A STUDY OF THE CONFORMATIONAL STATES OF CYCLOPEPTIDE SYSTEMS

V. NMR SPECTRA OF CYCLOHEXAPEPTIDES CONSTRUCTED FROM ALANINE AND GLYCINE RESIDUES, THE SPIN–SPIN COUPLING CONSTANTS OF THE NH–CH PROTONS, AND THE "PLEATED-SHEET" STRUCTURE*

V. T. Ivanov, S. L. Portnova,
T. A. Balashova, V. F. Bystrov,
V. V. Shilin, Ya. Bernat,†
and Yu. A. Ovchinnikov

UDC 547.96

In a preceding communication [1], on the basis of the chemical shifts and their dependence on the temperature and the rate of deuterium exchange for the NH groups of cyclohexapeptides (Fig. 1), it was shown that the dominating conformation of these compounds is the "pleated-sheet" with two transannular H bonds. Characteristic for this is a rapid dynamic equilibrium expressed in a migration of the H bonds within the ring (Fig. 2). For a number of compounds it was impossible to identify the dominating structure and to show the amino acid residues forming the transannular H bonds. More definite information on the dominating structures and preferred orientation of the methyl groups of the alanine residues has been obtained in the present work by considering the spin–spin coupling constants of the NH–CH protons of the alanine fragments of the cyclohexapeptides (Tables 1 and 2 in [1]) together with the results of an experimental investigation and a theoretical analysis of similar cyclopeptide systems.

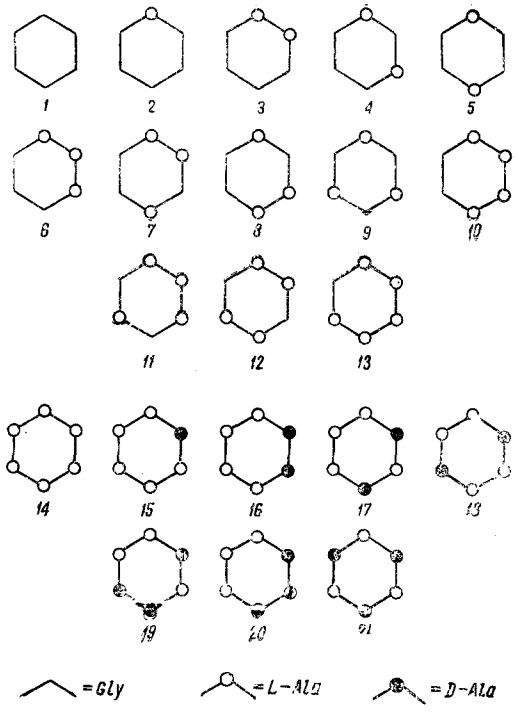

Fig. 1. Cyclic hexapeptides constructed from L- and D-alanine and glycine residues (direction of acylation clockwise).

The Coupling Constant J_{NH-CH} and the Orientation of the Methyl Groups of the Alanine Residues. It has been established [2-4] that large values of $^3J_{NH-CH}$ (8-10 Hz) correspond to a trans or cis orientation of the protons in the NH–C^{α}H fragments ($\Phi \sim 60°$ or $\sim 240°$ for the L-amino acid residues) and small constants (0-3 Hz) correspond to the gauche orientation ($\Phi \sim 150°$ or $\sim 300°$). However, for the compounds that we have studied, as a rule, intermediate values of $^3J_{NH-CH}$ are found, which (taking the participation of several preferred conformations in the confor-

* For the preceding communication, see [1].

† Chemical Faculty of the Polytechnic Institute (Gdansk, Poland).

M. M. Shemyakin Institute of the Chemistry of Natural Compounds of the Academy of Sciences of the USSR. Translated from Khimiya Prirodnykh Soedinenii, No. 3, pp. 339-346, May-June, 1971. Original article submitted February 8, 1971.

