eters. The final R factors for the ethanol and chloroform adducts are, respectively, 8.0 and 13.0% for all the data and 6.7 and 10.2% if the weakest reflections ($|F_o| \leq 5.0$) and those having large extinction errors are omitted.[14]

Finding a Cl atom at $z = 1/2$, the waist of the cage, suggested the possibility of including a larger molecule which could extend throughout the entire length of the cage. No such molecules were included in published studies.[6] However, an investigation on the occupancy of various isomers of heptane has been made by Goldup.[15] The cross section of the waist of the cage is large enough to accommodate CH_2 groups arranged in a zigzag chain. Crystals of an adduct with n-heptyl alcohol, which has an overall molecular length of ~10 Å, have been prepared and the structure is currently under investigation.

(14) For a complete list of observed and calculated structure factors order Document NAPS-00958 from ASIS National Auxiliary Publications Service, % CCM Information Corp., 909 Third Ave., New York, N. Y. 10022, remitting $1.00 for microfiche or $3.00 for photocopies. Make checks payable to CCMIC-NAPS.

(15) A. Goldup, British Petroleum Co., Ltd., private communication.

The Conformation and Crystal Structure of the Cyclic Polypeptide ┌Gly-Gly-D-Ala-D-Ala-Gly-Gly┐ · $3H_2O$

I. L. Karle, J. W. Gibson, and J. Karle

Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, D. C. 20390. *Received November 12, 1969*

Abstract: A single-crystal X-ray diffraction analysis has been made on the structure of the cyclic polypeptide ┌Gly-Gly-D-Ala-D-Ala-Gly-Gly┐ · $3H_2O$. If the CH_3 groups in the alanine residues are disregarded, the molecule very nearly has a center of symmetry. All the peptide units are planar and in the *trans* conformation. In order to close the ring, the residues, from N to N, have the conformations *cis,trans,cis,cis,trans,cis*. The 18-membered ring is stabilized by two intramolecular hydrogen bonds and by hydrogen bonding to H_2O molecules. Each NH and CO moiety participates in one or more hydrogen bonds. The material crystallizes in the orthorhombic space group $P2_12_12_1$ with cell parameters $a = 12.662$, $b = 18.102$, and $c = 8.678$ Å. The X-ray intensity data were collected with an automatic diffractometer and refined to an $R = 6.5\%$. The crystal structure was solved by the symbolic addition procedure for phase determination for noncentrosymmetric crystals.

The conformation of cyclic polypeptides is of fundamental importance in understanding the relationship between the structure and function of natural products.[1] Detailed structural information obtained from single-crystal X-ray diffraction analyses has been published for the cyclic tetradepsipeptide ┌D-HyIv-L-MeIleu-D-HyIv-L-MeLeu┐,[2] for ferrichrome-A · $4H_2O$,[3] and for cyclohexaglycyl · $1/2H_2O$.[4] In the cyclic tetradepsipeptide, the two peptide units assumed the *cis* conformation, contrary to the general observation from linear polypeptides that the *trans* conformation prevails. In the crystal of the cyclic hexaglycyl, there are four distinct conformers in the same unit cell. Only one of the conformers has intramolecular hydrogen bonds. Accordingly, it seemed appropriate to investigate the crystal structure of another cyclic hexapeptide in order to compare its conformation with that of the cyclohexaglycyl. The material reported in this investigation is the cyclic polypeptide I. It was prepared by

┌Gly-Gly-D-Ala-D-Ala-Gly-Gly┐

I

the method of Gerlach, Ovchinnikov, and Prelog[5] anp made available to us by Dr. Bernhard Witkop of the National Institutes of Health.

Experimental Section

The crystal was a thin plate, roughly hexagonal in shape. The b axis was perpendicular to the plate, and the crystal was mounted along the c axis. Cell constants and other physical data are listed in Table I.

