

The Crystal Structure of 4,6-dimethyl-2-hydroxypyrimidine. I

By G. J. PITT

*Birkbeck College Research Laboratory (University of London), 21 Torrington Square,
London W.C. 1, England*

(Received 6 February 1948)

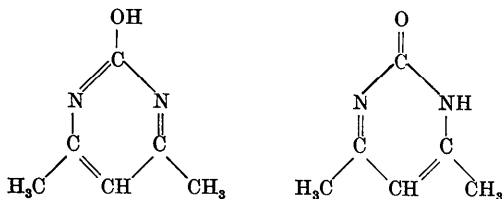
Crystals of 4,6-dimethyl-2-hydroxypyrimidine dihydrate, $C_6H_8ON_2 \cdot 2H_2O$, are monoclinic with $a = 6.80$ A., $b = 13.45$ A., $c = 9.18$ A., $\beta = 101^\circ$.

The cell contains four molecules. The space group is $P2_1/n$, but is strongly pseudo- $A2/a$. Coordinates for the atoms are derived and refined by two-dimensional Fourier syntheses and the method of steepest descents. The pyrimidine molecule is found to be planar within the limits of present accuracy and the bond lengths are C-C: 1.52, 1.56 A.; C=N: 1.34, 1.35, 1.38, 1.39 A.; C=OH: 1.25 A.; C=CH: 1.37, 1.39 A.; (all ± 0.04 A.). Two water molecules form hydrogen bridges to the hydroxyl group, and each water is also linked to a nitrogen atom in the next molecule and to another water molecule; all hydrogen bridges are of moderate strength, the lengths lying between 2.7 and 3.0 A.

Introduction

The pyrimidines form a group of compounds of interest to the chemist and to the biologist studying the processes occurring in the living cell, but it appears that until recently no attempts have been made to elucidate their structures crystallographically. During the past two years work has been in progress elsewhere on 2-amino-4-methyl-6-chloropyrimidine and 2-amino-4,6-dichloropyrimidine (Clews & Cochran, 1947), but as yet no data on bond lengths have been published.* The present investigation of 4,6-dimethyl-2-hydroxypyrimidine-dihydrate is independent of this work, so that a comparison of findings at a later stage should permit more general conclusions to be drawn about the form of the pyrimidine ring.

The synthesis of this compound by de Haan (1908) gave two crystalline forms which were thought to be composed of tautomers:



but Jaeger (1908) concluded that the two forms were identical except for a trace of impurity in one. The question of which formula represents the compound has not apparently been finally cleared up in the literature. This, together with the possibility of

hydrogen bridge formation, makes the compound an interesting one from a crystallographic standpoint.

The present report covers the preliminary investigation and two-dimensional syntheses which have been carried out to date, leading to values for bond lengths.

Preliminary investigation

Crystals of both de Haan's forms were prepared at the Chester Beattie Laboratory of the Royal Cancer Hospital, and samples were kindly supplied for this analysis. The identity of the two types of crystals was deduced from microscopic inspection and from comparison of Weissenberg photographs before Jaeger's observations could be traced; and the two sets of axial ratios agreed closely. Subsequent work has all been carried out on the colourless form, since it was judged to be of higher purity; it has not yet proved possible to isolate the impurity in the other form which was pale yellow-green in colour.

The crystals obtained by recrystallization from water or water-alcohol solutions are monoclinic, elongated parallel to the x axis, with the general habit shown in Fig. 1. Very rarely other $(0kl)$ faces have been found, but no $(hk0)$ or (hkl) faces. (Jaeger's a and c axes correspond to the present c and a axes respectively. He observed very small (031) , (001) , (110) and (210) faces—referring to new axes—on very well-developed crystals.) (100) is a good cleavage.

If exposed to sunlight, the crystals become slightly yellow but do not otherwise deteriorate, whereas the

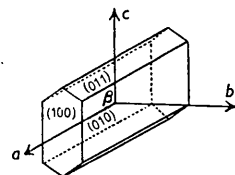


Fig. 1. Clinographic projection of the crystal.

* Note by Editor. A further communication has now appeared (Clews & Cochran, 1948).

anhydrous compound (to be reported later) is hygroscopic, and crystals of it obtained by sublimation transform into a polycrystalline mass in a few hours if exposed to air.

