residues, has not shown the presence of lower molecular weight species which could be attributed to peptide bond rupture at the tryptophanyl residues.[48]

In conclusion, the studies with thioglycolate described herein have shown the reduced human serum albumin to be a homogeneous protein of the same molecular weight as the native molecule. The disulfide bonds are therefore all present as intrachain crosslinks. The reason for the difference between the oxidized and reduced preparations is as yet unclear.

Acknowledgments.—The authors would like to thank Dr. J. L. Oncley for his invaluable help, encouragement and criticisms. They are also deeply indebted to Dr. Cyril M. Kay for performing the light scattering measurements.

BOSTON, MASS.

(48) M. E. Reichmann and R. Colvin, *Can. J. Chem.*, **34**, 411 (1956).

[CONTRIBUTION FROM THE CHEMICAL LABORATORY, FACULTY OF SCIENCE, TOKYO UNIVERSITY]

Near Infrared Spectra of Compounds with Two Peptide Bonds and the Configuration of a Polypeptide Chain. VII. On the Extended Forms of Polypeptide Chains

BY MASAMICHI TSUBOI, TAKEHIKO SHIMANOUCHI AND SAN-ICHIRO MIZUSHIMA

RECEIVED JULY 22, 1958

Infrared absorptions of the five acetylamino acid N-methylamides (where amino acids are glycine, alanine, valine, norleucine and proline) in the 3 μ region have been observed with a grating spectrometer. The results are interpreted in terms of the molecular configurations of these compounds (the folded and the extended forms). Our previous conclusion for the folded form in these solutions has been confirmed. As to the extended form more accurate measurement has shown that this form of acetylglycine N-methylamide is different from that of acetylalanine (or valine or norleucine) N-methylamide. Taking this difference into account, the molecular configurations of polyglycine and other polypeptides consisting of L-amino acid residues have been discussed.

Introduction

The near infrared spectra of acetylamino acid N-methylamides, $CH_3CONHCHRCONHCH_3$, have been measured in carbon tetrachloride and chloroform solutions.[1-8] Based on these results, the molecular structure of these substances has been discussed in relation to the configuration of a polypeptide chain.

Recently spectrometers with higher resolving power and higher accuracy become available and many of the spectra have now been re-examined. A number of new facts have been observed, most of which have confirmed our former conclusions, and none of which have led to any essential alteration of them. However, some of the new results have enabled us to discuss in more detail the molecular structure of acetylamino acid N-methylamides and hence also of polypeptides.

In this paper, the results of the re-examinations of the near infrared spectra of five acetylamino acid N-methylamides, where, the amino acids are glycine, DL-alanine, L-valine, DL-norleucine, and L-proline in carbon tetrachloride will be given. The solutions are so dilute that intermolecular hydrogen bonding is negligible.

Re-examinations of the spectra of more concentrated solutions will be given later.

Experimental

The samples used are acetylglycine N-methylamide, acetyl-DL-alanine N-methylamide, acetyl-L-valine N-methylamide, acetyl-DL-norleucine N-methylamide and acetyl-L-proline N-methylamide, abbreviated as AGlyMA, A-DL-AlaMA, A-L-ValMA, A-DL-NorlMA, and A-L-ProMA, respectively. The samples are the same as those used in our previous researches.[1,3,5]

The infrared absorption measurements were carried out by a Perkin-Elmer 112G spectrometer with a 75 lines/mm. grating and a KBr fore-prism. The grating is blazed at 12 μ, and its fourth order diffraction was used in the 3 μ region measurements. The wave length calibration for the spectrometer was made with known frequencies of 296 absorption lines reported in a previous paper.[9] The frequency given for any sharp line observed in the present experiment should be accurate up to the figure on the place of 1 cm.$^{-1}$. The effective slit width was made 1.7–2.9 cm.$^{-1}$ (0.2–0.4 mm.) and the resolving power of the spectrometer was more than sufficient for the present purpose.

Dilute carbon tetrachloride solutions of the compounds were placed in a cell of 10 cm. path length, made of fused silica, with plane parallel silica windows on both ends.[10] The cell was heated electrically.

For each sample, measurements were made at 30 and 60°, and at two different concentrations, c_1 and c_2, where c_2 is just half of c_1. The optical path length (l_2) for c_2 was made twice as long as that (l_1) for c_1, so that $c_1 \times l_1 = c_2 \times l_2$ (actually, l_1 = 10 cm. and l_2 = 20 cm.). This was achieved by placing between the source unit and the fore-prism unit of the spectrometer an attachment which consists of two spherical mirrors and two plane mirrors, arranged as shown in Fig. 1. The infrared beam goes forward and backward through the 10 cm. silica cell, and hits the entrance slit with the same direction and with the same aperture as it does without the attachment.

