Only Weak Dependence of the Protected Peptides Solubility in Organic Solvents on Their Amino Acid Sequence

Sho-kichi OH-UCHI, Jin-Yi YANG, Jin-Shik LEE,
Yuka MURAKAWA, and Mitsuaki NARITA[†]

*Department of Biotechnology, Faculty of Technology, Tokyo University of Agriculture
and Technology, Nakamachi, Koganei, Tokyo 184, Japan*

ABSTRACT: The adequacy of the hypothesis that the protected peptide solubility only weakly depended on their amino acid sequence was confirmed by the evaluation of the β-sheet-structure stability of protected peptides. As the β-sheet-structure stability of protected peptides is related with their solubility, it is investigated by the solvent–titration method using the protected host–guest peptides, Boc-(X-Glu(OBzl)-Ala-Leu-Gly)$_n$-OPac ($n = 1$, 2) which have the same amino acid composition with Boc-(X-Ala-Glu(OBzl)-Leu-Gly)$_n$-OPac ($n = 1$, 2) studied previously. In fact, the results show concurrently that the β-sheet-structure stability of pentapeptides is useful for the estimation of protected peptide solubility in organic solvents and only weakly dependent on their amino acid sequence. It is also little weakly dependent on the solvation mechanism, by which peptides they solvate with an electron acceptor or donor solvent. The β-sheet-structure stability of protected decapeptides is also strongly dependent on their amino acid composition and only weakly dependent on their amino acid sequence. The significance of the hypothesis for the prediction of the protected protein solubility is briefly discussed.

KEY WORDS Protected Peptide Solubility / Peptide Aggregation / Oligopeptide Synthesis / β-Sheet-Structure Stability / Donor Acceptor Interaction / Heteroselective Solvation /

For the protein chemical synthesis in liquid phase, it is indispensable that a protected protein intermediate dissolves in organic solvents for the coupling and the deblocking reactions to proceed completely.[1-3] Actually the insolubility of a protected protein intermediate gives rise often through aggregation of peptide backbones and causes the incompletion of the reactions when protected peptides. Usually, the synthetic route of peptides is determined in trial and error because of the difficulty to design the synthetic route of peptides to avoid the insolubility. For the design of the synthetic route, any new concept is needed for predicting the solubility of protected peptides and proteins. On the basis of the concept, a new solubility prediction method for protected peptides and proteins will be available for the design of polypeptide synthetic routes. The sufficient solvation of peptide backbone results in the disruption of a β-sheet structure of peptides and promotes the complete contact of N- and C-termini of peptide chains with chemical reagents.

So far the solubility has been considered to depend on the amino acid sequence and its prediction, to be difficult. When the solubility of peptides is dependent on their amino acid sequence, we must predict the solubility for so many kinds of peptides, and the solubility prediction becomes very difficult. For example, the number of the decapeptides which consists of ten different amino acids take $10! = 3.628 \times 10^6$. It will be impossible that their solubility in organic solvents is predicted. On the other hand, when it is independent of their sequence and dependent only on their amino acid composition, we need to predict the solubility for the only one peptide according to their amino acid composition. Namely, if it is assumed that the protected peptides solubility does not depend on their amino acid sequence but on their amino acid composition, we can consider the peptides solubility to be very simple.

In the early paper,[4] our solubility prediction method for protected protein intermediates is based on two hypotheses, namely, the concept of "peptide segment separation" (PSS) and only weak dependence of the protected peptides solubility on their amino acid sequence. To confirm the adequacy of the latter hypothesis we previously investigated the solubility of the octapeptides containing three peptide fragments, *i.e.*, Gln-Val-Gly, Asn-Ala-Ile, and Phe-Leu.[5] The results suggested that their solubility and their potential for the β-sheet formation in the solid state were not strongly dependent on their amino acid sequence but on their amino acid composition.

