Polymer Bulletin 38, 227–234 (1997)

Polymer Bulletin

Theoretical conformational analysis of tetrapeptide Ac-Cys-Pro-Ala-Cys-NHMe with disulfide linkage

Jun Yoshimoto[1], Akira Nishinaga[2], Masahito Oka[3,*], Toshio Hayashi[3]

[1] Graduate School of Human Informatics, Nagoya University, Chikusa-ku, Nagoya 464-01, Japan
[2] Department of Applied Chemistry, Osaka Institute of Technology, Asahi-ku, Osaka 535, Japan
[3] Research Institute of Advanced Science and Technology, Osaka Prefecture University, Sakai, Osaka 593, Japan

Received: 11 September 1996/Revised version: 5 November 1996/Accepted: 19 November 1996

Summary

Theoretical conformational analysis was carried out for the acyclic and cyclic tetrapeptides Ac-Cys-Pro-Ala-Cys-NHMe using ECEPP and optimization procedure for investigating the conformational preference of peptides having disulfide linkage. Calculated results indicate that cyclic Ac-Cys-Pro-Ala-Cys-NHMe forms compactly fold conformations with type III-III double bend at the Pro-Ala-Ala portion, and also show fairly good agreement with experimental results of the NMR spectroscopy for the tetrapeptides having Cys-Pro-Ala-Cys sequence.

Introduction

From the viewpoint of designing artificial functional proteins using the conformational preference of peptides with disulfide linkage, in the previous work(1), theoretical conformational analysis was carried out for the acyclic and cyclic tetrapeptides Ac-Cys-Pro-Gly-Cys-NHMe using ECEPP(2). Calculated results indicate that cyclic Ac-Cys-Pro-Gly-Cys-NHMe dominantly forms compactly fold conformations with type II β-bend at the Pro-Gly portion. It means that the -Cys-Pro-Gly-Cys- sequence with S-S linkage is a very useful amino-acid sequence for designing the turn structure with the type II β-bend. In this work, as the second step of investigating the conformational preference of peptides with disulfide linkage, theoretical conformational analysis was carried our for the acyclic and cyclic tetrapeptides Ac-Cys-Pro-Ala-Cys-NHMe using ECEPP and optimization procedure.

Theoretical

All conformational energy calculations were carried out with the energy functions of ECEPP. During minimizations, all ϕ of Pro,(ϕ, ψ, χ^1) of Ala, and (ϕ, ψ, χ^1, χ^2) of cystein (abbreviate as CyH), were allowed to vary. ϕ of Pro was fixed to -75°. All other backbone dihedral angles were fixed to 180°. Conformational energy of tetrapeptide Ac-Ala-Pro-Ala-Ala-NHMe, which is a model peptide of acyclic Ac-CyH-

* Corresponding author

Pro-Ala-CyH-NHMe based on the Ala-residue approximation, was minimized using all combinations of the single-residue minima of Ala and Pro residues(9 and 4, respectively). As the first step of minimization of cyclic tetrapeptide Ac-Cys-Pro-Ala-Cys-NHMe with disulfide bridge, conformational energy of acyclic tetrapeptides, Ac-CyH-Pro-Ala-CyH-NHMe, was minimized using all combinations of the minima of Ac-CyH-Pro-NHMe(obtained in the previous paper(1)) and the single-residue minima of Ala and CyH residues(9 and 47, respectively) as starting conformations. As the final step, conformational energy of cyclic tetrapeptide Ac-Cys-Pro-Ala-Cys-NHMe with disulfide bridge was minimized using stable minima of acyclic Ac-CyH-Pro-Ala-CyH-NHMe.

A bend (occurring at i +1 and i +2th residues) is defined as a conformation in which $R \leqq 7$ Å (R is the distance between i th C^{α} and i +3th C^{α} atoms.) and also classified into eleven types given in Table I of ref 3. A polar hydrogen atom and oxygen or nitrogen atom with an interatomic distance of less than 2.3 Å are regarded to be hydrogen-bonded. Vicinal NH-C^{α}H coupling constants $^{3}J_{NH\text{-}C^{\alpha}H}$ of Cys and Ala residues for Ac-Cys-Pro-Ala-Cys-NHMe were computed using the equation derived by Bystrov *et al*(4) and normalized Boltzmann factor(v) for all minima with ΔE<3kcal/mol. Conformational space is divided into 16 regions with the conformational letter codes shown in Figure 2 of ref 5.

