  Bull. Chem. Soc. Jpn., 59, 2445—2449 (1986)

Infrared Absorption Study of Human Proinsulin C-Peptide Fragments in Dichloromethane[1)]

Mitsuaki NARITA,* Toshihiko OGURA, Kazuhiro SATO, and Shinya HONDA
Department of Industrial Chemistry, Faculty of Technology, Tokyo University of Agriculture and Technology, Koganei 184
(Received March 20, 1986)

In connection with the relationship between the conformation and solubility of peptide intermediates having polar side chains, IR spectroscopic conformational analysis of human proinsulin C-peptide fragments was performed in dichloromethane. The polar amino acid residues involved in the peptide fragments are Glu(OBzl), Asp(OBzl), Gln, and Ser(Bzl). Especially, the N-H stretching absorption spectra have been explored over a wide range of concentration. Based on the concentration dependence of the N-H stretching bands, it has been shown that the characteristic N-H stretching band around 3330 cm^{-1} is due to the intramolecular hydrogen bond. Intermolecular hydrogen bonding also occurs to a small extent in these peptides, giving rise to a slight concentration dependence of the N-H stretching bands. Conformational behaviors of oligopeptides having protected polar side chains of the Glu(OBzl) and Asp(OBzl) residues are just the same as those of the corresponding homooligo(Leu)s, indicating that the protected polar side chains do not disturb conformational behaviors of the oligopeptides. Conformational analysis of the peptide in dichloromethane, which contains the Pro residue and has a high potential for the formation of an α-helical structure, distinctly suggests that the Pro residue near the N-terminal in the peptide has the great ability to promote successive intramolecular hydrogen bonds, probably corresponding to the α-helical structure of the peptide.

In connection with the design of the synthetic route for peptides and proteins based on the solubility prediction method,[2-4)] it is quite interesting to investigate conformational behaviors of peptide intermediates in solution since the investigation offers useful information for the solubility prediction of peptide intermediates. In the previous paper,[4)] we reported IR spectroscopic conformational analysis in the solid state of human proinsulin C-peptide fragments. Here, we report their IR spectroscopic conformational analysis in dichloromethane, which is commonly used in conventional solid-phase peptide synthesis.[5,6)] Especially, in this paper, we give attention to the role of the Pro residue in solubility improvement of peptide fragments containing the Pro residue.

The Pro residue is well recognized to a strong helix and β-sheet breaker and plays an important role in stopping the development of helix and β-sheet structures in peptides and proteins.[7-10)] In fact, the insertion of the Pro residue into oligo(Leu)s caused "the peptide segment separation" in the solid state and achieved remarkable solubility improvement of peptides in a variety of organic solvents.[11-13)] The effect of the Pro residue on increasing solubility of peptide intermediates was also confirmed in human proinsulin C-peptide fragments.[2)]

On the other hand, the values of the backbone dihedral angles ϕ and ψ of the Pro residue are severely restricted to be $\phi=-60°\pm10°$, $\psi=-30°\pm20°$, and $\phi=-60°\pm10°$, $\psi=-120°\pm20°$ due to the low flexibility of the pyrrolidine ring in the Pro residue.[14)] In our recent papers,[15,16)] we demonstrated the great ability of the Aib residue to promote helical folding in peptides due to the short-range interactions by van der Waals repulsions and subsequently proposed that the restriction of the dihedral angles was one of important initiation mechanisms of helical folding in the natural proteins. One of our purposes in this paper is also to investigate the ability of the Pro residue to promote helical folding in oligopeptides in solution.

Experimental

Materials. Samples of human proinsulin C-peptide fragments are those prepared in the previous papers.[2)] The purity of the peptides was confirmed by the amino acid and elemental analyses. The peptides also gave a single peak on HPLC. The peptide fragments used in this study are in the following: Boc-Glu(OBzl)-Ala-Glu(OBzl)-Asp(OBzl)-Leu-OPac (**1**), Boc-Leu-Ala-Leu-Glu(OBzl)-Gly-OPac (**2**), Boc-Gly-Gly-Pro-Gly-Ala-Gly-OPac (**3**), Boc-Ser(Bzl)-Leu-Gln-Pro-Leu-Ala-Leu-Glu(OBzl)-Gly-OPac (**4**), and Boc-Ser(Bzl)-Leu-Gln-OPac (**5**).

