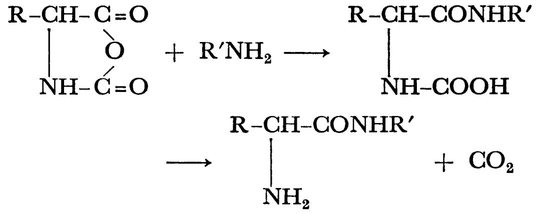
The Crystal and Molecular Structure of *N*-Carboxy-*γ*-benzyl-L-glutamate Anhydride

Hitoshi KANAZAWA,* Tohru KAWAI, Yuji OHASHI,** and Yoshio SASADA**
Department of Polymer Technology, Tokyo Institute of Technology, O-okayama, Meguro-ku, Tokyo 152
***Laboratory of Chemistry for Natural Products, Tokyo Institute of Technology, Nagatsuta, Midori-ku, Yokohama 227*

(Received January 26, 1978)

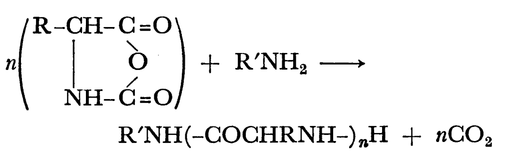
The structure of *N*-carboxy-*γ*-benzyl-L-glutamate anhydride has been determined by X-ray analysis. The crystal is orthorhombic; space group, $P2_12_12_1$; $Z=4$ with $a=7.766(1)$, $b=27.470(4)$, and $c=5.948(1)$ Å. The structure was deduced by the direct method and refined to a final R value of 0.057 for 1504 observed reflections by the block-diagonal least-squares method. The hydrogen bond is formed between the imino group and the carbonyl group of the side chain. The dimension of the five-membered ring is affected by the intermolecular hydrogen bond as compared with the crystal structures of the other *N*-carboxy anhydrides of α-amino acid.

The *N*-carboxy anhydrides of α-amino acids (amino acid NCAs) are often used for the syntheses of polypeptides. The polymerization mechanism differs with the kind of initiator. When a primary amine is used as an initiator, polymerization is considered to proceed as follows,[1] the first step being the formation of an amino acid derivative by interaction of the NCA and the initiator.



(1)

The amino acid derivative formed in (1) can react successively with amino acid NCA in a similar manner, leading to polymeric products. The over-all process is written as follows.



(2)

The polymerization of amino acid NCA was carried out not only in the homogeneous system but also in the heterogeneous system and in crystalline state. When butylamine (an initiator) is added into the suspension of the crystals of amino acid NCA in hexane, polymerization occurs in the solid state and a polypeptide of a considerably high molecular weight can be obtained. Polymerizability differs a great deal from that in a heterogeneous system using acetonitrile as a solvent of NCA.[2] In order to interpret the polymerization of amino acid NCAs in the solid state, we have attempted to analyze their crystal structures. In previous studies, the crystal structures of glycine NCA[3] and L-alanine NCA[4] were reported. This paper deals with the crystal structure of *γ*-benzyl-L-glutamate NCA, which is most extensively used for the study of synthetic polypeptides.

Experimental

γ-Benzyl-L-glutamate NCA (BLG NCA) was synthesized in the same way as that for glycine NCA, and purified by crystallization with ethyl acetate–hexane several times. BLG NCA thus obtained (about 100 mg) was dissolved in an appropriate vessel with 5 ml of ethyl acetate. The vessel was placed in a covered bottle containing hexane and allowed to stand at room temperature for a few days. Colorless plate crystals were obtained. The unit-cell dimensions were determined by the least-squares refinement with 25 high-angle reflections. The crystal data are summarized in Table 1.

