Acta Cryst. (1954). **7**, 313

The Crystal Structure of Uracil

By G. S. Parry*

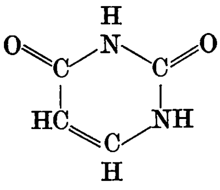
Department of Inorganic and Structural Chemistry, The University, Leeds 2, England

(*Received* 5 *November* 1953)

Uracil crystallizes in the monoclinic system with cell dimensions $a = 11{\cdot}82$, $b = 12{\cdot}35$, $c = 3{\cdot}62$ Å, $\beta = 120°$ and space group $P2_1/a$. The structure was solved by trial and refined by successive three-dimensional differential syntheses, a correction being applied to allow for series termination effects. The molecules are held together in layers by NH · · · O hydrogen bonds of length 2·81 and 2·86 Å. The planar molecule is found to contain two different C–N distances, two bonds having a probable length of 1·379±0·006 Å and the other two, of 1·342±0·008 Å. The two C–C bonds are not significantly different and have a probable length of 1·409±0·009 Å. The probable length of the C–O bonds is 1·236±0·007 Å. Attention is drawn to the close similarity between the dimensions of the pyrimidine ring in adenine hydrochloride and in uracil.

Introduction

In the last few years, detailed studies have been made of the crystal structures of both the purine bases obtained from the hydrolysis of nucleic acids (adenine HCl (Broomhead, 1948; Cochran, 1951) and guanine HCl (Broomhead, 1951)). No similar studies have yet been reported for the pyrimidine bases uracil, thymine and cytosine, which are also product of hydrolysis, although some information about cytosine can be inferred from the results of Furberg (1950) for cytidine. The present work was undertaken to provide accurate information regarding the crystal structure and molecular dimensions of uracil.



(I)

Gilpin & McCrone (1950) report that uracil (I) crystallizes in the monoclinic system with cell dimensions $a = 11{\cdot}4$, $b = 12{\cdot}38$, $c = 3{\cdot}63$ Å, $\beta = 113°$. The crystals twin easily on (100), and $(\bar{3}01)$ is a good cleavage. The optical birefringence is large (0·43) and negative. Considerable difficulty was found in recrystallizing uracil and in obtaining untwinned specimens. Sublimation at 300° C. was the only method found to give crystals of a reasonable size, and many of these were twinned.

Preliminary investigation

The present work has confirmed those findings reported above, particularly the difficulty in obtaining suitable crystals, which was thought at one time to preclude any further detailed examination. Recrystallization from a range of solvents and under a variety of conditions produced no crystals of a suitable size, and, although the first attempts at sublimation gave larger crystals, these were still flat plates about 0·05 mm. thick and all were twinned in the manner described by Gilpin & McCrone. Variation of the sublimation temperature showed that the best crystals were grown in the temperature range 290–310° C. very close to the melting point 338° C. Successive sublimates were examined for crystals of a suitable thickness (0·1 mm.) which were untwinned and gave sharp optical extinction directions. X-ray diffraction photographs of those crystals which satisfied these conditions showed that in many cases the spots were multiple or were extended along arcs, indicating that these crystals were imperfect. In one extreme case a crystal that showed perfect optical extinction on (010) gave a photograph on which appeared a few sharp spots, some wide arcs and a very heavy background scatter, indicating that there was considerable disorder within the crystal.

As a result of this search, however, one crystal was found that gave a very good X-ray photograph with no tendency for the spots to develop into arcs at high ϱ values. The shape of the crystal was not particularly suitable for accurate intensity work and it was twinned, although it proved possible to reduce the size of one of the twins considerably by cutting. As judged by subsequent diffraction photographs, the volume of one twin was only about 1% of that of the larger crystal and no difficulties were encountered in indexing the photographs. The crystal was a plate of about 0·1 mm. in thickness and about 0·2 mm. in the other two directions. It was therefore well below the optimum size ($\mu = 12{\cdot}1$ cm.$^{-1}$, $1/\mu = 0{\cdot}8$ mm.) and absorption effects were neglected in the subsequent analysis. This crystal was used to collect all the intensity data for uracil and also to measure its cell dimensions, which

---

* Now at University of Pittsburgh, Pittsburgh 13, Pa., U. S. A.

were obtained from high-order reflexions on zero layer lines of oscillation photographs. The Straumanis method of film loading was used for the cell-dimension measurements, and photographs were taken with Cu $K(\lambda_{\alpha_1} = 1{\cdot}5405$ Å) and Ni $K(\lambda_{\alpha_1} = 1{\cdot}6578$ Å) radiations. The nickel radiation was employed to increase the sin θ value of some high-order reflexions. No attempt was made to determine the axial lengths by extrapolation techniques. Zero-level Weissenberg photographs about the b axis were used to measure β. The cell dimensions obtained were

$$a = 11{\cdot}82 \pm 0{\cdot}01,\quad b = 12{\cdot}35 \pm 0{\cdot}01,\quad c = 3{\cdot}62 \pm 0{\cdot}005 \text{ Å},$$
$$\beta = 120 \pm 0{\cdot}5^\circ.$$

The $h0l$ and $0k0$ reflexions were found to be absent for h and k odd respectively so that the space group is $P2_1/a$. The density measured by flotation in a mixture of liquids is 1·617 g.cm.$^{-3}$ so that there are 4 molecules (calculated 3·98) in the unit cell.

