Bull. Chem. Soc. Jpn., ***56***, *3005—3008 (1983)*

Studies on the *β*-Turn of Peptides. V.[1] Turn Preference of Peptide Sequences Analyzed by Conformational Energy Calculation on Dipeptide Models and by CD Spectra of Chromophoric Derivatives

Masao Kawai,*,† Kazuki Sato, Rika Sugawara, and Ukon Nagai
Mitsubishi-Kasei Institute of Life Sciences, Minamiooya 11, Machida, Tokyo 194
†*Department of Synthetic Chemistry, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466*
(Received April 8, 1983)

Empirical conformational energy calculations were performed on Ac–L-Pro–D-Ala–$NHCH_3$ and Ac–D-Ala–L-Pro–$NHCH_3$. In order to study the turn tendency of the dipeptide sequences turn probabilities and average interterminal distances were calculated. Much larger turn preference of D-Ala containing sequences was indicated compared with the corresponding L-Ala containing sequences. The turn tendency of the dipeptide sequences were shown to be in the order: L-Pro–D-Ala>D-Ala–L-Pro>L-Pro–L-Ala>L-Ala–L-Pro. This order coincides with the order of the magnitude of the Cotton effects in the CD spectra of the chromophoric derivatives having 2,4-dinitrophenylglycyl and glycine-*p*-nitroanilide at the *N*- and the *C*-termini, respectively, of the corresponding dipeptide sequences. This agreement supports the validity of the chiroptical method to study the turn tendency of the peptide sequences using chromophoric derivatives reported earlier.

CD spectra of *N*-(2,4-dinitrophenyl)tetrapeptide *p*-nitroanilides (Dnp–tetrapeptide–pNA)[2] exhibit characteristic Cotton effects when they take *β*-turn conformation. Exciton coupling of the electric transition moments of the two chromophores were assumed to explain the spectra and the mgnitude of the Cotton effects around 310 and 350 nm were shown to reflect the *β*-turn preference of the derivatives.[3] A question was raised, however, whether such chiroptical method using chromophoric derivatives as mentioned above reveals well the conformational characteristics of the original peptide sequences lacking those chromophoric groups. For the purpose of answering the question we have attempted empirical conformational energy calculations on model peptides having acetyl and *N*-methylamide groups at the *N*- and the *C*-termini, respectively. In this paper the conformational energy studies on model dipeptides are described and the results are compared with the CD spectral data of tetrapeptide derivatives having Dnp–Gly and Gly–pNA moieties at the *N*- and the *C*-termini, respectively.

Results and Discussion

The CD spectra of Dnp–Gly–X–Y–Gly–pNA are shown in Fig. 1 where X–Y represents a dipeptide sequence consisting of L- or D-Ala and L-Pro residues. Marked difference in the magnitude of the Cotton effects observed for this series of tetrapeptide derivatives can be attributed to the divergent *β*-turn tendency of the central dipeptide sequences X–Y. Theoretical study has been attempted to elucidate the turn preference of the X–Y sequences by means of empirical conformational energy calculation on the model dipeptides Ac–X–Y–$NHCH_3$. Computations were carried out using the program ECEPP[4] with the Powell's minimization subroutine.[5] The conformational properties of Ac–L-Pro–L-Ala–$NHCH_3$ and Ac–L-Ala–L-Pro–$NHCH_3$ had been already elucidated by Zimmerman and Scheraga[6] by use of ECEPP. Essentially same procedure as reported by them was followed for the calculation of Ac–L-Pro–D-Ala–$NHCH_3$ and Ac–D-Ala–L-Pro–$NHCH_3$ except that the contribution of librational entropy is not included in our calculation.

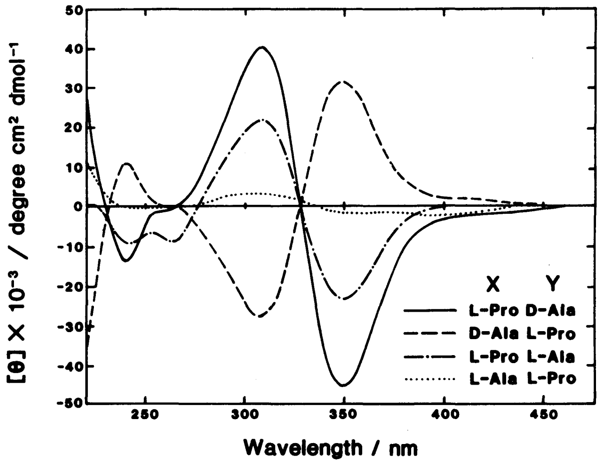

Fig. 1. CD spectra of Dnp–Gly–X–Y–Gly–pNA in methanol.

