Polymer Journal, Vol. 21, No. 12, pp 1011—1017 (1989)

Theoretical Analysis of the Conformational Preference of the Phe–Phe–Phe Tripeptide Sequence

Masahito OKA,* Yoshihiro BABA, Akihiro KAGEMOTO, and Akio NAKAJIMA†

Departments of General Education and † Applied Chemistry, Osaka Institute of Technology, Omiya, Asahi-ku, Osaka 535, Japan

(Received June 2, 1989)

ABSTRACT: To investigate local interactions in tripeptide sequences composed of amino acids having aromatic side chains, conformational analysis of *N*-acetyl-*N'*-methylamide of the Phe–Phe–Phe tripeptide was theoretically carried out using the conformational energy-minimization procedure. The calculated results show that extended conformations are favorable conformations and β-bend and α-helical conformations are unstable ones for the Phe–Phe–Phe sequence. These results accord with the previous results that extended conformation of the Tyr21–Phe22–Tyr23 sequence in BPTI is stabilized by short-range interaction and also that the tripeptide sequence composed of Phe and Tyr residues found in other proteins also have a tendency to take extended conformations. Experimental results on oligopeptides composed of the Phe residue in the solid state and solutions also support our calculated results. It was also shown that inter-residue interactions indicate minor effects on the mean-square end-to-end distance of a tripeptide.

KEY WORDS Conformational Energy / ECEPP / Phe–Phe–Phe / Tripeptide / Extended Structure / β-Bend / Short-Range Interaction / Protein /

The conformational state of each residue in protein is stabilized by short-, medium-, and long-range interactions. Through analysis using minimization procedure of conformational energy for peptide sequence of BPTI (bovin pancreatic trypsin inhibitor),[1] it was shown that extended conformations were stabilized by short- and long-range interactions. That is, the calculated lowest-energy conformation of half the residue situated in the extended region (Glu7–Tyr10, Ala16–Asn24, and Leu29–Tyr35) corresponds to the native conformation within short-range interactions, and that of another half residue does not correspond to the native conformation within short- and medium-range interactions. But, taking the effects of long-range interactions into consideration, the calculated lowest-energy conformation of all residues in extended region corresponds to the native conformation. The extended conformation of three residues (Tyr21, Phe22, and Tyr23) in BPTI was correctly predicted with short-range interactions, *i.e.*, inter-residue interactions within tripeptide sequence. The sequence Tyr21–Phe22–Tyr23 was found in other homologous proteins[2-4] mutated between Phe and Tyr residues with only one exception, that Tyr21 was mutated by Ile in Russel's viper toxin.[2] It was supposed that these kinds of mutations between Phe and Tyr residues at Tyr21–Phe22–Tyr23 caused no change in the three-dimensional structure of BPTI.[2]

In the previous analysis of tripeptide-

* To whom all correspondence should be addressed.

sequence of BPTI,[1] conformations of the first and third residues of a tripeptide were fixed to those of the refined crystal structure of BPTI,[5] and only ϕ, ψ, and χ of the second residue of the tripeptide were selected as variables in conformational-energy minimizations. That is, as conformations of Arg20 and Phe22 in Arg20–Tyr21–Phe22, Tyr21 and Tyr23 in Tyr21–Phe22–Tyr23, and Phe22 and Asn24 in Phe22–Tyr23–Asn24 were fixed to extended conformations, the conformational preference of the second residue of these tripeptides might be affected by the restricted conformation of the nearest-neighbor residues. It is very interesting to elucidate the conformational preference of tripeptide sequences composed of Phe and Tyr residues, and also to evaluate the effects of short-range interactions in these sequence for stabilizing the extended conformation. In this work, conformational energy calculations were tried on the Phe–Phe–Phe tripeptide as a model of the tripeptide sequence composed of Phe and Tyr residues based on homologous properties between Phe and Tyr residues.[2,6] Theoretical analyasis was carried out over the entire conformational space of the backbone and side chain.