TABLE 1. Regions of Permitted Values of Φ and Ψ and the Corresponding Calculated Values of $^3J_{NH-CH}$ for L-Amino Acid Residues in Positions 2 and 3 of the "Pleated-Sheet" Conformation of the Cyclohexapeptides [5]

Parameter	Type of conformation			
	I	II	I'	II'
Φ_2	110–130	110–130	230–250	230–250
Ψ_2	130–180	250–310	180–330	50–110
$^3J_{NHCH}$	1,5–4,5	1,5–4,5	8,5–10,0	8,5–10,0
Φ_3	60–130	240–330	230–300	30–120
Ψ_3	140–220	140–220	140–220	140–220
$^3J_{NHCH}$	1,5–11,0	0–10,0	2,7–10,0	3,0–11,0

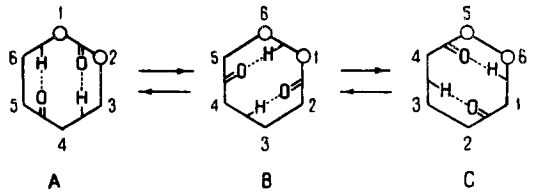

Fig. 2. Dynamic equilibrium of the "pleated-sheet" structures of a cyclohexapeptide (as an example, cyclo-Ala–Ala–Gly–Gly–Gly–Gly is shown), with the migration of the transannular H bonds.

mational equilibrium into account) considerably broadens the field of possible values of the angles Φ and complicates conformational analysis.

It may be observed that in analogy with the chemical shifts [1], in CF_3COOH solutions a somewhat lower (as compared with solutions in DMSO) differentiation of the NH signals with respect to the $^3J_{NH-CH}$ constants is observed (Tables 1 and 2 in [1]), which reflects the high conformational freedom of the cyclopeptides in this solvent. For the cyclopeptides (3)-(13), (14)-(17), (18) and (19) in $(CD_3)_2SO$ solutions the formation of an intramolecular H bond considerably limits the possibility of internal rotation, which permits a number of conclusions concerning the conformation of the molecules. Below we consider information for solutions in $(CD_3)_2SO$.

In the "pleated-sheet" model of the cyclohexapeptides (Fig. 3) it is possible to isolate two groups of amino acid residues. In the first of them there are two residues located on the "fold" of the sheet (1 and 4), the CO and NH groups of which form an intramolecular H bond, and in the second there are four residues (2, 3, 5, and 6) which do not participate in H bonding and form the "corners" of the sheet.

It can be seen from a consideration of molecular models that the methyl groups of the alanine residues in positions 1 and 4 [$Ala_{(1)}$ and $Ala_{(4)}$] may assume the pseudoaxial or pseudoequatorial orientation with respect to the mean plane of the ring (the angle Φ_1 or Φ_2 20-60° or 300-340°, respectively). In the first case one must expect greater values of the $^3J_{NH-CH}$ constants (5.6-10.7 Hz) than in the second case (0-2.9 Hz); values of $^3J_{NH-CH}$ in the range from 3.5 to 6.5 Hz indicate a realization of intermediate orientations of the methyl groups or an equilibrium of the axial and equatorial forms. The NH signals relating to positions 1 and 4 of the alanine residues were determined on the basis of the information given above on intramolecular H bonds. As follows from Tables 1 and 2 in the previous paper [1], the NH groups of the $Ala_{(1)}$ and $Ala_{(4)}$ residues (the signals of which are located in a stronger field and have a smaller temperature gradient) are characterized in the majority of cases by high values of $^3J_{NH-CH}$ (6.5-7.9 Hz), which show the preferred pseudoaxial orientation of the corresponding methyl groups.

So far as concerns the conformation of the ten-membered ring stabilized by H bonds, which is determined by the rotational isomerism of the amino acid residues in positions 2, 3, 5, and 6, as a theoretical analysis has shown [5], the condition for the closure of a ten-membered ring is satisfied by the four permitted regions of the angles Φ and Ψ (I, II, I', and II') given together with the calculated values of $^3J_{NH-CH}$ in Table 1. The results of a consideration of the conformations of the peptide systems in which ten-membered rings stabilized by H bonds have been observed experimentally (see Table 2) have shown that if there are L-amino acid residues in positions 2 and 3 (5 and 6), structures of types I and II are realized (Φ_2 and Φ_5 110-130°; Φ_3 and Φ_6 50-90° or 260°), and conformations I' and II', which are their mirror images (from the point of view of the spatial structure of the peptide skeleton) are absent. A calculation of the potential energies of linear and cyclic peptide systems also leads to the conclusion that structures of types I' and II' have higher energies than I and II and therefore, as a rule, should not take a substantial part in the conformational equilibrium (Table 3) [11, 12]. Thus, the alanine residues in positions 2 and 5 of the cyclohexapeptides studied must be characterized by comparatively low $^3J_{NH-CH}$ constants (1.0-4.5 Hz), while for the $Ala_{(3)}$ and $Ala_{(6)}$ residues the range of possible values of $^3J_{NH-CH}$ corresponding to the permitted regions of Φ (see Table 3) are considerably broader, although higher values of $^3J_{NH-CH}$ (7.0-10.5 Hz) are most likely for them.

