Pergamon

Bioorganic & Medicinal Chemistry, Vol. 5, No. 11, pp 2063–2067, 1997

PII: S0968-0896(97)00135-1
0968-0896/97 $17.00 + 0.00

Conformational Preference for Segetalins G and H, Cyclic Peptides with Estrogen-like Activity from Seeds of *Vaccaria segetalis*[1]

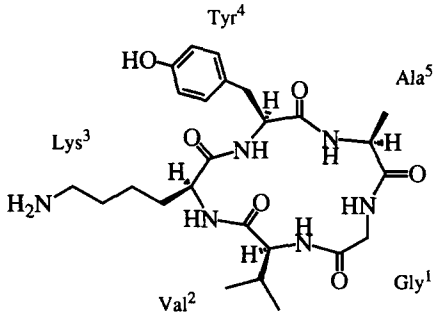
Hiroshi Morita, Young Sook Yun, Koichi Takeya and Hideji Itokawa*
Department of Pharmacognosy, School of Pharmacy, Tokyo University of Pharmacy and Life Science, 1432-1 Horinouchi, Hachioji, Tokyo 192-03, Japan

Abstract—Three-dimensional structures in DMSO-d_6 of segetalins G [*cyclo* (-Gly-Val-Lys-Tyr-Ala-)] and H [*cyclo* (-Gly-Tyr-Arg-Phe-Ser-)], cyclic pentapeptides from seeds of *Vaccaria segetalis*, showing estrogen-like activity, were determined by the distance geometry calculation and restrained energy minimization from NMR data. The backbone structure of segetalin G contains one β-turn: a β II-like turn at Tyr[4]-Ala[5], and that of segetalin H one β-turn: a β II′ turn at Gly[1]-Tyr[2] and one γ-turn at Arg[3]-Phe[4]-Ser[5] sequence. The results of distance comparison analysis proposed a pharmacophore model of estrogen-like cyclic peptides, segetalins. © 1997 Elsevier Science Ltd.

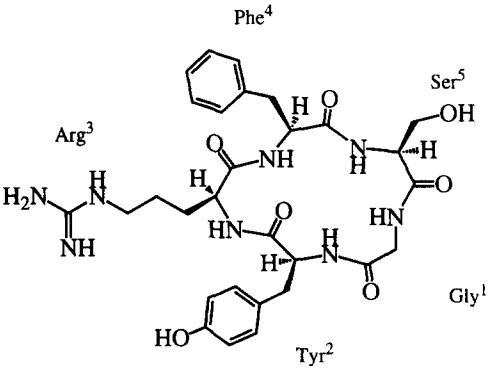
Introduction

In our investigation of bioactive cyclic peptides from higher plants, we reported the structures of segetalins A–H, cyclic peptides isolated from the seeds of *Vaccaria segetalis* (Caryophyllaceae),[2–4] and the solid and solution conformations of segetalins A and B.[5] Segetalins A, B, G, and H, which are cyclic hexa and penta peptides, showed an estrogenic activity in ovariectomized rats.[4,6] It is interesting that segetalins are the only naturally occurring cyclic peptides reported to have an estrogenic activity, which they lose when they are enzymatically digested to acyclic segetalins, suggesting that their conformations play an important role in showing the activity.

In the present studies, the solution state conformations of cyclic pentapeptides, segetalins G and H (Fig. 1) were elucidated by a distance geometry (DG)-molecular dynamics (MD) procedure using the distance constraints from phase sensitive ROESY experiments[7] and the NMR data such as the temperature effects on NH protons and the torsion angles calculated from the $^3J_{NH-C\alpha H}$ coupling constants. In addition, distance comparison (DISCO) analysis[8] was performed for establishing a pharmacophore model of estrogenically active segetalins, which might be useful for elucidation of conformationally induced structure–activity relationships. Our present results describe the preferred conformations of segetalins G and H in DMSO-d_6 and propose a pharmacophore model of estrogenic segetalins on the basis of DISCO.



segetalin G



segetalin H

Figure 1. Structures of cyclic pentapeptides, segatalin G and H. Glycine was provisionally numbered as a first amino acid.