Results and Discussion

There were 734 energy minima for Ac-Ala-Pro-Ala-Ala-NHMe with ΔE<10.0kcal/mol, and 10 of them are shown in Table I. The lowest-energy conformation(Figure 1) is a compactly fold conformation(DAAA conformation) which takes type III-III double bend at Pro-Ala-Ala portion stabilized by the hydrogen bond (Ala_1)CO•••HN(NHMe). The 2nd low-energy conformation(ΔE=0.68kcal/mol) is DAAC one which takes type III-I double bend at Pro-Ala-Ala portion, and the 3rd low-energy conformation(ΔE=1.23kcal/mol) is DAAD one which takes type III-IV double bend at Pro-Ala-Ala portion. Whole backbone structures of the 2nd and 3rd low-energy conformations are very resemble to that of the lowest-energy conformation except that they have not the hydrogen bond (Ala_1)CO•••HN(NHMe). The 4th low-energy conformation(ΔE=1.35kcal/mol) is a non-fold conformation(DCCC one) with continuous C_7 structure formed by three hydrogen bonds (Ala_1)CO•••HN(Ala_3), (Pro)CO•••HN(Ala_4) and (Ala_3)CO•••HN(NHMe). These results indicate that the compactly fold conformations like the lowest-energy one are essentially stable conformations for Ac-Ala-Pro-Ala-Ala-NHMe. In the previous work(1), it was shown that compactly fold conformations are also favorable for Ac-Ala-Pro-Gly-Ala-NHMe, however, fold modes of backbone conformations between them are remarkably different because of the difference in the stable region of (ϕ, ψ)-plane between Gly and Ala residues, e.g., C* region (40° < ϕ <110° , -130° < ψ <-50°) and D* region (110° < ϕ <180° , -110° < ψ <-20°) are stable regions for the Gly residue but not for the Ala residue. So, the low-energy conformations of Ac-Ala-Pro-Gly-Ala-NHMe such as DCC*C, DCC*F and so on are changed to the unstable ones. Then, the DAAA conformation, which is the 11th low-energy one(ΔE=1.05kcal/mol) for Ac-Ala-Pro-Gly-Ala-NHMe, was appeared as the lowest-energy conformation for Ac-Ala-Pro-Ala-Ala-NHMe.

Table I. Calculated Minimum Energy Conformations[a] of Ac-Ala-Pro-Ala-Ala-NHMe

Conformational Letter Code	ΔE^b (kcal/mole)	v^c	Bend Type[d]	ϕ Ala1	ψ Ala1	ϕ Pro	ψ Ala3	ϕ Ala3	ψ Ala4	ϕ Ala4
DAAA	0.00	0.289	III III	-151	80	-23	-68	-42	-77	-46
DAAC	0.68	0.093	III I	-151	79	-45	-66	-51	-94	69
DAAD	1.23	0.036	III IV	-151	80	-21	-72	-42	-151	94
DCCC	1.35	0.030	- -	-152	80	79	-83	79	-84	79
DCAC	1.44	0.026	- I	-152	80	78	-69	-46	-86	76
DCGD	1.46	0.025	IV VII	-153	79	73	-163	-57	-155	45
DFA*G	1.53	0.022	II VII	-153	79	140	54	46	-163	-57
DCA*E	1.60	0.020	II -	-152	79	78	54	50	-156	154
DACC	1.60	0.020	I -	-151	80	-51	-87	72	-85	79
DACG	1.62	0.019	I IV	-151	80	-50	-83	73	-160	-57

[a]All minima with $\Delta E < 1.68$ kcal/mole.
[b]E_0 =-7.56 kcal/mole, $\Delta E = E - E_0$
[c]Normalized Boltzmann factor at 300K.
[d]Bend type for Pro-Ala and Ala-Ala.

