Gdn solution at this temperature is essentially supersaturated and must be handled carefully in order to prevent precipitation in the cuvette. A dramatic increase in PPII content was observed. This increase was due to not only the prolines adopting a more rigid PPII structure, but presumably also to the C-terminal -Pro-Gly-Tyr triplet becoming more extended. The maxima at Gdn concentrations of 8.0 and 8.4 M are similar (Figure 5), however, it is possible that further increases would occur if higher concentrations were feasible at this temperature, and the maxima are not yet asymptotic with respect to increasing Gdn concentration. Since a Gdn concentration of 8.4 M cannot be exceeded, the $[\theta]_{max} = 7600$ deg dmol^{-1} cm^2 (at 228 nm) measured at this concentration is taken as the upper bound for PPII helical content.

Obtaining an estimate for 0% PPII helix content is more problematic. The PPP peptide does not lose significant structure even when heated to 90 °C (Figure 4). Toumadje and Johnson (*36*) used a peptide of sequence $(TSDSR)_3$ to model completely disordered states. At 80 °C, this peptide

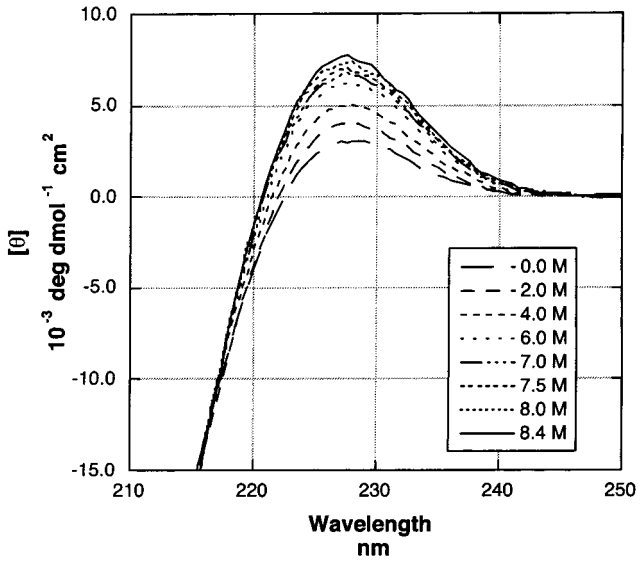

FIGURE 5: CD spectra around the maxima characteristic of PPII helix measured at 5 °C for PPP in increasing concentrations of Gdn.

has $[\theta]_{max} = -4300$ deg dmol^{-1} cm^2 (estimated from a shoulder corresponding to the characteristic maximum). However, Bienkiewicz et al. (*35*), using ellipticities at 200 and 222 nm, estimate that this peptide possesses 13% PPII content at 80 °C. Taking 7600 deg dmol^{-1} cm^2 as the upper limit and -4300 deg dmol^{-1} cm^2 as indicative of 13% PPII helix, the ellipticity for 0% PPII helix is estimated to be -6100 deg dmol^{-1} cm^2. Estimates of PPII helix content can then be obtained using

$$\% \text{ PPII} = \frac{[\theta]_{max} + 6100}{13\,700} \times 100 \qquad (2)$$

where $[\theta]_{max}$ is the molar ellipticity at the characteristic maximum. Note that the position of the characteristic maximum will vary slightly from the 228 nm of poly(proline) depending on the ratio of tertiary to secondary to primary amines in each peptide (*7*), as well as upon the nature of other structures adopted. This model should be used with caution. It has been derived for use with short proline-rich sequences. The models of Park et al. (*14*) and Bienkiewicz et al. (*35*) are more suitable for use with proline-poor sequences.

No claims as to the accuracy of the above model are made, as we have derived it simply as a means for estimating the relative PPII helical content of our peptides. We are assuming that 100% PPII helix content corresponds to a rigid PPII helix, which one could debate. Our lower bound is estimated from data for a peptide containing no proline (*36*) and, furthermore, makes use of other, approximate models for estimating PPII content (*14*, *35*). As noted above, these prior models are not suitable for use with our poly(proline)-based peptides. Ideally, one would want to obtain both the upper and lower bounds using the one poly(proline)-based peptide. However, as illustrated in Figure 4, all of the peptides examined retain significant PPII helix-character when heated to 90 °C. Consequently, we have little choice but to use two different peptides for the upper and lower bounds, leading to the approximate model derived above. Fortunately, errors in the upper and lower boundary estimates will simply move the estimates for all our peptides up or down without changing the relative ordering.