Previous chemical evidence indicates that there are two molecules of water of crystallization per molecule of pyrimidine (i.e. the formula is $C_6H_8ON_2 \cdot 2H_2O$). The water of crystallization is driven off on heating to about $120^\circ C$. and the anhydrous compound remaining melts at $197^\circ C$., though considerable sublimation occurs below that temperature.

The birefringence is large and negative, $\gamma \parallel b$, and α makes an angle of $2-3^\circ$ with the a axis in the acute angle. $2V = 30 \pm 4^\circ$.

Oscillation and Weissenberg photographs confirmed that the crystals are monoclinic, and the following cell dimensions were derived from reflexions with high θ values:

$$\begin{aligned} a &= 6.79_7 \pm 0.01 \text{ \AA.}, & b &= 13.45_3 \pm 0.02 \text{ \AA.}, \\ c &= 9.17_8 \pm 0.01_5 \text{ \AA.}, & \beta &= 101^\circ \pm \frac{1}{4}^\circ. \end{aligned}$$

Inspection of all reflexions obtained showed the following absences: ($0k0$) when k is odd; ($h0l$) when $(h+l)$ is odd. There are no other systematic absences, so the space group is $P2_1/n$, but, in general, reflexions with $(k+l)$ odd are noticeably weaker than neighbouring ones with $(k+l)$ even; this represents a pseudo-centring on the A face (pseudo-space group: $A2/a$). The density as measured by flotation is 1.29_3 g.cm^{-3} (cf. Jaeger: 1.291 g.cm^{-3}), hence there are four (calculated 4.00_4) molecules of $C_6H_8ON_2 \cdot 2H_2O$ per cell.

The optical properties are consistent with the molecules being approximately planar and arranged with their planes roughly parallel to $(30\bar{1})$ or normal to the a axis; (200) and $(60\bar{2})$ are very strong reflexions agreeing with this conclusion.

The pseudo-face-centring can be interpreted in this way. If the space group had really been $A2/a$ (eight general positions in the cell), the molecule must have had some symmetry in itself since there are only four molecules in the cell. This molecular symmetry could only be a diad axis, since the other alternative (a centre of symmetry) is clearly not possible in a tri-substituted six-member ring. It seems reasonable therefore to expect the molecule to lie with the molecular pseudo-diad axis (as shown dotted in Fig. 2) near to the line $x = \frac{1}{2}, z = 0$, which would be the diad axis in the $A2/a$ space group; the possibility that the molecule is turned in its own plane through 120° either way cannot be ruled out at this stage, however. This, together with the deductions from the optical properties, forms a useful basis for the interpretation of the Patterson projections.

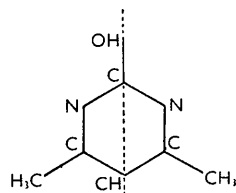


Fig. 2. Molecule showing pseudo-diad axis - - - - -.

Patterson and Fourier projections

For intensity measurements oscillation photographs about the three principal axes were obtained, using a crystal roughly $\frac{1}{2}$ mm. cube exposed to $Cu K$ radiation. The reflexions were estimated visually in the usual manner and converted to an absolute scale later by comparison of observed and calculated structure factors.

Contour diagrams of the Patterson F^2 -function projected parallel to the three principal axes are shown in Figs. 3-5; all show very clearly the marked pseudo-symmetry. The b projection, with its series of peaks lying along the traces of the $(60\bar{2})$ planes, can be readily interpreted in the light of the model postulated above, peak α being due to intramolecular vectors such as N-C, C-CH₃, C-CH, etc., projected parallel to b , peak β to vectors such as N-N, C-C, CH-CH₃, etc., peak γ to CH₃-C, CH₃-N, and so on. The intermolecular vectors fall in the group centred about the peak (π) at $x = \frac{1}{2}, z = 0$. Rough x and z co-ordinates can be assigned to all the atoms except those of the water, the positions of which are not clear.