This conclusion is in good agreement with the observed values of $^3J_{NH-CH}$ and enables spectral assignments to be carried out and the structures of a number of cyclopeptides containing alanine and glycine

TABLE 2. Experimentally Established Conformations of Peptide Systems with 4→1 H Bonds

Compound	Fragment	$\Phi_2(\Phi_5)$	$\Psi_2(\Psi_5)$	$\Phi_3(\Phi_6)$	$\Psi_3(\Psi_6)$	Literature reference
BrZGlyProLeuGlyOH	-Pro-Leu-	117	151	74	195	[6]
[Gly$_6$]	Gly$_{(2)}$-Gly$^*_{(3)}$-	111	151	86	188	[12]
[Gly$_6$]	Gly$_{(5)}$-Gly$_{(6)}$-	111	151	88	184	[12]
[*D*-Ala-*D*-Ala-Gly$_4$]	-*D*-Ala-*D*-Ala†	—114	—164	—49	—212	[8]
The same	Gly-Gly-	—250*	—195	—287	—163	[8]
Ferrichrome A	Ser-Ser-	123	311	262	178	[7]
Gramicidin S	*D*-Phe-Pro-†	—125	—290	120‡	140	[10]

* For the —Gly—Gly— fragment the conformations with positive and negative coordinates Φ and Ψ are energetically equivalent.
† The energetically equivalent conformations of the D- and L-amino acid residues have values of Φ and Ψ of opposite signs.
‡ The coordinate Φ in the L-Pro residues is fixed at ~120° because the $N-C^{\alpha}$ bond forms part of a pyrrolidine ring.

TABLE 3. Calculated Conformations of Peptide Systems with 4→1 Hydrogen Bonds Corresponding to the Potential Energy Minima

Compound	Fragment	$\Phi_2(\Phi_5)$	$\Psi_2(\Psi_5)$	$\Phi_3(\Phi_6)$	$\Psi_3(\Psi_6)$	Lit. ref.
Ac-Gly-GlyNHCH$_3^*$	-Gly-Gly-	130	140	110	159	
		128	276	250	210	
		141	127	79	218	
AcAlaAlaNHCH$_3^*$	-Ala-Ala-	130	138	120	146	[11]
		132	294	230	216	
		138	129	76	220	
		238	230	235	215†	
		230	110	82	152‡	
[Gly$_6$]	-Gly-Gly-	120—130	290	230—250	200—220	[12]
[Ala$_6$]	-Ala-Ala-	120	290—300	230—240	200—220	[9]
		120	140	70	210	

* The hydrogen bond is formed by the acetyl C=O group and by the methylamide NH group.
† This conformation corresponds to a considerably higher energy than the other conformations given.
‡ In this conformation, hydrogen bonds of the 4→1 and 3→1 types exist simultaneously.

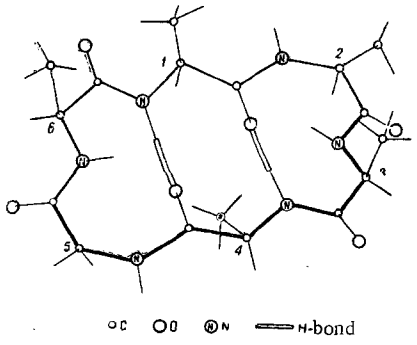

Fig. 3. Preferred conformation of the "pleated-sheet" structure of the cyclohexapeptide (13).

residues to be refined. The majority of the assignments are based on the fact that in the alanine residues that do not participate in intramolecular H bonds, as a rule, it is possible to distinguish the signals relating to the $Ala_{(2)}$ and $Ala_{(5)}$, $Ala_{(3)}$ and $Ala_{(6)}$ residues: in the first case, the $^3J_{NH-CH}$ constants are lower (5.4-6.8 Hz) than in the second case (6.8-8.9 Hz). At the same time, it must be observed that the difference mentioned is not very great and in some cases this considerably complicates the interpretation of the spectra. The averaging of the $^3J_{NH-CH}$ constants is apparently explained by the following factors: 1) for the cyclohexapeptides there is an equilibrium of forms A, B, and C (see Fig. 2) connected with the migration of the H bonds, and 2) in each of the forms A, B, and C in its turn an equilibrium of several conformers exists in which the methyl groups of the $Ala_{(1)}$ and $Ala_{(4)}$ residues occupy different orientations.

