  Bull. Chem. Soc. Jpn., 59, 3547—3552 (1986)

The Ability of the Proline Residue to Promote Successive-Intramolecular Hydrogen Bonds in Oligopeptides[1)]

Mitsuaki NARITA, Shizuko ISOKAWA, Masamitsu DOI, and Ryuhei WAKITA
Department of Industrial Chemistry, Faculty of Technology, Tokyo University of Agriculture and Technology, Koganei 184
(Received May 15, 1986)

In order to investigate the ability of the Pro residue to promote helical folding in oligopeptides, the IR spectroscopic conformational study of the oligo(Leu)s containing the Pro residue was performed in dichloromethane. Their IR difference spectra suggested the occurrence of successive-intramolecular hydrogen bonds just like incipient α-helical structures formed by one, two, three, and so forth $i \rightarrow i-4$ hydrogen-bonding patterns, alluding to the ability of the Pro residue to promote helical folding in the peptides. However, the intensities of hydrogen-bonded absorption bands in the amide A region of each peptide strongly suggested that conformations of the peptides investigated were also contributed by other intramolecularly hydrogen-bonded folded structures, which had lower absorption coefficients. The initiation mechanism of successive-intramolecular hydrogen bonding in the peptides is apparently attributed to the disturbance of the β-sheet formation by the rotation of the tertiary peptide bond plane (Leu-Pro), that is, the restriction of the values of the backbone dihedral angles ϕ and ψ of the Pro residue due to the low flexibility of the pyrrolidine ring.

In connection with the design of the synthetic route for peptides and proteins based on the solubility prediction method,[2-5)] it is of particular importance to investigate the nucleation site of helical folding in peptide intermediates since the β-sheet→helical conformational transformation gives rise to the great improvement on solubility of peptide intermediates.[6-8)] The investigation of the initiation mechanism of helical folding in proteins is also one of the quite interesting problems for understanding the mechanism of protein folding and also for predicting protein folding from amino acid sequences. However, it has been well recognized that the determination of nucleation sites of helical folding in peptides and proteins is quite difficult.[9)] In our recent papers,[6-8)] we demonstrated the great ability of the Aib residue to promote helical folding in peptides and subsequently proposed that the restriction of the backbone dihedral angles ϕ and ψ of an amino acid residue to the helical region is one of important initiation mechanisms of helical folding in natural proteins. Therefore, it is of particular interest to investigate the ability of the Pro residue to promote helical folding in peptides since the most frequently observed conformations for the Pro residue lie in the region of the backbone dihedral angles $\phi=-60\pm10°$ and $\psi=-30\pm20°$ (helical region), and $\phi=-60\pm10°$, $\psi=120\pm20°$ due to the low flexibility of the pyrrolidine ring in the Pro residue.[10)]

Thus, in this paper, we give attention to the Pro residue, which is hitherto emphasized to be a strong helix and β-sheet breaker.[11-14)] As for conformations of oligo(Leu)s containing the Pro residue,[14)] in the solid state, the insertion of the Pro residue into homooligo(Leu)s caused the disturbance of the β-sheet structure by the rotation of the tertiary peptide bond plane (Leu-Pro). On the other hand, in methanol, the CD patterns of the oligo(Leu)s containing the Pro residue suggested the existence of a detectable amount of the α-helical structure. In this paper, we investigate the ability of the Pro residue to promote helical folding in oligo(Leu)s containing the Pro residue using IR difference spectra in dichloromethane and discuss the role of the Pro residue in remarkable improvement on solubility of peptide intermediates containing the Pro residue.

Experimental

The samples of Boc-Leu$_2$-OBzl **1**, Boc-Leu$_3$-OBzl **2**, Boc-Leu$_4$-OBzl **3**, and Boc-Leu$_5$-OBzl **4** were prepared by the method described in the previous paper.[14)] The other peptides **5**—**10** examined in this study were also prepared by the essentially same method described previously for