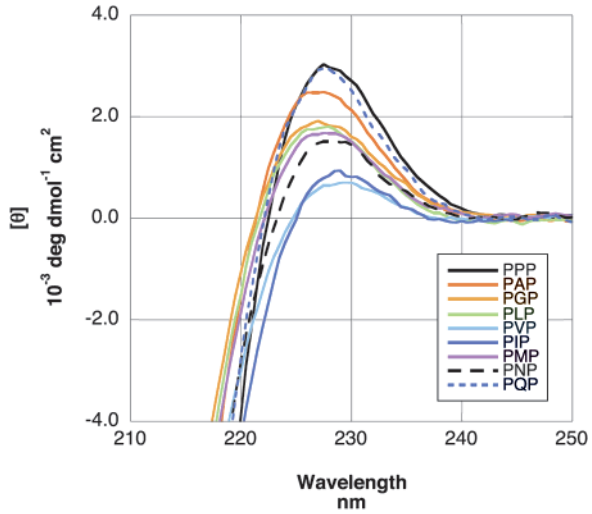

FIGURE 6: CD spectra around the maxima characteristic of PPII helix measured at 5 °C for PPP, PAP, PGP, PLP, PMP, PIP, PVP, PNP and PQP.

Table 2: Measured Maximum Ellipticities, Wavelength (λ_{max}) at Which the Maximum Occurs, PPII Helix Contents, and Estimated PPII Propensities, P_{PPII} for Residues in PPII Helices Found in Known Protein Structures

peptide	$[\theta]_{max}$ (deg dmol^{-1} cm^2)	λ_{max} (nm)	% PPII content	P_{PPII}[a]
PPP	3000	228.0	66 ± 2	5.06
PAP	2500	227.5	63 ± 2	0.84
PGP	1900	227.0	58 ± 2	0.27
PLP	1800	227.5	58 ± 2	0.91
PMP	1700	227.5	57 ± 2	0.86
PIP	700	229.5	50 ± 1	0.69
PVP	600	229.0	49 ± 1	0.82
PQP	2900	228.0	65 ± 2	1.24
PNP	1500	228.0	55 ± 2	0.71
PAAP	1300	227.0	54 ± 2	

[a] From Stapley and Creamer (*9*).

Host−Guest Measurements. CD spectra around the characteristic maxima for PPII helices, measured at 5 °C, are shown in Figure 6 for the peptides with single guest residues, and PPII helix contents estimated using eq 2 are given in Table 2. Errors listed for PPII content estimates are derived from estimated errors in measured molar ellipticities. The best PPII helix-former is proline, as might be expected. PPP is, on average, 66% PPII helical. One would not expect the C-terminal tyrosine to possess significant PPII character, although, as discussed below, this may not be true for the C-terminal glycine. The most C-terminal proline is not restricted by a following proline, can adopt any conformation usually available to proline (*17*), and will contribute less to $[\theta]_{max}$ than the other prolines. The implication is then that the first six prolines of PPP are predominantly in the PPII conformation, as predicted by earlier calculations (*17*).

The range of PPII helix contents is quite narrow, a consequence of the host peptide design. The two peptides with lowest PPII helix content, PIP and PVP, are around 50% PPII helix. The host peptide possesses three prolines on either side of the guest site. Two prolines in each of these triplets would be expected to be predominantly PPII helical at 5 °C due to steric interactions between prolyl rings (*17*). This results in a lower limit of ∼44% PPII helix content assuming that these four prolines are the sole contributors.