The cell chosen differs from that of Gilpin & McCrone in the direction of the a axis. A result of this change is to transform the indices of the cleavage plane from $(\bar{3}01)$ to (001). All the optical and physical properties of uracil suggest that it is a layer structure with the layers oriented parallel to the new (001); the cell given above is therefore the most suitable for a detailed crystallographic study.

The solution and refinement of the structure

The diffraction spots about the a and c axes were recorded on Weissenberg photographs with Cu $K\alpha$ radiation ($\lambda = 1{\cdot}542$ Å), using the multiple-film technique. The intensity of each spot was estimated visually. No photographs for intensity purposes were taken about the b axis for the shape of the crystal was not altogether suitable and only a few reflexions were observed with ϱ values greater than 1·55. This rather sharp cut-off in the difffraction intensities leads to a severe reduction in the amount of information which can be obtained, and is due to a high temperature factor in a direction normal to the molecular layers. There also appeared to be rather heavy background scatter, which probably indicates that even this crystal was not entirely free from the internal disorder mentioned earlier. Of about 950 possible independent reflexions (440 within a sphere of radius 1·55 r.u.) only 303 were actually observed and, as will be seen later, this resulted in relatively large standard deviations of the atomic parameters, especially in the z direction.

An attempt was made to put the observed intensities on an absolute scale by the Wilson method (1942). It failed because it proved impossible to get a straight-line plot of ln $(\Sigma F^2/\Sigma f^2)$ against $\sin^2\theta$ by dividing the observed F^2 into any reasonable number of groups (values of Hartree f functions were taken from *Internationale Tabellen* (1935), vol. 2). Fig. 1 shows the mean intensity F^2 for each of six groups plotted against the mean value of $\sin^2\theta$ for the group. Each group contained upwards of 40 reflexions, and absent reflexions were given an intensity equal to half the minimum observable. There is a noticeable maximum in the curve at a $\sin^2\theta$ value of 0·425, which corresponds to a spacing of 1·2 Å. It is easy to show qualitatively that the reason for the failure of the Wilson statistics was that the 32 atoms in the unit cell were not arranged at random, but fell roughly on a plane hexagonal network. One might expect the Fourier transform of such a structure to bear some resemblance to the molecular transform of benzene and, in particular, to have an excess of strong planes of about 1·2 Å spacing corresponding to the first ring of high peaks in the benzene transform (Knott, 1940). Under these circumstances it was not possible to derive reasonable scale and temperature factors until calculated F values for an approximately correct structure had been obtained.

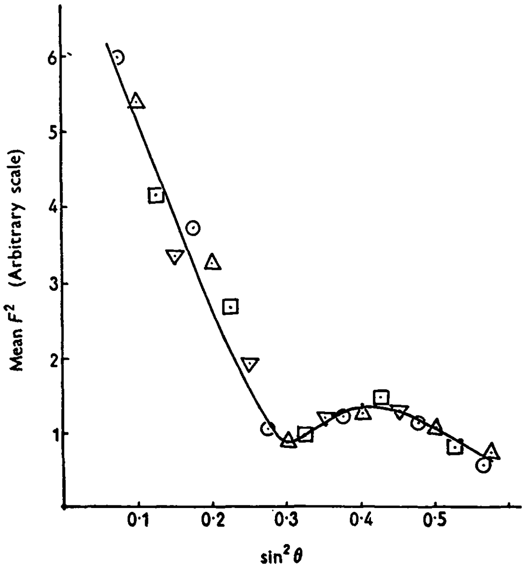


Fig. 1. The mean value of F^2 averaged over a range of 0·1 in $\sin^2\theta$ plotted against $\sin^2\theta$. The initial mean value of $\sin^2\theta$ in the four overlapping ranges is as follows:

○ $\sin^2\theta = 0{\cdot}075$, △ = 0·100, □ = 0·125, ▽ = 0·150.

The preliminary investigation established that in the structure of uracil the molecules are arranged in sheets parallel to (001), and the great intensity of the $00l$ reflexions confirmed this. As the z co-ordinates in the space group $P2_1/a$ must be either z or $\bar{z}$, layer structures of this type are possible only if $z \simeq 0$ or $\frac{1}{2}$. These two possibilities are structurally equivalent and the first was chosen. The problem was then a two-dimensional one: to determine the position and orientation of the molecule in this sheet. Models of the molecule were drawn to scale on cardboard using interatomic distances based on values found in other pyrimidine derivatives (Pitt, 1948; Clews & Cochran, 1948, 1949), and these were arranged on a sheet of paper to give a plausible structure. It was noticed that

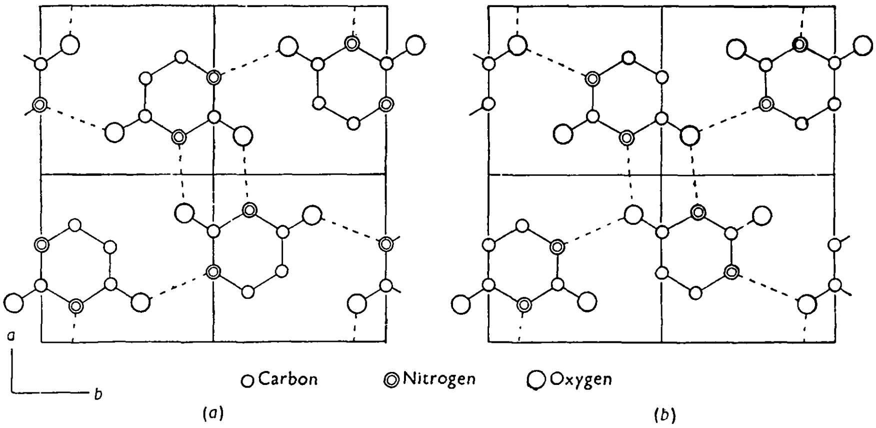


Fig. 2. The two trial structures used in the analysis. (*a*) The structure assumed initially in which each oxygen atom makes one hydrogen bond to a nitrogen atom. (*b*) The correct structure, as indicated by the refined electron-density maps, in which one oxygen atom makes two hydrogen bonds to nitrogen atoms while the other makes no hydrogen bonds at all.

