Infrared Absorption Study of Human Proinsulin C-Peptide Fragments in the Solid State[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 15, 1986)

In order to evaluate the propriety of the solubility prediction method for peptide intermediates having polar side chains, IR absorption spectroscopic analysis of human proinsulin C-peptide fragments was performed in the solid state. The polar amino acid residues involved in the peptide fragments are in the following: Glu(OBzl), Asp(OBzl), Gln, and Ser(Bzl). All peptide fragments except for ones containing the Pro residue exhibit a high potential for the formation of a β-sheet structure, indicating that, even when the polar side chains of the Glu, Asp, and Ser residues are protected by the Bzl group, the coil conformation parameter P_c for each amino acid residue is useful for estimating the potential for the formation of a β-sheet structure of peptide fragments smaller than critical size for development of α-helical structure in the solid state. As in the case of hydrophobic peptides, the β-sheet aggregation clearly plays an important role in reducing solubility of the peptide fragments having polar side chains. The effect of the Pro residue on the disturbance of the β-sheet structure was also observed. The ability of the short-range (polar side chain-backbone) or intraresidue interactions due to the low flexibility of the pyrrolidine ring of the Pro residue to promote helical folding in peptide fragments in the solid state was not observed for the peptide fragments corresponding to helical regions of human proinsulin.

In a previous paper,[2)] the usefulness of the solubility prediction method proposed previously by us[3)] was demonstrated taking the synthesis of human proinsulin C-peptide fragments as an example. On the basis of the solubility prediction of peptide intermediates, subsequently, a strategy suitable for the design of the synthetic route for peptides and proteins was proposed. Insolubility of peptide intermediates is essentially caused by a β-sheet aggregation formed by peptides equal to or larger than an octapeptide sequence level and their solubility is independent on their amino acid sequences and compositions.[4)] Thus, the solubility prediction method is principally based on both the randomness of peptide intermediates in the solid state and the existence of tertiary peptide bond, that is, the disturbance of the β-sheet structure by the rotation of the tertiary peptide bond plane. In order to evaluate the structural randomness of peptide chains, we have used the coil conformation parameter, P_c, for each intact amino acid residue which is determined according to the definition of Chou and Fasman.[5)] One of our purposes in this study is to examine the propriety of evaluation of the structural randomness of peptide fragments with protected polar side chains using the coil conformation parameter, P_c, obtained for each intact amino acid residue.[3)]

The correctness of the concept of the solubility prediction method was moreover verified by the another evidence that solubility improvement on oligopeptides was easily achieved by a β-sheet → helix conformational transition using the ability of the Aib residue to promote helical folding in oligopeptides.[4,6)] The ability of the Aib residue to promote helical folding was distinctly demonstrated using oligo(Leu)s and was clearly attributed to the restriction of the values of the backbone dihedral angles ϕ and ψ of the Aib residue to helical regions due to steric hindrance.[6)] These result also suggested that the restriction of the values of the backbone dihedral angles ϕ and ψ of an amino acid residue to the right-handed α-helical region ($\phi=-57°$ and $\psi=-47°$) was one of important initiation mechanisms of α-helical folding in natural proteins.[6)] In connection with the results, it is quite interesting to investigate the ability of the short-range (polar side chain-backbone) or intraresidue interactions due to the low flexibility of the pyrrolidine ring of the Pro residue to promote helical folding in peptide fragments since two helical regions of human proinsulin are included in fragments of the C-peptide.

In this paper, we report IR absorption spectroscopic analysis of human proinsulin C-peptide fragments with protecting groups in the solid state and investigate the points described above.

Experimental

Materials. Samples of human proinsulin C-peptide fragments are those prepared in the previous paper.[2)] The solid samples of the peptides **1** and **5—9** were obtained by repeated recrystallization from appropriate solvents.[2)] Those of the peptides **2—4** and **10** were obtained by repeated washing with hot methanol or ethanol. The purity of the peptides **1—10** was confirmed by the amino acid and elemental analyses. The peptides **1—3** and **5—10** which were soluble in *N,N*-dimethylformamide also gave a single peak on HPLC.

IR Measurements. The IR absorption spectra of the solid samples were recorded with a Jasco Model DS-701G spectrometer in a nujol mull.

