The Crystal and Molecular Structure of N-Acetylactinobolin: the α-Helix in a Small Peptide

By Robert B. Von Dreele

*Department of Chemistry, Arizona State University, Tempe, Arizona* 85281, *U.S.A.*

A crystal structure analysis of N-acetylactinobolin has revealed a molecular conformation which includes an N-acetyl-L-alanyl side chain in the α-helix conformation ($\varphi = -83\cdot6°$, $\psi = -24\cdot1°$) and is controlled by molecular packing and intermolecular hydrogen bonding. The conformation of the isocoumarin ring system is in accord with the interpretation of NMR studies in both water and dimethyl sulfoxide. The structure was solved by direct methods of analysis of data from a crystal with space group $P2_12_12_1$ and $a = 8\cdot708$ (1), $b = 8\cdot851$ (1), $c = 21\cdot718$ (4) Å, $Z = 4$ and density $\varrho_{calc} = 1\cdot364$ g cm^{-3}. Anisotropic least-squares refinement converged to a conventional residual of $R = 0\cdot064$ for 2427 independent observed reflections recorded with Mo $K\alpha$ radiation on an automatic four-circle diffractometer.

Introduction

The broad-spectrum antibiotic actinobolin (I) was first isolated as the sulfate salt (Haskell & Bartz, 1959). Its structure was established from studies of its acid hydrolysis and the base hydrolysis of N-acetylactinobolin (II) (Munk, Nelson, Antosz, Herald & Haskell, 1968; Antosz, Nelson, Herald & Munk, 1970). More recently the structure of the hydroiodide salt of actinobolin was reported (Wetherington & Moncrief, 1975). This extremely complex structure with four molecules in each asymmetric unit confirmed the chemical studies and displayed an extensive hydrogen-bonding scheme. Since it was clear from the cell parameters that N-acetylactinobolin has a simpler structure than actinobolin.HI, the present study was undertaken with the expectation of finding a different hydrogen-bonding scheme with possible concomitant changes in the molecular conformation. Moreover, a large number of high, quality diffraction data was available which provided the opportunity to establish the molecular parameters of this important antibiotic with great precision.

(I)   (II)

Experimental

Excellent crystals of N-acetylactinobolin* were grown by slow cooling of an ethanol–water solution. Preces-

---

* The author is indebted to M. E. Munk for a sample of this material.

sion photographs displayed orthorhombic symmetry, *mmm*, with the systematic absences $h = 2n + 1$ for $h00$, $k = 2n + 1$ for $0k0$, $l = 2n + 1$ for $00l$, which uniquely correspond to the space group $P2_12_12_1$. A total of 15 reflections in the angular range $10\cdot8 \le 2\theta \le 22°$ for Cu $K\alpha$ radiation was automatically centered on a Syntex P$\bar{1}$ autodiffractometer; a least-squares refinement of the angular settings yielded the lattice parameters $a = 8\cdot708 \pm 0\cdot001$, $b = 8\cdot851 \pm 0\cdot001$, $c = 21\cdot718 \pm 0\cdot004$ Å, which for $Z = 4$ gives $\varrho_{calc} = 1\cdot364$ g cm^{-3} ($\varrho_{obs} = 1\cdot356$ g cm^{-3}).

The diffraction intensities were measured on a $0\cdot18 \times 0\cdot30 \times 0\cdot40$ mm crystal using Zr-filtered Mo $K\alpha$ radiation with the diffractometer operating in the variable speed θ–2θ scan mode. For each reflection the scan speed, between 1 and 12° min^{-1}, was determined from the intensity found in a rapid sampling scan. The scans were taken over the range $2\theta K\alpha_1 - 0\cdot8°$ to $2\theta K\alpha_2 + 0\cdot8°$ with background counts for $0\cdot5$ of the scan time taken at each end of the scan. A total of 2802 independent reflections were investigated ($2\theta \le 60\cdot0°$) with the Mo tube operating at 20 mA and 50 kV. The low-angle data ($2\theta \le 22°$) were reinvestigated at 10 mA and 50 kV to obtain more reliable intensities for the strongest reflections. After applying a least-squares scaling procedure (Hamilton, Rollett & Sparks, 1965), a total of 2427 reflections was retained as objectively observed with $|F_o| > 0\cdot675\sigma_F$; $\sigma_F = 0\cdot02|F_o| + (C + Bk^2)^{1/2}R/(2|F_o|Lp)$, where C is the total count in a scan taken at the rate R, and k is the ratio of scanning time to the time for the total background count B. Periodic monitoring of three reflections showed a maximum 3% random variation in intensity during the time of data collection. Corrections were applied for Lorentz and polarization effects but absorption and extinction effects proved to be negligible. After a number of attempts to solve the structure with this data set, a second data set was collected in the same fashion with Ni-filtered Cu $K\alpha$ radiation. A total of 1826 independent reflections was investigated ($2\theta \le 140°$) within the limits of the diffractometer. The intensity of the 14
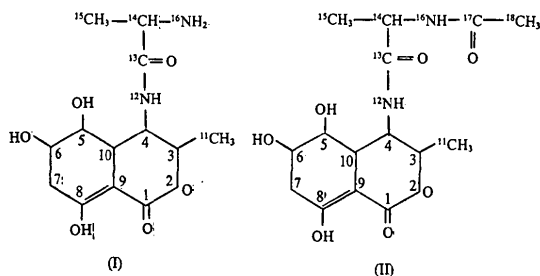
strongest reflections exceeded the capacity of the counting system and these were left out of the data set. After corrections for Lorentz and polarization effects this data set was successfully used in the direct-methods analysis. The data set collected with Mo $K\alpha$ radiation was used in the refinement of the structure.