Principally, the insolubility of protected peptides in organic solvents originates from their aggregation through two types of intermolecular hydrogen bonding and van der Waals interactions. Especially, the former causes the formation on an intermolecular hydrogen bonded β-sheet structure and plays an important role in insolubility. The stability of the β-sheet aggregation in organic solvents is influenced by two factors: the natures of the organic solvents and the protected peptides. In regards to the former, we found that the electron-acceptor and -donor numbers of organic solvents were useful for estimating their disrupting potentials of β-sheet struc-

---

ture.[6]

On the other hand, with respect to the nature of the protected peptide,[7] we have demonstrated that the β-sheet-structure-stabilizing potential, SPβi, of 20 kinds of common amino acids was evaluated by the solvent titration method[8-13] using the protected host–guest pentapeptides, Boc-X-Ala-Glu(OBzl)-Leu-Gly-OPac and Boc-X-Val-Asp(OBzl)-Ala-Gly-OPac in which X stands for 20 kinds of guest amino acid residues (Table I). The $\langle SP\beta \rangle$ value which was defined as the arithmetic average of the β-sheet-structure-stabilizing potential, SPβi, of the protected peptide properly reflected the β-sheet-structure stability of peptides up to a hexapeptide in a variety of organic solvents regardless of the amino acid sequence,[7] strongly supporting the adequacy hypothesis that the protected peptides solubility only weakly depends on their amino acid sequence. Nevertheless the chemical structure of the protected host–guest pentapeptides, Boc-X-Ala-Glu(OBzl)-Leu-Gly-OPac, changes only in the substituent of C_β of the N-terminal amino acid residue except for Gly, their β-sheet-structure stability is different remarkably.[6] Its property was evaluated by the solvent titration method.[8-13] Besides the protected peptides solubility in organic solvents was apparently related to their $\langle SP\beta \rangle$ values.[7] And, we demonstrated that the β-sheet-structure stability of protected peptide is almost regardless of the kind of their protecting group.[14] Therefore, in this study, we confirm the hypothesis of weak dependence of the protected peptides solubility on their amino acid sequence by the solvent titration method using Boc-(X-Glu(OBzl)-Ala-Leu-Gly)$_n$-OPac ($n=1$, 2) which have the same amino acid composition with Boc-(X-Ala-Glu(OBzl)-Leu-Gly)$_n$-OPac ($n=1$, 2)[15] in which

X stands for guest amino acid residues as follows: Ala, Asp(OBzl), Gln, Gly, Ile, Leu, Lys(Z), Phe, Ser(Bzl), Trp(CHO), and Val. In the research, the host–guest approach was devised to evaluate the order of the β-sheet-structure-stabilizing potential of the twelve common amino acid residues in the protected peptides, and only weak dependence of the protected peptides solubility on their amino acid sequence was confirmed.
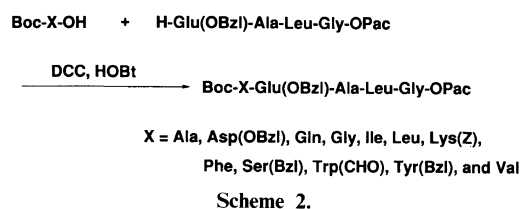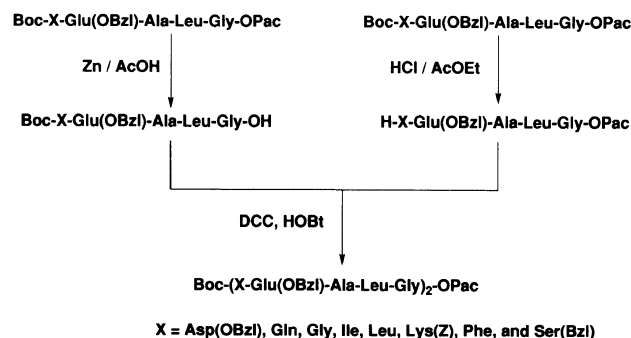
EXPERIMENTAL

