Theoretical Analysis of α -Helix Hairpin Structures Constructed by Two Right-Handed α -Helices

Yuichirou ISHIKAWA, Masahito OKA,[†] Toshio HAYASHI, and Akira Nishinaga*

Research Institute for Advanced Science and Technology, University of Osaka Prefecture, Gakuen-cho, Sakai, Osaka 593, Japan

* Department of Applied Chemistry, Osaka Institute of Technology, Omiya, Asahi-ku, Osaka 535, Japan

(Received October 18, 1995)

ABSTRACT: Theoretical conformational analysis was carried out for $\text{Ac}-(\text{Ala})_n\text{-Gly-Gly-Gly}-(\text{Ala})_n\text{-NHMe}$ ($n=5, 10-15$, and 20) and $\text{Ac}-(\text{Ala})_n\text{-Gly-Gly}-(\text{Ala})_n\text{-NHMe}$ ($n=5, 10, 15, 20$, and 25) for analyzing the stability of the α -helix hairpin structure constructed by two right-handed α -helices. Calculated results indicate that α -helix hairpin structures can exist as stable super secondary structures by the favorable interhelix interactions for $\text{Ac}-(\text{Ala})_n\text{-Gly-Gly-Gly}-(\text{Ala})_n\text{-NHMe}$ ($n \geq 13$), and also that their orientation angles almost correspond to those experimentally found for adjacent neighbor helices in square four- α -helix bundle-type proteins. The Gly-Gly dipeptide sequence is not sufficient for forming stable α -helix hairpin structure as the loop portion, but the Gly-Gly-Gly tripeptide sequence is the shortest peptide sequence enough to form stable α -helix hairpin structures as the flexible loop portion.

KEY WORDS Super Secondary Structure / α -Helix Hairpin Structure / Polypeptide / Artificial Protein / Molecular Mechanics / Empirical Conformational Energy Program for Peptides / ECEPP /

In native proteins, there are many super secondary structures¹ such as 4- α -helix bundle, β -sheet, β -barrel, Rosmann fold, Tim barrel, and leucine-zipper structures constructed by the specific spatial combination of α -helix and β -strand, and these super secondary structures have very important roles as backbone structures supporting the molecular functions of proteins. It is very important to analyze super helical structures of polypeptides for designing new artificial proteins exhibiting interesting molecular functions exceeding those of native proteins.

As one of the process for these viewpoints, polypeptides designed to assemble into an antiparallel 4- α -helix bundle and a dimerized 2- α -helix bundle were synthesized, and their conformational properties in solution were characterized by the measurements of NMR, circular dichroism (CD) and fluorescence spectra.²⁻⁷ However, it is not clear that designed amphiphilic polypeptides take intramolecular or intermolecular bundle in aqueous solutions, and also that the relative orientation of the helix axes in solutions is similar to those of the native proteins in the solid states analyzed by the X-ray crystallography.^{8,9} Theoretical conformational analysis was also carried out for the intermolecular assembled system constructed by two and four $\text{Ac}-(\text{Ala})_{10}\text{-NHMe}$ molecules which takes α -helix by Chou *et al.*^{9,10} However, it is not clear that intermolecular arrangement of α -helices proposed by them can directly correspond to intramolecular arrangement of α -helices which is usually found in square 4- α -helix bundle-type proteins. It is very important to investigate the relative stabilities of an α -helix bundle compared to a single α -helix, and intramolecular arrangement of α -helices in a 4- α -helix bundle for designing the new artificial proteins based on the 4- α -helix bundle structures.

In this work, as the first step of the theoretical investigation of the α -helix bundle structure, conformational preference of the α -helix hairpin structures (*i.e.*,

2- α -helix bundle) was theoretically analyzed based on the molecular-mechanics calculations and the conformational-energy optimization procedures.