IR Measurements. The IR absorption spectra of the samples in dichloromethane were recorded at room temperature with a JEOL Model JIR-100 FT-IR spectrometer by employing 0.5 mm- and 5 mm-path length cells with potassium bromide windows.

Results

IR Absorption Study of the Pentapeptide Fragments 1 and 2 in Dichloromethane. Conformational analysis of peptide fragments having polar side chains is quite interesting in order to elucidate the role of protected polar amino acid residues in conformational behaviors of peptide fragments in dichloromethane. As reported in the previous paper,[2)] the pentapeptides **1** and **2** are easily soluble in dichloromethane, and their IR absorption spectra are easily obtained over a wide range of concentration (1.0×10^{-4}—10^{-2} M†). In Fig. 1, we show the N-H stretching bands of both peptides in dichloromethane and

† 1 M=1 mol dm^{-3}.

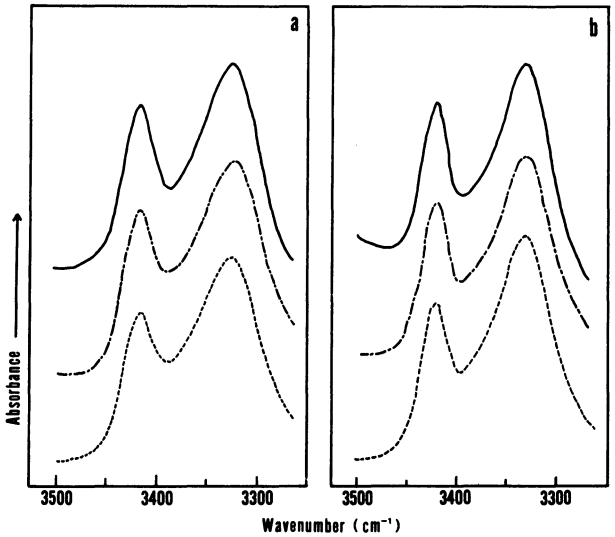

Fig. 1. Concentration dependence of the N-H stretching bands in the IR absorption spectra of the peptides **1** and **2** in dichloromethane. -----: 1.0×10^{-2} M, —·—·—: 1.0×10^{-3} M, ——: 1.0×10^{-4} M. a: The peptide **1**; b: the peptide **2**.

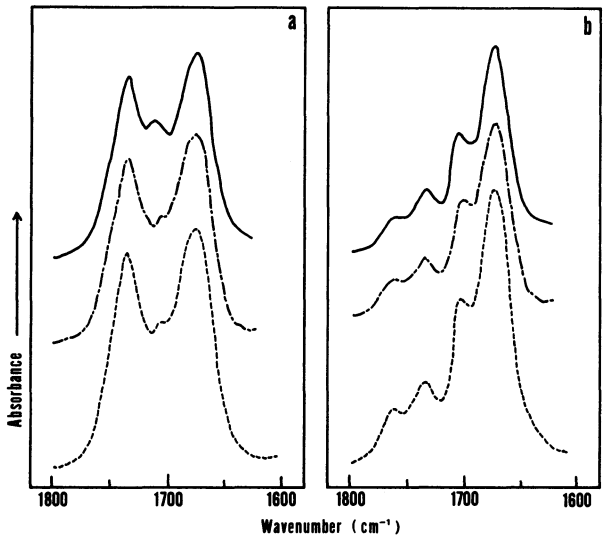

Fig. 2. Concentration dependence of the IR absorption bands in the amide I region of the peptides **1** and **2** in dichloromethane. -----: 1.0×10^{-2} M, —·—·—: 1.0×10^{-3} M, ——: 1.0×10^{-4} M. a: The peptide **1**; b: the peptide **2**.