Materials

Boc-Ala, -Asp(OBzl), -Gln, -Glu(OBzl), -Gly, -Ile, -Leu, -Lys(Z), -Phe, -Ser(OBzl), -Trp(CHO), -Tyr(Bzl), and -Val were prepared starting with amino acids in general procedures.[16,17] Boc-X-Glu(OBzl)-Ala-Leu-Gly-OPac was obtained by the stepwise chain-elongation in previous studies.[6] The host–guest pentapeptides were prepared in a mixture of CH_2Cl_2 and N,N-dimethylformamide (DMF) by coupling reactions of Boc-X-OH with HCl·H-Glu(OBzl)-Ala-Leu-Gly-OPac, which were obtained from the corresponding Boc-derivatives by treatment with 3.6 M (1 M $= 1$ mol dm^{-3}) HCl/AcOEt as shown in Scheme 1. The coupling reaction was repeated until the Kaiser test became negative, using excess amounts of Boc-X-OH and DCC. The decapeptides were obtained by the fragment condensation using the DCC–HOBt method[18] in liquid phase as illustrated in Scheme 2. After the usual work-up procedures, all the products were purified by recrystallization or precipitation. They gave a single peak on silica gel HPLC by use of a mixture of CH_2Cl_2 and HFIP as an eluent. Acid hydrolyses of the host–guest pentapeptides were performed with propionic acid/12 M HCl (volume ratio, 2/1) at 115°C for 5 days.[19]

IR Absorption Spectra Measurements

The IR absorption spectra of the host–guest peptides in solution or in the suspended state were recorded at room temperature with a JEOL Model JIR-100 FT-IR spectrometer by employing 0.5 mm-path-length cells with sodium chloride windows.[6] The peptides were dissolved

Table I. The β-sheet-structure-stabilizing potential of the 20 amino acid residues

SPβi	Conclusive classification
6	Arg(Mts), Asn, Val
5	Ala, Gln, Gly, His(Bom), Ile
4	Phe, Trp(CHO), Tyr(Bzl)
3	Cys(Bzl), Glu(OBzl), Leu, Lys(Z), Met(O)
2	Ser(Bzl), Thr(Bzl)
1	Asp(OBzl), Pro

Scheme 1.

Boc-X-OH + H-Glu(OBzl)-Ala-Leu-Gly-OPac

DCC, HOBt ⟶ Boc-X-Glu(OBzl)-Ala-Leu-Gly-OPac

X = Ala, Asp(OBzl), Gln, Gly, Ile, Leu, Lys(Z), Phe, Ser(Bzl), Trp(CHO), Tyr(Bzl), and Val

Scheme 2.

X = Asp(OBzl), Gln, Gly, Ile, Leu, Lys(Z), Phe, and Ser(Bzl)

Scheme 3.

or suspended in CH_2Cl_2 or diethyl ether containing a variety of concentrations of DMSO, TFE, or THF. The concentrations of each peptide were kept near 5.0×10^{-2} M for the pentapeptides and near 1.6×10^{-2} M for the decapeptides.

RESULTS AND DISCUSSION

The protected host–guest peptides, Boc-(X-Glu(OBzl)-Ala-Leu-Gly)$_n$-OPac ($n = 1$, 2), having the same amino acid composition and the different amino acid sequence with Boc-(X-Ala-Glu(OBzl)-Leu-Gly)$_n$-OPac ($n = 1$, 2)[6,15] were newly prepared by the general procedure for this study (Schemes 1, 2). The amino acid analysis of the peptides, Boc-X-Glu(OBzl)-Ala-Leu-Gly-OPac, is summarized in Table II. The amino acid ratios of the acid hydrolysates were in good agreement with their calculated values.

By means of the solvent titration method, the β-sheet-structure stability of the peptides was evaluated. The presence of the β-sheet structure of peptides through intermolecular hydrogen bonds was observed by the IR absorption band around $1630 \, cm^{-1}$ in the amide I region in CH_2Cl_2 containing a variety of concentrations of DMSO.[12,13,20] The β-sheet structure of protected pentapeptides, Boc-X-Glu(OBzl)-Ala-Leu-Gly-OPac, was disrupted in CH_2Cl_2 along with increasing amount of DMSO, resulting in the successive decrease of the intensity of the band around $1630 \, cm^{-1}$, and the complete disruption was observed at a high concentration of DMSO. As a typical example, the behavior of the β-sheet-structure disruption of Boc-Val-Glu(OBzl)-Ala-Leu-Gly-OPac is shown in Figure 1. The β-sheet-structure disruption of peptides means their sufficient solvation. Namely, their β-sheet-structure stability is finely related with their solubility. The solvent titration curves of the protected host–guest peptides in CH_2Cl_2 using DMSO as a titrating solvent are depicted for the twelve kind of guest amino acid residues in Figure 2. On the basis of the solvent titration curves of Boc-X-Glu(OBzl)-Ala-Leu-Gly-OPac, the scale for the β-sheet-structure-stabilizing potential of the twelve kinds of guest amino acid residues in the peptides can be derived as follows: Ala > Gly > Gln > Phe > Ser(Bzl) > Ile > Val > Trp(CHO) > Leu > Lys(Z) > Tyr(Bzl) > Asp(OBzl). In the previous result,[6] the scale derived from the solvent titration curves of Boc-X-Ala-Glu(OBzl)-Leu-Gly-OPac

