Polymer Journal, Vol. 16, No. 7, pp 553—558 (1984)

Theoretical Conformational Analysis on *N*-Acetyl-*N'*-methylamide of L-Ala–D-Ala Dipeptide

Masahito OKA and Akio NAKAJIMA*

Department of Polymer Chemistry, Kyoto University, Kyoto 606, Japan

(Received March 16, 1984)

ABSTRACT: To investigate the short-range interactions in gramicidin A, conformational energy calculations using an Empirical Conformational Energy Program for Peptides (ECEPP) were carried out on *N*-acetyl-*N'*-methylamide of L-Ala–D-Ala dipeptide. The conformational energy was minimized from the starting conformations which included all combinations of low-energy single-residue minima. It was found that the calculated bend probability of the L-Ala–D-Ala sequence is significantly higher than that of the L-Ala–L-Ala sequence, and that type II β-bend is the most favorable conformation.

KEY WORDS Conformational Analysis / ECEPP / L-Ala–L-Ala / L-Ala–D-Ala / Dipeptide / Type II β-Bend / Gramicidin A /

The molecular force field method is useful for investigating the molecular conformations of globular and fibrous proteins, and peptide hormones. For this purpose, an Empirical Conformational Energy Program for Peptides (ECEPP) was offered by Momany *et al.*,[1] and the conformational analysis of oligopeptides[2-9] with energy minimization procedures was carried out. The local conformational feature on backbone conformations was analyzed as statistically averaged properties in the ensemble composed with low-energy minimums. As one of the statistically averaged properties, β-bend forming tendencies (bend probabilities) of the dipeptide unit were investigated for dipeptides,[3-6] tripeptides[5,7,8] and tetrapeptides.[9] Through these works, it is shown that bend probabilities depend on the amino acid pairs in the dipeptide unit and are also influenced by interactions with the nearest and next-to-nearest neighboring residues.[3-9]

Bandekar *et al.*[10] carried out the conformational analysis of cyclo(L-Ala–D-Ala–Aca),[11] and showed a type II β-bend to be an energetically favorable conformation for the L-Ala–D-Ala dipeptide sequence whose distance between the ith and $i+3$th[12] C^{α} atoms are kept to close to those forming the β-bend.

Gramicidin A is a linear pentadecapeptide with the alternating L- and D-hydrophobic amino acid sequence and also forms the transmembrane channels which select transport cations across a lipid bilayer. Conformations of gramicidin A in solutions were analyzed by Urry *et al.*,[13] Blout *et al.*[14] and Ivanov *et al.*,[15] and different model systems were proposed. Moreover, the head-to-head dimerized single stranded β-helix was presented as a reasonable conformation compatible with the experimental results[16-18] in the lipid bilayer membrane.

In this paper, conformational analysis of Ac–L-Ala–D-Ala–NHMe is carried out as the first step to predict the stable conformations of the polypeptide with the repeating sequence L-Ala–D-Ala as a model polypeptide of gramicidin A by the same procedure used

* To whom correspondence should be addressed.

for the conformational analysis of the elastin model polypeptide Ac-(Val-Pro-Gly-Gly)$_6$-NHMe.[9] Gramicidin A has not the L-Ala-D-Ala repeating sequence but L-Val-Gly-L-Ala-D-Leu-L-Ala-D-Val-L-Val-D-Val-L-Trp-D-Leu-L-Trp-D-Leu-L-Trp-D-Leu-L-Trp. The stable conformations of the model polypeptide having the L-Ala-D-Ala repeating sequence might not exactly coincide with those of gramicidin A. However, intra- and inter-residue interactions in this model polypeptide are supposed to play an important role to stabilize the backbone conformations of gramicidin A, and their backbone conformations would be slightly changed to more stable ones with the side-chain/backbone and side-chain/side-chain interactions owing to the actual amino acid sequence of gramicidin A. An ensemble of low-energy conformations, used as starting conformations in the following steps, was obtained for the Ac-L-Ala-D-Ala-NHMe dipeptide. Moreover, with the calculated bend probabilities and conformational distribution in the (ϕ, ψ) space, intra- and inter-residue interactions in this dipeptide are discussed.