Results and Discussion

Conformational analysis of segetalins G and H

For analysis of the preferential conformations for the structures of segetalins G and H, a computational method, which does not depend on the starting structures, should be useful. The initial 293 structures for segetalin G and 296 structures for segetalin H, satisfying the experimental restraints, were embedded by distance geometry (DG) calculations. The structural calculations using the simulated annealing (SA) protocol with the program SYBYL[9] were carried out and the produced conformers were then subjected to the restrained energy minimization with the AMBER all-atom force field.[10,11] In the SA simulation, each system was equilibrated for 5000 fs in a thermal bath at 800 K, and then successively for 2700 fs, the temperature was decreased in 54 steps to a final temperature of 100 K. Each of the frozen conformations was finally minimized.

NOE relationship of segetalins G and H in DMSO-d_6 are summarized in Figures 2 and 3. Interatomic distances were calculated from the integrated volumes of the ROESY cross peaks and classified into three ranges, 1.8–2.5, 1.8–3.5 and 1.8–5.0 Å, corresponding to strong, medium and weak ROEs, respectively. The distance constraints of the positions of the lack of stereospecific assignments were relaxed by means of the pseudoatom corrections (+1.0 Å for methylene protons and +1.5 Å for methyl protons). Since segetalins G and H are composed of all L amino acid residues and all *trans* amide bonds, the chiral constraints for Cα positions and torsional constraints for amide bonds were also taken into consideration, but the hydrogen bonding constraints were not.

The conformational determination of segetalins G and H by distance geometry and by restrained simulated annealing gave two structural families to each: conformers A and B to segetalin G, and conformers C and D to segetalin H (Table 1).

Conformers A and B for segetalin G

Out of 293 structures, 18 conformers A and 57 conformers B were generated by the DG method. The averaged conformers of them, followed by the energy minimization, are shown in Figure 4. Mean RMS ROE of the restraint violations was 0.01 Å for the two conformers. The overall heavy atomic RMSDs between the individual structures and the mean coordinate positions were 0.62 Å for the backbone atoms of conformer A and 0.96 Å for those of conformer B. The total energy of conformer A (59.323 kcal/mol) was slightly higher than that of conformer B (54.976 kcal/mol). Conformer A contains one intramolecular hydrogen bond between Ala^5-NH and Lys^3-CO (N-H...O 2.046 Å, 141.8°). However, the temperature coefficient[12] of the Ala^5 amide proton of segetalin G measured in DMSO-d_6 was 3.3 ppb/K (Table 2) which suggested that the amide proton was exposed to the solvent and not involved in the intramolecular hydrogen bond. In contrast, the amide proton of Ala^5 in conformer B does not participate in hydrogen bonding as the temperature coefficient of the amide proton of Ala^5 implies. The low temperature coefficient of the amide proton of Lys^3 (2.3 ppb/K) in conformer B was

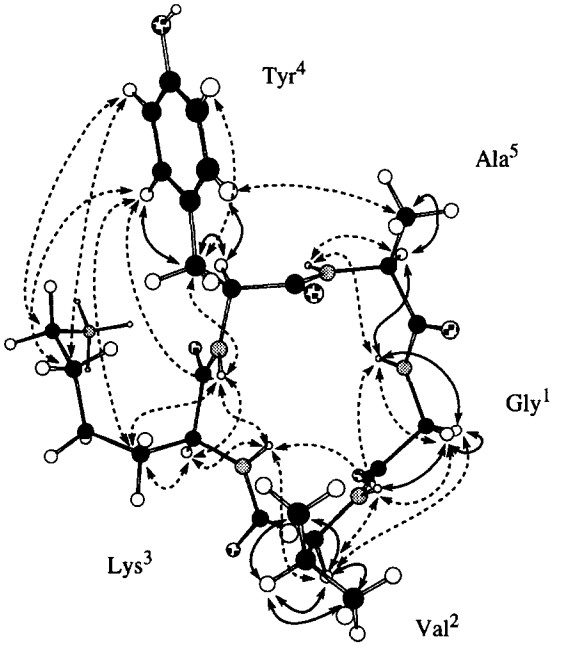


Figure 2. Proposed conformation of segetalin G in solution. Arrows show strong ROE relationship and broken arrows show medium or weak ROE relationship.