Structure determination and refinement

The structure of N-acetylactinobolin was readily solved when the $MULTAN74$ system of computer programs (Germain, Main & Woolfson, 1971; Declercq, Germain, Main & Woolfson, 1973; Koch, 1974) was applied to the data set collected with Cu $K\alpha$ radiation. A set of normalized structure factors, E_{hkl}, was obtained by means of a Debye curve calculated from the molecular scattering factors of the substituted isocoumarin ring system found in actinobolin and the two peptide groups in the side chain. A total of 250 reflections with $E_{hkl} \geq 1\cdot45$ was expanded over 2000 $\sum_2$ interactions. These were subjected to a convergence analysis to give the starting set 0014 (π), 860 (π), 701 $(\pi/2)$, 704 (0), 093 $(\pi/2)$, 460 $(0$ or $\pi)$, 377 $(\pm \pi/4)$ and 364 $(\pm \pi/4$ or $\pm 3\pi/4)$. The first two phases were determined by $\sum_1$ relationships, the next three fix the origin (Hauptman & Karle, 1956) and the last three are variable with the 377 determining the enantiomorph. A multiple-solution tangent refinement of the 16 possible starting sets gave the highest absolute figure of merit of 1·069 and the lowest $\psi(0) = 279$ (Cochran & Douglas, 1955) for the best solution. A Fourier synthesis of these phases revealed the positions of 21 atoms. These positions were used to phase $(R = 37\%)$ a difference

Fourier synthesis* which revealed the positions of the remaining three atoms. A similar sequence of calculations using the more complete Mo $K\alpha$ data set did not sharply distinguish between enantiomorphs and gave no clear indication of the true position of the side chain. The presence of several $00l$ reflections with large E values in the Mo $K\alpha$ data set may have magnified the 'chicken wire' effect caused by the parallel stacking along c of the isocoumarin ring systems, thus obscuring the correct structure in the E syntheses. These reflections were among those left out of the Cu $K\alpha$ data set because of their intensity.

The model was refined with isotropic thermal parameters by full-matrix least-squares analysis using the data set collected with Mo $K\alpha$ radiation restricted to $\sin \theta/\lambda \leq 0\cdot5$ with each reflection assigned a weight, $w = 1/\sigma_F^2$, and with atomic scattering factors for C°, N°, O°, and H° (International Tables for X-ray Crystallography, 1974). At convergence the standard residual was $R = 0\cdot125$ and the weighted residual, $R_w = [\sum w(|F_o| - |F_c|)^2/\sum w|F_o|^2]^{1/2}$, was $0\cdot160$. After correcting an inadvertent switch between an oxygen atom and a carbon atom the model with anisotropic temperature factors was refined by full-matrix least squares (217 parameters) on all the observed data to obtain the residuals $R = 0\cdot102$ and $R_w = 0\cdot106$ at convergence. A difference Fourier synthesis based on these results gave the positions of all the hydrogen atoms. The C–H atoms were placed at ideal positions and all other hydrogen

_____

* From this point on all calculations were performed with the '$CRYSTALS$' system of computer programs adapted for the Univac 1110 (Rollett & Carruthers, 1974).

Table 1. Fractional coordinates ($\times 10^4$) and thermal parameters ($\times 10^4$) for N-acetylactinobolin

The thermal parameters are expressed in the form $T = \exp[-2\pi^2(U_{11}h^2a^{*2} + \ldots + 2U_{23}klb^*c^*)]$, where the U_{ij} values are in Å^2.
The estimated standard deviations are listed in parentheses.

	x	y	z	U_{11}	U_{22}	U_{33}	U_{12}	U_{13}	U_{23}
C(1)	4376 (3)	6082 (4)	3245 (1)	245 (14)	441 (19)	365 (17)	−26 (14)	−6 (13)	−57 (15)
O(1)	3135 (2)	6767 (3)	3192 (1)	229 (10)	599 (16)	597 (15)	37 (12)	−34 (11)	−24 (14)
O(2)	4301 (2)	4611 (3)	3374 (1)	292 (11)	461 (13)	579 (15)	−113 (12)	−8 (11)	30 (12)
C(3)	5676 (4)	3661 (4)	3464 (1)	365 (16)	320 (16)	412 (18)	−43 (15)	15 (14)	−54 (14)
C(4)	7070 (3)	4601 (3)	3658 (1)	283 (14)	275 (14)	312 (15)	35 (13)	21 (12)	−4 (12)
C(5)	8745 (3)	6787 (3)	3269 (1)	233 (13)	363 (16)	275 (14)	−28 (13)	−11 (11)	−28 (13)
O(5)	9997 (2)	5817 (3)	3120 (1)	207 (10)	508 (15)	582 (15)	7 (11)	20 (10)	−63 (13)
C(6)	8686 (3)	8174 (4)	2863 (1)	284 (15)	442 (18)	361 (16)	−115 (15)	−19 (13)	67 (15)
O(6)	10120 (3)	8921 (3)	2902 (2)	370 (13)	594 (16)	653 (17)	−220 (13)	−31 (12)	193 (14)
C(7)	7381 (4)	9195 (4)	3076 (1)	442 (19)	341 (18)	524 (22)	−59 (16)	−81 (16)	81 (16)
C(8)	5897 (3)	8341 (4)	3123 (1)	299 (15)	417 (18)	330 (16)	54 (15)	−60 (13)	2 (14)
O(8)	4647 (2)	9233 (3)	3117 (1)	391 (13)	447 (15)	616 (16)	124 (12)	−64 (12)	−13 (13)
C(9)	5856 (3)	6817 (3)	3175 (1)	248 (13)	348 (16)	276 (14)	16 (14)	17 (12)	−20 (13)
C(10)	7289 (3)	5862 (3)	3182 (1)	222 (12)	326 (16)	250 (13)	−20 (12)	2 (11)	−6 (13)
C(11)	5196 (4)	2471 (4)	3913 (2)	568 (21)	374 (19)	599 (23)	−118 (20)	92 (19)	55 (18)
N(12)	6923 (3)	5210 (3)	4280 (1)	290 (12)	329 (13)	300 (12)	43 (12)	20 (10)	−27 (11)
C(13)	7663 (3)	4571 (4)	4756 (1)	328 (16)	361 (17)	347 (16)	2 (15)	12 (13)	31 (14)
O(13)	8519 (3)	3494 (3)	4711 (1)	592 (16)	626 (17)	483 (14)	316 (15)	−34 (13)	46 (13)
C(14)	7318 (4)	5253 (4)	5395 (1)	358 (16)	410 (18)	375 (17)	−26 (16)	−36 (14)	0 (15)
C(15)	6140 (4)	4255 (4)	5724 (2)	672 (25)	642 (25)	438 (19)	−93 (24)	185 (19)	25 (19)
N(16)	6786 (3)	6799 (3)	5380 (1)	239 (12)	454 (16)	452 (15)	7 (13)	−39 (11)	−75 (14)
C(17)	7734 (4)	7987 (4)	5348 (1)	345 (16)	456 (19)	372 (17)	9 (16)	−13 (14)	−89 (15)
O(17)	9144 (2)	7830 (3)	5366 (1)	269 (11)	564 (16)	726 (17)	−49 (12)	59 (12)	−60 (14)
C(18)	6989 (4)	9509 (4)	5302 (2)	493 (22)	462 (22)	893 (30)	9 (20)	−29 (22)	−60 (22)