THEORETICAL

The nomenclature and conventions adopted are those recommended by the IUPAC-IUB nomenclature commission.[19] Conformational energy calculations of the Ac-L-Ala-L-Ala-NHMe and Ac-L-Ala-D-Ala-NHMe dipeptides were carried out with ECEPP,[1] using the standard geometry for bond lengths and bond angles, except that the C^{α}–H bond length was increased[20] from 1.00 to 1.09 Å.

Minimization was carried out with the Powell algorithm[21] until the conformational energy did not change by more than 0.001 kcal mol^{-1} between successive iterations. During minimizations, all ϕ, ψ, and χ^1 were allowed to vary. All other dihedral angles were held fixed at 180°. All combinations of single-residue minima[22] were used as starting conformations. The following conformations, $(\phi, \psi) = (-75°, 140°)$ for L-Ala and $(\phi, \psi) = (75°, -140°)$ for D-Ala, were also used as single-residue minima.

The same definitions of bend and hydrogen-bonded conformations as ref 3 were used. The classification of the bend type followed Table I of ref 3. The normalized Boltzmann factor v_i, the relative conformational energy ΔE_i, the distance R, the total bend probability P_b, and the type j bend probability $P_{b,j}$ are defined in ref 3. Conformational space is devided into 16 regions with the conformational letter codes shown in Figure 1 of ref 22.

RESULTS AND DISCUSSION

All minimum energy conformations of Ac-L-Ala-L-Ala-NHMe and Ac-L-Ala-D-Ala-NHMe with $\Delta E < 3$ kcal mol^{-1} are listed in Tables I and II, respectively.

Calculated energy minimums of Ac-L-Ala-L-Ala-NHMe in Table I almost correspond to those of the previous works[4] (Supplementary Table IM of reference 4) in which 1.00 Å was used for the C^{α}–H bond length, (the revised value 1.09 Å is used in this work), except for the following two points. 1) Three new minimums, FE ($\Delta E = 1.03$ kcal mol^{-1}), FF ($\Delta E = 1.21$ kcal mol^{-1} and A*F ($\Delta E = 2.86$ kcal mol^{-1}) were found, and 2) the relative stabilities of several minimums were slightly changed. Table I also indicates that there are no favorable conformations and that the ensemble of Ac-L-Ala-L-Ala-NHMe are composed of many stable conformations.

For Ac-L-Ala-D-Ala-NHMe, 37 minimum conformations were found with values of $\Delta E < 3$ kcal mol^{-1}. All of the four lowest-energy conformations (CA*, CC*, CD*, and FC*) were β-bend conformations, three of them type II β-bend, and the other, type V, which is analogous to type II β-bend. Their total Boltzmann factor was 0.516. This corresponds to the conformational preference of cyclo(L-Ala-D-Ala-Aca) peptide whose two lowest minimums (CD* with $\Delta E = 0.00$

