Polymer Bulletin 38, 109–115 (1997)

Polymer Bulletin

Theoretical conformational analysis of tetrapeptide Ac-Cys-Pro-Gly-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: 26 August 1996/Revised version: 18 September 1996/Accepted: 23 September 1996

Summary

Theoretical conformational analysis was carried out for the acyclic and cyclic tetrapeptides Ac-Cys-Pro-Gly-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-Gly-Cys-NHMe forms compactly fold conformations with type II β-bend at the Pro-Gly portion, and also show fairly good agreement with experimental results of the NMR spectroscopy for the tetrapeptides having Cys-Pro-Gly-Cys sequence.

Introduction

Disulfide linkage formed between two Cys residues is an important factor for stabilizing three dimensional structures of proteins as well as non-covalent bonded interactions such as hydrogen bonds, hydrophobic interactions, and salt bridges between oppositely charged residues. From the viewpoint of designing artificial functional proteins, it is very important to analyze the conformational preference of peptides with disulfide linkage. In this work, as a first 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-Gly-Cys-NHMe using ECEPP(1) 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) of Gly, (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-Gly-Ala-NHMe, which is a model peptide of acyclic Ac-CyH-Pro-Gly-CyH-NHMe based on the Ala-residue approximation, was minimized using all combinations of the single-residue minima of Ala, Pro, and Gly residues(9, 4 ,

* Corresponding author

and 9, respectively). As the first step of minimization of cyclic tetrapeptide Ac-Cys-Pro-Gly-Cys-NHMe with disulfide bridge, conformational energy of dipeptide Ac-CyH-Pro-NHMe was minimized using all combinations of the single-residue minima of CyH and Pro residues(47 and 4, respectively). Then, conformational energy of acyclic tetrapeptides, Ac-CyH-Pro-Gly-CyH-NHMe, was minimized using all combinations of the minima of Ac-CyH-Pro-NHMe and the single-residue minima of Gly and CyH residues as starting conformations. As the final step, conformational energy of cyclic tetrapeptide Ac-Cys-Pro-Gly-Cys-NHMe with disulfide bridge was minimized using stable minima of acyclic Ac-CyH-Pro-Gly-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 2. 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 Gly residues for Ac-Cys-Pro-Gly-Cys-NHMe were computed using the equation derived by Bystrov *et al*(3) and normalized Boltzmann factor(v) for all minima with ΔE<3kcal/mol. Conformational space is divided into 16 regions with the conformational letter codes(CLC) shown in Figure 2 of ref 4.

Results and Discussion

There were 811 energy minima for Ac-Ala-Pro-Gly-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(DCC*C conformation) which takes type V-V' double bend at Pro-Gly-Ala portion stabilized by the hydrogen bonds (Ala_1)CO•••HN(Gly) and (Pro)CO•••HN(Ala_4). The 2nd low-energy conformation(ΔE=0.11kcal/mol) is DCC*F one which takes type V β-bend at Pro-Gly portion, and is also a compactly fold conformation stabilized by the (Ala_1)CO•••HN(Gly) and (Pro)CO•••HN(Ala_4). As the atom-pair distance of (Pro)C^{α}•••C(NHMe) (7.3 Å) is very close to the critical value of 7.0 Å of defining β-bend structures, this conformation is also regarded as a conformation analogous to the double-bend structures. All of 24 conformations with ΔE<2.0kcal/mol are also compactly fold conformations with β-bend at Pro-Gly portion or with double bend at Pro-Gly-Ala portion.