TABLE 1. CRYSTAL DATA

$C_{13}NO_5H_{13}$	Space group: $P2_12_12_1$
$M.W.$: 263	$Z=4$
Orthorhombic	$D_x=1.40\ \mathrm{g\cdot cm^{-3}}$
$a=7.766(1)$ Å	$D_m=1.38\ \mathrm{g\cdot cm^{-3}}$
$b=27.470(4)$	$\mu(\mathrm{Mo}\ K\alpha)=1.30\ \mathrm{cm^{-1}}$
$c=5.948(1)$	
$U=1268.9(4)$ Å^3	

With use of a 0.3×0.4×0.4 mm crystal, intensity data up to $2\theta=55°$ were collected on a Rigaku four-circle diffractometer with graphite monochromated Mo $K\alpha$ radiation. The usual ω-2θ scan technique was employed with a scan speed 4° $(2\theta)\mathrm{min}^{-1}$. A total of 1735 independent reflections were obtained, in which 231 with $|F_o|<3\sigma(|F_o|)$ were regarded as zero reflections. The usual Lorentz and polarization corrections were made. Absorption was uncorrected.

Structure Determination

The structure was determined by the direct method using the program MULTAN. Refinement of the structure was carried out by the block-diagonal least-squares method. After several cycles of refinement, two strong reflections, 020 and 060, were excluded since they seem to undergo secondary extinction. All the hydrogen atoms were revealed on a difference map. The final refinement was performed including these hydrogen atoms with isotropic temperature factors. The weighting scheme: $w=0.25$ for $|F_o|<1.95$ and $|F_o|>15.62$, and $w=(5.959-1.127|F_o|+0.0642|F_o|^2)^{-1}$ for the other reflections. The final R value became 0.057 for all the non-zero reflections. In the final stage, no peaks higher than 0.2 eÅ^{-3} were found on the difference map. The atomic scattering factors were taken from "International Tables for X-Ray Crystallography."[5] The final atomic parameters and their standard devia-

TABLE 2. FINAL ATOMIC COORDINATES AND THERMAL PARAMETERS FOR γ-BENZYL-L-GLUTAMATE NCA WITH THEIR ESTIMATED STANDARD DEVIATIONS

Anisotropic thermal parameters ($\times 10^5$) are in the form: $\exp\{-(B_{11}h^2+B_{22}k^2+B_{33}l^2+B_{12}hk+B_{13}hl+B_{23}kl)\}$.

Atom	x	y	z	B_{11}	B_{22}	B_{33}	B_{12}	B_{13}	B_{23}
C(1)	0.4302(4)	0.2372(1)	0.1049(5)	1443(55)	126(4)	1930(84)	1(3)	89(127)	163(32)
C(2)	0.2635(3)	0.2530(1)	0.4070(5)	1207(47)	82(3)	1988(77)	46(2)	−230(116)	106(31)
C(3)	0.2523(3)	0.1989(1)	0.3630(4)	1057(46)	83(3)	1740(75)	−53(2)	−107(112)	6(28)
C(4)	0.3077(4)	0.1702(1)	0.5698(5)	1448(54)	84(3)	1776(76)	6(2)	−136(126)	125(29)
C(5)	0.3144(4)	0.1153(1)	0.5284(5)	1789(61)	86(4)	2213(88)	−8(3)	−440(138)	56(31)
C(6)	0.3813(4)	0.0888(1)	0.7299(5)	1572(54)	84(3)	2535(91)	3(2)	186(142)	155(32)
C(7)	0.4930(5)	0.0149(1)	0.8678(6)	2589(82)	119(5)	2505(106)	294(3)	25(171)	175(37)
C(8)	0.5531(4)	−0.0340(1)	0.7847(6)	1664(59)	81(4)	2864(104)	−23(2)	95(146)	287(33)
C(9)	0.6522(5)	−0.0614(1)	0.9291(7)	2319(80)	113(4)	3998(134)	54(3)	−1689(204)	40(44)
C(10)	0.7050(5)	−0.1078(1)	0.8708(7)	2399(80)	112(4)	5278(170)	274(3)	−2123(224)	180(48)
C(11)	0.6611(5)	−0.1272(1)	0.6674(7)	2251(81)	96(4)	4855(152)	135(3)	151(212)	−67(44)
C(12)	0.5623(5)	−0.1005(1)	0.5230(7)	3164(98)	132(5)	3573(124)	192(4)	−480(214)	230(45)
C(13)	0.5080(5)	−0.0538(1)	0.5815(6)	2871(87)	126(5)	3110(119)	235(3)	−1051(196)	20(42)
N(1)	0.3655(3)	0.1945(1)	0.1706(4)	1637(46)	91(3)	1565(61)	40(2)	311(109)	−124(25)
O(1)	0.5216(3)	0.2473(1)	−0.0494(4)	2246(51)	195(4)	2761(81)	−103(3)	1850(117)	353(33)
O(2)	0.3727(3)	0.2735(1)	0.2533(3)	1439(35)	87(2)	2493(61)	−57(2)	275(91)	156(23)
O(3)	0.1968(3)	0.2765(1)	0.5499(4)	2040(46)	98(3)	2639(67)	123(2)	855(106)	−179(26)
O(4)	0.3992(4)	0.1053(1)	0.9147(4)	3690(66)	102(3)	2148(64)	362(2)	−556(132)	−61(24)
O(5)	0.4221(3)	0.0426(1)	0.6818(4)	2550(54)	83(3)	2450(64)	208(2)	−606(116)	70(23)