THEORETICAL

Two kinds of polypeptides, $\text{Ac}-(\text{Ala})_n\text{-Gly-Gly-Gly}-(\text{Ala})_n\text{-NHMe}$ ($n=5, 10-15$, and 20; abbreviated as $\alpha_n\text{-GGG-}\alpha_n$) and $\text{Ac}-(\text{Ala})_n\text{-Gly-Gly}-(\text{Ala})_n\text{-NHMe}$ ($n=5, 10, 15, 20$, and 25; abbreviated as $\alpha_n\text{-GG-}\alpha_n$) were used as model polypeptides for analyzing α -helix hairpin structures. Ala residue was selected as a fundamental residue taking α -helical conformations, and Gly residue was selected as a desirable residue having the largest flexibility in the conformational space among the 20 naturally occurring amino-acid residues. All conformational energy calculations were carried out with the energy functions of ECEPP.¹¹

For the case of the first model polypeptides $\alpha_n\text{-GGG-}\alpha_n$, at first, conformational optimization of the Gly-Gly-Gly portion was carried out for the whole conformational space of the model peptide, $\text{Ac-Gly-Gly-Gly-NHMe}$ using all combinations of the single residue minima of Gly residue as starting conformations. During minimization, (ϕ, ψ) of Gly residue were allowed to vary. All other backbone dihedral angles were fixed at 180° . Then, conformational energies of $\text{Ac}-(\text{Ala})_5\text{-Gly-Gly-Gly}-(\text{Ala})_5\text{-NHMe}$ (*i.e.*, $\alpha_5\text{-GGG-}\alpha_5$) were optimized using all local minima of $\text{Ac-Gly-Gly-Gly-NHMe}$ as starting conformations for the Gly-Gly-Gly portion. During minimization, (ϕ, ψ) of Gly residue were allowed to vary. (ϕ, ψ, χ^1) of Ala residue were fixed to those of the energy-minima in the α -helical region of poly(Ala), *i.e.*, $(\phi_{\text{Ala}}, \psi_{\text{Ala}}, \chi_{\text{Ala}}^1) = (-68^\circ, -38^\circ, 61^\circ)$, and all other backbone dihedral angles were fixed at 180° . Further, conformational energies of $\alpha_{10}\text{-GGG-}\alpha_{10}$, $\alpha_{15}\text{-GGG-}\alpha_{15}$, and $\alpha_{20}\text{-GGG-}\alpha_{20}$ were minimized using all local minima of $\alpha_5\text{-GGG-}\alpha_5$, $\alpha_{10}\text{-GGG-}\alpha_{10}$, and $\alpha_{15}\text{-GGG-}\alpha_{15}$ as starting conformations, respectively. Other minimization

[†] To whom correspondence should be addressed.

conditions were the same as those for α_5 -GGG- α_5 . Additional optimizations were also carried out for four polypeptides, α_n -GGG- α_n ($n=11, 12, 13$, and 14), for further analysis of the relative stabilities between α -helix hairpin structure and single α -helical structure using 20 low-energy minima of α_{10} -GGG- α_{10} as starting conformations.

For the case of the second model polypeptides α_n -GG- α_n , at first, conformational optimization of the Gly-Gly portion was carried out for the whole conformational space of the model peptide, Ac-Gly-Gly-NHMe following the same optimization procedure as that for Ac-Gly-Gly-Gly-NHMe. Then, conformational energies of α_{10} -GG- α_{10} , α_{15} -GG- α_{15} , α_{20} -GG- α_{20} , and α_{25} -GG- α_{25} were minimized in a similar manner for α_n -GGG- α_n .

Conformational space is divided into 16 regions with the conformational letter codes shown in Figure 1 of ref 12. A bend (occurring at $i+1$ and $i+2$ th residues) is defined as a conformation in which $R \leq 7 \text{ \AA}$ (R is the distance between i th C^α and $i+3$ th C^α atoms.) and also classified into eleven types given in Table I of ref 13. A polar hydrogen atom and an oxygen or nitrogen atom with an interatomic distance of less than $2.3 \text{ \AA}$ are regarded to be hydrogen-bonded. The conformational energy per whole molecule, ΔE , is defined by $\Delta E = E - E_0$, E_0 is the value of E at the global minimum on the potential energy surface of the particular molecules, and ΔE_{res} is defined by $\Delta E_{\text{res}} = \Delta E/m$, where m is number of residue of the molecule. The relative orientation of two α -helices is described in terms of the orientation angle Ω_0 defined in ref 10, and taken to lie between 0° and $\pm 180^\circ$. The orientation angle is positive if, starting from an initial parallel orientation of the helices ($\Omega_0 = 0^\circ$), the 2nd helix far from the viewer is rotated clockwise relative to the 1st one near the viewer; it is negative if the 2nd helix is rotated in the counterclockwise sense. Following this definition, the conformation of the α_n -GGG- α_n and α_n -GG- α_n molecules having $\Omega_0 = 0^\circ$ corresponds to a single α -helix. An end-to-end distance of the molecule, R_0 , is defined as the distance between the methyl carbon atom of acethyl group and the methyl carbon atom of N -methyl amide