atoms were placed as found in the map. The parameters for the non-hydrogen atoms were again refined by full-matrix least squares to yield $R = 0.068$ and $R_w = 0.059$ at convergence. All hydrogen atoms were included with fixed parameters and a fixed value of $U_{iso} = 0.05$ Å². The positions of the hydrogen atoms were then refined in two cycles of least squares and the refinement of the non-hydrogen parameters was repeated to give the final residuals $R = 0.064$ and $R_w = 0.052$.*

Results

Final atomic coordinates and thermal parameters for N-acetylactinobolin are presented in Tables 1 and 2 along with the estimated standard deviations derived from the least-squares analysis. The stereoscopic view shown in Fig. 1 displays the essential configurational and conformational features of the molecule. Each non-hydrogen atom is represented by an ellipsoid consistent with the anisotropic thermal parameters in Table 1. The six chiral centers are (R)-C(3), (R)-C(4), (R)-C(5), (R)-C(6), (R)-C(10) and (S)-C(14). Since the absolute configuration of actinobolin has been determined (Antosz, Nelson, Herald & Munk, 1970; Wetherington & Moncrief, 1975) the molecule is shown in the correct enantiomorphous form.

Table 2. *Fractional coordinates* ($\times 10^3$) *and bond distances* (Å) *for the hydrogen atoms*

Values in parentheses are estimated standard deviations in the last figure

	x	y	z	d
H(C3)	589 (3)	321 (3)	304 (1)	1·02 (3)
H(C4)	792 (3)	392 (3)	365 (1)	0·95 (3)
H(C5)	884 (3)	716 (3)	370 (1)	1·00 (3)
H(C6)	849 (3)	781 (3)	244 (1)	0·98 (3)
H(C7)$_{ax}$	758 (3)	967 (3)	344 (1)	0·91 (3)
H(C7)$_{eq}$	725 (3)	1003 (3)	278 (1)	0·99 (3)
H(C10)	733 (3)	534 (3)	279 (1)	0·96 (3)
H(O5)	1080 (3)	632 (3)	318 (1)	0·84 (3)
H(O6)	1015 (3)	959 (3)	265 (1)	0·81 (3)
H(O8)	380 (3)	860 (3)	314 (1)	0·93 (3)
H(C11)	428 (3)	189 (3)	376 (1)	1·01 (3)
H(C11)	481 (3)	287 (3)	432 (1)	1·01 (3)
H(C11)	592 (3)	174 (3)	396 (1)	0·91 (3)
H(N12)	622 (3)	595 (3)	435 (1)	0·91 (3)
H(C14)	836 (3)	527 (3)	565 (1)	1·06 (3)
H(C15)	649 (3)	322 (3)	576 (1)	0·97 (3)
H(C15)	602 (3)	466 (3)	614 (1)	0·97 (3)
H(C15)	505 (3)	430 (3)	552 (1)	1·05 (3)
H(N16)	585 (3)	697 (3)	532 (1)	0·84 (3)
H(C18)	588 (3)	946 (3)	531 (1)	0·97 (3)
H(C18)	725 (3)	1021 (3)	567 (1)	1·04 (3)
H(C18)	737 (3)	1004 (3)	495 (1)	0·96 (3)

Table 3. *Bond distances and angles for N-acetylactinobolin*

Values in parentheses are estimated standard deviations in the last figure

C(1)–O(1)	1·245 (3) Å	O(1)—C(1)—O(2)	116·9 (6)°
C(1)–O(2)	1·333 (4)	O(1)—C(1)—C(9)	122·8 (5)
C(1)–C(9)	1·451 (4)	O(2)—C(1)—C(9)	120·2 (5)
O(2)–C(3)	1·476 (4)	C(1)—O(2)—C(3)	123·0 (4)
C(3)–C(4)	1·530 (4)	O(3)—C(3)—C(4)	111·7 (5)
C(3)–C(11)	1·495 (5)	O(3)—C(3)—C(11)	105·1 (6)
		C(4)—C(3)—C(11)	115·2 (6)
C(4)–C(10)	1·533 (4)	C(3)—C(4)—C(10)	108·0 (5)
C(4)–N(12)	1·460 (3)	C(3)—C(4)—N(12)	112·6 (5)
		C(10)-C(4)—N(12)	111·4 (5)
C(5)–O(5)	1·425 (3)	O(5)—C(5)—C(6)	112·5 (5)
C(5)–C(6)	1·512 (4)	O(5)—C(5)—C(10)	106·6 (5)
C(5)–C(10)	1·521 (4)	C(6)—C(5)—C(10)	109·6 (5)
C(6)–O(6)	1·416 (3)	C(5)—C(6)—O(6)	108·4 (5)
C(6)–C(7)	1·523 (4)	C(5)—C(6)—C(7)	109·3 (6)
		O(6)—C(6)—C(7)	111·3 (6)
C(7)–C(8)	1·501 (4)	C(6)—C(7)—C(8)	111·4 (6)
C(8)–O(8)	1·345 (3)	C(7)—C(8)—O(8)	113·6 (5)
C(8)–C(9)	1·354 (4)	C(7)—C(8)—C(9)	122·0 (5)
		O(8)—C(8)—C(9)	124·4 (5)
C(9)–C(10)	1·507 (4)	C(1)—C(9)—C(8)	118·6 (5)
		C(1)—C(9)—C(10)	118·8 (5)
		C(8)—C(9)—C(10)	122·5 (5)
		C(4)—C(10)-C(5)	114·3 (5)
		C(4)—C(10)-C(9)	108·2 (5)
		C(5)—C(10)-C(9)	113·0 (5)
N(12)–C(13)	1·345 (4)	C(4)—N(12)—C(13)	121·0 (5)
C(13)–O(13)	1·214 (4)	N(12)-C(13)—O(13)	124·2 (5)
C(13)–C(14)	1·542 (4)	N(12)-C(13)—C(14)	115·8 (5)
		O(13)-C(13)—C(14)	120·1 (5)
C(14)–C(15)	1·530 (4)	C(13)—C(14)—C(15)	108·9 (6)
C(14)–N(16)	1·445 (4)	C(13)—C(14)—N(16)	114·4 (5)
		C(15)—C(14)—N(16)	110·0 (6)
N(16)–C(17)	1·339 (4)	C(14)—N(16)—C(17)	123·2 (5)
C(17)–O(17)	1·237 (3)	N(16)-C(17)—O(17)	121·5 (6)
C(17)–C(18)	1·499 (5)	N(16)-C(17)—C(18)	116·3 (6)
		O(17)-C(17)—C(18)	122·2 (6)