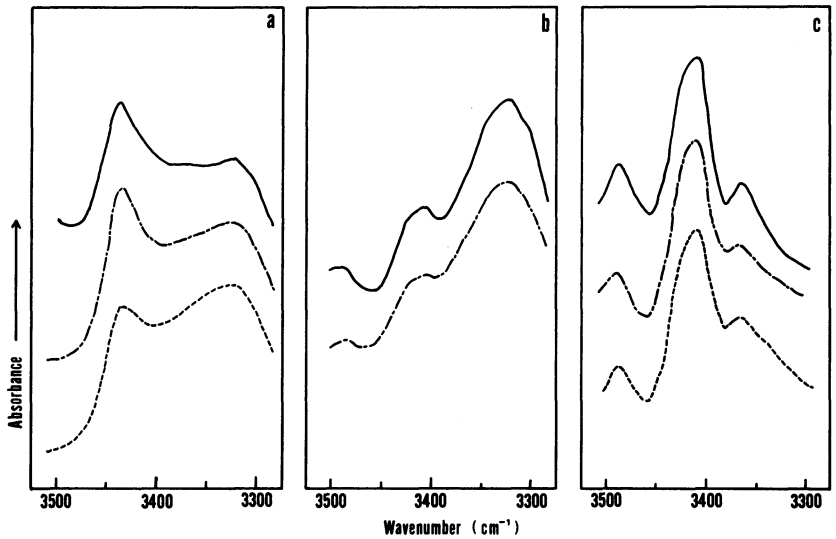

Fig. 3. Concentration dependence of the N-H stretching bands in the IR absorption spectra of the peptides **3**—**5** in dichloromethane. -----: 1.0×10^{-2} M, —·—·—: 1.0×10^{-3} M, ——: 1.0×10^{-4} M. a: The peptide **3**; b: the peptide **4**; c: the peptide **5**.

clearly see two distinct bands at 3420 and 3330 cm^{-1} over the entire concentration range. The 3420 cm^{-1} band is clearly due to the free N-H stretching vibration, while the 3330 cm^{-1} band is due to a hydrogen-bonded species.[17-19] Since the intensities of these two bands show little dependence on concentrations, the 3330 cm^{-1} band is attributed mainly to the intramolecular hydrogen-bonded species, accompanied by a slight contribution from intermolecular hydrogen bonding. The IR absorption spectra in the amide I region of the peptides **1** and **2** in dichloromethane are also illustrated in Fig. 2. The IR spectra of both peptides measured over the wide range of concentrations show the bands around 1760, 1733, 1706, and 1675 cm^{-1}, corresponding to the Pac ester carbonyl, Bzl ester carbonyl, Boc urethane carbonyl, and amide carbonyl groups, respectively. The Pac ketone carbonyl group is estimated to have the absorption band around 1685 cm^{-1}, overlapped with the amide carbonyl group. The amide carbonyl band also showed little dependence on concentration (1.0×10^{-4}—10^{-2} M).

IR Absorption Study of the Hexapeptide 3, Nonapeptide 4, and Tripeptide 5 in Dichloromethane. The IR absorption spectra in the N-H stretching region of the peptides **3**—**5** in dichloromethane (1.0×10^{-4}—10^{-2} M) are shown in Fig. 3. We also see two

