PROTEINS: Structure, Function, and Genetics 33:218–226 (1998)

Left-Handed Polyproline II Helix Formation Is (Very) Locally Driven

Trevor P. Creamer*
Kentucky Center for Structural Biology, Department of Biochemistry, University of Kentucky, Lexington, Kentucky

ABSTRACT **The left-handed polyproline II helix (PPII) is believed to be the preferred conformation for proline-rich regions of sequence in proteins. Such regions have been postulated to be protein-protein interaction domains. The formation of this structure is studied here using simple Monte Carlo computer simulations employing the hard sphere potential. It is found that polyproline sequences adopt only the PPII structure in the simulations. Non-proline, non-glycine residues inserted as guests into polyproline host peptides are conformationally restricted by the following proline residues and tend to be part of the PPII helix. It is found through insertion of two alanine residues into polyproline that the PPII structure is not propagated through more than one non-proline residue. This finding calls into question the hypothesis that proline-rich regions will preferentially adopt this structure since many such sequences are comprised of less than 50% proline residues. Proteins 33:218–226, 1998.** © 1998 Wiley-Liss, Inc.

Key words: proline; polyproline; protein structure; local structure; secondary structure

INTRODUCTION

Proline-rich regions (PRR) of sequence are becoming of increasing interest as potential protein interaction domains.[1,2] Perhaps the most commonly studied example is the binding of PRR's by SH3 domains.[3–5] SH3 domains bind PRR's of approximately ten residues in length containing the consensus sequence HPpHP (H = hydrophobe, P = proline and p = usually proline).[4] Interestingly it has recently been found that another small modular domain, the WW domain, competes with SH3 domains for the same PRR's.[6,7]

Examples of other PRR's abound: RNA polymerase II possesses a long C-terminal PRR that has been implicated in interactions with various transcription factors.[8] This PRR "tail" has a repeating heptad sequence of $(YSPTSPS)_{26}$ and is thought to interact with the PRR's of various transcription factors.[1,8] The structure of the actin-binding protein profilin, in complex with a ten-residue polyproline peptide, has recently been solved.[9] The Huntington's disease gene product has been found to contain a PRR, also rich in glutamine, that is thought to interact with other proteins, perhaps by binding SH3 domains.[10] The rich variety of proteins containing PRR's has recently been reviewed by Williamson.[1]

PRR's are commonly believed to preferentially adopt the left-handed polyproline II helical conformation (PPII).[1,2] This conformation is illustrated for a polyproline sequence in Figure 1. In the PPII conformation the peptide adopts average backbone dihedrals of $(\phi,\psi) = (-75°, +145°)$ and does not form any intramolecular backbone hydrogen bonds.[11] As such it can be considered to be a form of regular secondary structure.[11,12] This is a very extended helical structure with three residues per turn with the proline residues in the *trans* conformation (Fig. 1). Notably the carbonyl oxygens of each proline residue are solvent exposed, allowing for hydrogen bonding by solvent, a potentially stabilizing interaction for the PPII conformation. By contrast, the polyproline I helical conformation has backbone dihedrals centered around $(-75°,+160°)$, involves *cis* proline and is right-handed.[11]

Sreerama and Woody[12] have estimated that about 10% of all protein residues adopt the PPII conformation. PPII helices of four or more residues are not uncommon—around half of proteins of known structure possess one,[11] comprising approximately 2% of all residues. However, the vast majority of those observed in globular proteins are short: more than 90% are only four or five residues in length.[11] This is in contrast to many PRR's which range from around ten residues for those bound by SH3 domains[4] to well over a hundred residues in human RNA polymerase II.[1] Notably, it has been found that PPII helices in globular proteins do not necessarily have to contain any proline residues, although most do.

Formation of the PPII conformation in PRR's is generally believed to arise as a result of conformational restrictions imposed by proline upon the preceding residue.[1] It is well known that proline restricts the preceding residue in sequence to the β-region of (ϕ,ψ) dihedral space.[13] Proline residues

*Correspondence to: Trevor P. Creamer, Kentucky Center for Structural Biology, Department of Biochemistry, University of Kentucky, 800 Rose Street, Lexington, KY 40536–0298. E-mail: trevor@euripides.gws.uky.edu
Received 28 April 1998; Accepted 26 June 1998

© 1998 WILEY-LISS, INC.

