Acta Cryst. (1953). **6**, 591

The Electron Distribution in Crystalline α-Pyridone

By Bruce R. Penfold*

Crystallographic Laboratory, Cavendish Laboratory, Cambridge, England

(*Received* 30 *January* 1953)

The electron density in crystals of α-pyridone has been measured with sufficient accuracy in two projections to establish the existence of the molecule in the pyridone form which differs from its tautomer, the hydroxy form, only in the position of one hydrogen atom. The molecular electron density is satisfactorily explained within the experimental error by an assemblage of isolated atoms, each with its own temperature factor, and possessing the electron distribution calculated by McWeeny using analytical approximations to atomic wave functions. Bond lengths have been measured with a standard deviation of 0·010 Å and can be explained in terms of a resonance structure in which polar forms of the molecule play an important part. The hydrogen atom responsible for tautomerism is concerned in an intermolecular N–O hydrogen bond, which, repeated throughout the structure, links molecules into endless helices.

1. Introduction

The potential N–C double bond in pyridine allows a tautomerism between two distinct forms of some of its derivatives, e.g. the α- and γ-hydroxy pyridines (Elderfield, 1950). In the case of α-hydroxy pyridine, one of the two possible tautomers (I and II in Fig. 1) may more correctly be described as α-pyridone (I). This name has been used in the title because it was eventually shown that in the crystalline state the compound exists as form I. Although two separate tautomers have not been isolated, derivatives of both are known. The predominant form in solution, as determined from measurements of the ultra-violet absorption spectra, depends largely on the pH. To gain the desired information as to which tautomeric form exists in the crystal, it is necessary to determine with some precision the positions of the hydrogen atoms, because it is to the alternative positions of one of these that the tautomerism is due. The techniques employed were similar to those used in the determination of the electron distributions in adenine hydrochloride (Cochran, 1951) and in salicylic acid (Cochran, 1953). In this case, however, it was possible to determine experimentally all three coordinates of each hydrogen atom, and hence to calculate interatomic distances involving hydrogens. No accurate structural analysis has previously been made of a compound containing a pyridine ring, and the information obtained about the dimensions of the molecule is therefore also of value.

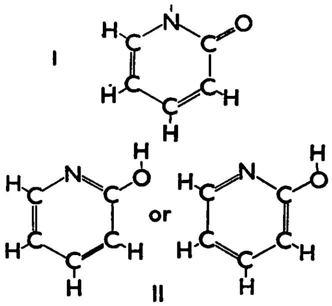


Fig. 1. Tautomeric forms of the molecule.

2. Experimental

The material used was kindly supplied by Prof. L. Hunter of University College, Leicester. It consisted of a sandy-coloured mass of badly formed crystals of α-pyridone. It was purified by crystallization from benzene. Subsequent recrystallization from ethyl alcohol produced a crystalline aggregate containing individual crystals measuring up to several millimetres across, none of which was well formed. Cleavage occurred readily only in planes parallel to the *c* crystallographic axis. Therefore, although nearly cylindrical crystals for rotation about this axis were easily cut, only two crystals suitable for rotation about the *b* axis were obtained.

Crystallographic and physical data

α-pyridone C_5NOH_5; m.p. 104–105° C.; orthorhombic,

$$a = 13{\cdot}63,\ b = 5{\cdot}89,\ c = 5{\cdot}67\ \text{Å}\ (\text{all } \pm 0{\cdot}3\%).$$

Absent spectra: $h00$, $0k0$, $00l$ when h, k, l respectively is odd.

Space group: $P2_12_12_1$. Four molecules per unit cell.

Density (calc.) = 1·391 g.cm.$^{-3}$. Linear absorption coefficient for Cu $K\alpha$ radiation, $\mu = 9{\cdot}3$ cm.$^{-1}$.

Using Cu $K\alpha$ radiation, complete sets of $hk0$ and $h0l$ relative intensities were obtained by use of standard Weissenberg techniques and visual comparison

---

* Now at Canterbury University College (University of New Zealand), Christchurch C. 1, New Zealand.

of spots with an intensity scale. These photographic data were used in the preliminary determination of the crystal structure.

For precision measurements of electron distribution in the crystal, eye-estimated intensities are not sufficiently accurate. The intensities of *hk*0 and *h*0*l* spectra were therefore redetermined using the Geiger-counter spectrometer technique described by Cochran (1950). The range $\sin\theta/\lambda = 0\text{–}0{\cdot}50$ for each zone was covered using Cu $K\alpha$ radiation and, for the *hk*0 zone, the range 0·50–0·765 using Mo $K\alpha$ radiation. Intensities of *h*0*l* spectra in the range 0·50–0·65 used in final refinement procedures were those derived from eye-estimation of spots on Weissenberg photographs. Measurements in the Cu $K\alpha$ range for *hk*0 spectra were made on three crystals having the following maximum and minimum dimensions at right angles to the axis of rotation:

Crystal 1	0·8 × 0·6 mm.
Crystal 2	0·22 × 0·18 mm.
Crystal 3	0·26 × 0·16 mm.

Absorption corrections were made for crystal (1) only, by the method of Albrecht (1939), calculations showing that absorption factors for reflexions from the other two crystals would in no case differ from the mean by more than 2%. The effects of secondary extinction were minimized by using mean intensity values from the two smallest crystals for about 20 of the strongest reflexions. For all other reflexions, mean values from all three crystals were used.

The only suitable crystal available for measuring *h*0*l* intensities with the spectrometer had maximum and minimum cross-sectional dimensions of 0·37 mm. and 0·32 mm. No absorption corrections were made. To minimize extinction the crystal was dipped in liquid air and the intensity values used for the strongest reflexions were those of the dipped crystal. Further allowances for extinction effects in both zones will be mentioned in § 4.

3. Determination of the structure from Patterson projections

Patterson functions projected on (001) and (010) were computed, using relative intensity values which had first been sharpened by dividing them by the factor $S(\sin\theta) = \overline{\mathcal{J}}/K \exp[-2\sin^2\theta]$, where $\overline{\mathcal{J}}$ is the average intensity of reflexions which have nearly the same value of $\sin\theta$, and K is an arbitrary constant.