Through theoretical analysis of poly(L-phenylalanine) and poly(L-tyrosine)[6] based on intra-residue interactions, it was shown that extended conformations were most stable for Phe and Tyr residues with favorable intra-residue side-chain/backbone interactions, and also that the characteristic ratios of poly(L-phenylalanine) and poly(L-tyrosine) were larger than those of poly(L-alanine),[6] poly(L-glutamine) and poly(L-glutamic acid),[7] poly(L-leucine),[8] and poly(L-valine).[9] Good agreement between theoretical[6,7] and experimental[7,10-13] results indicate the validity of the treatments based on the intra-residue interactions and further demonstrates that the characteristic ratio can be decided by intra-residue interactions. However, it is supposed that aromatic residues such as Phe and Tyr, which present favorable intra-residue side-chain/backbone interactions, may also present favorable side-chain/side-chain interactions between adjacent residues. It is thus interesting to investigate how characteristic ratios of polypeptides with aromatic side chain are affected by the side-chain/side-chain interactions. To investigate the effects of these interactions within a tripeptide on the mean-square end-to-end distance, the statistically averaged square of the end-to-end distance was calculated using the results of energy minimizations of Ac–Phe–Phe–Phe–NHMe and compared with the mean-square end-to-end distance evaluated by the statistically averaged transformation matrix $\langle T \rangle$ of Phe residue.[6]

THEORETICAL

The nomenclature and conventions adopted are those recomended by IUPAC-IUB nomenclature commission.[14] Conformational energy calculations of the Ac–Phe–Phe–Phe–NHMe tripeptide were carried out with ECEPP (Empirical Conformational Energy Program for Peptides).[15] Energy minimization was carried out with the Powell algorithm[16] until conformational energy did not change more than 0.001 kcal mol^{-1} between successive iterations. During minimizations, all ϕ, ψ, χ^1, and χ^2 were allowed to vary. All other dihedral angles were held fixed at 180°. All combinations of single-residue minima[17] of Phe were used as starting conformations (*i.e.*, 3375 starting conformations).

The same definitions of bend and hydrogen-bonded conformations in ref 18 were used. The classification of the bend type follows Table I of ref 18. The relative conformational energy ΔE_i, end-to-end distance R, and total bend probability P_b are defined in ref 18. Conformational space was devided into 16 regions with the conformational letter codes shown in Figure 1 of ref 17. The statistical average of the conformation-dependent property A for the ensemble is defined by eq 3 of

Table I. Calculated minimum energy conformations[a] of Ac–Phe–Phe–Phe–NHMe

Conformational letter code	ΔE^b / kcal mol^{-1}	v^c	R / Å	ϕ_{Phe1}	ψ_{Phe1}	ϕ_{Phe2}	ψ_{Phe2}	ϕ_{Phe3}	ψ_{Phe3}
ECC	0.00	0.129	12.35	−156	149	−81	101	−86	74
CDE	0.16	0.099	12.81	−87	125	−144	64	−144	155
ECC	0.21	0.091	11.95	−158	165	−61	109	−83	76
CDD	0.31	0.077	11.39	−87	128	−143	53	−141	46
EDE	0.55	0.052	12.28	−151	113	−145	47	−157	159
ECA	0.68	0.041	11.79	−157	166	−64	106	−79	−33
FCC	0.69	0.041	11.87	−74	143	−82	103	−87	74
CDD	0.73	0.038	11.15	−85	113	−128	53	−141	47
EFE	0.74	0.037	13.49	−155	152	−78	152	−141	152
ECA	0.79	0.035	11.91	−156	147	−80	101	−81	−29
EDC	0.91	0.028	11.82	−153	134	−151	31	−109	60
ECF	0.94	0.027	12.28	−157	163	−63	110	−82	148
EFE	1.10	0.021	13.57	−155	152	−80	156	−145	152
EDF	1.12	0.020	12.85	−153	133	−154	35	−112	156
FDC	1.20	0.017	11.73	−100	132	−149	30	−107	65
EEE	1.24	0.016	14.23	−156	153	−147	154	−154	159
FCA	1.36	0.013	11.26	−72	141	−81	104	−83	−29
FDE	1.38	0.013	12.70	−102	131	−152	33	−110	154
FFE	1.44	0.012	13.04	−73	138	−79	153	−129	147
ECC	1.48	0.011	12.26	−156	154	−81	91	−81	82
EEC	1.52	0.010	12.20	−157	150	−154	150	−83	85
CDE	1.55	0.010	12.07	−105	120	−153	35	−152	157
ECF	1.61	0.009	12.50	−157	154	−80	86	−74	144
EDE	1.62	0.009	12.18	−150	123	−155	37	−153	157
FEE	1.64	0.008	13.74	−86	149	−144	153	−159	158
EEE	1.67	0.008	14.22	−155	151	−143	154	−159	158
EFE	1.69	0.008	13.61	−156	141	−96	146	−159	159
CCG	1.80	0.006	10.44	−83	77	−84	76	−149	−53
FFE	1.84	0.006	13.05	−79	148	−76	157	−145	152
EFE	1.90	0.005	13.49	−157	156	−93	147	−158	159
AAA	1.91	0.005	6.00	−69	−31	−61	−40	−83	−39
EEE	1.96	0.005	14.23	−158	161	−156	149	−146	154
EEE	2.00	0.005	14.26	−156	155	−149	158	−159	158