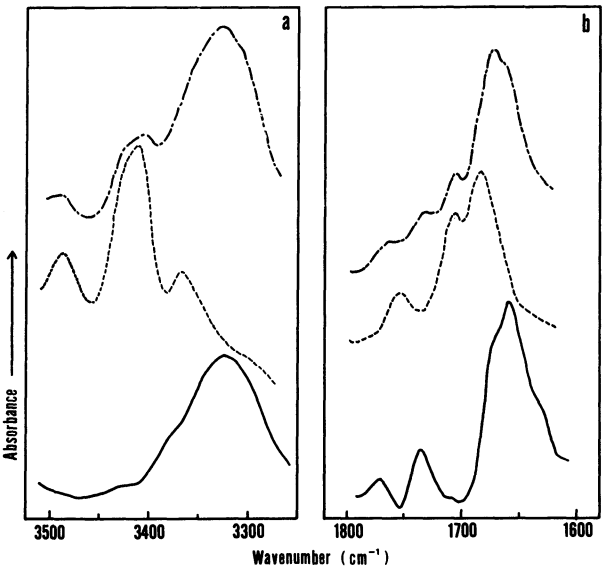

Fig. 4. IR absorption spectra in the amide A and amide I regions of the peptides **4** and **5** in dichloromethane (1.0×10^{-4} M). —·—·—: The peptide **4**, ------: the peptide **5**, ——: IR difference spectra for the peptides **4** and **5**. a: The amide A region; b; the amide I region.

distinct bands at 3435—3425 cm^{-1} and around 3370 or 3330 cm^{-1} over the entire concentration range. The peptides **4** and **5** have the additional absorption band at 3487 cm^{-1} corresponding to the side-chain amide group, which is free from hydrogen bonding. Again, the 3430 cm^{-1} band is due to the free N-H stretching vibration, while the bands around 3370—3330 cm^{-1} are due to hydrogen-bonded species. The intensities of the hydrogen-bonded band of the peptide **3** (Fig. 3a) are more dependent on concentration than the intensities of the band at 3330 cm^{-1} of the peptides **1**, **2**, **4**, and **5** (Figs. 1, 3b, and 3c), indicating that the peptide **3** is accompanied by more contribution from intermolecular hydrogen bonding than the peptides **1**, **2**, **4**, and **5**. Figures 4a and 4b show the IR absorption spectra in the amide A and amide I regions of the peptides **4** and **5** in dichloromethane (1.0×10^{-4} M). The amide carbonyl band also showed little dependence on concentration (1.0×10^{-4}—10^{-2} M) (not shown). Assignment of each band in the amide I region of the peptides **4** and **5** is essentially the same with that of the peptides **1** and **2**. In Figs. 4a and 4b, the IR difference spectra for both peptides are also represented. The spectra indicate the disappearance of the free N-H stretching vibration and clear-cut features of absorption bands at 3323 $^{-1}$ and 1658 cm^{-1}, probably assigned to the α-helical structure.[17-19]

Discussion

Contrary to the fact that the pentapeptides **1** and **2** in the solid state have the β-sheet structure,[4] in dichloromethane, the clear conformational transformation from intermolecular to intramolecular hydrogen bonding takes place in both peptides. The N-H stretching absorption band at 3330 cm^{-1} of the peptides **1** and **2** shows little dependence on concentration, indicating a slight contribution from inter molecular hydrogen bonds, which are apparently different from the intermolecular hydrogen bonds in the solid state assigned by the band at 3270 cm^{-1} (β-sheet). Their conformational behaviors both in the solid state and in dichloromethane are just the same as those of Boc-Leu_4-OBzl and Boc-Leu_5-OBzl,[20] indicating that the protected polar side chains of the Glu(OBzl) and Asp(OBzl) residues do not disturb the conformational behaviors of the pentapeptides **1** and **2** both in the solid state and in dichloromethane. Since the IR difference spectrum in dichloromethane for Boc-Leu_5-OBzl and Boc-Leu_3-OBzl suggests that the conformation formed by the intramolecular hydrogen bonds of Boc-Leu_5-OBzl appears to be a helical structure in dichloromethane,[20] the conformation of the peptides **1** and **2** is assumed to be the helical structure in dichloromethane. A little concentration dependence of the peptides **1** and **2** is assumed to be mainly due to intermolecular hydrogen bonds between the free N-H and C=O groups in the helical structure. On the other hand, Boc-Gln-Val-Glu(OBzl)-Leu-Gly-OPac **6**, which is also the partial sequence of the C-peptide, has the β-sheet structure in the solid state and it is insoluble in dichloromethane.[2,4] The insolubility represents that, in dichloromethane, the conformational transformation from intermolecular to intramolecular hydrogen bonding does not take place in the pentapeptide **6**. The insolubility is probably due to intermolecular hydrogen bonds through the Gln side-chain amide as reported for oligopeptides containing the Asn residue,[21] and the intermolecular hydrogen bonds may interrupt the conformational transformation from intermolecular to intramolecular hydrogen bonding in dichloromethane.