There were 13877 energy minima for Ac-CyH-Pro-Gly-CyH-NHMe with ΔE<6.9kcal/mol, and 15 of them are shown in Table II. The lowest-energy conformation is a compactly fold conformation(DCC*F conformation) which takes type V β-bend at Pro-Gly portion stabilized by the hydrogen bonds (CyH_1)CO•••HN(Gly) and (Pro)CO•••HN(CyH_4) as shown in Figure 2. As the atom-pair distance of (Pro)C^{α}•••C(NHMe) (7.1 Å) is very close to the critical value of 7.0 Å of defining β-bend structures, this conformation is also regarded as a conformation analogous to the double-bend structures. The 2nd low-energy conformation(ΔE=0.04kcal/mol) is also DCC*F one. Conformational difference between them is only found in χ^2 of CyH_4, i.e., χ^2=180° and -65° for the lowest- and 2nd low-energy conformations, respectively. The 3rd and 4th low-energy conformations(ΔE=0.13 and 0.15kcal/mol, respectively) are DCC*C ones which take type V-V' double bend at Pro-Gly-CyH portion stabilized by the hydrogen bonds (CyH_1)CO•••HN(Gly) and (Pro)CO •••HN(CyH_4). Their overall backbone conformations are very resemble to the DCC*F ones. All of 30 stable

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

Conformational Letter Code	ΔE^b (kcal/mole)	v^c	Bend Type[d]	ϕ Ala1	ψ Ala1	ϕ Pro	ϕ Gly	ψ Gly	ϕ Ala4	ψ Ala4
DCC*C	0.00	0.155	V V'	-152	79	75	80	-78	-86	80
DCC*F	0.11	0.130	V -	-153	80	82	86	-69	-80	146
DCB*E	0.42	0.077	II -	-154	79	79	103	-47	-151	157
DCC*D	0.43	0.076	V IV	-151	80	84	86	-70	-149	73
DCD*A	0.59	0.058	VII III	-153	80	79	161	-56	-82	-38
DFC*A	0.63	0.054	II III	-152	80	157	85	-66	-69	-47
DFC*C	0.71	0.047	II V	-152	80	141	80	-77	-87	76
DCA*E	0.89	0.035	II -	-152	79	83	75	53	-151	158
DCC*D	0.96	0.031	V IV	-151	80	78	80	-78	-154	52
DCD*C	1.02	0.028	IV IV	-152	80	76	154	-60	-88	75

[a]All minima with $\Delta E < 1.04$ kcal/mole.
[b]E_0 =-9.59 kcal/mole, $\Delta E = E - E_0$
[c]Normalized Boltzmann factor at 300K.
[d]Bend type for Pro-Gly and Gly-Ala.

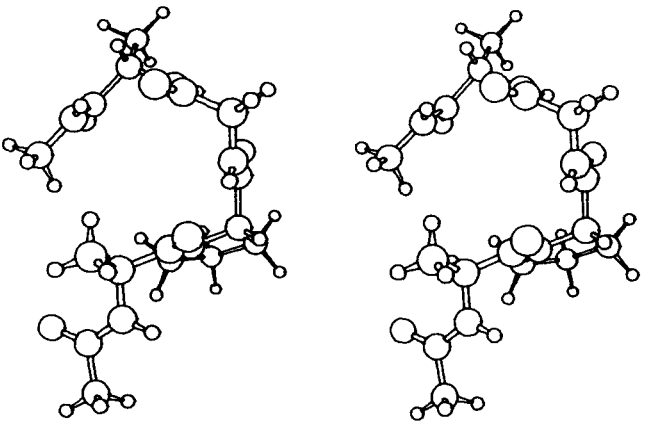

Fig. 1. The lowest-energy conformation(DCC*C) of Ac-Ala-Pro-Gly-Ala-NHMe.

conformations with ΔE<0.5kcal/mol are one of the following 6 conformation-types whose CLC are DCC*F(type V β-bend), DCC*C (type V-V' double bend), DFC*A(type II-II' double bend), DCC*A(type V-II' double bend), DCC*E(type V β-bend) and DCC*D(type V-IV double bend), and number of them are 5, 5, 9, 6, 3 and 2, respectively. Moreover, the atom-pair distances of (Pro)C^{α}•••C(NHMe) of DCC*F and DCC*E conformations are very close to the critical value of defining β-bend structures. That is, overall conformational feature of Ac-CyH-Pro-Gly-CyH-NHMe is compactly fold conformation with double-bend or double-bend like structure taking type V(which is a modified bend of type II) or type II β-bends at Pro-Gly portion. Above conformational feature is very resemble to that of Ac-Ala-Pro-Gly-Ala-NHMe. That is, all above 6 types of backbone conformations of Ac-CyH-Pro-Gly-CyH-NHMe are also found as the stable