[a] All minima with $\Delta E < 2$ kcal mol^{-1}.
[b] $E_0 = -10.58$ kcal mol^{-1}.
[c] Values at 300 K.

ref 19. The averaged second moment of the end-to-end distance of tripeptide (the distance between C-atom of acetyl group and that of NHMe group) based on the intra-residue interactions was calculated by the following equation.[20]

$$\langle R^2 \rangle_{intra} = l^2[4(E_3 + \langle T \rangle)(E_3 - \langle T \rangle)^{-1} - 2\langle T \rangle(E_3 - \langle T \rangle^4)(E_3 - \langle T \rangle)^{-2}]_{11} \quad (1)$$

where, E_3 is 3×3 unit matrix, the subscript 11 denotes that the 1,1-element of matrix, l is the virtual bond length, respectively. The statistically averaged transformation matrix $\langle T \rangle$ of Phe was obtained in ref 6 taking average over the whole (ϕ, ψ, χ^1, χ^2) space for the 15° interval.

RESULTS AND DISCUSSION

All minimum energy conformations of Ac–

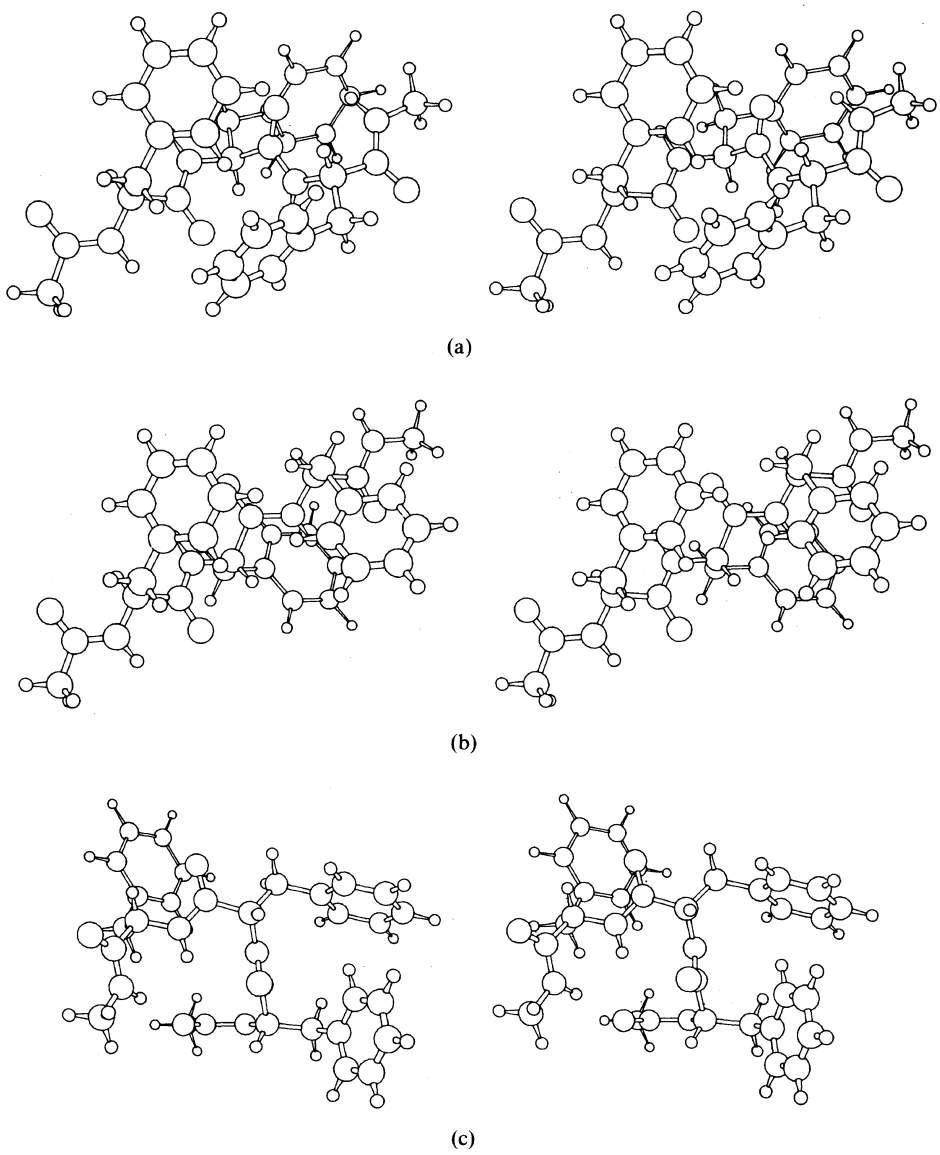

Figure 1. Minimum energy conformations of Ac–Phe–Phe–Phe–NHMe.
(a) Lowest-energy conformation with $\Delta E = 0.00\,\text{kcal mol}^{-1}$. The ϕ and ψ of the three residues are ($-156°$, $149°$, $-81°$, $101°$, $-86°$, $74°$) and the corresponding conformational letter code is ECC.
(b) 16th low-energy conformation with $\Delta E = 1.24\,\text{kcal mol}^{-1}$. The ϕ and ψ of the three residues are ($-156°$, $153°$, $-147°$, $154°$, $-154°$, $159°$) and the corresponding conformational letter code is EEE.
(c) 79th low-energy conformation with $\Delta E = 2.95\,\text{kcal mol}^{-1}$. The ϕ and ψ of the three residues are ($-68°$, $-34°$, $-102°$, $40°$, $-152°$, $-54°$) and the corresponding conformational letter code is ABG.

Phe–Phe–Phe–NHMe with $\Delta E < 2\,\text{kcal mol}^{-1}$ are listed in Table I (Phe1, Phe2, and Phe 3 are abbreviated as the first, second and third residues of this tripeptide). The lowest-energy conformation is the extended conformation with the letter code ECC and (ϕ_1, ψ_1, ϕ_2, ψ_2, ϕ_3, ψ_3)=($-156°$, $149°$, $-81°$, $101°$, $-86°$, $74°$) as shown in Figure 1a. The same EC

