Polymer Bulletin 30, 647–654 (1993)

Polymer Bulletin

Conformational analysis of periodic polypeptides

I. Helical conformations of poly(Gly-Pro)

Masahito Oka* and Akio Nakajima

Department of General Education, Osaka Institute of Technology, Omiya, Asahi-ku, Osaka 535, Japan

SUMMARY

As one of the steps for investigating the conformational properties of the periodic poly(dipeptide)s composed of the Pro residue, theoretical conformational analysis was carried out for poly(Gly-Pro) using ECEPP and the conformational minimization procedure. Calculated results showed that right-handed $\beta^{6.8}$-helix is the most stable helical conformation of poly(Gly-Pro). Obtained conformational preference of poly(Gly-Pro) indicates that poly(dipeptide)s composed of Pro residue can be expected as one of the useful periodic polypeptides for designing the functional polypeptides.

INTRODUCTION

From the viewpoint of molecular design of functional polymers, periodic polypeptides which have the repetitive amino acid sequence are very interesting molecules in the following points. The first points is that many interesting helical conformations, which construct the specific spatial arrangement of functional groups for inducing the useful molecular functions, will be designed by selecting the suitable repetitive amino-acid sequence. It is well known that α-helix and extended conformation are favorable for the polypeptides composed of the alanine-type residues(i.e., all amino acid residues except for the Pro and Gly residues in 20 naturally occurring residues) and the Gly residue. However, other type helices are also known as energetically favorable ones for the polypeptides having specific amino-acid sequence, i.e., β^6-helix[1,2] for poly(Ala-D-Ala) which is a model polypeptide of gramicidin A with the repetitive L- and D-amino-acid sequences, and the triple-stranded helix[3] for poly(Gly-Pro-Pro) which is a model polypeptide of collagen with the repetitive tripeptide sequence containing the Pro residue, and γ-helix[4], β-helix[5] and β-spiral[6] for elastin model polypeptides such as poly(Val-Pro-Gly-Gly) and poly(Val-Pro-Gly-Val-Gly), and a hairpin conformation[7] for model polypeptide of bactenecin(Bac 7). Above multiformity in helical conformations of the polypeptide with the repetitive specific amino-acid sequence support the possibility of constructing the various useful helical structures by designing the amino-acid sequences. Moreover, periodic proteins composed of the specific repetitive amino-acid sequence have been frequently found, ex., bactenecin(Bac 5)[8], RNA polymerase II[9], circumsporozoite protein[10], ice nucleation protein[11], proline-rich precursor protein of the sheath from microfilariae[12], octopus rhodopsin[13], and salivary basic proline-rich protein[14]. These results suggest that helical structures constructed by the repetitive specific amino-acid sequence play important roles in the biological phenomenon in nature, that is, periodic polypeptides are very important molecules for design the functional molecules with biomimetic properties.

*Corresponding author

The second point is that it is possible for polypeptides to form the specific helical conformations in the non-crystal states such as solution or lipid bilayer membranes, in addition to the crystal state. For general synthetic polymers, it is difficult to form the helical conformations in the non-crystal states. They usually take random-coil conformations in such states. Above different properties between polypeptides and synthetic polymers were caused by the following difference in the characters of backbone of them. Polypeptides have the peptide bond on the backbone, then they form helical conformations supporting by the favorable intramolecular hydrogen-bonds. However, general synthetic polymers usually do not have such functional groups interacting strongly in backbone. This character of polypeptides suggests that polypeptides are more appropriate molecules than general synthetic polymers in the meaning that the phase state of polypeptides is not restricted to the solid state for constructing the functional structures and for exhibiting the molecular functions.

The third point is the conformational responsibility of polypeptides for the external conditions, i.e., their conformations are transformed by the change of temperature[15], pH[16], and dielectric constant of solvent[17], and by irradiation of ultraviolet or visible rays[18]. For the case of homopolypeptides and copolypeptides with non-periodicity, conformational states are restricted to the several ones such as right- or left-handed α-helices and, extended and random-coiled conformations. However, for the case of periodic polypeptides, conformational states can be unristrically selected by designing the suitable amino-acid sequence as already mentioned above. It means that periodic polypeptides have desirable molecular properties such as the multiple and sensitive responsibility for the change of external conditions.

