Acta Cryst. (1964). **17**, 496

Crystal and Molecular Structure of 3-Indolylacetic Acid

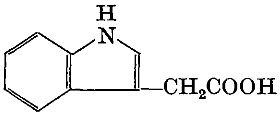
BY ISABELLA L. KARLE, KATHLEEN BRITTS AND P. GUM
U.S. Naval Research Laboratory, Washington, D.C., U.S.A.

(*Received* 27 *June* 1963)

The new symbolic addition procedure for direct phase determination has been applied to the solution of the structure of 3-indolylacetic acid. The material crystallizes in the space group $P2_1/c$ with cell dimensions: $a = 17{\cdot}99$, $b = 5{\cdot}28$, $c = 9{\cdot}61$ Å and $\beta = 106°\ 05'$. The molecules exist as dimers with hydrogen bonds between the acid groups. However, the NH in the indole group is not involved in any hydrogen bonding. The molecule is characterized by two planes, one through the indole group and one through the acid group. The dihedral angle between the planes is 62° 52′.

Introduction

3-Indolylacetic acid,



is an important plant growth hormone. We were motivated to investigate its crystal structure because of its biological significance, and to gain further experience with a new direct procedure for phase determination which has been designated 'symbolic addition' (Karle & Karle, 1963 and to be published).

Experimental

Crystals of 3-indolylacetic acid were grown in the form of diamond-shaped plates which cleaved very readily into very thin platelets. The cleavage was parallel to (100). Since a fairly thin plate had to be used to collect the intensity data, the observed intensities have inaccuracies due to the difficulty of estimating the density of the resulting irregular diffraction spots. Equi-inclination Weissenberg photographs were taken along the b and c axes and zero level precession photographs on the a and b axes.

The space group was determined to be $P2_1/c$ with unit-cell dimensions:

$$a = 17{\cdot}99 \pm 0{\cdot}03,\ b = 5{\cdot}28 \pm 0{\cdot}01,\ c = 9{\cdot}61 \pm 0{\cdot}02 \text{ Å};$$
$$\beta = 106°\ 05' \pm 10' .$$

The density calculated on the basis of four molecules in the unit cell is 1·321 g.cm^{-3}. The density measured by flotation in mixed solvents is 1·317 g.cm^{-3}.

The intensities were estimated visually with a calibrated film strip. These data were punched on IBM cards directly from the primary data sheets and processed on the IBM 7090 computer (Norment, 1962). The computer output yielded the data in structure factor F and normalized structure factor E form where

$$E_{\mathbf{h}}^2 = F_{\mathbf{h}}^2/\varepsilon \sum_{j=1}^{N} f_{j\mathbf{h}}^2 . \qquad (1)$$

For space group $P2_1/c$, $\varepsilon = 1$ except $\varepsilon = 2$ when $\mathbf{h}$ is $h0l$ or $0k0$, and f_j is the atomic scattering factor for the jth atom. For this crystal there were 1289 data and 865 non-zero data. The statistical averages for the normalized structure factors are listed in Table 1.

Table 1. *Experimental and theoretical values of normalized structure factors*

	Experimental	Centrosym.	Non-centrosym.
$\|E\|$	0·769	0·798	0·886
$\|E_2-1\|$	0·934	0·968	0·736
$\|E\|^2$	1·031	1·000	1·000

The experimental values are compared with the theoretical values computed for crystals with randomly distributed atoms. The distribution of the normalized structure factors was 0·15% for $|E|>3$, 3·3% for $|E|>2$ and 36·1% for $|E|>1$. These values can be compared with those computed for a theoretical centrosymmetric crystal with randomly distributed atoms, *i.e.* 0·3%, 5·0% and 32·0%, respectively.

Phase determination

The phases were determined by the symbolic addition procedure described in a previous publication (Karle & Karle, 1963). The origin was specified by assigning signs to three linearly independent reflections (Hauptman & Karle, 1953). In addition, the signs of four other reflections were specified with a letter. These seven assignments which form the basic starting set for applying the Σ_2 formula are shown in Table 2. The Σ_2 formula is

$$sE_{\mathbf{h}} \sim s\sum_{\mathbf{k}} E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}} \qquad (2)$$

where s means 'sign of'.