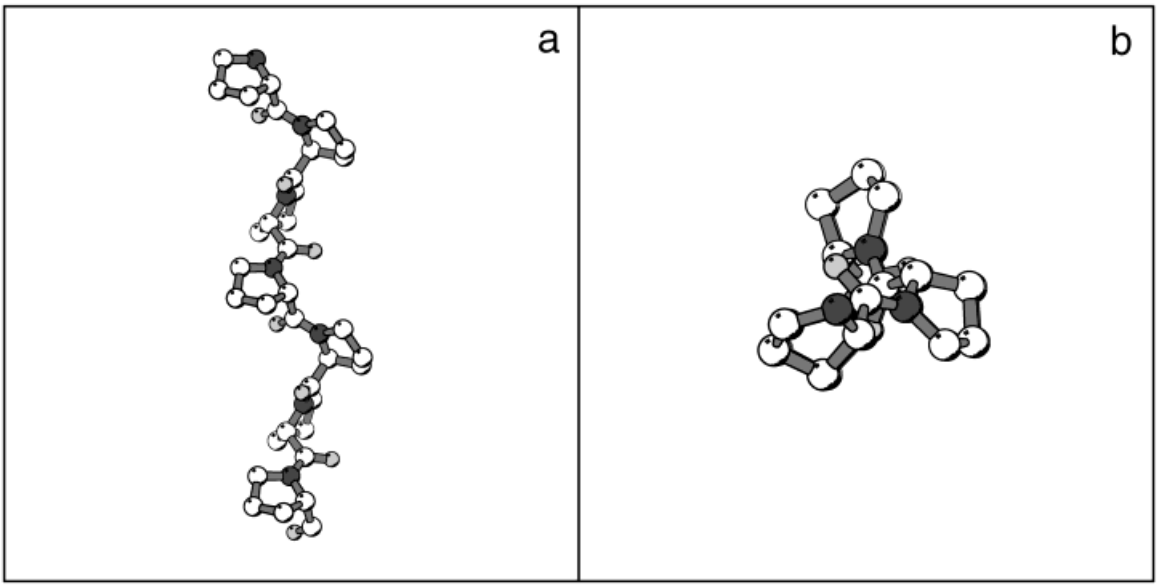

Fig. 1. Ideal seven-residue all-proline PPII helix with $\phi = -75°$ and $\psi = +145°$. Panel a shows the helix side-on. Panel b shows the helix looking down the long axis. The three perfect three-fold symmetry is clear in this view. The figure was generated using MOLSCRIPT.

are generally restricted to two narrow regions of dihedral space centered at (ϕ,ψ) = ($-61°$, $-35°$) and ($-65°$, $+150°$).[13] A proline that is followed in sequence by another proline will consequently be restricted to the region around (ϕ,ψ) = ($-65°$, $+150°$). A pure polyproline peptide will then be restricted to an extended conformation with all residues (except the last) in this dihedral region. The last residue does not have a following proline to restrict its conformational space.

In this work the formation of the PPII helical conformation is investigated. Using computer simulations employing the hard sphere potential it is shown that pure polyproline sequences do indeed adopt the PPII conformation preferentially and that this is a result of steric interactions between the prolyl rings and the peptide backbone. A series of computer "host-guest" experiments demonstrates that the PPII conformation is adopted by single, non-proline, non-glycine residues in the center of a polyproline host peptide. As with the polyproline sequence, this is a result of steric interactions. Notably, it is found that the PPII conformation is not propagated through more than one non-proline residue inserted into a polyproline sequence. This calls into question the belief that PRR's, many of which contain less than 50% proline (e.g. RNA polymerase II), will preferentially adopt the PPII structure in solution.

METHODS

Monte Carlo Computer Simulations

Distributions of peptide conformations are readily obtained from Monte Carlo computer simulations using Metropolis sampling.[14] The Metropolis algorithm is a widely used method of exploring the conformational space available to a molecular system by sampling from the Boltzmann distribution. The algorithm involves making random perturbations to a system, starting from some initial configuration, in order to generate a new trial configuration. The difference in energy, ΔE, between the trial and initial configurations is then calculated. If (*i*) $\Delta E < 0$, or (*ii*) $\Delta E > 0$ and a normal random deviate between 0 and 1 is less than the Boltzmann-weighted energy, $e^{-\Delta E/RT}$, the trial configuration is accepted and becomes the new initial configuration. If neither set of conditions is satisfied, the trial configuration is rejected, and the initial configuration is retained.

The above sampling scheme is repeated iteratively in cycles. A cycle is defined in this work as N_{cycle} attempts at generating trial configurations. N_{cycle} is equal to the number of rotatable dihedral angles in the peptide being simulated. Each attempt to generate a new trial configuration involves making a randomly generated rotation ($\leq \pm 180°$) about a dihedral that is chosen at random. This type of perturbation has been employed in previous work.[15,16] Data is collected at the end of each cycle. The resulting distribution of configurations will satisfy the Boltzmann distribution.

The simulations here employ the hard sphere potential with atomic radii as used in previous work.[16] This potential makes use of the united atom approximation: the hydrogens attached to carbon atoms are not treated explicitly, rather the radii of carbon atoms are inflated to account for attached hydrogens. In this potential, each atom is treated as

TABLE I. Peptides Simulated and Related Simulation Information

Simulation name	Peptide sequence[a]	Simulation equilibration period (cycles)	Simulation length (cycles)	Cycles between collected conformations
A9	Ace-Ala_9-NMe	10^4	2×10^6	500
APA	Ace-Ala-Pro-Ala-NMe	10^3	5×10^5	100
P9	Ace-Ala-$(Pro)_9$-Ala-NMe	2×10^4	5×10^6	1000
PAP	Ace-Ala-$(Pro)_3$-Ala-$(Pro)_3$-Ala-NMe	10^4	2×10^6	500
PVP	Ace-Ala-$(Pro)_3$-Val-$(Pro)_3$-Ala-NMe	10^4	2×10^6	500
PLP	Ace-Ala-$(Pro)_3$-Leu-$(Pro)_3$-Ala-NMe	10^4	2×10^6	500
PFP	Ace-Ala-$(Pro)_3$-Phe-$(Pro)_3$-Ala-NMe	10^4	2×10^6	500
PGP	Ace-Ala-$(Pro)_3$-Gly-$(Pro)_3$-Ala-NMe	10^4	2×10^6	500
PA2P	Ace-Ala-$(Pro)_2$-$(Ala)_2$-$(Pro)_3$-Ala-NMe	10^4	2×10^6	500
A4P4	Ace-$(Ala)_4$-$(Pro)_4$-Ala-NMe	10^4	4×10^6	200
P4A4	Ace-Ala-$(Pro)_4$-$(Ala)_4$-NMe	10^4	4×10^6	200

[a]All peptide models were acetylated (Ace-) and methyl-amidated (-NMe).

having only excluded volume effects with no attractive components. The hard sphere potential has only two possible energy states: zero when there are no atomic overlaps, or infinite energy when atoms overlap. Thus all allowed conformations have the same energy (zero). Solvent is not treated in any manner in these simulations: since there are no electrostatic interactions in the hard sphere potential it would be somewhat misleading to say the solvent is treated as a dielectric continuum.