Table I. Physical Data

Mol formula	$C_{14}H_{22}O_6N_6 \cdot 3H_2O$
Mol wt	424.4
Habit	Thin tabular (010)
Crystal size	0.7 × 0.08 × 0.7 mm
Space group	$P2_12_12_1$
a	12.662 ± 0.003 Å
b	18.102 ± 0.005 Å
c	8.678 ± 0.002 Å
V	1989.06 Å^3
ρ calcd	1.417 g/cm^3
Radiation	Cu Kα, 1.5418 Å
No. of independent reflections	1975
Molecules per unit cell	4

The X-ray intensity data were collected from one crystal on a four-circle automatic diffractometer using the $\theta,2\theta$ (moving crystal–

(1) See, *e.g.* M. M. Shemyakin and Yu. A. Ovchinnikov, *Recent Develop. Chem. Natur. Carbon Compounds*, **2**, 1 (1967); *Chem. Abstr.*, **68**, 87510 (1968); and Conformation of Biopolymers, International Symposium in Madras, G. N. Ramachandran, Ed., 1967.

(2) J. Konnert and I. L. Karle, *J. Amer. Chem. Soc.*, **91**, 4888 (1969).

(3) A. Zalkin, J. D. Forrester, and D. H. Templeton, *ibid.*, **88**, 1810 (1966).

(4) I. L. Karle and J. Karle, *Acta Crystallogr.*, **16**, 969 (1963).

(5) H. Gerlach, Yu. A. Ovchinnikov, and V. Prelog, *Helv. Chim. Acta*, **47**, 2294 (1964).

Table II. Phase Assignments for Specifying the Origin and Enantiomorph and Implementing Equation 1

$\vec{h}$	$\phi_{\vec{h}}$	$\lvert E_{\vec{h}}\rvert$
11 11 0	$+\pi/2$	3.06
2 0 1	$+\pi/2$	2.43
4 1 0	0	2.39
0 5 6	$+\pi/2$	2.32
0 3 2	$p\ (\pm\pi/2)$	3.41
3 0 1	$q\ (\pm\pi/2)$	2.40

Table III. Fractional Coordinates and Thermal Parameters[a]

Atom	x	y	z	$\beta_{11} \times 10^4$	$\beta_{22} \times 10^4$	$\beta_{33} \times 10^4$	$\beta_{12} \times 10^4$	$\beta_{13} \times 10^4$	$\beta_{23} \times 10^4$
N_1	0.2190	0.1027	1.0343	51	12	69	−3	5	−1
C_1^{α}	0.1695	0.1502	1.1482	83	17	63	−2	21	0
C_1'	0.1602	0.2296	1.0956	42	14	66	−1	−1	−3
O_1	0.1421	0.2792	1.1904	75	20	57	11	8	−7
N_2	0.1709	0.2428	0.9450	45	11	60	2	10	0
C_2^{α}	0.1661	0.3162	0.8807	43	15	82	3	13	3
C_2'	0.2752	0.3471	0.8456	44	17	56	2	1	−4
O_2	0.3574	0.3160	0.8806	43	20	85	3	9	13
N_3	0.2712	0.4128	0.7743	42	11	78	5	0	1
C_3^{α}	0.3660	0.4549	0.7350	46	15	86	−1	−13	2
C_3^{β}	0.3343	0.5310	0.6807	62	16	217	1	0	7
C_3'	0.4365	0.4186	0.6160	38	13	75	4	10	6
O_3	0.5277	0.4408	0.5983	46	26	111	13	2	2
N_4	0.3952	0.3617	0.5369	38	16	69	−5	5	−3
C_4^{α}	0.4557	0.3179	0.4238	34	21	65	0	5	−1
C_4^{β}	0.4241	0.3390	0.2579	87	21	71	1	20	9
C_4'	0.4416	0.2368	0.4551	32	19	73	0	1	−3
O_4	0.4438	0.1903	0.3496	84	19	61	4	1	−4
N_5	0.4335	0.2168	0.6023	41	16	65	−1	4	−4
C_5^{α}	0.4332	0.1396	0.6544	38	20	86	3	7	2
C_5'	0.3245	0.1131	0.7019	43	17	55	−2	4	−5
O_5	0.2430	0.1485	0.6744	35	20	104	5	−2	5
N_6	0.3235	0.0487	0.7774	43	13	86	2	3	−3
C_6^{α}	0.2237	0.0147	0.8235	50	14	120	−2	18	−7
C_6'	0.1636	0.0536	0.9519	50	12	81	0	2	6
O_6	0.0713	0.0385	0.9755	38	26	121	−5	13	−4
W (1)	0.4109	0.1810	1.0251	50	26	77	−5	−8	−3
W (2)	0.1991	0.2811	0.5049	52	23	93	−6	−10	−1
W (3)	0.0590	0.4497	0.6925	81	30	467	−3	−114	26
				Standard Deviations					
C	0.0005	0.0004	0.0008	5	2	9	2	6	4
N	0.0004	0.0003	0.0006	3	2	7	2	5	3
O	0.0004	0.0003	0.0005	4	2	7	2	4	3
W (1–2)	0.0004	0.0003	0.0006	3	2	7	2	4	3
W (3)	0.0005	0.0003	0.0011	5	2	22	3	10	6