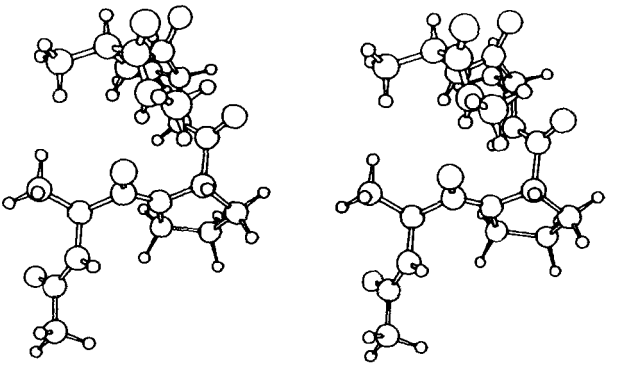

Fig. 1. The lowest-energy conformation(DAAA) of Ac-Ala-Pro-Ala-Ala-NHMe.

There were 14597 energy minima for Ac-CyH-Pro-Ala-CyH-NHMe with ΔE<7.5kcal/mol, and 15 of them are shown in Table II. The lowest-energy conformation is a compactly fold conformation(DAAA conformation) which takes III-III double bend at Pro-Ala-CyH portion stabilized by the hydrogen bond (CyH_1)CO•••HN(NHMe) as shown in Figure 2. All 8 low-energy conformations with ΔE<0.24kcal/mol are also DAAA ones having conformational difference in side-chain conformations of CyH residues. The 2nd type of low-energy conformation is a DACC one taking type III β-bend at Pro-Ala portion and C_7 structure at Ala-CyH_4 portion with hydrogen bonds or energetically favorable hydrogen-bond like interaction, (Pro)CO•••HN(CyH_4) and (Ala)CO•••HN(NHMe).The 3rd type of low-energy conformation is a DCCC one taking non-fold conformation with continuous C_7 structure formed by hydrogen bonds

Table II. Calculated Minimum Energy Conformations[a] of Ac-CyH-Pro-Ala-CyH-NHMe

Conformational Letter Code	ΔE^b (kcal/mole)	v^c	Bend Type[d]	ϕ CyH1	ψ CyH1	ψ Pro	ϕ Ala	ψ Ala	ϕ CyH4	ψ CyH4
DAAA	0.00	0.017	III III	-153	83	-19	-64	-41	-84	-48
DAAA	0.00	0.017	III III	-153	83	-18	-62	-45	-86	-47
DAAA	0.00	0.017	III III	-153	83	-19	-64	-41	-84	-48
DAAA	0.09	0.015	III III	-153	83	-19	-63	-41	-85	-48
DAAA	0.09	0.015	III III	-153	83	-19	-63	-41	-85	-48
DAAA	0.11	0.014	III III	-153	83	-17	-61	-45	-86	-47
DAAA	0.13	0.014	III III	-153	82	-17	-61	-45	-86	-47
DAAA	0.21	0.012	III III	-153	82	-17	-61	-45	-87	-47
DACC	0.25	0.011	III -	-153	84	-45	-82	88	-81	85
DAAA	0.27	0.011	III III	-153	83	-19	-64	-41	-84	-48
DACC	0.29	0.011	III -	-153	84	-46	-83	88	-81	85
DAAA	0.36	0.009	III III	-153	83	-19	-63	-42	-85	-48
DACC	0.38	0.009	III -	-153	84	-47	-83	88	-82	85
DACC	0.44	0.008	III -	-153	84	-47	-83	89	-82	85
DACC	0.45	0.008	III -	-153	84	-45	-85	89	-82	85

[a]All minima with $\Delta E < 0.46$ kcal/mole.
[b]E_0 =-9.09 kcal/mole, ΔE= E-E_0
[c]Normalized Boltzmann factor at 300K.
[d]Bend type for Pro-Ala and Ala-CyH.

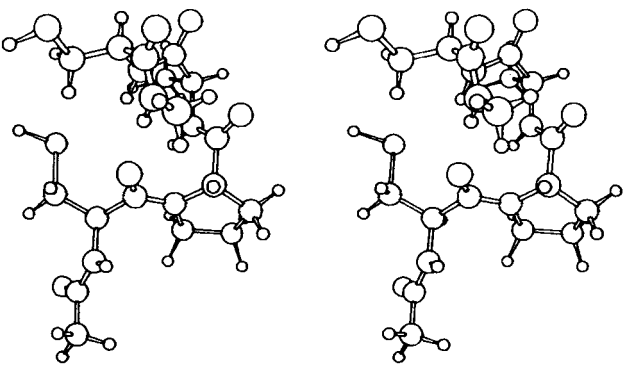

Fig. 2. The lowest-energy conformation(DAAA) of Ac-CyH-Pro-Ala-CyH-NHMe.

