  Bull. Chem. Soc. Jpn., 60, 3255—3259 (1987)

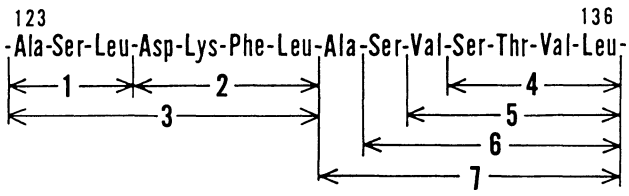
Infrared Absorption Study of Human Hemoglobin α-Chain (123-136) Fragments in Dichloromethane[1)]

Mitsuaki NARITA,* Masamitsu DOI, and Takao NAKAI
Department of Industrial Chemistry, Faculty of Technology,
Tokyo University of Agriculture and Technology, Koganei 184
(Received February 3, 1987)

In connection with the relationship between the conformation and solubility of peptide intermediates having polar side chains, IR spectroscopic conformational analysis of peptide fragments of human hemoglobin α-chain (123-136) was performed in dichloromethane. In dichloromethane, the conformational behavior of the peptide fragments was, in essence, dependent on their amino acid sequences, and Boc-Ser(Bzl)-Thr(Bzl)-Val-Leu-OPac through Boc-Ala-Ser(Bzl)-Val-Ser(Bzl)-Thr(Bzl)-Val-Leu-OPac had predominantly unordered structures including some intramolecular hydrogen bonds, while Boc-Ala-Ser(Bzl)-Leu-Asp(OBzl)-Lys(Z)-Phe-Leu-OPac had successive intramolecular hydrogen bonds, probably corresponding to an α-helical structure. On the other hand, in carbon tetrachloride, the conformation of Boc-Val-Ser(Bzl)-Thr(Bzl)-Val-Leu-OPac and Boc-Ser(Bzl)-Val-Ser(Bzl)-Thr(Bzl)-Val-Leu-OPac had a typical β-sheet structure. These results are discussed on the basis of the amino acid sequences of the peptide fragments.

One of the serious problems in protein synthesis by the classical solution method is the low yields of the fragment condensation reactions, along with the decrease in the solubility of peptide intermediates with the increase in peptide chain length. In fact, Hofmann and coworkers[2)] reported that the azide condensation of tetraeicosapeptide with heptapentacontapeptide gave ribonuclease T_1 (24-104) only in 5% yield. The extreme decline in coupling yields with increasing peptide chain length was presumed to be caused by conformational factors. The reactive ends of the peptides may become sterically hindered because of folding into the peptide chain. Since conformational stability can be expected to increase with increasing chain length, conformation may become the overriding factor influencing the coupling yields. In synthetic studies of peptides and proteins, however, conformational analysis in solution has seldom been performed,[3-5)] since the role of the polar amino acid residues in the conformational behavior of peptide fragments is complicated. This is in contrast with the fact that conformational analyses of homo- and sequential peptides have been widely carried out.[6-8)]

In relation to the problem of the extreme decrease in the solubility and reactivity of peptide intermediates with increasing peptide chain length, we have developed the solubility-improvement[9,10)] and solubility-prediction methods[11)] and proposed the design method of the synthetic route for peptides and proteins based on the solubility prediction method.[12)] In connection with the relationship between the conformation and solubility of peptide intermediates having polar side chains, we here report IR spectroscopic conformational analysis of peptide fragments of human hemoglobin α-chain (123-136) in dichloromethane, which is most commonly used in solid phase peptide synthesis and is the best solvent for coupling reactions when peptide intermediates are soluble. Our concern in this paper is to elucidate the role of polar amino acid residues in the conformational behavior of peptide fragments.



1 Boc-Ala-Ser(Bzl)-Leu-OPac
2 Boc-Asp(OBzl)-Lys(Z)-Phe-Leu-OPac
3 Boc-Ala-Ser(Bzl)-Leu-Asp(OBzl)-Lys(Z)-Phe-Leu-OPac
4 Boc-Ser(Bzl)-Thr(Bzl)-Val-Leu-OBzl
5 Boc-Val-Ser(Bzl)-Thr(Bzl)-Val-Leu-OBzl
6 Boc-Ser(Bzl)-Val-Ser(Bzl)-Thr(Bzl)-Val-Leu-OBzl
7 Boc-Ala-Ser(Bzl)-Val-Ser(Bzl)-Thr(Bzl)-Val-Leu-OBzl

Fig. 1. The amino acid sequence of human hemoglobin α-chain (123-136) and the peptides **1—7** used in this study.