Table 4. *Selected torsion angles* (°) *for N-acetylactinobolin*

		IUPAC designation
C(1)O(2)C(3)C(4)	23·8	–
O(2)C(3)C(4)C(10)	– 54·1	–
C(3)C(4)C(10)C(9)	60·0	–
C(4)C(10)C(9)C(1)	– 37·8	–
C(10)C(9)C(1)O(2)	6·7	–
C(9)C(1)O(2)C(3)	1·0	–
C(10)C(5)C(6)C(7)	65·0	–
C(5)C(6)C(7)C(8)	– 52·1	–
C(6)C(7)C(8)C(9)	20·5	–
C(7)C(8)C(9)C(10)	– 0·7	–
C(8)C(9)C(10)C(5)	12·8	–
C(9)C(10)C(5)C(6)	– 44·3	–
C(3)C(4)N(12)C(13)	102·0	–
C(10)C(4)N(12)C(13)	– 136·4	–
C(4)N(12)C(13)C(14)	– 175·7	ω_1
N(12)C(13)C(14)N(16)	– 24·1	ψ
C(13)C(14)N(16)C(17)	– 83·6	ϕ
C(14)N(16)C(17)C(18)	176·8	ω_2

Bond lengths and angles within the molecule are systematically recorded in Table 3. The set of torsion angles listed in Table 4 fully characterizes all the conformational features of the molecule. A stereoscopic drawing of the unit-cell contents is shown in Fig. 2.

Discussion

In the crystal structure of N-acetylactinobolin, intermolecular hydrogen bonds between the isocoumarin ring systems hold the molecules in sheets with the
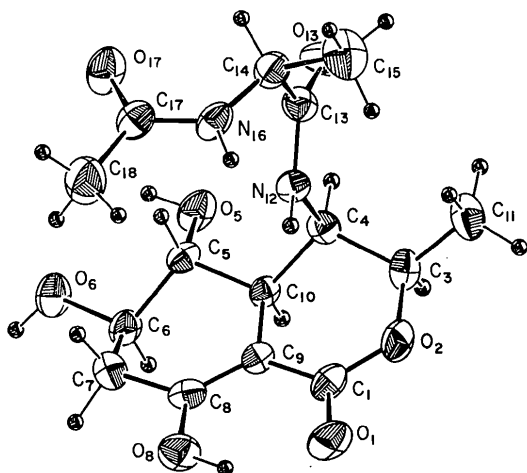

Fig. 1. A perspective representation of the structure of N-acetylactinobolin.

N-acetyl-L-alanyl side chains extending on both sides of each sheet. The sheets are packed together with concomitant folding of the side chains and held together by intermolecular hydrogen bonds between the side chains. The details of the hydrogen-bonding network in the sheets and its effects on the isocoumarin ring system in each molecule as well as the details of the folding and hydrogen bonding of the side chains will each be discussed in turn.

As seen in Fig. 3, each isocoumarin ring system forms intermolecular hydrogen bonds to its four neighbors from the hydroxyl group O(5)H to the carbonyl C(1)O(1) and between the hydroxyl groups O(6)H and O(5)H. From a consideration of O$\cdots$O distances only, there appears to be a bifurcated hydrogen bond (Donohue, 1968) since both O(1) and O(6) are at normal hydrogen-bonding distances from O(5), 2·864 (3) and 2·790 (3) Å respectively. However, the hydrogen atom, which was clearly observed in the difference Fourier synthesis and subsequently refined, is much closer to O(1), 2·08 (3), than O(6), 2·45 (3) Å. In addition, the O(5)–H$\cdots$O(1) angle, 157°, is well within the normal range for hydrogen bonds (Donohue, 1968). There appears to be little attraction between the hydrogen atom and O(6) since the C(5)–O(5)–H angle is normal, 106 (5)°; however, a firm conclusion cannot be made because of the inherent lack of precision in the hydrogen positions. Also of interest is the fact that although O(5) is 0·26 Å out of the plane of the carbonyl, C(5)C(6)O(6)C(7), the hydrogen atom is exactly in the plane (0·03 Å out of plane). This hydrogen bond links
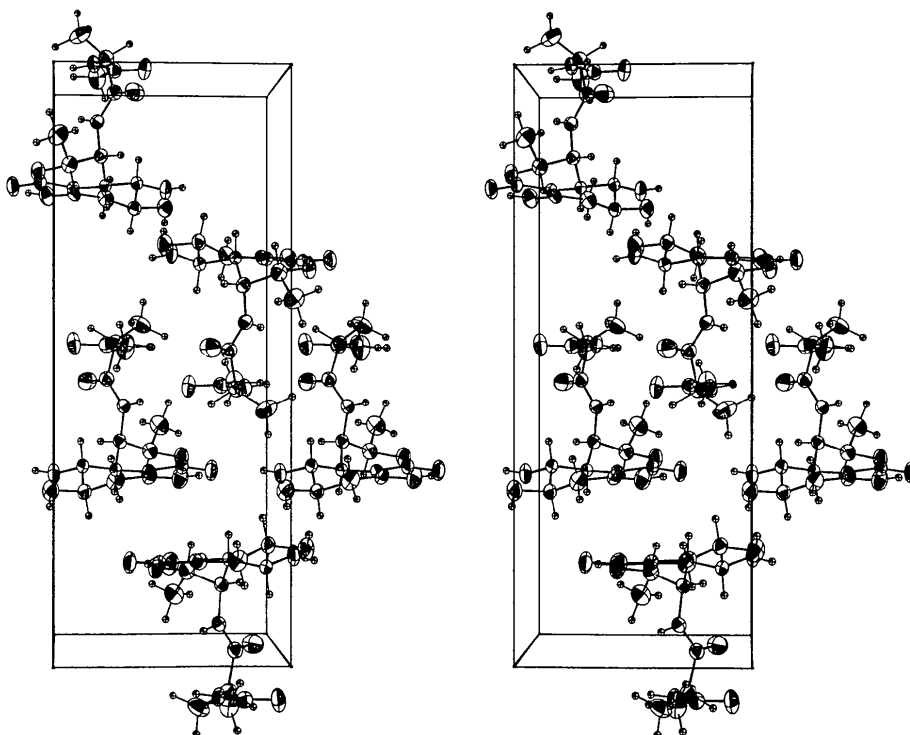