conformation at Phe1–Phe2 of a tripeptide was found as the 6th conformation of Ac–Phe–Phe–NHMe (M. Oka and A. Nakajima, unpublished data). However, the same CC conformation at Phe2–Phe3 of tripeptide was not found as a stable conformation of Ac–Phe–Phe–NHMe with $\Delta E < 10\,\mathrm{kcal\,mol^{-1}}$. An analogous CC conformation, which has a different value of $\chi^1{}_1$ ($=-62°$), was found as the 28th conformation of Ac–Phe–Phe–NHMe with $\Delta E = 2.10\,\mathrm{kcal\,mol^{-1}}$ ($\chi^1{}_2 = 178°$ for Ac–Phe–Phe–Phe–NHMe). The ECC conformation also appeared as the 3rd conformation which has almost the same dihedral angles as the lowest-energy ECC conformation with one exception $\chi^1{}_1 = 63°$ ($\chi^1{}_1 = 178°$ for the lowest-energy ECC conformation). The above two ECC conformations each have a (Phe2)CO $\cdots$HN(NHMe) hydrogen bond. This hydrogen bond is a common type for the C conformation.[17] The CC conformation usually has two hydrogen bonds of this type.[21,22] However, the third and most stable ECC conformations lack the (Phe1)CO$\cdots$HN(Phe3) hydrogen bond because of small deviation of the backbone conformation of the Phe2 residue due to interactions between the side chains of Phe1 and Phe3 residues. A typical extended conformation with the letter code EEE was found as the 16th conformation with $\Delta E = 1.24\,\mathrm{kcal\,mol^{-1}}$ as shown in Figure 1b. Three phenyl groups are alternatively situated on the opposite sides of the backbone of the tripeptide, but their orientation has incomplete regularity, *i.e.*, $\chi^1{}_1$ is almost the same as to $\chi^1{}_2$, but not $\chi^1{}_3$. Only one EEE conformation (35th conformation) having the regularity of the side-chain orientations was found with $\Delta E < 3\,\mathrm{kcal\,mol^{-1}}$, but it is energetically unfavorable in comparison with four EEE conformations having irregularities of the side-chain orientations in Table I.