[a] The thermal parameters are expressed in the form $T = \exp -(\beta_{11}h^2 + \beta_{22}k^2 + \beta_{33}l^2 + 2\beta_{12}hk + 2\beta_{13}hl + 2\beta_{23}kl)$.

moving counter) technique with a 1.8° + $2\theta(\alpha_2) - 2\theta(\alpha_1)$ scan over 2θ. The intensities were corrected for Lorentz and polarization factors and normalized structure factors $|E|$ were derived. The values of the statistical averages for $\langle|E|\rangle$ and for $\langle||E|^2 - 1|\rangle$ for the noncentrosymmetric reflections were 0.880 and 0.775, respectively, as compared with the theoretical values of 0.886 and 0.736.

Structure Determination

Phases were derived directly from the measured X-ray intensities by means of the symbolic addition procedure for noncentrosymmetric crystals.[6] Table II lists the assignment of phase values made to reflections with large $|E|$ magnitudes in order to specify the origin, choose the enantiomorph, and implement the sum of angles formula

$$\phi_{\vec{h}} \approx \langle \phi_{\vec{k}} + \phi_{\vec{h}-\vec{k}} \rangle_{\vec{k}_r} \qquad (1)$$

With the assignments in Table II, phases for 57 additional reflections with large $|E|$ values were derived with eq 1. The phase values for these 63 reflections were refined and about 600 additional phases for reflections with $|E| > 1.0$ were obtained with the tangent formula[7] (2) for the four cases in which p and q were given the values $+\pi/2$ or $-\pi/2$. No meaningful structure was apparent in any of the four E maps based on the derived phases.

$$\tan \phi_{\vec{h}} = \frac{\sum_{\vec{k}} E_{\vec{k}} E_{\vec{h}-\vec{k}} \sin (\phi_{\vec{k}} + \phi_{\vec{h}-\vec{k}})}{\sum_{\vec{k}} E_{\vec{k}} E_{\vec{h}-\vec{k}} \cos (\phi_{\vec{k}} + \phi_{\vec{h}-\vec{k}})} \qquad (2)$$

It was suspected that the cause of the difficulty lay in an incorrect indication from eq 1. Such an incorrect indication, if it occurred early in the phase determination, would generate many more incorrect phases. Accordingly, an auxiliary phase-determining formula, an alternative form of $B_{3,0}$,[8] was used to corroborate each of the first twenty phase indications from eq 1. The

		$\lvert E\rvert$	ϕ
$\vec{k}$	0 8 8	2.21	$+\pi/2 + p$
$\vec{h} - \vec{k}$	0 3 $\bar{2}$	3.41	$-p$
$\vec{h}$	0 11 6	3.30	$+\pi/2$

(6) I. L. Karle and J. Karle, *Acta Crystallogr.*, **17**, 835 (1964); J. Karle and I. L. Karle, *ibid.*, **21**, 849 (1966).

(7) J. Karle and H. Hauptman, *ibid.*, **9**, 635 (1956).

(8) See eq 4 in J. Karle, *ibid.*, in press.