Orientation of the pyridine ring

If the pyridine ring of α-pyridone is considered to be a regular hexagon of side 1·4 Å, the longest interatomic vector involving ring atoms is 2·8 Å, the distance between opposite corners of the ring. Therefore all such intra-ring (i.-r.) vectors will be represented on the Patterson projection by peaks within 2·8 Å of the origin, and, of these, all but opposite-atom vectors will fall within 2·2 Å of the origin. The only other vectors expected within this latter range are those between the oxygen atom and ring atoms, but these would not seriously obscure the pattern of i.-r. vectors because in most cases they will coincide with the latter. By means of a cardboard i.-r. vector model it was found possible to assign all the peaks within 2·2 Å of the origin of the (001) projection uniquely to their appropriate vectors. The orientation of the ring was thus immediately established. The ring is inclined to (001) at an angle of 56°.

Translation of the molecule

The method for determining the coordinates of the centre of the approximately symmetrical pyridine ring was that suggested by Patterson (1949) and used by McDonald & Beevers (1952) in the structure analysis of α-glucose. Peaks on the two Patterson projections corresponding to a coincidence of vectors between atoms of centrosymmetrically related rings were readily found, and approximate coordinates were assigned to the centre of the ring. It then remained only to determine the position of the oxygen atom, there being initially six possible points of attachment to the ring. The correct position was found by trial with the aid of structure-factor graphs.

4. Refinement of the structure

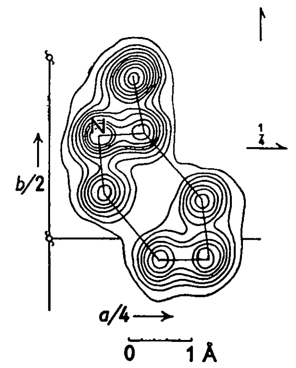
Preliminary refinement of coordinates was carried out by application of the (F_o-F_c) technique (Cochran, 1951) to both projections, using the eye-estimated intensity data but making no allowance for hydrogen atoms. The factor $R = \Sigma|F_o-F_c| \div \Sigma|F_o|$ was reduced to about 0·18 for both the *hk*0 and *h*0*l* zones. From that stage all refinement was made using the Geiger-counter measured intensities. Projections on (001) and (010) of the observed electron density are shown in Figs. 2 and 3.

Refinement of atomic coordinates, including those of hydrogen atoms, proceeded simultaneously with the assignment of temperature-factor parameters. Each atom was treated separately as far as its thermal vibration was concerned and the best temperature factor was deduced after successive approximation from maps of $(\varrho_o-\varrho_c)$, (D maps).

Hydrogen atoms were postulated for a first approximation as being attached to the four available carbon atoms of the pyridine ring at a distance of 1·05 Å from each, and so that each H–C bond would make equal angles with the two C–C (or C–N) bonds it meets, and would lie in the plane of the ring. The fifth hydrogen, which, from consideration of the relationship of N and O atoms about screw axes perpendicular to (001), is likely to be concerned in hydrogen bonding between them, was placed on the line joining this pair of atoms and at a distance of

1·0 Å from the nitrogen. From later evidence discussed in § 5 it was clear that these proposed hydrogen coordinates were correct to within 0·2 Å and they were used in all future refinement, adjusted slightly in accordance with shifts to the atoms of the ring.

Final 'greater than average' temperature factors for both projections are listed in Table 1, the form of the expression for anisotropic vibrations being that proposed by Hughes (1941). The two projections were treated independently and the discrepancies observed between the two sets of figures are indications that the final values are only approximations.



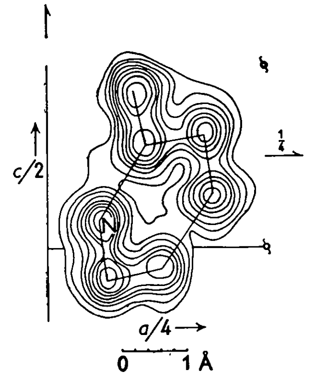


Fig. 2. Final electron-density projection of one molecule on (001). All terms up to $\sin\theta/\lambda = 0{\cdot}765$ have been included in the synthesis. Contours at an interval of 1·0 e.Å^{-2}, starting at 2·0 e.Å^{-2}.

Fig. 3. Final electron-density projection of one molecule on (010). Contours at an interval of 1·0 e.Å^{-2}, starting at 2·0 e.Å^{-2}.

Table 1. *Atomic temperature factors*

	Atom	'Greater than average' temperature factor
(001) projection	C_1	Average
	C_2	$\exp[-\{0{\cdot}4\sin^2(\varphi+26^\circ)\}\sin^2\theta]$
	C_3	$\exp[-\{0{\cdot}4\sin^2(\varphi+39^\circ)\}\sin^2\theta]$
	C_4	$\exp[-0{\cdot}2\sin^2\theta]$
	C_5	$\exp[-0{\cdot}2\sin^2\theta]$
	N	Average
	O	$\exp[-\{0{\cdot}2+0{\cdot}4\sin^2(\varphi+39^\circ)\}\sin^2\theta]$
(010) projection	C_1	Average
	C_2	Average
	C_3	$\exp[-\{0{\cdot}5\sin^2(\varphi+25^\circ)\}\sin^2\theta]$
	C_4	$\exp[-0{\cdot}3\sin^2\theta]$
	C_5	$\exp[-0{\cdot}3(\tfrac{1}{2}lc^*)^2]$
	N	Average
	O	$\exp[-0{\cdot}4\sin^2\theta]$

The fixed angle denoted is that between the direction of maximum vibration and the *b* axis (for the (001) projection) or the *c* axis (for the (010) projection) (Hughes, 1941).

