  Bull. Chem. Soc. Jpn., 62, 3572—3576 (1989) [Vol. 62, No. 11

The β-Sheet Structure-Stabilizing Potential of Twenty Kinds of Amino Acid Residues in Protected Peptides[1)]

Mitsuaki NARITA* and Yoshihisa KOJIMA
Department of Industrial Chemistry, Faculty of Technology, Tokyo University of Agriculture and Technology, Koganei 184
(Received June 14, 1989)

A host-guest approach was devised to investigate the β-sheet structure-stabilizing potential of the 20 common amino acid residues in protected peptides where functional groups in side chains were protected by suitable groups commonly used in peptide synthesis. The potential was evaluated using solvent-titration curves for the disruption of the β-sheet structure of cross-linked polystyrene resin-bound host-guest hexa- and octapeptides in CH_2Cl_2 by HFIP. The β-sheet structure of host-guest hexapeptides, Boc-Val-Ile-X_2-Val-Ile-$NHCH_2$-resins, in which X stands for guest amino acid residues, was disrupted in CH_2Cl_2 by adding increasing amounts of HFIP; the disrupted behaviors were strongly dependent on the nature of guest amino acid residues. A scale for the β-sheet structure-stabilizing potential of the 20 kinds of amino acid residues in CH_2Cl_2 could be derived at 0.25 M of HFIP concentration as follows: Ile, Val ≫ Asn > His(Tos) or His, Gln, Cys(Bzl), Tyr(Bzl) > Phe, Gly, Lys(Z), Arg(Tos), Leu > Thr(Bzl), Ala, Met(O), Glu(OBzl), Trp(CHO) ≫ Ser(Bzl), Asp(OBzl) ≫ Pro. To confirm the scale, their β-sheet structure-stabilizing potential was further examined using host-guest octapeptides, Boc-X_2-Val-Ile-X_2-Val-Ile-$NHCH_2$-resins. The disrupted behaviors of the octapeptides having Gly, Ala, and Thr(Bzl) residues as guests were remarkably different from those of the corresponding hexapeptides; the β-sheet structure-stabilizing potential of the Gly, Ala, and Thr(Bzl) residues became significantly high with an increase in their ratios to host residues. The potential of the Gly, Ala, and Thr(Bzl) residues appears to reflect the ratios of guest amino acid residues to hosts, the repeated sequence, and/or neighbor interactions in the sequence. The significance of the present study in the design of synthetic routes for peptides and proteins is briefly discussed.

One of the most serious obstacles in peptide and protein syntheses is the insolubility of protected peptides in organic solvents, which causes difficulty in successive reactions. In previous papers[2,3)] we proposed a predictive method for the solubility of protected peptides in the highly polar solvents commonly used in peptide synthesis and showed its usefulness for the design of synthetic routes for peptides and proteins. In the method we assumed that the insolubility of protected peptides was caused by a β-sheet aggregation consisting of peptide segments, i.e., sequence units separated by tertiary peptide bonds,[2,4)] as long as or larger than an octa- or nonapeptide sequence and that the β-sheet structure-forming potential of peptide segments was estimated using their average coil conformation $\langle P_c \rangle$ values when a peptide segment is smaller than the critical chain length for the development of a stable helix in the solid state. In the solubility prediction for protected peptides,[2)] the $\langle P_c \rangle$ value was practically not used for a peptide segment but was used for a whole peptide; the adequacy of the above assumption was widely confirmed by the application of the solubility prediction method to various kinds of oligopeptides.[2-8)] Here, the $\langle P_c \rangle$ value of a protected peptide is defined as an arithmetic average of coil conformational parameters, P_c, of the amino acid residues composing the protected peptide. As mentioned in a previous paper,[9)] $\langle P_c \rangle$ values of protected peptides must reflect their β-sheet structure-forming potential because the P_c parameters were determined on the basis of three-dimensional structures of 29 kinds of native globular proteins[2)] and must reflect the nature of the 20 kinds of amino acid residues in native globular proteins. However, it should be emphasized that the P_c parameters are not for the 20 kinds of amino acid residues in protected peptides but for those in native globular proteins. Thus, for the solubility prediction of protected peptides in organic solvents, it is essentially important to evaluate the P_c parameters for the 20 kinds of amino acid residues in protected peptides in organic solvents. Especially, the P_c parameters for amino acid residues having functional groups in side chains should be determined for side chains-protected amino acid residues.