Table IV. Approximate Coordinates for Hydrogen Atoms

H atom attached to	x	y	z
N_1	0.302	0.120	1.008
C_1^α	0.085	0.138	1.163
C_1^α	0.183	0.147	1.235
N_2	0.195	0.200	0.915
C_2^α	0.117	0.350	0.933
C_2^α	0.078	0.328	0.828
N_3	0.217	0.445	0.782
C_3^α	0.410	0.463	0.825
C_3^β	0.277	0.540	0.613
C_3^β	0.292	0.553	0.757
C_3^β	0.392	0.555	0.658
N_4	0.318	0.340	0.565
C_4^α	0.540	0.325	0.430
C_4^β	0.333	0.325	0.275
C_4^β	0.423	0.387	0.242
C_4^β	0.458	0.318	0.187
N_5			
C_5^α	0.493	0.128	0.725
C_5^α	0.483	0.112	0.587
N_6	0.383	0.028	0.808
C_6^α	0.225	−0.032	0.843
C_6^α	0.168	0.015	0.742

E map computed with the assignments $p = +\pi/2$ and $q = -\pi/2$ had 21 large peaks which could be associated with 21 atoms in the molecule. Phases based on this partial structure were used in a recycling procedure[9] with eq 2 to refine phase values and obtain additional ones. An E map computed with the new phases showed 24 atoms of the cyclopeptide and oxygen atoms for two H_2O molecules very distinctly. A difference Fourier map located the two methyl carbon atoms in the alanine moieties and the remaining H_2O of crystallization.

The atomic coordinates and thermal parameters for the 29 C, N, and O atoms were refined by full-matrix least squares. The function minimized was $\Sigma_w(|F_o| - |F_c|)^2$ with $w = 0.5$ for $|F_o| = 0$, $w = 1$ for $|F_o| < 30$, and $w = 30/|F_o|$ for $|F_o| \geqslant 30$. Atomic scattering factors used were those listed in the "International Tables for X-Ray Crystallography." A difference map computed after the anisotropic refinement, $R = 8.9\%$, revealed the approximate positions of 21 of the 22 hydrogen atoms associated with the peptide molecule. None of the six hydrogen atoms associated with the three water molecules was found. When the

Table V. Bond Lengths in Angstrom Units[a]

	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	Av	Rms deviation
$N_jC_j^\alpha$	1.452	1.442	1.463	1.475	1.469	1.462	1.460	0.010
$C_j^\alpha C_j^\beta$			1.511	1.543				
$C_j^\alpha C_j'$	1.513	1.522	1.514	1.504	1.515	1.522	1.515	0.006
$C_j'O_j$	1.238	1.222	1.232	1.243	1.237	1.218	1.232	0.009
$C_j'N_{j+1}$	1.335	1.341	1.344	1.331	1.338	1.339	1.338	0.004

[a] The standard deviations for the individual bond lengths, as derived from the least-squares refinement of the structural parameters, are near 0.010 Å.

Table VI. Bond Angles in Degrees[a]

	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	Av	Rms deviation
$C_{j-1}'N_jC_j^\alpha$	122.1	122.6	122.6	123.4	123.6	120.6	122.5	1.0
$N_jC_j^\alpha C_j'$	113.0	112.2	114.6	110.1	112.7	116.0	113.1	1.9
$N_jC_j^\alpha C_j^\beta$			109.3	110.7				
$C_j'C_j^\alpha C_j^\beta$			109.8	112.3				
$C_j^\alpha C_j'O_j$	120.1	123.6	117.8	121.6	122.8	120.0	121.0	1.9
$C_j^\alpha C_j'N_{j+1}$	117.2	112.6	116.9	116.5	114.7	115.9	115.6	1.6
$O_jC_j'N_{j+1}$	122.7	123.8	123.4	121.7	122.5	124.1	123.0	0.8

[a] The standard deviations of the individual bond angles, as derived from the least-squares refinement of the structural parameters, are near 0.6°.