group. All molecular diagrams are described by the PEPCON program¹⁴ which is a modified version of the NAMOD program.¹⁵

RESULTS AND DISCUSSION

Stable Conformation of Ac-Gly-Gly-Gly-NHMe

A total of 501 minima of Ac-Gly-Gly-Gly-NHMe was found in $\Delta E_{\text{res}} < 1.88 \text{ kcal mol}^{-1}$. All 30 low-energy minima with $\Delta E_{\text{res}} < 0.50 \text{ kcal mol}^{-1}$ are very compact conformations induced by the double-bend structures, and all value of end-to-end distances, R_0 , of them are less than $6 \text{ \AA}$. However, number of conformations having large value of R_0 increases with an increase in ΔE_{res} . That is, the overall conformational feature of Ac-Gly-Gly-Gly-NHMe is that relatively low-energy conformations with small ΔE_{res} take compact conformations and also that relatively high-energy conformations with large ΔE_{res} take extended conformations.

Stable Conformation of α_{20} -GGG- α_{20}

In Table I, all 14 stable conformations of α_{20} -GGG- α_{20} with $\Delta E_{\text{res}} < 0.17 \text{ kcal mol}^{-1}$ are shown. The lowest-energy conformation, which has $(\phi_{\text{Gly}_{21}}, \psi_{\text{Gly}_{21}}, \phi_{\text{Gly}_{22}}, \psi_{\text{Gly}_{22}}, \phi_{\text{Gly}_{23}}, \psi_{\text{Gly}_{23}}) = (104^\circ, 51^\circ, -76^\circ, -37^\circ, 76^\circ, -99^\circ)$, is a typical α -helix hairpin structure with orientation angle $\Omega_0 = -166^\circ$ having the structural feature that the 1st helix is rotated counterclockwise relative to the 2nd one as shown in Figure 1. It takes A*AC* conformation at the Gly-Gly-Gly loop portion with non-bend at Ala₂₀-Gly₂₁, Gly₂₁-Gly₂₂, and Gly₂₂-Gly₂₃ portions and type II' β -bend at Gly₂₃-Ala₂₄ portion, and forms (Ala₁₇)CO $\cdots$ HN(Gly₂₂) hydrogen bond together with usual the intrahelical hydrogen-bonds forming α -helix. The A*AC* conformation is found as the 344th energy minimum for Ac-Gly-Gly-Gly-NHMe with $\Delta E_{\text{res}} = 1.36 \text{ kcal mol}^{-1}$. That is, A*AC* conformation is not so stable as the conformation for the Gly-Gly-Gly portion. It suggest that the lowest-energy conformation of α_{20} -GGG- α_{20} is stabilized by the interhelical interactions. All other 12 stable conformations of α_{20} -GGG- α_{20} with $\Delta E_{\text{res}} < 0.15 \text{ kcal mol}^{-1}$ are also α -helix hairpin structures having the similar

Table I. Calculated minimum energy conformations^a of Ac-(Ala)₂₀-Gly-Gly-Gly-(Ala)₂₀-NHMe