With respect to the role of the Pro residue in conformational behaviors of peptides containing the Pro residue, as reported in the previous papers,[2,4] the peptides **3** and **4** in the solid state, which contain the Pro residue in central positions in the peptide chains, have strong bands at 3290—3270 and 1635—1630 cm^{-1}, accompanied by strong broad shoulders at 1700—1650 cm^{-1} in the amide I region, indicating the disturbance of the β-sheet structure by the rotation of the tertiary peptide bond planes (Gly-Pro and Gln-Pro). In the solid state, the role of the Pro residue in conformational behaviors of the peptides **3** and **4** is just the same as that of oligo(Leu)s as reported in previous papers.[11-13] In dichloromethane, each Pro residue in the hexapeptide **3** and the nonapeptide **4** clearly plays a different role in conformational behaviors of both peptides. When we compare

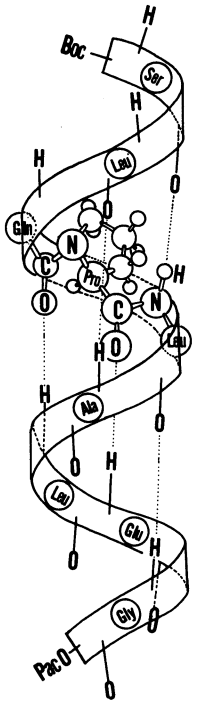

Fig. 5. α-Helical structure of the peptide **4**. Circles containing abbreviations for amino acids represent C^{α} atoms for each amino acid.

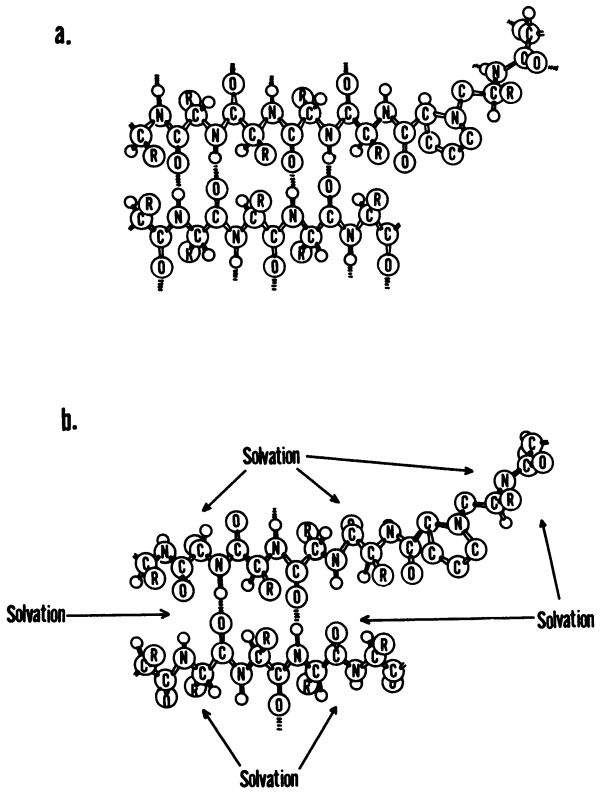

Fig. 6. Role of the Pro residue in dissolving process of the peptide **4**. a: In the solid state; b: solvating process.