Results and Discussions

The results of the measurements are shown graphically in Fig. 2. For each of the five com-

(1) S. Mizushima, T. Shimanouchi, M. Tsuboi, T. Sugita, E. Kato and E. Kondo, THIS JOURNAL, **73**, 1330 (1950).

(2) S. Mizushima, T. Shimanouchi, M. Tsuboi and R. Souda, *ibid.*, **74**, 270 (1952).

(3) S. Mizushima, T. Shimanouchi, M. Tsuboi, T. Sugita, K. Kurosaki, N. Mataga and R. Souda, *ibid.*, **74**, 4639 (1952).

(4) S. Mizushima, T. Shimanouchi, M. Tsuboi, K. Kuratani, T. Sugita, N. Mataga and R. Souda, *ibid.*, **75**, 1863 (1953).

(5) S. Mizushima, M. Tsuboi, T. Shimanouchi, T. Sugita and T. Yoshimoto, *ibid.*, **76**, 2479 (1954).

(6) S. Mizushima, M. Tsuboi, T. Shimanouchi and M. Asai, *ibid.*, **76**, 6003 (1954).

(7) S. Mizushima, T. Shimanouchi, M. Tsuboi and T. Arakawa, *ibid.*, **79**, 5357 (1957).

(8) M. Tsuboi, *J. Chem. Soc. Japan*, **76**, 380 (1955).

(9) S. Mizushima, *et al.*, "Report on Perkin-Elmer Grating Spectrophotometer Model 112G," Nankodo Co. Ltd., Tokyo, 1957.

(10) M. Tsuboi, *Bull. Chem. Soc. Japan*, **24**, 75 (1951).

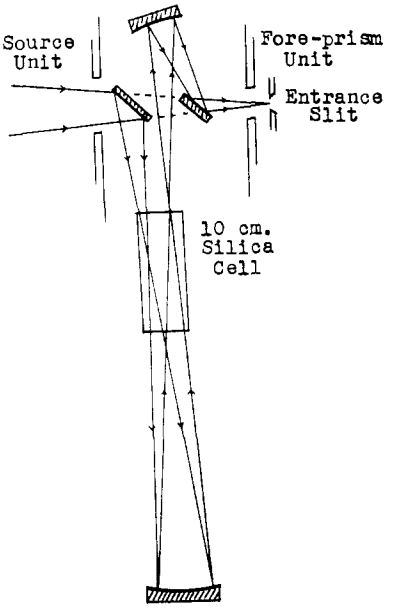

Fig. 1.

pounds, four curves are given corresponding to the four measurements described above. The ordinate (per cent. transmission) of any of the lower curves is shifted by 20% from that of the next upper one. Of the experimental results obtained here the following two points are worthy of note.

(I) **Hydrogen-bonded NH Bands.**—In the first place, our previous conclusion of the intramolecular hydrogen bonding for the NH bands have been confirmed. Even at high dilution the bonded NH band in the 3360–3330 cm.$^{-1}$ region never disappears, and at 60° and any concentration less than 0.0003 mole/l., the ratio of the intensity of the bonded NH band to that of the free NH band in the 3460–3420 cm.$^{-1}$ region remains constant as can be seen in Fig. 2 (see the first two curves for each compound). This would not be the case, if the bonded band were assigned to the NH vibration of the intermolecular hydrogen bond.

However, the intensity ratio considerably depends on the temperature[1,3,5,7] and solvent.[5,7] This was considered in our previous papers[1–7] to be due to the coexistence of two molecular forms, one with an intramolecular hydrogen bond and the other without this bond. The former is the folded form shown in Fig. 3B and the latter the extended form shown in Fig. 3A. The abundance ratio of the two forms changes as the temperature or the solvent is changed.

As can be seen in Fig. 2, the relative intensity of the NH bands in the 3360–3330 cm.$^{-1}$ region is nearly the same for AGlyMA, A-DL-AlaMA, A-L-ValMA and A-DL-NorlMA. Therefore, the abundance ratio of the extended and the folded forms should be almost the same for these compounds. In A-L-ProMA, however, this is not the case and the folded form is predominant. This is as expected, since in AProMA the folded form would be formed much more easily than in the other four substances.