(CyH_1)CO•••HN(Ala) and (Pro)CO•••HN(CyH_4) and hydrogen-bond like interaction (Ala)CO•••HN(NHMe). There are 20, 17 and 6 minima with ΔE<1.0kcal/mol corresponding to the DAAA, DACC and DCCC ones, and their occurring probabilities among 59 minima with ΔE<1.0kcal/mol are 0.48, 0.29 and 0.07, respectively. All other 16 minima with ΔE<1.0kcal/mol are also classified into above three types of conformations DAAA(double bend), DACC(bend-C_7 structure) and DCCC(continuous

Table III. Calculated Minimum Energy Conformations[a] of Ac-Cys-Pro-Ala-Cys-NHMe

Conformational Letter Code	ΔE^b (kcal/mole)	V^c	Bend Type[d]	ϕ Cys1	ψ Cys1	ϕ Pro	ϕ Ala	ψ Ala	ϕ Cys4	ψ Cys4
DAAA	0.00	0.406	III III	-152	90	-19	-80	-21	-75	-47
DAAC	0.62	0.144	III III	-152	90	-15	-85	-15	-82	81
EADE	1.03	0.072	VII -	-156	152	-43	-157	98	-154	157
DFA*E	1.10	0.064	II -	-152	187	138	55	59	-126	143
DFA*D	1.16	0.058	II -	-152	85	139	55	57	-124	85
DADC	1.36	0.071	VII -	-149	89	-38	-155	64	-85	84
FADE	1.57	0.029	VII -	-78	151	-43	-159	98	-154	156
DADF	1.64	0.026	VII -	-154	81	-43	-150	37	-81	153
EACE	1.64	0.026	I -	-160	150	-66	-98	98	-166	155
DADC	1.80	0.020	VII -	-150	95	-48	-148	64	-85	83

[a]All minima with $\Delta E < 2.08$ kcal/mole.
[b]E_0 =-4.62 kcal/mole, $\Delta E = E - E_0$
[c]Normalized Boltzmann factor at 300K.
[d]Bend type for Pro-Ala and Ala-Cys.

C_7 structure). These results indicate that compactly fold conformations such as DAAA one taking type III-III double bend structures are very energetically favorable for Ac-CyH-Pro-Ala-CyH-NHMe. By comparing the conformational characteristics of Ac-CyH-Pro-Ala-CyH-NHMe with those of Ac-Ala-Pro-Ala-Ala-NHMe, it is suggested that backbone conformations of Ac-CyH-Pro-Ala-CyH-NHMe are dominantly governed without interactions related to SH group of CyH residues except for the relative stability-change of DACC and DCCC conformations. For the case of Ac-CyH-Pro-Gly-CyH-NHMe, DAAA conformation are not so stable one, e.g., ΔE of the lowest-energy one among them is 1.32 kcal/mol(1). On the contrary, all stable conformations of Ac-CyH-Pro-Gly-CyH-NHMe having conformational letter code C* or D* for Gly residue were not found as stable ones for Ac-CyH-Pro-Ala-CyH-NHMe. That is, stable conformations of Ac-CyH-Pro-Ala-CyH-NHMe are remarkably different from those of Ac-CyH-Pro-Gly-CyH-NHMe. It is caused by the difference in the stable region of (ϕ, ψ)-plane between Ala and Gly residues, indicating that conformational property of the Xaa residue within intraresidue interactions has a great role for stabilizing each conformation of Ac-CyH-Pro-Xaa-CyH-NHMe.