TABLE 3. FINAL ATOMIC COORDINATES AND THERMAL PARAMETERS B(Å^2) FOR THE H ATOMS WITH THEIR ESTIMATED STANDARD DEVIATIONS

Atom	x	y	z	B
H(1)	0.382(4)	0.166(1)	0.097(5)	4.6(7)
H(2)	0.138(3)	0.192(1)	0.329(4)	2.6(5)
H(3)	0.418(4)	0.182(3)	0.616(4)	3.0(6)
H(4)	0.231(4)	0.177(1)	0.694(5)	4.5(7)
H(5)	0.200(4)	0.104(1)	0.499(6)	6.2(9)
H(6)	0.399(4)	0.110(1)	0.408(6)	6.2(9)
H(7)	0.398(4)	0.011(1)	0.968(5)	4.5(7)
H(8)	0.589(4)	0.034(1)	0.938(6)	5.6(8)
H(9)	0.688(4)	−0.046(1)	1.078(6)	5.5(8)
H(10)	0.783(4)	−0.124(1)	0.964(6)	7.4(10)
H(11)	0.702(5)	−0.161(1)	0.614(6)	7.3(10)
H(12)	0.541(4)	−0.112(1)	0.376(6)	7.5(10)
H(13)	0.451(4)	−0.038(1)	0.473(7)	6.7(9)

tions for non-hydrogen atoms and hydrogen atoms are given in Tables 2 and 3, respectively. A list of the observed and calculated structure factors is kept in the office of the Chemical Society of Japan (Document No. 7821). Computation was carried out on the HITAC 8700 computer at this Institute.

Description of the Structure

Bond distances and angles are shown in Fig. 1, together with the numbering system. The dimensions involving hydrogen atoms are given in Table 4. There is a significant difference between the distances of C(1)–O(2) and C(2)–O(2) in the *N*-carboxy anhydride ring (NCA ring). This feature is similar to that of L-alanine NCA[4] but differs from that of glycine NCA[3] and L-leucine NCA,[6] in which C(1)–O(2) has nearly the same distance as C(2)–O(2). The other bond distances, C(2)–C(3), C(1)–N(1), and C(3)–N(1), are comparable to the corresponding lengths in glycine NCA, L-alanine NCA, and L-leucine NCA. The characteristic dimensions in the NCA ring of the present compound are discussed in detail and compared with those of the other amino acid NCAs. The fluctuation of the distances in the benzene ring can be explained by the thermal vibration around C(8)–C(11).