is as follows: Ala > Gly > Gln > Trp(CHO) > Phe > Ser(Bzl) > Ile > Val > Leu > Lys(Z) > Tyr(Bzl) > Asp(OBzl). Except for the Trp(CHO), both of the scales obtained for the above host–guest peptides were in a fine agreement with each other. These results indicate that the β-sheet-structure stability of protected peptides is strongly dependent on their amino acid composition and only weakly dependent on their amino acid sequence. That is to say, the solubility of the protected pentapeptides depends only weakly on the amino acid se-
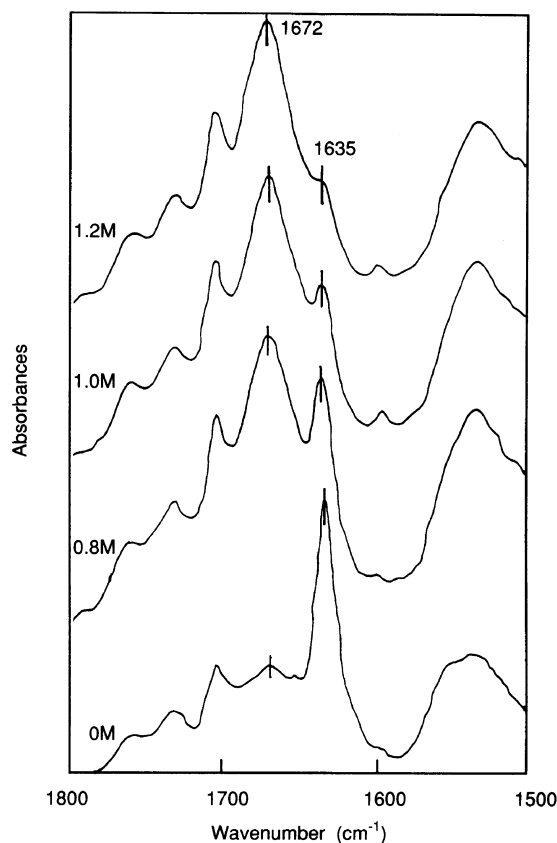

Figure 1. The chemical structure of the protected host–guest pentapeptides, used in this study, Boc-X-Ala-Glu(OBzl)-Leu-Gly-OPac and Boc-X-Glu(OBzl)-Ala-Leu-Gly-OPac.
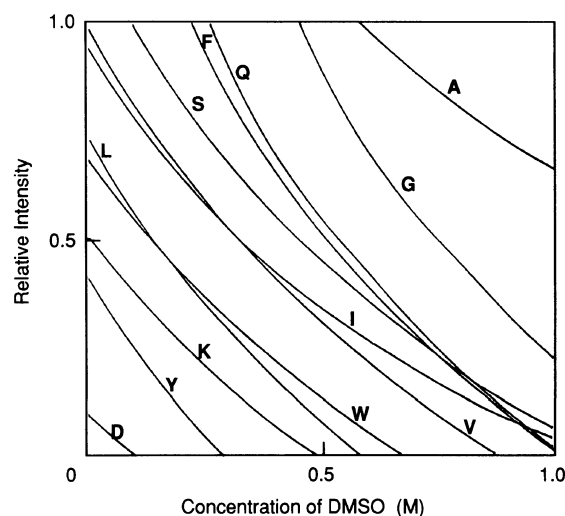

Figure 2. The representative example of typical IR absorption spectra in the amide I region of a host–guest pentapeptide, Boc-Val-Glu(OBzl)-Ala-Leu-Gly-OPac, in the mixed solvent, CH_2Cl_2-DMSO. The concentrations in the figure represent those of DMSO.