The amino-acid sequences composed of the repetitive Xaa-Pro sequence were found in several native proteins and peptides, ex., bactenecin[8], myosin L1 and L4 of chicken, outer membrane protein A precursor of *Escherichia coli*, early 32K, 26K and 13K proteins of adenovirus 2 and 5, gene *tonB* protein of *E.coli*[19], β-casein[20], and β-casomorphin[21] and it is already known that some of them have biological functions such as antimicrobial or opioid activity, suggesting that polypeptides having the repetitive Xaa-Pro sequence are expected as a new material exhibiting interesting molecular functions. So it is very important to known what kinds of helical conformations are energetically favorable for the polypeptides composed of the repetitive Xaa-Pro sequence. As Gly residue has the most favorable conformational flexibility in (ϕ, ψ)-space[22,23] among 20 naturally occurring amino-acid residues, it is reasonable to select the Gly-Pro sequence for investigate the conformational characters in the $(\phi_{Xaa}, \psi_{Xaa}, \phi_{Pro}, \psi_{Pro})$-space of the periodic polypeptides, poly(Xaa-Pro). In this work, as one of the step for searching interesting helical structures for molecular design of functional polypeptides, theoretical conformational analysis based on the molecular mechanics was tried for the periodic polypeptides with the repetitive Gly-Pro sequences, i.e., poly(Gly-Pro).

THEORETICAL

All conformational energy calculations were carried out with the energy functions of ECEPP[24]. During minimizations, all (ϕ, ψ) of Gly and ϕ of Pro were allowed to vary. All other backbone dihedral angles were fixed to 180° except for ϕ_{Pro}=-75°. Optimization of the helical structure was tried by two methods, i.e., the three-steps method and the grid method, as already tried in the previous optimization of poly(Ala-D-Ala)[2] and poly(Ala-Gly)[25]. The three-steps method was carried out by the following three steps of minimizations. The first step was the minimization for the dipeptide, Ac-Gly-Pro-NHMe. All combinations of the single residue minima[23] of Gly and Pro residues(i.e., 9 and 4, respectively) were used as starting conformations. The second step was the minimization for the tetrapeptides having two repetitive sequences of Gly-Pro, i.e., Ac-(Gly-Pro)$_2$-NHMe.

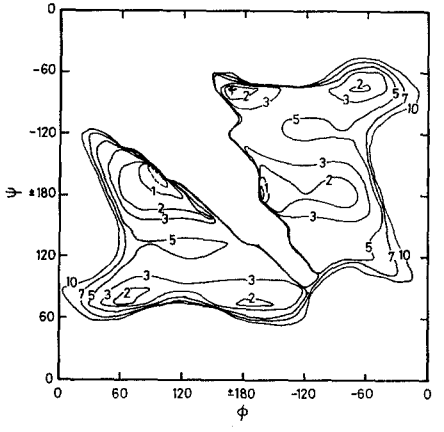
All minima in the first step were used as starting conformations of the second step. The third step was the minimization for the tetraeicosapeptide having twelve repetitive sequences of Gly-Pro, i.e., Ac-(Gly-Pro)$_{12}$-NHMe(abbreviated as poly(Gly-Pro)). All minima found in the second step were also used as starting conformations of the third step. During the 2nd and 3rd minimization step, the condition of helical conformation was used in a similar manner as the previous optimization works[2,4,25]. The grid method was tried as shown in the previous works[2,25]. Conformational energy of poly(Gly-Pro) was calculated by changing ϕ_{Gly} and ψ_{Gly} at 15° intervals and fixing ϕ_{Pro} to the energy minima of Ac-Pro-NHMe(i.e., ϕ_{Pro} =-48° , 79° , 159°). Then all local minima found in (ϕ_{Gly}, ψ_{Gly}, ψ_{Pro}) space were used as starting conformations of minimization of poly(Gly-Pro) in this method.