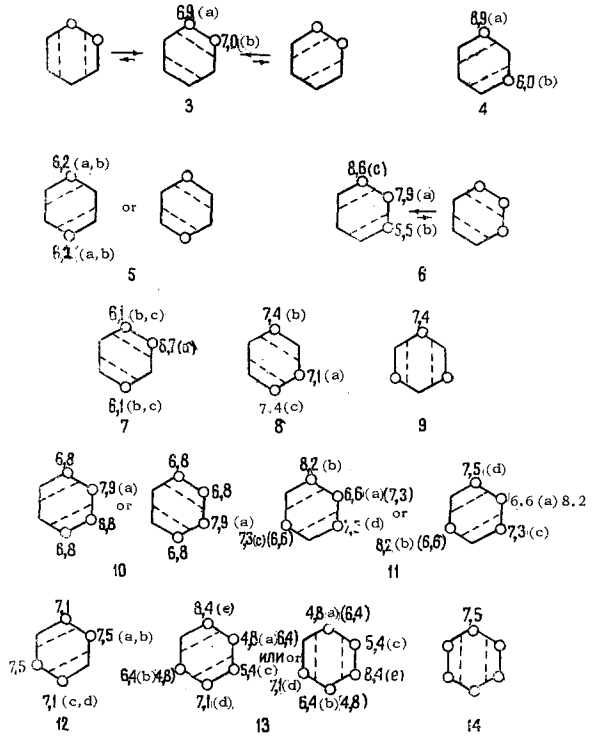

Fig. 4. Positions of the transannular H bonds in the dominating "pleated-sheet" structures of the cyclohexapeptides (3)-(14) and the $^3J_{NH-CH}$ values of the alanine residues.

Dominating "Pleated-Sheet" Structures of the Cyclohexapeptides. In a preceding communication [1], on the basis of the values of the chemical shifts and their dependence on the temperature, dominating "pleated-sheet" structures with transannular H bonds were established (see Fig. 9 in [1]). However, for a number of compounds some indeterminacy remained in the choice of the positions of the H bonds. The taking into account of the considerations given above on the ^3J constants of the NH–CH fragments of the "pleated sheet" makes it possible to assign the signals in the NMR spectra and to draw some conclusions concerning the conformation of the dominating structures. The information obtained is given in Fig. 4. A consideration of the individual cyclohexapeptides follows.

a. cyclo-Gly–Ala–Gly–Gly–Gly–Ala (4). Since in this compound the H bonds are formed between the glycine residues (see Fig. 4), the alanine residues occupy positions 2 and 6. In view of what has been said above, the higher constant (8.9 Hz) must correspond to the $Ala_{(6)}$ residue and the lower one (6.0 Hz) to $Ala_{(2)}$ (see Fig. 4).

b. cyclo-Ala–Ala–Gly–Ala–Gly–Gly (7). In view of the structural similarity, the analogous position of the hydrogen bonds, and the close values of the two $^3J_{NH-CH}$ constants in relation to the cyclopeptide (4) considered above [6.0 and 8.9 Hz for (4), and 6.1 and 8.7 Hz for (7)], it may be considered that the corresponding signals in the spectrum of (7) relate to $Ala_{(2)}$ and $Ala_{(6)}$, and the third signal with $^3J_{NH-CH}$ = 6.1 Hz to $Ala_{(5)}$ (Fig. 4).

c. cyclo-Ala–Ala–Ala–Ala–Ala–Gly (13). In the spectrum of compound (13), the NH signals with constants of 4.8 and 6.4 Hz relate to the $Ala_{(1)}$ and $Ala_{(4)}$ residues. The NH signals with constants of 7.1 and 8.4 Hz must be assigned to the $Ala_{(3)}$ and $Ala_{(6)}$ residues, and that with a constant of 5.4 Hz to $Ala_{(2)}$. On this basis, it may be assumed that of the two possible forms (see Fig. 4) the preferred one is that with the glycine residue in position 5. In the dominating conformation of compound (13), which is shown in Fig. 3, the methyl side chain in position 2 must have the pseudoaxial orientation and those in positions 3 and 6 the pseudoequatorial orientation. These orientations of the side chains are characteristic for the majority of alanylglycine cyclohexapeptides.