The amino acid sequence of human hemoglobin α-chain (123–136) corresponds to the partial sequence of the helical H region (Fig. 1).[13,14] Figure 1 also illustrates the protected peptide fragments **1—7** used in this study.

Experimental

Materials. The preparation of the samples of human hemoglobin α-chain (123–136) fragments will be reported elsewhere. The purity of the peptides was confirmed by amino acid and elemental analyses. The peptides also gave a single peak on HPLC.

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 Spectra of the Tripeptide 1, the Tetrapeptides 2 and 4, and the Pentapeptide 5 in Dichloromethane. The peptides **1**, **2**, **4**, and **5** are easily soluble in dichloromethane, and their IR absorption spectra were obtained over a wide range of concentration (0.10—10 mM (1 M=1 mol dm^{-3})). In Fig. 2, we show the N-H stretching bands of the peptides **1**, **2**, **4**, and **5** in dichloromethane and clearly see two distinct bands around 3425 and 3360 cm^{-1} over the entire concentration range. The 3425-cm^{-1} band is clearly due to the free N-H stretching vibration, while the 3360-cm^{-1} band is due to a hydrogen-bonded species.[3,4,6–8,15] Since the intensities of these two bands show little dependence on concentrations, the 3360-cm^{-1} band is mainly attributed to the intramolecular hydrogen-bonded species. The band around 3440 cm^{-1} of the peptide **2** is assigned to the free N-H stretching band of the side chain urethane group of the Lys residue. The IR absorption spectra in the amide I region of the peptides **1**, **2**, **4**, and **5** are also illustrated in Fig. 3. Their IR spectra measured over the wide range of concentration show the bands around 1755—1740, 1711—1707, and 1678—1670 cm^{-1}, corresponding to the Pac and Bzl ester carbonyl, Boc urethane carbonyl, and amide carbonyl groups, respectively. The absorption band around 1690 cm^{-1} of the peptide **1** may be attributed to the Pac ketone carbonyl group. The band around 1718 cm^{-1} of the peptide **2** is also assigned to Z urethane carbonyl group free from hydrogen bonding. The amide carbonyl bands of the peptides **1**, **2**, **4**, and **5** also showed little dependence on concentration (0.10—10 mM).

IR Absorption Study of the Hexapeptide 6 in Dichloromethane. The IR absorption spectra in the N-H stretching region of the peptide **6** in dichloromethane (0.10—10 mM) are shown in Fig. 4A. Those in the amide I region are also shown in Fig. 4B. At lower concentrations, the spectral patterns of the peptide **6** are essentially the same as those of the

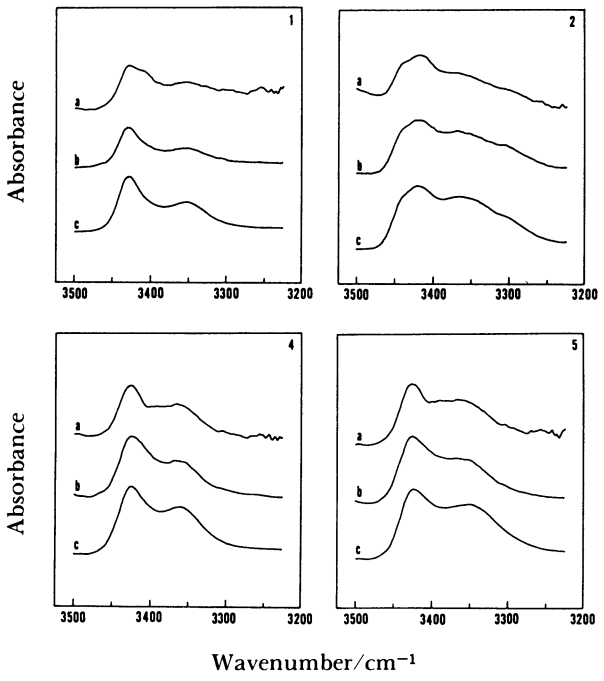


Fig. 2. IR absorption spectra in the amide A region of the tripeptide **1**, the tetrapeptides **2** and **4**, and the pentapeptide **5** in dichloromethane over the wide range of concentrations (0.10—10 mM). **a**: 0.10 mM, **b**: 1.0 mM, **c**: 10 mM.