Fig. 2. A stereo diagram of the unit cell and five molecules of N-acetylactinobolin viewed along the b axis.

molecules related by unit-cell translations along the a axis. The other intermolecular hydrogen bond O(6)–H$\cdots$O(5) can readily be assigned from the O(6)$\cdots$O(5) distance, 2·784 (3) Å. The hydrogen atom was clearly observed in the difference Fourier synthesis and subsequently refined; it is 2·00 (3) Å from O(5) and the hydrogen-bonding angle, O(6)–H$\cdots$O(5), is 165°. Again the C(6)–O(6)–H angle, 109 (5)°, is apparently undisturbed from the ideal tetrahedral angle. This hydrogen bond links molecules related by the 2_1 screw axis parallel to the b axis. This results in a nearly square (cf. a and b axis lengths) network of hydrogen-bonded isocoumarin ring systems which forms a pleated sheet with rows of parallel side chains along a extending in an alternating fashion from opposite sides of the sheet along b.

The other hydroxyl group in the isocoumarin ring, O(8)H, forms a typical intramolecular hydrogen bond with the neighboring carbonyl C(1)O(1). This hydrogen atom was also readily identified and refined; it forms quite a short bond, 1·72 (3) Å, to the carbonyl but lies well out of the carbonyl plane, 0·24 Å, as does the hydroxyl oxygen, 0·31 Å. This is probably due to a combination of effects from the molecular conformation (see below) and the minimum permissible hydrogen-bonding distance (Donohue, 1968) which is reflected in the O(6)$\cdots$O(8) distance, 2·554 (3) Å. As was the case for the intermolecular hydrogen bonds the C(8)–O(8)–H angle, 107 (5)°, is apparently unaffected by the formation of the hydrogen bond.

Despite the variety and complexity of the hydrogen bonding for the isocoumarin ring system, it has little or no effect on the bonding or conformation of this part of the molecule. The bonds in the β-ketolactone system centered about the atom chain C(8)C(9)C(1) are all typical. In particular, the double bond at C(8)–C(9), 1·354 Å, shows the effect of conjugation as does the bond C(1)–O(1), 1·245 Å. Each one of these sp^2 carbon atoms and the three atoms bound to each form accurate planes; the average deviations are 0·001, 0·002, and 0·005 Å for the planes at C(1), C(8) and C(9),

respectively. The two planes joined by the double bond C(8)–C(9) are coplanar as expected [average deviation C(7)C(8)O(8)C(9)C(10)C(1) is ±0·012 Å], but the plane containing the carbonyl C(1)–O(1) is twisted with respect to the other planes by 7·4° imparting a slight right-handed helical character to the β-ketolactone group. The other bonds in the isocoumarin ring system are normal C–C and C–O single bonds and are similar to those found for the four independent molecules in actinobolin.HI (Wetherington & Moncrief, 1975) despite the completely different molecular packing and hydrogen-bonding scheme found in that structure.

The ring system in N-acetylactinobolin is rigid; the r.m.s. thermal displacements are quite small and fall with few exceptions in the range 0·15–0·20 Å. The sub-
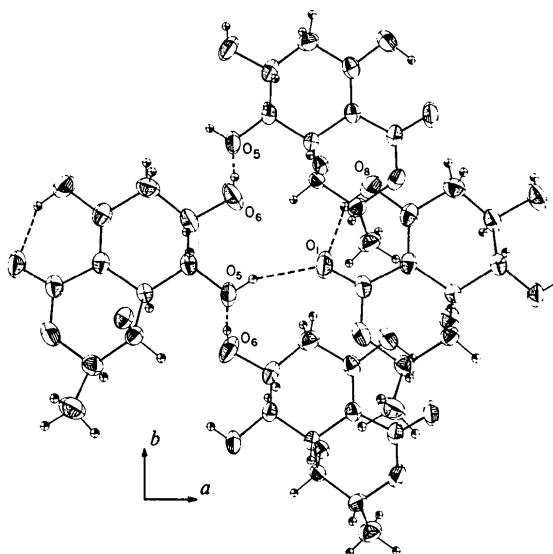

Fig. 3. Hydrogen-bonded network between four neighboring isocoumarin ring systems in the structure of N-acetylactinobolin. The a and b axis directions are indicated.

Table 5. *Conformational parameters in N-acetyl-L-alanine side chain of N-acetylactinobolin and other peptides*

	N-Acetyl-actinobolin	L-AANMA[e]	DL-APHNMA[f]	DL-ALNMA[g]	β-Poly-L-alanine	α-Poly-L-alanine
$\varphi(°)$	−83·6	−86·0	−105·5	−93·8	−139·5	−57·4
$\psi(°)$	−24·1	156·9	107·9	140·4	140·5	−47·5
$\omega_1(°)$	−175·7	−173·3	178·1	177·2	−176·3	−179·8
$\omega_2(°)$	176·8	172·6	−175·0	174·1	−176·3	−179·8
$\theta(°)^{(a)}$	83·1	89·8	69·2	71·6	43·3	
$C_{\alpha1}$–$C_{\alpha2}(\text{Å})^{(b)}$	4·57	6·84	6·20	6·65	6·95	6·01
$P(\text{Å})^{(c)}$	5·44	4·79	4·55	5·26	6·80	5·40
$n^{(d)}$	3·57	1·49	1·98	1·62	1·96	3·61
References	(1)	(2)	(3)	(4)	(5)	(6)

(a) θ: dihedral angle between the two peptide groups. (b) $C_{\alpha1}$–$C_{\alpha2}$: chain repeat distance between successive α-carbon atoms. (c) P: repeat distance per turn along helix axis. (d) n: number of peptide units per turn. (e) L-AANMA: N-acetyl-L-alanyl-N-methylamide; averaged values for the two independent molecules. (f) DL-APHNMA: N-acetyl-DL-phenylalanyl-N-methylamide. (g) DL-ALMNA: N-acetyl-DL-leucyl-N-methylamide.