For the calculation of Ac–L-Pro–D-Ala–$NHCH_3$ all combination of low energy conformations of Ac–L-Pro–$NHCH_3$[7] (C −75°, −79°; F −75°, 159°; A −75°, 48°; A −75°, −23°)[8] and of Ac–D-Ala–$NHCH_3$[7] (C* 84°, −79°; E* 154°; −153°; D* 150°, −72°; F* 75°, −145°; A* 74°, 45°; G* 158°, 85°; A −54°, −57°)[8] were used as the starting conformations for energy minimization. All *ω* values were fixed to 180°. In Table 1 are listed 19 low-energy conformations obtained through energy minimization of Ac–L-Pro–D-Ala–$NHCH_3$. The interatomic distance *R*, defined as the distance between the methyl carbons of the acetyl and the *N*-methylamide groups, and the O(Ac)···N($NHCH_3$) distance *h* are given along with the relative conformational energy ΔE and the normalized statistical weight *v* of each conformation. For each conformation with the *R* value less than 7.0 Å the bend type is given according to the classification by Zimmerman and Scheraga.[6] In the case of energy minimization of Ac–D-Ala–L-Pro–$NHCH_3$ some starting conformations having *cis* peptide linkage which were taken from the minimum energy conforma-

TABLE 1. MINIMUM-ENERGY CONFORMATIONS[a)] OF Ac–L-Pro–D-Ala–NHCH$_3$

Conformation letter code[b)]	ΔE[c)]/kcal mol^{-1}	v[b)]	h[d)]/Å	R[d)]/Å	Bend[b)] type	Dihedral angles[e)]/°			
						ψ_{Pro}	$\phi_{\mathrm{D\text{-}Ala}}$	$\psi_{\mathrm{D\text{-}Ala}}$	$\chi_{\mathrm{D\text{-}Ala}}$
C A*	0.00	0.278	3.59	4.81	II	82	78	42	59
C C*	0.14	0.219	4.87	5.93	V	78	85	−77	59
C D*	0.30	0.168	3.39	5.33	V	80	153	−47	59
C F*	0.79	0.074	5.72	7.19		77	73	−142	58
F C*	0.95	0.057	4.29	6.37	V	159	85	−76	59
C E*	1.10	0.044	5.11	7.13		80	154	−149	61
C G*	1.33	0.030	3.80	6.34	IV	76	155	58	66
A C*	1.43	0.025	6.32	8.83		−47	84	−80	59
A E*	1.51	0.022	7.23	8.65		−47	154	−155	60
F E*	1.55	0.021	6.47	9.60		161	155	−152	61
F D*	1.69	0.016	4.97	8.22		154	149	−68	61
A F*	1.97	0.010	6.75	9.21		−46	76	−138	58
A F*	2.02	0.009	6.71	9.00		−27	77	−139	58
A D*	2.10	0.008	6.80	8.15		−47	150	−68	61
C A	2.38	0.005	5.43	8.67		78	−54	−58	55
A A*	2.56	0.004	6.50	8.47		−45	75	44	59
A A*	2.56	0.004	6.24	7.94		−25	75	43	59
A G*	2.69	0.003	6.22	5.79	IV	−23	159	58	67
A G*	2.75	0.003	5.91	6.65	IV	−44	159	58	66

a) All minimum-energy conformations with $\Delta E < 3$ kcal mol^{-1} are listed. b) See Ref. 6. c) Relative energy taking the energy value ($E = -14.58$ kcal mol^{-1}) of the lowest-energy conformation as 0.0 kcal mol^{-1}. d) See text. e) Only variable dihedral angles are listed; all others are 180°, except $\phi_{\mathrm{Pro}} = -75°$.