Table 1. Elemental Analysis of the Peptides **5**—**10**

Compound	Formula	Found (Calcd)		
		C (%)	H (%)	N (%)
5	$C_{45}H_{83}N_7O_{10}\cdot H_2O$	60.23 (60.04)	9.48 (9.52)	11.03 (10.89)
6	$C_{35}H_{56}N_4O_7$	65.31 (65.19)	8.71 (8.75)	8.93 (8.69)
7	$C_{41}H_{67}N_5O_8$	64.95 (64.97)	8.95 (8.91)	9.29 (9.24)
8	$C_{47}H_{78}N_6O_9$	64.52 (64.80)	9.07 (9.03)	9.66 (9.65)
9	$C_{53}H_{89}N_7O_{10}\cdot H_2O$	63.51 (63.51)	8.77 (9.15)	9.71 (9.78)
10	$C_{59}H_{100}N_8O_{11}\cdot 0.5H_2O$	64.04 (64.04)	9.22 (9.20)	10.02 (10.13)

		Boc-Leu$_2$-Aib-Leu$_4$-OH	5
Boc-Leu$_2$-OBzl	1	Boc-Leu-Pro-Leu$_2$-OBzl	6
Boc-Leu$_3$-OBzl	2	Boc-Leu-Pro-Leu$_3$-OBzl	7
Boc-Leu$_4$-OBzl	3	Boc-Leu-Pro-Leu$_4$-OBzl	8
Boc-Leu$_5$-OBzl	4	Boc-Leu$_2$-Pro-Leu$_4$-OBzl	9
		Boc-Leu$_3$-Pro-Leu$_4$-OBzl	10

Fig. 1. The peptides **1**—**10** examined in this study.

Boc-Leu$_4$-Aib-Leu$_4$-OBzl[6] and Boc-Leu$_3$-Pro-Leu$_3$-OBzl.[14] The data of elemental analysis of the peptides **5**—**10** are assembled in Table 1 and they are in good agreement with the calculated values. Each peptide also gave a single peak on HPLC. The IR spectra in dichloromethane of the peptides **1**—**10** were recorded at room temperature with a JEOL JIR-100 FT-IR spectrophotometer over a wide range

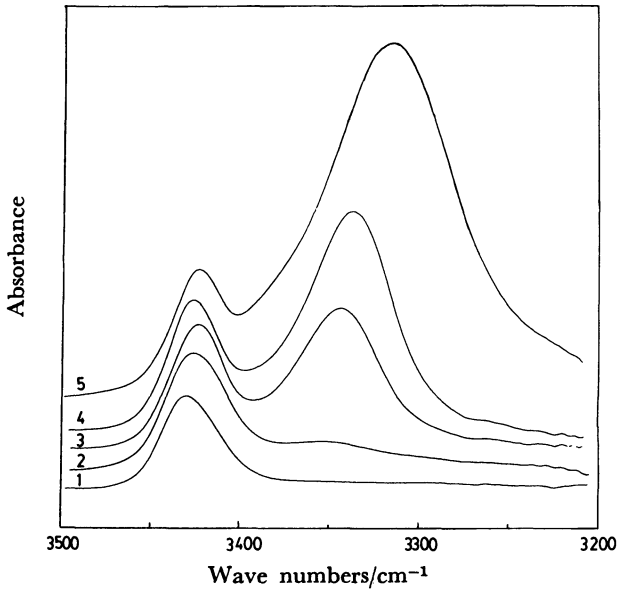

Fig. 2. IR absorption spectra in the amide A region of the peptides **1**—**5** in dichloromethane (1.0 mM).

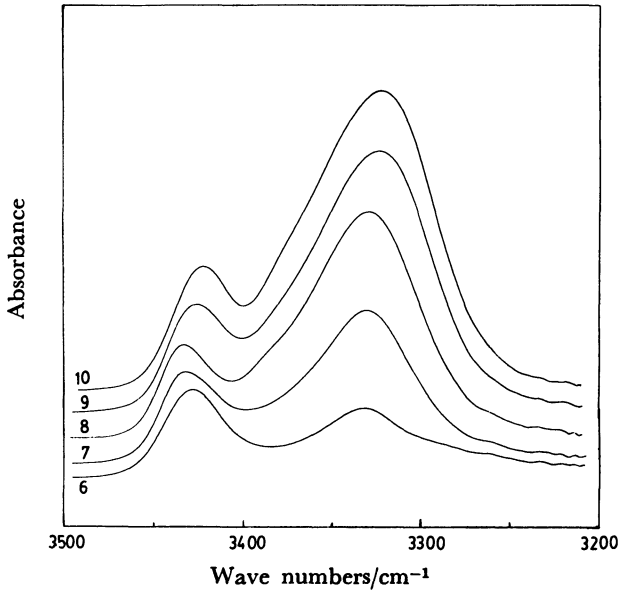

Fig. 3. IR absorption spectra in the amide A region of the peptides **6**—**10** in dichloromethane (1.0 mM).