d. cyclo-Ala–Ala–Ala–Gly–Gly–Gly (6). For compound (6) it is not possible to perform an analogous assignment of the NH signals. However, in view of the fact that the NH signal with $^3J_{NH-CH}$ = 7.9 Hz relates to $Ala_{(1)}$ and also in view of the structural similarity of compounds (4) and (6), it may be assumed that the smaller constant of 5.5 Hz belongs to the signal from the $Ala_{(2)}$ residue and the larger of 8.6 Hz to $Ala_{(6)}$. A comparatively low temperature gradient is possessed by the NH protons (a, c) (2.9 and $3.3 \cdot 10^{-3}$ ppm/deg, respectively [1]) and by one glycine NH proton (k) ($3.3 \cdot 10^{-3}$ ppm/deg [1]). Unfortunately, it was not possible to determine the value of $\Delta\delta/\Delta T$ for the NH(*l*) proton of the glycine residue which, in all probability, is close to the value for NH (a, c, k). On this basis, it may be concluded that some proportion of the second structure also takes part in the conformational equilibrium.

e. cyclo-$(Ala–Ala–Gly)_2$ (12). By studying the spectra of the specially synthesized deuterated derivative (see preceding communication), we established that in the dominating conformation of compound (12) the alanine residues with $^3J_{NH-CH}$ = 7.1 Hz are in positions 3 and 6 (see Fig. 4).

f. cyclo-Gly–Ala–Gly–Gly–Ala–Gly (5). As shown previously [1], one of the two structures illustrated in Fig. 4 is realized. The comparatively low value of $^3J_{NH-CH}$ (6.2 Hz) permits it to be assigned to $Ala_{(2)}$ and $Ala_{(5)}$ and the second of the two structures shown in Fig. 4 to be regarded as the more likely.

However, this has not yet been firmly shown if one takes into account the fact that, according to the optical rotatory dispersion and circular dichroism curves [13], the conformation of the cyclopeptide (5) differs substantially from the conformations of the other cyclopeptides of this group.

g. cyclo-Ala–Gly–Ala–Ala–Gly–Gly (8). On the basis of the temperature gradients of the chemical shifts of the NH protons (Table 3 in [1]), in compound (8) the H bond is formed by the oppositely placed alanine and glycine residues (see Fig. 4).

h. cyclo-$(Ala–Gly)_3$ (9) and cyclo-$(L\text{-}Ala)_6$ (14). In view of the nature of compounds (9) and (14), all three forms A, B, and C with transannular H bonds for them are equally likely. Consequently, the observed $^3J_{NH-CH}$ constants (7.4 and 7.5 Hz, respectively) are practically equal and are averaged over positions 1, 2 (5), and 3 (6) of the alanine residues.

i. cyclo-Ala–Ala–Ala–Ala–Gly–Gly (10). The chemical shifts show that in compound (10) the alanine and the glycine residues participate in the H bond. Consequently, two structures are possible for it (see Fig. 4). In view of the closeness of the $^3J_{NH-CH}$ constants for the NH (b, c, d) signals and their intermediate value (6.8 Hz), it is impossible to choose between these structures.

j. cyclo-Ala–Ala–Ala–Gly–Ala–Gly (11). It follows from the high values of $\Delta\delta/\Delta T$ of the glycine protons of compound (11) that positions 1 and 4 in the ring are occupied by alanine residues. Under these conditions, several assignments of the $^3J_{NH-CH}$ constants are possible for the given structure that correspond in equal measure to the analysis that we have performed (see Fig. 4).

k. cyclo-Ala–Ala–Gly–Gly–Gly–Gly (3). Compound (3), in which the $Ala_{(2)}$ and $Ala_{(3)}$ residues are characterized by practically identical $^3J_{NH-CH}$ values (6.9 and 7.0 Hz), behaves somewhat unusually. The lowest temperature gradient is observed for the NH (k) glycine proton (Table 3 in [1]). For the NH (a) proton of the alanine residue and the NH (*l*) proton of the glycine residue the temperature gradients are practically equal. This shows that considerable amounts of forms B and C participate in the conformational equilibrium (see Fig. 2), although the equilibrium is also shifted in the direction of the production of a form corresponding to the crystalline structure of this compound [8]. It is possible that the mirror forms I′ and II′, to position 2 of which correspond the high values of $^3J_{NH-CH}$ (see Table 1), are also present in this case. Such a hypothesis is in harmony with the considerably lower intensities of the Cotton effects in (3) as compared with the other cyclopeptides [13].