TABLE 2. MINIMUM-ENERGY CONFORMATIONS[a)] of Ac–D-Ala–L-Pro–NHCH$_3$

Conformation letter code[b)]	ΔE[c)]/kcal mol^{-1}	v[b)]	h[d)]/Å	R[d)]/Å	Bend type[b)]	Dihedral angles[e)]/°				
						$\phi_{\mathrm{D\text{-}Ala}}$	$\psi_{\mathrm{D\text{-}Ala}}$	$\omega_{\mathrm{D\text{-}Ala}}$	$\chi_{\mathrm{D\text{-}Ala}}$	ψ_{Pro}
D*C	0.00	0.243	6.76	6.73	IV	151	−72	173	61	78
F*C	0.13	0.198	3.95	5.63	II′	70	−143	175	60	76
E*C	0.31	0.144	6.67	7.02		153	−146	179	64	79
D*F	0.41	0.122	7.78	7.83		152	−71	172	61	164
F*A	0.68	0.078	3.18	5.36	II′	71	−140	175	59	−18
A C	0.68	0.078	3.24	4.88	I	−53	−69	169	57	75
E*F	0.99	0.046	7.95	8.66		154	−147	180	65	161
D*A	1.22	0.032	5.95	5.59	III	150	−72	172	61	−46
E*A	1.48	0.020	6.33	7.32		153	−147	−179	64	−48
D*F	1.69	0.014	3.88	6.62	VI	147	−78	−4	62	160
D*F	1.69	0.014	5.79	7.61		80	−145	180	60	158
A F	2.36	0.005	4.69	6.17	I	−53	−69	169	56	151
A A	2.53	0.003	3.47	4.06	III	−52	−68	171	57	−39
D*A	2.70	0.003	5.46	7.88		153	−76	−7	62	−46

a) All minimum-energy conformations with $\Delta E < 3$ kcal mol^{-1} are listed. b) See Ref. 6. c) Relative energy taking the energy value (-13.12 kcal mol^{-1}) of the lowest-energy conformation as 0.0 kcal mol^{-1}. d) See text. e) Only valuable dihedral angles are listed; all others are 180°, except $\phi_{\mathrm{Pro}} = -75°$.

tions with *cis* Ac–Pro bond of Ac–L-Pro–NHCH$_3$[7)] (A −75°, −48°; F −75°, 162°)[8)] were also included. The results of energy minimization are given in Table 2.

In order to evaluate the turn tendency of the dipeptide sequences the bend probability P_b and the averaged interterminal distance $\langle R \rangle$ were calculated for each model peptide. As defined by Zimmerman and Scheraga[6)] the quantity P_b is the probability of occurrence (or mole fraction) of bend conformations ($R < 7$ Å) in a particular dipeptide and the quantity $\langle R \rangle$ is the statistical average of the interterminal distance R.[6)] In Table 3 are shown the calculated P_b and $\langle R \rangle$ values for the D-Ala containing model peptides along with the reported values[6)] for Ac–L-Pro–L-Ala–NHCH$_3$ and Ac–L-Ala–L-Pro–NHCH$_3$. The larger P_b values and the smaller $\langle R \rangle$ values for the D-Ala containing dipeptides compared with the corresponding L-Ala derivatives indicated the larger turn tendency of D-Ala containing sequences. Stability of turn conformation of LL and LD sequences and the role of proline in the formation of the LL and the LD bends were studied by Chandrasekaran *et al.*[9)] Our results afforded semiquantitative evaluation of turn tendency of these sequences.

As can be seen from Table 3 theoretical conforma-

TABLE 3. MAGNITUDE OF COTTON EFFECTS OF Dnp–Gly–X–Y–Gly–pNA AND BEND PROBABILITY P_b[a)] AND AVERAGE INTERTERMINAL DISTANCE $\langle R \rangle$[a)] of Ac–X–Y–NHCH$_3$

X	Y	CD Spectral data $\lvert [\theta]^{max}_{\approx 350} - [\theta]^{max}_{\approx 310} \rvert$	Calculated value	
			P_b	$\langle R \rangle$/Å
L-Pro	D-Ala	84.7 × 10^3	0.76	6.06
D-Ala	L-Pro	58.8 × 10^3	0.65	6.51
L-Pro	L-Ala	44.8 × 10^3	0.26[b)]	7.99[b)]
L-Ala	L-Pro	3.3 × 10^3	0.02[b)]	9.13[b)]

a) See Ref. 6. b) Values taken from Ref. 6.