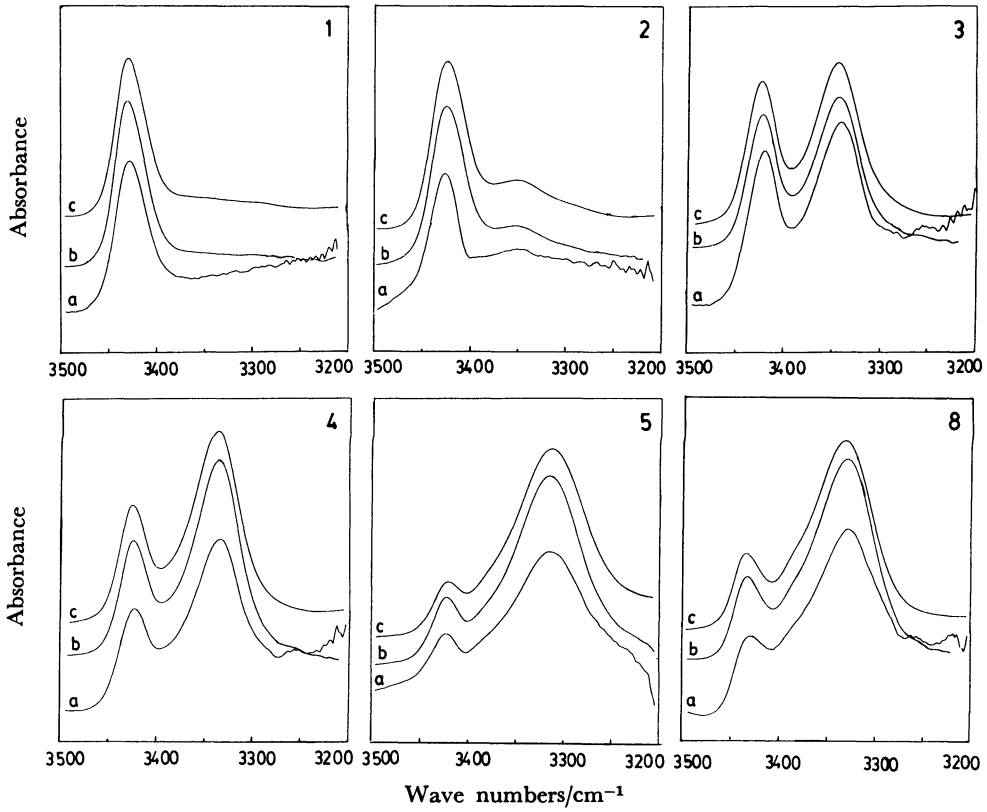

Fig. 4. IR absorption spectra in the amide A region of the peptides **1**—**4** and **8** in dichloromethane over the wide range of concentrations (0.10—10 mM). a: 0.10 mM, b: 1.0 mM, c: 10 mM.

of concentrations by employing 0.5 mm- and 5 mm-path length cells with potassium bromide windows.

Results and Discussion

The peptides **1**—**10** examined in this study are listed in Fig. 1. The IR spectra in the N-H stretching region of the peptides **1**—**5** and **6**—**10** in dichloromethane (1.0 mM (1 M=1 mol dm^{-3})) are shown in Figs. 2 and 3, respectively. The spectra in Figs 2, 3, 5, 7, and 9 are shown with the same scale of absorbance. The spectra in Figs. 2 and 3 show two distinct bands at 3430—3420 and 3350—3320 cm^{-1} although the latter band of the dipeptide **1** is negligible. The bands at 3430—3420 cm^{-1} are clearly due to the free urethane and peptide N-H stretching vibration, and the bands at 3350—3320 cm^{-1} can arise from the hydrogen-bonded N-H groups.[5,15-18] Figure 4, as typical examples, shows the IR spectra of the peptides **1**—**5** and **8** over the wide range of concentrations (0.10—10 mM). Intensity ratios of the hydrogen-bonded N-H absorption to the free N-H absorption (A_H/A_F) for each peptide do not vary much with concentration, indicating that, especially at concentrations lower than 1.0 mM, the hydrogen-bonded N-H absorptions are due almost exclusively to intramolecularly hydrogen-bonded folded species as reported for Boc-Leu$_n$-OCH_3 (n=3—6).[17]