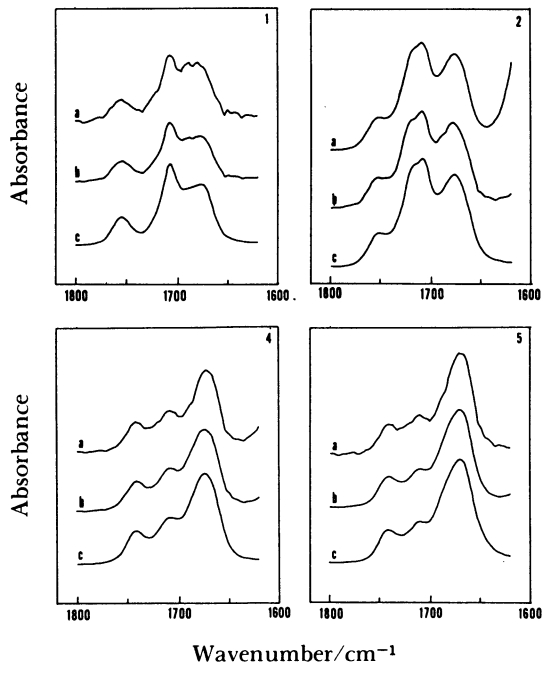


Fig. 3. IR absorption spectra in the amide I region of the tripeptide **1**, the tetrapeptides **2** and **4**, and the pentapeptide **5** in dichloromethane over the wide range of concentrations (0.10—10 mM). **a**: 0.10 mM, **b**: 1.0 mM, **c**: 10 mM.

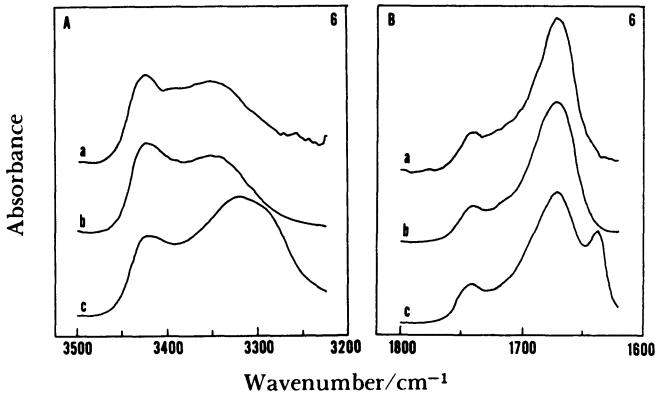


Fig. 4. IR absorption spectra of the hexapeptide **6** in dichloromethane over the wide range of concentrations (0.10—10 mM). A: The amide A region, B: the amide I region. **a**: 0.10 mM, **b**: 1.0 mM, **c**: 10 mM.

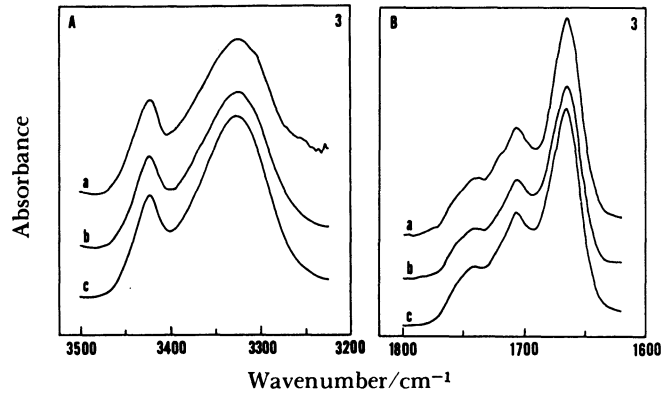


Fig. 5. IR absorption spectra of the heptapeptide **3** in dichloromethane over the wide range of concentrations (0.10—10 mM). A: The amide A region, B: the amide I region. **a**: 0.10 mM, **b**: 1.0 mM, **c**: 10 mM.

