Syntheses and Properties of Tertiary Peptide Bond Containing Polypeptides. 7.[1] Conformational Studies of Sequential Polypeptides Containing the Pro-Pro Sequence by ^{13}C and ^{1}H NMR

Shizuko Isokawa,[†] Ichiro Tominaga,[†] Tetsuo Asakura,[‡] and Mitsuaki Narita*[†]

Department of Industrial Chemistry and Department of Polymer Engineering, Faculty of Technology, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184, Japan.
Received May 7, 1984

ABSTRACT: The conformations of Boc-$(Leu_3$-Pro_2-$Gly)_n$-OBzl **2** (n = 2), **3** (n = 3), and **4** (n = 4) and Boc-Leu_3-Pro_2-Gly-Leu_3-OBzl (**5**) in Me_2SO-d_6 have been studied by ^{13}C and ^{1}H NMR spectroscopy. They are almost fully solvated in this solvent. The local conformations of N-terminal Leu_3 segments and C-terminal Pro_2-Gly segments in the peptides **2–4** are similar to those of the corresponding segments in the hexapeptide **1** (n = 1), in which ca. 15% of the cis isomer about the Pro-Pro bond exists. The internal peptide chains, $-(Pro_2$-Gly-$Leu_3)_m-$, of the peptides **2–4** have a repeating local conformation characteristic to the internal hexapeptide segment, Pro_2-Gly-Leu_3, in which 6–8% of the cis isomer about each Pro-Pro bond exists.

Introduction

The reactivity of a peptide fragment in peptide synthesis by fragment condensation is strongly affected by its conformation, especially at the amino and carboxyl terminals of the fragment concerned in the coupling reaction. As a matter of fact, the physicochemical properties of oligopeptides depend to a large extent on the conformational preferences of the peptide chain. In particular, the solubility of the peptide and the reactivity of the terminal amino group are significantly reduced in oligomers exhibiting β-structures compared to oligomers in α-helical or randomly coiled conformations.[2,3] On the other hand, a peptide fragment in a randomly coiled structure was found to maintain good solubility and high reactivity in the preparation of sequential polypeptides Boc-$(Leu_3$-Pro_2-$Gly)_n$-OBzl (n = 1, 2, 4, 6, 8, 10, and 12).[4] By CD and molar rotation measurements, the peptides have been demonstrated to exist as a predominantly randomly coiled structure in strongly polar solvents.[5] The maintenance of a randomly coiled structure in spite of chain elongation was attributed to "peptide segment separation" which helps to facilitate solvation of the peptide chain via the insertion of tertiary peptide bonds at suitable intervals.[4-7] The N- or C-terminal deprotected types of these peptides have also been shown to maintain high reactivity independently of chain length.[4,8] The effect of the insertion of guest amino acid residues such as Gly and Pro on the solid-state and solution conformations of host peptides has also been studied by a number of workers[9-15] using IR, CD, and NMR spectroscopy. It is the objective of this study to investigate, by the use of ^{13}C and ^{1}H NMR, the conformation of peptides **2–4** and the relationship between the chain length of the peptide and the conformations of individual residues.

So far, several conformational studies of sequential polypeptides by NMR have been reported.[16-19] Each pentapeptide sequence of tropoelastin, Boc-(Val-Pro-Gly-Val-Gly)-OMe and HCO-(Val-Pro-Gly-Val-Gly)$_n$-OMe (n = 2 and 3), has been reported to exist as a β-turn structure, with the sequence Pro-Gly at the corner, by examination of the temperature and solvent dependencies of the NH proton chemical shifts.[19] Consequently, in this study we also examine, by use of ^{13}C and ^{1}H NMR, the cis contents about individual Pro-Pro bonds of the peptides **2–4** in Me_2SO-d_6. In addition, in order to delineate the conformation about individual residues or sequences in this solvent, this study will investigate the temperature and concentration dependencies of NH chemical shifts of the peptides **2–4** in Me_2SO-d_6. The conformation of **5** will also be investigated to delineate the conformation of the Pro_2-Gly segment in the internal peptide chain. Conformational studies of hexapeptide **1** by NMR have been reported in detail in the previous paper,[20] and the NMR data developed there are used in this study.