Peptides Simulated

The peptides simulated are listed in Table I, along with the equilibration period and length of each simulation in cycles. The number of cycles between conformations collected for analysis is also listed for each peptide. Due to the nature of the hard sphere potential, these simulations did not require extensive equilibration periods or large numbers of cycles between collected conformations. Note that the ϕ dihedral of proline is fixed at $-75°$ in the hard sphere potential employed and only *trans* proline conformations were considered: *cis* conformations were not generated in the simulations. For technical reasons each peptide has alanine residues in the first and last positions in the sequence.

RESULTS

"All-Proline" Peptide

The P9 [Ace-Ala-$(Pro)_9$-Ala-NMe] peptide was simulated in order to study its conformational behavior and determine whether it preferentially adopts the PPII helical conformation. This peptide is found to adopt only the PPII structure. The average ψ dihedral angles for each proline residue in this peptide are listed in Table II. Note that the ϕ dihedral angles of the proline residues are fixed at $-75°$. As can be seen from the average ψ angles, Pro_2 through Pro_9 are restricted to the PPII region of (ϕ,ψ) space (approximately $+143 \pm 14°$). Pro_{10} possesses more conformational freedom (see Table II), as would be expected given that the conformational restrictions imposed by a proline are upon the preceding residue.[13] The lack of conformational freedom is illustrated in Figure 2. This is a MOLSCRIPT depiction[17] of ten conformations of the P9 peptide drawn randomly from the computer simulation. The conformations are superimposed upon the central proline residue using the MSI InsightII (San Diego, California) software package. Careful inspection shows that these ten structures diverge relatively slowly from each other moving along the sequence in either direction from the central proline.

TABLE II. Average ψ Dihedrals and Standard Deviations of All Proline Residues in the P9 Simulated Peptide[†]

Residue	$\langle\psi\rangle \pm$ Standard deviation
Pro_2	$+142 \pm 16°$
Pro_3	$+142 \pm 16°$
Pro_4	$+143 \pm 14°$
Pro_5	$+144 \pm 13°$
Pro_6	$+145 \pm 13°$
Pro_7	$+144 \pm 13°$
Pro_8	$+143 \pm 15°$
Pro_9	$+142 \pm 15°$
Pro_{10}	$+91 \pm 95°$

[†]The ϕ dihedrals are fixed at $-75°$.

Host-Guest Experiments

A series of "host-guest" simulations were run to examine the effect of introducing a non-proline residue into a polyproline host peptide. The residues Ala, Val, Leu, Phe and Gly were each substituted into position Xaa of the host peptide Ace-Ala-$(Pro)_3$-Xaa-$(Pro)_3$-Ala-NMe. The average (ϕ,ψ) dihedral angles for the guest residues (from peptides PAP, PVP, PLP, PFP, and PGP) are shown in Table III. For comparison, corresponding data for peptides A9 and APA are also listed. The alanine, valine, leucine and phenylalanine guest residues are all found to be restricted

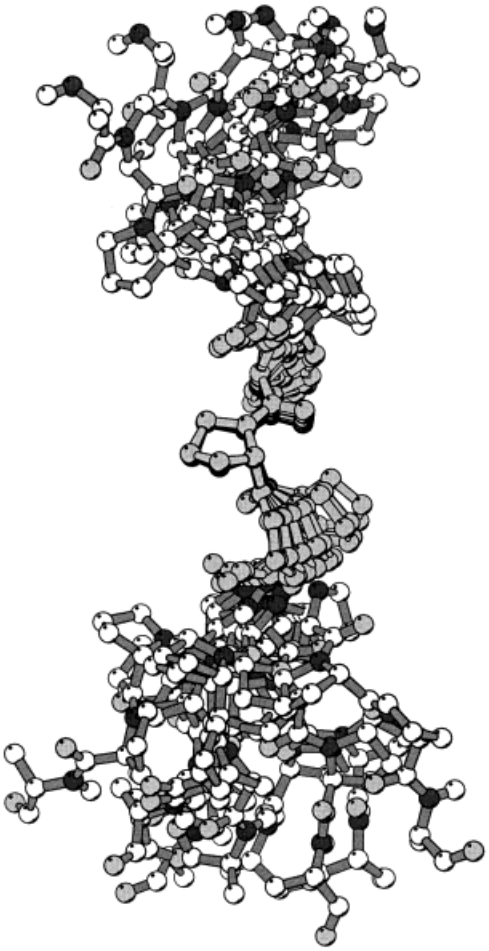

Fig. 2. Ten conformations of the simulated P9 peptide superimposed on the central proline residue. The central three residues are colored gray in order to highlight the narrow range of conformations available to this triad. The structures were drawn randomly from the set of conformations generated in the Monte Carlo computer simulation. Superimpositions were generated using MSI InsightII (San Diego, California) software. The figure was generated using MOLSCRIPT.

to the β-region of (ϕ,ψ) space. Figure 3 is a Ramachandran plot of the (ϕ,ψ) backbone dihedral distribution for Ala_5 in PAP. Figure 4 consists of histograms of the distribution of ϕ (panel a) and ψ (panel b) dihedral angles for this residue. For comparison the corresponding data for Ala_5 in the A9 peptide are included in both Figures 3 and 4. The ϕ distribution of Ala_5 in PAP shows a slight increase in population in the range −170° to −120° relative to Ala_5 in A9 (Figure 4, panel a). There is a clear preference for ψ dihedrals in the β-region, with a particular increase in the population of residues in the region ψ = +110° to +170° (see Figure 4, panel b). The leucine and phenylalanine guest residues are found to behave almost identically to the alanine guest (data not shown).