the planes 360 (1·91 Å spacing) and 0,10,0 (1·23 Å) were the strongest high-order reflexions on the *hk*0 zone and that although 200 and 400 were very strong there were no strong high-order *h*00 reflexions. This suggested that one side of the pyrimidine ring was roughly parallel to the *a* axis. The position of the molecule in the cell was found by ensuring that two NH · · · O hydrogen bonds of length about 2·9 Å were formed from any one molecule to its neighbours and that no atom approached closer than 1·5 Å to a centre of symmetry. These considerations lead to the structure shown in Fig. 2(*a*). Trial sets of *hk*0 structure factors, which were calculated on the assumption that all ring atoms were carbon, confirmed that this was basically the correct structure, and an electron-density map was computed with the calculated phases. This map resolved all the atoms in the molecule and showed that alternate atoms in the ring were higher than the rest. Two of these atoms were nitrogen atoms on the existing model but the third large peak, which was 20% higher than the other carbon peaks, could not be explained. Because of this, the structure factors for the next stage of refinement were again computed with all the ring atoms of the same scattering power and this time the corresponding electron-density map was unambiguous. The two nitrogen atoms had peak heights about 25% greater than those of the four carbon atoms. The position of one of these nitrogen atoms corresponded not to the model in Fig. 2(*a*) but to that in Fig. 2(*b*), where the molecule has now been inverted about a line parallel to the *a* axis. As a consequence of this change one oxygen atom makes two hydrogen bonds to neighbouring molecules, while the other forms no hydrogen bonds at all.

From this point the refinement of this projection proceeded smoothly, the scale and temperature factors being derived at each stage by plotting ln $(\Sigma F_c/\Sigma F_o)$ against $\sin^2\theta$. The final alterations in the atomic parameters were made using difference maps computed with (F_o-F_c) as coefficients. These also gave some indication of hydrogen atoms attached to the carbons and to one of the nitrogens, but the position of the fourth hydrogen was uncertain. The other nitrogen atom formed a hydrogen bond to an oxygen across a centre of symmetry and the hydrogen peak in the difference map was so diffuse that at this stage it was impossible to decide to which of those atoms the hydrogen was attached. The contribution to the calculated structure factors of these hydrogen atoms whose position could be determined improved slightly the agreement with the observed structure factors.

An attempt was next made to obtain the *z* coordinates from the other two projections, but it soon became obvious that such a procedure could not give accurate results because most of the calculated structure factors were relatively insensitive to small changes in the *z* parameters. Only those 0*kl* reflexions with *k* odd were at all sensitive and as this group contained only eight measured intensities there was little difficulty in adjusting the *z* parameters to give reasonable agreement. However, as it was known that all the *z* parameters must be approximately zero it proved possible to use the *hkl* data directly.

The expression for $F(hkl)$ in the space group $P2_1/a$ is given by

$$F(hkl) = \pm\, 4 \sum_i f_i \begin{matrix}\cos\\ \sin\end{matrix} 2\pi(hx_i+lz_i) \begin{matrix}\cos\\ \sin\end{matrix} 2\pi k y_i\,.$$

As z_i is approximately zero we can write

$$\cos 2\pi l z_i = 1\,,$$
$$\sin 2\pi l z_i = 2\pi l z_i\,;$$

and expanding the trigonometrical terms we get

$$F(hkl) = F(hk0) - 8\pi l \sum_i f_i z_i \begin{matrix}\sin\\ \cos\end{matrix} 2\pi h x_i \begin{matrix}\cos\\ \sin\end{matrix} 2\pi k y_i\,.$$

Thus, provided z_i is small, the phase of $F(hkl)$ will be the same as that of $F(hk0)$, and in any case the magnitude of $F(hkl)$ should alter smoothly with systematic variation of the l index. By plotting $F(hkl)$ against l it was found possible to infer about 90% of the hkl phases in terms of the known $hk0$ phases. Planes for which a phase could not be deduced included $F(0kl)$ with k odd, $F(hkl)$ where the corresponding $F(hk0)$ was absent, and several planes for which the method was not sufficiently sensitive.