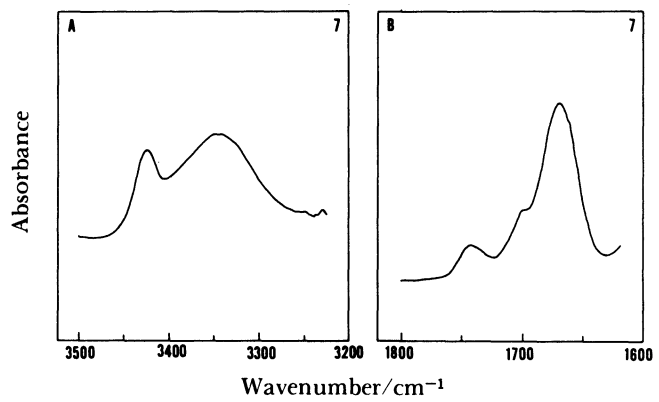


Fig. 6. IR absorption spectra of the heptapeptide **7** in dichloromethane (below 0.1 mM, saturated concentration). A: The amide A region, B: the amide I region.

peptides **4** and **5**, while at the high concentration, the concentration-depending absorption bands appear around 3300 cm^{-1} and at 1637 cm^{-1}, suggesting the onset of a β-sheet structure.[3,4,6-8,15] In fact, the IR absorption spectra of the peptides **5** and **6** in carbon tetrachloride (**5**: 2 mM, **6**: ca. 0.1 mM, saturated concentration) have the bands at 3280—3278, 1695, and 1635—1631 cm^{-1}, indicating an antiparallel β-sheet structure.[3,4,6-8,15,16]

IR Absorption Spectra of the Heptapeptides 3 and 7 in Dichloromethane. The IR absorption spectra in the N-H stretching region of the peptide **3** in dichloromethane (0.10—10 mM) are shown in Fig. 5A. We also see two distinct bands at 3425 cm^{-1} and around 3325 cm^{-1} over the entire concentration range. Again, the 3425-cm^{-1} band is due to the free N-H stretching vibration, while the band around 3325 cm^{-1} is due to a hydrogen-bonded species. The intensity of the hydrogen-bonded band (Fig. 5A) is little dependent on concentration, indicating it is nearly free from intermolecular hydrogen bonding. The positions and intensity ratios of the carbonyl bands (Fig. 5B) are also independent of concentration (0.10—10 mM). Assignment of each band at 1743, 1707, and 1666 cm^{-1} is essentially the same as that of the peptides **1**, **2**, **4**, and **5**. Figures 6A and 6B show the IR absorption spectra in the amide A and amide I regions of the peptide **7** in dichloromethane (below 0.1 mM, saturated concentration), respectively. The free N-H stretching band also appears at 3425 cm^{-1}, while the hydrogen-bonded band of the peptide **7** shifts to a higher frequency (3344 cm^{-1}) than that of the peptide **3** (3325 cm^{-1}). The absorption bands in the amide I region appear at 1743, 1701, and 1670 cm^{-1}.

Discussion

With respect to conformational analysis of Boc-Leu_4-OBzl and Boc-Leu_5-OBzl in dichloromethane, in a previous paper,[17] we showed that the positions and intensity ratios of the bands at 3430—3420 and 3350—3320 cm^{-1} are nearly independent of concentration and that the absorptions of the three free N-H groups in Boc-Leu_4-OBzl and Boc-Leu_5-OBzl are almost clearly comparable with similar absorption in Boc-Leu_3-OBzl. On the basis of these results, we suggested the occurrence of successive intramolecular hydrogen bonds like incipient α-helical structures in Boc-Leu_4-OBzl and Boc-Leu_5-OBzl. In Fig. 2, the positions and intensity ratios of the bands around 3425 and 3360 cm^{-1} are also nearly independent of concentrations, and the 3360-cm^{-1} band is thus to be attributed to intramolecular hydrogen-bonded species. Compared to the intensity ratios of the hydrogen-bonded N-H absorption to the free N-H absorption (A_H/A_F) for Boc-Leu_4-OBzl and Boc-Leu_5-OBzl, however, those for the peptides **2**, **4**, and **5** are relatively small, indicating that the hydrogen-bonded