For the solubility prediction of protected peptides, it is also important to evaluate their helix-forming potential when they are at a sufficient peptide chain length for the development of a stable helix in the solid state, since their β-sheet → helix conformational transformation results in a remarkable solubility improvement.[6,10)] In previous papers[9,11,12)] we showed that the average helix and β-sheet conformation values, $\langle P_\alpha \rangle$ and $\langle P_\beta \rangle$, of protected peptides are useful for estimating their potential for the β-sheet → helix transformation with an increasing peptide chain length or with an application of strong shear stress. Again, $\langle P_\alpha \rangle$ and $\langle P_\beta \rangle$ values of protected peptides are defined as arithmetic averages of the helix and β-sheet conformational parameters (P_α and P_β) of the amino acid residues composing them, respectively, which are determined by Chou and Fasman on the basis of three-dimensional structures of 15 kinds of native globular proteins.[13)] It should again be

emphasized that, for the solubility prediction of protected peptides, it is essentially important to evaluate the conformational parameters (P_α and P_β) for the 20 kinds of amino acid residues in protected peptides where functional groups in side chains are protected by suitable groups commonly used in peptide synthesis.

In a recent study[9,14,15] we found that a β-sheet structure of a variety of resin-bound peptides was easily disrupted by adding increasing amounts of HFIP and that the stability of the β-sheet structure was strongly dependent on the amino acid composition of protected peptides. The result suggests that the investigation into the disruption of the β-sheet structure of resin-bound host-guest peptides by HFIP makes it possible to evaluate the β-sheet structure-stabilizing potential of guest amino acid residues in protected peptides. A host-guest approach was introduced by Mutter and coworkers[16] to explore the relationship between the sequence and the secondary structure of protected peptides; it was shown to be useful for the evaluation of the conformational preference of guest amino acid residues in host-guest peptides. Thus, in the present study, the cross-linked polystyrene resin-bound host-guest peptides are designed to be suitable for the evaluation of the β-sheet structure-stabilizing potential of the 20 kinds of guest amino acid residues in protected peptides in CH_2Cl_2. The designed host-guest peptides have a sequence of Boc-Val-Ile-X_2-Val-Ile-$NHCH_2$-resin and Boc-X_2-Val-Ile-X_2-Val-Ile-$NHCH_2$-resin, in which X stands for guest amino acid residues as follows: Ala, Arg(Tos), Asn, Asp(OBzl), Cys(Bzl), Gln, Glu(OBzl), Gly, His(Tos) or His, Ile, Leu, Lys(Z), Met(O), Phe, Ser(Bzl), Thr(Bzl), Trp(CHO), Tyr(Bzl), and Val.

Experimental

Materials. Copoly(styrene-1% divinylbenzene) beads of 200—400 mesh, Bio-Beads S-Xl, were purchased from Bio-Rad Laboratories. The aminomethylation was carried out according to a method described in the literature.[17] The aminomethyl content was 66 μmol g^{-1} of resin. The synthesis of the host-guest peptide resins was performed in CH_2Cl_2 substantially according to the general procedures described previously.[12] For the assembly of host-guest peptides, Boc-Val-Ile-OH and Boc-X-OH were used as carboxyl components and the coupling reactions were carried out using DCC and HOBt as coupling reagents.[18] After a double coupling reaction, the completion of the reaction was checked by a ninhydrin test. When the test was positive, a small amount of HFIP (5 vol% to CH_2Cl_2) was added and the reaction was continued an additional day. When the test was still positive, a triple coupling reaction was performed to force the coupling reaction to completion. A termination reaction using pyridine and acetic anhydride was not performed. Acid hydrolyses of host-guest peptide resins were carried out with propionic acid /12 M HCl (2/1 v/v) (1 M=1 mol dm^{-3}) at 115 °C for 10 days.[19] The amino acid ratios except for Ser, Thr, and Trp residues of acid hydrolysates were in agreement with the calculated values. Acid hydrolyses host-guest peptide resins having Ser(Bzl), Thr(Bzl), and Trp(CHO) residues as guests were examined with propionic acid/3 M TosOH (2/1 v/v) at 115 °C for 10 days;[20] however, they did not proceed sufficiently.