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

Conformational Letter Code	ΔE^b (kcal/mole)	V^c	Bend Type[d]	ϕ CyH1	ϕ CyH1	ϕ Pro	ϕ Gly	ϕ Gly	ϕ CyH4	ϕ CyH4
DCC*F	0.00	0.013	V -	-155	82	73	80	-75	-84	150
DCC*F	0.04	0.013	V -	-155	81	73	80	-74	-85	150
DCC*C	0.13	0.011	V V'	-155	81	72	76	-82	-88	97
DCC*C	0.15	0.010	V V'	-155	83	72	77	-82	-88	98
DFC*A	0.21	0.010	II II'	-154	83	164	84	-68	-72	-50
DFC*A	0.23	0.009	II II'	-154	80	164	84	-68	-71	-50
DCC*C	0.23	0.009	V V'	-155	83	72	76	-84	-90	90
DCC*A	0.24	0.009	V II'	-153	83	81	79	-96	-82	-52
DFC*A	0.24	0.009	II II'	-154	83	164	84	-68	-79	-49
DCC*A	0.26	0.009	V II'	-152	83	81	79	-96	-82	-52
DCC*E	0.26	0.009	V -	-154	82	79	89	-62	-150	158
DFC*A	0.26	0.009	II II'	-154	83	164	84	-69	-72	-49
DCC*C	0.28	0.008	V V'	-155	80	72	76	-83	-90	90
DCC*A	0.34	0.008	V II'	-152	83	82	80	-96	-81	-53
DFC*A	0.35	0.008	II II'	-154	83	164	85	-67	-73	-49

[a]All minima with $\Delta E < 0.36$ kcal/mole.
[b]E_0 =-11.18 kcal/mole, $\Delta E = E - E_0$
[c]Normalized Boltzmann factor at 300K.
[d]Bend type for Pro-Gly and Gly-CyH.

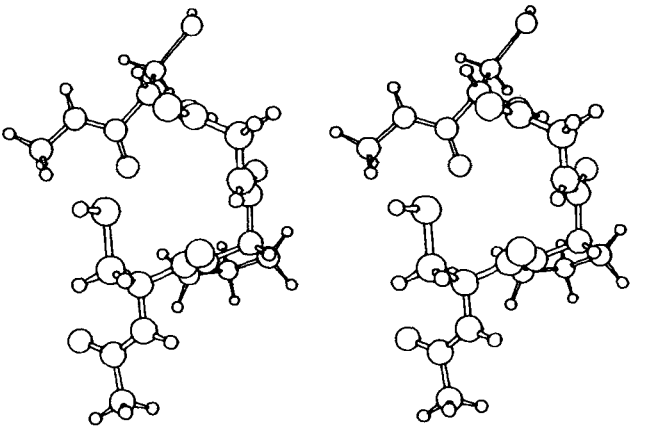

Fig. 2. The lowest-energy conformation(DCC*F) of Ac-CyH-Pro-Gly-CyH-NHMe.

conformations of Ac-Ala-Pro-Gly-Ala-NHMe in spite of the change in relative stability of each conformation. It suggests that the SH group of the cystein residue has not so important roles for stabilizing the backbone conformation of Ac-CyH-Pro-Gly-CyH-NHMe as a whole.