Experimental Section

The syntheses and physical properties of Boc-$(Leu_3$-Pro_2-$Gly)_n$-OBzl **1** (n = 1), **2** (n = 2), and **4** (n = 4) have been reported previously.[4] The synthesis of Boc-$(Leu_3$-Pro_2-$Gly)_3$-OBzl **3** was carried out by using Boc-Leu_3-Pro_2-Gly-OH and HCl·$(Leu_3$-Pro_2-$Gly)_2$-OBzl by the usual method described previously.[4] The product was recrystallized from acetone/ethyl acetate. Anal. Found: C, 59.28; H, 8.26; N, 12.04. Calcd. for $C_{102}H_{166}O_{21}N_{18}\cdot 5H_2O$: C, 59.16; H, 8.57; N, 12.18. Amino acid analysis: Pro, 1.83; Gly, 1.10; Leu, 3.00. (The recovery of Pro was 93%.) Boc-Leu_3-Pro_2-Gly-Leu_3-OBzl (**5**) was synthesized from Boc-Leu_3-Pro_2-Gly-OH and HCl·Leu_3-OBzl by the same method.[4] The product was recrystallized from ethyl acetate/n-hexane. Anal. Found: C, 62.10; H, 8.56; N, 10.84%. Calcd. for $C_{60}H_{100}O_{12}N_9\cdot H_2O$: C, 62.26; H, 8.88; N, 10.89. Amino acid analysis: Gly, 1.03; Pro, 2.14; Leu, 6.00. Tetramethylsilane (Me_4Si) from Merck and Co. and Me_2SO-d_6 (99.8%) from CEA were used as an internal standard and a solvent for NMR measurements, respectively. ^{13}C NMR spectra were obtained on a Jeol FX200 spectrometer at room temperature, with 45° pulse, 16K data points on a spectral width of 10 000 Hz, acquisition time 0.82 s, and with delay time 1.5 s. Up to 21 000 scans were made to improve signal-to-noise ratios for ^{13}C NMR measurements. ^{1}H NMR spectra were also obtained on a Jeol FX200 spectrometer at room temperature; 100–200 scans were made for ^{1}H NMR measurements. Approximately 0.2 M solutions per unit of hexapeptide in Me_2SO-d_6 were prepared and used for NMR measurements except for the experiments on concentration dependencies of the NH chemical shifts. All variations of the NH chemical shifts with temperature were found to be linear. Chemical shifts were recorded by using an internal Me_4Si reference. The amino acid residues were numbered from the N terminal of the peptide chain.

Results and Discussion

Conformational Analysis by ^{13}C NMR for Pro-Pro Segments. Figure 1 shows ^{13}C NMR spectra for the peptides **1–5** in the spectral region between 25 and 67 ppm. The α carbons of the residues preceding Pro (attached to the imino nitrogen) have been known to resonate upfield

[†] Department of Industrial Chemistry.
[‡] Department of Polymer Engineering.