Scaling of structure factors

The set of relative intensity figures derived from Geiger-counter measurements are proportional to the true integrated reflexion intensities (Cochran, 1950). The fall off with $\sin\theta$ of $\overline{G}_o$, the local average value of their square roots, can therefore be assumed to be due entirely to the normal angular fall off of the scattering factors of stationary atoms and to their thermal vibrations. The shape of the $\overline{G}_o/\sin\theta$ curve was therefore taken to be that of the average effective atomic scattering factor of the molecule, and structure factors F_s, calculated for atoms with their 'greater than average' temperature factor, if any, were scaled to fit this curve. When $\log(\overline{G}_o/\overline{F}_c)$ was plotted against $\sin^2\theta$, the points, as expected, fell very close to a straight line which had the same slope (1·10) for both zones. Thus $\overline{G}_o \approx C\exp[-1{\cdot}10\sin^2\theta]\overline{F}_s$. In Tables 8 and 9 are listed observed and calculated structure factors, where

$$F_o = G_o/C \quad \text{and} \quad F_c = F_s(\overline{F}_o/\overline{F}_s)\,.$$

Values of F_s were at first calculated using f's listed by James & Brindley (1931) but in the final stages were recalculated using the f's recently published by McWeeny (1951), who has given evidence that they are superior to the former. Bacon (1952) has also found that much better agreement with observed powder diffraction intensities from graphite is obtained using McWeeny's f curve for carbon. The f values used for oxygen were $f_{\mathrm{O}} = \frac{1}{3}(f^{\parallel}+2f^{\perp})$, $f^{\parallel}$ and $f^{\perp}$ being listed in McWeeny's table.

After rescaling F_s to G the only significant changes to be expected in the values of (F_o-F_c) are in the lowest-order spectra because it is only for values of $\sin\theta/\lambda <$ about 0·25 that there is a change in the *differences* between $f_{\mathrm{C}}, f_{\mathrm{N}}$ and f_{O} in going from the James & Brindley to the McWeeny figures. In fact the only marked change observed was in the agreement for the four lowest-angle reflexions for both zones. For the ($hk0$) zone, the decrease in $\Sigma|F_o-F_c|$ for these four was from 10·0 to 5·7, where the sum for all reflexions measured by the spectrometer with Cu radiation is 55·8. For the ($h0l$) zone the corresponding decrease was from 11·0 to 8·7.

Extinction

Towards the end of the refinement it was noticed that the values of $|F_o-F_c|$ for the two strongest reflexions in each zone ranged from five to nine times the r.m.s. value, and that in each case F_o was too small. Such large discrepancies must be due to systematic errors and it was assumed that the four reflexions concerned were all observed too low because of secondary extinction. The (F_o-F_c) values concerned were therefore given zero weight in deriving D maps, used for determining final coordinates. In all calculations on accuracy which involve numerical values of $|F_o-F_c|$, these reflexions were assigned the r.m.s. values.

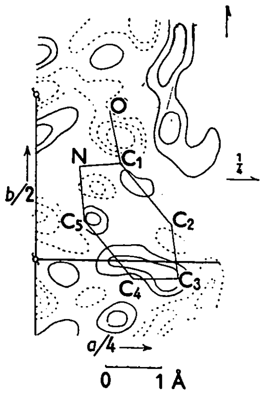
The values of the discrepancy factor R at the end of refinement, for each zone, and for different ranges of $\sin\theta/\lambda$, are listed in Table 2. Final D maps are shown in Figs. 4 and 5 and coordinates which produced them are listed in Table 3. In Figs. 6 and 7 are shown the electron distributions resulting from the subtraction of the calculated electron density of only C, N and O atoms from that observed. They do in fact represent the electron density due to hydrogen atoms in the structure (cf. Cochran, 1951) and will be referred to as D_H maps. Hydrogen coordinates in Table 3 were derived directly from them.

Table 2. *Final values of* $R = \Sigma|F_o-F_c| \div \Sigma|F_o|$

Range of $\sin\theta/\lambda$	$R(hk0)$	$R(h0l)$
0—0·5	0·058	0·059
0—0·65	0·065	0·067
0·65—0·765	0·29	—
0—0·765	0·085	—

Table 3. *Atomic coordinates*

Atom	x (Å)	y (Å)	z (Å)
C_1	1·522	1·742	1·621
C_2	2·454	0·647	1·760
C_3	2·568	−0·312	0·840
C_4	1·768	−0·333	−0·334
C_5	0·910	0·724	−0·496
N	0·795	1·684	0·425
O	1·344	2·653	2·437
H_1	3·04	0·69	2·50
H_2	3·46	−1·08	0·85
H_3	1·70	−1·08	−1·01
H_4	0·38	0·69	−1·32
H_5	0·00	2·31	0·38



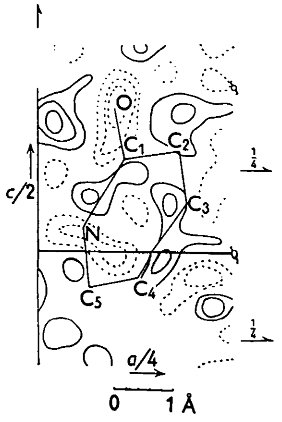
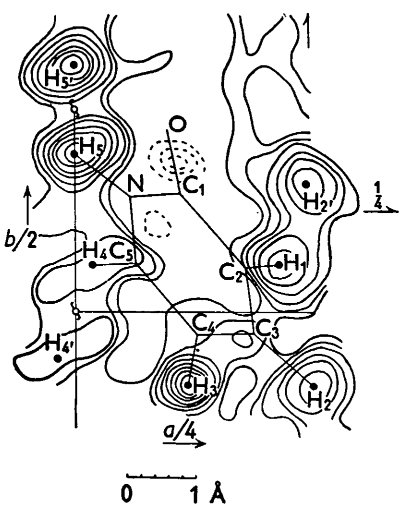


Fig. 4. Final D map showing $(\varrho_o-\varrho_c)$ projected on (001). Contours at an interval of 0·1 e.Å^{-2}, starting at 0·1 e.Å^{-2}; negative contours broken.

Fig. 5. Final D map showing $(\varrho_o-\varrho_c)$ projected on (010). Contours at an interval of 0·1 e.Å^{-2}, starting at 0·1 e.Å^{-2}; negative contours broken.