Conformational letter code	ΔE_{res}^b kcal mol ⁻¹	Ω_0^c	R_0^d Å	$\phi_{\text{Gly}_{21}}$	$\psi_{\text{Gly}_{21}}$	$\phi_{\text{Gly}_{22}}$	$\psi_{\text{Gly}_{22}}$	$\phi_{\text{Gly}_{23}}$	$\psi_{\text{Gly}_{23}}$
A*AC*	0.00	-166	5.9	104	51	-76	-37	76	-99
C*CD	0.05	-165	6.2	84	-76	-92	68	-133	81
A*CH*	0.07	-165	6.1	85	20	-94	87	84	93
C*CD	0.09	-164	6.3	86	-66	-80	75	-168	60
DE*F	0.10	-164	8.0	-164	100	167	-168	-79	412
E*C*C*	0.11	-167	5.6	132	-117	71	-92	121	-80
G*FG	0.11	-164	5.9	116	54	-98	160	-127	-82
C*EE*	0.11	-163	7.6	89	-88	-144	158	134	-176
A*AE	0.12	-165	5.9	68	28	-77	-40	-152	111
G*GE*	0.13	-162	5.9	149	84	-115	-69	141	-178
DCE	0.13	-165	5.4	-132	100	-51	114	-153	131
G*EG	0.14	-166	5.4	119	82	-130	157	-124	-83
HEA*	0.14	-164	7.3	-93	-99	-160	-175	88	85
AAA	0.15	-2	64.0	-66	-46	-65	-39	-69	-40

^a All minima with $\Delta E_{\text{res}} < 0.17 \text{ kcal mol}^{-1}$. ^b $E_0 = -0.75 \text{ kcal mol}^{-1}$, $\Delta E_{\text{res}} = (E - E_0)/43$. ^c The angle of orientation between the axes of the two helices. ^d End-to-end distance.

Figure 1. The lowest-energy conformation of Ac-(Ala)₂₀-Gly-Gly-Gly-(Ala)₂₀-NHMe (α_{20} -GGG- α_{20}). It takes typical α -helix hairpin structure with orientation angle $\Omega_0 = -166^\circ$. (a) a stereo side-view; (b) a view from the loop portion.

conformational feature to the lowest-energy one. The average value of Ω_0 ($= -164.6^\circ \pm 3.0^\circ$) of 13 low-energy minima almost corresponds to those experimentally found for adjacent neighbor helices in square 4- α -helix bundle-type proteins.^{8,9} Moreover, it also corresponds to the calculated results for the intermolecular assembled systems of two and four Ac-(Ala)₁₀-NHMe molecules

which take α -helix.^{9,10}

A single α -helix, which takes α -helical conformation at Gly-Gly-Gly portion with $(\phi_{\text{Gly}_{21}}, \psi_{\text{Gly}_{21}}, \phi_{\text{Gly}_{22}}, \psi_{\text{Gly}_{22}}, \phi_{\text{Gly}_{23}}, \psi_{\text{Gly}_{23}}) = (-66^\circ, -46^\circ, -65^\circ, -39^\circ, -67^\circ, -40^\circ)$, was found as the 14th low-energy conformation with $\Delta E_{\text{res}} = 0.15 \text{ kcal mol}^{-1}$ and $\Omega_0 = -2^\circ$. This result indicates that α -helix hairpin structures are energeti-

cally more favorable than a single α -helix for α_{20} -GGG- α_{20} even though they lose stabilization energy by breaking three hydrogen bonds formed between the carbonyl oxygen of the i th residue and the amide proton of the $i+4$ th residue compared to a single α -helix, and also that the interhelix-interaction energy exceeds such three-hydrogen-bonds energy. It means that α -helix hairpin structures can exist as essentially stable conformations for polypeptides in the level of the interhelical interactions designed with the Ala-residue approximation¹⁶ without support of the hydrophobic interactions induced by the amphiphilic distribution of the side-chain groups along helical axis. Obtained results give the distinct theoretical basis to the experimentally proposed concept that de-

signed amphiphilic polypeptides can fold into self-associated 4- α -helix bundle structures.²⁻⁷

In the energy region with $E_{\text{res}} > 0.17 \text{ kcal mol}^{-1}$, a total of 290 local energy minima was found. Four of them (*i.e.*, the 16th, 23th, 24th, and 25th low-energy minima with $\Delta E_{\text{res}} = 0.19, 0.22, 0.25,$ and $0.25 \text{ kcal mol}^{-1}$, respectively) have large value of $|\Omega_0|$ ($|\Omega_0| > 175^\circ$) presenting almost antiparallel packing of two α -helices. Most of others are distorted α -helix hairpin structures having slightly large values of the end-to-end distance R_0 or non-hairpin structures having energetically unfavorable interhelical interactions with small value of $|\Omega_0|$.