Electron-density line syntheses were next computed through the centre of the two oxygen atoms and the carbon atom C_5, using those planes for which the phases had been determined by the above method. The maxima along these lines gave the z parameter of the atoms and enabled the equation of the molecular plane to be established from which all other z parameters could be obtained. Structure factors calculated for these parameters gave an R factor of 0·28. This large disagreement was due in part to the presence of anisotropic thermal motion of the molecule, the isotropic temperature factor used on the c-axis projection being inadequate to allow for this. A correction for this anisotropy was therefore made before proceeding to further refinement. As all the molecules were effectively in the plane $z = 0$ (the mean displacement from this plane was about 0·1 Å) it was assumed that there were two principal temperature factors, B_2 normal to (001) and B_1 in the plane (001), and that the temperature factor B for a plane whose normal made an angle φ with the normal to (001) was given by

$$B = B_1 + (B_2 - B_1)\cos^2\varphi \qquad (1)$$

where $\cos\varphi = (ha^*\cos\beta + lc^*)/2\sin\theta$ (Helmholz, 1936; Hughes, 1941).

The observed planes were divided into four major groups depending on the angle φ, and each major group into five or six minor groups according to $\sin^2\theta$. Scale- and temperature-factor plots were then made separately for each major group. Some of the groups contained too few reflexions for the method to be very reliable but the fact that all four lines had to converge on the same scale factor enabled reasonable mean values for B to found for these different directions. The constants B_1 and B_2 in (1) were adjusted to give a fit with these experimental values, which are given in Table 1 and compared with those calculated from the relation

$$B = 2{\cdot}78 + 1{\cdot}31\cos^2\varphi \quad (B_1 = 2{\cdot}78,\ B_2 = 4{\cdot}09\ \text{Å}^2)\,.$$

Table 1

Group	φ (°)	B obs. (Å^2)	B calc. (Å^2)
1	0–22·5	4·07	4·04
2	22·5–45·0	3·69	3·69
3	45·0–67·5	3·16	3·19
4	67·5–90·0	2·85	2·83

A table was constructed giving B as a function of φ and so the appropriate f functions could be calculated for each plane.

As a result of these corrections the hkl agreement index dropped to 0·21. Further refinement was carried out by computing differential syntheses for both observed and calculated structure factors. These and subsequent calculations were made on the electronic computer at Manchester University using programmes developed by Ahmed & Cruickshank (1953*b*). The termination-of-series effects were reduced by the back-correction method of Booth (1948). Those planes excluded from the line syntheses were included in the differential syntheses as their phases were established. The agreement index after three stages of refinement was 0·169. A difference map was next computed using the observed and calculated $F(hk0)$ and this showed that the hydrogen which previously could not be placed was in fact attached to the nitrogen atom, thus confirming that, when crystallized, uracil contains two keto oxygens. The position of the hydrogen atoms for structure-factor calculation was not taken from this difference map as it was obvious from the appearance of the map that it was not very accurate. Instead, the hydrogen atoms were placed in the molecular plane so that C–H = 1·1 Å and N–H = 1·0 Å, and in the subsequent structure-factor calculations the hydrogen temperature factor was assumed to vary between 4·3 and 6·4 Å^2. The agreement index for these structure-factor calculations was 0·158. An examination of the atomic peak heights and curvatures showed that, excluding the two oxygen atoms, the observed values were on the average about 5% too low, and it was obvious that an increase in the scale factor was necessary. The discrepancies were even greater in the case of the two oxygen atoms, the observed values being about 15% too low, and it was felt that some increase in the temperature factor of the oxygen was desirable. The observed structure factors were therefore scaled up by 5% and the temperature factors for

Table 2. *Atomic peak densities and curvatures*

	ϱ		$-A_{hh}$		$-A_{kk}$		$-A_{ll}$		A_{hl}	
Atom	obs.	calc.	obs.	calc.	obs.	calc.	obs.	calc.	obs.	calc.
C_1	6·21	6·11	35·1	32·9	56·7	53·9	30·6	29·5	7·1	6·8
C_2	5·60	5·59	28·5	26·5	48·2	51·2	22·7	23·1	3·7	2·9
C_4	6·27	6·24	37·1	34·3	60·9	57·4	29·0	28·4	7·8	7·3
C_5	5·64	5·63	33·7	31·1	45·0	46·4	24·9	24·8	7·4	6·5
N_3	7·22	7·30	40·9	38·4	56·3	59·8	32·6	31·6	8·1	7·1
N_6	7·44	7·54	42·3	39·1	67·1	65·4	30·5	33·1	7·7	7·7
O_7	8·24	8·45	42·8	40·8	74·9	72·0	29·4	33·3	8·2	7·4
O_8	7·65	8·13	39·1	38·6	68·2	69·1	26·7	31·4	8·3	7·7

the oxygen atoms changed to $B_1 = 3{\cdot}2$, $B_2 = 4{\cdot}7$ Å^2. Structure factors calculated for these new scale and temperature factors reduced R to 0·142, and the results of the subsequent differential synthesis showed that better overall agreement was now obtained although it appeared that in some cases there had been an over-correction. An examination of the results given in Table 2 shows that there are still further improvements that could be incorporated in the thermal parameters, but it was felt that, apart from assigning individual scale and temperature factors to each atom (Cochran 1951, 1953), there was little point in pursuing the refinement further. The final differential synthesis indicated atomic shifts of about 0·003 Å in the ring atoms, the effect being to contract the ring as a whole. These changes were mainly due to the alteration in the termination-of-series correction when the hydrogen atoms were included in the calculated structure factors and indicated that as a consequence of its high temperature factor the hydrogen atom still gave a positive slope at the position of the heavier atom. Final values of the electron density and peak curvatures are given in Table 2 and the atomic parameters in Table 3.