Table II. Amino acid analyses of the host–guest peptides, Boc-X-Glu(OBzl)-Ala-Leu-Gly-OPac

X of the host–guest peptides	Found (Calcd)				
	X	Glu	Ala	Leu	Gly
Ala	—	0.89 (1)	1.89 (2)	0.97 (1)	1.00 (1)
Gly	—	0.81 (1)	1.00 (1)	1.10 (1)	2.11 (2)
Ile	0.88 (1)	0.84 (1)	1.00 (1)	1.09 (1)	1.09 (1)
Leu	—	1.26 (1)	0.98 (1)	2.22 (2)	1.00 (1)
Lys(Z)	1.20 (1)	1.02 (1)	0.98 (1)	0.90 (1)	1.00 (1)
Phe	1.01 (1)	1.02 (1)	0.98 (1)	0.88 (1)	1.00 (1)
Ser(Bzl)	0.90 (1)	0.84 (1)	1.00 (1)	1.08 (1)	1.08 (1)
Trp(CHO)	0.91 (1)	1.00 (1)	1.05 (1)	0.89 (1)	1.01 (1)
Tyr(Bzl)	0.95 (1)	1.00 (1)	1.05 (1)	0.87 (1)	1.01 (1)
Val	1.02 (1)	0.80 (1)	0.96 (1)	1.00 (1)	1.04 (1)

Positions Solvated with Donor Solvents

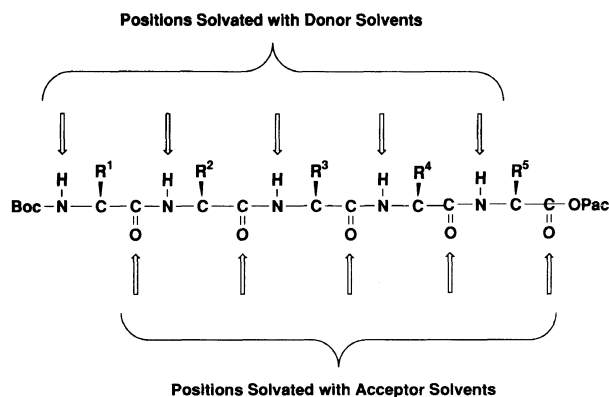**Positions Solvated with Acceptor Solvents**

Figure 3. The solvent titration curves of protected pentapeptides, Boc-X-Glu(OBzl)-Ala-Leu-Gly-OPac, in CH_2Cl_2 using DMSO as a titrating solvent. X is the guest amino acid residues and stands for the following amino acid residues: A = Ala, D = Asp(OBzl), Q = Gln, G = Gly, I = Ile, L = Leu, K = Lys(Z), F = Phe, S = Ser(Bzl), W = Trp(CHO), and V = Val.
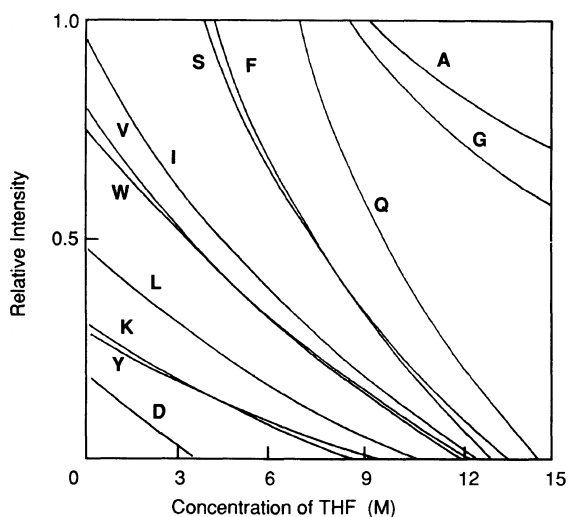