A bend (occurring at *i+1* and *i+2* th residues) is defined as a conformation in which $R \leqq 7$ Å (R is the distance between ith C^{α} and $i+3$ th C^{α} atoms.) and also classified into eleven types given in Table I of ref 26. A polar hydrogen atom and an oxygen or nitrogen atom with an interatomic distance of less than 2.3 Å are regarded to be hydrogen-bonded. Conformational space is divided into 16 regions with the conformational letter codes shown in Figure 1 of ref 23. The conformational energy per whole molecule, ΔE is defined by $\Delta E=E-E_0$, where E_0 is the value of E at the global minimum on the potential energy surface of the particular molecules, and ΔE_{res} is defined by $\Delta E_{res}=\Delta E / m$, where m is number of residues of the molecule. β^X-helix is defined as a helix which has a spiral structure with X residue per turn. γ-Helix is defined as helical conformation which cannot be categorized to the α-helix and β^X-helix. This definition of γ-helix is extended from the previous one which was used for naming a new helical conformation found as the lowest-energy conformation of poly(Val-Pro-Gly-Gly) composed with six repetitive amino-acid sequences, Val-Pro-Gly-Gly[4]. Two helical parameters, n and h, are the number of residues per turn and rise per residue, respectively. All molecular diagrams are drown by the molecular graphic program PEPCON[27,28].

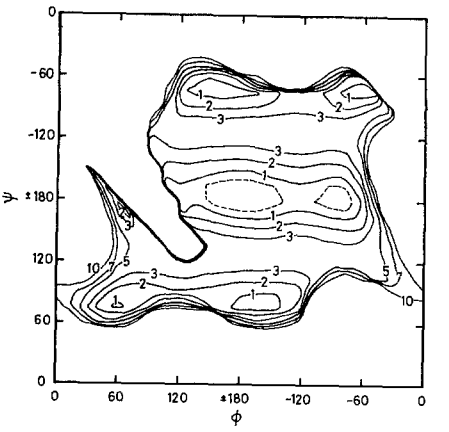
RESULTS

Local Minima in (ϕ_{Gly}, ψ_{Gly}) Maps of Poly(Gly-Pro) for the Specified Backbone Conformation of Pro Residue

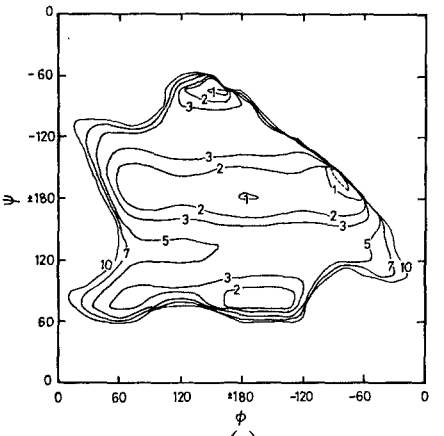
Three (ϕ, ψ) maps of Gly residue of poly(Gly-Pro) with the specified value of(ϕ_{Pro}, ψ_{Pro})=(-75° ,79°), (-75° ,159°), and (-75° ,-48°) under the condition of helical conformation are shown in Figures 1a, 1b, and 1c, respectively. Contour lines of Gly residue are represented by the energy difference from the lowest-energy in each (ϕ, ψ) map, and energy difference is also designated as the value per residue. Stable conformational regions in (ϕ_{Gly}, ψ_{Gly}) maps are restricted to more narrow regions than those of (ϕ_{Gly}, ψ_{Gly}) maps of Ac-Gly-Pro-NHMe whose (ϕ_{Pro}, ψ_{Pro}) are specified to the corresponding values of poly(Gly-Pro). For the case of (ϕ_{Pro}, ψ_{Pro})=(-75° ,79°), the region around (ϕ_{Gly}, ψ_{Gly}) =(120° ,-120°) is destabilized by the interactions within the two repetitive units, and the region around (ϕ_{Gly}, ψ_{Gly}) =(180° ,180°) is destabilized by the interactions within the three repetitive units. The energetically unfavorable regions are extended, and the potential valleys around the local minima are steepened with an increase in number of the repetitive units. However, the energetically favorable regions in (ϕ_{Gly}, ψ_{Gly}) space of Ac-(Gly-Pro)$_6$-NHMe almost correspond to those of Ac-(Gly-Pro)$_{12}$-NHMe. It means that the global character of (ϕ_{Gly}, ψ_{Gly}) space is decided by the medium-range interactions. Similar characters in (ϕ_{Gly}, ψ_{Gly}) space are also shown for the specified values of (ϕ_{Pro}, ψ_{Pro})=(-75° ,159°) and (-75° ,-48°). For the former case, the regions from (ϕ_{Gly}, ψ_{Gly}) =(60



