# THEORETICAL CONFORMATIONAL ANALYSIS OF THE METHYLAMIDE OF N-ACETYL-L-ALANYL-L-PROLYL-L-ALANINE. II

É. P. Gromov, V. Z. Pletnev, and E. M. Popov

UDC 547.962

The main difficulty in the theoretical conformational analysis of oligopeptides is due to the presence in many natural amino-acid residues of a large number of stable conformations of similar energies. If it is considered that on an average each residue of a protein chain possesses ten approximately equivalent forms, even a tripeptide fragment is an extremely complex subject for analysis from the aspects of a monopeptide. In view of this, it appeared of interest to compare the conformational states of the residues in a tripeptide (n − 1, n, n + 1) with the states in the two pairs of adjacent dipeptides (n − 1, n, and n + 1). In this case, when the most preferred conformations of the tripeptides are combinations of the most preferred forms of the corresponding dipeptides, the possibility appears of making an a priori prediction of a set of low-energy forms of the tripeptide on the basis of the known conformational states of the dipeptides. The fulfillment of this condition means that interactions between neighboring residues (n − 1 and n, n and 1 + 1) have a more fundamental influence on the conformation of the tripeptide than interactions between extreme members (n − 1 and n + 1), or the latter are not in contradiction with the former and can exist with definite conformations of suitable dipeptide forms. A complete conformational analysis of a dipeptide is a relatively simple task at the present time.

In a preceding communication [1] we considered the conformational possibilities of the dipeptide molecules Ac-L-Ala-L-Pro-NHMe and Ac-L-Pro-L-Ala-NHMe. The influence exerted by the preceding and following Ala residues separately on the conformational state of the Pro was shown. In the present work, independently of the results obtained for dipeptides, a conformational analysis is made of the tripeptide Ac-L-Ala-L-Pro-L-Ala-NHMe. In this molecule, an action on the state of the central Pro residue is exerted by Ala residues simultaneously from both sides. The main aim of the investigations performed consisted in comparing the conformational states of the residues in the tripeptide and in the two dipeptides.

The calculation of the Ac-L-Ala-L-Pro-L-Ala-NHMe molecule was performed with reference to the conditions of a polar medium without taking intramolecular hydrogen bonds into account. The analysis was performed in two variants — with the trans and with the cis configuration of the tertiary peptide group preceding the Pro. To evaluate the dependence of the energy on the size of the angle $\varphi_2$ of the proline, all the calculations were made with a variation $\varphi$ from −55 to 70° in steps of 5°.

A model of the molecule with the symbols for the variable dihedral angles of rotation is shown in Fig. 1. The method of calculation, the potential functions, and the parametrization have been described in the preceding paper [1].

In the extended conformations of the Ac-L-Ala-L-Pro-L-Ala-NHMe molecule with $\omega_1 \sim 180°$ and $\sim 0°$ (in the trans and cis configurations of the tertiary peptide groups), the first Ala residue may be present in three forms — B, L, and R (B', L', and R' in the case of $\omega_1 \sim 0°$), the Pro residue in the B and R forms, and the third, Ala, residue in the R, B, and L forms [1, 2]. Thus, in a polar medium where the formation of intramolecular hydrogen bonds is

M. M. Shemyakin Institute of Bioorganic Chemistry, Academy of Sciences of the USSR.
Translated from Khimiya Prirodnykh Soedinenii, No. 5, pp. 626-630, September-October, 1975. Original article submitted June 14, 1974.

TABLE 1. Energies of the Optimum Extended Conformations of the Ac-L-Ala-L-Pro-L-Ala-NHMe Molecule with $\omega_1 \sim 180°$ (trans-)

| Type of conformation | Energy, kcal/mole | Forms of dipeptides composing conformation of the tripeptides* | Total energy of dipeptide forms, kcal/mole |
|---|---|---|---|
| B–R–R | 0 | 2–1 | 0,1 |
| L–R–R | 0,2 | 1–1 | 0 |
| B–R–B | 1,5 | 2–3 | 1,2 |
| R–R–R | 1,7 | 5–1 | 2,3 |
| B–R–L | 1,8 | 2–2 | 1,0 |
| L–R–L | 1,8 | 1–2 | 0,9 |
| L–B–L | 1,8 | 3–4 | 1,8 |
| L–R–B | 1,8 | 1–3 | 1,1 |
| B–B–L | 1,9 | 4–4 | 2,1 |
| L–B–R | 2,5 | 3–5 | 2,2 |
| B–B–R | 2,7 | 4–5 | 2,5 |
| R–R–L | 2,8 | 5–2 | 3,2 |
| R–R–B | 2,9 | 5–3 | 3,4 |
| L–B–B | 3,0 | 3–6 | 2,7 |
| B–B–B | 3,3 | 4–6 | 3,0 |
| R–B–R | 4,7 | 6–5 | 6,0 |
| R–B–L | 4,9 | 6–4 | 5,6 |
| R–B–B | 5,2 | 6–6 | 6,5 |

*The order of numbering the forms of the dipeptides corresponds to an increase in energy [1].