Judging by the standard deviations in average (ϕ,ψ) dihedrals (Table III), valine is somewhat more restricted than the other non-glycyl residues. Figure 5 consists of a Ramachandran plot (panel a) and ψ distribution plot (panel b) for the valine guest residue in PVP. Once again, data for Ala_5 from A9 is included for comparison. The Ramachandran plot (panel a) demonstrates that the valine is somewhat more conformationally restricted than alanine in a polyproline host. More notably, the ψ distribution for valine (panel b) is greatly enhanced in the range +130° to +150° (compare the scales for panels b in Figures 4 and 5). The ϕ distribution for Val_5 is almost identical to that for Ala_5 in PAP (data not shown).

TABLE III. Average (ϕ, ψ) Dihedrals and Standard Deviations of Selected Residues

Peptide	Residue	⟨ϕ⟩ ± Standard deviation	⟨ψ⟩ ± Standard deviation
A9	Ala_5	−113 ± 39°	+77 ± 83°
APA	Pro_2	−75°	+100 ± 95°
PAP	Ala_5	−120 ± 37°	+118 ± 28°
PVP	Val_5	−121 ± 30°	+124 ± 24°
PLP	Leu_5	−116 ± 37°	+115 ± 26°
PFP	Phe_5	−116 ± 37°	+115 ± 27°
PGP	Gly_5	−1 ± 128°	+7 ± 130°
PA2P	Ala_5	−119 ± 39°	+118 ± 29°
PA2P	Ala_4	−113 ± 37°	+74 ± 84°
A4P4	Ala_4	−119 ± 39°	+118 ± 28°
A4P4	Ala_3	−112 ± 40°	+75 ± 83°
P4A4	Ala_6	−112 ± 38°	+74 ± 85°
P4A4	Ala_7	−113 ± 40°	+73 ± 86°

Glycine in PGP does not appear to be significantly effected by the polyproline host. As can be seen in Figure 6, the distribution of (ϕ,ψ) dihedrals for Gly in PGP on a Ramachandran plot appear to reproduce that of glycine residues in any sequence. In addition, the distribution of ψ dihedral angles is not enhanced in the β-region of (ϕ,ψ) space (data not shown). This is in keeping with the findings of MacArthur and Thornton[13] for glycine residues preceding a proline.

The proline residue following each guest site, Pro_6, possesses average backbone dihedrals of (ϕ,ψ) = (−75°, +145 ± 13°), indicating that this is in the PPII helical region. The proline preceding each guest site, Pro_4, has average dihedrals of (ϕ,ψ) = (−75°, +99 ± 97°), indicating that it is not restricted to be within the PPII helical conformation. The guest residue would then appear to introduce some flexibility in the PPII helix at the preceding residue.

PPII Structure Direction And Propagation

Simulations of three peptides, A4P4, P4A4, and PA2P, were run to examine the sequence direction of PPII helix formation and its possible propagation through more than one non-proline residue. The A4P4 and P4A4 peptides were designed to study both of these phenomena in the N- and C-terminal directions respectively. The PA2P peptide was designed to look at the effects of having two non-proline

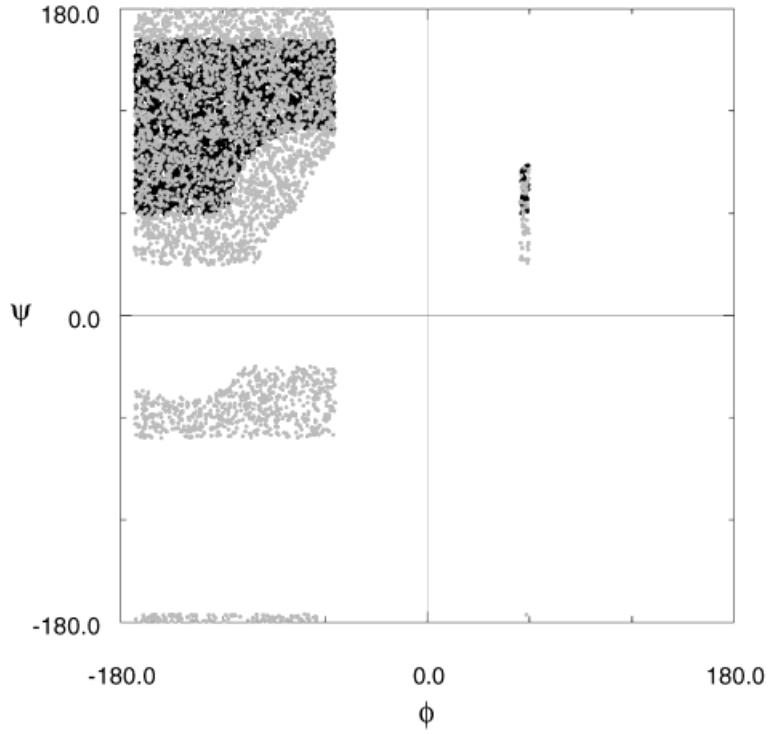

Fig. 3. Ramachandran plot of the (ϕ,ψ) dihedrals of Ala_5 in the PAP peptide (black dots) along with the corresponding data for Ala_5 in the A9 peptide for comparison (gray dots).