(a)



(b)



(c)

Figure 1.

Energy contour (ϕ , ψ)maps of the Gly residue of poly(Gly-Pro) for the specified (ϕ , ψ) of the Pro residue;
(a) (-75° ,79°), (b) (-75° ,159°),
(c) (-75° ,-48°).
The contour lines are labeled as energy in kcal/mol per residue above the minimum energy point of each energy map. The dashed lines indicate the 0.5 kcal/mole energy contour lines.

° ,-120°) to (100° ,-210°) are destabilized by the interactions within four repetitive units, and the latter case, the regions around (ϕ_{Gly}, ψ_{Gly}) =(-60° ,-60°) are destabilized by the interactions within two repetitive units.

Stable Helical Conformations of Poly(Gly-Pro)

A total of 20 energy minima of poly(Gly-Pro) was found in $\Delta E_{res} < 3$ kcal mole^{-1}, and 10 of them are shown in Table I. The lowest-energy conformation is a right-handed β [6,8]-helix(EC conformation) with (ϕ_{Gly}, ψ_{Gly}, ϕ_{Pro}, ψ_{Pro})=(-166° ,-175° ,-75° ,77 °)(Figure 2). This helix has the hydrogen bond, (Gly_i)NH···OC(Gly_i) and (Gly_i)CO··· HN(Gly_{i+1}) and takes non-bend structure at both of Gly-Pro and Pro-Gly portions. The backbone dihedral-angles of this helix almost corresponds to those of the lowest-energy minima of Ac-Gly-Pro-NHMe, (178° ,-175° ,-75° ,80°), indicating that the short-range interaction is essentially important for stabilizing the helical structure of poly(Gly-Pro). However, as shown in the (ϕ_{Gly}, ψ_{Gly}) maps of poly(Gly-Pro) for the specified (ϕ_{Pro}, ψ_{Pro})=(-75° ,79°) (Figure 1a), the region around (ϕ_{Gly}, ψ_{Gly})=(180° ,180°), which is the region containing global minimum of Ac-Gly-Pro-NHMe, is significantly destabilized by

Table I. Calculated Minimum Energy Conformations[a] of Poly(Gly-Pro)

Conformational Letter Code	ΔE_{res}[b] (kcal mole^{-1})	Helix[c] Type	h[d]	ϕ Gly	ϕ Gly	ϕ Pro
E C	0.00	$\beta^{6.8}$(R)	0.69	-166	-175	77
E*C	0.45	$\beta^{5.6}$(L)	0.89	134	-173	71
F*C	0.83	$\beta^{5.2}$(L)	1.05	109	-162	76
E*C	0.86	$\beta^{6.1}$(L)	0.82	156	172	68
FA	1.49	$\beta^{5.8}$(L)	0.92	-109	-157	-25
FA	1.65	$\beta^{6.3}$(L)	0.95	-84	-168	-37
D*A	1.69	$\beta^{5.4}$(L)	1.34	174	-72	-52
D*C	1.90	$\beta^{4.3}$(R)	1.70	164	-73	78
A*C	2.56	γ	2.31	68	81	79
AF	2.63	$\beta^{5.9}$(L)	1.40	-55	-71	176

[a]All 10 minimum-energy conformnations with $\Delta E_{res} < 2.63$ kcal mol^{-1}.

[b]E_o=-168.90 kcal mol^{-1}, $\Delta E_{res}=(E-E_o)/24$.