Figure 4. The solvation mechanism of the peptide backbone with a donor or an acceptor solvent.
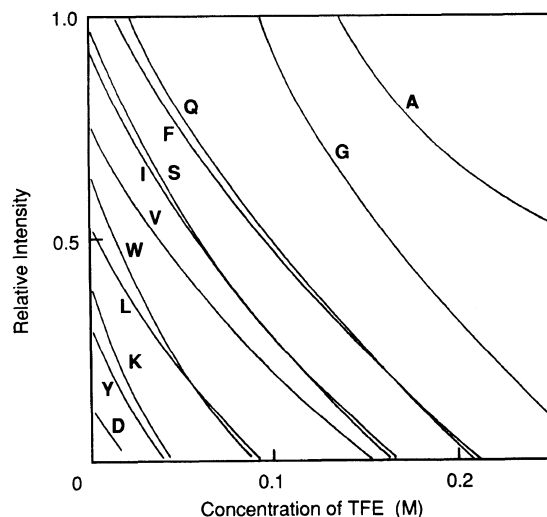

Figure 5. The solvent titration curves of protected pentapeptides. a) Boc-X-Glu(OBzl)-Ala-Leu-Gly-OPac in diethyl ether using THF as a titrating solvent. b) Boc-X-Glu(OBzl)-Ala-Leu-Gly-OPac in CH_2Cl_2 using TFE as a titrating solvent.

quence. But it is also suggested that the solubility of the Trp(CHO) containing peptide significantly depends on the amino acid sequence.

In a previous study,[21] we demonstrated that the solvation mechanism of protected peptides in mixed solvents was explained by the heteroselective solvation through both of the interactions. One is the electron donor-acceptor interaction between the peptide backbone and one component of the mixed solvents. The other is the van der Waals interaction between the peptide side chain and the other component of the solvents. Especially, the former plays the important role in the β-sheet-structure disruption of protected peptides in the mixed solvents. Therefore, we further investigated whether the scale obtained above is influenced or not by the interaction of the peptide backbone with a donor or an acceptor solvent. The solvation mechanism of the peptide backbone is depicted in Figure 3.

For the purpose, the solvent titration was carried out either in diethyl ether using THF as a titrating solvent (Figure 4) or in CH_2Cl_2 using TFE as a titrating solvent

(Figure 5). In these experiments, THF functions as a donor solvent and disrupts the β-sheet structure of the host–guest peptides through the donor–acceptor interaction of the oxygen atom of THF with the hydrogen atom of the peptide bond, while TFE functions as an acceptor solvent and the hydrogen atom of TFE interacts with the oxygen atom of the peptide bond. In a mixture of DMSO and CH_2Cl_2, peptide backbones interact with DMSO as shown in Figure 3, and the scale derived is as follows: Ala > Gly > Phe > Ser(Bzl) > Ile > Val > Trp-(CHO) > Leu > Lys(Z) > Tyr(Bzl) > Asp(OBzl). Thus the order of the β-sheet-structure stability of the protected peptides as derived from the solvent titration curves (Figures 2, 4, and 5) is not changed with the different solvation mechanism by the donor-acceptor interaction. It is expected from the solvent titration curves that the β-sheet-structure-stabilizing potentials of 20 kinds of common amino acid residues are not dependent on the solvation mechanism remarkably.