The reflections chosen to have signs a, b, c, d were selected on the basis that they had many relationships

Table 2. *Assignment of three origin-specifying reflections and four other reflections as a starting set for the application of* Σ_2

Sign	*hkl*	$\lvert E\rvert$
+	$\bar{8}35$	2·71
+	$\overline{15},1,2$	2·34
+	$\overline{21},0,6$	3·15
a	$\bar{6}24$	3·33
b	925	2·35
c	$\bar{7}36$	2·20
d	$\overline{16},4,6$	2·18

applicable to the Σ_2 formula among themselves and with the origin-specifying reflections. In addition, they are all associated with large values for $|E|$, giving very high probabilities that the indications from the Σ_2 formula are correct. In the course of application, as many Σ_2 combinations as possible were used to determine each phase. The squares of the letter symbols were not carried through the determination since each symbol represents a positive or negative sign and, therefore, the square of a symbol is always positive. Those Σ_2 combinations which were solely given in terms of the square of a symbol indicated a plus sign. The only reflections considered were those with $|E|\geq 1{\cdot}5$. Three-dimensional reflections ($hkl \neq 0$) were used in general to determine other phases, since it has been found that one- and two-dimensional reflections often cause many more inconsistencies in the Σ_2 formula than do three-dimensional reflections. To facilitate the application of Σ_2, an IBM 7090 program was prepared to list the sets of interacting pairs for each *hkl* arranged in order of decreasing $|E|$ values (Norment, 1963). One hundred and twenty-three phases as a function of *a*, *b*, *c* and *d* were determined in this manner.

In the absence of any other phase information, at most 2^n (where $n=4$, the number of unknown symbols) different density maps would be required in order to find the structure. At this point other phase determining formulae were applied to decrease the number of possibilities, or at least, to indicate the most probable values of *a*, *b*, *c* and *d*. For example, an application of Σ_1 (Hauptman & Karle, 1953),

$$sE_{2h2k2l} \sim s(E^2_{hkl}-1)\;, \qquad (3)$$

to the reflection $\bar{6}24$ gave a probability of 0·91 that the sign is positive. (The probability that a certain sign is positive was computed from

$$P_+(E_{2\mathbf{h}}) = \tfrac{1}{2} + \tfrac{1}{2}\tanh\,(1/2\sqrt{N})|E_{2\mathbf{h}}|(E^2_{\mathbf{h}}-1)\;, \qquad (4)$$

a formula analogous to one derived by Woolfson (1954) for Σ_2 and by Hauptman & Karle (1953)). Hence the symbol *a* which was assigned to $\bar{6}24$ can be considered to be most likely positive. An application of another Σ_1 relationship,

$$sE_{2h0 2l} \sim s\sum_k\,(-1)^{k+l}(E^2_{hkl}-1)\;, \qquad (5)$$

to the reflections 14,0,0 and $\overline{16},0,4$ gave the probabilities of 0·71 and 0·60, respectively, that the combination *ac* should be positive. Although these are not definitive probabilities, they serve to indicate the order in which density maps should be computed. Actually, *a* and *c* were both found to be positive.

These simple Σ_1 relations did not afford much information concerning the values of *b* and *d*. One indication for the reflection 628 showed that the combination *abd* should be positive with a probability of 0·60. It was noted, in addition, that the Σ_2 relations had fewest inconsistencies if *b* and *d* had the same sign. On the basis of this simple analysis, the most probable combinations for *a*, *b*, *c* and *d* were (+ + + +) and (+ − + −). The combination (+ + + +) is possible in this space group, $P2_1/c$, because negative signs are obtained from relationships among the structure factors, *e.g.* using two of the reflections from Table 1, $\overline{15},1,2(+)$ and $\bar{6}24(a)$, one has the combination

$$\begin{array}{cccc} 15 & 1 & \bar{2} & - \\ 6 & 2 & 4 & a \\ \hline 9 & 3 & 2 & -a\,. \end{array}$$

A three-dimensional density map computed with 123 normalized structure factors *E* as coefficients (about 10 terms per atom in the asymmetric unit), and the phases based on the combination that *a*, *b*, *c* and *d* were all positive, revealed the positions of the thirteen atoms immediately. Sections of this three-dimensional map are projected on the (010) plane in Fig. 1. It was determined later that two signs out of 123 in this initial set were incorrectly determined.