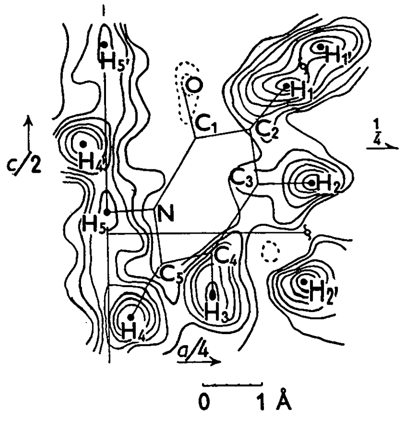


Fig. 6. Final D_H map projected on (001) representing the difference between the observed electron density and that calculated for isolated 'McWeeny' C, N and O atoms. Final assumed positions for H atoms are indicated by black dots. Contours at an interval of 0·1 e.Å^{-2}, starting at 0·2 e.Å^{-2}; negative contours broken.

Fig. 7. Final D_H map projected on (010). Description as for Fig. 6.

5. Accuracy of the analysis

The standard deviation (s.d.) in electron density $\sigma(\varrho_o)$ was estimated by the expression given by Cochran (1948) and the s.d. in the coordinates of C, N and O atoms, $\sigma(x_i)$, by the equation of Cox & Cruickshank (1948). Results were:

$$\sigma(\varrho_o) = 0{\cdot}14\ \text{e.Å}^{-2} \text{ for both projections.}$$

Hence the limit of error in electron density is taken to be $2{\cdot}5\sigma = 0{\cdot}35$ e.Å^{-2}. In fact the final D maps of Figs. 4 and 5 both reach a negative value of 0·40 e.Å^{-2} in the neighbourhood of the oxygen atom. This high value may be attributed to the use of incorrect temperature factors for this atom, for the value 0·35 e.Å^{-2} is exceeded nowhere else in the structure. For the s.d. in coordinates, two values for $\sigma(x)$ and one each for $\sigma(y)$ and $\sigma(z)$ were obtained for each type of atom. For each type, the difference between maximum and minimum values for these independent measurements was 0·0005 Å. The maximum value was used in each case for calculating s.d.'s in interatomic distances, and is denoted $\sigma(\bar{x}, \bar{y}, \bar{z})$.

$$\sigma(\bar{x}, \bar{y}, \bar{z}) = 0{\cdot}0073 \text{ Å for carbon,}$$
$$0{\cdot}0062 \text{ Å for nitrogen,}$$
$$0{\cdot}0064 \text{ Å for oxygen.}$$

Corresponding s.d.'s in interatomic distances were calculated to be 0·010±0·001 Å for C–C, C–N and C–O bonds. Hence we may fix the limit of error for all bond lengths as 0·025 Å, with a corresponding limit of error in bond angle of about 2°.

Hydrogen coordinates

In this discussion 'hydrogen coordinate' is taken to mean 'the point of maximum electron density of the hydrogen atom'.

The errors in hydrogen coordinates cannot be readily estimated by the method used for the heavier atoms because the rather irregular shape of the peaks in the D_H maps precludes a reliable estimation of their curvature, and there will be errors associated with determination of peak maxima which are not allowed for in the equation of Cox & Cruickshank. There do remain however the following ways of estimating the s.d.'s:

First, the x coordinates of the two projections may be compared. The maximum observed difference between the two projections for the five hydrogens is 0·16 Å, and the r.m.s. value is 0·11 Å. The corresponding r.m.s. error in coordinate is then given approximately by $0{\cdot}11/\sqrt{2} = 0{\cdot}08$ Å.

Secondly, C–H bond lengths may be computed from the observed coordinates and compared with standard values. In Table 4 are listed these calculated bond lengths and the N–H bond length, together with the deviations Δr from the standard lengths of 1·07 Å for C–H and 1·00 Å for N–H. These standards are the sums of covalent radii and are also the mean values of a large number of experimental electron-diffraction measurements which are listed by Allen & Sutton (1950). Considering only the C–H bonds (because we are not sure of the 'standard' N–H length in a hydrogen bond), the maximum value of Δr is 0·12 Å and the r.m.s. value is 0·097 Å ≈ 0·10 Å, which value we may regard to a first approximation as being due entirely to the error in hydrogen coordinate. Although, with only four observations, such figures cannot be referred to as 'standard deviations', nevertheless their magnitudes are a fair indication of the errors to be expected. We shall assume from these arguments a limit of error for both C–H and N–H bonds of 0·30 Å.

Table 4. *Bond lengths involving hydrogen atoms*

Bond	Length	Deviation (Δr) from standard length
C_2–H_1	0·95 Å	−0·12 Å
C_3–H_2	1·18	0·11
C_4–H_3	1·02	−0·05
C_5–H_4	0·98	−0·09
N–H_5	1·02	0·02

6. Description of the structure

In Fig. 8 are shown the bond lengths and bond angles of the α-pyridone molecule (excluding those in which hydrogen participates) calculated from the coordinates listed in Table 3.

Although the prydine ring departs considerably from regular hexagonal shape, it is, however, planar within experimental error. Representing the equation of the plane of the ring by $Ax+By+Cz = 1$, the best values

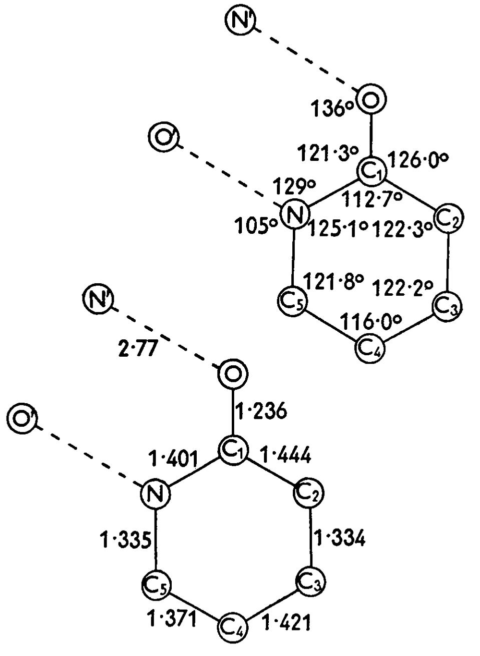


Fig. 8. Bond lengths and bond angles. O′ and N′ are the atoms to which N and O respectively are hydrogen-bonded. The drawings are not to scale.