[c]Helix sense is abbreviated as R or L for right- or left-handded, respectively.

[d]Rise per residue.

the medium-range interactions. That is, the value of ϕ Gly of poly(Gly-Pro) is slightly shifted 16° from the local minima of Ac-Gly-Pro-NHMe by the favorable medium-range interactions in poly(Gly-Pro), indicating the medium-range interaction is also important for specifying the precise helical structure of poly(Gly-Pro). The 2nd, 3rd and 4th low-energy conformations are left-handed $\beta^{5.6}$, $\beta^{5.2}$and $\beta^{6.1}$-helices with ΔE_{res}=0.45, 0.83, and 0.86 kcal mole^{-1}, respectively. These three energy minima situate in a common ravine located in the E* and F* regions of Figure 1a, and one of them can be transformed to other two local minima passing through the low energy pass whose energy height is less than 1 kcal mole^{-1} per residue. These three helices have a common conformational preference that they takes compact structure, i.e., bend or quasi-bend structure at Gly-Pro and Pro-Gly portions, and also have the hydrogen bond (Gly$_i$)CO···HN(Gly$_{i+1}$). It means that they seem to behave as a single helical conformation in solution under thermal equibrium. The energy barriers between the right-handed $\beta^{6.8}$-helix with the lowest-energy and the above 2nd, 3rd and 4th low-energy left-handed helices is almost 5 kcal mole^{-1} height per residue. It also means that the lowest-energy right-handed $\beta^{6.8}$-helix is isolated from the above three left-handed helices in the conformational space. All local minimum conformations besides above four low-energy ones are energetically unstable as shown Table I, i.e., the 5th low-energy conformations has ΔE_{res}=1.49 kcal mole^{-1}. The lowest-energy β^4-helix and γ-helix are found as the 8th and 9th low-energy ones with ΔE_{res}=1.90 and 2.56 kcal mole^{-1}, respectively. No conformations corresponding to the α-helix were found as a local energy minima. These results indicate that the β^6-type and β^5-type helices are favorable conformations for poly(Gly-Pro) even though they present the distribution in the values of helical parameters n and h.

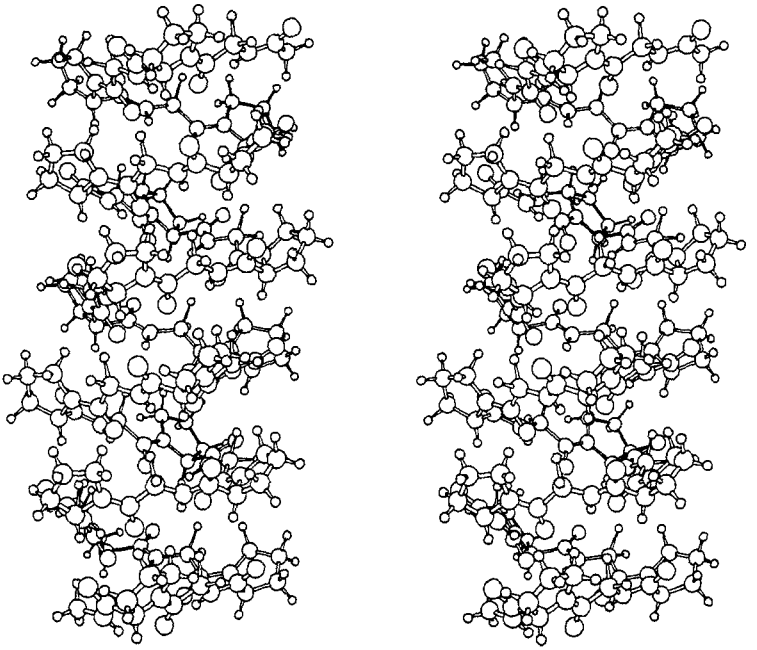


Figure 2. A right-handed $\beta^{6.8}$-helix(EC conformation) of Poly(Gly-Pro) with the lowest-energy.