Only 7 bend conformations were found among 87 conformations with $\Delta E < 3\,\mathrm{kcal\,mol^{-1}}$. All of them are bend at Phe1–Phe2 and five of them are bend at Phe2–Phe3, and three of them are double-bend with type III–III which is a part of the α-helical conformation. The typical double-bend conformation (79th conformation with $\Delta E = 2.95\,\mathrm{kcal\,mol^{-1}}$), which is stabilized by the hydrogen bond between (Ac)CO$\cdots$HN(NHMe), is shown in Figure 1c.

Table II. Calculated conformation-dependent properties[a] of Ac–Phe–Phe–Phe–NHMe

	Ac–Phe–Phe–NHMe[b]	Ac–Phe–Phe–Phe–NHMe
P_b		
Phe1–Phe2	0.040	0.013
Phe2–Phe3	—	0.010
$\langle R\rangle_{inter}$ (Å)	9.84	12.25
$\langle R^2\rangle_{inter}$ (Å)	97.9	151.2
$\langle R^2\rangle_{intra}$ (Å)	94.8	157.4

[a] Values at 300 K.
[b] Unpublished data (M. Oka and A. Nakajima).

As shown by the large value of R in Table I, most of the stable conformations of Phe–Phe–Phe tripeptide sequence are extended conformations, that is the Phe–Phe–Phe tripeptide sequence has a tendency to form an extended conformation with intra- and inter-residue interactions of the tripeptide sequence. These properties are clearly different from the previous theoretical results for X–Pro–Y tripeptides.[23]

A detailed comparison of the backbone conformations of each Phe residue in Ac–Phe–Phe–Phe–NHMe with those of Phe residue in Ac–Phe–NHMe[17] indicates that C and D conformations of the second and third residues and F conformation of the first residue of tripeptide are stabilized by the inter-residue interactions, but that the E conformation of the second residue is destabilized by inter-residue interactions. That is, the relative conformational stabilities of the residue in tripeptide are affected by inter-residue interactions. However, the mean-square end-to-end distance based on the intra-residue interac-

Table III. Backbone conformations of tripeptide-sequence composed of Phe and Tyr residues in proteins

Protein	Sequence	Conformational letter code	ϕ_1	ψ_1	ϕ_2	ψ_2	ϕ_3	ψ_3
Trypsin inhibitor	Tyr21 –Phe22 –Tyr23	EEF	−117	135	−138	134	−73	141
Insulin	Phe24 –Phe25 –Tyr26	EED	−155	171	−128	139	−124	105
Thermolysin	Tyr27 –Tyr28 –Tyr29	ECF	−147	150	−89	98	−92	153
Subtilisin	Phe261–Tyr262–Tyr263	BBA	−158	−28	−131	−18	−103	−38

tions $\langle R^2 \rangle_{intra}$ calculated with $\langle T \rangle_{Phe}$ in ref 6 is almost in good agreement with that based on the intra- and inter-residue interactions $\langle R^2 \rangle_{inter}$. That is, the relative stabilities of each conformation of Phe residue are changed by inter-residue interactions, but conformation-dependent properties, especially mean-square end-to-end distance, are not so affected by these inter-residue interactions as whole. This is also supported by the facts that the theoretically evaluated characteristic ratios of polypeptides based on the intra-residue interactions show good agreement with those experimentally evaluated.[6,7]