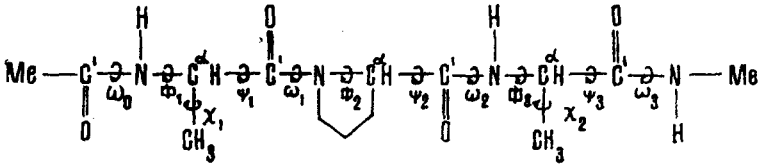

Fig. 1. Calculation model of the molecule of the methylamide of N-acetyl-L-alanyl-L-propyl-L-alanine.

unlikely,* the conformational equilibrium of the tripeptide can be represented in the general case by 18 extended forms. Table 1 gives the values of the energies of all the conformations with a transoid peptide group preceding the Pro ($\omega_1 \sim 180°$). Table 2 gives the values of the dihedral angles of rotation of the most favorable conformations, the energy of which is within the range from 0 to 2 kcal/mole. The analogous figures for the conformations of molecules with the cis configuration of the tertiary peptide group ($\omega_1 \sim 0°$) the energy of which does not exceed 2 kcal/mole are given in Table 3. The optimum forms were obtained by minimizing the energy with the variation of eleven variables: $\omega_0$, $\varphi_1$, $\psi_1$, $\omega_1$, $\chi_1$, $\psi_2$, $\omega_2$, $\varphi_3$, $\psi_3$, $\omega_3$, $\chi_3$. The calculation was performed with the following values of the angle $\varphi_2$: –55, –60, –65, and –70°. Tables 2 and 3 give the values of $\varphi_2$ corresponding to the conformations of lowest energy.

Figure 2 gives an idea of the sensitivity of the energy to the angle $\varphi_2$:

The energies of the conformations with $\omega_1 \sim 180°$ fall into the range from 0 to 5.2 kcal/mole, and of those with $\omega_1 \sim 0°$ into the range from 1.1 to 9.6 kcal/mole. The number of low-energy conformations (0-2 kcal/mole) in the first case is 9, and in the second case 4. The energy of the global form with a transoid peptide group is less than the energy of the form with a cisoid group by 1.1 kcal/mole. The observed enthalpy and entropy preference of the tripeptide conformations with $\omega_1 \approx 180°$ as compared with $\omega_1 \approx 0°$ is more pronounced than

*The forms with hydrogen bonds within the tripeptide fragment (3–1, 4–1, and 5–1) are of the greatest interest in the conformational analysis of cyclic oligopeptides.

TABLE 2. Geometric Parameters of the Low-Energy Conformations of the Ac-L-Ala-L-Pro-L-Ala-NHMe Molecule with $\omega_1 \sim 180°$ (trans-)

| Angles of rotation, deg | Type of conformation | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | B-R-R | L-R-R | B-R-B | R-R-R | B-R-L | L-R-L | L-B-L | L-R-B | B-B-L |
| $\omega_0$ | 180 | 180 | 173 | 178 | 180 | 180 | 180 | 180 | 180 |
| $\varphi_1$ | −66 | 53 | −66 | −59 | −67 | 53 | 53 | 53 | −66 |
| $\psi_1$ | 151 | 69 | 153 | −53 | 152 | 66 | 67 | 67 | 152 |
| $\omega_1$ | 178 | 178 | 174 | 179 | 176 | 180 | −178 | 178 | 176 |
| $\chi_1$ | 56 | 60 | 59 | 22 | 60 | 60 | 60 | 60 | 60 |
| $\varphi_2$ | −60 | −60 | −55 | −70 | −60 | −65 | −70 | −55 | −70 |
| $\psi_2$ | −49 | −43 | −40 | −33 | −41 | −45 | 108 | −43 | 122 |
| $\omega_2$ | 176 | 174 | 174 | 179 | 179 | 180 | −178 | 174 | −177 |
| $\varphi_3$ | −63 | −66 | −68 | −56 | 54 | 54 | 54 | −64 | 52 |
| $\psi_3$ | −38 | −41 | 115 | −32 | 36 | 36 | 37 | 110 | 36 |
| $\omega_3$ | 180 | 180 | 180 | 180 | 180 | 180 | −179 | 180 | −178 |
| $\chi_3$ | 60 | 60 | 59 | 58 | 57 | 57 | 57 | 59 | 57 |
| $E_{tot}$, kcal/mole | 0 | 0.2 | 1.5 | 1.7 | 1.8 | 1.8 | 1.8 | 1.8 | 1.9 |