tional study clearly demonstrated the turn tendency of the dipeptide sequences to be in the order: L-Pro–D-Ala>D-Ala–L-Pro>L-Pro–L-Ala>L-Ala–L-Pro. This order coincides with the order of the magmitude of the Cotton effects in the CD spectra of Dnp–Gly–X–Y–Gly–pNA as summarized in Table 3. In other words turn preference of dipeptide sequences studied by the chiroptical approach has been shown to agree with the results of theoretical conformational analysis, indicating the usefulness of the chiroptical method. Though librational entropy was not calculated in our study, the sequence dependence of turn tendency obtained in this study is assumed to be valid since the deviation of the $\langle R \rangle$ and P_b values due to neglection of librational entropy is at most 0.36 Å and 0.14, respectively, among the 12 proline containing dipeptides.[6)]

Conformational properties of the tetrapeptide derivatives studied by their CD spectra were discussed with relation to the intrinsic turn tendency of the central dipeptide units disclosed by theoretical study. Conformational energy calculations were performed on the dipeptide models which contain acetyl and *N*-methylamide groups in place of Dnp–Gly and Gly–pNA moieties, respectively, of the tetrapeptide derivatives. These achiral terminal moieties in the chromophoric tetrapeptide derivatives, however, make some contribution to the absolute turn tendency of these molecules. Acetyl and *N*-methylamide groups of the dipeptide models correspond to the CH$_2$CO group of the 1st Gly residue and the NHCH$_2$ group of the 4th Gly residue, respectively, but Dnp–NH and CO–pNA groups were not included in the calculation. It was shown that intramolecular interaction between the two chromophoric groups stabilizes turn conformers without changing relative turn prefernce of the peptide sequences of the same series of compounds.[10)] Replacement of the 1st and the 4th Gly residues of Dnp–Gly–X–Y–Gly–pNA by chiral amino acids markedly changes the turn tendency,[11)] which is not dealt with in this paper.

In summary intrinsic turn tendency of proline containing dipeptide sequences were estimated semiquantitatively by means of empirical conformational energy calculation on model peptides. Calculation on Ac–L-Pro–D-Ala–NHCH$_3$ and Ac–D-Ala–L-Pro–NHCH$_3$ demonstrated the larger turn tendency of these D-Ala containing dipeptides than the corresponding L-Ala analogs, which is consistent with the CD spectral data of the tetrapeptide derivatives having Dnp–Gly and Gly–pNA groups. We consider that the results described here proved the validity of the chiroptical method using Dnp,pNA derivatives.[3)] NMR study of acetyltetrapeptide *N*-methylamide also supported the chiroptical analysis as described elsewhere.[10)]

Experimental

Materials. Synthesis of Dnp–Gly–L-Pro–L-Ala–Gly–pNA,[12)] Dnp–Gly–L-Pro–D-Ala–Gly–pNA,[12)] and Dnp–Gly–D-Ala–L-Pro–Gly–pNA[13)] were already described.

Synthesis of Dnp–Gly–L-Ala–L-Pro–Gly–pNA. All the melting points were measured on a Yanagimoto micromelting point apparatus and were uncorrected. TLC's were carried out on Merck silica gel 60 F$_{254}$ plates with the following solvent systems: R_f^1, CHCl$_3$–MeOH (5 : 1, v/v); R_f^2, CHCl$_3$–MeOH–AcOH (95 : 5 : 1, v/v); R_f^3, *n*-BuOH–AcOH–pyridine–H$_2$O (4 : 1 : 1 : 2, v/v). Optical rotations were measured on an Union automatic polarimeter PM-201.

Boc–L-Pro–Gly–pNA. To a chilled solution of Boc–L-Pro–OH (1.075 g, 5 mmol) and H–Gly–pNA (0.975 g, 5 mmol) in DMF (20 ml) was added EDC·HCl (0.959 g, 5 mmol). The mixture was stirred at 0 °C for 1 h and at room temperature overnight and evaporated under reduced pressure. The residue was dissolved in EtOAc, and the solution was washed successively with 10% citric acid, 4% NaHCO$_3$, and water, dried (Na$_2$SO$_4$), and evaporated. The residue was solidified by addition of ether and petroleum ether, and the product was recrystallized from EtOAc–ether–petroleum ether; yield, 1.523 g (77%); mp 204—205 °C; $[\alpha]_D^{22}$ −18.8° (*c* 0.5, MeOH); R_f^1 0.70, R_f^2 0.41, R_f^3 0.83.