Next, as the solubility of protected peptides is related with the stability of their β-sheet structure, the stability of the β-sheet structure of Boc-(X-Glu(OBzl)-Ala-Leu-Gly)$_2$-OPac was also evaluated by monitoring their IR absorption band around $1630 cm^{-1}$ in CH_2Cl_2 containing a variety of concentrations of DMSO. The β-sheet structure of the protected decapeptides was disrupted in CH_2Cl_2 along with increasing amount of DMSO, resulting in the successive increase in the intensity of the band around $1670 cm^{-1}$. The behavior of the β-sheet-structure disruption of the decapeptides also depended on their guest amino acid, namely, their amino acid composition. The results were summarized in Table III. The typical IR absorption spectra in the amide I regions of peptides are shown in Figure 6. In order to summarize the results of IR absorption spectra in Table III, the IR absorption spectra of the peptides were classified into five groups as shown in Figure 6. The first group (Figure 6a) showed a strong band around $1630 cm^{-1}$, being
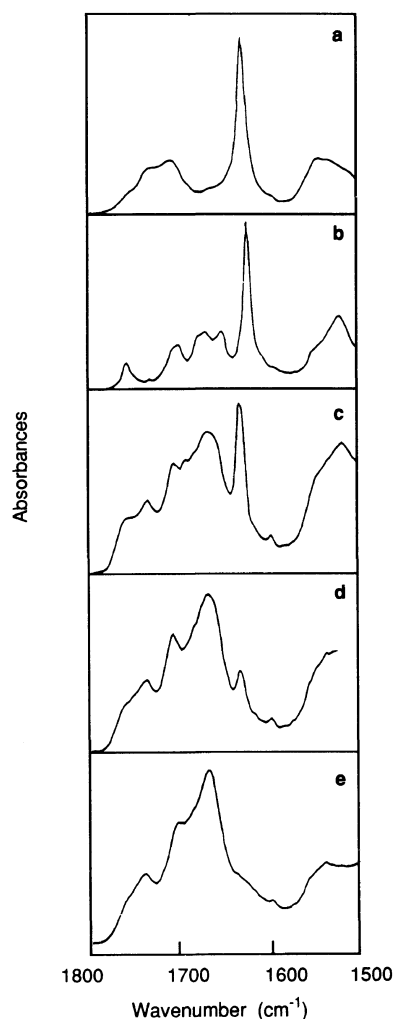
1036

Figure 6. The classification of typical IR absorption spectra in the amide I region of host–guest decapeptides in mixed solvents. Explanations for a—e, see the text.

Table III. The stability of the β-sheet-structure[a] of Boc-(XEALG)$_2$-OPac and Boc-(XAELG)$_2$-OPac[5] in mixed solvents[b]

Protected decapeptides	Mixed solvent		
	1	2	3
(DEALG)$_2$	—	r	—
(KEALG)$_2$	—	r	—
(QEALG)$_2$	β	β(sh)/r	r
(LEALG)$_2$	β	β/r(sh)	r
(SEALG)$_2$	β	β/r(sh)	r
(FEALG)$_2$	β	β/r(sh)	β(sh)/r
(IEALG)$_2$	β	β/r(sh)	β(sh)/r
(GEALG)$_2$	β	β	β
(DAELG)$_2$	—	r	—
(KAELG)$_2$	—	r	—
(QAELG)$_2$	β	β(sh)/r	r
(LAELG)$_2$	β	β/r(sh)	r
(SAELG)$_2$	β	β	r
(FAELG)$_2$	β	β	β(sh)/r
(IAELG)$_2$	β	β	β(sh)/r
(GAELG)$_2$	β	β	β

[a] Structure: β, β-sheet structure; r, random and/or μ-helix structure. [b] Mixed solvents: 1, DMSO : CH$_2$Cl$_2$ = 1 : 4 (v/v); 2, DMSO : CH$_2$Cl$_2$ = 2 : 3 (v/v); 3, DMSO : CH$_2$Cl$_2$ = 3 : 2 (v/v).

accompanied with no or weak shoulder band around 1670 cm^{-1}. They are assigned to β in Table III. The second group (Figure 6b) exhibited a strong band around 1630 cm^{-1} and a medium shoulder band around 1670 cm^{-1} (β/r(sh) in Table III), the third (Figure 6c), medium bands around both 1630 cm^{-1} and 1670 cm^{-1} (β/r in Table II), and the fourth (Figure 6d), a medium shoulder band around 1630 cm^{-1} and a strong band around 1670 cm^{-1} (β(sh)/r in Table III). The fifth (Figure 6e) showed a strong band around 1670 cm^{-1}, being accompanied with no or weak shoulder band around 1630 cm^{-1} (r in Table III).