For the purpose of elucidating intramolecularly hydrogen-bonded folded species, the IR difference spectra between the dipeptide **1** and the peptides **6** through **9** are examined in the amide A region. As shown in Fig. 5, the absorptions of the free urethane and peptide N-H groups in the peptides **6**—**9** are almost clearly compensated with those absorptions in the peptide **1**. When we presume that the absorption coefficients of the free urethane and peptide N-H groups in the unordered structure for the dipeptide **1** are nearly equal to those absorption coefficients in the α-helical structures for the peptides **6**—**9**, the results in Fig. 5 suggest the occurrence of incipient α-helices formed by one, two, three, and four $i \rightarrow i-4$ hydrogen-bonding patterns in the peptides **6**—**9**, respectively, as shown in Fig. 6 schematically. These conformational

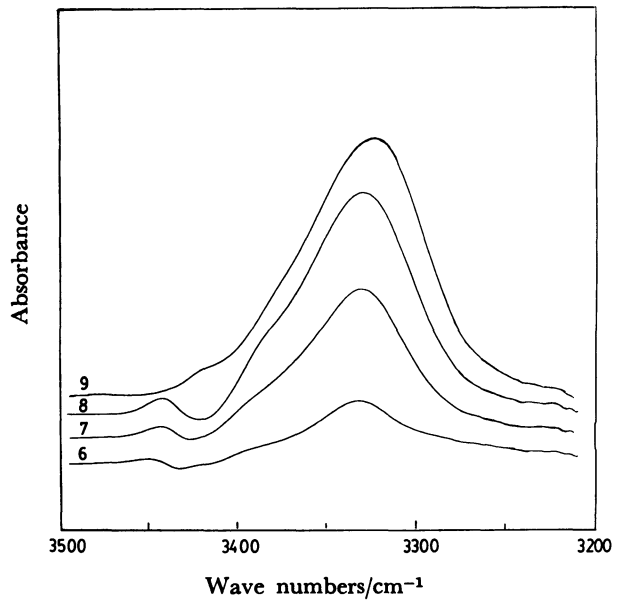

Fig. 5. IR difference spectra between the dipeptide **1** and the peptides **6** through **9** in the amide A region (1.0 mM).

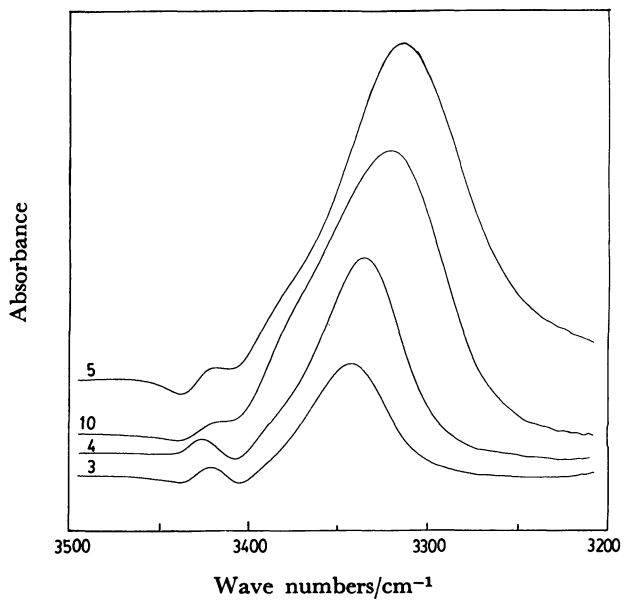

Fig. 7. IR difference spectra between the tripeptide **2** and the peptides **3**—**5** and **10** in the amide A region (1.0 mM).