(II) **Free NH Bands.**—In the second place the spectra in the free NH region (3470–3420 cm.$^{-1}$)

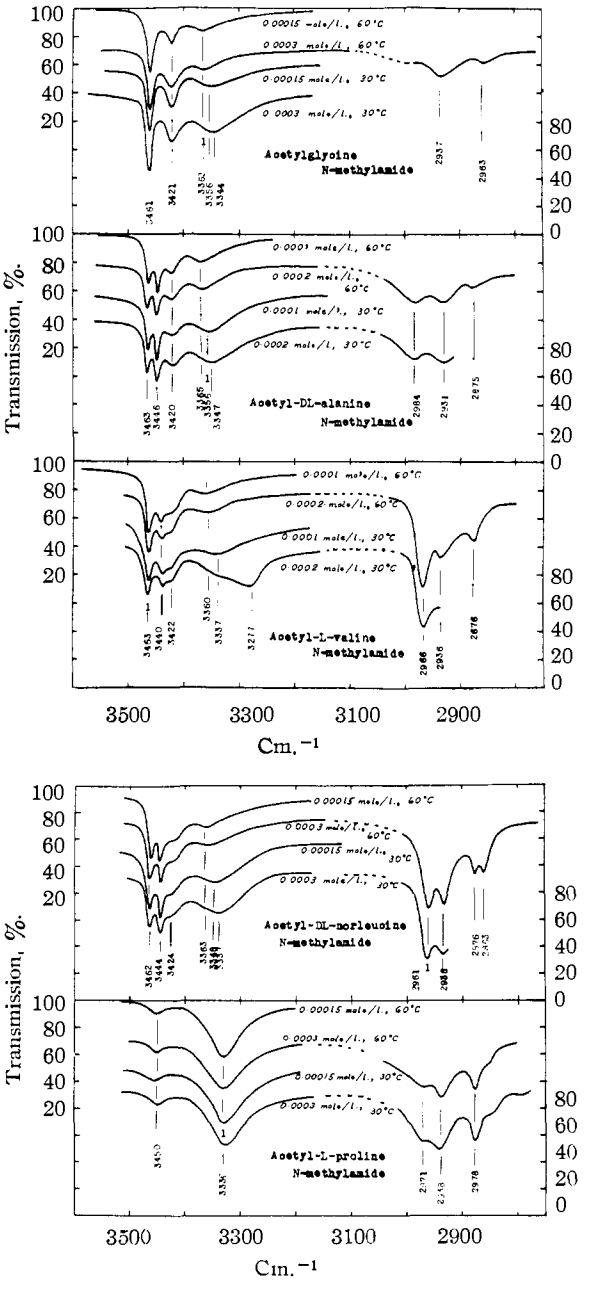

Fig. 2.—Infrared absorptions of acetylamino acid N-methylamides in dilute carbon tetrachloride solutions.

have been observed with a resolution much higher than that of our earlier measurements.

Previously only one peak was observed for AAlaMA, AValMA and ANorlMA in this region. However, in the present experiments three bands have been observed at 3460, 3440 and 3420 cm.$^{-1}$ for all of these three compounds. For AGlyMA only two bands, at 3461 and 3421 cm.$^{-1}$, have been observed. The band at 3461 cm.$^{-1}$ of this substance is much stronger than that of the above three compounds. AGlyMA shows no band at 3440 cm.$^{-1}$, where the above three give another NH band. AProMA has shown a weak NH band at 3450 cm.$^{-1}$, which was not observed previously in the carbon tetrachloride solution.

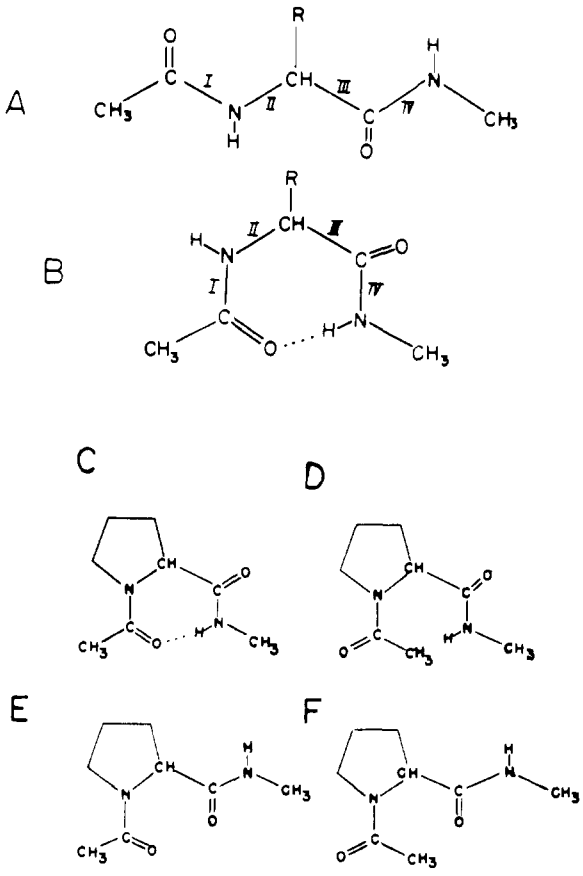

Fig. 3.—Extended (A) and folded (B) forms of acetylamino acid N-methylamide and various forms of acetylprolone N-methylamide (C-F).