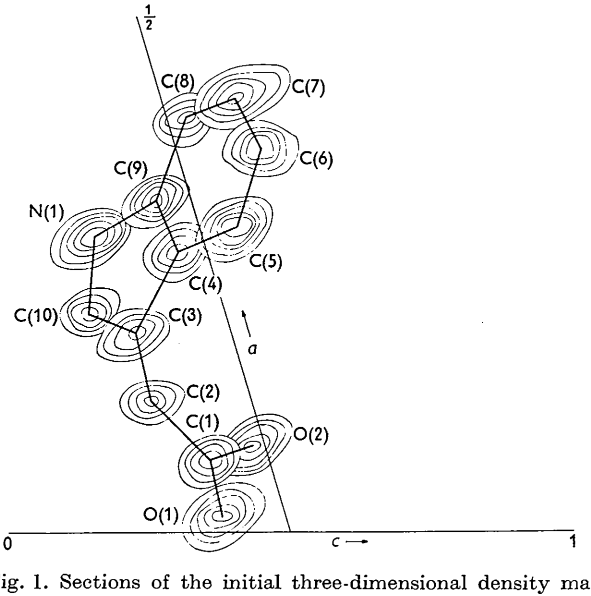


Fig. 1. Sections of the initial three-dimensional density map computed with 123 *E* values as coefficients. The contours are drawn at equal intervals on an arbitrary scale.

Structure

The coordinates of the thirteen atoms as read from the initial density map were subjected to a least-squares refinement by the ORFLS program (Busing, Martin & Levy, 1962) which was adapted in our laboratory to the IBM 7030 (STRETCH). Two cycles with isotropic and four cycles with anisotropic temperature factors were performed during which the R index decreased from 25·5% to 18·2%.* The unfavorable shape of the crystal contributed to the relatively high R factor. The refined atomic parameters are listed in Table 3. A final electron density map computed with all the data is shown in Fig. 2.

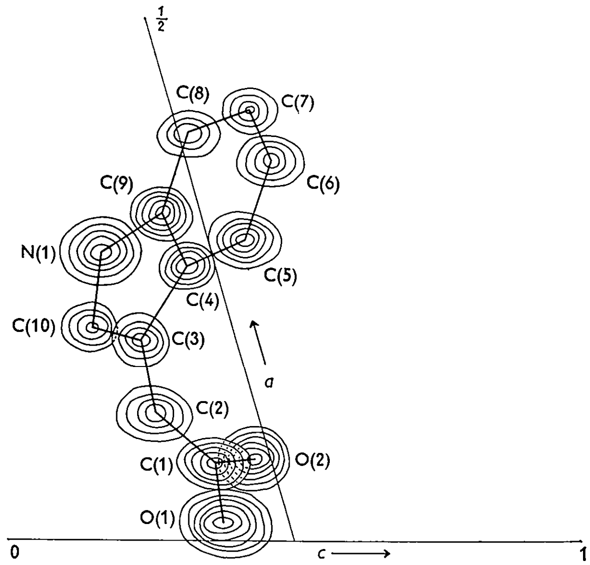


Fig. 2. Sections of the final three-dimensional electron density map projected on the (010) plane. The contours are at intervals of 1 e.Å^{-3}, beginning with the 2 e.Å^{-3} contour.

In the crystal, the molecules exist as dimers. Hydrogen bonds are formed between the acid groups across centers of symmetry. The NH in the indole group is not involved in any hydrogen bonding. The molecule is characterized by two planes, one through the indole group and the other through the acid group. The dihedral angle between the planes is 60° 52′. The equations of the least-squares planes are, for the indole group,

$$6{\cdot}6549x+2{\cdot}4284y-5{\cdot}0709z=2{\cdot}1076 \qquad (6)$$

and for the acid group,

$$4{\cdot}8492x+1{\cdot}9430y+4{\cdot}8934z=4{\cdot}4506\ . \qquad (7)$$

The deviations of the individual atoms from the planes are listed in Table 4. The r.m.s. deviation from the indole group is 0·0098 Å and from the acid group 0·012 Å. The two carboxyl groups are hydrogen bonded across a center of symmetry, and are therefore in parallel planes. In this molecule, the two carboxyl groups are nearly coplanar since the planes are separated by only 0·023 Å. The separation of such planes