Table I. Calculated minimum energy conformations[a] of Ac-L-Ala-L-Ala-NHMe

Conformational letter code	ΔE[b] / kcal mol^{-1}	v[c]	R / Å	Bend type	Number of hydrogen bonds	$\phi_{\text{L-Ala}_1}$	$\psi_{\text{L-Ala}_1}$	$\phi_{\text{L-Ala}_2}$	$\psi_{\text{L-Ala}_2}$
CC	0.00	0.166	8.83		2	−83	79	−84	79
AC	0.47	0.076	5.74	I	1	−69	−46	−86	76
EC	0.47	0.075	8.68		2	−154	154	−84	80
CF	0.58	0.062	9.24		1	−83	80	−75	141
CD	0.59	0.062	7.97		1	−83	77	−153	63
CE	0.62	0.059	8.76		2	−83	81	−153	154
CG	0.73	0.048	6.79	IV	1	−81	79	−159	−57
CA	0.75	0.047	8.48		1	−83	77	−76	−45
EE	0.81	0.043	10.69		2	−155	154	−154	154
DC	0.87	0.038	9.16		1	−152	72	−84	80
EF	0.97	0.032	9.41		1	−154	154	−75	141
FE	1.03	0.029	10.05		1	−74	137	−154	155
AE	1.13	0.025	6.94	VII	1	−73	−44	−154	153
DE	1.16	0.024	9.16		1	−152	68	−155	153
AD	1.18	0.023	5.01	VII	0	−69	−41	−147	62
FF	1.21	0.022	9.11		0	−75	137	−75	141
ED	1.24	0.021	9.97		1	−155	154	−151	72
AA	1.29	0.019	4.78	III	0	−69	−43	−72	−42
DD	1.46	0.014	8.48		0	−151	68	−152	72
EA	1.47	0.014	8.93		1	−155	154	−74	−45
GC	1.48	0.014	6.90	IV	1	−158	−59	−84	80
DF	1.50	0.013	9.70		0	−151	74	−75	139
CA*	1.60	0.011	5.20	II	0	−79	87	55	55
CA*	1.61	0.011	4.86	II	0	−68	109	55	49
GE	1.72	0.009	8.16		1	−158	−58	−154	155
DA	1.84	0.008	8.78		0	−150	71	−74	−45
GF	1.96	0.006	7.99		0	−159	−59	−77	141
EG	2.00	0.006	9.80		1	−154	153	−158	−58
A*C	2.28	0.004	8.68		1	54	57	−83	81
GD	2.29	0.004	6.63	VII	0	−158	−58	−151	74
DG	2.30	0.004	7.45		0	−150	73	−158	−58
A*E	2.30	0.004	8.53		1	54	56	−155	155
EA*	2.65	0.002	7.00	VII	1	−154	153	54	58
GA	2.70	0.002	5.93	III	0	−158	−59	−74	−45
AG	2.83	0.001	6.28	VII	0	−74	−44	−158	−58
A*F	2.86	0.001	9.00		0	54	57	−76	136
A*D	2.93	0.001	8.25		0	54	56	−152	77
DA*	2.98	0.001	6.24	III′	0	−150	77	54	57

[a] All minimums with $\Delta E < 3$ kcal mol^{-1}.
[b] $E_0 = 0.37$ kcal mol^{-1}.
[c] Values at 300 K.

kcal mol^{-1} and CA* with $\Delta E = 0.74$ kcal mol^{-1}) also corresponded to the type II β-band.

As shown in Table III, the total bend probability of Ac-L-Ala-D-Ala-NHMe (0.614) is significantly higher than that of Ac-L-Ala-L-Ala-NHMe (0.237). This bend probability which increased by substituting residues from L to D was also found for the calculated results of Ac-L-Pro-L-Ala-NHMe ($P_b = 0.255$)[3] and Ac-L-Pro-D-Ala-NHMe ($P_b = 0.758$).[6] These tendencies are supported by the CD spectrum