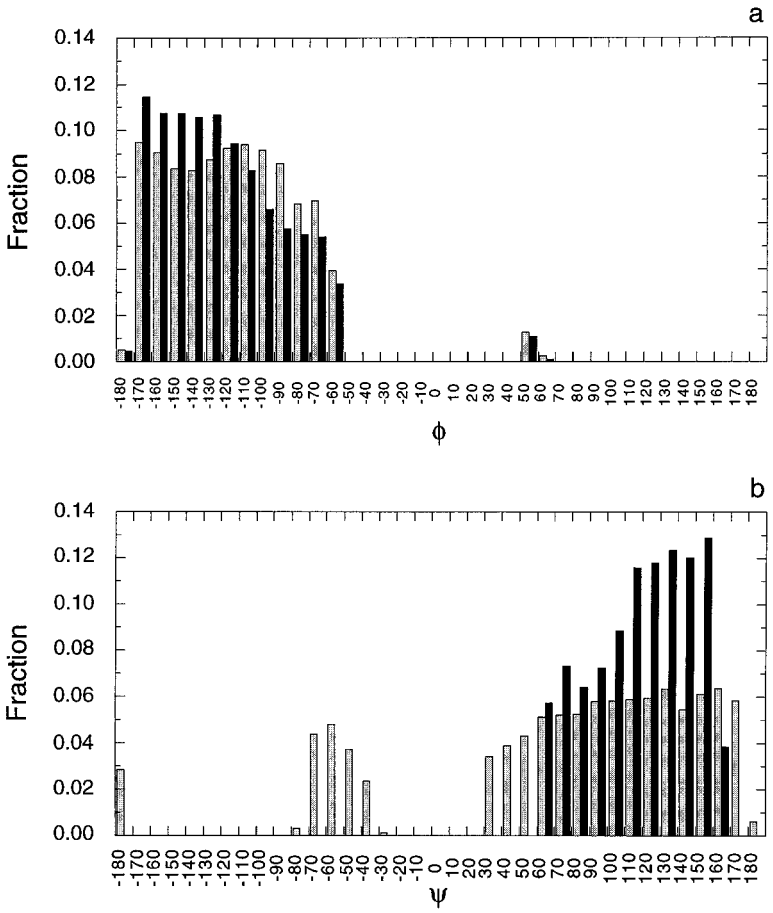

Fig. 4. **Panel a** is a histogram of the distribution of ϕ dihedral angles for Ala_5 in PAP (black bars) and Ala_5 in A9 (gray bars). **Panel b** shows the distribution of ψ dihedral angles.

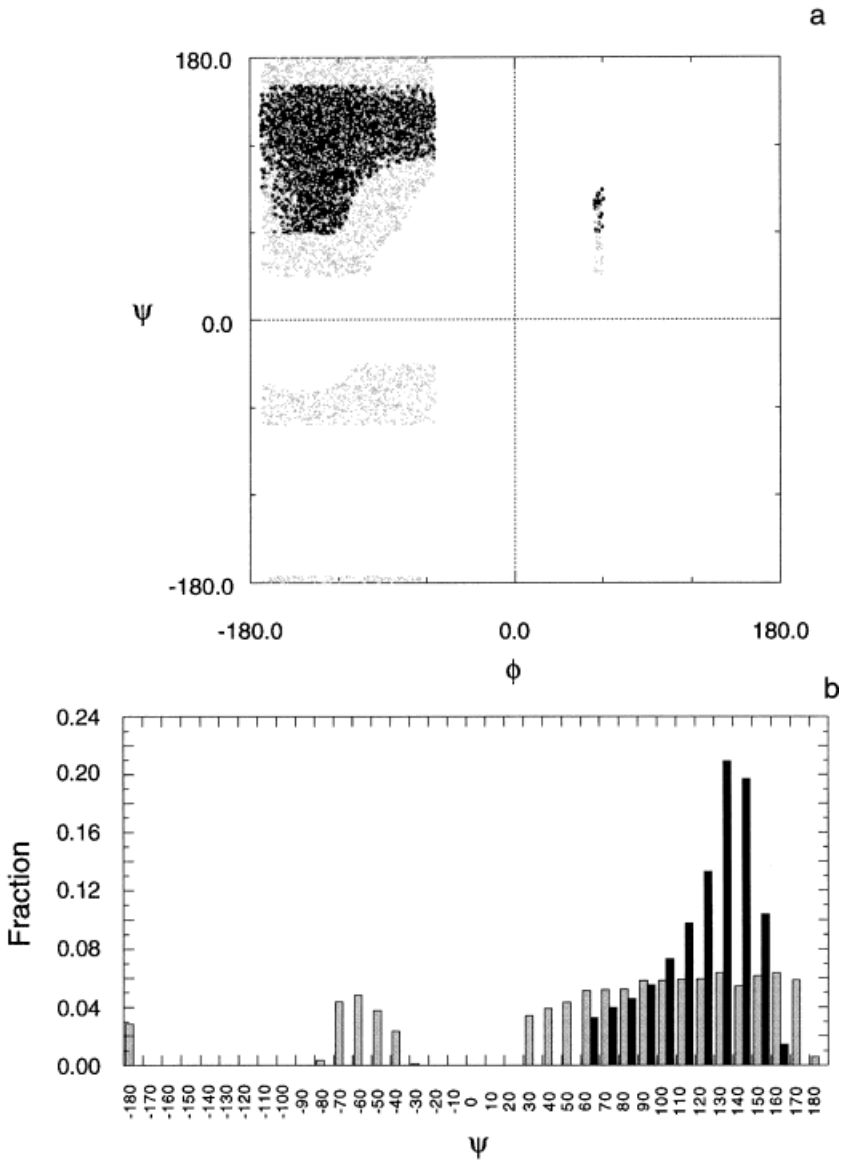

Fig. 5. **Panel a** is a Ramachandran plot of the (ϕ,ψ) dihedrals of Val_5 in the PVP peptide (black dots), along with the corresponding data for Ala_5 in the A9 peptide for comparison (gray dots). **Panel b** shows the distribution of ψ dihedral angles for Val_5 in PVP (black bars) and Ala_5 in A9 (gray bars).

residues, in this case alanines, bounded on each side by proline sequences.