Table 4. *Distances of atoms from the least-squares planes through the indole group and the acid group in 3-indolylacetic acid*

Indole group		Acid group	
Atom	Δ	Atom	Δ
C(3)	−0·0021 Å	C(1)	+0·0206 Å
C(4)	−0·0083	C(2)	−0·0056
C(5)	−0·0007	O(1)	−0·0071
C(6)	−0·0006	O(2)	−0·0079
C(7)	+0·0129		
C(8)	−0·0005		
C(9)	−0·0153		
C(10)	+0·0189		
N(1)	−0·0056		

Table 3. *Fractional coordinates for 3-indolylacetic acid*

The thermal parameters are of the form $T=\exp\,[-(\beta_{11}h^2+\beta_{22}k^2+\beta_{33}l^2+2\beta_{12}hk+2\beta_{13}hl+2\beta_{23}kl)]$.
Each thermal parameter is multiplied by 10^4

Atom	x	y	z	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
C(1)	0·0744	1·1056	0·4010	13	339	89	−10	12	40
C(2)	0·1272	1·1827	0·3127	19	451	147	−15	33	105
C(3)	0·1992	1·0215	0·3354	18	215	99	−15	20	−8
C(4)	0·2717	1·0600	0·4502	14	207	112	0	25	17
C(5)	0·2965	1·2380	0·5665	21	289	137	−24	25	−21
C(6)	0·3730	1·2133	0·6548	31	318	110	−44	10	−41
C(7)	0·4232	1·0286	0·6298	19	427	158	−11	11	50
C(8)	0·4007	0·8496	0·5172	24	260	186	− 8	29	40
C(9)	0·3231	0·8726	0·4293	23	86	115	−13	21	16
C(10)	0·2109	0·8275	0·2537	35	232	145	−51	22	−35
N(1)	0·2856	0·7296	0·3097	35	232	166	−25	37	−24
O(1)	0·0174	1·2618	0·3898	22	332	178	28	29	141
O(2)	0·0807	0·9062	0·4681	17	350	153	15	22	138
Standard error									
C	0·0008	0·0030	0·0017	5	61	22	14	9	35
O	0·0005	0·0021	0·0012	3	47	18	10	6	26
N	0·0007	0·0025	0·0016	5	50	22	14	9	32

* The y coordinates have been chosen so that they can be directly substituted into equations (5) and (6).

---

* A microfilm containing the list of the computed and observed structure factors has been deposited with the Library of Congress, Washington, D.C., U.S.A. The reel number is 8497.