Table VII. Conformational Angles for cyclo-(Gly-Gly-D-Ala-D-Ala-Gly-Gly)

	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
ϕ_j	73°32′	−78°54′	−114°20′	−49°05′	74°39′	109°52′
ψ_j	−163°50′	−6°01′	−164°47′	148°53′	11°38′	164°54′
ω_j	2°16′	1°54′	2°06′	4°06′	−3°40′	−8°19′

following relationship among the phases was shown to be wrong. It has been observed in the past that when three reflections of the type $0kl$ (or $h0l$ or $hk0$) are combined, the phase indication is frequently wrong by π, even though each $|E_{\vec{h}}|$ magnitude is large and the associated probability that the phase is correct is near 100%. Accordingly, the phase for 0 11 6 was changed to $-\pi/2$ and the phase determination repeated. An

21 hydrogen atom positions were included as constant parameters in a final refinement, the R factor was reduced to 6.5%.[10]

(9) J. Karle, *Acta Crystallogr.*, *B*, 24, 182 (1968).

(10) Observed andc alculated structure factors may be obtained by ordering Document NAPS-00949 from ASIS National Auxiliary Publications Service, c/o CCM Information Corp., 909 Third Ave., New York, N. Y., 10022, remitting $1.00 for microfiche and $3.00 for photocopies. Make checks payable to CCMIC-NAPS.

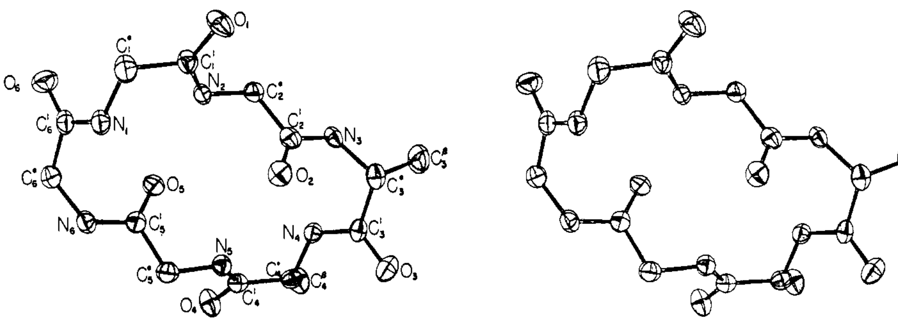

Figure 1. A stereodrawing depicting the configuration of Gly-Gly-Gly-Gly-D-Ala-D-Ala. The ellipsoids are related to the thermal motion of each atom. They correspond to 50% probability. The figure was drawn by a computer from a program by C. K. Johnson (Oak Ridge National Laboratory, Oak Ridge, Tenn.) and should be viewed with a three-dimensional viewer for printed stereophotographs.

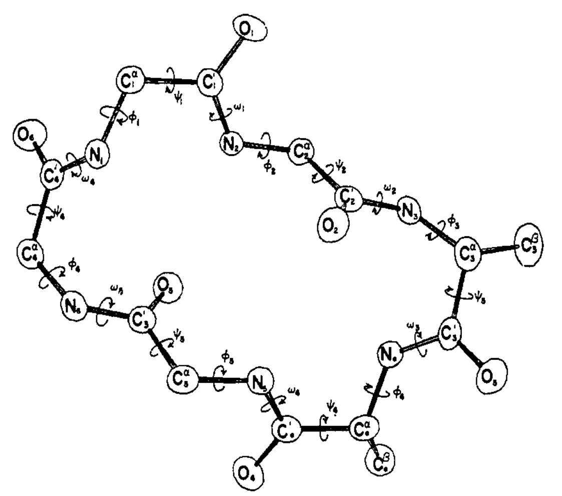

Figure 2. Labeling of atoms and rotational angles.

The refined coordinates and thermal parameters of the heavy atoms and the approximate coordinates of the hydrogen atoms are listed in Tables III and IV. Phase values computed from the coordinates will differ from those assigned in Table II since the enantiomorph was arbitrarily specified by a phase at the beginning of the investigation. The resulting structure was the mirror image of the known stereoconfiguration of amino acids. Accordingly all z coordinates were changed to $-z$ to depict the proper configuration for the D-alanine moieties.