Of the three NH bands observed for AAlaMA, AValMA and ANorlMA, two bands at 3460 and 3440 cm.$^{-1}$ are considered to arise from the two free NH vibrations of the extended form (Fig. 3A). The third one at 3420 cm.$^{-1}$ is assigned to the free NH vibration of the folded form (Fig. 3B). This NH bond will be weaker than that of the completely free NH of the extended form because of the electron migration towards the adjacent carbonyl group, which is involved in the intramolecular hydrogen bonding. This seems to be the reason why the free NH frequency of the folded form is lower than that of the extended form. The lowering of the NH frequency is also found in AGlyMA, and this can be assigned to the "free" NH of the folded form of this compound. The extended form of AGlyMA, however, seems to give only one free NH band. A possible explanation will be given in the next section.

For AProMA, no band is found at 3420 cm.$^{-1}$. This is compatible with the above assignment because this molecule has no "free" NH group in its folded form (see Fig. 3C). AProMA cannot take the extended form since the internal rotation about the N–C axis is frozen because it forms a part of the five-membered ring. Therefore, the weak free NH band observed at 3450 cm.$^{-1}$ must be assigned to a form which is neither extended nor folded. This will be one of the forms shown in Fig. 3D, E and F. The existence of such a form was considered in our earlier investigation on the chloroform solution.[5] It has now been shown that even in the carbon tetrachloride solution a small fraction of the molecules take such a form.

It may be mentioned here that we observed *single* free NH band for acetylsarcosine N-methylamide, $CH_3CON(CH_3)CH_2CONHCH_3$, and for acetyl-N-methyl-DL-norleucine N-methylamide, $CH_3CON(CH_3)CH(CH_2)_3CH_3CONHCH_3$ (at 3452 cm.$^{-1}$ and 3454 cm.$^{-1}$, respectively) with the same spectroscopic resolution to that with which we observed *two* and *three* free NH bands for AGlyMA and ANorlMA, respectively. This fact furnishes a support to the above assignments of the free NH bands in question.

Possible Difference in the Configuration of the Extended Form of Acetylglycine N-Methylamide from Those of Other Three Acetylamino Acid N-Methylamides.—It has been found that each of AAlaMA, AValMA and ANorlMA shows two NH bands at about 3460 and 3440 cm.$^{-1}$ with almost equal intensity whereas AGlyMA shows only one band at 3461 cm.$^{-1}$ which is about twice as strong as any of the other three. One of the explanations for this is that the two NH groups of the extended form of AAlaMA, AValMA and ANorlMA show absorptions at two frequencies different from each other, whereas the two NH groups of the extended form of AGlyMA give absorption at the same frequency.

This will be the case, if the extended forms of the former three substances have the structure shown in Fig. 4A, where the N–H bond in the first peptide linkage takes almost the same direction with the C=O bond in the second peptide linkage, and the extended form of AGlyMA has the structure shown in Fig. 4B, in which the directions of the two bonds are quite different from each other. In the former structure (Fig. 4A), the N–H bond in the first peptide linkage may form a weak hydrogen bond with the C=O group in the second peptide linkage and would exhibit a stretching frequency lower than that of the N–H bonds in the second peptide linkage free from such a hydrogen bonding.[11] In the latter structure (Fig. 4B), the two N–H bonds will show one and the same stretching frequency.

The stability of the structures A and B in Fig. 4 can be understood easily from a consideration of the stable positions in the internal rotation about the N–C and C–C single bonds as axes.[12] In the extended form of a polypeptide chain the azimuthal angles of internal rotation about these axes are all about 180.°[13] Actually, however, the extended forms of many polypeptide chains are not the fully extended form (see the next section). This is

(11) The weak hydrogen bond of this type (with a five-membered ring) will give rise to a frequency decrease of about 20 cm.$^{-1}$ as is actually observed. One of the two O–H groups of glycol or the O–H group in $C_2H_5O(CH_2)_2OH$ which forms a hydrogen bond with another O atom in a way similar to that of the N–H group under investigation shows a frequency about 30 cm.$^{-1}$ lower than the completely free O–H group. (L. P. Kuhn, THIS JOURNAL, **74**, 2492 (1952)).