TABLE 3. Geometric Parameters of the Low-Energy Conformations of the Ac-L-Ala-L-Pro-L-Ala-NHMe Molecule with $\omega_1 \sim 0°$ (cis-)

| Angles of rotation, deg | Type of conformation | | | |
|---|---|---|---|---|
| | B'-R-B | B'-R-R | B'-B-R | B'-B-B |
| $\omega_0$ | 178 | −179 | −176 | −177 |
| $\varphi_1$ | −163 | −165 | −170 | −170 |
| $\psi_1$ | 90 | 82 | 93 | 93 |
| $\omega_1$ | 2 | − 13 | 3 | 3 |
| $\chi_1$ | 58 | 58 | 58 | 58 |
| $\varphi_2$ | − 65 | − 70 | − 65 | − 65 |
| $\psi_2$ | − 43 | − 41 | 138 | 139 |
| $\omega_2$ | 179 | 180 | 175 | 175 |
| $\varphi_3$ | −133 | − 92 | − 66 | − 67 |
| $\psi_3$ | 164 | − 44 | − 64 | 130 |
| $\omega_3$ | 180 | 180 | 180 | 180 |
| $\chi_3$ | 60 | 60 | 60 | 60 |
| $E_{tot}$, kcal/mo | 1.1 | 1.2 | 1.4 | 1.9 |

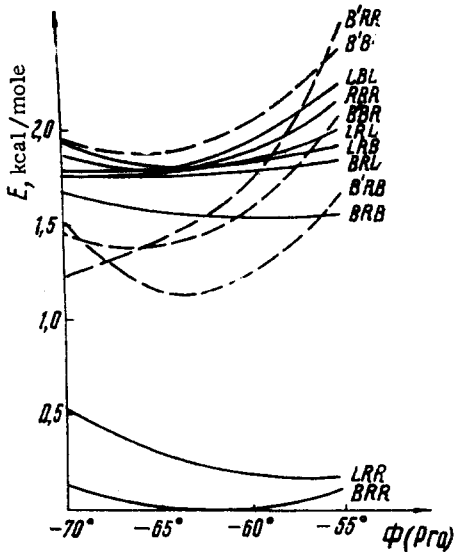


Fig. 2. Dependence of the energy of the conformations of Ac-L-Ala-L-Pro-L-Ala-NHMe on the angle φ of the proline residue.

for the corresponding dipeptide conformations [1]. Thus, lengthening the peptide chain promotes the stabilization of the trans configuration of the tertiary amide group.

The fourth column of Table 1 shows the conformations of the dipeptide molecules Ac-L-Ala-L-Pro-NHMe and Ac-L-Pro-L-Ala-NHMe the superposition of which forms the conformation of the tripeptide Ac-L-Ala-L-Pro-L-Ala-NHMe. Each conformation of a tripeptide is denoted by a number corresponding to the position which is occupied in the set of stable forms arranged in order of increasing energy (see Tables 1 and 2 in [1]). As can be seen from the table, the most favorable conformations of the tripeptide, located in its top part, represent combinations of the preferred forms of the dipeptides. The very first positions in the table, with practically identical values of the energy are occupied by the structures B–R–R and L–R–R; the conformational states of the dipeptide fragments in these structures possess forms also possessing the lowest energies.

TABLE 4. Conformational Distribution in Proteins of the Tripeptide Fragments X-Pro-Y

| Type of conformation | Distribution in proteins |
|---|---|
| B—R—R | 6 |
| B—R—B | 2 |
| R—R—R | 2 |
| B—B—L | 1 |
| B—B—R | 5 |
| B—B—B | 6 |