By sequence analysis of proteins whose three-dimensional structures are shown, three tripeptide sequences composed of Phe and Tyr residues are found in insulin (B-chain, Phe24–Phe25–Tyr26), thermolysin (Tyr27–Tyr28–Tyr29), and subtilisin (Phe261–Tyr262–Tyr263) in addition to Tyr21–Phe22–Tyr23 of BPTI. All the backbone conformations of these four tripeptide-sequences composed of Phe and Tyr residues are shown in Table III. The conformations of Tyr21–Phe22–Tyr23 in BPTI, Phe24–Phe25–Tyr26 in insulin (B-chain), and Tyr27–Tyr28–Tyr29 in thermolysin are extended conformations, *i.e.*, EEF, EED, and ECF conformations, respectively, two former sequences are situated in the central region of β-structure and the last one is situated in the starting point of β-structure. However, the conformation of Phe261–Tyr262–Tyr263 in subtilisin is a non-extended conformation (BBA conformation), *i.e.*, distorted short α-helix situated in the turn region of the backbone of subtilisin.

Similar backbone conformations of three tripeptide-sequences taking extended conformations in crystal state are found in the ensemble of minimum-energy conformations of Ac–Phe–Phe–Phe–NHMe. Two EEF conformations similar to Tyr21–Phe22–Tyr23 of BPTI appeared at the 59 and 80th conformations ($\Delta E = 2.68$ and $2.96\,kcal\,mol^{-1}$, respectively) with $\Delta E < 3\,kcal\,mol^{-1}$. Three EED conformations similar to Phe24–Phe25–Tyr26 of insulin (B-chain) appeared at the 142, 246, and 292th conformations ($\Delta E = 3.55, 4.41$, and $4.63\,kcal\,mol^{-1}$, respectively) with $\Delta E < 5\,kcal\,mol^{-1}$ (no EED conformations with $\Delta E < 3\,kcal\,mol^{-1}$), and 11 analogous type EEE conformations exist with $\Delta E < 3\,kcal\,mol^{-1}$ (the lowest-energy conformation is $\Delta E = 1.24\,kcal\,mol^{-1}$). Four ECF conformations similar to Tyr27–Tyr28–Tyr29 of thermolysin appeared at the 12, 23, 51, and 70th conformations ($\Delta E = 0.94, 1.61, 2.56$, and $2.85\,kcal\,mol^{-1}$, respectively). The lowest-energy and third conformations are ECC conformations analogous to the ECF one. Phe261–Tyr262–Tyr263 of subtilisin has (ϕ_1, ψ_1, ϕ_2, ψ_2, ϕ_3, ψ_3) = ($-158°$, $-28°$, $-131°$, $-18°$, $-103°$, $-38°$), *i.e.*, BBA conformation. No BBA conformations have been found in the energy minima of tripeptides[23] and tetrapeptides.[19,24-29] The (ϕ, ψ) for Phe261 and Tyr262 are those of unstable conformations with intra-residue interactions[6] and the conformational energy of Ac–Phe–Phe–Phe–NHMe which has the same dihedral angles as Phe261–Tyr262–Tyr263 of subtilisin is very

high ($\Delta E = 44.32\,\text{kcal mol}^{-1}$). Thus BBA conformations are stabilized by medium- and long-range interactions in subtilisn. Based on the about discussion, it is supposed that the tripeptide-sequence composed of Phe and Tyr residues has a tendency to form extended structures with internal interactions of the tripeptide-sequence. That is, the native conformations of these sequences in proteins may be assigned as extended conformations with short-range interactions.

IR absorption spectroscopy and CD measurements showed that oligopeptides composed of Phe residue take non-folded conformations in solution[30-32] and β-structures in the solid state,[31-33] and also that oligopeptides composed of Tyr residue take β-structures in solutions and the solid state.[34] These experimental results agree with our calculated results that energetically favorable conformations of Ac–Phe–Phe–Phe–NHMe are extended ones.