Thus, the present work has shown that cyclohexapeptides constructed from L(D)-alanine and glycine residues connected with one another by trans amide bonds do not possess a fixed spatial structure in polar solvents but take part in a complex conformational equilibrium. A characteristic feature of them is the existence of two transannular IMHBs whose stability in polar solvents is explained by the correspondence of the found conformation to the minimum of the potential energy of the nonvalent interactions in the cyclohexapeptide systems. In solutions of the cyclopeptides studied there is a rapid equilibrium of three different hydrogen-bonded structures. Of the latter, only the dominating one can be isolated in the majority of cases. The lateral substituents of the L-alanine residues forming the hydrogen bonds have the pseudoaxial orientation with respect to the mean plane of the ring. The fragments of the $NH-C^{\alpha}H$ residues of $Ala_{(2)}$ adopt the preferential gauche orientation (Φ 110-130°), and $Ala_{(3)}$ the trans or cis orientation (Φ 50-90° or 230-260°).

SUMMARY

1. On the basis of an analysis of the experimental and calculated data on peptide systems containing IMHBs of the 4→1 type, it has been shown that positions 2 and 5 in the preferred conformations of the cyclohexapeptides must be characterized by low values of the $^3J_{NH-CH}$ constants (1-5 Hz), and positions 3 and 6 by large values of these constants (7-10.5 Hz).

2. The positions of the IMHBs in cyclohexapeptides including L-alanine and glycine residues have been refined, and the signals of the NH groups in the NMR spectra have been assigned to the individual amino acid residues.

3. In the dominating conformations of the cyclohexapeptides, the lateral methyl groups of the alanine residues generally occupy the pseudoaxial orientation in positions 1, 3, 4, and 6 and the pseudoequatorial orientation in positions 2 and 5.

LITERATURE CITED

1. S. L. Portnova, T. A. Balashova, V. F. Bystrov, V. V. Shilin, R. Bernat, V. T. Ivanov, and Yu. A. Ovchinnikov, Khim. Prirodn. Soedin., 7, 323 (1971).
2. S. L. Portnova, V. F. Bystrov, V. I. Tsetlin, V. T. Ivanov, and Yu. A. Ovchinnikov, Zh. Obshch. Khim. 38, 428 (1968).
3. V. F. Bystrov, S. L. Portnova, V. I. Tsetlin, V. T. Ivanov, and Yu. A. Ovchinnikov, Tetrahedron, 25, 413 (1969).
4. V. F. Bystrov, S. L. Portnova, T. A. Balashova, P. V. Kostetskii, V. I. Tsetlin, V. T. Ivanov, and Yu. A. Ovchinnikov, Tetrahedron Lett., 1969, 5225.
5. M. Venkatachalam, Biopolymers, 6, 1425 (1968).
6. T. Ueki, T. Ashid, M. Kakudo, Y. Sasada, and Y. Katsube, Nature, 216, 1205 (1967).
7. A. Zalkin, J. D. Forrester, and D. H. Templeton, J. Amer. Chem. Soc., 88, 1810 (1966).
8. I. L. Karle, J. W. Gibson, and J. Karle, J. Amer. Chem. Soc., 92, 3755 (1970).
9. C. Ramakrishnan and K. P. Sarathy, Int. J. Protein Research, 1, 103 (1969).
10. Yu. A. Ovchinnikov, V. T. Ivanov, A. V. Evstratov, V. F. Bystrov, N. D. Abdullaev, E. M. Popov, G. M. Lipkind, S. A. Arkhipova, E. S. Efremov, and M. M. Shemyakin, Biochem. Biophys. Res., Commun., 37, 668 (1969).
11. G. M. Lipkind, S. F. Arkhipova, and E. M. Popov, Mol. Biol., 4, 509 (1970).
12. I. L. Karle and J. Karle, Acta Cryst., 16, 969 (1963).
13. V. T. Ivanov, G. A. Kogan, E. N. Mescheryakova, V. V. Shilin, and Yu. A. Ovchinnikov, Khim. Prirodn. Soedin., 7, 309 (1971).