solubility of the peptide **3** with that of Boc-Gly_6-OBzl,[22] the solubility improvement in the peptide **3** by the Pro residue is quite significant. The Pro residue in the peptide **3** clearly disturbs the development of the β-sheet structure in dichloromethane. The peptide **3** contains the four Gly residues, and each Gly residue can explore a large conformational space,[22,23] thus inducing a high flexibility to the peptide chain. Actually, the wide range of hydrogen bonded N-H stretching band and concentration dependence of the peptide **3** indicate that the peptide **3** in dichloromethane contains many kinds of intermolecular and intramolecular hydrogen bonds and exists in more than one conformation. On the other hand, the conformational behaviors of the peptide **4** in dichloromethane, which has polar side chains, are just the same as those of Boc-Leu_3-Pro-Leu_4-OBzl (**7**),[20] and the IR difference spectrum for the peptides **4** and **5** (Fig. 4) suggests that the peptide **4** appears to have successive-intramolecular hydrogen bonds, probably corresponding to the α-helical structure as shown in Fig. 5. This is in remarkable contrast with the fact that Boc-Leu-Ala-Leu-Glu(OBzl)-Gly-Ser(Bzl)-Leu-Gln-OPac (**8**), which is also the partial sequence of the C-peptide and has the very close amino acid sequence with the peptide **4**, has the β-sheet structure in the solid state and is insoluble in dichloromethane.[2,4] The peptides **4** and **7** have high parameters for α-helix formation, and their $<P_{\alpha}>$ values[24] are 1.12 and 1.25, respectively. In fact, the peptide **4** is included in one of helical regions in human proinsulin.[25] In comparison with the conformational behaviors of the peptide **8**, the IR difference spectra for the peptides **4** and **5** in dichloromethane clearly suggest that the Pro residue, which exists in near the N-terminal portion in the peptide having a high potential for the formation of the α-helical structure, plays an important role in helical folding in the peptide. Figure 6 represents the role of the Pro residue in dissolving process of the peptide **4**. The Pro residue in the peptide **4** clearly disturbs the development of the β-sheet aggregation in the solid state by the rotation of the tertiary peptide bond plane (Gln-Pro) and facilitates solvation of the peptide chain. Thus, the restriction of the values of the backbone dihedral angles ϕ and ψ of the Pro residue appears to promote helical folding in dichloromethane of the peptide **4**.

References

1) This paper forms Part III of "Design of the Synthetic Route for Peptides and Proteins Based on the Solubility Prediction Method" series. For part II of this series, see M. Narita, T. Ogura, K. Sato, and S. Honda, *Bull. Chem. Soc. Jpn.*, **59**, 2439 (1986). The abbreviations for amino acids are those recommended by the IUPAC-IUB Commission on Biochemical Nomenclature, *J. Biol. Chem.*, **247**, 977 (1972). Amino acid symbols except for Gly denote the L-configuration. Additional abbreviations used are the following: IR, infrared; OBzl, benzyl ester; Bzl, benzyl; Aib, α-aminoisobutyric acid residue; HPLC, high-performance liquid

chromatography; Boc, *t*-butoxycarbonyl; Pac, phenacyl.

2) M. Narita, T. Ogura, K. Sato, and S. Honda, *Bull. Chem. Soc.* **59**, 2433 (1986).

3) M. Narita, K. Ishikawa, J.-Y. Chen, and Y. Kim, *Int. J. Peptide Protein Res.*, **24**, 580 (1984).

4) M. Narita, T. Ogura, K. Sato, and S. Honda, *Bull. Chem. Soc. Jpn.*, **59**, 2439 (1986).

5) R. B. Merrifield, *J. Am. Chem. Soc.*, **85**, 2149 (1963).

6) M. Narita, Y. Tomotake, S. Isokawa, T. Matsuzawa, and T. Miyauchi, *Macromolecules*, **17**, 1903 (1984).

7) P. Y. Chou and G. D. Fasman, *Biochemistry*, **13**, 211 (1974); *ibid.*, **13**, 222 (1974); P. Y. Chou and G. D. Fasman, *Advan. Enzymol.*, **47**, 45 (1978).