Results

IR Absorption Spectroscopic Analysis of Each Peptide Fragment in the Solid State. The peptide fragments **1—10**, investigated in this study are shown in Fig. 1, which are corresponding to the C-peptide (1-5), (1-8), (6-13), (1-13), (14-19), (9-19), (20-22), (24-28), (20-28), and (24-31), respectively.[2)] The numbers in

1. Boc-Glu(OBzl)Ala Glu(OBzl)Asp(OBzl)Leu-OPac
2. Boc-Glu(OBzl)Ala Glu(OBzl)Asp(OBzl)Leu Gln Val Gly-OPac
3. Boc-Gln Val Gly Gln Val Glu(OBzl)Leu Gly-OPac
4. Boc-Glu(OBzl)Ala Glu(OBzl)Asp(OBzl)Leu Gln Val Gly-Gln Val Glu(OBzl)Leu Gly-OPac
5. Boc-Gly Gly Pro Gly Ala Gly-OPac
6. Boc-Gln Val Glu(OBzl)Leu Gly Gly Gly Pro Gly Ala Gly-OPac
7. Boc-Ser(Bzl)Leu Gln-OPac
8. Boc-Leu Ala Leu Glu(OBzl)Gly-OPac
9. Boc-Ser(Bzl)Leu Gln Pro Leu Ala Leu Glu(OBzl)Gly-OPac
10. Boc-Leu Ala Leu Glu(OBzl)Gly Ser(Bzl Leu Gln-OPac

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

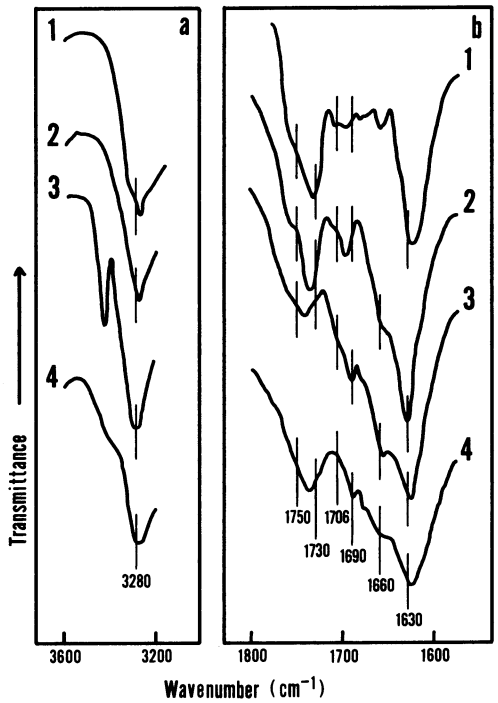

Fig. 2. IR absorption spectra in the solid state of the peptides **1**—**4**. a: The amide A region; b: the amide I region.

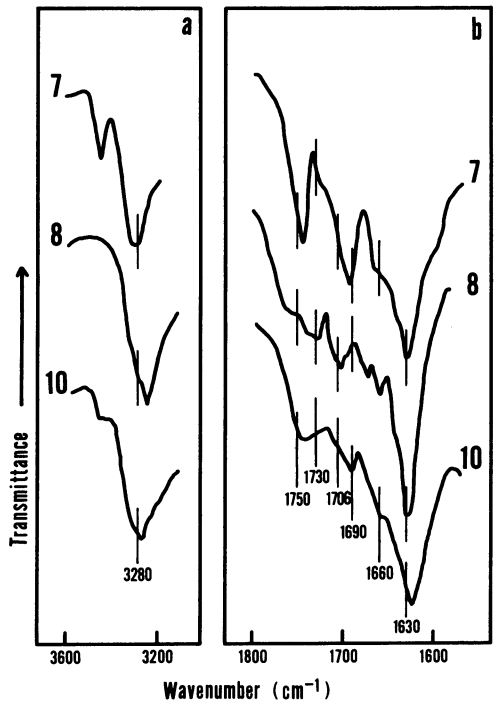

Fig. 3. IR absorption spectra in the solid state of the peptides **7**, **8**, and **10**. a: The amide A region; b: the amide I region.