(12) S. Mizushima, "Structure of Molecules and Internal Rotation," Academic Press, Inc., New York, N. Y., 1954.

(13) T. Shimanouchi and S. Mizushima, *Bull. Chem. Soc. Japan*, **21**, 1 (1948).

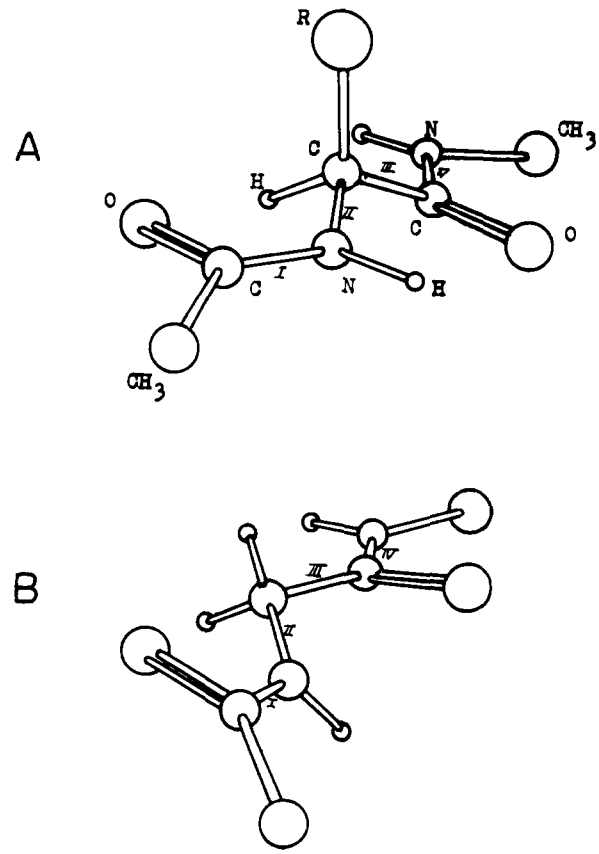

Fig. 4.—Extended forms of acetyl-L-alanine (or valine, or norleucine) N-methylamide (A) and acetylglycine N-methylamide (B).

probably due to the following fact. If the azimuthal angles were exactly 180°, the N–H bond and the ^{III}C—C bond should come to the *cis* position with respect to each other in the internal rotation about the ^{II}N—C axis (see Fig. 5A) and ^{II}C—N bond and C=O bond should come to the *cis* position in the internal rotation about the ^{III}C—C axis (see Fig. 5D). The *cis* positions are not usually stable due to the steric repulsions, although in the present case the repulsive force is not very large. Thus, a small deviation may take place from the position of exactly 180°.

In the extended forms of AAlaMA, AValMA and ANorlMA, the deviations may take place only in definite directions. About the ^{II}N—C axis, the ^{I}C–N bond may approach the C–H bond but not the C–R bond because of the steric repulsion between the C and R of both sides of the ^{II}N—C axis (Fig. 5B). Similarly, about the ^{III}C—C axis, the ^{IV}C–N bond may approach the C–H bond but not the C–R bond (Fig. 5E). These two possible deviations result in the structure shown in Fig. 4A. On the other hand, the extended form of AGlyMA, the deviations may take place in different ways. About the ^{II}N—C axis, the ^{I}C–N bond may approach any one of the two C–H bonds, and around the ^{III}C—C axis, the ^{IV}C—N bond may also approach any one of the two C–H bond (Fig. 5B′ and E′). Therefore, it is quite possible that the ^{IV}C—N bond approaches the other C–H bond than the one to which the ^{I}C–N bond approaches. If this is the case we have the structure shown in Fig. 4B and the result of the present experiment is satisfactorily explicable.

Molecular Configurations of Polyglycine-I and Other Polypeptides in β-Form.—Based on the experimental result obtained above, new models for the configurations of the polypeptide chains in polyglycine-I and β-poly-L-alanine will be presented.