Table 3. *Atomic parameters and B values*

Atom	x	y	z	B_1 (Å^2)	B_2 (Å^2)
C_1	0·1807	0·0119	0·0398	2·78	4·09
C_2	0·3480	0·1310	0·0907	—	—
C_4	0·1530	0·2082	0·0040	—	—
C_5	0·3068	0·0242	0·0943	—	—
N_3	0·2735	0·2166	0·0562	—	—
N_6	0·1106	0·1038	0·0054	—	—
O_7	0·1293	−0·0773	0·0177	3·2	4·7
O_8	0·0833	0·2857	−0·0267	—	—
H_9	0·367	−0·047	0·124	4·3	6·4
H_{10}	0·445	0·142	0·131	—	—
H_{11}	0·311	0·299	0·058	—	—
H_{12}	0·022	0·094	−0·030	—	—

Extinction effects were noticeable when the final values of F_c were compared with F_o. The observed value of $F(001)$ is little more than half its calculated value but none of the other planes is seriously affected although, for the five strongest reflexions, the calculated value is always greater than that observed.

The accuracy of the analysis

The standard deviation of the x, y and z parameters was estimated by the methods of Cruickshank (1949):

$$\sigma(A_h) = \frac{2\pi}{aV}\{\sum_3 4h^2(\Delta F)^2\}^{\frac{1}{2}} ,$$

the summations being over independent $F(hkl)$ and $\Delta F = F_o - F_c$, with similar expressions for $\sigma(A_k)$ and $\sigma(A_l)$. Then

$$\sigma(x) = \frac{\{[A_{ll}\sigma(A_h)]^2-[A_{hl}\sigma(A_l)]^2\}^{\frac{1}{2}}}{A_{hh}A_{ll}-A_{hl}^2} ,$$

$$\sigma(y) = \frac{\sigma(A_k)}{A_{kk}} ,$$

$$\sigma(z) = \frac{\{[A_{hh}\sigma(A_l)]^2-[A_{hl}\sigma(A_h)]^2\}^{\frac{1}{2}}}{A_{hh}A_{ll}-A_{hl}^2} .$$

Table 4. *Estimated standard deviations of atomic positions*

Atom	$\sigma(x)$ (Å)	$\sigma(y)$ (Å)	$\sigma(z)$ (Å)
C_1	0·0074	0·0064	0·0142
C_2	0.0092	0.0075	0.0182
C_4	0.0069	0.0060	0.0151
C_5	0.0075	0.0080	0.0177
N_3	0.0063	0.0065	0.0134
N_6	0.0060	0.0054	0.0142
O_7	0.0059	0.0049	0.0149
O_8	0.0064	0.0053	0.0166

Values of $\sigma(x)$, $\sigma(y)$ and $\sigma(z)$ are given in Table 4. The error in the z direction is much greater than those in x and y, which, as one might expect, are about equal. As the molecule lies approximately parallel to (001), the bond lengths are only slightly dependent on the z parameter. The standard deviation of the bond length was therefore calculated from the relation

$$\sigma^2(AB) = [\sigma_A^2(x)+\sigma_B^2(x)]\cos^2\alpha + [\sigma_A^2(y)+\sigma_B^2(y)]\cos^2\beta + [\sigma_A^2(z)+\sigma_B^2(z)]\cos^2\gamma ,$$

where cos α, cos β and cos γ are the direction cosines of the bond AB (Ahmed & Cruickshank, 1953*a*). Values of $\sigma(AB)$ are given in Table 5.

Table 5. *Standard deviation of bond lengths*

Bond	$\sigma(AB)$ (Å)	Bond	$\sigma(AB)$ (Å)
C_1–C_5	0·0136	C_4–N_6	0·0082
C_5–C_2	0·0111	N_6–C_1	0·0097
C_2–N_3	0·0117	C_1–O_7	0·0095
N_3–C_4	0·0122	C_4–O_8	0·0101

The deviations of the bond angles ABC were computed from the equation given in the above paper by Ahmed & Cruickshank and were found to vary from 0·6° to 0·8°.

The crystal structure and molecular dimensions of uracil

The molecules are arranged in hydrogen-bonded layers parallel to (001). A typical layer is shown in Fig. 3; part of the next layer is also illustrated. It has already been observed that an unexpected feature of this structure is that one oxygen atom (O_7) forms two hydrogen bonds to nitrogen atoms in neighbouring molecules while the other forms no hydrogen bonds at all, unless the approach of two CH groups at 3·2 Å is regarded as weak hydrogen or polarization bonding. The length of the hydrogen bond from O_7 to N_6 in a molecule related by a centre of symmetry is 2·81 Å, while that to N_3 in a molecule related by a screw axis is 2·86 Å. The NH · · · O distances agree well with previous results for similar compounds (see Brown, 1951). The hydrogen-bond lengths are intermediate in value between that in urea, 3·10 Å (mean) (Vaughan & Donohue, 1952) and that in L$_s$-threonine, 2·67 Å (mean) (Shoemaker *et al.*, 1950); they agree closely with the mean value of 2·84 Å in DL-alanine (Donohue,