Table II. Calculated minimum energy conformations[a] of Ac–L-Ala–D-Ala–NHMe

Conformational letter code	ΔE[b] / kcal mol^{-1}	v[c]	R / Å	Bend type	Number of hydrogen bonds	$\phi_{\text{L-Ala}_1}$	$\psi_{\text{L-Ala}_1}$	$\phi_{\text{D-Ala}_2}$	$\psi_{\text{D-Ala}_2}$
CA*	0.00	0.192	4.66	II	1	−71	99	77	36
CC*	0.13	0.154	5.98	V	2	−83	79	84	−78
CD*	0.34	0.109	5.42	II	1	−80	83	153	−49
FC*	0.69	0.061	5.72	II	1	−70	137	86	−75
CF*	0.70	0.059	7.21		1	−84	76	74	−143
EC*	0.79	0.051	7.24		2	−154	154	84	−79
CE*	0.96	0.039	7.18		2	−84	79	154	−151
DC*	1.00	0.036	6.93	IV	1	−152	69	83	−80
EE*	1.06	0.033	10.09		2	−155	154	155	−153
AC*	1.20	0.026	8.73		1	−73	−44	84	−80
CG*	1.26	0.023	6.36	IV	1	−83	75	156	58
AE*	1.32	0.021	8.43		1	−72	−44	154	−155
FE*	1.39	0.019	8.97		1	−75	140	155	−151
DF*	1.39	0.019	8.08		0	−152	65	75	−143
ED*	1.40	0.018	8.95		1	−154	154	152	−73
DE*	1.41	0.018	7.98		1	−152	66	154	−154
EF*	1.44	0.017	8.30		1	−155	153	75	−138
A*C*	1.60	0.013	5.33	I′	1	53	58	86	−75
AF*	1.71	0.011	9.10		0	−73	−44	76	−140
EA*	1.74	0.010	7.25		1	−154	152	74	45
GC*	1.79	0.010	9.19		1	−158	−58	84	−80
DD*	1.80	0.009	6.49	IV	0	−151	72	152	−70
DA*	1.92	0.008	5.86	II	0	−150	74	74	44
AD*	1.96	0.007	8.05		0	−72	−45	151	−76
GE*	1.98	0.007	9.33		1	−158	−58	155	−154
CA	2.31	0.004	8.66		1	−83	79	−54	−58
EG*	2.32	0.004	9.71		1	−155	153	158	58
GF*	2.35	0.004	9.72		0	−158	−58	76	−140
GD*	2.49	0.003	8.72		0	−158	−58	151	−74
DG*	2.49	0.003	7.09		0	−150	69	158	58
AA*	2.54	0.003	8.37		0	−74	−44	75	44
AG*	2.63	0.002	6.40	VII	0	−71	−42	159	58
A*A*	2.64	0.002	4.47	III′	0	54	56	72	44
A*E*	2.71	0.002	6.90	VII	1	54	55	154	−152
AA	2.83	0.002	5.24	III	0	−68	−44	−54	−52
GA*	2.91	0.001	8.83		0	−159	−58	74	44
EA	2.92	0.001	8.64		1	−155	153	−54	−58

[a] All minimums with $\Delta E < 3$ kcal mol^{-1}.
[b] $E_0 = 0.12$ kcal mol^{-1}.
[c] Values at 300 K.

measurements[6,23] of Dnp–Gly–L-Pro–X–Gly–*p*NA (X=L-Ala and D-Ala) which have two chromophores at N- and C-terminal (dinitrophenyl and *p*-nitroanilide, respectively). That is, more significant Cotton effects around 310 and 350 nm were observed for the L-Pro–D-Ala containing peptide than the L-Pro–L-Ala containing peptide. The most energetically favorable five conformations of Ac–L-Ala–D-Ala–NHMe were quite similar to those of Ac–L-Pro–D-Ala–NHMe,[6,23] *i.e.*, three of them, type II bend and one, type V bend.[24] These results suppose that the high bend probability of the dipeptide consisting of residues with

opposite optical activities is essentially independent of the type of amino acid.

L-Ala–D-Ala sequence exists in the HC-toxin isolated from culture filtrates of *Helminthosporium carbonum*. HC-toxin, a cyclic peptide composed of the L-Ala–D-Ala–L-Aeo–D-Pro sequence[25] and the L-Ala–D-Ala sequence is fixed to the β-bend forming structure by geometrical constraint. Conformations of HC-toxin in $CDCl_3$ have recently been proposed on the basis of the results of the vicinal coupling constant $^3J_{NH-C^{\alpha}H}$ and the observed Nuclear Overhauser Effect (NOE). For the L-Ala–D-Ala sequence, these results show that ($\phi_{L\text{-}Ala}$, $\psi_{L\text{-}Ala}$, $\phi_{D\text{-}Ala}$, $\psi_{D\text{-}Ala}$)=($-110°$, 110°, 60°, $-50°$)[26] and ($\phi_{L\text{-}Ala}$, $\psi_{L\text{-}Ala}$, $\phi_{D\text{-}Ala}$)=($-120°$, 95°, 110°).[27] Both of these conformations are of type II β-bond, and consistent with our calculated four low-energy conformations. HC-toxin in $CDCl_3$ is stabilized by two hydrogen bonds[26,27] (*i.e.*, (L-Ala)-NH···OC(Aeo) and (Aeo)NH···OC(L-Ala)). The latter corresponds to the (NHMe)NH···OC(L-Ala) hydrogen bond, and the former cannot be formed in Ac–L-Ala–D-Ala–NHMe. Seven conformations have (NHMe)NH···OC(L-Ala) hydrogen bonds, and the calculated probability for these conformations is 0.35. These high hydrogen-bond forming tendencies at (NHMe)NH···OC(L-Ala) are also consistent with the experimental results in $CDCl_3$.[26,27] Moreover, it is expected that this type of hydrogen-bond might exist in the local conformations of gramicidin A as favorable short-range interactions in the L-Ala–D-Ala sequence.