Since PIP and PVP are around 50% PPII helix in content, it is reasonable to conclude that isoleucine and valine do not spend much time in the PPII conformation, despite being restricted to the β-region of (ϕ,ψ)-space by the following proline (*17*, *37*). Leucine and methionine behave similarly and are better PPII helix formers than the two β-branched residues (Table 2 and Figure 5), indicating that the low propensities of isoleucine and valine are not due purely to hydrophobicity. Interestingly, glycine is as good as leucine at forming PPII helices. It is known that poly(glycine) can adopt left-handed helical structures similar to the PPII helix, suggesting that glycine has a relatively high intrinsic propensity to adopt this structure (*38*). Surprisingly, alanine has almost as high a PPII-forming propensity as proline in the poly(proline)-based host peptide (Table 2 and Figure 5). The high propensities of alanine and glycine suggest that, for nonproline residues, possession of a side chain that extends past the β-carbon may be detrimental to PPII formation.

Petrella et al. (*32*) have examined proline undecamers with alanine and glycine substituted into the central position. They find that these peptides have a lower PPII helix content, as adjudged from CD spectroscopy at 20 °C, than an all-proline peptide, and that alanine and glycine substitutions result in similar spectra. These findings are in agreement with our measurements for PAP and PGP.

In earlier work, we had found that the residue with the second highest frequency of occurrence in PPII helices in known protein structures was glutamine (*9*). A number of the glutamines found in these PPII helices made a side chain to backbone carbonyl oxygen hydrogen bond with the following residue in sequence. It was hypothesized that such hydrogen bonds could stabilize this conformation. To examine this we have measured the PPII helical content of peptides containing glutamine (PQP) and asparagine (PNP) guest residues. Asparagine is used as a control since modeling indicates that it is too short to make such a hydrogen bond. Glutamine is estimated to have a PPII helix-forming propensity almost as high as proline in our poly(proline) host peptide at 5 °C (Figure 5 and Table 2). Asparagine has a much lower propensity, slightly lower than that of methionine. Clearly glutamine in this context is a very good PPII helix former. Whether this high propensity is due in part to a side chain to backbone hydrogen bond remains to be seen. It is not clear that such a hydrogen bond would be particularly favorable in 55 M water. We do find that the PQP peptide loses PPII structure rapidly as temperature is increased (Figure 4), consistent with the idea of such a hydrogen bond.

Physical Determinants of PPII Structure. The relative PPII helix-forming propensities of the residues examined form a scale of order proline ≥ glutamine > alanine > (glycine, leucine, methionine) > asparagine > (isoleucine, valine). The origin of this ordering is not immediately clear. Proline is a very good PPII helix former simply because it is conformationally restricted by its own prolyl ring and is further restricted by steric interactions with the following proline (*17*, *37*). In addition, a proline guest will also restrict the preceding proline in our host peptide. An alanine guest is somewhat limited by steric interactions with the following proline but would not be expected to have much effect upon the preceding proline (*17*, *37*). Thus, the proline preceding the alanine guest would be expected to adopt all conformations generally accessible to proline unless interactions that extend past nearest neighbor residues favor the PPII helical conformation. The latter is not considered a possibility since prolines on either side of the guest site are too far apart to interact. One is left to conclude that an isolated alanine has a high propensity to adopt this conformation. The propensity for multiple alanines to favor the α-helical conformation is well-known, due in large part to intramolecular hydrogen bond formation and good van der Waals interactions (*39*). Such hydrogen bonds, and van der Waals interactions, are precluded in the peptides employed here. However, the backbone of the alanine residue is well-exposed to solvent in the PPII conformation, allowing for formation of multiple hydrogen bonds with water molecules (*40*). This suggests that alanine will tend to adopt the PPII conformation when it cannot form intramolecular hydrogen bonds.

The remarkably high propensity of alanine is reflected in the binding of proline-rich peptides to SH3 domains. Such ligands are bound in the PPII conformation (*2*) and possess considerable tendency to adopt this conformation in the unbound state (*41*). Wittekind et al. (*42*) performed alanine-scanning mutations through a proline-rich ligand known to bind to the N-terminal SH3 domain from the Grb2 adaptor protein. They found that mutations of proline to alanine in positions that did not directly contact the domain had very little effect upon the free energy of binding. If alanine were to significantly perturb the PPII conformation in the unbound state, one would expect the free energy of binding to become more unfavorable.