Figure 2. The 2nd low-energy conformation of Ac-(Ala)₂₅-Gly-Gly-(Ala)₂₅-NHMe (α_{25} -GG- α_{25}) with $\Delta E_{\text{res}} = 0.04 \text{ kcal mol}^{-1}$. It takes slightly distorted α -helix hairpin structure with orientation angle $\Omega_0 = -168^\circ$. (a) a stereo side-view; (b) a view from the loop portion.

Helix Length Dependence of the Stability of α -Helix Hairpin Structures in α_n -GGG- α_n

For the case of α_5 -GGG- α_5 , the lowest-energy conformation is a single α -helix, and most of low-energy minima is not typical α -helix hairpin structures, *i.e.*, they have small values of $|\Omega_0|$. The lowest-energy α -helix hairpin structure appeared as the 22th low-energy one with $(\phi_{\text{Gly}_6}, \psi_{\text{Gly}_6}, \phi_{\text{Gly}_7}, \psi_{\text{Gly}_7}, \phi_{\text{Gly}_8}, \psi_{\text{Gly}_8}) = (-64^\circ, -37^\circ, 132^\circ, -10^\circ, -76^\circ, 76^\circ)$, $\Delta E_{\text{res}} = 0.69 \text{ kcal mol}^{-1}$, and $\Omega_0 = -165^\circ$. The A*AC* conformation, which is the lowest-energy one of α_{20} -GGG- α_{20} , is the 71th minimum with $(\phi_{\text{Gly}_6}, \psi_{\text{Gly}_6}, \phi_{\text{Gly}_7}, \psi_{\text{Gly}_7}, \phi_{\text{Gly}_8}, \psi_{\text{Gly}_8}) = (104^\circ, 49^\circ, -75^\circ, -34^\circ, 69^\circ, -99^\circ)$, $\Delta E_{\text{res}} = 0.82 \text{ kcal mol}^{-1}$, and $\Omega_0 = -171^\circ$. It means that interhelical interactions within 5-residues-helix are not sufficient for stabilizing α -helix hairpin structures.

For the case of α_{10} -GGG- α_{10} , the lowest-energy conformation is also a single α -helix. However, the A*AC* conformation is stabilized to the 2nd minimum with $(\phi_{\text{Gly}_{11}}, \psi_{\text{Gly}_{11}}, \phi_{\text{Gly}_{12}}, \psi_{\text{Gly}_{12}}, \phi_{\text{Gly}_{13}}, \psi_{\text{Gly}_{13}}) = (104^\circ, 48^\circ, -75^\circ, -36^\circ, 72^\circ, -96^\circ)$, $\Delta E_{\text{res}} = 0.11 \text{ kcal mol}^{-1}$, and $\Omega_0 = -168^\circ$. The energy difference between a single α -helix and the A*AC* conformation decreases with an increase in number of residues of each α -helix n . For $n \geq 13$, α -helix hairpin structure changes to the lowest-energy conformation and its relative stability increases with an increase in n . These results indicate that $n = 13$ is the critical number of residues of each α -helix for stabilizing α -helix hairpin structures in α_n -GGG- α_n .