There were 41 energy minima for Ac-Cys-Pro-Gly-Cys-NHMe with ΔE<3.0kcal/mol, and 15 of them are shown in Table III. The lowest-energy conformation

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

Conformational Letter Code	ΔE^b (kcal/mole)	v^c	Bend Type[d]	ϕ Cys1	ψ Cys1	ϕ Pro	ϕ Gly	ψ Gly	ϕ Cys4	ψ Cys4
DCA*E	0.00	0.201	II -	-153	85	129	69	51	-126	144
DFA*D	0.10	0.169	II -	-153	83	133	66	46	-120	84
DAAA	0.44	0.096	III III	-152	90	-19	-82	-17	-75	-47
DCA*F	0.49	0.088	II -	-155	76	76	72	25	-72	153
DAAC	0.80	0.053	III -	-152	90	-15	-87	-10	-83	82
ECH*E	0.94	0.041	II -	-157	150	70	69	90	-150	157
DCD*C	0.97	0.039	II -	-154	92	84	153	-29	-57	129
ECA*E	1.16	0.029	II -	-158	155	98	76	81	-155	154
DCD*A	1.31	0.022	II -	-154	93	84	154	-30	-58	-52
EADE	1.35	0.021	VII -	-157	151	-24	-175	96	-155	157
DADC	1.37	0.020	VII -	-148	87	-22	-172	61	-84	85
A*CA*E	1.37	0.020	II -	61	82	127	71	51	-124	146
A*FA*D	1.41	0.019	II -	61	81	132	68	46	-118	81
FCH*E	1.46	0.017	II -	-74	148	70	69	92	-150	156
DFA*G	1.48	0.017	I -	-153	83	134	67	43	-118	-58

[a]All minima with $\Delta E < 1.65$ kcal/mole.
[b]E_0 =-6.15 kcal/mole, $\Delta E = E - E_0$
[c]Normalized Boltzmann factor at 300K.
[d]Bend type for Pro-Gly and Gly-Cys.

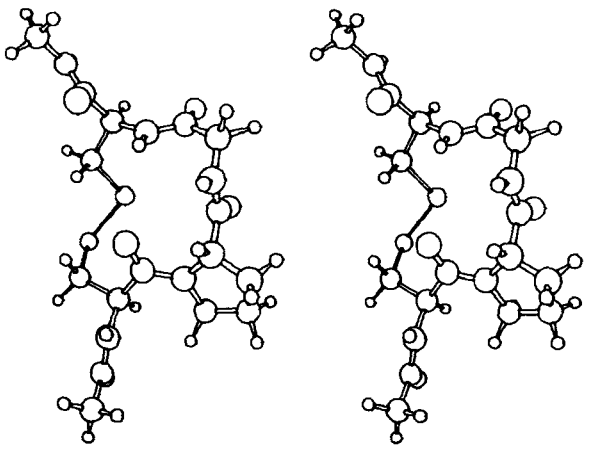

Fig. 3. The lowest-energy conformation(DCA*E) of Ac-Cys-Pro-Gly-Cys-NHMe.

is a DCA*E conformation(Figure 3) taking type II bend at Pro-Gly portion. As shown in Figure 3, the atom-atom pairs (Pro)C$^{\alpha}$H•••HN(Gly), (Gly)NH•••HN(Cys4) and (Gly)C$^{\alpha}$H•••HN(Cys4) present very close contact, i.e., their distance are 2.1, 2.7 and 2.8 Å, respectively. These short interatomic distances show fairly good agreements with NOE(5) observed in the two-dimensional NMR spectroscopy of Ac-Cys-Pro-Gly-Cys-NHMe in aqueous solution. That is, a very strong NOE was observed between Pro C$^{\alpha}$H