0024-9297/85/2218-0878$01.50/0

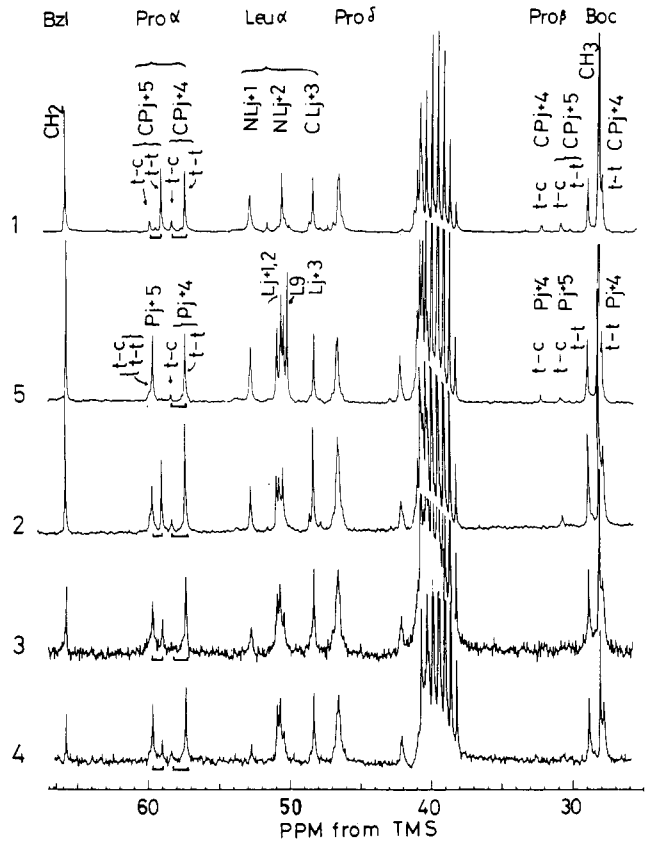

Figure 1. Partial ^{13}C NMR spectra of peptides **1–5** in Me_2SO-d_6 at 50 MHz together with the assignment. Chemical shifts were represented in ppm from Me_4Si.

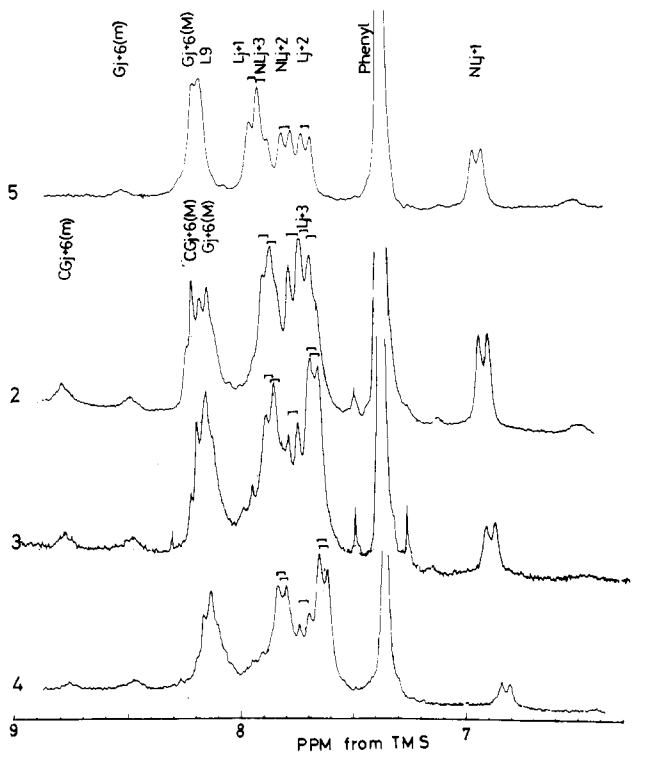

Figure 2. Amide regions of ^1H NMR spectra of peptides **2–5** in Me_2SO-d_6 at 200 MHz together with the assignment. Chemical shifts were represented in ppm from Me_4Si.