TABLE 4. BOND LENGTHS AND ANGLES INVOLVING HYDROGEN ATOMS IN γ-BENZYL-L-GLUTAMATE NCA

The estimated standard deviations are 0.03—0.04 Å in distances and 2—3° in angles.

Bond lengths (l/Å)			
N(1)–H(1)	0.91	C(7)–H(8)	1.00
C(3)–H(2)	0.93	C(9)–H(9)	1.02
C(4)–H(3)	0.96	C(10)–H(10)	0.93
C(4)–H(4)	0.97	C(11)–H(11)	1.03
C(5)–H(5)	0.96	C(12)–H(12)	0.95
C(5)–H(6)	0.99	C(13)–H(13)	0.90
C(7)–H(7)	0.95		

Bond angles (ϕ/°)			
C(1)–N(1)–H(1)	124	C(4)–C(5)–H(5)	109
C(3)–N(1)–H(1)	123	C(6)–C(5)–H(6)	106
C(2)–C(3)–H(2)	107	C(8)–C(7)–H(7)	111
N(1)–C(3)–H(2)	113	O(5)–C(7)–H(8)	109
C(3)–C(4)–H(4)	110	C(11)–C(12)–H(12)	120
H(3)–C(4)–H(4)	106	C(10)–C(11)–H(11)	123
C(4)–C(5)–H(6)	107	C(8)–C(13)–H(13)	124
H(5)–C(5)–H(6)	116	C(8)–C(7)–H(8)	112
C(3)–C(4)–H(3)	108	H(7)–C(7)–H(8)	112
C(5)–C(4)–H(3)	111	C(10)–C(9)–H(9)	122
C(6)–C(5)–H(5)	108	C(11)–C(10)–H(10)	120
O(5)–C(7)–H(7)	104	C(12)–C(11)–H(11)	117
C(5)–C(4)–H(4)	109	C(12)–C(13)–H(13)	115
		C(13)–C(12)–H(12)	120

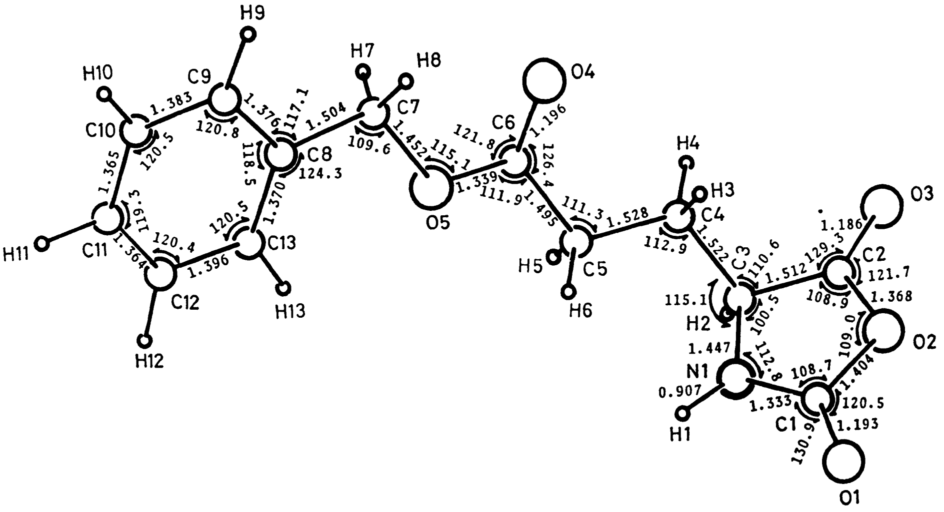


Fig. 1. A perspective view of γ-benzyl-L-glutamate NCA along the c axis with the numbering of the atoms, bond lengths(l/Å), and angles(ϕ/°). The estimated standard deviations are 0.004—0.006Å in distances and 0.2—0.4° in angles.