There were 27 energy minima for Ac-Cys-Pro-Ala-Cys-NHMe with ΔE<3.0kcal/mol, and 10 of them are shown in Table III. The lowest-energy conformation is a DAAA conformation(Figure 3) which takes type III-III double bend at Pro-Ala-Cys portion. Type III β-bend at Pro-Ala portion is supported by the experimental results(6) that no NOE were observed between Pro $C^{\alpha}H$ and Ala NH of Boc-Cys-Pro-Ala-Cys-NHMe in $CDCl_3$. As shown in Figure 3, two amido protons of the Cys_4 residue and the NHMe group interact with the Cys_1 carbonyl oxygen, and two amido protons of the Cys_1 and Ala residues present no such interactions with carbonyl oxygen. These results correspond to the experimental results, e.g., the temperature dependence of the chemical shift of the amido protons for the Cys_1, Ala, Cys_4 residues and NHMe group, 0.0073, 0.0051, 0.0021 and 0.0019 ppm/K, respectively, in $(CD_3)_2SO$, and the order(NH(Ala) $>$ NH(Cys_1) $\gg$ NH(NHMe) $>$ NH(Cys_4)) of rates of hydrogen-deuterium (H-D)

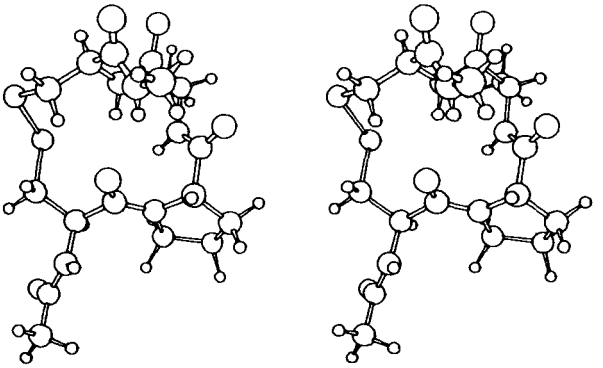

Fig. 3. The lowest-energy conformation(DAAA) of Ac-Cys-Pro-Ala-Cys-NHMe.

Table IV. Vicinal Coupling Constant($^3J_{NH-C^{\alpha}H}$)of Ac-Cys-Pro-Ala-Cys-NHMe

	Cys1	Ala	Cys4
Calculated	8.0	7.6	7.3
Experimental[a]	9.0	9.0	5.5

[a]Boc-Cys-Pro-Ala-Cys-NHMe in $(CD_3)_2SO$ from ref 6.

exchange in $CDCl_3$-D_2O mixtures(6). That is, two amido protons of the Cys4 residue and the NHMe group are strongly solvent shielded in $CDCl_3$ and $CDCl_3$- D_2O solution. The 2nd low-energy conformation(ΔE=0.62kcal/mol) is a DAAC one which also takes III-III double bend at Pro-Ala-Cys portion. This conformation is almost corresponding to the DAAA one except for the value of ϕ, i.e., only the direction of the C-terminal peptide-group is different between DAAA and DAAC. So, the hydrogen bond (Cys1)CO$\cdots$HN(NHMe), which is shown in DAAA, is not formed, but new hydrogen bond (Ala)CO$\cdots$HN(NHMe) is formed. That is, amido proton of the NHMe group is also placed in the situation strongly interacted with CO group in spite of the difference in the hydrogen-bond pattern. The DAAA and DAAC conformations are only two stable ones with ΔE<1kcal/mol. Moreover, all 27 minima with ΔE<3kcal/mol take β-bends at Pro-Ala portion. Most of them take type III-III double bend at Pro-Ala-Cys portion, or type VII or I β-bend at Pro-Ala portion, with theoretically estimated occurring probabilities 0.57, 0.20, and 0.08, respectively, and minor part of them takes type II β-bend at Pro-Ala portion with theoretically estimated occurring probability 0.16. As the types VII and I β-bends have very resemble structures to that of the type III β-bend, the type III(or VII, or I) β-bend structure is essentially favorable conformation for the Pro-Ala portion of Ac-Cys-Pro-Ala-Cys-NHMe. These overall conformational features present good agreement with those estimated by experimental works(6). As shown in Table IV, these

points are also supported by the fairly good agreement of the calculated vicinal NH-$C^{\alpha}H$ coupling constants $^{3}J_{NH\text{-}C^{\alpha}H}$ of Cys and Ala residues with experimentally evaluated ones for Boc-Cys-Pro-Ala-Cys-NHMe in $(CD_3)_2SO$ solution(6).

Comparing with Tables II and III, it is suggested that forming the S-S linkage stabilizes the type III-III double bend structures and destabilizes the bend-C_7 structure(DACC) and the continuous C_7 structure(DCCC). It means that forming the S-S linkage is a very useful method designing the double-bend structures at the turn portion of artificial proteins.