In the case of the protected decapeptides, Boc-(X-Glu(OBzl)-Ala-Leu-Gly)$_2$-OPac and Boc-(X-Ala-Glu(OBzl)-Leu-Gly)$_2$-OPac, their β-sheet-structure stability is also dependent on the nature of guest amino acid residues, and the orders of the β-sheet-structure-stabilizing potential of the guest amino acid residue in the protected peptides are in agreement with each other as follows: Gly > Ile > Phe > Ser(Bzl) > Leu > Gln > Lys-(Z) > Asp(OBzl). These results also show that the β-sheet-structure stability of protected peptides is dependent on their amino acid composition and weakly dependent on their amino acid sequence. The difference of the orders between penta- and decapeptides was discussed in elsewhere.[16,17]

CONCLUSIONS

These results show concurrently that the β-sheet-structure stability is useful for the estimation of the solubility of protected peptides in organic solvents and weakly dependent on the amino acid sequence and the nature of organic solvents, namely, the solvation mechanism. We can expect that the evaluation of the β-sheet-structure stability of protected peptides on the basis of the hypothesis of only weak dependence of the protected peptides solubility on their amino acid sequence enables us to design the peptide synthetic route to avoid the problem of the protected peptide insolubility. Furthermore, on the basis of the concept of PSS, we can also design the protein synthetic route to avoid the problem of the protected protein insolubility.

REFERENCES

1. R. C. de L. Milton, S. C. F. Milton, and P. A. Adams, *J. Am. Chem. Soc.*, **112**, 6039 (1990).
2. G. B. Fields and C. G. Fields, *J. Am. Chem. Soc.*, **113**, 4202 (1991).
3. T. Johnson, M. Quibell, D. Owen, and R. C. Sheppard, *J. Chem. Soc., Chem. Commun.*, 369 (1993).
4. M. Narita, K. Ishikawa, J.-Y. Chen, and Y. Kim, *Int. J. Peptide Protein Res.*, **24**, 580 (1984).
5. M. Narita, H. Umeyama, and C. Sasaki, *Int. J. Peptide Protein Res.*, **33**, 340 (1989).
6. M. Narita, J.-S. Lee, S. Hayashi, Y. Yamazaki, and M. Hitomi, *Bull. Chem. Soc. Jpn.*, **66**, 494 (1993).
7. M. Narita, J.-S. Lee, S. Hayashi, and M. Hitomi, *Bull. Chem. Soc. Jpn.*, **66**, 489 (1993).
8. P. H. von Dreele, D. Poland, and H. A. Scheraga, *Macromolecules*, **4**, 396 (1971).
9. P. H. von Dreele, N. Lotan, V. S. Ananthanarayanan, R. H. Andreatta, D. Poland, and H. A. Scheraga, *Macromolecules*, **4**, 408 (1971).

10. V. S. Ananthanarayanan, R. H. Andreatta, D. Poland, and H. A. Scheraga, *Macromolecules*, **4**, 417 (1971).

11. M. Sueki, S. Lee, S. P. Powers, J. B. Denton, Y. Konishi, and H. A. Scheraga, *Macromolecules*, **17**, 148 (1984).

12. M. Mutter, F. Maser, K.-H. Altmann, C. Toniolo, and G. M. Bonora, *Biopolymers*, **24**, 1057 (1985).

13. M. Mutter and K. H. Altmann, *Int. J. Peptide Proein Res.*, **26**, 373 (1985).

14. S. Oh-uchi, J.-S. Lee, and M. Narita, *Bull. Chem. Soc. Jpn.*, **69**, 1303 (1996).

15. J.-S. Lee, Y. Murakawa, K. Fujino, and M. Narita, *Bull. Chem.*

16. T. Nagasawa, K. Kuroiwa, K. Narita, and Y. Isowa, *Bull. Chem. Soc. Jpn.*, **46**, 1269 (1973).

17. S. Honda, *Chem. Express*, **6**, 743 (1991).

18. W. König and R. Geiger, *Chem. Ber.*, **99**, 110 (1970).

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

20. T. Miyazawa, "Poly-α-Amino Acids," G. D. Fasman, Ed., Dekker Press, New York, N.Y., 1967, pp 69.

21. M. Narita, J.-S. Lee, S. Hayashi, Y. Yamazaki, and T. Sugiyama, *Bull. Chem. Soc. Jpn.*, **66**, 500 (1993).

Soc. Jpn., **66**, 2283 (1993).