References: (1) Present study. (2) Harada & Iitaka (1974b). (3) Harada & Iitaka (1974a). (4) Ichikawa & Iitaka (1969). (5) Marsh, Corey & Pauling (1955). (6) Arnott & Dover (1967).

stituent atoms show somewhat more thermal motion, mostly parallel to the c axis. As a consequence, a riding correction (Busing & Levy, 1964) was applied to the bond lengths for these atoms. The corrected lengths are C(5)–O(5), 1·440; C(6)–O(6), 1·434; C(8)–O(8), 1·360; C(1)–O(1), 1·261; and C(3)–C(11), 1·511 Å.

The isocoumarin ring system is in the double chair conformation as predicted by Antosz, Nelson, Herald & Munk (1970) based upon a careful NMR study in water and dimethyl sulfoxide (Antosz, 1969) and also found in actinobolin.HI (Wetherington & Moncrief, 1975). In particular, the apparent coupling constants (9–10 Hz) between the protons in *trans*–axial positions at C(5)–C(6), C(5)–C(10) and C(6)–C(7) are in excellent agreement with the torsion angles (Table 4) for these linkages. In addition, the apparent coupling constants (2–3 Hz) between the protons in the *cis*–equatorial positions C(3)–C(4) and C(4)–C(10) are also in excellent agreement with the observed torsion angles. However, the apparent coupling constant (6·5 Hz) between the equatorial proton on C(7) and the proton on C(6) is much larger than would be expected from the torsion angle between these protons observed in the solid. This apparent discrepancy may arise from the effects of both the conjugated system starting at C(8) and the electronegative hydroxyl at C(6), as well as slight molecular changes required to accommodate the hydrogen bonding in the solid.

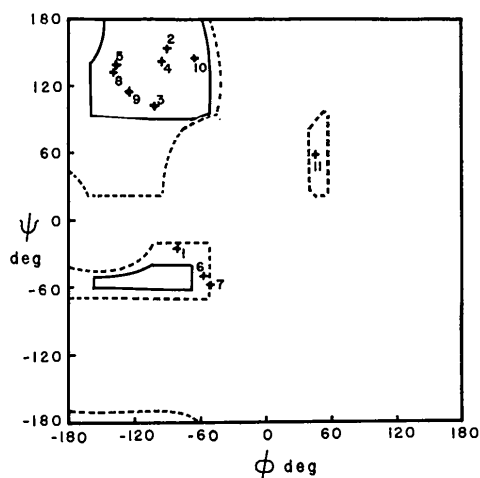In contrast to the situation for the isocoumarin ring

Fig. 4. The Ramachandran plot showing the torsion angles (φ, ψ) for various peptides. The solid line encloses fully allowed regions and the broken line encloses the absolute minimum contact regions. (1) *N*-Acetylactinobolin (this work); (2) *N*-acetyl-L-alanyl-*N*-methylamide (Harada & Iitaka, 1974b); (3) *N*-acetyl-DL-phenylalanyl-*N*-methylamide (Harada & Iitaka, 1974a); (4) *N*-acetyl-DL-leucyl-*N*-methylamide (Ichikawa & Iitaka, 1969); (5) β-poly-L-alanine (Marsh, Corey & Pauling, 1955); (6) α-poly-L-alanine (Arnott & Dover, 1967); (7) right-handed α-helix (Pauling & Corey, 1951); (8) antiparallel chain pleated sheet (Pauling & Corey, 1953); (9) parallel chain pleated sheet (Pauling & Corey, 1953); (10) three-stranded collagen helix (Sasisekharan, 1962); (11) left-handed α-helix.

system, the conformation of the *N*-acetyl-L-alanyl side chain is considerably affected by crystal packing and hydrogen bonding. This is best seen in a comparison of torsion angles (Table 5) for the side chain as found in *N*-acetylactinobolin and as an individual molecule in *N*-acetylalanyl-*N*-methylamide (Harada & Iitaka, 1974b) and *N*-acetylphenylalanyl-*N*-methylamide (Harada & Iitaka, 1974a). In the latter cases the peptide chain is in an extended form while in *N*-acetylactinobolin it is folded to accommodate more efficient packing of the molecular sheets.

From a comparison of φ and ψ torsion angles for various small peptides it is found that the values obtained invariably fall in the β region of a φ–ψ chart (Ramachandran, Ramakrishnan & Sasisekharan, 1963) as shown in Fig. 4. The resulting molecular conformations are then essentially the same as the parallel or antiparallel β-pleated-sheet protein structures (Pauling & Corey, 1953) or the three stranded collagen helix structure (Sasisekharan, 1962). In none of these studies has there been observed a set of φ, ψ torsion angles which fall in the right-handed α-helix (Pauling & Corey, 1951) stability region. However, the folding of the *N*-acetyl-L-alanyl side chain in *N*-acetylactinobolin is such that the φ, ψ angles do fall within the α-helix stability region. This residue from C(4) to C(18) then comprises roughly one half turn of an α-helix and thus is the shortest recognizable segment of a helix possible. It is then of considerable interest to compare in detail the conformation of this side chain, as well as its bonding parameters, with various examples of α-helices.

The usual torsion angles for an α-helix are in the ranges $-67 \le \varphi \le -48°$ and $-57 \le \psi \le -44°$ (Ramakrishnan & Ramachandran, 1965; Arnott & Wonacott, 1966) and, in particular, α-poly-L-alanine has φ, ψ angles of $-57·4$ and $-47·5°$ (Arnott & Dover, 1967). However, for *N*-acetylactinobolin these angles are $-83·6$ and $-24·1°$, respectively, which are somewhat outside the normal range for an α-helix but they still lie within the region defined by the absolute minimum contact distances (Ramakrishnan & Ramachandran, 1965). The extra twisting of the side chain is probably required for effective hydrogen bonding between the side chains on adjacent molecules. It is interesting to note that the values for the repeat distance, 5·44 Å, and the number of residues per turn, 3·57, calculated (Schellman & Schellman, 1964) from the φ, ψ angles for *N*-acetylactinobolin are identical to those for an ideal α-helix.