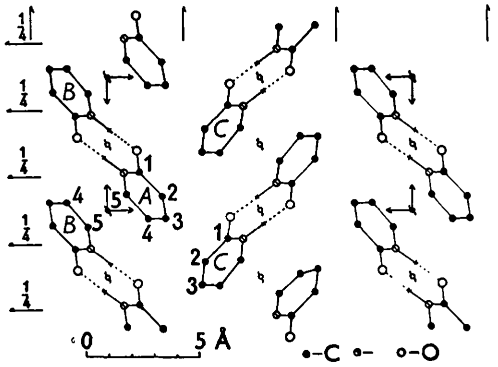
for A, B and C were found by least-squares solution of the equations $Ax_1+By_1+Cz_1 = 1$ etc. for all six atoms. They were, $A = 0{\cdot}5666$, $B = 0{\cdot}42305$, $C = -0{\cdot}3752$ when the atomic coordinates were expressed in Ångström units. The greatest perpendicular distance of any of the ring atoms from this plane is 0·017 Å and the r.m.s. value is 0·011 Å. The oxygen atom lies off the plane by 0·039 Å but the $O(C_1C_2N)$ group must be planar, for the sum of the three angles about C_1 is 360·0°.

The closest intermolecular approach distances, including some involving hydrogen atoms, are listed in Table 5 and the structure of the whole unit cell as viewed along the c and b axes is shown in Figs. 9 and 10 respectively. The only approaches short enough to require explanation other than in terms of van der Waals forces are those of 2·77 Å between N(A) and O(B) and 1·80 Å between H_5(A) and O(B). Evidently there must exist quite a strong hydrogen bond between N(A) and O(B). This bond, which is repeated throughout the crystal by the operation of the screw axes parallel to the c crystallographic axis, links individual molecules into endless helices.

Table 5. *Intermolecular distances*

N–O(B)	2·77 Å
H_5–O(B)	1·80
C_5–O(B)	3·37
H_4–O(B)	3·20
C_5–C_5(B)	3·67
H_4–C_5(B)	2·77
C_3–C_2(C)	3·73
H_2–C_2(C)	3·02
C_3–H_1(C)	3·30

The first named atom is in each case from molecule A (Figs. 9 and 10).



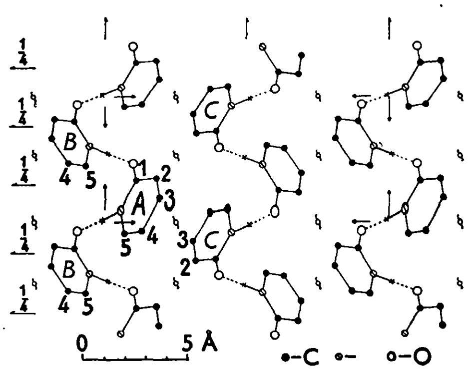


Fig. 9. Structure viewed along c axis. The H atom attached to N is represented by a cross, and the hydrogen bonds are represented by broken lines.

Fig. 10. Structure viewed along b axis. The H atom attached to N is represented by a cross, and the hydrogen bonds by broken lines.

7. Discussion

Distribution of electron density

The first point to be made about the evidence from final D and D_H maps is that there is no indication of any systematic concentration of electron density between bonded atoms. The average value of $(\varrho_o-\varrho_c)$ at the centre of bonds is $< 0{\cdot}1$ e.Å^{-2} in either projection and the residual density does in fact appear to be randomly distributed except for a trough in the neighbourhood of the oxygen atom which has already been mentioned. This would indicate that an assemblage of isolated atoms, each with an appropriate temperature factor and with a McWeeny electron distribution, is as close an approximation to the true distribution in the molecule as can be obtained with the experimental data available. Such an assemblage of individual atoms but with 'Hartree' electron distribution was found by Cochran (1951) similarly to explain the observed electron density in adenine hydrochloride. However, in a more recent structure determination of salicylic acid in which a more detailed allowance was made for anisotropic thermal motion in one projection, Cochran (1953) found a significant electron displacement away from the centres of the benzene ring towards the centres of bonds to the extent of about 0·05 e. per bond. The McWeeny f's correspond to a greater spread in the electron density of carbon atoms relative to oxygens and nitrogens than the James & Brindley f's (derived from 'Hartree' atoms) so that D maps derived from these different sets are not directly comparable. However, the very fact that the α-pyridone molecule is inclined considerably to both (001) and (010) would tend to obscure fine features of the electron density in space when projected on to these planes. It is therefore unwise to draw conclusions about electron distribution in bonds from this structure until a full three-dimensional analysis has made possible a determination of the distribution in the plane of the molecule and in parallel planes.

An additional check of the general correctness of our representation of the molecular electron density may be obtained by consideration of the number of electrons associated with each hydrogen atom in the D_H maps, and the corresponding peak values of density. The electron distribution in an isolated hydrogen atom in this structure may be calculated by forming the two-dimensional Fourier transform of

$f_H \exp [-1·10 \sin^2 \theta]$. The result indicates that the theoretical peak value of electron density is 0·78 e.Å^{-2} and that only about 5% of the projected 'electron' lies outside a circle of radius 1·0 Å drawn about the atomic centre (Cochran, 1951). Such circles were described on D_H maps and the electron density in the areas formed was integrated in the way Cochran describes. The results for the number of electrons in each such area n(H), and the peak values of observed electron density ϱ_m(H) are listed in Table 6. The r.m.s. value of the electron count for one atom is 0·94 e. for the (001) and 0·92 e. for the (010) projection. Corresponding r.m.s. values for the peak electron density are 0·80 and 0·76 e.Å^{-2} respectively. Such close agreement of these mean values with theory is fortuitous, but the result supports the other indications of the general correctness of the assumed electron distribution in carbon and nitrogen atoms.