The equations of the least-squares planes, plane I through the NCA ring, plane II through the carbonyl group and plane III through the benzene ring, are given in Table 5, together with the deviations of the atoms from the planes. Three oxygen atoms, O(1), O(2), and O(3) deviate distinctly from plane I. This is also the case in the other amino acid NCA crystals. Plane I makes angles of 56.8 and 69.7° with planes II and III, respectively. Plane II makes an angle of 13.0° with plane III.

TABLE 5. EQUATION OF THE LEAST-SQUARES PLANES AND THE DEVIATIONS OF ATOMS FROM THE PLANES

Coordinates X, Y, Z are in Å referred to $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$.

Atom	Deviation
Plane I: $-0.7834X+0.1448Y-0.6044Z+2.0546=0$	
Atom	Deviation
C(1)	0.005
C(2)	−0.004
C(3)	0.007
N(1)	−0.008
*O(1)	0.044
*O(2)	−0.035
*O(3)	−0.019
*H(1)	0.043
*C(4)	−1.188
Plane II: $-0.9305X-0.2835Y+0.2322Z+2.4389=0$	
Atom	Deviation
C(5)	−0.000
C(6)	0.001
O(4)	−0.001
O(5)	−0.000
Plane III: $-0.8289X-0.3821Y+0.4087Z+1.2926=0$	
Atom	Deviation
C(8)	−0.002
C(9)	−0.001
C(10)	0.004
C(11)	−0.004
C(12)	0.001
C(13)	0.002
*C(7)	0.074

*Atoms not included in the least-squares caluculations.

Conformations of the side chain by Newman projections are shown in Fig. 2. Both of the conformations about C(4)–C(3) and C(5)–C(4) are trans.

A stereoscopic view of the crystal structure is shown in Fig. 3. The hydrogen bond (broken lines) is formed between N(1)–H(1)···O(4′). The length of N(1)···O(4′)

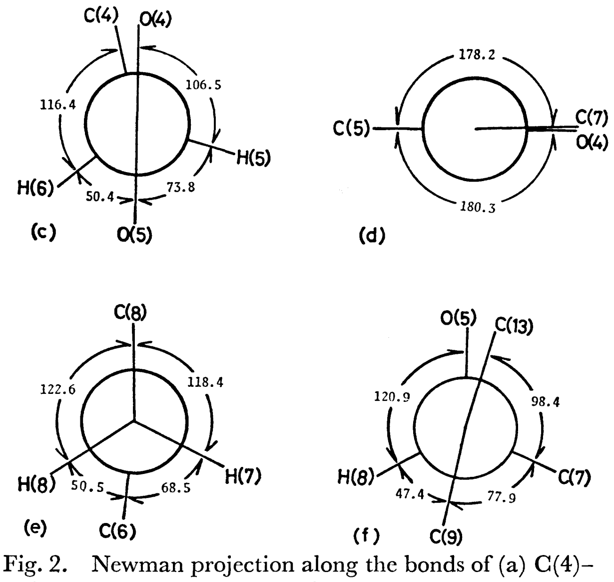


Fig. 2. Newman projection along the bonds of (a) C(4)–C(3), (b) C(5)–C(4), (c) C(6)–C(5), (d) O(5)–C(6), (e) C(7)–O(5), and (f) C(8)–C(7) in the side chain.