by 1.0 ppm compared to the resonance position of the α carbons placed at other points in the peptide chain.[21] Thus, peaks for Leu and Pro α carbons were assigned from the chemical shifts as shown in Figure 1. The peak at the lowest field in the spectral region of the α carbon of Leu was assigned to the α carbon of N-terminal Leu from the chemical shift as reported in the previous paper.[20] The hexapeptide **1** exhibited two peaks attributable to the Pro(4) α carbon; a major peak at 57.46 ppm (M) due to the trans(Leu-Pro bond)–trans(Pro-Pro bond) form and a minor peak at 58.42 ppm (m) due to the trans(Leu-Pro bond)–cis(Pro-Pro bond) form.[20] The cis isomer about the Pro-Pro bond was estimated to be 15% from the relative intensity of the two peaks. The nonapeptide **5** similarly exhibits two peaks for the α carbon of Pro(4) at 57.38 (M) and 58.41 (m) ppm. The peptide **5** also exhibits two minor peaks for β carbons of Pro(4) and Pro(5) at 32.23 and 30.82 ppm, respectively. The chemical shifts coincide with those of the minor peaks for β carbons of Pro(4) (at 32.21 ppm) and Pro(5) (at 30.80 ppm) of **1**. These results indicate that the minor peak for the α carbon of Pro(4) and the minor peaks for β carbons of Pro(4) and Pro(5) in the spectrum of **5** are of the trans(Leu-Pro)–cis(Pro-Pro) form. The cis content was estimated to be 9–11%. Thus, the Pro_2-Gly segment placed between two Leu_3 segments exhibits cis content about the Pro-Pro bond different from that for the Pro_2-Gly segment placed at the carboxyl terminal in the hexapeptide **1**. The peptide **5** exhibits a peak having a shoulder attributable to the α carbon of Pro(5) at 59.67 ppm. The peak position is at lower field by 0.5 ppm compared to the major peak for Pro(5) of **1**. The dodecapeptide **2** exhibits two major peaks at 59.04 and 59.71 ppm due to α carbons of Pro(11) and Pro(5), respectively. Major (at 57.41 ppm) and minor (at 58.37 ppm) peaks are observed for the α carbons of Pro(4) and Pro(10) of the same peptide. The latter peaks show a pattern similar to the peaks for α carbons of Pro(4) of **1** and **5**. These results indicate that the Pro(4)-Pro(5) segment of the dodecapeptide **2** has a local conformation similar to that of the Pro(4)-Pro(5) segment of **5** and that the Pro(10)-Pro(11) segment of **2** also has a conformation similar to that of the Pro(4)-Pro(5) segment of **1**. The cis contents about two Pro-Pro bonds cannot be estimated separately because both the minor and major peaks of each α carbon of Pro(4) and Pro(10) resonate at the same field. Two major peaks are observed at 59.04 and 59.74 ppm in the spectrum of **3**. The intensity of the peak at 59.74 ppm is approximately twice the intensity of the peak at 59.04 ppm. The peptide **4** also exhibits two major peaks at 59.04 and 59.69 ppm, where the peak at 59.69 ppm is higher and larger than the peak at 59.04 ppm. These results indicate that the Pro-Pro segments at internal positions in peptide chains have the same conformation as that of the Pro-Pro segment of **5**. The Pro-Pro segments existing within the peptide chains have a similar conformation. But, the Pro-Pro segments existing at the C-terminal Pro_2-Gly sequences have another conformation similar to that of **1**. The cis contents about the Pro-Pro bonds of **2–4** will be estimated from the ^1H NMR spectra as described below.

Conformational Analysis by ^1H NMR. Figure 2 shows ^1H NMR spectra for the peptides **2–5**, in the spectral region between 6.3 and 9 ppm. The high-field NH resonances were assigned to the N-terminal Leu NH protected by the Boc group.[22] In the previous paper,[20] we have demonstrated that the $C^\alpha H$ proton of Leu preceding a Pro residue in the Leu_2-Pro_2-Gly or Leu_3-Pro_2-Gly sequences resonates at 4.55 ppm and that the $C^\alpha H$ protons of the other Leu within the peptide chains resonate at 4.35 ppm. Thus, from the result of decoupling experiments, the peak at 7.84 ppm for **5** is attributable to the NH proton of Leu(3). The assignment of the other NH peaks in the spectrum of **5** was performed from comparisons of the chemical shifts and the temperature dependencies with

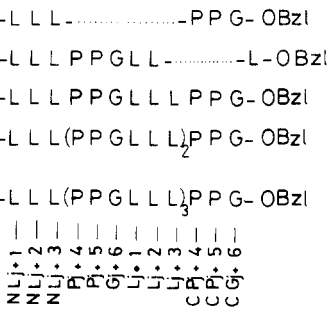

Figure 3. Differentiation of the amino acid residues in the sequential polypeptides Boc-$(Leu_3\text{-}Pro_2\text{-}Gly)_n$-OBzl 1–4 ($n$ = 1–4) and Boc-Leu_3-Pro_2-Gly-Leu_3-OBzl (5).