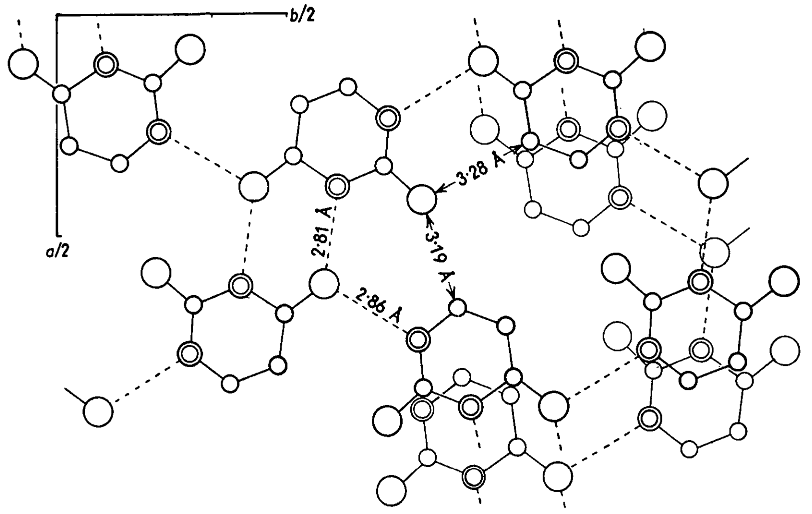


Fig. 3. The structure of uracil projected on (001). One half of the drawing shows the molecular arrangement in a typical layer while the other half shows the superposition of two such layers.

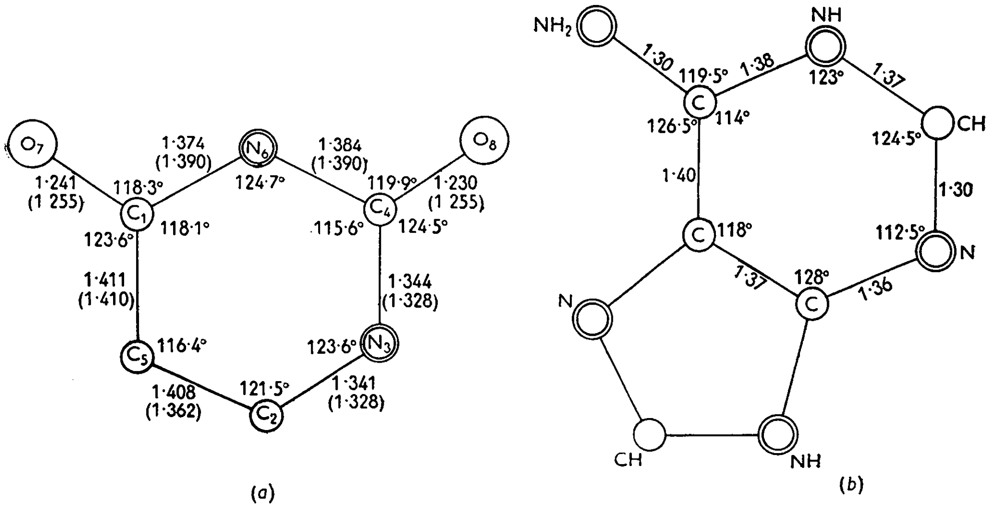


Fig. 4. (a) The dimensions and bond angles of the uracil molecule. Bond lengths predicted by valence bond methods are shown in parentheses. The difference in the bond C_2–C_5 is significant. (b) Dimensions of the pyrimidine ring in adenine hydrochloride.

1950). The oxygen atom O_8 is approached by carbon atoms in different molecules, the interatomic distances being 3·19 and 3·28 Å. The closest approach between layers is 3·34 Å.

The molecule is found to be planar to within experimental error; the atomic displacements from the plane

$$0{\cdot}120x-0{\cdot}027y-0{\cdot}992z = 0{\cdot}156$$

(x, y and z in Ångström units) is never greater than $2\sigma(z)$.

The molecular dimensions are given in Fig. 4(a). The C–N bonds adjacent to N_6 are indistinguishable with a probable length of 1·379±0·006 Å, while those adjacent to N_3 are also indistinguishable with a probable length of 1·342±0·008 Å. The probability that the difference in these two different C–N bond lengths is due solely to experimental error is 2×10^{-3}, and the difference is therefore significant. The two C–C bond lengths and also the two C=O bond lengths are not significantly different, the probable lengths being 1·409±0·009 Å and 1·236±0·007 Å respectively. There is no indication that the different intermolecular environment is affecting the C=O bond lengths in any way. Details of the statistical analysis upon which these conclusions are based are given in Table 6.

The internal angles of the ring show that the nitrogen bond angles are the same with a mean value of 124·1°, which is significantly greater than 120°. The carbon bond angles on the average are less than 120°. The angle formed between the C=O and the longer of the two C–N is somewhat less than 120°, whereas that formed with the shorter C–N or with the C–C has a mean value of 124°, which is again significantly greater than 120°.

Table 6. *Comparison of bond lengths in the uracil molecule*

Bonds		Individual estimated s.d. (Å)		Total estimated s.d. $\sigma(t)$ (Å)	Difference in bond length Δ (Å)	$\Delta/\sigma(t)$	P	Conclusion
N_6–C_1	N_6–C_4	0·0097	0·0082	0·0143	0·010	0·70	0·48	Not significant
N_3–C_4	N_3–C_2	0·0122	0·0117	0·0182	0·003	0·16	0·88	Not significant
N_6–C_1	N_3–C_4	0·0097	0·0122	0·0156	0·030	1·92	0·054	Possibly significant
N_6–C_4	N_3–C_2	0·0082	0·0117	0·0143	0·043	3·01	0·003	Significant
N_6–C_4	N_3–C_4	0·0082	0·0122	0·0157	0·040	2·55	0·010	Significant
N_3–C_4	N_3–C_2	0·0097	0·0117	0·0152	0·033	2·17	0·030	Possibly significant
C_1–C_5	C_5–C_2	0·0136	0·0111	0·0175	0·003	0·17	0·87	Not significant
C_1–O_7	C_4–O_8	0·0095	0·0101	0·0139	0·011	0·79	0·42	Not significant
Test of hypothesis that N_3–C_4 N_3–C_2 are different from N_6–C_1 N_6–C_4		—	—	0·0119	0·0365	3·07	0·002	Significant

Summary of formula used above

Non-adjacent bonds:

$\sigma^2(t) = \sigma^2(AB)+\sigma^2(CD)$.