Corey and Pauling[14] gave the dimensions of the fully extended polypeptide chain derived from crystal structure analysis of many simple substances and from other experimental data. According to them, the identity period along the chain is 7.23 Å. Actually, however, the fiber identity period determined by the X-ray investigation[14a] is 6.96–7.05 Å. for polyglycine-I and about 6.9 Å. for some other polypeptides of L-aminoacids in the β-form. One of the probable interpretations for this difference is that the chains of polyglycine-I and other β-polypeptides are not fully extended, *e.g.*, the azimuthal angles of internal rotation about the ^{II}N—C axis and about the ^{III}C—C axis are less than 180°. In 1951 Pauling and Corey[15] presented a couple of models for β-polypeptides, which they called "new pleated sheet structures," in which the polypeptide chain is not fully extended. In their models, C–H is *trans* to N–H about the ^{II}N—C axis and C–H *trans* to C=O about the ^{III}C—C axis, in other words, the azimuthal angle between ^{I}C–N and ^{III}C—C about the N–C axis is about 120° and that between ^{II}N—C and ^{IV}C—N about the ^{III}C—C is about 120° (see Fig. 5C and F).

In 1953, they[16] revised their models, so as to make all the inter-chain N–H. . .O hydrogen bonds linear. This revision involved departure from the above mentioned orientations about the single bonds. In their revised "antiparallel-chain pleated sheet," the azimuthal angles are both about 150°; and in their revised "parallel-chain pleated sheet," these are a little less than 120°.

In our models, the stable azimuthal angles are both larger than 120° but smaller than 180°, and are probably about 140°. In this respect our models are not essentially different from the "antiparallel-chain pleated sheet" of Pauling and Corey.

In our models polyglycine-I has a configuration quite different from that of β-poly-L-alanine and other β-L-polypeptides. As may be seen from Figs. 4 and 6, the difference is similar to that which we

(14) R. B. Corey and L. Pauling, *Proc. Roy. Soc.* (*London*), **B141**, 10 (1953).

(14a) C. H. Bamford, L. Brown, A. Elliott, W. E. Hanby and I. F. Trotter, *Nature* **171**, 1149 (1953).

(15) L. Pauling and R. B. Corey, *Proc. Nat. Acad. Sci.*, **37**, 729 (1951).

(16) L. Pauling and R. B. Corey, *ibid.*, **39**, 253 (1953).

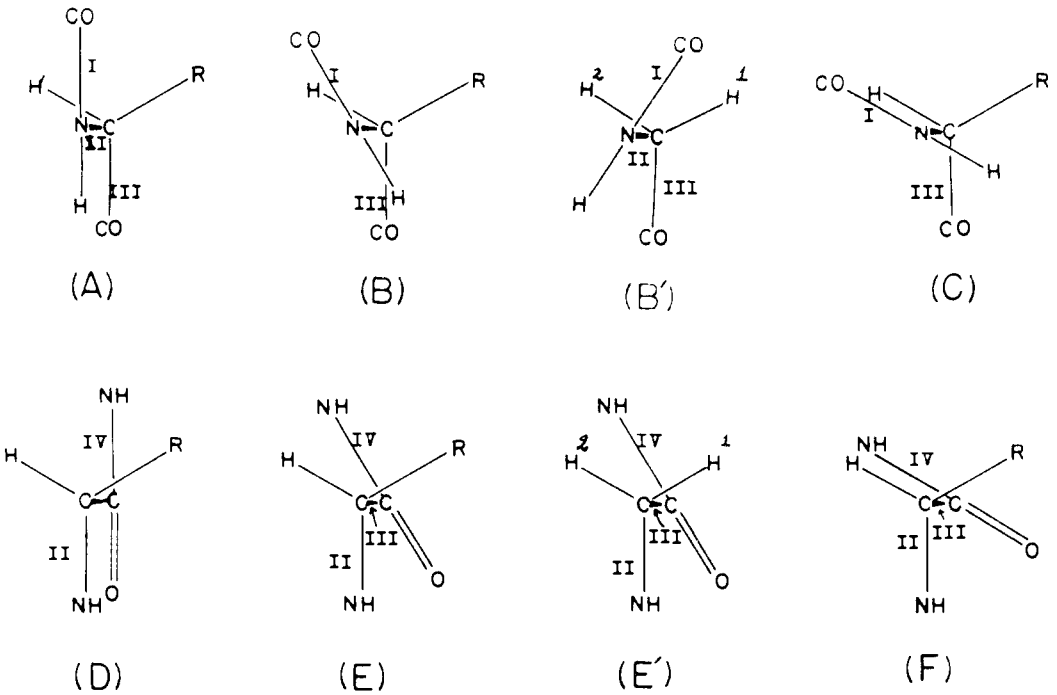

Fig. 5.—Various orientations of the groups on both ends of the N^{II}-C and C^{III}-C single bonds in the polypeptide chain.