The propensities of the other residues can also be explained in part by interactions with solvent. The backbone of the glycine guest is well-exposed in the PPII conformation, allowing for efficient hydrogen bonding. However, restricting glycine to a single conformation carries a large entropic penalty, reducing its propensity to be part of a regular structure. Computer modeling reveals that the β-branched side chains of isoleucine and valine partially occlude their own backbones when in the PPII conformation (data not shown), preventing effective solvent interactions. The more flexible leucine and methionine side chains are less prone to occluding their own backbones, although they will do so in some side-chain conformations, making these residues better PPII helix formers than isoleucine and valine, but not as good as alanine. It should be noted that none of the apolar side chains can bury significant hydrophobic surface area in these peptides, even if prolines were to adopt the cis conformation (data not shown). We propose that backbone−solvent interactions are an important component of observed propensities and are testing this hypothesis in other work.

As noted above, the high propensity of glutamine might be explained in part by a side chain to backbone hydrogen bond. The observed temperature dependence of the positive band in the CD spectrum for PQP is consistent with this hypothesis (Figure 4). On the other hand, as noted above, the value of a single side chain to backbone hydrogen bond in 55 M water is of some debate. The lower propensity of asparagine could be due to the fact that it is incapable of forming an equivalent side chain to backbone hydrogen bond.

Propensities versus Occurrence in Protein PPII Helices. It is of interest to examine whether measured propensities correlate with occurrence of these residues in PPII helices

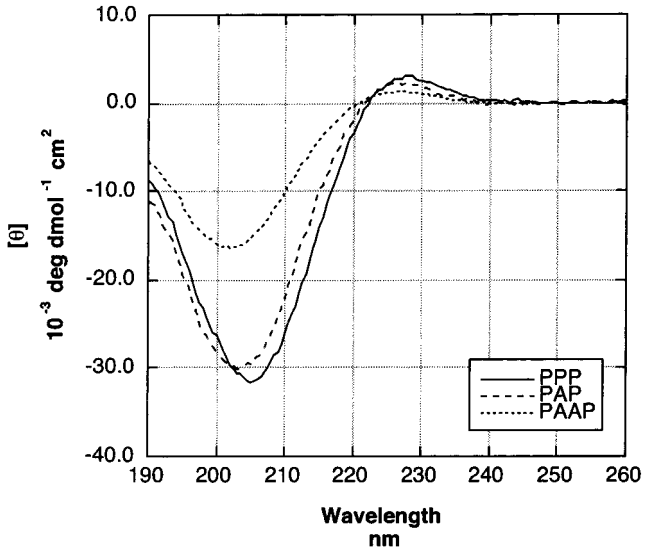

FIGURE 7: CD spectra measured at 5 °C for PPP, PAP, and PAAP.

in proteins of known structure. Stapley and Creamer (*9*) have estimated the Chou-Fasman tendencies for residues to be in PPII helices, P_{PPII}, using a dataset of 275 high-resolution protein structures. These are listed in Table 2. Visual inspection reveals little correlation. Glycine has the lowest P_{PPII} value and valine is as common as alanine in protein PPII helices. In our peptides however glycine has a relatively high propensity and valine a low propensity. Glutamine was found to have the highest tendency after proline, and this is reflected in the measured propensities. The lack of correlation between measured propensities and residue occurrence is not surprising since, as noted by Stapley and Creamer (*9*) and Adzhubei and Sternberg (*8*), PPII helices found in globular proteins tend to be very short (≤6 residues) and many are likely to arise as a result of tertiary interactions rather than as a result of intrinsic propensities.

Our hypothesis that backbone solvation is important may also lend some explanation as to the lack of correlation. Although the PPII helices in known protein structures tend to be on the surface and are more solvent exposed on average than residues in other structures (*8*, *9*), the residue backbones are still somewhat buried. Backbone polar atom solvent accessible surface area of apolar residues in protein PPII helices ranges from ∼8 to 12 Å^2 (*9*), significantly less than that estimated for the same residues in an isolated ideal PPII helix (∼20 to 30 Å^2). The backbone of glycine behaves similarly (12 Å^2 in protein PPII helices versus 33 Å^2 in an ideal PPII helix). Backbone−backbone hydrogen bonds were excluded from the survey of protein PPII helices, so these residues are hydrogen bonded to solvent, to side chains, or not at all (*9*). This reduces the potential influence of backbone solvation, and concomitantly increases the influence of tertiary interactions, upon residue occurrence.