Adjacent bonds:

$\sigma^2(t) = \sigma^2(AB)+\sigma^2(BC)-2\operatorname{cov}(AB.BC)$, where $\operatorname{cov}(AB.BC) = \sigma^2(B)\cos\theta$. $\sigma(B)$ should be the estimated s.d. of atom B along a line joining B to the centre of the circle passing through A, B and C. In the present case, this was taken to be along a line passing through B and the midpoint of AC. θ is the bond angle ABC.

Test of hypothesis that bonds AB, BC are significantly different from CD, DE

$$\sigma^2(t) = \tfrac{1}{4}\{\sigma^2(AB)+\sigma^2(BC)+\sigma^2(CD)+\sigma^2(DE)-2\operatorname{cov}(AB.BC)-2\operatorname{cov}(BC.CD)-2\operatorname{cov}(CD.DE)\}\,.$$

In this case, bond length difference $\Delta = \tfrac{1}{2}(AB+BC)-\tfrac{1}{2}(CD+DE)$.

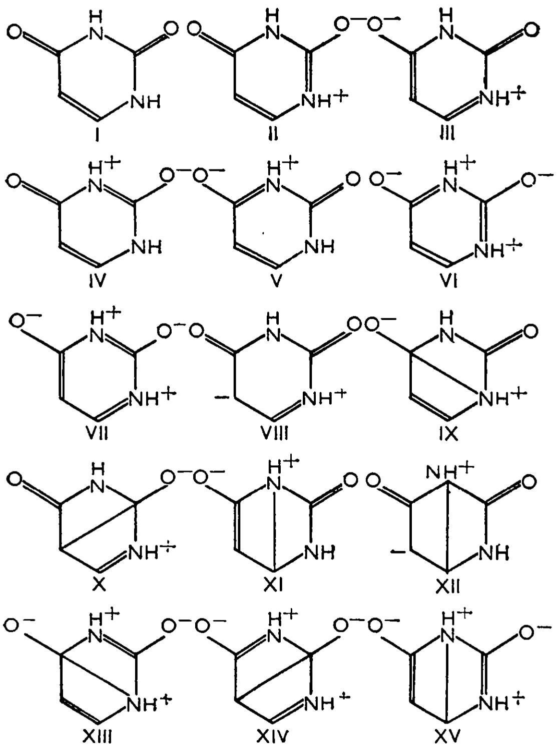
Discussion

It has been established in the present work that the pyrimidine ring in uracil contains two different C–N bond lengths; the shorter of these, 1·342±0·008 Å is in agreement with the accurate determinations of the C–N bond length in the following compounds: melamine, 1·343 Å (mean of ring values) (Hughes, 1941); urea, 1·335 Å (Vaughan & Donohue, 1952); N-acetylglycine, 1·323 Å (peptide C–N) (Carpenter & Donohue, 1950). The value is also in agreement with the estimated value of 1·33 Å for the C–N bond distance in pyrimidine given by Clews & Cochran (1949) as a result of their study of pyrimidine derivatives. The longer of the C–N bonds, 1·379±0·006 Å, is likely to be significantly different from all the above values. This length agrees with the mean of the C–N bonds in a similar position in the pyrimidine ring of adenine HCl (Cochran, 1951); (in this analysis, the standard deviation of the C–N and C–C bonds is given as 0·01 Å). A comparison of the pyrimidine rings in uracil and adenine HCl is included in Fig. 4. The replacement of the hydrogen in adenine by the oxygen in uracil does not affect the two adjacent C–N bonds equally. There is a significant lengthening of one C–N bond by 0·04 Å but the other bond lengthens by only 0·01 Å, which is not significant. This preferential lengthening of one bond in uracil, tends to equalize the two bonds, even though they are still significantly different. In adenine these bonds differ by 0·07 Å, in uracil by only 0·04 Å. The C(NH)C angle is about 124° in both molecules, compared with a value of 112·5° for the CNC angle in adenine. A comparison of the bond lengths around the ring shows that in no case is the difference in bond length more than 2·7σ, which is only just significant and indicates a close resemblance in the two rings despite the fact that in adenine an imidazo ring is attached to the pyrimidine ring.

α-Pyridone also contains a long C–N, 1·401 Å (Penfold, 1953), which is adjacent to a C=O of length 1·236 Å. In this case the third bond to the carbon atom, C–C of length 1·444 Å, is somewhat longer than that in uracil, 1·411 Å.

The C=O distance of 1·236±0·007 Å, assuming both bonds in uracil are equivalent, is the same as in α-pyridone (see above) and agrees with the value of 1·243 Å in N-acetylglycine, 1·23 Å in β-glycylglycine (Hughes & Moore, 1949) and the value found in most other molecules containing a keto oxygen, with the exception of the carboxylic acids.