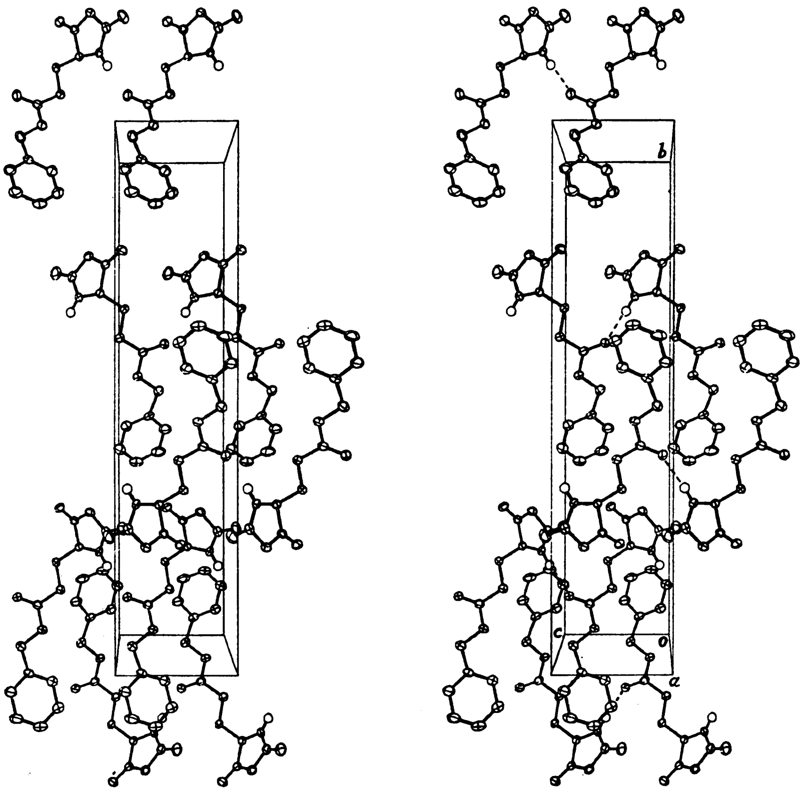


Fig. 3. Stereoscopic drawing of the crstal structure of γ-benzyl-L-glutamate NCA viewed along the a axis. The thermal ellipsoids are drawn at 50% probability level. Hydrogen atoms are omitted for clarity except for H(1) with circles. Hydrogen bond are shown by broken lines.

and the angle of N(1)–H(1)···O(4′) are 2.898(4) Å and 174(3)°, respectively. This type of hydrogen bond is in marked contrast to that in the crystals of the other amino acid NCAs, in which the oxygen atom of either carbonyl group in the NCA ring participates in the hydrogen bond. There is no unusually short contact among the other intermolecular distances.

Discussion

The bond distances in the NCA rings of four amino acid NCAs are given in Table 6. A significant difference in the distances C(1)–O(2) and C(2)–O(2) is seen among these compounds. In glycine NCA and L-leucine NCA, the distances of two bonds are similar to each other. On the other hand, C(1)–O(2) is significantly greater than C(2)–O(2) in L-alanine NCA and BLG NCA. The difference in L-alanine NCA is more distinct than that in BLG NCA. It is plausible that the difference is caused by the hydrogen bond in the crystalline state.

The hydrogen bonds of these amino acid NCA crystals are divided into three types:

(1) N(1)–H(1)···O(1′) as observed in glycine NCA and L-leucine NCA,

(2) N(1)–H(1)···O(3′) in L-alanine NCA,

(3) N(1)–H(1)···O(4′) in BLG NCA.

Since no oxygen atoms of carbonyl groups in the NCA ring form the hydrogen bond with N(1)–H(1) in BLG NCA, it is reasonable to take the dimensions of this NCA ring as the standard ones.

It was pointed out by Leiserowitz and Nader⁷⁾ that the oxygen atom of the amide group is more negative than that of the carbonyl group in the amide-dicarboxylic acid complexes. This implies that the limiting structure (II), shown in Fig. 4, might contribute to the resonance state in the NCA ring. However, in BLG NCA, C(2)–O(2) is shorter than C(1)–O(2), C(1)–O(1) having the same distance as C(2)–O(3). This suggests that the limiting structures (II) and (IV), which is given in a simplified form by (V), should contribute to the state in the NCA moiety.

Table 6. Comparison of the bond distances (Å) in the four NCAs
Estimated standard deviations are in parentheses.