Propagation through Multiple Nonproline Residues. PRRs generally contain numerous nonprolines (*6*). These are often found adjacent to one another or even in stretches of three or more. Given the high apparent propensity for alanine to be in the PPII helical conformation in a poly(proline)-based peptide, it is of interest to ask whether this conformation is propagated through two adjacent alanines. To this end, we have measured the CD spectrum of the peptide PAAP at 5 °C, which is presented in Figure 7 along with equivalent spectra for PPP and PAP. The estimated PPII helical content of PAAP at this temperature is given in Table 2. The decrease in PPII content going from PPP to PAP to PAAP is clearly nonlinear with respect to the number of alanines, with the difference between PAP and PAAP (9%) being much larger than that between PPP and PAP (3%). A nonlinear decrease is not surprising. The C-terminal of the two alanines will be restricted to the β-region of (ϕ,ψ)-space by the following proline and, consequently, should contribute somewhat more to the measured PPII content. The N-terminal alanine, however, is not restricted by the preceding proline, as noted in earlier calculations (*17*). Consequently, it has the possibility of behaving much like an alanine in a non-proline-rich environment, leading to more flexibility and the increased possibility of turn-like intrapeptide hydrogen bonds. Notably, however, PAAP still possesses significant PPII character, indicating that alanine does indeed have a relatively high intrinsic propensity to adopt this structure. The finding that a PPII helix can propagate through two adjacent non-prolines is in keeping with the work of Park et al. (*14*) who found that an alanine-rich peptide with a single proline inserted in the middle of the sequence possesses significant PPII character at low temperatures. In addition, Toumadje and Johnson (*36*) have demonstrated that systemin, a hormone-like peptide from plants, forms a PPII helix, despite only four of 18 residues being proline.

SUMMARY

We have presented here a host−guest system useful for determining the intrinsic propensities of residues to adopt the PPII helical conformation when in a proline-rich environment. When coupled with our model for estimating PPII helix content from CD spectra, it is possible to use this system to derive a scale of propensities. We have made a start on such a scale and found that glutamine, alanine, and glycine are surprisingly good PPII helix formers in this context. The β-branched valine and isoleucine appear to disfavor this conformation. We hypothesize that the resulting propensities are driven to a significant extent by the accessibility of the backbone of the guest residue when in this conformation and perhaps by a side chain to backbone hydrogen bond in the case of glutamine. In addition, we have used our system to demonstrate that proline-rich PPII helices propagate through two adjacent alanines, demonstrating that proline is not essential for the formation of this structure.

ACKNOWLEDGMENT

We thank David Rodgers and Ben Stapley for assistance with portions of this work, and Neville Kallenbach and Jeffrey Seale for useful discussions. Many thanks are due to Paul Bummer for use of his spectropolarimeter and for technical guidance.

REFERENCES

1. Mahoney, N. M., Janmey, P. A., and Almo, S. C. (1997) *Nat. Struct. Biol. 4*, 953−960.
2. Kay, B. K., Williamson, M. P., and Sudol, M. (2000) *FASEB J. 14*, 231−241.
3. Jardetzky, T. S., Brown, J. H., Gorga, J. C., Stern, L. J., Urban, R. G., Strominger, J. L., and Wiley, D. C. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 734−8.