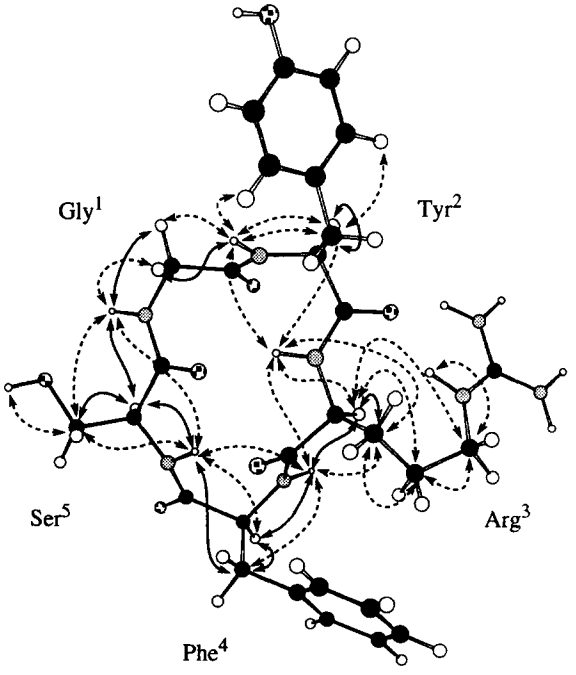


Figure 3. Proposed conformation of segetalin H in solution. Arrows show strong ROE relationship and broken arrows show medium or weak ROE relationship.

Table 1. Results of distance geometry calculations for segetalins G and H

Structural parameters	Segetalin G		Segetalin H	
	Conformer A	Conformer B	Conformer C	Conformer D
No. of constraint				
Distance	48		39	
Chiral	5		5	
Torsion	5		5	
No. of calculated conformer	293		296	
No. of converged conformer	18	57	74	23
Mean energy (kcal/mol)	59.323	54.976	13.305	15.299
Mean RMS ROE	0.01	0.01	0.02	0.04
RMSD for heavy atoms of mean structures (Å)	0.62	0.96	1.89	1.65

considered to be due to its location in the peptide backbone structure. Low energy conformer B, obtained by DG method using ^{1}H NMR information is the proposed solution conformation for segetalin G.

Conformers C and D for segetalin H

Out of 296 structures, 74 conformers C and 23 conformers D structures were generated. The average RMSDs of the restraint violations were 0.02 Å for conformer C and 0.04 Å for conformer D. Figure 4 shows the mean structures of the backbone heavy atoms. Comparison between the backbone atoms and the average structures gave average RMSDs of 1.89 Å for conformer C and 1.65 Å for conformer D. The total energy of conformer D (15.299 kcal/mol) was a little higher than that of conformer C (13.305 kcal/mol).

Both conformers C and D take a γ-turn at Arg^3-Phe^4-Ser^5 sequence. The difference in the two conformers is that conformer C takes another γ-turn at Ser^5-Gly^1-Tyr^2, whereas, conformer D takes a type II′ β-turn at Gly^1-Tyr^2. In conformer C, the amide proton of Tyr^2 was involved in the formation of hydrogen bonding with Ser^5-CO (N-H...O 1.897 Å, 128.8°). However, the temperature coefficient of Tyr^2-NH (7.0 ppb/K) denied the possibility of the amide proton being involved in intramolecular hydrogen bonding. The temperature coefficients[12] (Table 2) indicated that the amide protons of Arg^3 and Ser^5 were involved in intramolecular hydrogen bonds which agreed with the structure of conformer D. The distances between Arg^3-NH and Ser^5-CO, and between Ser^5-NH and Arg^3-CO, and the angles (N-H...O between Arg^3-NH and Ser^5-CO 2.029 Å at 146.4° and N-H...O between Ser^5-NH and Arg^3-CO 1.944 Å at 151.9°), also confirmed the presence of intramolecular hydrogen bondings involving Arg^3 and Ser^5. It has been known that artifacts incorporating γ-turns caused by vacuum effects may be produced, when calculations are carried out in vacuo.[13] The production of conformer C may be due to the vacuum effects. In the final results of our DG calculations, the preferred conformation of segetalin H is conformer D, involving one β-turn structure (type II′ β-turn between Gly^1 and Tyr^2 at two corners), and one γ-turn at Arg^3-Phe^4-Ser^5 sequence.