The last column of Table 1 is the sums of the relative energies of the dipeptide conformations ($E_I + E_{II}$) which, in the absence of an interaction between the terminal Ala residues, must be extremely close to the energy of the Ac-L-Ala-L-Pro-L-Ala-NHMe molecule ($E_{III}$). The changes in $E_{III}$ are basically symbatic to the changes in $E_I + E_{II}$, which shows the dominating role of the interactions of adjacent residues as compared with the interactions of residues not connected directly with one another. Nevertheless, the production of tripeptide conformations from the corresponding dipeptide forms cannot be considered as completely additive. Thus, the relative energy of the conformation R—R—R (1.7 kcal/mole) is considerably less than the relative sum of the energies of the R—R forms of the two dipeptides (2.3 kcal/mole). This is due to the stabilizing dispersion interactions of the terminal peptide groups in the α-helical form of the chain. Conversely, the energy of the B—R—L, L—R—L, and L—R—B conformations (1.8 kcal/mole) is somewhat greater than the sums $E_I + E_{II}$ (1.0, 0.9, and 1.1 kcal/mole, respectively), which is due to destabilizing nonvalent interactions of the terminal Ala residues with the R conformation of the central Pro residue. However, these deviations from additivity are not very great, and the values of $E_{III}$ are fairly close to $E_I + E_{II}$. In agreement with this, the differences between the geometric parameters of the residues in monotypical conformations of the dipeptides and of the tripeptide are also small (compare Tables 2 and 3 with Tables 1 and 2 in [1]).

Thus, the possibility of the formation of the most preferred conformations of Ac-L-Ala-L-Pro-L-Ala-NHMe from suitable conformations of the dipeptides Ac-L-Ala-L-Pro-NHMe and Ac-L-Pro-L-Ala-NHMe is determined by the inconsiderable nature of the interactions between the terminal residues. This conclusion, of course, is not a general one for any amino-acid sequences. For the tripeptide considered it proved to be valid because of the absence of fairly voluminous side chains form the residues. The only possibility in this case of an additional stabilizing interaction of the terminal group in the R—R—R conformation is, as already mentioned, realized within the framework of combinations of low-energy forms of the dipeptide.

The Ac-L-Ala-L-Pro-L-Ala-NHMe molecule models the maximum conformational possibilities of the main chains of the protein tripeptide fragment X-Pro-Y, where X and Y are any amino-acid residues with the exception of Gly. In view of this, it is of interest to compare the calculated optimum forms of the tripeptide with the conformational states of the corresponding tripeptide fragments of proteins of known structure. We have used information on the tertiary structures of six proteins (myoglobin, lysozyme, α-chymotrypsin, carboxypeptidase A, cytochrome C, and insulin), in which there is a total of 22 fragments of the X-Pro-Y type. The forms of the main chains found in them are given in Table 4. A comparison of Tables 1 and 4 shows that the number of observed types of conformations is considerably less than the set of optimum forms of the free tripeptide. Thus, conformations in which the residues adjacent to proline assume the L form (L—R—R, B—R—L, etc.) are practically absent, although they also belong to the low-energy conformations of the tripeptide. This is due to the unsuitability of form L in residues having a $C^\beta$ atom when they are present not at the end of the chain but are surrounded by other residues. The low probability of the realization of conformations of the tripeptide fragment with the L form of the central residue follows from the calculation performed in our laboratory of the molecule Ac-L-Ala-L-Ala-L-Ala-NHMe.

Of the six types of conformations of X-Pro-Y observed in proteins, the first four (see Table 4) belong to the most preferred forms of the Ac-L-Ala-L-Pro-L-Ala-NHMe molecule, and their energies are in the range from 0 to 2 kcal/mole (see Tables 1 and 2). The other two types, B—B—R and B—B-B, correspond to considerably less favorable forms (2.7 and 3.3 kcal/mole, respectively). The facts that such conformations of the main chains are nevertheless encountered fairly frequently in X-Pro-Y protein fragments is due to the stabilizing influence of the system of hydrogen bonds of β-structural sections.

## SUMMARY

1. Spatial forms of Ac-L-Ala-L-Pro-L-Ala-NHMe with trans peptide bonds are the most preferred. Elongation of the peptide chain promotes the stabilization of the trans configuration of the tertiary amide group.

2. In the structure of the compound investigated, the dominating role is played by the interaction of the neighboring residues.

3. Of the six types of conformations of X-Pro-Y observed in proteins, the first four belong to the preferred forms of the Ac-L-Ala-L-Pro-L-Ala-NHMe molecule.

## LITERATURE CITED


1. V. Z. Pletnev, É. P. Gromov, and D. M. Popov, Khim. Prirodn. Soedin., No. 5, p. 618 (1975).
2. E. M. Popov, G. M. Lipkind, S. F. Arkhipova, and V. G. Dashevskii, Molekul. Biol., 3, 622 (1968).