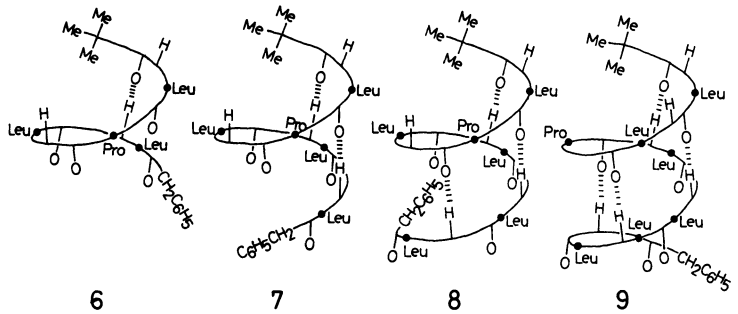

Fig. 6. α-Helical structures for the peptides **6**—**9**.

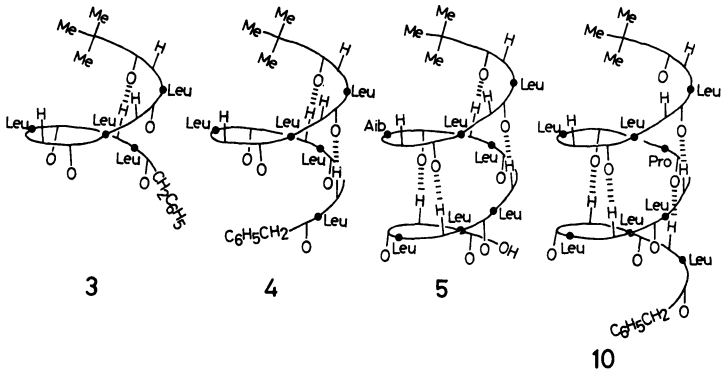

Fig. 8. α-Helical structures for the peptides **3**—**5** and **10**.

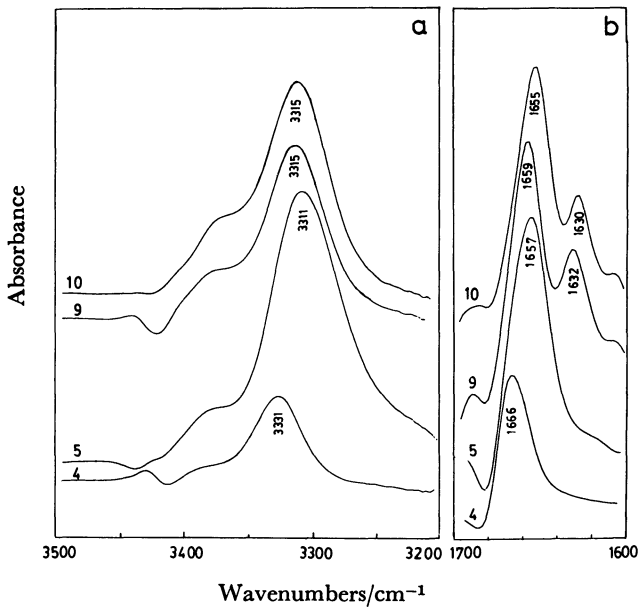

Fig. 9. IR difference spectra between the tetrapeptide **3** and the peptides **4**, **5**, **9**, and **10** (1.0 mM). a: The amide A region, b: the amide I region.

behaviors are just the same with those of the oligo(Leu)s containing the Aib residue as reported previously.[6,7] Furthermore, although the tripeptide **2** possesses low amounts of intramolecularly hydrogen-bonded folded species, the IR difference spectra between the tripeptide **2** and the peptides **3**—**5** and **10** are also examined in the amide A region (Fig. 7). When we also presume the α-helical structures for the peptides **3**—**5** and **10** (Fig. 8), the three free N–H groups in the helical structures for each peptide look like to be almost compensated with the three N–H groups of the peptide **2**. Small positive and negative peaks around 3440 and 3420 cm^{-1} in Figs. 5 and 7 are probably caused by the difference in the band positions of the free N–H groups.