parentheses in the following C-peptide represent the positions of the first and last amino acids from the N-terminal in the C-peptide. Figures 2 and 3 show the IR absorption spectra of the peptides **1**—**4**, **7**, **8**, and **10** in the solid state in the most significant spectral regions for the conformational assignments (3600—3200 cm^{-1}, amide A; 1800—1600 cm^{-1}, amide I). All of the peptides show strong bands at 3290—3270 cm^{-1} and 1635—1625 cm^{-1}, assigned to a typical β-sheet structure.[7] The absorption band at 1690 cm^{-1} of the peptides **3**, **4**, and **10** also suggests the antiparrallel β-sheet structure. The absorption band or shoulder around 1750 cm^{-1} is assigned to the Pac ester carbonyl group, while the bands around 1730, and 1706 cm^{-1} are assigned to the Bzl ester- and Boc urethane-carbonyl groups, respectively. These absorption bands were assigned by IR measurement in dichloromethane of the peptides **1**, **7**, and **8**. The Pac ketone carbonyl group is estimated to show the band around 1685 cm^{-1}, overlapped with the amide carbonyl group. Although the absorption band characteristic to the Gln side-chain amide can not clearly be assigned, the band at 1660—1655 cm^{-1} of the peptide **2**—**4**, **7**, and **10** is incidentally assigned to the Gln side-chain amide. Similar assignment was reported for the Asn side-chain amide.[8] Figure 4 represents the IR absorption spectra in the solid state of the peptides **5**, **6**, and **9**, which contain the Pro residue in central positions of the peptide chains. These IR absorption spectra are accompanied by additional strong broad shoulders at 1680—1650 cm^{-1}, indicating the disturbance of β-sheet structures by the rotation of the tertiary peptide bond planes (Gly-Pro and Gln-Pro) as reported in previous papers.[9-13] The peptide **9** happens to have, except for the Pro residue, the very close amino acid sequence with the peptide **10**, clearly showing the effect of the Pro residue in a central position of the peptide chain on the disturbance of the β-sheet structure in the peptide

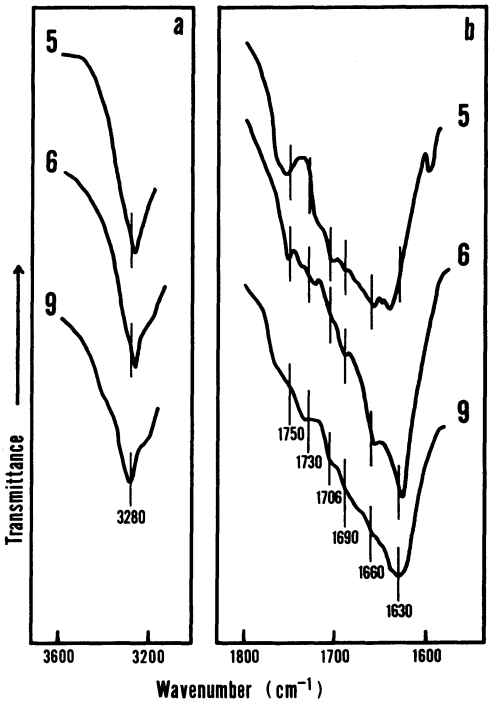

Fig. 4. IR absorption spectra in the solid state of the peptides **5**, **6**, and **9**. a: The amide A region; b: the amide I region.

9.

Discussion

All peptide fragments **1**—**4**, **7**, **8**, and **10**, except for the peptides **5**, **6**, and **9** containing the Pro residue, exhibit a high potential for the formation of the β-sheet structure. Their chain lengths range from a tripeptide to a tridecapeptide. The results indicate that, with respect to peptides having protected or intact polar amino acid residues such as the Glu(OBzl), Asp(OBzl), Gln, and Ser(Bzl) residues, the polar side chain-backbone (short-range) interactions within neighboring residues can not give rise to the disturbance of the formation of the β-sheet structure of the peptides in the solid state. The typical β-sheet structure of the higher oligopeptides **2**—**4** and **10** is also supported by the fact that these peptides have low solubility as reported previously.[2] If the short-range interactions play an important role in determing the conformation of the peptides, it is estimated that they conduct the peptides to random coil, turn, or helical structure, resulting in high solubility of the peptides.