those of the hexapeptide 1 and is shown in Figure 2. Similarly, the peaks for NH protons of peptides 2–4 were assigned from decoupling experiments, chemical shift data, and the temperature coefficient of the chemical shift. The spectra of 2 and 5 exhibit a minor peak at about 8.0 ppm (at 30 °C), and the minor peak would be assigned to Lj+1 NH (differentiation of the amino acid residues is shown in Figure 3) proton in the cis configuration about the Pro-Pro bonds. The spectra of 3 and 4 also exhibit two minor peaks at about 8.0 and 7.9 ppm (at 30 °C), and they would also be due to Lj+1 NH (Figure 3) protons in the cis configuration about the Pro-Pro bonds.

The NMR parameters are summarized in Table I. Figure 3 shows differentiation of the amino acid residues in the peptides 1–5. The symbols N and C represent the residue placed at the N- and C-terminal regions, respectively, and symbols L, G, and P represent Leu, Gly, and Pro residues, respectively. The numbers 1–6 give the positions of the amino acid residue from the N-terminal of each hexapeptide unit Leu_3-Pro_2-Gly and j = 0, 6, or 6(n − 1). It is clear that the residues distinguished in Figure 3 have approximately the same chemical shift and temperature coefficient of the chemical shift as for all the peptides examined in this paper (Table I). As is evident from Table I, the same characteristics have been observed for the coupling constants $J_{NH\text{-}C^{\alpha}H}$ of these peptides. These results indicate that chain elongation does not affect the conformation of the hexapeptide Pro_2-Gly-Leu_3 segments at internal positions in these peptides and that both the N-terminal Leu_3 and C-terminal Pro_2-Gly segments have a conformation characteristic to the terminal positions. The conformations for N-terminal Leu_3 and C-terminal Pro_2-Gly segments are different from that of the internal sequences. These conclusions can be supported from the ^{13}C NMR data concerning the Pro_2 segments as mentioned above.

Most of the temperature coefficients of NH chemical shifts are larger than 4×10^{-3} ppm/°C, indicating that the NH protons are almost fully solvated.[18,22-24] Exceptions are the NLj+2 and Lj+2 NH protons and several CGj+6(m) and Gj+6(m) NH protons. We have reported that the NH protons of Leu(2) in the hexapeptide 1 and Boc-Leu_3-Pro-OBzl have smaller temperature coefficients than the NH protons of the other Leu's in these peptides in Me_2SO-d_6.[20] This has been attributed to the presence of the bulky side chains of the neighboring residues. The same cause may be attributed to the small temperature dependencies of NLj+2 and Lj+2 NH chemical shifts observed in the present study.

The temperature dependencies of Gj+6 NH (m) and Gj+6 NH (M) protons are indicative of a slight shielding of the Gj+6 NH protons from the solvent. The magnitude of the temperature coefficients of Gj+6 NH (m) of 2–5 are comparable to those of the Gj+6 NH (M). The shielding

Table I
^1H NMR Parameters of Boc-$(Leu_3\text{-}Pro_2\text{-}Gly)_n$-OBzl 1–4 ($n$ = 1–4) and Boc-Leu_3-Pro_2-Gly-Leu_3-OBzl (5) in Me_2SO-d_6