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

	Cys_1	Gly	Cys_4
Calculated	7.8	12.9	8.5
Experimental[a]	8.0	-	7.0

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

and Gly NH. A strong NOE was observed between Gly NH and Cys_4 NH. A NOE was also observed between Gly $C^\alpha H$ and Cys_4 NH. Moreover, Cys_1 carbonyl oxygen and Cys_4 amido proton present the close contact. This structural situation is also supported by the small temperature dependence of the chemical shift of the Cys4 amido proton(less than 4 ppb/℃) in aqueous solution(5). The 2nd low-energy conformation(ΔE=0.10kcal/mol) is a DFA*D one. This conformation almost corresponds to the lowest-energy one with slight difference in the direction of peptide group at the C-terminal. Moreover, most of stable conformations also takes type II β-bend at the Pro-Gly portion as shown in Table III. The 3rd low-energy conformation is DAAA conformation taking type III-III double bend at the Pro-Gly-Cys portion. The 5th low-energy conformation is DAAC one taking type III β-bend at the Pro-Gly portion, and its (Pro)C^α•••C(NHMe) is 7.1 Å indicating that this conformation is analogous to the double-bend structures. As both of DAAA and DAAC conformations don't show short inter-atom distances for (Pro)C^αH•••HN(Gly) and (Gly)C^αH•••HN(Cys_4), they are not expected as stable conformations in aqueous solution. The (ϕ, ψ)-values of Gly residue in the most of stable conformations of Ac-CyH-Pro-Gly-CyH-NHMe take those in C* region(40° $< \phi <$110° , -130° $< \psi <$-50°). However, there are no conformations taking (ϕ, ψ)-values of Gly residue in C* region with ΔE<3.0kcal/mol for Ac-Cys-Pro-Gly-Cys-NHMe. As shown in Table III, the (ϕ, ψ) of Gly residue in the most of stable conformations of Ac-Cys-Pro-Gly-Cys-NHMe take those in A* region(40° $< \phi <$110° , 10° $< \psi <$90°) or A region(-110° $< \phi <$-40° , -90° $< \psi <$-10°). Such a difference in the stable region of (ϕ, ψ)-plane between Ac-CyH-Pro-Gly-CyH-NHMe and Ac-Cys-Pro-Gly-Cys-NHMe appeares as a difference in the conformational feature between them, i.e., double-bend structures are stable conformations for Ac-CyH-Pro-Gly-CyH-NHMe, however, they are not stable ones for Ac-Cys-Pro-Gly-Cys-NHMe. All stable conformations in Table III, except the 3rd and 5th low-energy ones, take β-bends at the Pro-Gly portion and take non-bend structures at the Gly-Cys portion. Moreover, as clearly shown in Figures 2 and 3, the relative direction of the peptide backbone extending out of the N-terminal and the C-terminal of each peptide is remarkably different with or without the disulfide linkage. That is, whole conformational feature of Ac-Cys-Pro-Gly-Cys-NHMe is remarkably different from that of Ac-CyH-Pro-Gly-CyH-NHMe, indicating that forming the disulfide linkage between two Cys residues significantly affects the conformational character of Ac-Cys-Pro-Gly-Cys-NHMe.

Calculated occurring probability indicates that conformations taking type II β-bend at the Pro-Gly portion are essentially favorable in the whole ensemble of Ac-Cys-Pro-Gly-Cys-NHMe. It also corresponds to Falcomer et al.'s conclusion(6) that Ac-Cys-Pro-Gly-Cys-NHMe takes type II β-bend at the Pro-Gly portion in aqueous solution.

Calculated vicinal NH-$C^{\alpha}H$ coupling constants $^{3}J_{NH\text{-}C^{\alpha}H}$ of Cys_1, Gly and Cys_4 residues are shown in Table IV. They show good agreement with the experimental results for Boc-Cys-Pro-Gly-Cys-NHMe in $(CD_3)_2SO$ solution(6).

Good agreements between calculated and experimental results strongly indicate that Cys-Pro-Gly-Cys sequence has a tendency forming type II β-bend at the Pro-Gly portion by forming disulfide bond between two Cys residues. These results suggest that the Cys-Pro-Gly-Cys sequence is very useful for designing the turn or hairpin structures in artificial proteins.

References

1. Momany FA, McGuire RF, Burgess AW, Scheraga HA (1975) J Phys Chem 79: 361
2. Zimmerman SS, Scheraga HA, (1977) Biopolymers 16: 811
3. Bystrov VF, (1976) Prog NMR Spectroscopy 10:41
4. Zimmerman SS, Pottle MS, Nemethy G, Scheraga HA, (1977) Macromolecules 10: 1
5. Falcomer CM, Meinwald YC, Choudhary I, Talluri S, Milburn PJ, Clardy J, Scheraga HA, (1992) J Am Chem Soc 114: 4036
6. Ravi A, Balaram T, (1984) Tetrahedron 40: 2577