IR Absorption Spectra Measurements. The IR absorption spectra of resin-bound peptides were recorded at room temperature with a JEOL Model JIR-100 FT-IR spectrometer. IR measurements in the swollen state were performed by holding the samples between potassium bromide windows after resin-bound peptides were swollen in CH_2Cl_2 containing a variety of molar concentrations of HFIP.

Results

The host-guest hexa- and octapeptides were prepared by a general procedure[12] and the addition of HFIP in a coupling reaction in CH_2Cl_2 was significantly effective for its completion. The IR absorption spectra of the resin-bound host-guest hexa- and octapeptides swollen in CH_2Cl_2 alone showed strong bands around 3280 cm^{-1} in the amide A region and around 1630 cm^{-1}

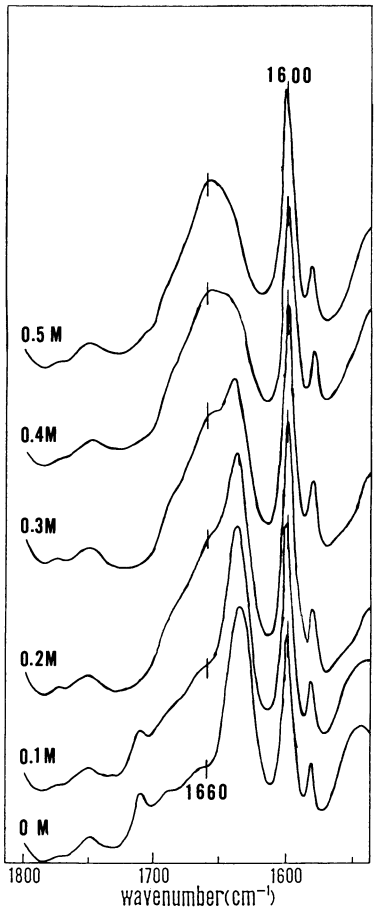

Fig. 1. IR absorption spectra of a resin-bound host-guest hexapeptide (X=Phe) in CH_2Cl_2 containing a variety of molar concentrations of HFIP. The numerals in Fig. 1 indicate molar concentrations of HFIP.