8) S. Lifson and C. Sander, *Nature*, **282**, 109 (1979).

9) K. Nagano, *Kagakunoryoiki*, **36**, 120 (1982).

10) C. Toniolo, G. M. Bonora, M. Mutter, and V. N. R. Pillai, *Makromol. Chem.*, **182**, 1997 (1981); *ibid.*, **182**, 2007 (1981).

11) M. Narita, T. Fukunaga, A. Wakabayashi, K. Ishikawa, and H. Nakano, *Int. J. Peptide Protein Res.*, **23**, 306 (1984); M. Narita, K. Ishikawa, H. Nakano, and S. Isokawa, *ibid.*, **24**, 14 (1984).

12) M. Narita, S. Nagasawa, J.-Y. Chen, H. Sato, and Y. Tanaka, *Makromol. Chem.*, **185**, 1069 (1984); M. Narita, N. Ohkawa, S. Nagasawa, and S. Isokawa, *Int. J. Peptide Protein Res.*, **24**, 129 (1984).

13) S. Isokawa, T. Asakura, and M. Narita, *Macromolecules*, **18**, 871 (1985); S. Isokawa, I. Tominaga, T. Asakura, and M. Narita, *ibid.*, **18**, 878 (1985).

14) D. F. De Tar and N. Luthra, *J. Am. Chem. Soc.*, **99**, 1232 (1977); V. Madison, *Biopolymers*, **16**, 2671 (1977); B. V. V. Prasad, H. Balaram, and P. Balaram, *ibid.*, **21**, 1261 (1982).

15) M. Narita, M. Doi, H. Sugasawa, and K. Ishikawa, *Bull. Chem. Soc. Jpn.*, **58**, 1473 (1985); M. Narita, K. Ishikawa, H. Sugasawa, and M. Doi, *ibid.*, **58**, 1731 (1985).

16) M. Narita, J.-Y. Chen, H. Sato, and Y. Kim, *Bull. Chem. Soc. Jpn.*, **58**, 2494 (1985).

17) T. Miyazawa and E. J. Blout, *J. Am. Chem. Soc.*, **83**, 712 (1961); Yu. N. Chirgadze and N. A. Nevskaya, *Biopolymers*, **15**, 627 (1976); M. Narita, Y. Tomotake, S. Isokawa, T. Matsuzawa, and T. Miyauchi, *Macromolecules*, **17**, 1903 (1984); M. Narita, S. Isokawa, Y. Tomotake, and S. Nagasawa, *Polymer J.*, **15**, 25 (1983), and references cited therein.

18) C. H. P. Rao, P. Balaram, and C. N. R. Rao, *Biopolymers*, **22**, 2091 (1983) and references cited therein.

19) G. Boussard and M. Marraud, *J. Am. Chem. Soc.*, **107**, 1825 (1985); C. Toniolo, G. M. Bonora, V. Borone, A. Bavoso, E. Benedetti, B. D. Blasio, P. Grimaldi, F. Lelj, V. Pavone, and C. Pedone, *Macromolecules*, **18**, 895 (1985).

20) M. Narita, M. Doi, R. Wakita, and S. Isokawa, *Bull. Chem. Soc. Jpn.*, in press.

21) T. Toniolo, G. M. Bonora, and W. M. M. Schaaper, *Int. J. Peptide Protein Res.*, **23**, 389 (1984).

22) M. Narita, M. Doi, K. Kudo, and Y. Terauchi, *Bull. Chem. Soc. Jpn.*, in press.

23) D. A. Brant, W. G. Miller, and P. J. Flory, *J. Mol. Biol.*, **23**, 47 (1967); R. T. Ingwall, E. A. Crurylo, and P. J. Flory, *Biopolymers*, **12**, 1137 (1973).

24) P. Y. Chou and G. D. Fasman, *Biochemistry*, **13**, 211 (1974).

25) C. R. Snell and P. G. Smyth, *J. Biol. Chem.*, **250**, 6291 (1975).