Both peptide groups are nearly flat; the average deviation from the plane for N(12)C(13)O(13)C(14) is $\pm 0·006$ Å with C(4) 0·085 Å out of plane and corresponds to an ω_1 torsion angle of 176·8°. The average deviation for N(16)C(17)O(17)C(18) is $\pm 0·003$ Å with C(14) 0·070 Å out of plane. The corresponding ω_2 torsion angle, $-175·7°$, gives a slight twist in this peptide of the opposite hand to the first peptide and if anything is opposite to what might be expected from the pattern of the hydrogen bonding between O(17) and N(16)H.

As can be seen in the packing diagram in Fig. 2, the two amide hydrogen atoms, H(N12) and H(N16), on each molecule form hydrogen bonds to the same carbonyl oxygen atom, O(17), on the side chain of the neighboring molecule related by a 2_1 operation along c. These intermolecular hydrogen bonds link molecules in adjacent sheets thus imparting three-dimensional rigidity to the crystal structure. The hydrogen-bonding distances between N(12)–H···O(17), 3·075 (3), and N(16)–H···O(17), 2·832 (3) Å, are typical; however, both protons are considerably out of the plane on opposite sides of the acceptor carbonyl group, 0·66 and 1·49 Å respectively, much as in benzamide (Penfold & White, 1959). This hydrogen-bonding scheme is quite different from that found in actinobolin.HI (Wetherington & Moncrief, 1975) where the amide hydrogen atoms on all four molecules bond to the carbonyl on the isocoumarin ring of an adjacent molecule. In addition, there is no interaction between the C(13)O(13) carbonyl and the O(5)H hydroxyl or any other hydroxyl in contrast to what had been found for one of the molecules in actinobolin.HI.

Because the conformation of the N-acetyl-L-alanyl side chain is a consequence of intermolecular hydrogen bonding and crystal packing effects there can be no reason to assume that this conformation is maintained in solution. The apparent coupling constant (9·5 Hz) between the protons on C(4) and N(2) obtained from the NMR spectrum of N-acetylactinobolin in dimethyl sulfoxide-d_6 (Antosz, 1969) is in accord with the torsion angles for the C(4)N(12) bond found in the crystal structure and thus is probably fixed by steric interactions between the side chain and the ring system. The other conformationally significant apparent coupling constant (7 Hz) between the protons on C(14) and N(16) is also consistent with the torsion angle (φ) found in the crystal; however, that coupling constant is also what would be obtained if there were rapid rotation about the C(14)N(16) bond as might be the case in solution. Therefore, it is apparent that the preferred orientation of the first peptide group C(4)N(12)C(13)-O(13)C(14) with respect to the isocoumarin ring system is the same in both the crystal structure and in dimethyl sulfoxide solution, but the preferred orientation of the second peptide group in solution is unknown.

Since the side chain on N-acetylactinobolin is held in place by intermolecular hydrogen bonding it is quite rigid. The r.m.s. displacements for the atoms in the chain are quite small and fall in the range 0·15–0·23 Å. As in the isocoumarin ring system the substituent atoms show somewhat more thermal motion, mostly parallel to the c axis, therefore riding corrections (Busing & Levy, 1964) have been applied to the bond lengths for these atoms. The corrected bond lengths are C(13)–O(13), 1·243; C(14)–C(15), 1·549; C(17)–O(17), 1·255; and C(17)–C(18), 1·520 Å.

Despite the drastic change in the conformation of the peptide side chain in N-acetylactinobolin, the bond distances and angles compare quite well with those for an average polypeptide chain. In particular the two $C_\alpha N$ distances, C(4)–N(12) (1·460) and C(14)–N(16) (1·445), are exactly the same as for the average polypeptide, 1·455 Å (Marsh & Donohue, 1967). The corrected carbonyl bond lengths for C(13)–O(13), 1·243, and C(17)–O(17), 1·255, also agree quite well with the average value, 1·24 Å. However, both bonds to the carbonyl carbon atom, NC' and $C_\alpha C'$, are longer than those given by Marsh & Donohue (1967). The NC' bonds, N(12)–C(13) (1·345), and N(16)–C(17) (1·339), differ from the average values, 1·325 and 1·32 Å given by Marsh & Donohue (1967) and Pauling & Corey (1953) respectively, but are the same as that found L-alanyl-L-alanine, 1·344 Å (Fletterick, Tsai & Hughes, 1971). Similarly, the $C_\alpha C'$ bonds, C(14)–C(13) (1·542) and C(18)–C(17) (1·520 corrected), differ from the average value, 1·51 Å, given by Marsh & Donohue (1967) but are the same as the older value, 1·53 Å, given by Pauling & Corey (1953) and are the same as that found in L-alanyl-L-alanine, 1·530 Å.

The angles in the side chain of N-acetylactinobolin show at most a 3° deviation from the average values of Marsh & Donohue (1967). The average values for similar angles for C_α–C'–O (121·2), N–C'–O (122·8), N–C'–C_α (116·0) and C_α–N–C' (122·1°) are exactly the same as the values given by Marsh & Donohue (1967). Only the angle at the α-carbon atom, C(13)–C(14)–N(16) (114·4), is markedly different from the typical value, 111°. This deviation might be a result of a close interaction accompanying the small ψ torsion angle.

References

ANTOSZ, F. J. (1969). Structure and Chemistry of Actinobolin. Thesis, Arizona State Univ.

ANTOSZ, F. J., NELSON, D. B., HERALD, D. L. JR & MUNK, M. E. (1970). J. Amer. Chem. Soc. 92, 4933–4942.

ARNOTT, S. & DOVER, S. D. (1967). J. Mol. Biol. 30, 209–212.

ARNOTT, S. & WONACOTT, A. J. (1966). J. Mol. Biol. 21, 371–383.

BUSING, W. R. & LEVY, H. A. (1964). Acta Cryst. 17, 142–146.