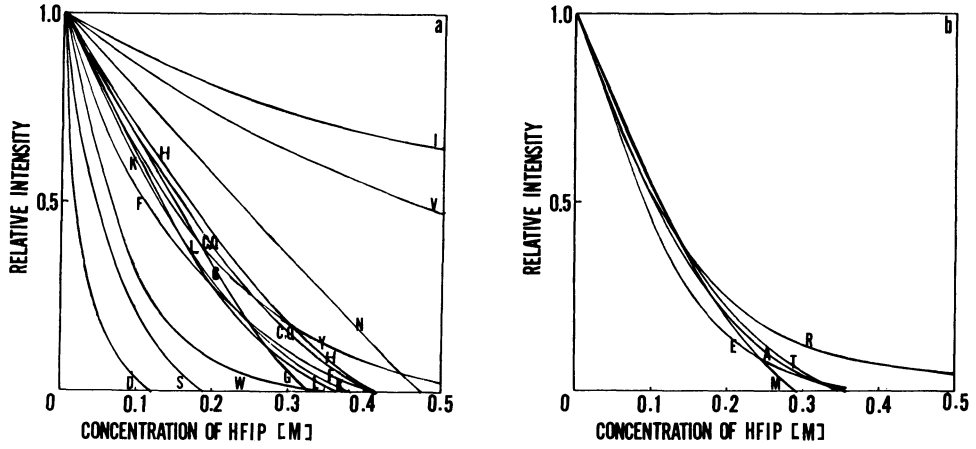

Fig. 2. The solvent-titration curves of resin-bound hexapeptides in CH_2Cl_2 using HFIP as a titrating solvent. A guest amino acid residue for each titration curve is shown in Fig. 2 using a one-letter symbol.[23] a: I=Ile, V=Val, N=Asn, H=His(Tos) or His, C=Cys(Bzl), Q=Gln, Y=Tyr(Bzl), F=Phe, K=Lys(Z), L=Leu, G=Gly, W=Trp(CHO), S=Ser(Bzl), and D=Asp(OBzl) and b: R=Arg(Tos), T=Thr(Bzl), A=Ala, M=Met(O), and E=Glu(OBzl).

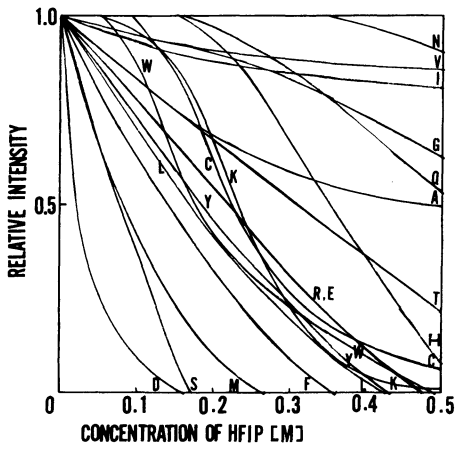

Fig. 3. The solvent-titration curves of resin-bound octapeptides in CH_2Cl_2 using HFIP as a titrating solvent. A guest amino acid residue for each titration curve is shown in Fig. 3 using a one-letter symbol.[23]

in the amide I region, assigned to a β-sheet structure,[12,21] and were accompanied with a weak band around 1660 cm^{-1} (Fig. 1, 0 M), probably assigned to an unordered structure. The disruption of the β-sheet structure by HFIP was investigated using a solvent titration method.[22] It was monitored by a successive decrease in the intensity of the band around 1630 cm^{-1}, together with a successive addition of HFIP. Figure 1 shows the typical IR absorption spectra of resin-bound host-guest hexapeptide (X=Phe) in CH_2Cl_2 containing a variety of molar concentrations of HFIP. The solvent-titration curves of the resin-bound host-guest hexa- and octapeptides in CH_2Cl_2 using HFIP as a titrating solvent are depicted in Figs. 2 and 3, respectively. The relative intensities of the bands around 1630 cm^{-1} shown in Figs. 2 and 3 were obtained using the band at 1603 cm^{-1} due to the aromatic ring of a polystyrene resin support as a standard and were normalized to be 1.0 for each relative intensity in CH_2Cl_2 alone. The relative intensities before the normalization reflected the difference in the intensities of 1660 cm^{-1}-band of host-guest peptide resins and dereased together with an increase in the intensities of the 1660 cm^{-1}-band, indicating the β-sheet structure-forming potential of guest amino acid residues in the resin-bound host-guest peptides. As shown in Figs. 2 and 3, a successive addition of HFIP induced a dramatic decrease in the strong band around 1630 cm^{-1} and an increase in the broad band around 1660 cm^{-1}, indicating that the β-sheet aggregation of the resin-bound hexa- and octapeptides was easily disrupted by increasing the amounts of HFIP.