Table III. Calculated bend probabilities of dipeptides

Bend type	L-Ala–L-Ala	L-Ala–D-Ala	L-Pro–L-Ala[a]	L-Pro–D-Ala[b]
I	0.076	—	0.072	—
II	0.023	0.370	0.022	0.503
III	0.021	0.002	0.040	—
IV	0.062	0.069	0.086	0.036
V	—	0.154	—	0.219
VII	0.054	0.004	0.035	—
I′	—	0.013	—	—
III′	0.001	0.002	—	—
Total	0.237	0.614	0.255	0.758

[a] From ref 3. [b] From ref 6.

In Table IV, the calculated probabilities of the occurrence of the conformations shown by the conformational letter code are given. The conformational probabilities of L-Ala_1 and L-Ala_2 of Ac–L-Ala_1–L-Ala_2–NHMe[3], and L-Ala

Table IV. Calculated probabilities of the occurrence of the various minimum-energy residue conformations of Ac–Ala–NHMe and Ac–C–Y–NHMe

Conformation	Single residue[b]	L-Ala_1–L-Ala_2		L-Ala–D-Ala		L-Pro–L-Ala[c]	L-Pro–D-Ala[d]
	Ala	L-Ala_1	L-Ala_2	L-Ala	D-Ala[e]	L-Ala	D-Ala[e]
A	0.073	0.143	0.089	0.072	0.216	0.114	0.286
C	0.484	0.467	0.372	0.580	0.350	0.374	0.301
D	0.145	0.102	0.125	0.092	0.147	0.132	0.192
E	0.255	0.193	0.192	0.135	0.138	0.148	0.087
F	—	0.051	0.137	0.079	0.109	0.115	0.093
G	0.034	0.034	0.059	0.025	0.033	0.095	0.036
A*	0.009	0.010	0.026	0.017	0.007	0.022	0.005

[a] Values at 300 K.
[b] From ref 22.
[c] From ref 3.
[d] From ref 6.
[e] Values are represented by the order of conformational letter codes, A*, C*, D*, E*, F*, G*, and A.

of Ac–L-Ala–D-Ala–NHMe are almost the same as those of Ac–L-Ala–NHMe.[22] However, L-Ala and D-Ala of Ac–L-Ala–D-Ala–NHMe, and D-Ala of Ac–L-Pro–D-Ala–NHMe[6] indicated somewhat deviated conformational distributions from those of Ac–L-Ala–NHMe. The above results suggest 1) that the nearest-neighbor inter-residue interactions are not so important as the intra-residue interactions for the residues in the dipeptide sequence with the same optical activity, but 2) that the conformational stabilities of the residues in the dipeptide sequence having opposite optical activities are influenced by the nearest-neighbor interactions. That is, the assumption[28] of independent rotation about the virtual bond between successive C^{α} atoms in the polypeptide chain can be held for the homopolypeptide chain, but a more accurate analysis is desirable for the copolypeptide chains composed of residues with different optical activities.

Acknowledgments. We should like to thank Professor H. A. Scheraga, Department of Chemistry, Cornell University, for presenting the program ECEPP and the minimization program POWELL. We also thank Dr. S. Yamaguchi for his helpful comments on this paper.