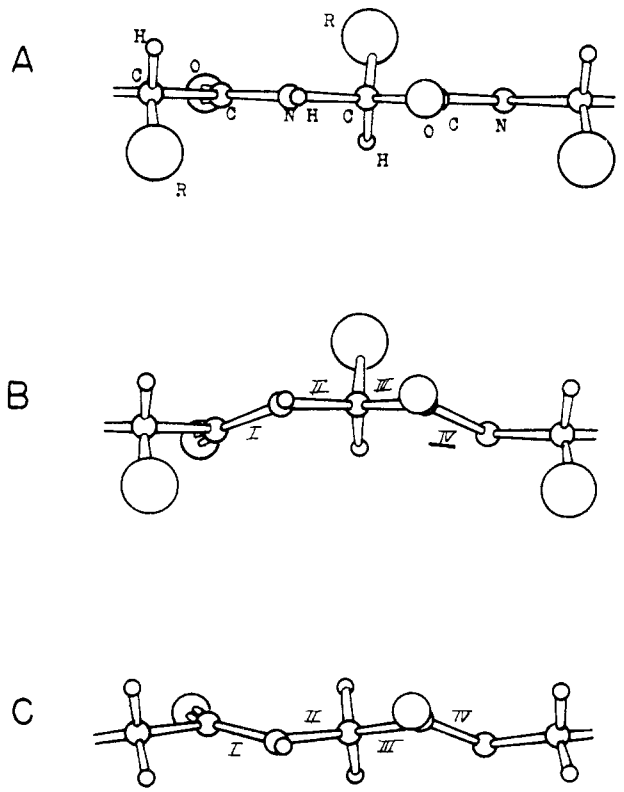

Fig. 6.—Drawings of the structures of the polypeptide chains. (A) Fully extended polypeptide chain viewed along the direction, which is perpendicular to the chain axis and in the zig-zag plane (back-bone view of the fully extended chain). (B) Our model of poly-L-alanine, or other polypeptides with L-amino acid residues, viewed along the same direction as (A). (C) Our model of polyglycine, viewed along the same direction as (A).

assumed between AGlyMA and the other three acetylamino acid N-methylamides. Namely, in polyglycine-I, the C^{I}–N bond takes a direction whose azimuthal angle about the N^{II}–C axis is about 140° from the C^{III}—C bond, about 20° from one of the two C–H bonds (let us call this C–H^1 bond) and about 80° from the other C–H bond (C–H^2 bond); and the C^{IV}—N bond takes a direction whose azimuthal angle about the C^{III}—C axis is about 140° from the N^{II}—C bond, about 20° from the C–H^2 bond and about 80° from the C–H^1 bond (see Fig. 5B′ and E′). On the other hand for β-poly-L-alanine and other β-L-polypeptides, the C^{I}–N bond takes a direction whose azimuthal angle about the N^{II}—C axis is about 140° from the C^{III}—C bond, about 20° from the C–H bond and about 80° from the C–R bond; and the C^{IV}—N bond takes a direction whose azimuthal angle about the C^{III}—C axis is about 140° from the N^{II}—C bond, about 20° from the C–H bond and about 80° from the C–R bond (see Fig. 5B and E). In short, in polyglycine, the C^{I}—N bond forms a smaller azimuthal angle with the C–H bond with which the C^{IV}—N bond forms a greater azimuthal angle, whereas in other polypeptides, the C^{I}–N bond forms a smaller azimuthal angle with the C–H bond with which the C^{IV}—N bond also forms a smaller azimuthal angle.

The new models have the following advantages:

(1) The models are favored by energy consid-

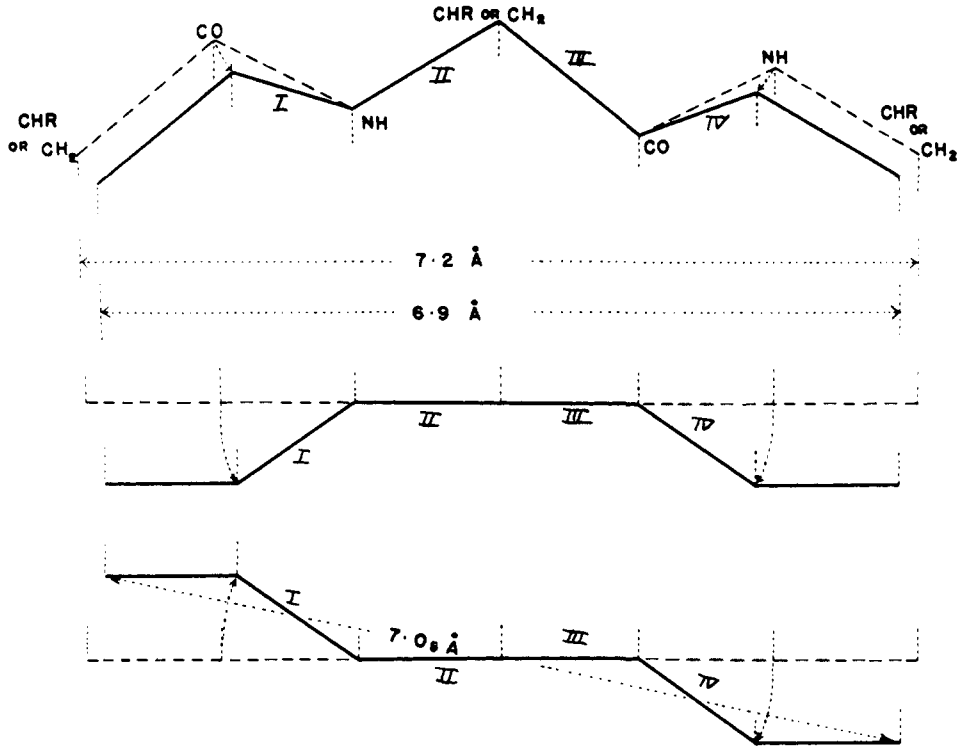

Fig. 7.—The structures of the polypeptide chains, viewed along the two directions perpendicular to the chain axis, *i.e.*, side view (upper one) and back-bone view (lower two). The broken lines show the fully extended chain and the block lines show our models. One of the two back-bone views shows our model for poly-L-alanine (or other polypeptides with L-amino acid residues), and the other shows our model for polyglycine, with the identity distances of 6.9 and 7.05 A., respectively.

erations of the internal rotation about a single bond as axis. Pauling and Corey assumed, in their earlier discussion[15] of the configurations of polypeptide chains, that one of the three adjacent single bond on the α-carbon atom should always lie in the plane of the amide group and they eliminated all other orientations. In other words, they accepted the orientations (A), (C), (D) and (F) in Fig. 5, but eliminated (B) and (E). Our studies on internal rotation show however, that the first factor determining stable positions in the internal rotation is the steric repulsion between the groups on both ends of the single bond under consideration and in general the *cis* position (where the azimuthal angle is zero) is unstable, unless there is a great stabilization of this position because of other factors. Therefore, it is probable that orientation (B) or (E), in which there is no *cis* relation, is more stable than (A), (C), (D) or (F) of Fig. 5.

(2) The models explain the identity period actually observed for β-polypeptides. The identity period expected for the fully extended polypeptide chain is 7.23 Å.[14]; that actually observed[14a] is close to the value expected for our models. This value is also expected from the antiparallel-chain pleated sheet of Pauling and Corey. A simple calculation shows that if the azimuthal angle between the $\overset{I}{C}$–N and $\overset{III}{C}$—C bonds about the $\overset{II}{N}$—C axis and that between the $\overset{II}{N}$—C and C—$\overset{IV}{N}$ bonds about the C–C axis are both about 140° and if all the bond lengths and bond angles take the values which Corey and Pauling[14] gave, the identity period of poly-L-alanine becomes about 6.9 Å.

(3) Our models explain quite naturally the small difference between the identity period of polyglycine-I and that of β-poly-L-alanine. In these models, the deformation of the molecule by the same amount of internal rotation from the fully extended form should result in the identity period of 6.9 Å. for β-poly-L-alanine and that of 7.05 Å. for polyglycine-I. The situation is shown in Fig. 7. Suppose that the N—$\overset{II}{C}$—$\overset{III}{C}$ be fixed and the azimuthal angles of internal rotation about the N—$\overset{II}{C}$ and C—$\overset{III}{C}$ axes are changed. In this case, the direction of the fiber axis of β-poly-L-alanine in our model is almost the same as the direction of the fiber axis of the fully extended polypeptide; while the direction of the fiber axis of polyglycine-I in our model makes an angle of about 12° with that of the fully extended polypeptide. Therefore, the identity period of polyglycine-I is calculated as 6.9 Å. cos 12° = 7.05 Å.

(4) By our models, the infrared spectra in the 3470–3420 cm.$^{-1}$ region of AGlyMA and other acetylamino acid N-methylamides in dilute carbon tetrachloride solutions are well explained as shown in the preceding section.

HONGO, TOKYO, JAPAN