DISCUSSIONS

The conformational preference of poly(Gly-Pro) are significantly different from those of poly(Ala-D-Ala) and poly(Ala-Gly) obtained by the optimization with the energy function of ECEPP[2,25]. A right- or left-handed α-helix is a energetically favorable conformation for poly(Ala-D-Ala)and poly(Ala-Gly), i.e., a right-handed α-helix is the lowest-energy conformation for poly(Ala-Gly), and the left- and right-handed α-helices are also found as the 5th and 6th low-energy conformations with ΔE_{res}=0.23 and 0.24 kcal mole^{-1}, respectively. However, no α-helical conformations are found as the local minimum-energy ones for poly(Gly-Pro). For the case of Ac-Gly-Pro-NHMe, the local minima with (ϕ_{Gly}, ψ_{Gly}, ϕ_{Pro}, ψ_{Pro})=(-62° ,-69° ,-75° -22°), which has the closest dihedral angles to those of the α-helical region in the (ϕ_{Gly}, ψ_{Gly}, ϕ_{Pro}, ψ_{Pro}) space, is found at ΔE_{res}=1.43 kcal mole^{-1}. However, by considering the short-range interactions within the tetrapeptide Ac-(Gly-Pro)$_2$-NHMe, this local minimum is shifted to the other local minimum with (ϕ_{Gly}, ψ_{Gly}, ϕ_{Pro}, ψ_{Pro})=(-51° ,-74° ,-75° ,4°) which is the outside of the α-helical region in the (ϕ_{Gly}, ψ_{Gly}, ϕ_{Pro}, ψ_{Pro}) space and ΔE_{res} increases to 3.94 kcal mole^{-1}. These results indicate
that α-helix in poly(Gly-Pro) is destabilized by the short-range interactions within tetrapeptide. This conformation is further destabilized by the medium-range interactions within the tetraeicosapeptide Ac-(Gly-Pro)$_{12}$-NHMe to the local minima with ΔE_{res}=5.94 kcal mole^{-1}.

Mattice and Manderkern[29] measured the CD spectra of poly(Gly-Pro) at various temperatures in water, ethylene glycol-water(2:1,v/v), and trifluoroethanol. Poly(Gly-Pro) exhibits a negative band of 201 nm at -48℃ in ethylene glycol-water(2:1,v/v), and raising the solution temperature from -48 to 50℃ reduced the magnitude of the 200-201 band and also caused the red-shift of the negative CD band from 201 to 204nm. The thermal transition of the CD spectra of poly(Gly-Pro) resemble those of certain all-β globular proteins such as soybean trypsin inhibitor and elastase which exhibit the different CD spectra from the typical all-β proteins[30]. That is, these two proteins have a negative CD band around 200 nm, and it drastically decreases in magnitude with increasing temperature around the melting temperature[30]. By the comparison with the CD spectra presented by Mattice et al[29] and Wu et al[30], it is suggested that poly(Gly-Pro) takes certain ordered conformations which are different from the well known conformations such as α-helix, and extended and random-coiled conformations, at low temperature in ethylene glycol-water(2:1,v/v). The CD spectra of poly(Gly-Pro) in trifluoroethanol don't indicate so significant temperature dependency, and they are similar to those at high temperature in ethylene glycol-water(2:1,v/v), suggesting that poly(Gly-Pro) takes disordered conformations in trifluoroethanol. These results indicate that poly(Gly-Pro)takes particular ordered conformations depending on the external conditions such as solvents and temperature. However, it is difficult to specify the type of ordered conformation in the relations of the calculated helical conformations presented in Table I, and also to show whether the poly(Gly-Pro) takes a single conformation or in conformational equilibrium of several low-energy conformations. Further experimental information are desired for make clear the relations between the repetitive amino-acid sequences and stable conformations.

Optimized results indicate that β^6- and β^5-helices are stable for poly(Gly-Pro). These helical characters are significantly different from those of periodic polypeptides composed of repetitive dipeptide sequence without the Pro residue. It means that the Pro residue has a very important role for designing the functional helical structure which would be useful as the basic backbone conformation in designing functional polypeptides.