Found: C, 55.34; H, 6.23%; N, 14.20%. Calcd for C$_{18}$H$_{24}$O$_6$N$_4$: C, 55.09; H, 6.17%; N, 14.28%.

H–L-Pro–Gly–pNA·HCl. Boc–L-Pro–Gly–pNA (785 mg, 2 mmol) was dissolved in 0.1 M (1 M=1 mol dm^{-3}) hydrogen chloride in formic acid (24 ml). The solution was allowed to stand at room temperature for 30 min and evaporated under reduced pressure to leave crystals, which were collected by filtration with the aid of ether; yield, 651 mg (99%); mp 217—219 °C; $[\alpha]_D^{22}$ −25.6° (*c* 1, MeOH); R_f^1 0.13, R_f^3 0.53.

Found: C, 47.40; H, 5.15; N, 17.12%. Calcd for C$_{13}$H$_{17}$O$_4$N$_4$Cl: C, 47.49; H, 5.21; N, 17.05%.

Boc–L-Ala–L-Pro–Gly–pNA. To a chilled solution of H–L-Pro–Gly–pNA·HCl (625 mg, 1.9 mmol) were added TEA (0.27 ml, 1.9 mmol) and Boc–L-Ala–ONSu (653 mg, 2.3 mmol). The mixture was treated as described for the synthesis of Boc–L-Pro–Gly–pNA. The product was recrystallized from MeOH–ether; yield, 511 mg (59%); mp 220—222 °C; $[\alpha]_D^{22}$ −12.6° (*c* 1, MeOH); R_f^1 0.64, R_f^3 0.23, R_f^3 0.76.

Found: C, 54.31; H, 6.25; N, 15.11%. Calcd for C$_{21}$H$_{29}$O$_7$N$_5$: C, 54.42; H, 6.31; N, 15.11%.

H–L-Ala–L-Pro–Gly–pNA·HCl. Boc–L-Ala–L-Pro–Gly–pNA (278 mg, 0.6 mmol) was treated with 0.1 M hydrogen chloride in formic acid (7 ml) as described for H–L-Pro–Gly–pNA·HCl. The product was highly hygroscopic crystals; yield, 240 mg (100%); $[\alpha]_D^{22}$ −11.0° (*c* 1, MeOH); R_f^1 0.07, R_f^3 0.54.

Found: C, 45.84; H, 5.84; N, 16.56%. Calcd for C$_{16}$H$_{22}$O$_5$N$_5$Cl·H$_2$O: C, 45.99; H, 5.79; N, 16.76%.

Boc–Gly–L-Ala–L-Pro–Gly–pNA. This compound was prepared from H–L-Ala–L-Pro–Gly–pNA·HCl (200 mg, 0.5 mmol) and Boc–Gly–ONSu (163 mg, 0.6 mmol) as described for Boc–L-Ala–L-Pro–Gly–pNA; yield, 151 mg (58%); mp 135—138 °C; $[\alpha]_D^{22}$ −34.2° (*c* 1, MeOH); R_f^1 0.52, R_f^2 0.09, R_f^3 0.77.

Found: C, 52.47; H, 6.04; N, 15.82%. Calcd for $C_{23}H_{32}O_8N_6 \cdot 1/2\ H_2O$: C, 52.17; H, 6.28; N, 15.87%.

*H-Gly-*L*-Ala-*L*-Pro-Gly-pNA·HCl.* Boc-Gly-L-Ala-L-Pro-Gly-pNA (78 mg, 0.15 mmol) was dissolved in 0.1 M hydrogen chloride in formic acid (1.8 ml). After 30 min, the solution was evaporated to leave an oil, which was used for the next reaction without further treatment; yield was quantitative; R_f^1 0.02, R_f^3 0.51.