Ala_4 of the A4P4 peptide behaves conformationally the same way as Ala_5 of PAP: it is restricted to the β-region of (ϕ,ψ) space by the following proline (Pro_5). A Ramachandran plot and (ϕ,ψ) distribution histogram plots for Ala_4 (not shown) of A4P4 would be essentially identical to Figures 3 and 4. The ψ distribution (panel a) for Ala_3 of A4P4 is shown in Figure 7 along with the corresponding data for Ala_5 of A9. As can be seen from panel a of this Figure, Ala_3 is not affected by the proline residues two and more residues towards the C-terminus of the A4P4 peptide. Panel b of Figure 7 is a plot of the ψ distribution of Ala_3 when Ala_4 is within the PPII region (ϕ,ψ) = (−75 ± 20°, +145 ± 20°). Comparison with the data for Ala_5 from A9 (also shown in panel b) shows that this residue does not appear to preferentially adopt the PPII structure when Ala_4 is in this conformation.

Ala_5 of the P4A4 peptide would appear to be unaffected by the preceding proline residue. Conformationally it behaves much as Ala_5 of the A9 peptide. This is also true of Ala_6. The restricting effect of proline would thus appear to be limited to the preceding residue, as noted by MacArthur and Thornton[13] and Adzhubei and Sternberg.[11]

The PA2P peptide [Ace-Ala-$(Pro)_2$-$(Ala)_2$-$(Pro)_3$-Ala-NMe] was simulated in order to determine if two alanines bounded on each side by proline residues in the PPII conformation would adopt this conformation. Ala_5 does: its conformational behavior is restricted by Pro_6 and is identical to that of Ala_5 in PAP. Ala_4, however, does not preferentially adopt the PPII

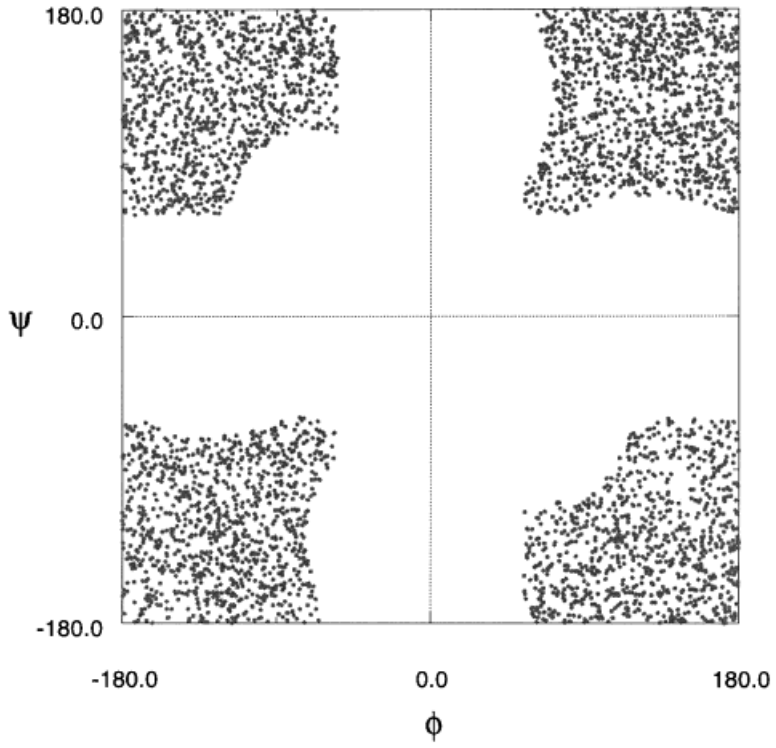

Fig. 6. Ramachandran plot of the (ϕ,ψ) distribution of Gly_5 in the PGP peptide.

conformation and behaves much as Ala_3 of the A4P4 peptide (see Fig. 7).

DISCUSSION

"All-Proline" Peptide

The "all-proline" peptide P9 (which is actually bounded by an alanine at each terminus) is found to adopt only the PPII helical structure in a hard-spheres Monte Carlo computer simulation (Figure 2 and Table II). The backbone dihedrals average (ϕ,ψ) = (−75°, +143 ± 14°) for Pro_2 through Pro_9—the ϕ dihedral of proline is fixed at −75° in this potential. Since the hard sphere potential consists purely of a steric term with no attractive interactions, it can then be concluded that PPII helix formation by an all-proline sequence is in large part driven by steric interactions. This is not entirely surprising: it is well known that a proline residue restricts the conformation space of the preceding residue due to steric interactions.[13] The preceding residue is predominantly confined to the β-region of (ϕ,ψ) dihedral space. Proline residues can normally only occupy two narrow regions of space clustered around (ϕ,ψ) = (−61°, −35°) in the β-region and (ϕ,ψ) = (−65°, +150°) in the β-region.[13] In a polyproline sequence each proline will be restricted to the β-region of space by the following residue, resulting in formation of a PPII helical structure. In the case of polyproline, PPII helix formation is driven by interactions between proline residues adjacent in sequence—longer-range interactions do not play a significant role in PPII helix formation. Polyproline PPII helix formation can thus be said to be a very local event.