	NH chem shift,[a] ppm				
residue	1	5	2	3	4
NLj+1	6.91	6.90	6.90	6.92	6.85
NLj+2(M)	7.76	7.76	7.79	7.77	7.73
NLj+2(m)	7.66				
NLj+3(M)	7.85	7.84	7.86	7.85	7.82
NLj+3(m)	8.02				
Gj+6(M)		8.14	8.18	8.17	8.14
Gj+6(m)		8.50	8.52	8.49	8.50
Lj+1(M)		7.87	7.88	7.89	7.83
Lj+1(m_1)		8.01	7.99	7.99	7.95
Lj+1(m_2)				7.95	7.89
Lj+2		7.66	7.67	7.68	7.64
Lj+3(M)			7.69	7.70	7.65
CGj+6(M)	8.21		8.22	8.22	8.21
CGj+6(m)	8.78		8.81	8.80	8.78
Leu(9)		8.01			
	temp coeff of NH chem shift, 10^3 ppm/°C				
residue	1	5	2	3	4
NLj+1	6.8	7.2	7.6	6.8	7.3
NLj+2(M)	3.3	3.3	3.6	3.1	3.3
NLj+2(m)	3.2				
NLj+3(M)	4.9	5.0	5.5	5.3	5.4
NLj+3(m)	5.8				
Gj+6(M)		4.1	4.3	4.2	4.2
Gj+6(m)		3.6	4.4	3.3	4.1
Lj+1(M)		5.2	5.3	4.4	4.4
Lj+1(m_1)		6.8	5.2	4.7	5.0
Lj+1(m_2)				4.7	4.8
Lj+2		3.2	3.7	3.6	3.5
Lj+3			5.6	5.7	6.0
CGj+6(M)	5.0		5.5	5.3	6.1
CGj+6(m)	3.9		4.7	4.2	4.1
Leu(9)		6.8			
	$J_{NH\text{-}C^{\alpha}H}$,[b] Hz				
residue	1	5	2	3	4
NLj+1	8.0	7.9	7.9	7.9	7.9
NLj+2	8.4	8.6	8.6	8.6	8.6
NLj+3	7.6	7.3	7.1	7.1	6.9
Lj+1		7.2	7.6	7.1	7.6
Lj+2		7.7	8.6	7.9	8.3
Lj+3			7.9	7.9	7.4
Leu(9)		7.6			

[a] Relative to Me_4Si at 30 °C. Calculated from temperature dependencies of NH chemical shifts. [b] Coupling constants.

of Gj+6 NH (m) and CGj+6 NH (m) protons from the solvent is probably due to intramolecular hydrogen-bond formation between the NH proton and the carbonyl of NLj+3 and Lj+3 (type VI β-turn). The CD spectra of 1–5 in methanol exhibited a small negative band with a maximum at 225–232 nm, and the difference spectra for 1–5 produced by dioxane referred to methanol were characterized by a negative band with a maximum at about 220–222 nm. From these results, the negative band at about 230 nm of 1–5 in methanol may be due to a small contribution of a γ-turn structure on the conformation of 1–5 in methanol.[25] Thus, it may be concluded that the shielding of Gj+6 NH (M) protons from the solvent is due to a small contribution of a γ-turn structure in which Gj+6 NH (M) protons are intramolecularly hydrogen bonded with carbonyl of Pj+4. However, the ordered structures have made only a small contribution to the overall conformation of the peptides.

From the relative intensities of CGj+6(M) and CGj+6(m) NH peaks, the cis contents about the Pro-Pro bond in the C-terminal Pro_2-Gly segments were estimated to be 15%, 15%, 15%, and 13% for the peptides 1–4, respec-

tively. Similarly, the cis contents about the Pro-Pro bonds existing in the internal peptide chains can be estimated from the intensity ratios of G*j*+6(m) NH to G*j*+6(M) NH to be, on the average, 6%, 7%, 8%, and 6% for the peptides **2**–**5**, respectively. A small difference observed for the values of the cis content of **5** between ^{13}C and ^{1}H NMR spectra would be due to measurement errors.[26]

Chemical shifts for the NH protons of **2** did not exhibit concentration dependence in Me_2SO-d_6 over the concentration range from 0.004 to 0.25 M. This fact indicates that, as expected, the NH protons of **2** are free from intermolecular hydrogen bonding in Me_2SO-d_6 at the levels of concentration examined.

In conclusion, the sequential polypeptides Boc-$(Leu_3$-Pro_2-Gly$)_n$-OBzl **1**–**4** (n = 1–4) exist in an almost fully solvated conformation (a randomly coiled structure) in Me_2SO, in which the internal peptide chains, -$(Pro_2$-Gly-$Leu_3)_m$-, have a repeating local conformation characteristic of the internal hexapeptide segment, and the N-terminal Leu_3 and C-terminal Pro_2-Gly segments have local conformations common to the sequential polypeptides. The above conclusion would help to explain a similar reactivity of the carboxyl-free sequential polypeptides Boc-$(Leu_3$-Pro_2-Gly$)_n$-OH.[3]

References and Notes

(1) Part 6 in this series coincides with ref 20. The abbreviations for amino acids were those recommended by the IUPAC-IUB Commission on Biochemical Nomenclature (*J. Biol. Chem.* **1972**, *247*, 977). The amino acid symbols except Gly denote the L configuration.