In the α-helical structures for the peptides **3**—**10** in Figs.6 and 8, three peptide carbonyl groups and one C-terminal ester carbonyl group are free from hydrogen bonding. Therefore, the IR difference spectra between the tetrapeptide **3** and the peptides **4**, **5**, **9**, and **10** are useful to investigate the absorption bands assigned to hydrogen-bonded peptide carbonyl groups. Figures 9a and 9b show their IR difference spectra in the amide A and amide I regions, respectively. Except for the peptide **9**, the three free N–H groups in the presumed α-helices for the peptides **4**, **5**, and **10** are again almost compensated with the free N–H groups in the presumed α-helix for the peptide **3** (Fig. 9a). On the other hand, the peptide **9** has only two free N–H groups in the α-helical structure, and the IR difference spectrum between the peptides **3** and **9** shows a negative band at 3420 cm^{-1}. Shoulder bands around 3370 cm^{-1} in Fig. 9a are probably explained by the difference in the band positions of the hydrogen-bonded N–H groups or the contribution of other intramolecularly hydrogen-bonded folded-structures. In Fig. 9b, the strong intensity bands at 1666—1655 cm^{-1} in the amide I region also suggest the α-helical structures.[15-18] The bands at 1632—1630 cm^{-1} are probably assigned to the hydrogen-bonded Pro carbonyl groups in the peptides **9** and **10**. The tendency that the hydrogen-bonded absorption bands in the amide A and amide I regions for the peptides **4**, **9**, and **10** shift to lower frequencies with increasing peptide chain length is quite resembled with that for oligo(Leu)s containing the Aib residue[6,7] and the homopeptides Z-Aib_n-OBu^t (n=3—11),[19] which have 3_{10}-helical structures. With respect to the heptapeptide **5** having the Aib residue, the ^1H NMR study also suggests that it has the α-helix in $CDCl_3$.[20]

The conformational analysis of the peptides **3**—**10** mentioned above is based on the assumption that the absorption coefficients in the amide A region of the free urethane and peptide N–H groups in the unordered structures for the peptides **1** and **2** are nearly equal to those absorption coefficients in the α-helical structures for the peptides **3**—**10**. However, the fully developed, stable α-helical structures in Figs. 6 and 8 are not supported by the intensities of

hydrogen-bonded N-H groups (Figs. 5 and 7). In Table 2, the relative intensities for hydrogen-bonded N-H groups of each peptide are summarized when we presume the α-helical structures for the peptides **3**—**10**. With respect to the homopeptides Z-Aib$_n$-OBu^t (n=3—11),[19] it was reported that a fully-developed, stable 3_{10}-helix was formed at about the octamer level. Therefore, we estimated the percentages[21] of α-helical structures for each peptide using the data of the A_H/A_F ratios for Z-Aib$_n$-OBu^t (n=8—11).[19] The results are also summarized in Table 2. The data strongly suggest that each peptide does not exist only in the fully-developed, stable α-helical structure but coexists in intramolecularly hydrogen-bonded folded-structures which have lower absorption coefficients. Since the assumption is probably not suitable that the absorption coefficients in the amide A region of the free urethane and peptide N-H groups in the unordered structures for the peptides **1** and **2** are nearly equal to those absorption coefficients in the α-helical structures for the peptides **3**—**10**, we should assume that conformations of the peptides **3**—**10** are contributed by the α-helical structures, other successively intramolecularly hydrogen-bonded folded-structures, and the unordered structures free from hydrogen bonding. This conclusion may be in compatible with the ^{1}H NMR result of the peptide **5** that its α-helical structure is not so tight around the C-terminal.[20]

High solubility in dichloromethane of the hexapeptide **8** and the higher oligopeptides **5**, **9**, and **10** is in remarkable contrast with the fact that homologous oligo(Leu) counterparts are practically insoluble in dichloromethane due to their β-sheet structures in the solid state.[5,14] The Pro residue in the peptide **10**, for example, clearly disturbs the development of the β-sheet aggregation by the rotation of the tertiary peptide bond plane (Leu-Pro), facilitating solvation of the peptide chain to cause helical folding. Thus, the ability of the Pro residue to promote helical folding in dichloromethane of the octapeptide **10** is attributed to the disturbance of the β-sheet aggregation, that is, the restriction of the values of the backbone dihedral angles ϕ and ψ of the Pro residue due to the low flexibility of the pyrrolidine ring. The fact that the Pro residue promotes helical folding in the peptides **6**—**10** suggests that the Pro residue included in the N-terminal, four amino acid residues of a helical region plays the role of the nucleation site in helical folding in natural proteins. Undoubtedly, the Pro residue can play the role of the nucleation site in helical folding only when the Pro residue is included in the sequence having high potential for helical formation. In fact, it is well recognized that the Pro residue is frequently found in N-terminal portions of helical regions in natural proteins.[11]

Furthermore, in connection with the design of the synthetic route for peptides and proteins based on the solubility prediction method,[2-5] the fact that the Pro residue promotes successive-intramolecular hydrogen bonds in oligopeptides is of particular interest since helical folding in peptide intermediates, which gives rise to the remarkable solubility improvement, is of quite importance for the design of the synthetic route for peptides and proteins. For example, we assume to be able to design successfully the synthetic route for globin proteins which contain several Pro residues in N-terminal portions of helical regions.