REFERENCES

1. M. Oka and H. A. Scheraga, unpublished data.
2. T. E. Creighton, *Nature*, **255**, 743 (1975).
3. A. Ritonja, B. Meloun, and F. Gubensek, *Biochim. Biophys. Acta*, **746**, 138 (1983).
4. C. Liu, T. Wu and T. Lo, *Int. J. Peptide Protein Res.*, **21**, 209 (1983).
5. H. Meirovich and H. A. Scheraga, *Macromolecules*, **14**, 1250 (1981).
6. M. Oka and A. Nakajima, *Polym. J.*, **16**, 693 (1984).
7. M. Oka, T. Hayashi, and A. Nakajima, *Polym. J.*, **17**, 621 (1985).
8. M. Oka, Y. Baba, A. Kagemoto, and A. Nakajima, *Polym. Bull.*, **21**, 377 (1989).
9. M. Oka, Y. Baba, A. Kagemoto, and A. Nakajima, *Polym. Bull.*, **21**, 385 (1989).
10. J. P. Vollmer and G. Spach, *Biopolymers*, **5**, 337 (1967).
11. M. Miyake, S. Akita, A. Teramoto, T. Norisuye, and H. Fujita, *Biopolymers*, **13**, 1173 (1974).
12. T. Norisuye, Ph. D. thesis, Osaka University (1973).
13. G. W. Chen, Ph. D. thesis, Kyoto University (1981).
14. IUPAC-IUB Commission on Biological Nomenclature, Biochemistry, **9**, 3471 (1970).
15. F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, *J. Phys. Chem.*, **79**, 2361 (1975).
16. M. J. D. Powell, *Computer J.*, **7**, 155 (1964).
17. S. S. Zimmerman, M. S. Pottle, G. Nemethy, and H. A. Scheraga, *Macromolecules*, **10**, 1 (1977).
18. S. S. Zimmerman H. A. Scheraga, *Biopolymers*, **16**, 811 (1977).
19. M. Oka, Y. Baba, A. Kagemoto, and A. Nakajima, *Polym. J.*, **21**, 585 (1989).
20. P. J. Flory, "Statistical Mechanics of Chain Molecules," John Wiley & Sons, New York, N.Y., 1969.
21. S. S. Zimmerman and H. A. Scheraga, *Biopolymers*, **17**, 1849 (1978).
22. M. Oka and A. Nakajima, *Polym. J.*, **16**, 553 (1984).
23. M. Oka, G. T. Montelione, and H. A. Scheraga, *J. Am. Chem. Soc.*, **106**, 7959 (1984).
24. J. C. Howard, A. Ali, H. A. Scheraga, and F. A. Momany, *Macromolecules*, **8**, 607 (1975).
25. I. Simon, G. Nemethy, and H. A. Scheraga, *Macromolecules*, **11**, 797 (1978).
26. S. Fitzwater, Z. I. Hodes, and H. A. Scheraga, *Macromolecules*, **11**, 805 (1978).
27. J. S. Anderson and H. A. Scheraga, *Macromolecules*, **11**, 812 (1978).
28. S. S. Zimmerman, R. Baum, and H. A. Scheraga, *Int. J. Peptide Protein Res.*, **19**, 143 (1982).
29. E. Lee, G. Nemethy, H. A. Scheraga, and V. S. Anantharayanan, *Biopolymers*, **23**, 1193 (1984).
30. M. H. Baron, C. de Loze, C. Toniolo, and G. D. Fasman, *Biopolymers*, **17**, 2225 (1978).
31. M. Palumbo, S. DaRin, G. M. Bonora, and C. Toniolo, *Makromol. Chem.*, **177**, 1477 (1976).
32. C. Toniolo, G. M. Bonora, M. Palumbo, E. Peggion, and E. S. Stevens, *Biopolymers*, **17**, 1713 (1978).
33. M. H. Baron, C. de Loze, C. Toniolo, and G. D. Fasman, *Biopolymers*, **18**, 411 (1979).
34. G. M. Bonora, V. Moretto, C. Toniolo, H. Anzinger, and M. Mutter, *Int. J. Peptide Protein Res.*, **21**, 336 (1983).