In Table 1, the conformational parameters $<P_c>$, $<P_\alpha>$, and $<P_\beta>$ of each peptide fragment are summarized. The values of $<P_c>$,[3] $<P_\alpha>$,[5] and $<P_\beta>$[5] were determined according to the literatures. The low values of average coil conformation parameters, $<P_c>$, of the peptide fragments **1**—**4** and **7**—**10** suggest that these peptides are highly structured. In actuality, all of the peptide fragments **1**—**4**, **7**, **8**, and **10**, except for the peptide **9** containing the Pro residue, exhibit the high potential for the formation of the β-sheet structure (Figs. 2 and 3). The conformational behaviors of the peptides **1**—**4**, **7**, **8**, and **10** indicate that, even when the polar side chains of the Glu, Asp, and Ser residues are protected by the Bzl group, the coil conformational parameters, P_c, obtained for each intact amino acid residue are useful for estimating the potential for the formation of a β-sheet structure of peptides smaller than critical size for development of α-helical structure in the solid state. On the contrary, the $<P_\beta>$ values of the peptides except for **3** and **7** are very low, indicating a low potential for the formation of the β-sheet structure of these peptides. In the case of the peptides **1**, **2**, **4**, and **8**—**10** which are smaller than a critical size for development of α-helical structure in the solid state, their high $<P_\alpha>$ values indicate a high potential for the formation of the β-sheet structure. Eventually, we can easily estimate the potential for the formation of the β-sheet structure of peptide intermediates only using their $<P_c>$ values.

Table 1. The Values of Average Conformational Parameters $<P_c>$, $<P_\alpha>$, and $<P_\beta>$ of the Peptides **1**—**10**

Peptide	$<P_c>$	$<P_\alpha>$	$<P_\beta>$
1	0.85	1.37	0.70
2	0.84	1.33	0.46
3	0.88	1.07	1.03
4	0.86	1.26	0.68
5	1.22	0.69	0.81
6	1.12	0.90	0.91
7	0.95	1.10	1.06
8	0.78	1.24	0.90
9	0.93	1.12	0.92
10	0.91	1.19	0.96

The β-sheet→helix conformational transition is provoked by the change of intermolecular hydrogen bonds to intramolecular hydrogen bonds, providing high solubility of peptides. In fact, we previously showed a dramatic solubility change of oligo(Leu)s by the complete β-sheet→helix conformational transition using the great ability of the Aib residue to promote helical folding in peptides.[4,6] As well known, the β-sheet→helix conformation transition is dependent on a peptide chain length, and the critical chain length for α-helix formation in the solid state in oligopeptides is not lower than n=10 although it depends on the nature of the constituent amino acid residues.[14] On the other hand, we have demonstrated that the restriction of the values of the backbone dihedral angles ϕ and ψ of the Aib residue to helical regions is an important initiation and stabilization mechanisms of helical folding in peptides.[4,6] In this case, the restriction of the values ϕ and ψ of the Aib residue is apparently attributed to the intraresidue interactions due to steric hindrance,[15] being favorable

to helical folding in peptides. Therefore, if nucleation sites for helical folding in peptide fragments is able to predict as in the case of the Aib residue, it is useful for predicting their solubility. In certain cases, turn structures, which are mainly promoted by backbone-backbone interactions within neighboring residues, are suggested to be nucleation sites for helix formation in proteins.[16] Therefore, the polar side chain-backbone interactions within neighboring residues have the possibility to promote helical folding in oligopeptides. However, unfortunately, the octapeptides **2**, **3**, and **10** having high $<P_\alpha>$ values show the β-sheet structure in the solid state, and no ability of the short-range interactions to promote helical folding in each peptide was observed.

The estimated structure of human proinsulin indicates that the protein contains three short helices, and two of them are included in the part of human proinsulin C-peptide.[17] These helical regions are also included in the peptides **4** and **9**. However, the tridecapeptide **4** also has the β-sheet structure in the solid state, indicating no ability of the short-range interactions to promote helical folding in the peptide **4**. In the sequence analysis of helical regions in proteins, several authors[18-21] have pointed out the tendency of charged side chains to cluster at the ends of helices. Blagdon and Goodman[16] have also suggested that the proper arrangement of charged side chains stabilize helices and they have denoted the route to helix formation as a polar mechanism of initiating helices. Therefore, one of causes of no ability of the peptide **4** to promote helical folding may be attributed to the negation of negative charges at the Glu and Asp residues by their protection with the Bzl group.