species of the latter are different from those of the former. The same conformational behavior is also observed in the hexapeptide **6** at lower concentrations (Fig. 4A), being in contrast to the fact that Boc-Leu-Pro-Leu$_4$-OBzl develops successive intramolecular hydrogen bonds in dichloromethane.[17] This different conformational behavior is clearly attributed to the low potential of the peptides **2** and **4**—**6** for the formation of an α-helical structure. The helical structure of the peptides **5** and **6** in dichloromethane may be unstable due to repulsion among C^{β}-branched Val and Thr(Bzl) residues, although the sequence of -Ser-Val-Ser-Thr-Val-Leu- is included in the helical H region of human hemoglobin α-chain.[14] Furthermore, interactions among protected polar side chains and peptide bonds disturb the formation of an α-helix or β-sheet structure, leading to an unordered structure. Thus, the conformations of the peptides **1**, **2**, and **4**—**6** are inferred to be unordered structures including intramolecular hydrogen bonds among protected polar side chains and peptide bonds. At the high concentration (10 mM), the peptide **6** in dichloromethane indicates the onset of a β-sheet structure, and the conformations of the peptides **5** and **6** in carbon tetrachloride of low polar solvent are indeed typical β-sheet structures. The results in carbon tetrachloride exhibit that the structures of the peptides **5** and **6** in carbon tetrachloride are the same as those in the solid state.[18] Further development of interpeptide-chain interactions through hydrogen bonds with increase in peptide chain length gives rise to the severe insolubility of peptides in carbon tetrachloride.

On the other hand, the conformational behavior of the heptapeptide **3** in dichloromethane is in contrast to that of the hexapeptide **6** and the heptapeptide **7**. Over the entire range of concentrations, the positions and spectral patterns of the N-H and C=O absorption bands are independent of concentrations and resemble those of Boc-Leu$_2$-Pro-Leu$_4$-OBzl in dichloromethane,[17] which is inferred to have a large contribution of α-helical structure. For the purpose of elucidating the intramolecularly hydrogen-bonded species, the IR difference spectra between the peptides **3** and **1** and between the peptides **3** and **2** were examined in the amide A and amide I regions (Figs. 7A and 7B). In the difference spectrum between the peptides **3** and **1**, the shoulder band around 3440 cm^{-1} is also assigned to the free N-H stretching band of the side chain urethane group of the Lys residue, and the band around 1720 cm^{-1} to the corresponding urethane carbonyl group. On the other hand, the IR difference spectrum between the peptides **3** and **2** shows a negative band at 1720 cm^{-1}, indicating that the Z urethane carbonyl group of the peptide **3** is subjected to hydrogen bonding more than that of the peptide **2**. Although both IR difference spectra in the amide A region still show the existence of a small amount of the free N-H stretching vibration, the bands at 3323—3321 and 1664—1662 cm^{-1} strongly suggest a large contribution of successive intramolecular hydrogen bonds similar to an α-helical structure.[17] With respect to the conformational behavior of the peptide **7**, the spectrum in the amide A region indicates an increase of hydrogen-bonded species, compared to the peptides **5** and **6**. Since the spectrum in the amide I region shows no band around 1635 cm^{-1}, the increase in the band at 3344 cm^{-1} probably corresponds to the increase in the intramolecular hydrogen-bonded species, suggesting the onset of an α-helical structure around N-terminal portions. The conformational behavior of the peptides **3** and **5**—**7**, which are equal