The two C–C distances are not significantly different from the benzene value of 1·39 Å. The equality of these two bonds is a notable feature of the structure as one is a single and the other a double bond in the formal valence bond representation. The valence bond formulae with zero, one and two charge separations are shown below. Considering first the Kekulé structures (I–VIII) the equality of the two C–C bonds means that form III, and to a lesser extent VII and possibly VIII, must contribute significantly to the valence structure. To account for the different C–N lengths found and for the equality of the C=O bonds, II and VI must contribute in roughly the same proportion as III and VII. Forms IV and V cannot make an



appreciable contribution as they would tend to shorten the C=O in I. If the Pauling relation between bond character and length is assumed (Pauling, 1948) and the values of the single and double C–C and C–N bond lengths are taken to be those given by Vaughan & Donohue (1952), then a satisfactory explanation of the observed bond lengths, with the exception of the C=C in I, can be given by assuming that forms I, II, III, VI and VII contribute to the resonant structure in the proportions 6:4:4:3:3. The bond lengths predicted are compared to those observed in Fig. 4(*a*). The observed magnitude of the formal double bond is significantly greater ($4 \cdot 14\sigma$) than its predicted value, but in all other cases the difference never exceeds 2σ. Forms VIII and XII predict a negative charge on the carbon atom and are therefore less likely to make a significant contribution. Although these forms would increase the length of the short C=C bond, there would be a general worsening in the agreement for the rest of the molecule. Of the three singly charged Dewar forms (IX–XI), two predict a lengthening of the double bond but would make agreement worse for the C=O and C–N bonds. The doubly charged Dewar forms XIII–XV would make all the ring bonds slightly larger but would do nothing to lengthen the short C=C appreciably. Hence it does not seem possible to account for the observed bond lengths in uracil by the valence bond method. Some less empirical form of theoretical treatment will be required. According to the valence bond interpretation, the difference in the C–N bonds is to be interpreted as a difference in charge on the two nitrogen atoms, N_3 carrying the greater formal positive charge. It is probable that this conclusion would be reached whatever the method used to interpret these results.

In the earlier part of this paper, it was mentioned that the crystal used to collect the intensity data was probably not completely free from the disorder present in so many uracil crystals. The effect of this disorder, if any, will be merely to raise the standard deviation of the measured atomic positions owing to the increased discrepancy between F_o and F_c. Likewise errors in the observed atomic positions due to those extinguished planes which were included in the final refinements with their observed values will also be allowed for by an increase in the estimated standard deviation. It is unlikely therefore that either of these two effects would invalidate any conclusions reached on a basis of the statistical analysis.

I wish to express my thanks to Prof. E. G. Cox for his interest and encouragement and to Dr D. W. J. Cruickshank for some helpful discussions on statistical matters. The three-dimensional refinement was carried out on the Manchester electronic computer and my thanks are due to the staff of the Computing Machine Laboratory for the use of the machine.

References

Ahmed F. R. & Cruickshank, D. W. J. (1953*a*). *Acta Cryst.* **6**, 385.
Ahmed, F. R. & Cruickshank, D. W. J. (1953*b*). *Acta Cryst.* **6**, 765.
Booth A. D. (1948). *Fourier Technique in X-ray Organic Structure Analysis.* Cambridge: University Press.
Broomhead, J. M. (1948). *Acta Cryst.* **1**, 324.
Broomhead, J. M. (1951). *Acta Cryst.* **4**, 92.
Brown, C. J. (1951). *Acta Cryst.* **4**, 100.
Carpenter, G. B. & Donohue, J. (1950). *J. Amer. Chem. Soc.* **72**, 2315.
Clews, C. J. B. & Cochran, W. (1948). *Acta Cryst.* **1**, 4.
Clews, C. J. B. & Cochran, W. (1949). *Acta Cryst.* **2**, 46.
Cochran, W. (1951). *Acta Cryst.* **4**, 81.
Cochran, W. (1953). *Acta Cryst.* **6**, 260.
Cruickshank, D. W. J. (1949). *Acta Cryst.* **2**, 65.
Donohue, J. (1950). *J. Amer. Chem. Soc.* **72**, 949.
Furberg, S. (1950). *Acta Cryst.* **3**, 325.
Gilpin, V. & McCrone, W. C. (1950). *Analyt. Chem.* **22**, 368.
Helmholz, L. (1936). *J. Chem. Phys.* **4**, 316.
Hughes, E. W. (1941). *J. Amer. Chem. Soc.* **63**, 1737.
Hughes, E. W. & Moore, W. J. (1949). *J. Amer. Chem. Soc.* **71**, 2618.
Knott, G. (1940). *Proc. Phys. Soc.* **52**, 229.
Pauling, L. (1948). *The Nature of the Chemical Bond*, 2nd ed. Ithaca: Cornell University Press.
Penfold, B. R. (1953). *Acta Cryst.* **6**, 591.
Pitt, G. J. (1948). *Acta Cryst.* **1**, 168.
Shoemaker, D. P., Donohue, J., Schomaker, V. & Corey, R. B. (1950). *J. Amer. Chem. Soc.* **72**, 2328.
Vaughan, P. & Donohue, J. (1952). *Acta Cryst.* **5**, 530.
Wilson, A. J. C. (1942). *Nature, Lond.* **150**, 152.