DISCO analysis

For preparation of a pharmacophore model of segetalins exhibiting estrogenic activity, 2-D structures of estrogenically active compounds must be converted into 3-D structures to translate the 2-D structure–activity information into 3-D requirements for the activity. In the present study, segetalins A, B, G and H were chosen as estrogenically active compounds. For each of the preferred conformations of segetalins G and H derived by the DG method, and those of segetalins A and B reported previously,[5] the points to be considered in the superposition step were chosen and their locations were determined by calculation. These points included not only the atoms in the molecule but also the central points of aromatic ring centroids, and the points locating between the hydrogen-bond donors and acceptors in hydrogen bondings. The points concerning with hydrogen bonds defined in this calculation are amide carbonyl oxygens normally interacted with hydrogen-bond donors in the directions of an angle of 120°, and the distance of 2.9–3.0 Å between the carbonyl oxygen and the hydrogen. On the basis of these chemical similarities such as hydrophobic, hydrogen-bond donating or accepting groups, the preferred conformers of segetalins A, B, G and H were examined

Table 2. Temperature coefficients, $-d\delta/dT$ (10^{-3} ppm/K), of NH protons of segetalins G(**1**), and H(**2**) in ten intervals over the range 300–330 K in DMSO-d_6

Compounds	Temperature coefficients				
Segetalin G (**1**)	4.0 (Gly^1)	4.3 (Val^2)	2.3 (Lys^3)	4.0 (Tyr^4)	3.3 (Ala^5)
Segetalin H (**2**)	6.6 (Gly^1)	7.0 (Tyr^2)	1.7 (Arg^3)	4.0 (Phe^4)	2.3 (Ser^5)