Discussion

For the evaluation of the β-sheet structure-stabilizing potential of guest amino acid residues in resin-bound host-guest peptides, the completion of a coupling reaction is indispensable at each peptide chain elongation, and the addition of HFIP in a coupling reaction was significantly effective for the quantitative reaction due to the strong β-sheet structure-disrupting potential of HFIP. For the assembly of host-guest peptides, DCC and HOBt were used as coupling reagents. Thus, it is uncertain for the His(Tos) residue to be incorporated in host-guest peptides as a His(Tos) or a His residue since the protecting Tos group for the His residue may be removed by HOBt under the reaction condition.[24] Host-guest peptide resins must also be designed to be

suitable for the evaluation of the β-sheet structure-stabilizing potential of guest amino acid residues. C^{β}-Branched Ile and Val residues are expected to stabilize the β-sheet structure of peptides,[9,13] and the sequence of Ile-X_2-Val resists helix formation due to the repulsion between Ile and Val residues in a helix.[11,14] Thus, host-guest peptides having the sequences of -Val-Ile-X_2-Val-Ile- and -X_2-Val-Ile-X_2-Val-Ile- are suitable for an evaluation of the β-sheet structure-stabilizing potential of guest amino acid residues. In fact, as shown in Fig. 2, the β-sheet structure-disrupted behaviors of resin-bound hexapeptides by HFIP were strongly dependent on the nature of guest amino acid residues. The scale for the β-sheet structure-stabilizing potential of the 20 kinds of guest amino acid residues in hexapeptides can be derived at 0.25 M of the HFIP concentration as follows:

Ile, Val≫Asn>His(Tos) or His, Gln, Cys(Bzl), Tyr(Bzl)>Phe, Gly, Lys(Z), Arg(Tos), Leu>Thr(Bzl), Ala, Met(O), Glu(OBzl), Trp(CHO)≫Ser(Bzl), Asp(OBzl)≫Pro.

The conformational behavior of peptide resins having Pro residues as guests is expected to be different from that of any other one due to the absence of an N-H bond in the Pro residue in peptides;[2,4,15,25] it is obvious that a Pro residue has the lowest β-sheet structure-stabilizing potential in the 20 kinds of amino acid residues. A scale for the β-sheet structure-stabilizing potential of the 20 kinds of guest amino acid residues in octapeptides (Fig. 3) can also be derived at 0.25 M of the HFIP concentration as follows:

Asn, Gln, Val, Ile, Gly, His(Tos) or His≫Ala, Thr(Bzl)≫Glu(OBzl), Arg(Tos), Lys(Z), Cys(Bzl), Tyr(Bzl), Trp(CHO), Leu, Phe≫Met(O), Ser(Bzl), Asp(OBzl)≫Pro.

The disrupted behaviors of the octapeptides having Gly, Ala, and Thr(Bzl) residues as guests are remarkably different from those of the corresponding hexapeptides, and the β-sheet structure-stabilizing potential of the Gly, Ala, and Thr(Bzl) residues becomes significantly high with an increase in their ratios to host residues. The potential appears to reflect the ratios of guest amino acid residues to hosts, the repeated sequence, and/or neighbor interactions in the sequence strongly. The Asn, Gln, and His(Tos) residues in octapeptides also have the same tendency that they stabilize the β-sheet structure of octapeptides.