NCA	C(1)–O(2)	C(2)–O(2)	C(2)–C(3)	C(1)–N(1)	C(3)–N(1)	C(1)–O(1)	C(2)–O(3)
Glycine	1.394(2)	1.381(2)	1.502(3)	1.326(3)	1.449(3)	1.212(3)	1.188(2)
L-Alanine	1.412(5)	1.354(4)	1.514(5)	1.327(5)	1.449(5)	1.203(5)	1.188(5)
L-Leucine	1.394(5)	1.380(5)	1.490(5)	1.319(5)	1.456(5)	1.199(5)	1.194(5)
BLG	1.404(4)	1.368(4)	1.512(4)	1.333(4)	1.447(4)	1.193(4)	1.186(4)

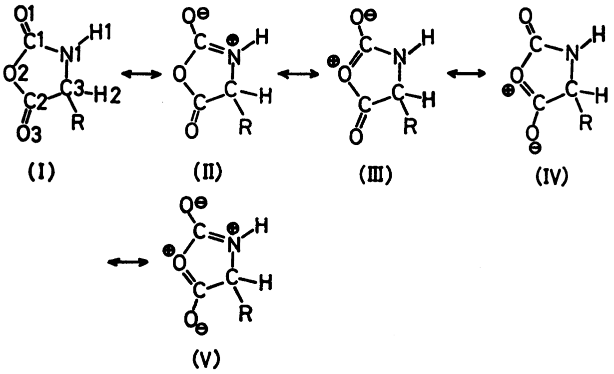


Fig. 4. The limiting canonical structures.

When O(1) forms the hydrogen bond with N(1)–H(1) as observed in glycine NCA and L-leucine NCA, the limiting structure (III) could contribute to the resonance state, on the assumption that the negative charge is attracted to the N–H···O hydrogen bond. In the glycine NCA and L-leucine NCA crystals, C(1)–O(2) is shorter and C(2)–O(2) is longer than the corresponding ones in BLG NCA. When O(3) forms the hydrogen bond with N(1)–H(1), the limiting structure (IV) would contribute for the same reason. In L-alanine NCA, C(1)–O(2) is longer and C(2)–O(2) shorter than the corresponding bonds in BLG NCA.

As for the C–O distances of carbonyl group and the N–C length, the difference is made obscure owing to the contribution of the limiting structure (II), except for the C–O distances in glycine NCA.

In the course of the polymerization process, the bonds of C(2)–O(2) and C(1)–N(1) are cleaved to release CO_2. Since the limiting structure (III) contributes to the resonance state in glycine NCA and L-leucine NCA, it is quite adequate to consider that the C(2)–O(2) bond of these amino acid NCAs should be more easily cleaved than that of the other amino acid NCAs. Contrary to such expectation, glycine NCA shows very low reactivity in the crystalline state, although the crystal of L-leucine NCA is observed to have the highest polymerizability.[2] This implies that the other factors should be taken into account. In the succeeding paper,[6] the effect of the crystalline field will be discussed.

References

1) C. H. Bamford, A. Elliot, and W. H. Hanby, "Synthetic Polypeptides," Academic Press, New York (1956), p. 62.

2) H. Kanazawa and T. Kawai, to be published.

3) H. Kanazawa, Y. Matsuura, N. Tanaka, M. Kakudo, T. Komoto, and T. Kawai, *Bull. Chem. Soc. Jpn.*, **49**, 954 (1976).

4) H. Kanazawa, Y. Matsuura, N. Tanaka, M. Kakudo, T. Komoto, and T. Kawai, *Acta Crystallogr., Sect. B*, **32**, 78 (1976).

5) "International Tables for X-Ray Crystallogaphy," Vol. IV, The Kynoch Press, Birmingham (1974), p. 72.

6) H. Kanazawa, T. Kawai, Y. Ohashi, and Y. Sasada, *Bull. Chem. Soc. Jpn.*, **51**, 2208 (1978).

7) L. Leiserowitz and F. Nader, *Acta Crystallogr., Sect. B*, **32**, 2719 (1977).