Table 6. *Numbers of electrons associated with hydrogen atoms and the peak values of hydrogen electron density*

	Atom	(001) projection	(010) projection	Mean
n(H) (electrons)	H_1	0·94	0·96	0·95
	H_2	1·01	0·77	0·89
	H_3	0·91	0·96	0·94
	H_4	0·85	1·03	0·94
	H_5	0·97	0·84	0·91
ϱ_m(H) (e.Å^{-2})	H_1	0·75	0·85	0·80
	H_2	0·80	0·75	0·78
	H_3	0·90	0·70	0·80
	H_4	0·50	0·90	0·70
	H_5	1·00	0·55	0·78

The hydrogen bond and tautomerism

N(A)–H_5–O(B), as we have seen, must be a hydrogen bond. The tautomeric form of the molecule then depends upon the precise location of H_5 between the other two atoms. The two interatomic distances are:

$$N(A)\text{–}H_5 = 1·02 \text{ Å}, \quad O(B)\text{–}H_5 = 1·80 \text{ Å} .$$

This difference of 0·78 Å is highly significant and it is certain that if H_5 is covalently bonded to either of N(A) or O(B), it must be to N. Now if H_5 were at the mid-point of the bond, its distance from N(A) would be 1·38 Å, a difference of 0·36 Å from that observed, which is greater than 0·30 Å, the limit deduced from a consideration of random errors. Therefore the evidence is conclusive that H_5 is non-symmetrically situated with respect to N(A) and O(B), and it is very likely that N–H_5 is close to the normal covalent bond length of 1·00 Å. This means that in the crystalline state, the stable form of the molecule is that of a pyridone (structure I, Fig. 1). The sum of the three angles ∠ H_5–N–C_1, ∠ H_5–N–C_5, ∠ C_1–N–C_5 is 358°, in satisfactory agreement with the expected value of 360° for a planar disposition of bonds about N.*

Resonance structure

There are a number of ways of writing the pyridone formula, involving different arrangements of valence bonds, the most likely forms being shown in Fig. 11. All forms except A have a positive charge on the nitrogen atom to allow for the increase of its covalency to four, and a negative charge on an oxygen or a carbon atom. Because the electronegativity trend is in the order O > N > C, structures B and C are expected to have much greater weights than those of D and E.

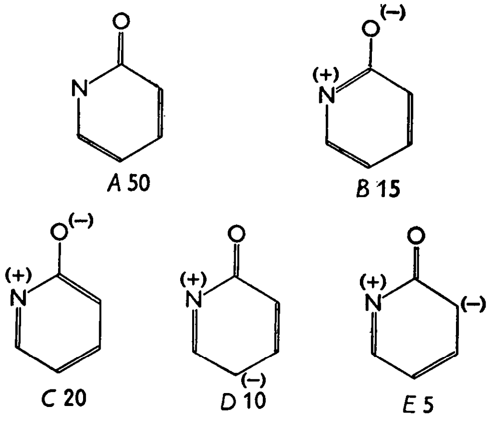


Fig. 11. Resonance contributors to the structure. Set (b) of the weights assigned to each are indicated as percentages.

For any particular set of weights assigned to the forms listed, the double-bond character of each bond in the resultant resonance structure may be read off and its bond length calculated from the equation

$$R = R_1 - (R_1 - R_2).3x/(2x+1) \quad \text{(Pauling, 1940)} ,$$

where R is the observed bond length, R_1 and R_2 are standard single- and double-bond lengths respectively, and x is the bond order expressed as the fraction of double-bond character.

For C–N, R_1 and R_2 were assigned values 1·48 and 1·24 Å respectively. For C–O, R_1 and R_2 were assigned

Table 7. *Observed bond lengths and those calculated for two resonance structures*

Structure (a): $A = 60\%$, $B = 15\%$, $C = 25\%$
Structure (b): $A = 50\%$, $B = 15\%$, $C = 20\%$, $D = 10\%$, $E = 5\%$

Bond	r_o	r_c from (a)	r_c from (b)
C_1–O	1·24 Å	1·24 Å	1·23 Å
C_1–C_2	1·44	1·44	1·45
C_2–C_3	1·33	1·35	1·355
C_3–C_4	1·42	1·44	1·44
C_4–C_5	1·37	1·35	1·37
C_5–N	1·335	1·36	1·33
N–C_1	1·40	1·40	1·40

---

* As H_5 is on a screw axis perpendicular to (001), there is an ambiguity between two possible values for its height relative to this plane, separated by $\frac{1}{2}c = 2·84$ Å. However, if it were moved by $\frac{1}{2}c$ from the position chosen, O–H and N–H distances would be 1·59 Å and 2·05 Å respectively, so that no covalent bond would be made at all, which is impossible.

Table 8. *Observed and calculated structure factors for the hk*0 *zone*

All spectra below the horizontal line in each group for which *k* is constant, lie outside the sphere of reflexion for Cu *Kα* radiation and were not considered in final stages of the refinement