Discussion

The six peptide units in the cyclic hexapeptide are very similar to each other. All are in the *trans* conformation. Bond lengths for corresponding atom pairs, shown in Table V, are generally within one standard deviation of each other. The averages of each set of six values are within 0.01 Å of the values derived from a weighted average of the results of three-dimensional crystal-structure analyses of linear di- and tripeptides.[11a] There is a greater spread of values for similar angles in the six peptide units as shown in Table VI. Nevertheless, the average values for each set of six are within 0.5° for the averages found in linear polypeptides[10] except for the $[N_jC_j^{\alpha}C_j']$ angles which average 2° larger in this cyclic polypeptide than in linear polypeptides.

The configuration of the molecule is illustrated in the stereodrawings, Figure 1, and in Figure 2. The atoms are labeled according to the convention proposed by Edsall, *et al.*,[11b] and the conformational angles are listed in Table VII. The fully stretched polypeptide chain is characterized by the rotational angles $\phi_j = \psi_j = \omega_j = 0$. An inspection of Figure 1 and Table VII shows that the peptide units are in the *trans* conformation with all the ω_j values close to zero. The two residues from N_2 to N_3 and from N_5 to N_6 are near the *trans* conformation with ψ_2 and ψ_4 near 0°, whereas the other four residues are near the *cis* conformation with ψ_j near 160°. If the two CH_3 groups which extend outward from the ring are disregarded, the molecule possesses an approximate center of symmetry. This is reflected in similar values and opposite signs for the rotational angles for the pairs $j = 1$ and 4, $j = 2$ and 5, and $j = 3$ and 6. The individual peptide units are almost planar with an average deviation of atoms of 0.02 Å from the least-squares plane of each unit and a maximum deviation of 0.06 Å. Dihedral angles between the planes of adjacent peptide units range from 67 to 110°.

The 4Gly·2D-Ala molecule crystallizes in an orthorhombic cell with one molecule per asymmetric unit and one conformation for the molecule. In contrast 6Gly crystallizes in the triclinic system with eight molecules in a unit cell and four different conformers in the asymmetric unit.[4] Figure 3a–d illustrates the four conformations assumed by 6Gly. Four molecules in a unit cell of 6Gly assume conformation a, two assume conformation b, and c and d occur once each. The conformation of the 4Gly·D-2Ala molecule is very similar to conformation a, the most prevalent form of 6Gly. Each has two internal NH···O hydrogen bonds. None of the other conformations of 6Gly has internal hydrogen bonds, only intermolecular hydrogen bonds. All the peptide units in each conformation of 6Gly are nearly planar and in the *trans* conformation, *i.e.*, ω_j is near zero. The differences occur in the arrangement of residues, *i.e.*, in the relationships between successive N atoms which are near the *cis*, *trans*, or skew[12] con-

(11) (a) R. E. Marsh and J. Donohue, *Advan. Protein Chem.*, **22**, 249 (1967); (b) J. T. Edsall, P. J. Flory, J. C. Kendrew, A. M. Liquori, G. Nemethy, G. N. Ramachandran, and H. A. Scheraga, *J. Mol. Biol.*, **15**, 399 (1966).