Table 2. Relative Intensities and Wave Numbers of Hydrogen-Bonded N-H Absorption Bands and Helical Contents of Each Peptide

Compound	Relative intensities of hydrogen-bonded N-H absorption bands[a]	Helical content/%[a,b]	Wavenumbers of hydrogen-bonded N-H absorption bands/cm^{-1} [c]
3	0.87	64	3344
4	0.84	62	3336
5	1.00	75	3315
6	0.56	42	3332
7	0.73	54	3331
8	0.80	59	3331
9	0.77	57	3325
10	0.85	63	3323

a) Relative intensities of R/n in Figs. 5 and 7, where R is the area of each N-H absorption bands, and n is the number of intramolecular hydrogen bonds of the α-helices formed by each peptide in Figs. 6 and 7. b) Calculated from the equation, helical content/% = R×100/5.5nA, where A is the half area of the free N-H absorption band of the peptide **1** in Fig. 2, and the coefficient, 5.5, is obtained as $2A_H/A_F/n'$ in Fig. 5 of Ref. 19. c) The band positions of each peptide in Figs. 2 and 3.

References

1) Design of the Synthetic Route for Peptides and Proteins. IV. For part III of this series, see M. Narita, T. Ogura, K. Sato, and S. Honda, *Bull. Chem. Soc. Jpn.*, **59**, 2445 (1986). Abbreviations used are the following: Pro, L-proline residue; Leu, L-leucine residue; IR, infrared; CD, circular dichroism; Aib, α-aminoisobutyric acid residue; Boc, *t*-butoxycarbonyl; OBzl, benzyl ester; HPLC, high-performance liquid chromatography; NMR, nuclear magnetic resonance; Z, benzyloxycarbonyl; OBu^t, *t*-butyl ester.

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

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

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

5) M. Narita, T. Ogura, K. Sato, and S. Honda, *Bull. Chem. Soc. Jpn.*, **59**, 2445 (1986).

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

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

8) M. Narita, J.-Y. Chen, H. Sato, and Y. Kim, *Bull. Chem. Soc. Jpn.*, **58**, 2494 (1985).
9) G. E. Schulz and R. H. Schirmer, "Principles of Protein Structure," Springer-Verlag GmbH & Co. KG (1979).
10) 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).
11) P. Y. Chou and G. D. Fasman, *Biochemistry*, **13**, 211, 222 (1974); P. Y. Chou and G. D. Fasman, *Adv. Enzymology*, **47**, 45 (1978).
12) S. Lifson and C. Sander, *Nature*, **282**, 109 (1979).
13) C. Toniolo, G. M. Bonora, M. Mutter, and V. N. R. Pillai, *Makromol. Chem.*, **182**, 1997, 2007 (1981).
14) 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).
15) M. Narita, Y. Tomotake, S. Isokawa, T. Matsuzawa, and T. Miyauchi, *Macromolecules*, **17**, 1903 (1984) and references cited therein.
16) Ch. P. Rao, P. Balaram, and C. N. R. Rao, *Biopolymers*, **22**, 2091 (1983); V. N. R. Pillai and M. Mutter, *Acc. Chem. Res.*, **14**, 122 (1981), and references cited therein.
17) C. Toniolo, G. M. Bonora, M. Palumbo, E. Peggion, and E. S. Stevens, *Biopolymers*, **17**, 1713 (1978).
18) T. Miyazawa and E. R. Blout, *J. Am. Chem. Soc.*, **83**, 712 (1961).
19) C. Toniolo, G. M. Bonora, V. Barone, A. Bavoso, E. Benedetti, B. D. Blasio, P. Grimaldi, F. Lelj, V. Pavone, and C. Pedone, *Macromolecules*, **18**, 895 (1985).
20) The ^{1}H NMR study will be reported elsewhere.
21) The percentages of α-helical structures for each peptide are obtained presuming the equilibrium between the fully developed α-helix and unordered structures.