Host-Guest Experiments

The proline-rich regions (PRR) involved in various protein-protein interactions are not, of course, pure polyproline.[1] It is of interest then to examine the effects of substituting various non-proline residues into the center of a polyproline host peptide. This was done in the case of the PAP (Ala), PVP (Val), PLP (Leu), PFP (Phe), and PGP (Gly) peptides (guest residues are noted in parentheses). Each guest residue was substituted into the Xaa position of the host Ace-Ala-$(Pro)_3$-Xaa-$(Pro)_3$-Ala-NMe. All of the guest residues, bar glycine, are found to be restricted to the β-region of (ϕ,ψ) space (see Table II and, Figures 3, 4, and 5), as might be expected from the work of MacArthur and Thornton.[13] MacArthur and Thornton note that Pro-Xaa-Pro sequences can adopt a β-turn motif. Such motifs are not observed in these simulations due to the presence of additional flanking proline residues. The glycine guest residue is found to occupy much the same regions of (ϕ,ψ) dihedral space as glycine residues in non-polyproline sequences (Table III and Figure 6).

The proline residue following the guest site, Pro_6, is restricted to the PPII helical conformation. This is to be expected since this proline is followed by two more prolines. The proline preceding the guest residue, Pro_4, however, behaves much like the proline in the peptide APA (see Table III). This is an indication

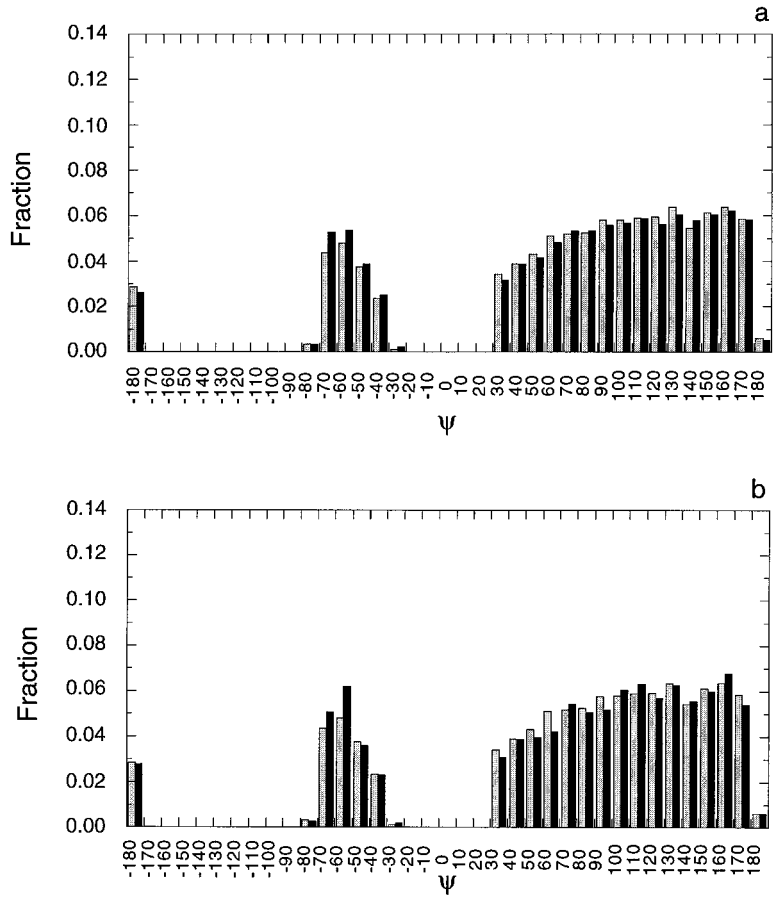

Fig. 7. **Panel a** is a histogram of the distribution of ψ dihedrals for Ala_3 in the A4P4 peptide (black bars). The ψ distribution for Ala_5 in A9 is shown for comparison (gray bars). **Panel b** shows the ψ distribution for Ala_3 in A4P4 when Ala_4 is restricted to the range (ϕ,ψ) = (−75 ± 20°, +145 ± 20°) (black bars). Again, the Ala_5 from A9 ψ distribution is given for comparison (gray bars).

that, at least in the hard sphere potential, the guest residue does not "coerce" the preceding proline into the PPII helical conformation, even though the guest residue itself may be restricted to this conformation. The steric interactions that result in the PPII conformation in these simulations must then involve interactions between the prolyl ring of proline with the preceding residue. Since similar conformational behavior is obtained with all of the guest residues bar glycine, the size and shape of the side chain would appear to be a minor factor. The majority of the interactions then are between the prolyl ring and the backbone and β-carbon of each guest.

PPII Structure Direction And Propagation

It is clear from the work of MacArthur and Thornton[13] and from the host-guest simulations in this work that proline has a profound effect upon the preceding residue. The A4P4 peptide was simulated in order to explore this finding further and also to further study the question of propagation of the PPII helical conformation through non-proline residues. The Ala_4 residue clearly is restricted to the β-region of (ϕ,ψ) dihedral space (Table III). In addition, the ψ dihedral has an enhanced population in the range +110° to +170° (Figure 4). Consequently, Ala_4 can be considered (a more flexible) part of the PPII helix itself. The preceding residue, Ala_3, is not effected by Ala_4's tendency to be in this conformation. This is demonstrated in Figure 7 (panel b): even when Ala_4 is restricted to being within the PPII region of dihedral space, (ϕ,ψ) = (−75 ± 20°, +145 ± 20°), the ψ distribution for Ala_3 is essentially identical to that for an alanine residue in the center of a polyalanine peptide. It appears then, that a PPII helix does not propagate through non-proline residues toward the N-terminus.