h	*k*	F_o	F_c
0	0	—	200
2		15·4	16·1
4		41·6	−47·7
6		18·1	−19·3
8		15·4	−14·5
10		8·1	7·9
12		2·5	− 2·6
14		6·3	− 5·1
16		9·3	9·2
18		4·1	2·9
20		1·6	− 0·2
1	1	9·8	9·6
2		43·3	−48·0
3		32·4	−32·3
4		35·1	−35·8
5		31·6	33·4
6		11·5	−11·3
7		3·0	2·9
8		6·0	6·2
9		13·6	−13·9
10		7·3	7·3
11		2·5	− 2·2
12		2·2	1·1
13		2·9	− 2·5
14		3·4	3·5
15		6·9	6·7
16		4·4	4·5
17		2·6	3·2
18		0·8	− 3·3
19		4·7	− 5·0
20		3·8	− 3·6
0	2	14·1	13·9
1		13·0	13·2
2		17·2	−16·8
3		18·8	18·8
4		13·4	−12·2
5		7·4	− 7·7
6		8·8	8·5
7		16·7	−17·6
8		3·1	3·2
9		3·6	2·7
10		9·3	9·8
11		8·5	8·4
12		2·6	2·2
13		0	− 0·4
14		1·6	− 1·7
15		1·4	− 1·1
16		1·3	− 1·0
17		1·5	− 1·3
18		4·5	− 4·5
19		1·3	− 1·7
20		0·9	− 0·8
1	3	2·7	2·0
2		3·1	2·2
3		2·1	− 2·3
4		3·4	− 3·0
5		2·2	0·9
6		2·9	− 1·9
7		3·4	3·4
8		5·6	5·1
9		1·3	0·9
10		6·9	7·0
11		4·0	− 4·7
12		2·1	− 2·2
13		3·3	− 3·6
14		6·7	− 6·7
15		2·1	2·2
16		0	0·2
17		0·6	1·5
18		1·1	1·3
19		1·0	0·6
0	4	0·8	− 0·1
1		5·9	5·8
2		3·2	2·9
3		0·8	− 0·5
4		4·9	− 4·6
5		8·2	− 8·0
6		4·4	− 4·6
7		2·2	− 3·7
8		6·3	5·8
9		1·6	− 1·7
10		1·3	1·5
11		3·7	3·5
12		4·3	− 3·6
13		2·1	1·8
14		1·0	− 0·7
15		1·5	1·9
16		1·7	− 1·1
17		2·3	2·4
18		0·7	0·8
1	5	13·0	−14·1
2		2·3	2·2
3		4·2	4·0
4		8·8	− 8·5
5		3·9	3·9
6		3·5	2·4
7		4·5	3·9
8		4·1	3·6
9		3·8	3·5
10		1·4	− 0·6
11		1·3	− 0·3
12		0·6	− 0·4
13		2·7	2·5
14		4·0	− 3·7
15		0	− 1·0
16		2·2	1·9
17		5·8	− 5·1
0	6	6·3	− 5·9
1		13·8	12·8
2		0	0·7
3		12·5	11·4
4		1·5	2·0
5		1·4	0·8
6		1·5	− 1·1
7		0	0
8		2·2	2·1
9		4·8	− 4·2
10		0	0·1
11		1·3	0·9
12		0·9	− 0·6
13		0·5	− 1·1
14		1·1	0·5
15		4·9	− 4·7
1	7	4·9	4·8
2		2·3	2·3
3		2·4	2·6
4		3·1	− 3·1
5		4·4	− 4·1
6		3·3	3·4
7		4·2	− 4·0
8		2·2	2·1
9		0	− 0·6
10		0·9	− 2·5
11		1·1	− 0·2
12		0·7	− 0·8
13		0·5	0·1
0	8	3·4	− 3·1
1		—	0·8
2		0	0·4
3		2·2	− 2·4
4		3·4	3·0
5		2·0	− 2·2
6		0·5	0·3
7		0	1·0
8		2·5	− 2·4
9		0	− 0·8
10		1·3	− 1·6

values 1·42 and 1·20 Å respectively. (These are the lengths used by Vaughan & Donohue (1952) in assessing double-bond character in a number of carboxylic acids and peptides. They obtained excellent agreement between observed and calculated values.) Calculated C–C bond lengths were obtained directly from Pauling's curve.

In Table 7 observed bond lengths are compared with those calculated from each of the two sets of weights which give best agreement with the observed lengths, (*a*) with only structures *A*, *B* and *C* included, and (*b*) with all five structures included. If a bond length is denoted by *r*, the r.m.s. value of $|r_o - r_c|$ for the (*a*) set is 0·015 Å and for the (*b*) set 0·011 Å. Both are of the order of the s.d. in observed bond length and no better agreement can be expected. In fact such agreement may be taken as much as a justification for the use of Pauling's equation (which is necessarily an approximation) as for the accuracy of the observations.

Table 9. *Observed and calculated structure factors for the h0l zone*

F_c values are referred to an origin displaced by $-\frac{1}{4}a$ from that for the *hk*0 zone

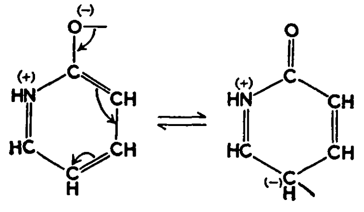
h	*l*	F_o	F_c
0	0	—	200
2		16·8	−16·2
4		44·3	−49·5
6		18·9	20·7
8		15·7	−14·9
10		8·2	− 9·2
12		2·3	− 2·5
14		5·5	5·8
16		10·0	9·5
1	1	15·2	−15·1
2		36·8	44·8
3		16·8	16·8
4		15·9	15·2
5		14·8	−15·2
6		5·7	5·3
7		18·9	19·8
8		8·2	8·8
9		5·2	− 5·3
10		6·4	− 5·7
11		9·5	− 9·2
12		13·2	−13·1
13		13·6	13·9
14		3·0	− 2·7
15		3·2	− 3·1
16		5·7	5·5
17		4·3	− 5·2
0	2	13·2	13·8
1		23·2	22·8
2		22·0	−22·7
3		15·5	16·3
4		4·8	− 5·5
5		26·1	−27·7
6		16·8	17·3
7		5·0	− 4·7
8		1·4	− 0·2
9		10·5	10·2
10		7·3	6·8
11		3·0	− 4·1
12		0·9	− 1·2
13		0·5	0·8
14		6·6	− 6·5
15		2·3	2·3
16		0	0·2
1	3	2·7	− 2·8
2		1·8	0·3
3		7·5	7·5
4		9·1	− 9·7
5		6·4	5·6
6		10·2	8·8
7		1·1	− 1·4
8		2·7	3·2
9		3·6	− 3·4
10		4·1	− 4·0
11		0	0·2
12		1·1	− 0·9
13		0	− 0·5
14		2·7	2·0
15		2·3	− 2·4
0	4	8·0	− 7·2
1		2·7	2·0
2		0·7	0·2
3		3·2	3·5
4		2·3	− 2·4
5		8·6	− 8·5
6		1·1	− 1·0
7		1·1	− 0·8
8		9·1	9·2
9		5·9	6·8
10		0·9	− 1·1
11		0	0
12		2·0	− 1·5
13		2·3	1·8
14		1·6	1·4
1	5	8·2	8·8
2		0·7	0·1
3		6·8	− 6·8
4		9·3	−10·3
5		0·9	0·7
6		1·6	− 1·3
7		4·1	− 4·3
8		4·3	3·2
9		2·7	− 2·4
10		0·9	0·5
11		2·3	2·2
12		0	0·6
0	6	7·3	− 7·3
1		5·5	5·5
2		8·0	8·0
3		8·0	8·0
4		4·1	3·8
5		1·6	− 1·6
6		8·2	− 8·0
7		2·3	− 1·9
8		0·9	− 0·3
9		2·3	2·8
10		2·0	− 2·4
1	7	3·0	− 2·5
2		1·6	1·5
3		2·3	− 1·4
4		1·6	1·3
5		3·2	3·8

Set (*b*) is to be preferred to set (*a*) because it requires a difference between the lengths of C_1–C_2 and C_3–C_4, and between C_2–C_3 and C_4–C_5, both of which pairs must be equal if *D* and *E* do not contribute. The observed difference of 0·037 Å between the lengths of the second pair is significant and indicates that *A*, *B* and *C* alone do not describe the structure completely. It may be observed that in neither set of weights do *B* and *C*, or *D* and *E* contribute equally. At first sight it might appear that the members of each of these pairs are equivalent forms, but in fact they are not, because of the non-symmetrical nature of the molecule, i.e. the bond C_1–N has a different environment from that of N–C_5.