*Dnp-Gly-*L*-Ala-*L*-Pro-Gly-pNA.* To a solution of H-Gly-L-Ala-L-Pro-Gly-pNA·HCl (69 mg, 0.15 mmol) in DMF (1 ml) were added TEA (0.063 ml, 0.45 mmol) and N_2ph-F (56 mg, 0.3 mmol). The reaction mixture was stirred at room temperature for 3 h and evaporated *in vacuo.* The residue dissolved in $CHCl_3$ (10 ml) was applied to a column (1.8×18 cm) of silica gel 60 (Merck) and the column was washed with $CHCl_3$. The desired product was eluted with a mixture of $CHCl_3$ and MeOH (5 : 1, v/v). The fractions containing the desired product were evaporated to leave an oil, which was crystallized by addition of ether and recrystallized from MeOH-ether; yield, 73 mg (83%); mp 162—163 °C; $[\alpha]_D^{22}$ −26.2° (*c* 1, DMF); R_f^1 0.54, R_f^2 0.09, R_f^3 0.74.

Found: C, 48.31; H, 4.59; N, 18.56%. Calcd for $C_{24}H_{26}O_{10}N_8 \cdot 1/2\ H_2O$: C, 48.40; H, 4.57; N, 18.82%.

Measurements. CD spectra were recorded in methanol solution on JASCO J-40 spectropolarimeter.

Computation. ECEPP program and minimization subroutine VA04A were obtained from Quantum Chemistry Program Exchange, Indiana University. Computations were performed with FACOM 260F computer. All bond lengths and bond angles were taken from the standard ECEPP data for residues and end groups. The proline ring was taken in the puckered down position ($\phi = -75°$).

The authors are grateful to Dr. Tadakazu Maeda in our Institute and the members of System Center, Mitsubishi-Kasei Industries, Ltd. for their assistance during computation and also to the members of System Engineering Laboratory, Mitsubishi-Kasei Industries, Ltd., for performing elemental analyses.

References

1) Part IV: K. Sato, N. Taki, U. Nagai, and T. Higashijima, *Bull. Chem. Soc. Jpn.*, **56**, 2476 (1983).

2) The abbreviations used in this paper are those recommended by IUPAC-IUB; *J. Biol. Chem.*, **247**, 977 (1972). Additional abbreviations: Boc, *t*-butoxycarbonyl; Dnp, 2,4-dinitrophenyl; pNA, p-nitroanilide; ONSu, *N*-hydroxysuccinimide ester; EDC·HCl, 1-ethyl-3-(3′-dimethylaminopropyl)-carbodiimide hydrochloride; N_2ph-F, 1-fluoro-2,4-dinitrobenzene; TEA, triethylamine; DMF, *N,N*-dimethylformamide; TLC, thin-layer chromatography.

3) K. Sato, M. Kawai, and U. Nagai, *Biopolymers*, **20**, 1921 (1981).

4) F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, *J. Phys. Chem.*, **79**, 2361 (1975).

5) M. J. D. Powell, *Computer J.*, **7**, 151 (1964).

6) S. S. Zimmerman and H. A. Scheraga, *Biopolymers*, **16**, 811 (1977).

7) S. S. Zimmerman, M. S. Pottle, G. Nemethy, and H. A. Scheraga, *Macromolecules*, **10**, 1 (1977).

8) Conformational letter code⁷⁾ and ϕ and ψ values are given for each conformation.

9) R. Chandrasekaran, A. V. Lakshminarayanan, U. V. Pandya, and G. N. Ramachandran, *Biochim. Biophys. Acta*, **303**, 14, (1973).

10) K. Sato, T. Higashijima, T. Miyazawa, R. Sugawara, and U. Nagai, "Peptide Chemistry 1982; Proceedings of the 20th Symposium on Peptide Chemistry," ed by S. Sakakibara, Protein Research Foundation, Osaka (1983), p. 299.

11) M. Kawai, K. Sato, and U. Nagai, "Peptide Chemistry 1982; Proceedings of the 20th Symposium on Peptide Chemistry," ed by S. Sakakibara, Protein Research Foundation, Osaka (1983), p. 317.

12) K. Sato, U. Nagai, and T. Higashijima, *Bull. Chem. Soc. Jpn.*, **56**, 1657 (1983).

13) K. Sato, M. Kawai, and U. Nagai, *Bull. Chem. Soc. Jpn.*, **56**, 1527 (1983).