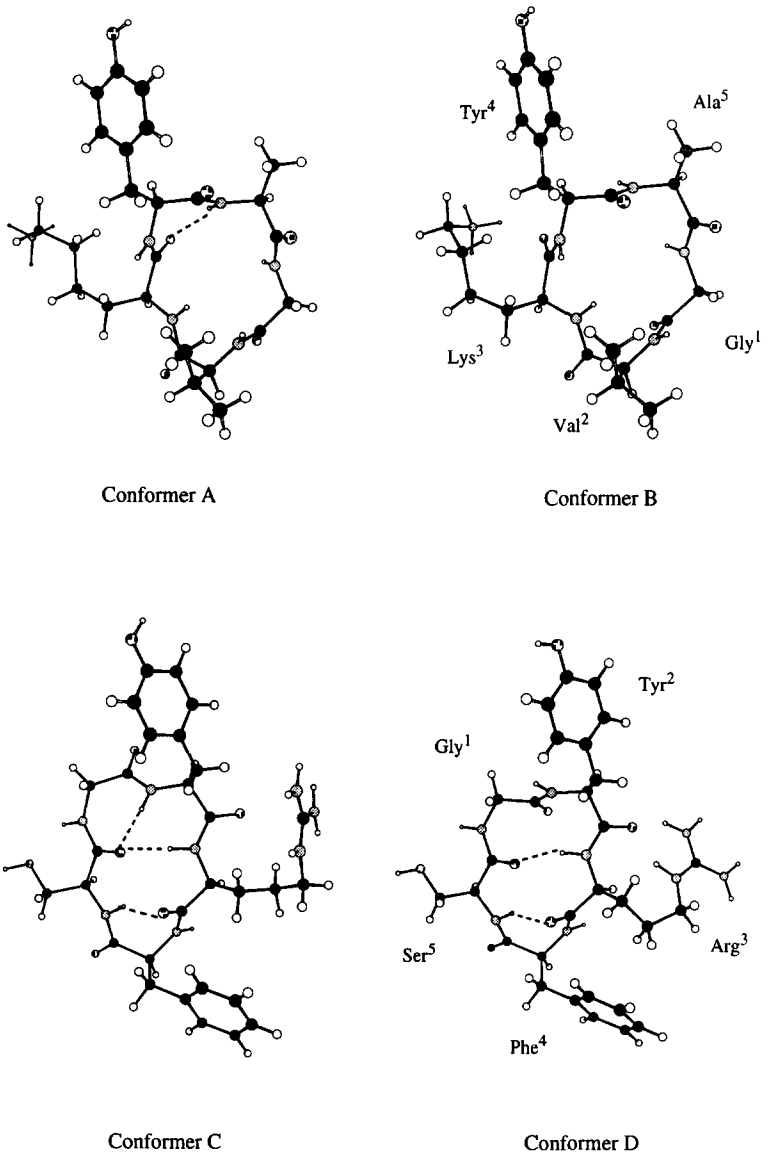


Figure 4. Each averaged conformers A–D calculated by DG methods.

for their superimposability by the program DISCO (DIStance COmparison).[8]

In the present survey, for the least squares fitting, three points, i.e. the aromatic ring centroid of Phe, Tyr, and Trp in segetalins, and the points locating between the hydrogen bonds involving the two electron lone pairs of the amide carbonyl oxygen atoms were chosen. DISCO is usually instructed to iterate the tolerance at which two inter-point distances are considered the same. The maximum tolerance of these three features, i.e. hydrophobic, acceptor atom and donor site shown in Figure 5, among the four conformers of segetalins A, B, G and H was 0.5 Å, indicated that the three-dimensional structures, including these three features, were superimposed (Fig. 5). The mean of the distances among the three features was 6.66 Å. The partial structure of segetalins including these features may be considered as an estrogenically active group.

The methods described in this study may find application to further understanding of the conformation and binding modes of various nonsteroidal estrogens. Solid state conformations and detailed NMR studies of segetalins G and H in reference to their activity are now in progress.

Experimental

Materials

Segetalins G and H were prepared according to the previous procedure.[2,3]

NMR

^{1}H and ^{13}C-NMR spectra were recorded on Varian Unity 400 spectrometers. For the homonuclear and heteronuclear NMR measurements, 10 mg each of segetalins G and H in a 5 mm tube (0.5 ml DMSO-d_6, degassed) was used. The spectra were recorded at 300

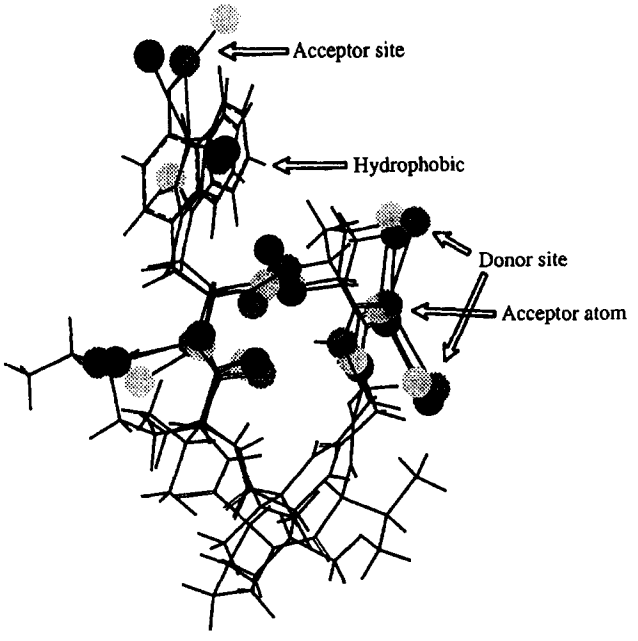


Figure 5. Superimposed structures of segetalins analysed by DISCO.