In connection with the fact that the most frequently observed conformations for the Pro residue lie in the region of the backbone dihedral angles $\phi=-60°\pm10°$ and $\psi=-30°\pm20°$, and $\phi=-60°\pm10°$, $\psi=-120\pm20°$ due to the low flexibility of the pyrrolidine ring in the Pro residue,[22] in which the former is corresponding to the helical region, it is quite interesting to investigate the ability of the Pro residue to promote helical folding in the peptide **9** due to the low flexibility of the pyrrolidine ring of the Pro residue, which is one of intraresidue interactions as in the case of the Aib residue. Unfortunately, the sample of the peptide **9** in the solid state mainly shows the β-sheet structure, accompanied by an unordered structure which is caused by the rotation of the tertiary peptide bond plane (Gln-Pro) due to steric hindrance. On the other hand, IR conformational analysis of the peptide **9** in dichloromethane solution clearly indicates that the peptide **9** has successive-intramolecular hydrogen bonds, probably corresponding to a stable α-helix in solution.[23] Therefore, another procedure for the preparation of the solid sample of the peptide **9** was examined. But IR conformational behaviors of the sample obtained by reprecipitation by pouring the dichloromethane solution into hexane were essentially the same with the sample before, indicating that the α-helical structure of the peptide **9** in dichloromethane solution was not so stable enough to tolerate the phase transformation from solution into the solid state.

References

1) This paper forms Part II of "Design of the Synthetic Route for Peptides and Proteins Based on the Solubility Prediction Method" series. For part I of this series, see M. Narita, T. Ogura, K. Sato, and S. Honda, *Bull. Chem. Soc. Jpn.*, **59**, 2433 (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: Aib, α-aminoisobutyric acid residue; Bzl, benzyl; Boc, *t*-butoxycarbonyl; Pac, phenacyl.

2) M. Narita, T. Ogura, K. Sato, and S. Honda, *Bull. Chem. Soc. Jpn.*, **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, J.-Y. Chen, H. Sato, and Y. Kim, *Bull. Chem. Soc. Jpn.*, **58**, 2494 (1985).

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

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, *Bull. Chem. Soc. Jpn.*, **58**, 1731 (1985).

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

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

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

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

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

12) S. Isokawa, T. Asakura, and M. Narita, *Macromolecules*, **18**, 871 (1985).

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

14) R. Katakai and Y. Nakayama, *J. Chem. Soc. Chem. Commun.*, **1977**, 805; C. Toniolo, G. M. Bonora, and M. Mutter, *Int. J. Biolog. Macromolecules*, **1**, 188 (1979) and references cited therein.

15) G. R. Marshall, "*Intra-Science Chemistry Report*," Vol. **5**, Ed. by N. Kharasch, Gordon and Breach, New York, pp. 305—316 (1971); A. W. Burgess and S. J. Leach, *Biopolymers*, **12**, 2599 (1973); Y. Paterson, S. M. Rumsey, E. Benedetti, G. Nemethy, and H. A. Scheraga, *J. Am. Chem. Soc.*, **103**, 2947 (1981) and references cited therein.

16) D. E. Blagdon and M. Goodman, *Biopolymers*, **14**, 241 (1975).

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

18) P. Y. Chou and G. D. Fasman, *Biochemistry*, **13**, 222

(1974).
19) P. N. Lewis, F. A. Momany, and H. A. Scheraga, *Proc. Nat. Acad. Sci. U.S.A.*, **68**, 2293 (1971).
20) J. L. Crawford, W. N. Lipscomb, and C. G. Schellman, *Proc. Nat. Acad. Sci. U.S.A.*, **70**, 538 (1971).
21) E. Ralston and J.-L. deCoen, *J. Mol. Biol.*, **83**, 393 (1974).
22) 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, *Biopolymers*, **21**, 1261 (1982).
23) M. Narita, T. Ogura, K. Sato, and S. Honda, *Bull. Chem. Soc. Jpn.*, **59**, 2445 (1986).