REFERENCES

1. D.W.Urry, *Proc.Nat.Acad.Sci.USA*,**69**,1610(1972).
2. M.Oka, Y.Baba, A.Kagemoto, and A.Nakajima, *Polym.J.*,**22**,185(1990).
3. W.H.Miller and H.A.Scheraga, *J.Polym.Sci.Polym.Symp.*,**54**,171(1976).
4. M.Oka, Y.Baba, A.Kagemoto, and A.Nakajima, *Polym.J.*,**22**,555(1990).
5. M.Oka and A.Nakajima, *Peptide Chem.*,**1992**,(1993)., in press.
6. C.M.Venkatachalam and D.W.Urry, *Macromolecules*,**14**,1225(1981).
7. M.Oka, Y.Baba, A.Kagemoto, and A.Nakajima, *Pepetide Chem.*,**1991**,99(1992).
8. R.W.Frank, R.Gennar, K.Schneider, M.Przybyski, and D.Romeo, *J.Biol.Chem.*, **265**,18871(1990).
9. L.A.Allison, J.K.Wang, V.D.Fitzpatrick, M.Moyle, and C.J.Ingles, *Mol.Cell.Biol.*,**8**,311(1988).
10. D.J.Kemp, R.C.Coppel, and R.F.Anders, *Ann.Rev.Microbiol*,**41**,181(1987).
11. G.Warren and C.Corotto, *Gene*,**85**,239(1989).
12. M.E.Selkirk, M.Yazdanbakhsh, D.Freedman, M.L.Blaxter, E.Cookson, R.E.Jenkins, and S.A.Williams, *J.Biol.Chem.*,**266**,11002(1991).
13. Yu.A.Ovchinnikov, N.G.Abdulaev, A.S.Zolotarev, I.D.Artamonov, I.A.Bespalov, A.E.Dergachev, and M.Tsuda, *Eurp.J.Biochem.*,**232**,69(1988).
14. A.M.Castle, L.E.Stahl, and J.D.Castle, *J.Biol.Chem.***267**,13093(1992).

15. P.Doty, and J.T.Yang, *J.Am.Chem.Soc.*,**78**,498(1956).
16. P.Doty, A.Wada, J.T.Yang, and E.R.Blout, *J.Polym.Sci.*,**23**,851(1957)..
17. G.D.Fasman, M.Idelson, and E.R.Bloutt, *J.Am.Chem.Soc.*,**83**,709(1961).
18. A.Ueno, K.Takahashi, J.Anzai, and T.Osa, *J.Am.Chem.Soc.*,**103**,6410(1981).
19. W.C.Baker, L.T.Hunt, B.C.Orcutt, D.G.George, L.S.Yeh, H.R.Chen, M.C.Blomquist, G.C.Johnson, E.i.Seibel-Ross, and M.O.Dayhoff, Protein Sequence Database(National Biomedical Research Foundation, Washington, DC).
20. B.Ribadeau Dumas, G.Brignon, F.Grosclaude, and J.-C.Mercier,*Eur.J.Biochem.*, **25**,505(1972).
21. R.Greenberg, M.L.Groves, and H.J.Dover, *J.Biol.Chem.*,**259**,5132(1984).
22. D.A.Brant, W.G.Miller, and P.J.Flory, *J.Mol.Biol.*,**23**,47(1967).
23. S.S.Zimmerman, M.S.Pottle, G.Nemethy, and H.A.Scheraga, *Macromolecules*,**10**,1(1977).
24. F.A.Momany, R.F.McGuire, A.W.Burgess and H.A.Scheraga, *J.Phys.Chem.*,**79**,2361(1975).
25. M.Oka, Y.Baba,A.Kagemoto, and A.Nakajima, *Polym.J.*,**22**,416(1990).
26. S.S.Zimmerman and H.A.Scheraga, *Biopolymers*,**16**,811(1977).
27. M.Sisido, private communications.
28. Y.Beppu, *Computers Chem.*,**13**,101(1989).
29. W.L.Mattice and L.Mandelkern, *Biochemistry*,**10**,1926(1971).
30. J.Wu, J.T.Yang, and C.C.Wu, *Anal.Biochem.*,**200**,359(1992).

Accepted May 11,1993 S