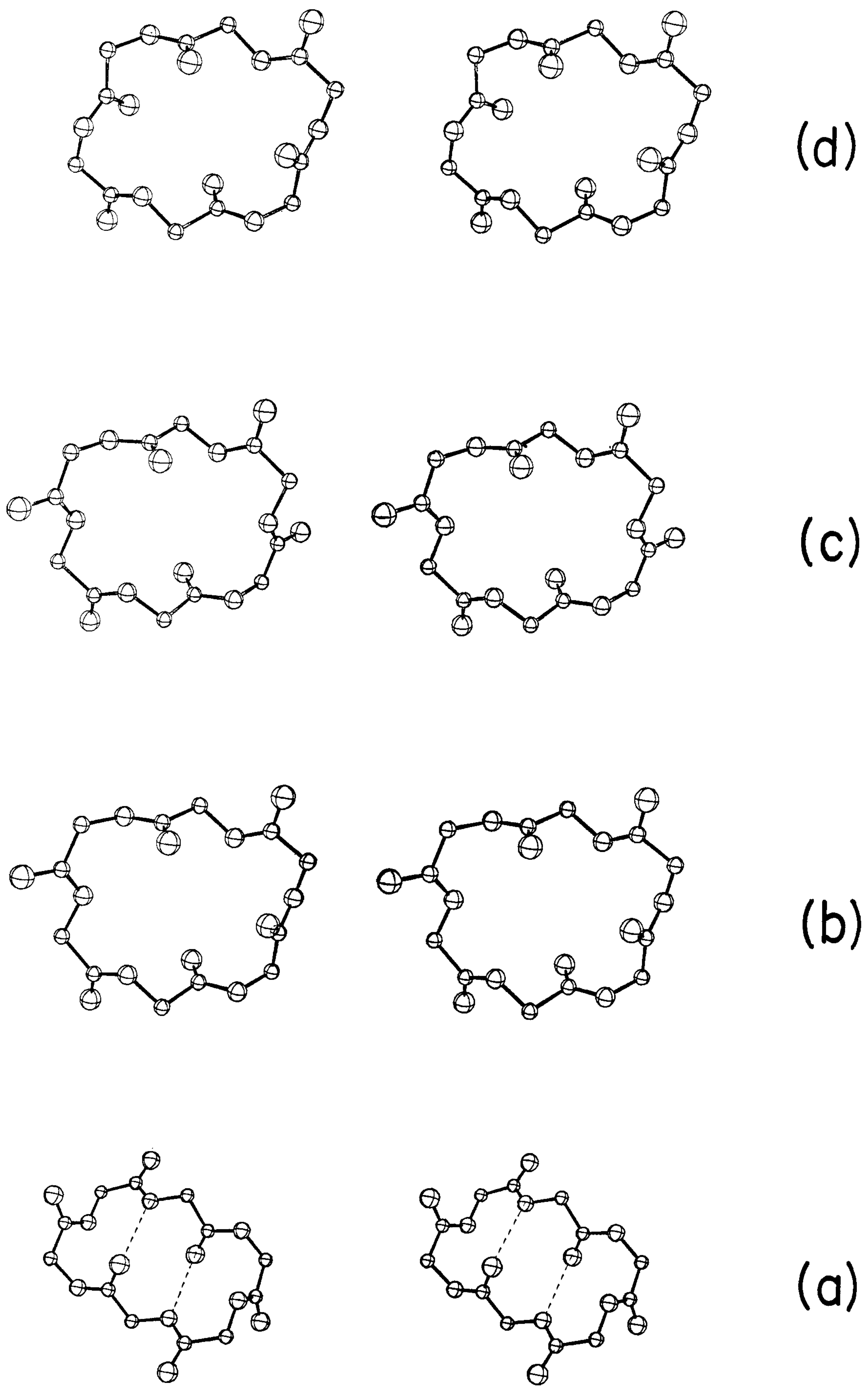

Figure 3. Four different conformers which exist in one unit cell of cyclohexaglycyl·$^1/_2H_2O$. Conformation a occurs four times, b occurs twice, and c and d occur once each. The different size spheres depict C, N, and O atoms in order of size.

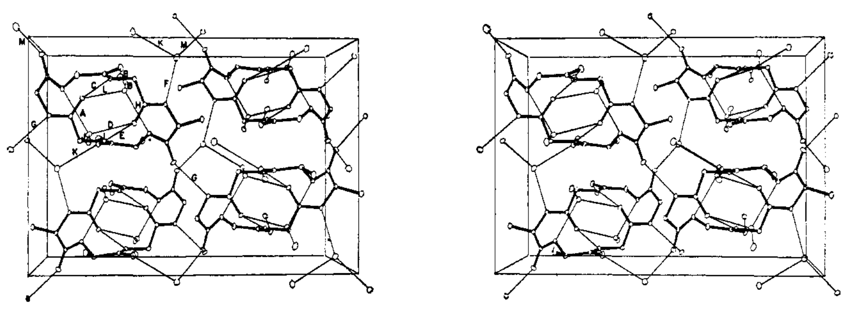

Figure 4. The molecular packing in a unit cell of 4-Gly·2-D-Ala $\cdot 3H_2O$. The axial directions are a ↓, b →, and c up out of the plane of the paper. Symbols A–M denote the twelve independent hydrogen bonds, which are drawn with light lines.