The P4A4 peptide was simulated in order to examine the directionality and propagation of PPII helices towards the C-terminus. In this simulation it was found that the alanine immediately following the last proline, Ala_6, showed no discernible preference for the PPII helical region of (ϕ,ψ) dihedral space (Table III). The PPII helical conformation appears to be determined in a C- to N-terminus direction. This does not imply that PPII helices must fold in this direction, just that the presence of a residue in a PPII helical conformation would appear to be determined by the proline residue immediately following it in sequence.

In order to determine whether the PPII conformation could be propagated through two non-proline

residues bounded on each side by proline sequences, the PA2P peptide was simulated. Ala_5, the guest alanine immediately preceding three proline residues, was found to be restricted to the β-region of (ϕ,ψ) space with an increase in the population of the PPII region. This is much the same behavior as that observed for Ala_5 in the PAP peptide (see Figures 3 and 4). Ala_4, the guest alanine preceding Ala_5, displays no discernible preference for the PPII conformation (see Figure 7). The PPII conformation does not then seem to be propagated through more than one non-proline residue in these simulations. This finding calls into question the belief that PRR's adopt the PPII conformation preferentially in solution.

CONCLUSIONS

From the hard sphere computer simulations performed in this work it would appear that PPII helix formation is driven by steric interactions arising between proline residues and the residues that immediately precede them. More specifically, the interactions appear to be between the prolyl ring of each proline and the backbone (including β-carbon) of the preceding residue. In this way, PPII helix formation can then be considered to be a very local folding event. In addition, PPII helix formation has a direction, in that the effect of proline is on the preceding, but not following, residue.

Somewhat troublesome is the indication that PPII helices are not propagated through non-proline residues. This calls into question the widely held belief that proline-rich regions of sequence predominantly adopt this structure.[1] Many such sequences are significantly less than 50% proline (e.g. RNA polymerase II, the PRR of which consists of a repeating YSPTSPS motif) and as such may well be too flexible to significantly populate the PPII conformation. Polar residues capable of hydrogen bonding to the polypeptide backbone (e.g. Ser, Thr, Gln) may help stabilize the PPII conformation and be better propagators than the non-polar residues studied in this work. Surveys of PPII structures in globular proteins of known structure have unearthed some PPII helices that contain no proline.[11] This raises the question of whether occurrences of this conformation in globular proteins arise as a result of the local sequence or more as a result of interactions with the body of the protein. A similar question can be asked of the proline-rich ligands bound by SH3 and WW domains, and by profilin: do these ligands adopt PPII helical conformations when in isolation, or as a result of binding? This indicated lack of propagation through non-proline residues may be due to deficiencies with the hard sphere potential used in this work, and the lack of solvent treatment. These issues will be explored in detail in future work.

ACKNOWLEDGMENTS

The author would like to thank Drs. Ed Lattman, George Rose, and Ben Stapley for their careful reading of the manuscript and helpful comments. This work was funded in part by Kentucky-NSF EPSCoR and the University of Kentucky Medical Center Research Fund.

REFERENCES

1. Williamson, M.P. The structure and function of proline-rich regions in proteins. Biochem. J. 297:249–260, 1994.
2. Siligardi, G., Drake, A.F. The importance of extended conformations and, in particular, the PII conformation for the molecular recognition of peptides. Biopolymers 37:281–292, 1995.
3. Pawson, T. SH2 and SH3 domains in signal transduction. Adv. Cancer Res. 64:87–110, 1994.
4. Pawson, T. Protein modules and signalling networks. Nature 373:573–580, 1995.
5. Schlessinger, J. SH2/SH3 signaling proteins. Curr. Opin. Genet. Dev. 4:25–30, 1994.
6. Chan, D.C., Bedford, M.T., Leder, P. Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. EMBO J. 15:1045–1054, 1996.
7. Sudol, M. The WW module competes with the SH3 domain? TIBS 21:161–163, 1996.
8. Sigler, P.B. Acid blobs and negative noodles. Nature 333: 210–212, 1988.
9. Mahoney, N.M., Janmey, P.A., Almo, S.C. Structure of the profilin-poly-L-proline complex involved in morphogenesis and cytoskeletal regulation. Nat. Struct. Biol. 4:953–960, 1997.
10. Liu, Y.F., Deth, R.C., Devys, D. SH3 domain-dependent association of huntingtin with epidermal growth factor receptor signaling complexes. J. Biol. Chem. 272:8121–8124, 1997.
11. Adzhubei, A.A., Sternberg, M.J.E. Left-handed polyproline II helices commonly occur in globular proteins. J. Mol. Biol. 229:472–493, 1993.
12. Sreerama, N., Woody, R.W. Poly(Pro)II helices in globular proteins: Identification and circular dichroism analysis. Biochemistry 33:10022–10025, 1994.
13. MacArthur, M.W., thornton, J.M. Influence of proline residues on protein conformation. J. Mol. Biol. 218:397–412, 1991.
14. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E. Equation of state calculations by fast computing machines. J. Chem. Phys. 21:1087–1092, 1953.
15. Wimley, W.C., Creamer, T.P., White, S.H. Solvation energies of amino acid side chains and backbone in a family of host-guest pentapeptides. Biochemistry 35:5109–5124, 1996.
16. Creamer, T.P., Srinivasan, R., Rose, G.D. Modeling unfolded states of peptides and proteins. Biochemistry 34: 16245–16250, 1995.
17. Kraulis, P. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. J. App. Cryst. 24:946–950, 1991.