K. A phase sensitive ROESY experiment was made with a mixing time of 200 ms. The NMR coupling constants (J) are given in Hz. The temperature effect on NH chemical shifts were measured to assess the solvent accessibilities to the amide protons at 10 intervals, over the range of 300–330 K, using a linear regression analysis.

Computational Methods

Computer modeling and all calculations were carried out by using the molecular-modeling software package SYBYL ver. 6.22 (Tripos, Inc., St. Louis, MO) on an IRIS 4D computer. Molecular mechanics and simulated annealing calculations were performed with the AMBER all-atom force field.[10,11] Dielectric constants (ε) were assumed to be proportional to the interatomic distances (r) as $\varepsilon=r$. Solvent molecules were not included in the calculations. The ROE relationships were taken into account in the calculations of the constrained minimizations and dynamics, with an extra harmonic term of the form $E=1/2k(d-d^{low})^2$ for $d<d^{low}$, $1/2k(d^{high}-d)^2$ for $d^{high}<d$ and $E=0.0$ for $d^{low}\leq d\leq d^{high}$ added to the force field [$k=200$ kcal/(mol)(°)^2]. Torsion constraints, with an extra harmonic form, $E=1/2k(\omega-\omega^0)^2$ [$k=0.01$ kcal/(mol)(Å^2)], were also added to the force field. The conformers were divided after minimization into two groups (conformers A and B for segetalin G; conformers C and D for segetalin H). Each energy minimization was carried out until the derivatives became less than 0.01 kcal mol^{-2} Å^{-1} using the MAXMIN program.

(*Received in Japan* 13 *May* 1997; *accepted* 24 *June* 1997)

DISCO analysis

The starting geometries of segetalins A and B were those given in the previous paper,[5] and those of segetalins G and H are given in this paper. Calculations of ring centroids, least square fitting, and excluded volume analyses were also carried out by using DISCO program in SYBYL ver. 6.22. The pharmacophore mapping strategy is the calculation of the location of the ligand and site points, followed by execution of DISCO to find pharmacophore maps.

Acknowledgements

We thank the Ministry of Education, Science and Culture, Japan, for financial support through Grant-in-Aid for General Scientific Research and the Uehara Memorial Foundation for financial support for the Encouragement of Young Scientists. We wish to thank Dr Liu Qian, Tripos Associates, Inc. and Mr Hiroshi Yamashita, Sumisho Electronics Co., Ltd, for support of the molecular simulation study.

References

1. Morita, H.; Yun, Y. S.; Takeya, K.; Itokawa, H. Cyclic peptides from higher plants. Part 44, Part 43. *Heterocycles* in press.

2. Morita, H.; Yun, Y. S.; Takeya, K.; Itokawa, H. *Tetrahedron Lett.* **1994**, *35*, 9593.

3. Morita, H.; Yun, Y. S.; Takeya, K.; Itokawa, H.; Yamada, K. *Tetrahedron* **1995**, *51*, 6003.

4. Yun, Y. S.; Morita, H.; Takeya, K.; Itokawa, H. *J. Nat. Prod.* **1997**, *60*, 216.

5. Morita, H.; Yun, Y. S.; Takeya, K.; Itokawa, H.; Shiro, M. *Tetrahedron* **1995**, *51*, 5987.

6. Itokawa, H.; Yun, Y. S.; Morita, H.; Takeya, K.; Yamada, K. *Planta Medica*. **1993**, *41*, 992.

7. Bother-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jenanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811.

8. Martin, Y. C.; Bures, M. G.; Danaher, E. A.; DeLazzer, J.; Lico, I.; Pavlik, P. A. *J. Comput-Aided Mol. Design* **1993**, *7*, 83.

9. Molecular-modeling software SYBYL ver. 6.22 (Tripos, Inc., St. Louis, MO).

10. Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghgio, C.; Algona, G.; Profeta, Jr. S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.

11. Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.

12. Kessler, H. *Angrew. Chem.* **1982**, *94*, 509; *Int. Ed. Engl.* **1982**, *21*, 512.

13. Kurz, M.; Mierke, D. F.; Kessler, H. *Angew. Chem. Int. Ed. Engl.* **1992**, *31*, 210.