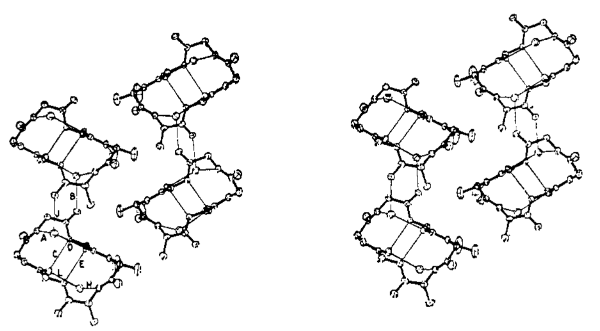

Figure 5. A stereoview of one layer of molecules at right angles to Figure 4. The axial directions are b → and c ↑. The hydrogen bonds are depicted by light lines.

figurations. For the four conformers of 6Gly the approximate arrangements are (a) *cis, cis, trans, cis, cis-trans;* (b) *cis, cis,* skew, *cis,* skew, skew; (c) *cis, cis,* skew, *cis, cis,* skew; (d) *cis,* skew, *cis,* skew, skew.

The crystal of 4Gly·2D-Ala contains three molecules of water of crystallization per molecule of cyclohexapeptide. An efficient scheme of hydrogen bonding exists which utilizes all the C=O and NH groups and has three bonds to each H_2O molecule. Figure 4 illustrates all the hydrogen bonds, Figure 5 shows most of them in a different view, and their lengths are listed in Table VIII. Of the twelve independent hydrogen bonds, nine of them involve the H_2O molecules, there is one intermolecular NH···O=C, bond G, and there are two intramolecular NH···O=C bonds, C and E. The bond lengths for C and E are 3.04 and 3.16 Å, somewhat longer than the value of 2.96 Å observed for similar bonds in hexaglycyl[4] and the median value of ~2.90 Å observed for intermolecular NH···O=C bonds in linear polypeptides.[10] If bond length is a measure of bond strength, then these intramolecular hydrogen bonds are relatively weak. The 18-membered ring of one molecule is held rigidly not only by the two intramolecular hydrogen bonds, C and E, but also by hydrogen bonds A and D which make the linkage

$$N_1H\cdots \underset{\displaystyle H}{\underset{|}{O}}H\cdots O_2{=}C$$

(with W_1) and hydrogen bonds H and L which make the similar linkage

$$NH_4\cdots \underset{\displaystyle H}{\underset{|}{O}}H\cdots O_5{=}C$$

(with W_2), Figure 4.

Table VIII. Hydrogen Bonds

Label	Donor	Acceptor	Symmetry operation on acceptor			Length, Å
A	N_1	W_1	x	y	z	2.814
B	W_2	O_1	x	y	$z-1$	2.824
C	N_2	O_5	x	y	z	3.043
D	W_1	O_2	x	y	z	2.829
E	N_5	O_2	x	y	z	3.159
F	N_3	W_3	x	y	z	2.857
G	N_6	O_3	$1-x$	$1/2+y$	$1\,1/2-z$	2.921
H	N_4	W_2	x	y	z	2.893
J	W_1	O_4	x	y	$1+z$	2.851
K	W_3	O_4	$-1/2+x$	$1/2-y$	$1-z$	2.947
L	W_2	O_5	x	y	z	2.870
M	W_3	O_6	$-x$	$1/2+y$	$1\,1/2-z$	2.727

(12) In this case, let skew be defined as ~60° away from the *trans* position.