4. Pauling, L., and Corey, R. B. (1951) *Proc. Natl. Acad. Sci. U.S.A. 37*, 272−281.
5. Ferris, P. J., Woessner, J. P., Waffenschmidt, S., Kilz, S., Drees, J., and Goodenough, U. W. (2001) *Biochemistry 40*, 2978−2987.
6. Williamson, M. P. (1994) *Biochem. J. 297*, 249−260.
7. Woody, R. W. (1992) *Adv. Biophys. Chem. 2*, 37−79.
8. Adzhubei, A. A., and Sternberg, M. J. E. (1993) *J. Mol. Biol. 229*, 472−493.
9. Stapley, B. J., and Creamer, T. P. (1999) *Protein Sci. 8*, 587−595.
10. Sreerama, N., and Woody, R. W. (1994) *Biochemistry 33*, 10022−10025.
11. Wilson, G., Hecht, L., and Barron, L. D. (1996) *Biochemistry 35*, 12518−12525.
12. Dukor, R. K., and Keiderling, T. A. (1991) *Biopolymers 31*, 1747−1761.
13. Tiffany, M. L., and Krimm, S. (1968) *Biopolymers 6*, 1379−1382.
14. Park, S. H., Shalongo, W., and Stellwagen, E. (1997) *Protein Sci. 6*, 1694−1700.
15. Blanch, E. W., Morozova-Roche, L. A., Cochran, D. A. E., Doig, A. J., Hecht, L., and Barron, L. D. (2000) *J. Mol. Biol. 301*, 553−563.
16. Tiffany, M. L., and Krimm, S. (1968) *Biopolymers 6*, 1767−1770.
17. Creamer, T. P. (1998) *Proteins 33*, 218−226.
18. Mutter, M., Wohr, T., Gioria, S., and Keller, M. (1999) *Biopolymers 51*, 121−128.
19. Chakrabartty, A., Kortemme, T., and Baldwin, R. L. (1994) *Protein Sci. 3*, 843−852.
20. Lyu, P. C., Liff, M. I., Marky, L. A., and Kallenbach, N. R. (1990) *Science 250*, 669−673.
21. Smith, C. K., Withka, J. M., and Regan, L. (1994) *Biochemistry 33*, 5510−5517.
22. Minor, D. L., Jr., and Kim, P. S. (1994) *Nature 367*, 660−663.
23. Persikov, A. V., Ramshaw, J. A. M., Kirkpatrick, A., and Brodsky, B. (2000) *Biochemistry 39*, 14960−14967.
24. Helbecque, N., and Loucheux-Lefebvre, M. H. (1982) *Int. J. Pept. Protein Res. 19*, 94−101.
25. Brandts, J. F., and Kaplan, L. J. (1973) *Biochemistry 12*, 2011−2024.
26. McRorie, D. K., and Voelker, P. J. (1993) *Self-Associating Systems in the Analytical Ultracentrifuge*, Beckman Instruments, Inc., Palo Alto.
27. Laue, T. M., and Stafford, W. F., III (1999) *Annu. Rev. Biophys. Biomol. Struct. 28*, 75−100.
28. Cohn, E. J., and Edsall, J. T. (1943) *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions*, Reinhold, New York.
29. Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) *Biophys. J. 36*, 575−588.
30. Braunschweiler, L., and Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.
31. Feng, Y., Melacini, G., Taulane, J. P., and Goodman, M. (1996) *J. Am. Chem. Soc. 118*, 10351−10358.
32. Petrella, E. C., Machesky, L. M., Kaiser, D. A., and Pollard, T. D. (1996) *Biochemistry 35*, 16535−16543.
33. Isemura, T., Okabayashi, H., and Sakakibara, S. (1968) *Biopolymers 6*, 307−321.
34. Okabayashi, H., Isemura, T., and Sakakibara, S. (1968) *Biopolymers 6*, 323−330.
35. Bienkiewicz, E. A., Moon Woody, A.-Y., and Woody, R. W. (2000) *J. Mol. Biol. 297*, 119−133.
36. Toumadje, A., and Johnson, W. C., Jr. (1995) *J. Am. Chem. Soc. 117*, 7023−7024.
37. MacArthur, M. W., and Thornton, J. M. (1991) *J. Mol. Biol. 218*, 397−412.
38. Bhatnagar, R. S., and Gough, C. A. (1996) in *Circular dichroism and the conformational analysis of biomolecules* (Fasman, G. D., Ed.) pp 183−199, Plenum Press, New York.
39. Aurora, R., Creamer, T. P., Srinivasan, R., and Rose, G. D. (1997) *J. Biol. Chem. 272*, 1413−1416.
40. Sreerama, N., and Woody, R. W. (1999) *Proteins 36*, 400−406.
41. Viguera, A. R., Arrondo, J. L. R., Musacchio, A., Saraste, M., and Serrano, L. (1994) *Biochemistry 33*, 10925−10933.
42. Wittekind, M., Mapelli, C., Lee, V., Goldfarb, V., Friedrichs, M. S., Meyers, C. A., and Mueller, L. (1997) *J. Mol. Biol. 267*, 933−952.

BI011043A