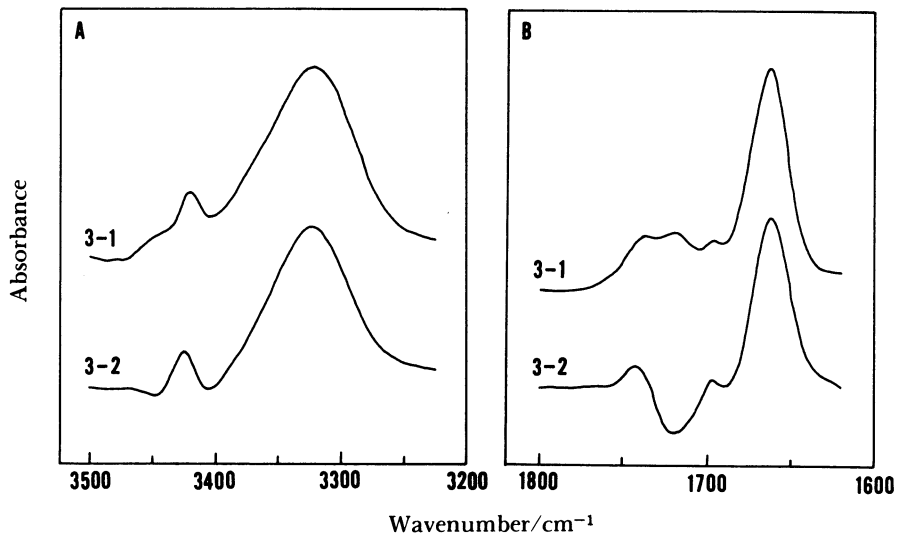


Fig. 7. IR difference spectra between the peptides **3** and **1** and between the peptides **3** and **2** (1.0 mM). A: The amide A region, B: the amide I region.

to or larger than a pentapeptide, clearly reflects their potential for the formation of an α-helical structure and are estimated using the average helix conformation $\langle P_\alpha \rangle$ values[19] of each peptide as discussed for the effect of shear stress on conformational behaviors of the peptides **3** and **5—7** in the solid state.[18]

The present work was partially supported by a Grant-in-Aid for Scientific Research (No. 61540393) from the Ministry of Education, Science and Culture.

References

1) This paper forms Part VII of "Design of the Synthetic Route for Peptides and Proteins Based on the Solubility Prediction Method" series. For part VI of this series, see M. Narita, M. Doi, and H. Takegahara, *Bull. Chem. Soc. Jpn.*, **60**, 2445 (1987). 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 denote the L-configuration. Additional abbreviations used are the following: IR, infrared; Boc, *t*-butoxycarbonyl; Bzl, benzyl; Pac, phenacyl; Z, benzyloxycarbonyl.

2) H. Romovacek, S. R. Dowd, K. Kawasaki, N. Nishi, and K. Hofmann, *J. Am. Chem. Soc.*, **101**, 6081 (1979).

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

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

5) C.-I. Niu, Q.-L. Shi, S.-Z. Yang, Y.-N. Chen, T.-J, Huang, and Z.-X. Lu, *Biopolymers*, **25**, S157 (1986).

6) R. Katakai, *Biopolymers*, **25**, 1 (1986) and references cited therein.

7) U. Narayanan, T. A. Keiderling, G. M. Bonora, and C. Toniolo, *J. Am. Chem. Soc.*, **108**, 2431 (1986) and references cited therein.

8) K.-H. Altmann, A. Florsheimer, and M. Mutter, *Int. J. Peptide Protein Res.*, **27**, 314 (1986), and references cited therein.

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

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

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

12) M. Narita, T. Ogura, K. Sato, and S. Honda, *Bull. Chem. Soc. Jpn.*, **59**, 2433, 2439, 2445 (1986).

13) G. Braunitzer, R. Gehring-Muller, N. Hilschmann, K. Hilsen, G. Hobom, V. Rudloff, and B. Wittmann-Liebold, *Z. Physiol. Chem.*, **325**, 283 (1961).

14) M. F. Perutz, M. G. Rossmann, A. F. Cullis, H. Muirhead, G. Will, and A. C. T. North, *Nature (London)*, **185**, 416 (1960); E. J. Heidner, R. C. Ladner, and M. F. Perutz, *J. Mol. Biol.*, **104**, 707 (1976).

15) M. Narita, S. Isokawa, Y. Tomotake, and S. Nagasawa, *Polymer J.*, **15**, 25 (1983), and references cited therein.

16) T. Miyazawa and E. J. Blout, *J. Am. Chem. Soc.*, **83**, 712 (1961); Yu. N. Chirgadze and N. A. Nevskaya, *Biopolymers*, **15**, 627 (1976).

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

18) M. Narita, M. Doi, and H. Takegahara, *Bull. Chem. Soc. Jpn.*, **60**, 2445 (1987).

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