Here, it is interesting to compare the scale for the β-sheet structure-stabilizing potential of guest amino acid residues in the hexapeptides with the β-sheet structure conformational parameters, P_{β}, obtained by Chou and Fasman,[13] although the media supporting peptide conformations are different from each other (the former, CH_2Cl_2 and the latter, water). The order of the P_{β} parameters is as follows: Met (1.67), Val (1.65), Ile (1.60)≫Cys (1.30), Tyr (1.29), Phe (1.28)>Gln (1.23), Leu (1.22), Thr (1.20), Trp (1.19)≫Ala (0.97), Arg (0.90)>Gly (0.81), Asp (0.80)>Lys (0.74), Ser (0.72), His (0.71), Asn (0.65), Pro (0.62)≫Glu (0.26). For Met(O), Trp(CHO), Arg(Tos), Gly, Lys(Z), His(Tos), Asn, and Glu(OBzl) residues, the former is remarkably different from the latter. Especially, Met(O), Trp(CHO), Arg(Tos), Lys(Z), and Glu(OBzl) residues strongly reflect the influence of their protecting groups and/or media supporting conformations on their β-sheet structure-stabilizing potential. The side-chain protection of Arg, Lys, and Glu residues stabilizes the β-sheet structure of host-guest peptides, while that of Met and Trp residues destabilizes it. Considering the easy disruption of the β-sheet structure of Boc-Val-Gly-Phe-Gly-Leu-Ile-Leu_2-OBzl[26] and Boc-(Val-Gly-Phe-Gly-Leu-Ile-$Leu_2)_n$-Ala-$NHCH_2$-resin (n=1, 2, and 3)[14] by HFIP, the Gly residue in the host-guest hexapeptide appears to reflect the influence of the sequence. It may have a much less stabilizing potential. For Val, Ile, Cys(Bzl), Tyr(Bzl), Phe, Gln, Leu, Thr(Bzl), Ala, Asp(OBzl), and Ser(Bzl) residues, the former is in fair accordance with the latter. As discussed in a previous paper,[9] the P_{β} parameters are assumed to reflect the nature for the β-sheet structure-stabilizing potential of the 20 kinds of corded amino acid residues in native globular proteins. On the other hand, the relative intensities of the 1630 cm^{-1}-band to the 1603 cm^{-1}-band of the host-guest hexapeptides in CH_2Cl_2 alone (data not shown) are assumed to reflect the β-sheet structure-forming potential of the 20 kinds of guest amino acid residues in protected peptides in CH_2Cl_2. A scale for the β-sheet structure-forming potential in the hexapeptides in CH_2Cl_2 alone can be derived: Arg(Tos)>Lys(Z)>Cys(Bzl)>Ile, Val, Leu>Ala, Met(O), Trp(CHO)>Glu(OBzl), Thr(Bzl)>Gly, His(Tos) or His, Phe, Tyr(Bzl)>Gln, Asn>Asp(OBzl), Ser(Bzl). Clearly, the β-sheet structure-forming potential is different from the β-sheet structure-stabilizing potential, and the forming potential of Arg(Tos), Lys(Z), and Cys(Bzl) residues is larger than that of Ile and Val residues.

For the design of synthetic routes for peptides and proteins, an estimation of the β-sheet structure-disrupted behaviors of protected peptides in organic solvents is essentially important since the insolubility of protected peptides due to the β-sheet aggregation causes difficulty in successive reactions. The present study shows that the β-sheet structure-disrupted behaviors of protected peptides are strongly dependent on the nature of amino acid residues composing protected peptides. In previous papers[9,14,15,26] we also demonstrated that the β-sheet structure-disruption by organic solvents is strongly dependent on their electron-acceptor numbers (AN) and electron-donor numbers (DN) and that solvents having large AN or DN values have high potential for the β-sheet structure-disruption. Thus, the scale for the β-sheet

structure-stabilizing potential of the 20 kinds of amino acid residues as well as the β-sheet structure-disrupting potential of solvents is useful for an estimation of the β-sheet structure-disrupted behaviors of protected peptides in organic solvents.

The authors wish to thank Central Glass Co., Ltd., Mitsubishi Kasei Co., Ltd., and Ajinomoto Co., Ltd., for the generous gift of reagents.

References

1) 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: HFIP, 1,1,1,3,3,3-hexafluoro-2-propanol; Boc, *t*-butoxycarbonyl; Tos, *p*-toluenesulfonyl; Bzl, benzyl; Z, benzyloxycarbonyl; OBzl, benzyl ester; CHO, formyl; IR, infrared; DCC, *N,N'*-dicyclohexylcarbodiimide; HOBt, 1*H*-1,2,3,-benzotriazol-1-ol.