The main facts which emerge from the above considerations are that the non-polar form of the molecule (*A*) is the most important resonance contributor in the crystalline state but that the importance of polar forms with oxygen bearing a negative charge is considerable. The existence of these latter forms has the effect of bestowing a high polarity on the N–H ⋯ O bond. Goldschmidt & Llewellyn (1950) found by similar reasoning that the observed bond lengths of isatin required non-polar forms to contribute to the extent of about 65%.

Relationship between observed structure and chemical properties

At the positions of the pyridine ring which are the most reactive with electrophilic reagents, small local concentrations of electron density are to be expected, but our measurements are not sufficiently accurate to reveal them directly if they do exist. However, the relative weights of contributing resonance forms, as derived from precise bond-length measurements, are a valuable means of deducing any unequal distribution of electrons among atoms. Thus the probable contribution, even in small amounts, of forms *D* and *E* means that, of all five carbon atoms, C_2 and C_4 have electrons most readily available for a reaction with an electrophilic reagent. They may be regarded as derived from forms *B* and *C* by operation of the 'mesomeric effect' described, e.g., by Remick (1943) in the following way:



In fact, halogenation and nitration of the ring is observed to occur most readily at the *ortho* and *para* positions with respect to the C=O group, i.e. at C_2 and C_4. Disubstitution in these positions is possible but it is very difficult to go beyond this stage. This structure analysis may therefore be said to provide direct evidence in one particular case of the existence of an effect which has been postulated to explain the reactivity of certain substituted aromatic compounds. Furthermore, the evidence is that this effect exists separately from the chemical reaction itself, and is not entirely brought about by the approach of reacting groups.

I should like to thank Prof. Sir Lawrence Bragg and Dr W. H. Taylor for facilities provided for this work and for their interest in it. I am greatly indebted to Dr W. Cochran for his encouragement and for much helpful advice. I also wish to thank Mrs L. Remnant for assistance with computation. I acknowledge gratefully the award of post-graduate scholarships by the University of New Zealand and the British Council during the tenure of which this work was carried out, and thank the Council of Canterbury University College, New Zealand, for the granting of three years leave of absence.

References

Albrecht, G. (1939). *Rev. Sci. Instrum.* **10**, 221.
Allen, P. W. & Sutton, L. E. (1950). *Acta Cryst.* **3**, 57.
Bacon, G. E. (1952). *Acta Cryst.* **5**, 492.
Cochran, W. (1948). *Acta Cryst.* **1**, 54.
Cochran, W. (1950). *Acta Cryst.* **3**, 268.
Cochran, W. (1951). *Acta Cryst.* **4**, 81.
Cochran, W. (1953). *Acta Cryst.* **6**, 260.
Cox, E. G. & Cruickshank, D. W. J. (1948). *Acta Cryst.* **1**, 92.
Elderfield, R. C. (1950). *Heterocyclic Compounds.* New York: Wiley.
Goldschmidt, G. H. & Llewellyn, F. J. (1950). *Acta Cryst.* **3**, 294.
Hughes, E. W. (1941). *J. Amer. Chem. Soc.* **63**, 1737.
James, R. W. & Brindley, G. W. (1931). *Phil. Mag.* (7), **12**, 81.
McDonald, T. R. R. & Beevers, C. A. (1952). *Acta Cryst.* **5**, 654.
McWeeny, R. (1951). *Acta Cryst.* **4**, 513.
Patterson, A. L. (1949). *Acta Cryst.* **2**, 339.
Pauling, L. (1940). *The Nature of the Chemical Bond*, 2nd ed. Ithaca: Cornell University Press.
Remick, A. E. (1943). *Electronic Interpretations of Organic Chemistry.* New York: Wiley.
Vaughan, P. & Donohue, J. (1952). *Acta Cryst.* **5**, 530.

Acta Cryst. (1953). **6**, 600

The Strength of the 10 Å Reflexions in Haemoglobin

By F. H. C. Crick

The Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems, The Cavendish Laboratory, Cambridge, England

(*Received* 14 *February* 1953)

A method is described for calculating the total intensity of X-ray reflexions at spacings near 10 Å that one would expect from models of a protein consisting mainly of α-helices, not necessarily parallel. A comparison with Perutz's observed data for haemoglobin shows that the observed intensity is compatible with such models. The comparison can be made only with moderate accuracy.

Introduction

All recent work (see Crick (1952*b*) for references) has confirmed that the structure of the synthetic polypeptides is based on the α-helix of Pauling, Corey & Branson (1951). Very recently, Crick (1952*b*) has pointed out that packing considerations might deform an α-helix so that it became a coiled-coil, and simultaneously Pauling & Corey (1953) have produced a definite model of this type for α-keratin (though based on different assumptions). This model explains qualitatively the main features of the α-keratin pattern, namely the 1·5 Å and 5·15 Å meridian reflexions, and the reflexions of about 10 Å on and near the equator. It is thus probable that the fibrous proteins of the α type are largely based on the α-helix.

The structure of the globular proteins is less satisfactory. Pauling & Corey (1951*b*) originally suggested that the radial distribution function of haemoglobin, which they derived in an approximate manner from Perutz's published data (Perutz, 1949), appeared to resemble that expected from the α-helix. This approach has been taken considerably further by Riley & Arndt (1952), who have compared the observed radial distribution function for bovine serum albumin in the form of a dry powder, with Pauling & Corey's