Conformational preference of Ac-Cys-Pro-Ala-Cys-NHMe is clearly different from that of Ac-Cys-Pro-Gly-Cys-NHMe in the following points. That is, Ac-Cys-Pro-Ala-Cys-NHMe has the tendency forming the type III-III double bend at the Pro-Ala-Cys portion, but Ac-Cys-Pro-Gly-Cys-NHMe has the tendency forming the type II bend at the Pro-Gly portion and non-bend structure at the Gly-Cys portion. These results indicate that the turn structures in the Ac-Cys-Pro-Ala-Cys-NHMe are more compactly fold ones than those in Ac-Cys-Pro-Gly-Cys-NHMe. As clearly shown in Figures 3 in the previous(1) and this works, the extending directions of the backbone out of the N- or C-terminal portion of the peptides are remarkably distinguishable between two peptides. Such conformational difference between Ac-Cys-Pro-Ala-Cys-NHMe and Ac-Cys-Pro-Gly-Cys-NHMe suggests that different fold structures could be designed as the turn portion of the artificial proteins by selecting the sequence, -Cys-Pro-Ala-Cys- or -Cys-Pro-Gly-Cys-. Moreover, the spatial arrangement of the polar groups in Ac-Cys-Pro-Ala-Cys-NHMe is also remarkably different from that in Ac-Cys-Pro-Gly-Cys-NHMe arising from bend-type differences between them. It means that such difference in the polar-group arrangement would induce the specific contribution of the -Cys-Pro-Ala-Cys- or -Cys-Pro-Gly-Cys- sequence for stabilizing the different three-dimensional structure in each protein.

In this work, an Ala residue is used as a model residue for the general alanine-type residues such as Val, Leu, etc. based on the Ala-residue approximation(7,8). As it was already shown that the Ala-residue approximation is a useful method for analyzing the conformational preference of the peptides and polypeptides as the primary procedure in the theoretical conformational analysis(9-15), above conclusion obtained in this work could be generally applicable for the -Cys-Pro-Xaa-Cys-(Xaa is a alanine-type residue) sequences except for the case with significant side-chain/backbone or side-chain/side-chain interactions. These points are experimentally supported by the X-ray diffraction measurements for the single-crystals of Ac-Cys-Pro-Val-Cys-NHMe and Ac-Cys-Pro-Ser-Cys-NHMe(16). Further works are now in progress for confirming the validity of the Ala-residue approximation in the -Cys-Pro-Xaa-Cys- systems.

References

1. Yoshimoto J, Nishinaga A, Oka M, Hayashi T, Polym Bull, in press.
2. Momany FA, McGuire RF, Burgess AW, Scheraga HA (1975) J Phys Chem 79: 361
3. Zimmerman SS, Scheraga HA, (1977) Biopolymers 16: 811
4. Bystrov VF, (1976) Prog NMR Spectroscopy 10:41
5. Zimmerman SS, Pottle MS, Nemethy G, Scheraga HA, (1977) Macromolecules 10: 1
6. Ravi A, Balaram T, (1984) Tetrahedron 40: 2577
7. Brant DA, Flory PJ, (1965) J Am Chem Soc 87:2791

8. Flory PJ, (1969) "Statistical Mechanics of Chain Molecules" John Wiley & Sons Inc, New York :248
9. Oka M, Montelione GT, Scheraga HA, (1984) J Am Chem Soc 106:7959
10. Oka M, Baba Y, Kagemoto A, Nakajima A, (1990) Polym J 22:135
11. Oka M, Baba Y, Kagemoto A, Nakajima A, (1990) Polym J 22:416
12. Oka M, Baba Y, Kagemoto A, Nakajima A, (1991) Polym Bull 25:701
13. Oka M, Nishikawa R, Baba Y, Kagemoto A, (1992) Peptide Chem 1991:253
14. Oka M, Nakajima A, Lee S, Sugihara G, (1993) Rept Prog Polym Phys Jpn 36:709
15. Oka M, Iwai S, Ishikawa Y, Hayashi T, Miyazawa M, (1996) Peptide Chem 1995:457
16. Falcomer CM, Meinwald YC, Choudhary I, Talluri S, Milburn PJ, Clardy J, Scheraga HA, (1992) J Am Chem Soc 114:4036