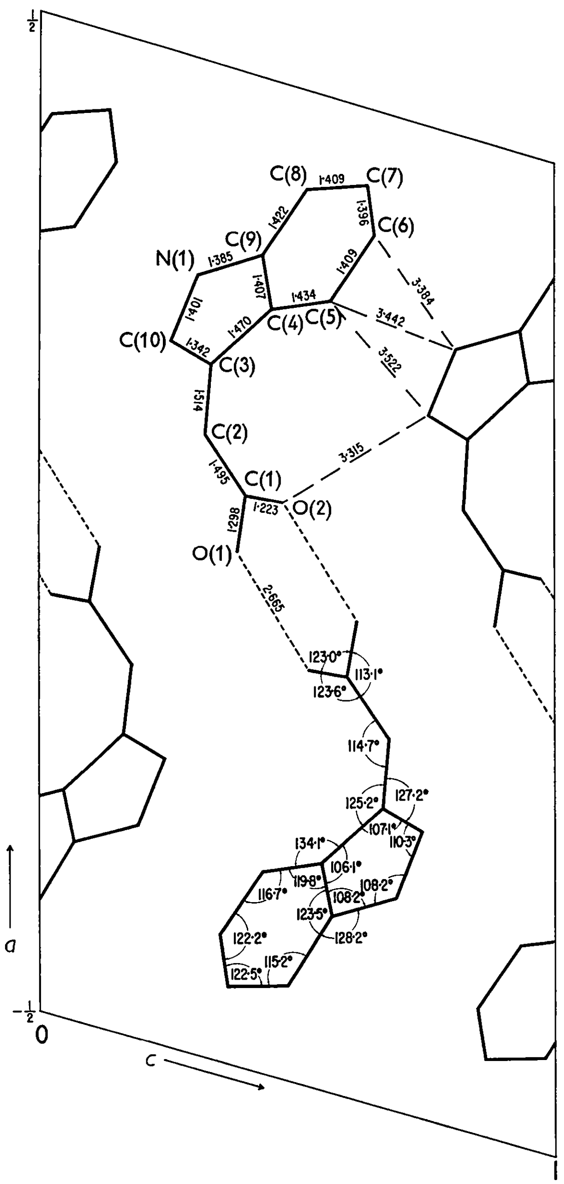


Fig. 3. Bond lengths, bond angles, and nearest intermolecular approaches in 3-indolylacetic acid. Standard deviation for bond lengths is 0·015–0·022 Å and for bond angles 1·3°.

is sometimes as large as 0·5 Å (Jeffrey & Sax, 1963).

The distances and angles in the molecule are shown in Fig. 3 and listed in Table 5. All bond lengths, bond angles and nearest intermolecular approaches have values consistent with previous experience. The six C–C bond lengths in the benzene ring average 1·412 Å.

Table 5. *Bond lengths and angles in 3-indolylacetic acid*

Bond		Angle	
C(1)–O(1)	1·298 Å	O(1)–C(1)–O(2)	123·0°
C(1)–O(2)	1·223	O(1)–C(1)–C(2)	113·1
C(1)–C(2)	1·495	O(2)–C(1)–C(2)	123·6
		C(1)–C(2)–C(3)	114·7
C(4)–C(5)	1·434	C(4)–C(5)–C(6)	116·7
C(5)–C(6)	1·409	C(5)–C(6)–C(7)	122·2
C(6)–C(7)	1·396	C(6)–C(7)–C(8)	122·5
C(7)–C(8)	1·409	C(7)–C(8)–C(9)	115·2
C(8)–C(9)	1·422	C(8)–C(9)–C(4)	123·5
C(9)–C(4)	1·407	C(9)–C(4)–C(5)	119·8
C(3)–C(4)	1·470	C(4)–C(3)–C(10)	107.1
C(3)–C(10)	1·342	C(3)–C(10)–N(1)	110·3
N(1)–C(10)	1·401	C(10)–N(1)–C(9)	108·2
N(1)–C(9)	1·385	N(1)–C(9)–C(4)	108·2
		C(9)–C(4)–C(3)	106·1
		C(2)–C(3)–C(4)	125·2
		C(2)–C(3)–C(10)	127·2
		C(3)–C(4)–C(5)	134·1
		N(1)–C(9)–C(8)	128·2

In the five-membered ring, the two C–N bond lengths are 1·385 and 1·401 Å, the C–C bond length is 1·470 Å and the C=C is 1·342 Å. The acetic group has two C–C bonds at 1·514 and 1·495 Å. The two C–O bonds in the carboxyl group are at 1·223 and 1·298 Å and the OH ··· O bond is 2·665 Å. These values are of the same order of magnitude as found in many other acid dimers. A tabulation of dimensions concerning carboxyl groups has been made by Nardelli, Fava & Giraldi (1962).

We are indebted to Dr J. Karle for stimulating discussions and to Mr Stephen Brenner who performed many of the calculations.

References

Busing, W. R., Martin, K. O. & Levy, H. A. (1962). ORFLS, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.

Hauptman, H. & Karle, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetric Crystal.* A.C.A. Monograph No. 3. Brooklyn: Polycrystal Book Service.

Jeffrey, G. A. & Sax, M. (1963). *Acta Cryst.* **16**, 430.

Karle, I. L. & Karle, J. (1963). *Acta Cryst.* **16**, 969.

Nardelli, M., Fava, G. & Giraldi, G. (1962). *Acta Cryst.* **15**, 737.

Norment, H. G. (1962). U.S. Naval Research Laboratory Report No. 5739. Washington, D. C., U.S.A.

Norment, H. G. (1963). *A Collection of FORTRAN Programs for Crystal Structure Analysis.* U. S. Naval Research Laboratory Report No. 5885. Washington, D. C., U.S.A.

Woolfson, M. M. (1954). *Acta Cryst.* **7**, 61.