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. Fukunaga, A. Wakabayashi, K. Ishikawa, and H. Nakano, *Int. J. Peptide Protein Res.*, **23**, 306 (1984).

5) M. Narita, J.-Y. Chen, H. Sato, and Y. Kim, *Bull. Chem. Soc. Jpn.*, **58**, 2494 (1985); M. Narita, M. Doi, T. Nakai, H. Takegahara, *Int. J. Peptide Protein Res.*, **32**, 200 (1988).

6) 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).

7) M. Narita, M. Doi, K. Kudo, and Y. Terauchi, *Bull. Chem. Soc. Jpn.*, **59**, 3553 (1986).

8) M. Narita, S. Honda, and H. Umeyama, *Bull. Chem. Soc. Jpn.*, **60**, 4127 (1987).

9) M. Narita, S. Isokawa, S. Honda, H. Umeyama, H. Kakei, and S. Obana, *Bull. Chem. Soc. Jpn.*, **62**, 773 (1989).

10) R. Katakai, *Macromolecules*, **14**, 613 (1981); V. N. R. Pillai and M. Mutter, *Acc. Chem. Res.*, **14**, 122 (1981).

11) M. Narita, M. Doi, and H. Takegahara, *Bull. Chem. Soc. Jpn.*, **60**, 2445 (1987); M. Narita, T. Ogura, K. Sato, and S. Honda, *Bull. Chem. Soc. Jpn.*, **59**, 2439, 2445 (1986).

12) M. Narita, S. Isokawa, Y. Tomotake, and S. Nagasawa, *Polym. J.*, **15**, 25 (1983); M. Narita, Y. Tomotake, S. Isokawa, T. Matsuzawa, and T. Miyauchi, *Macromolecules*, **17**, 1903 (1984).

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

14) M. Narita, H. Umeyama, S. Isokawa, S. Honda, C. Sasaki, and H. Kakei, *Bull. Chem. Soc. Jpn.*, **62**, 780 (1989).

15) M. Narita, H. Umeyama, and T. Yoshida, *Bull. Chem. Soc. Jpn.*, **62**, 3577, 3582 (1989).

16) M. Mutter, F. Maser, K.-H. Altmann, C. Toniolo, and G. M. Bonora, *Biopolymers*, **24**, 1057 (1985); M. Mutter and K.-H. Altmann, *Int. J. Peptide Protein Res.*, **26**, 373 (1985).

17) A. R. Mitchell, S. B. H. Kent, M. Engelhard, and R. B. Merrifield, *J. Org. Chem.*, **43**, 2845 (1978).

18) W. König and R. Geiger, *Chem. Ber.*, **103**, 788 (1970).

19) F. C. Westall, J. Scotchler, and A. B. Robinson, *J. Org. Chem.*, **37**, 3363 (1972).

20) T. Y. Liu and Y. H. Chang, *J. Biol. Chem.*, **246**, 2842 (1971).

21) T. Miyazawa and E. R. Blout, *J. Am. Chem. Soc.*, **83**, 712 (1961); T. Miyazawa, "Poly-α-Amino Acids," ed by G. D. Fasman, Marcel Dekker, New York (1967), pp. 69—103.

22) C. Toniolo, G. M. Bonora, M. Mutter, and F. Maser, *J. Chem. Soc., Chem. Commun.*, **1983**, 1298.

23) IUPAC-IUB Commission on Biochemical Nomenclature, *Biochemistry*, **7**, 2703 (1968).

24) T. Fujii and S. Sakakibara, *Bull. Chem. Soc. Jpn.*, **47**, 3146 (1974).

25) M. Narita, S. Isokawa, M. Doi, and R. Wakita, *Bull. Chem. Soc. Jpn.*, **59**, 3547 (1986).

26) M. Narita, S. Honda, and S. Obana, *Bull. Chem. Soc. Jpn.*, **62**, 342 (1989).