COCHRAN, W. & DOUGLAS, A. S. (1955). Proc. Roy. Soc. A 227, 486–500.

DECLERCQ, J. P., GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1973). Acta Cryst. A 29, 231–234.

DONOHUE, J. (1968). Structural Chemistry and Molecular Biology, edited by A. RICH & N. DAVIDSON, pp. 443–465. San Francisco: Freeman.

FLETTERICK, R. J., TSAI, C. C. & HUGHES, R. E. (1971). J. Phys. Chem. 75, 918–922.

GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1971). Acta Cryst. A 27, 368–376.

HAMILTON, W. C., ROLLETT, J. S. & SPARKS, R. A. (1965). Acta Cryst. 18, 129–130.

HARADA, Y. & IITAKA, Y. (1974a). *Acta Cryst.* B30, 726–730.

HARADA, Y. & IITAKA, Y. (1974b). *Acta Cryst.* B30, 1452–1459.

HASKELL, T. H. & BARTZ, Q. R. (1959). *Antibiot. Ann.* pp. 505–509.

HAUPTMAN, H. & KARLE, J. (1956). *Acta Cryst.* 9, 45–55.

ICHIKAWA, T. & IITAKA, Y. (1969). *Acta Cryst.* B25, 1824–1833.

International Tables for X-ray Crystallography (1974). Vol. IV, pp. 99–101. Birmingham: Kynoch Press.

KOCH, M. H. J. (1974). *Acta Cryst.* A30, 67–70.

MARSH, R. E., COREY, R. B. & PAULING, L. (1955). *Acta Cryst.* 8, 710–715.

MARSH, R. E. & DONOHUE, J. (1967). *Advanc. Protein Chem.* 22, 235–256.

MUNK, M. E., NELSON, D. B., ANTOSZ, F. J., HERALD, D. J. JR & HASKELL, T. H. (1968). *J. Amer. Chem. Soc.* 90, 1087–1089.

PAULING, L. & COREY, R. B. (1951). *Proc. Natl. Acad. Sci. U.S.* 37, 235–240.

PAULING, L. & COREY, R. B. (1953). *Proc. Natl. Acad. Sci. U.S.* 39, 253–256.

PENFOLD, B. R. & WHITE, J. C. B. (1959). *Acta Cryst.* 12, 130–135.

RAMACHANDRAN, G. N., RAMAKRISHNAN, C. & SASISEKHARAN, V. (1963). *J. Mol. Biol.* 7, 95–99.

RAMAKRISHNAN, C. & RAMACHANDRAN, G. N. (1965). *Biophys. J.* 5, 909–933.

ROLLETT, J. S. & CARRUTHERS, R. S. (1974). Private communication.

SASISEKHARAN, V. (1962). *Collagen*, edited by N. RAMANATHAN, pp. 39–78. New York: Interscience.

SCHELLMAN, J. A. & SCHELLMAN, C. (1964). *The Proteins*, edited by H. NEURATH, Vol. II, pp. 1–371. New York: Academic Press.

WETHERINGTON, J. B. & MONCRIEF, J. W. (1975). *Acta Cryst.* B31, 501–511.

X-ray Crystal Structures of Hexabromotellurates of Organic Ions.
I. Crystal Structure of the Hexabromotellurate of DL-α-Ammonio-*n*-butyric Acid

By FRANÇOISE DAHAN AND ODETTE LEFÈBVRE-SOUBEYRAN

*Laboratoire de Cristallochimie, CNRS, ER N° 63, Tour 44, 11 Quai Saint-Bernard, 75005-Paris, France*

The structure of the hexabromotellurate of DL-α-ammonio-*n*-butyric acid has been determined by the heavy-atom method and refined by full-matrix least squares to a final R of 0·041 for 1049 independent reflexions. Crystals are monoclinic, space group $P2_1/c$, with $a=11·443$ (4), $b=7·881$ (3), $c=12·886$ (4) Å, $\beta=112·19$ (5)°, $Z=2$. The crystalline cohesion is controlled by N–H$\cdots$Br and O–H$\cdots$O hydrogen bonds and by intermolecular van der Waals forces.

Introduction

Gillespie & Nyholm (1957) predicted that hexacoordinated complexes of Te^{IV} and Se^{IV} should not possess a regular octahedral structure but rather one based on seven coordination with a lone pair of electrons occupying the seventh position. Howewer, infrared and Raman spectroscopy (Adams & Morris, 1967), nuclear quadrupole resonance spectroscopy (Greenwood, 1970) and electronic spectra (Couch, Wilkins, Rossman & Gray, 1970) showed no distortion of TeX_6^{2-} octahedra (X = Cl or Br). The general conclusion made by Johnstone, Jones & Vasudev (1972) is that the Te atom employs essentially pure p orbitals in bonding, the $5s^2$ electrons being stereochemically inactive. Thus the TeX_6^{2-} ions are octahedral and have relatively long weak Te–X bonds when electronegativity of the ligand increases. No significant distortion of $TeBr_6^{2-}$ octahedra was observed in the few X-ray structures reported, K_2TeBr_6 (Brown, 1964), $(NH_4)_2TeBr_6$ and Cs_2TeBr_6 (Das & Brown, 1966).

Crystalline salts containing $TeBr_6^{2-}$ anions, and cations of protonated amino acids or protonated amides, have been synthesized (Dobrowolski & Pastuszak, 1970; Bujewski & Dobrowolski, 1973). Professor Dobrowolski suggested an X-ray study of a compound of each kind in order to compare the regularity of the $TeBr_6^{2-}$ ions and the influence of the different organic cations.

The structure described here is the hexabromotellurate of DL-α-ammonio-*n*-butyric acid

$$[CH_3-CH_2-CH-COOH]_2.TeBr_6.$$
$$|$$
$$NH_3$$
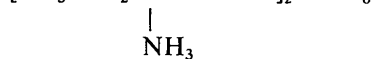
Experimental

A single crystal $0·15 \times 0·075 \times 0·075$ mm was selected for the X-ray investigation.

The intensities were collected on a CAD-3 Enraf–Nonius diffractometer in the θ–2θ scan mode. The range of each scan consisted of the base width of 1·0° at $2\theta=0°$ and an increment, $\Delta(2\theta)=0·5\tan\theta°$, to al-