Computations were carried out by an FACOM M-380 Computer at the Data Processing Center, Kyoto University.

REFERENCES

1. F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, *J. Phys. Chem.*, **79**, 2361 (1975).
2. G. Nemethy and H. A. Scheraga, *Ouart. Rev. Biophys.*, **10**, 239 (1977). Many papers were cited in this review.
3. S. S. Zimmerman and H. A. Scheraga, *Biopolymers*, **16**, 811 (1977).
4. S. S. Zimmerman and H. A. Scheraga, *Biopolymers*, **17**, 1849 (1978).
5. M. Oka, G. T. Montelione, and H. A. Scheraga, submitted to *J. Am. Chem. Soc.*
6. M. Kawai, K. Sato, R. Sugawara, and U. Nagai, *Bull. Chem. Soc. Jpn.*, **56**, 3005 (1983).
7. M. Oka and A. Nakajima, to be submitted to *Polym. J.*
8. M. Oka and A. Nakajima, unpublished data.
9. M. Oka and A. Nakajima, to be submitted to *J. Am. Chem. Soc.*
10. J. Bandekar, D. J. Evans, S. Krimm, S. J. Leach, S. Lee, J. R. McQuie, E. Minasian, G. Nemethy, M. S. Pottle, H. A. Scheraga, E. R. Stimson, and R. W. Woody, *Int. J. Peptide Protein Res.*, **19**, 187 (1982).
11. Cyclo(L-alanyl–D-alanyl–ε-aminocaproyl) is abbreviated cyclo(L-Ala–D-Ala–Aca). Acetyl and methylamide groups are also abbreviated as Ac and NHMe, respectively.
12. The *i*th and $i+3$th C^{α} atoms are defined as CA and CE in Figure 2 of ref. 10.
13. J. D. Glickson, D. F. Mayeres, J. M. Settine, and D. W. Urry, *Biochemistry*, **11**, 477 (1972).
14. S. Weinstein, B. A. Wallace, E. R. Blout, J. S. Morrow, and W. Veatch, *Proc. Natl. Acad. Sci. U.S.A.*, **76**, 4230 (1979).
15. V. T. Ivanov and S. V. Sychev, *Proc. Workshop Conference Hoechst*, **11**, 107 (1981).
16. B. A. Wallace, W. R. Veatch, and E. R. Blout, *Biochemistry*, **20**, 5754 (1981).
17. S. B. Hladky and D. A. Hydon, *Nature*, **225**, 451 (1970).
18. D. W. Urry, *Biochim. Biophys. Acta*, **265**, 115 (1972).
19. IUPAC-IUB Commission on Biological Nomenclature, *Biochemistry*, **9**, 3471 (1970).
20. E. Benedetti, C. Pedone, C. Toniolo, G. Nemethy, M. S. Pottle, and H. A. Scheraga, *Int. J. Peptide Protein Res.*, **16**, 156 (1980).
21. M. J. D. Powell, *Computer J.*, **7**, 155 (1964).
22. S. S. Zimmerman, M. S. Pottle, G. Nemethy, and H. A. Scheraga, *Macromolecules*, **10**, 1 (1977).
23. K. Sato, M. Kawai, and U. Nagai, *Biopolymers*, **20**, 1921 (1981).
24. The bend type of the 3rd CD* and the 5th FC* conformations in Table I of ref 6 is better to assigned to type II bend than to type V bend according to the definition of bend types.[3]
25. An unusual amino acid, 2-amino-8-oxo-9,10-epoxydecanoic acid, is abbreviated as Aeo.
26. M. Kawai, D. H. Rich, and J. D. Walton, *Biochem. Biophys. Res. Commun.*, **111**, 398 (1983).
27. P. Mascagni, M. Pope, W. A. Gibbons, L. M. Ciuffetti, and H. W. Knoche, *Biochem. Biophys. Res. Commun.*, **113**, 10 (1983).
28. D. A. Brant and P. J. Flory, *J. Am. Chem. Soc.*, **87**, 2791 (1965).