Stable Conformation of α_n -GG- α_n

For α_{25} -GG- α_{25} , a total of 43 energy minima was found with $\Delta E_{\text{res}} < 0.33 \text{ kcal mol}^{-1}$. The lowest-energy conformation is a single α -helix. A slightly distorted α -helix hairpin structure was found as the 2nd low-energy conformation taking B*C conformation at Gly-Gly portion with $(\phi_{\text{Gly}_{26}}, \psi_{\text{Gly}_{26}}, \phi_{\text{Gly}_{27}}, \psi_{\text{Gly}_{27}}) = (127^\circ, -14^\circ, -70^\circ, 83^\circ)$ and $\Delta E_{\text{res}} = 0.04 \text{ kcal mol}^{-1}$ (Figure 2). Its Ω_0 value is -168° which is a typical value of α -helix hairpin structure, however, end-to-end distance R_0 is larger than those of the typical α -helix hairpin structures of α_{20} -GGG- α_{20} . That is, its helix-helix contact is not so favorable as those of the α -helix hairpin structures of α_{20} -GGG- α_{20} . All other energy minima have small values of $|\Omega_0|$ (*i.e.*, $|\Omega_0| < 145^\circ$). Moreover, more than two thirds of them have $|\Omega_0| < 120^\circ$. It means that α -helix hairpin structures are not energetically favorable super secondary structures for α_{25} -GG- α_{25} .

Whole conformational feature of other three cases of

α_n -GG- α_n ($n = 5, 10, 15$, and 20) also corresponds to that of α_{25} -GG- α_{25} . The lowest-energy conformation of them are single α -helices, and only one slightly distorted α -helix hairpin structure, which takes B*C conformation at Gly-Gly portion, was found as the 22th, 2nd, 2nd, and 2nd local minima with $\Delta E_{\text{res}} = 0.96, 0.26, 0.11$, and $0.06 \text{ kcal mol}^{-1}$ for α_5 -GG- α_5 , α_{10} -GG- α_{10} , α_{15} -GG- α_{15} , and α_{20} -GG- α_{20} , respectively. Relative stability of such an α -helix hairpin structure increases with an increase in number of residues of α -helix portions. However, two helices interact with each other only in the helix portion close to the Gly-Gly sequence. Therefore, the extent of stabilizing the α -helix hairpin structure is not enough to change it to the essentially stable one by further increase in number of residues of helix. Moreover, all other energy minima also have small values of $|\Omega_0|$ as well as α_{25} -GG- α_{25} .

These results indicate that the Gly-Gly dipeptide sequence is not sufficient as the loop sequence for forming stable α -helix hairpin structures. Therefore, the Gly-Gly-Gly tripeptide sequence is the shortest peptide sequence enough to form stable α -helix hairpin structures as the flexible loop portion.

REFERENCES

1. C. Branden and J. Tooze, "Introduction to Protein Structure," Garland Publishing Inc., New York and London, 1991.
2. S. P. Ho and W. F. Degrad, *J. Am. Chem. Soc.*, **109**, 6751 (1987).
3. L. Regan and W. F. Degrad, *Science*, **241**, 976 (1988).
4. M. H. Hecht, J. S. Richardson, D. C. Richardson, and R. C. Ogden, *Science*, **249**, 884 (1990).
5. J. J. Osterhout, Jr., T. Handel, G. Na, A. Toumadje, R. C. Long, P. J. Connolly, J. C. Hoch, W. C. Jonson, Jr., D. Live, and W. F. Degrad, *J. Am. Chem. Soc.*, **114**, 331 (1992).
6. N. Nishino, H. Mihara, Y. Tanaka, and T. Fujimoto, *Tetrahedron Lett.*, **33**, 5767 (1992).
7. N. Nishino, H. Mihara, H. Uchida, and T. Fujimoto, *Chem. Lett.*, **53** (1993).
8. P. C. Weber and F. R. Salemme, *Nature*, **287**, 82 (1980).
9. K. C. Chou, G. M. Maggiora, G. Nemethy, and H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 4295 (1988).
10. K. C. Chou, G. Nemethy, and H. A. Scheraga, *J. Phys. Chem.*, **87**, 2869 (1983).
11. F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, *J. Phys. Chem.*, **79**, 2361 (1975).
12. S. S. Zimmerman, M. S. Pottle, G. Nemethy, and H. A. Scheraga, *Macromolecules*, **10**, 1 (1977).
13. S. S. Zimmerman and H. A. Scheraga, *Biopolymers*, **16**, 811 (1977).
14. M. Sisido, private communications.
15. Y. Beppu, *Computer and Chemistry*, **13**, 101 (1989).
16. P. J. Flory, "Statistical Mechanics of Chain Molecules", John Wiley & Sons Inc., New York, N.Y., 1969.