(2) Abd el Rahman, S.; Anzinger, H.; Mutter, M. *Biopolymers* **1980**, *19*, 173.

(3) Isokawa, S.; Matsuzawa, T.; Ohkawa, N.; Narita, M. Lecture at the 47th Congress of the Chemical Society of Japan, Kyoto, Japan, April 1983.

(4) Narita, M.; Nagasawa, S.; Chen, J. Y.; Sato, H.; Tanaka, Y. *Makromol. Chem.* **1984**, *185*, 1069.

(5) Narita, M.; Ohkawa, N.; Nagasawa, S.; Isokawa, S. *Int. J. Peptide Protein Res.* **1984**, *24*, 129.

(6) Narita, M.; Fukunaga, T.; Wakabayashi, A.; Ishikawa, K.; Nakano, H. *Int. J. Peptide Protein Res.* **1984**, *23*, 306.

(7) Narita, M.; Ishikawa, K.; Nakano, H.; Isokawa, S. *Int. J. Peptide Protein Res.* **1984**, *24*, 14.

(8) Narita, M.; Matsuzawa, T.; Miyauchi, T.; Isokawa, S. Lecture at the 32th Congress of the Society of Polymer Science, Kyoto, Japan, May 1983.

(9) Grathwohl, C.; Wüthrich, K. *Biopolymers* **1976**, *15*, 2025.

(10) Naider, F.; Becker, J. M.; Ribeiro, A. A.; Goodman, M. *Biopolymers* **1978**, *17*, 2213.

(11) Ribeiro, A. A.; Goodman, M.; Naider, F. *J. Am. Chem. Soc.* **1978**, *100*, 3903.

(12) Hagler, A. T.; Stern, P. S.; Sharon, R.; Becker, J. M.; Naider, F. *J. Am. Chem. Soc.* **1979**, *101*, 6842.

(13) Naider, F.; Sipzner, R.; Steinfeld, A. S.; Becker, J. M. *J. Chromatogr.* **1979**, *176*, 264.

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

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

(16) Kobayashi, Y.; Kyogoku, Y. *J. Mol. Biol.* **1973**, *81*, 337.

(17) Urry, D. W.; Mitchell, L. W.; Ohnishi, T. *Biochem. Biophys. Res. Commun.* **1974**, *59*, 62.

(18) Urry, D. W.; Ohnishi, T. *Biopolymers* **1973**, *13*, 1223.

(19) Urry, D. W.; Cunningham, W. D.; Ohnishi, T. *Biochemistry* **1974**, *13*, 609.

(20) Isokawa, S.; Asakura, T.; Narita, M. *Macromolecules*, preceding article in this issue.

(21) Deslauriers, R.; Becker, J. M.; Steinfeld, A. S.; Naider, F. *Biopolymers* **1979**, *18*, 523.

(22) Pysh, E. S.; Toniolo, C. *J. Am. Chem. Soc.* **1977**, *99*, 6211.

(23) Stimson, E. R.; Zimmerman, S. S.; Scheraga, H. A. *Macromolecules* **1977**, *10*, 1049.

(24) Iqbal, M.; Balaram, P. *J. Am. Chem. Soc.* **1981**, *103*, 5548.

(25) Cann, J. R.; Stewart, J. M.; Matsueda, G. R. *Biochemistry* **1973**, *12*, 3780.

(26) The extent of the measurement error of the cis content from the relative intensities of each set of two peaks of α, β, and γ carbons is estimated to be ±2.9% of the value of the cis content. The extent of the measurement error of the cis content from the areas of two peaks of Gly NH proton is estimated to be ±2.7% of the value of the cis content.