The a and c Patterson projections are less informative, but the principal intermolecular peaks (π) are recognizable. In these projections no very prominent intermolecular peak is to be expected between molecules related by the glide-plane, so it is possible to obtain only a very rough estimate of the positions of the molecules in the y direction. Again no clear evidence of vectors due to the water molecules could be deduced. ($h0l$) structure factors were therefore calculated ignoring the contributions from the water molecules, and the signs of these were used in a two-dimensional Fourier projection parallel to the b axis. Although the agreement between observed and calculated structure factors was poor and only about half of the available terms were used, additional peaks appeared in positions which seemed feasible for water molecules from consideration of a simple model. When the ($h0l$) structure factors were recalculated, including contributions from water molecules in these positions, some improvement in agreement occurred.

These z co-ordinates were then used in conjunction with tentative sets of y co-ordinates in calculating ($0kl$) structure factors. To begin with, the contributions from the water groups were again ignored, since their y co-ordinates were very uncertain, and after a number of trials two Fourier projections parallel to the a axis were chosen which appeared more promising than the rest; on both of these additional peaks appeared which could be interpreted as being due to water molecules (the main difference between the two sets of co-ordinates was a shift of the entire molecule about 2.5 \AA . in the y direction). Attempts to refine one of these sets of co-ordinates by successive Fourier projections met with no success, so the other was tried and found to improve slowly until after about eight successive corrections to

the co-ordinates the contour diagram shown in Fig. 6 was obtained; at this point hardly any terms changed sign and further repetitions would have produced no perceptible change.

Fourier projections parallel to the b and c axes were also computed, but showed very incomplete resolution, since in either case the pyrimidine ring is nearly parallel to the direction of projection; the water molecules are, however, distinguishable and their x co-ordinates can be determined with moderate accuracy. The b projection seems to indicate that the angle between the plane of the molecule and the y axis is very small, but the x co-ordinates of most of the atoms can be estimated only roughly (Fig. 7).

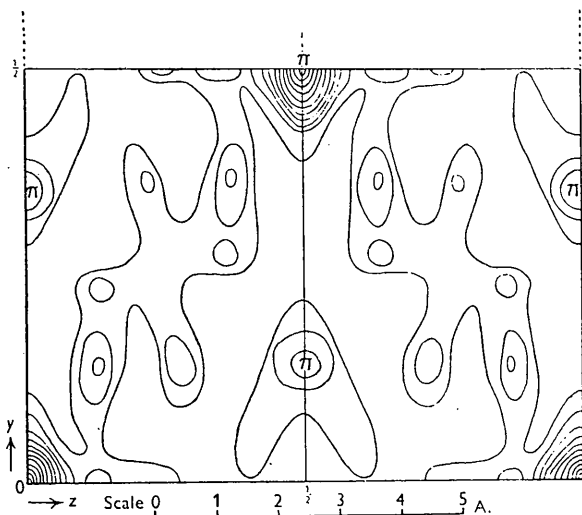


Fig. 3. Patterson projection in direction of a axis. (Contours at arbitrary heights; intermolecular peaks π .)

Steepest descents

Owing to the slowness with which the a projection had refined and the fact that the agreement between the observed and calculated structure factors was still not very good, it was decided to apply the method of steepest descents which has been recently suggested by Booth (1947). In this, a procedure is formulated for minimizing a function R or R_2 which measures the disagreement between observed and calculated structure factors; the smallness of R or R_2 is the generally accepted criterion for the correctness of a proposed structure.

If the original co-ordinates of the atoms x_{rj} correspond to a value of $R(=R_0)$, then the formula

$$\delta x_{rj} = -R_0 \left(\frac{\partial R}{\partial x_{rj}} \right)_0 \bigg/ \sum_{rj} \left(\frac{\partial R}{\partial x_{rj}} \right)_0^2$$

gives the corrections δx_{rj} to be applied to these co-ordinates as we proceed along the line of greatest slope of the $R(x_{rj})$ function to its intersection with the plane $R=0$; having applied these corrections we can calculate the new value of $R(=R_1)$. In order to locate the

position of the minimum the value of $R(=R_1)$ should also be calculated for corrections $\frac{1}{2}\delta x_{rj}$; parabolic interpolation between $R_0, R_{\frac{1}{2}}, R_1$ will give the required position. Clearly the line of greatest slope will not in

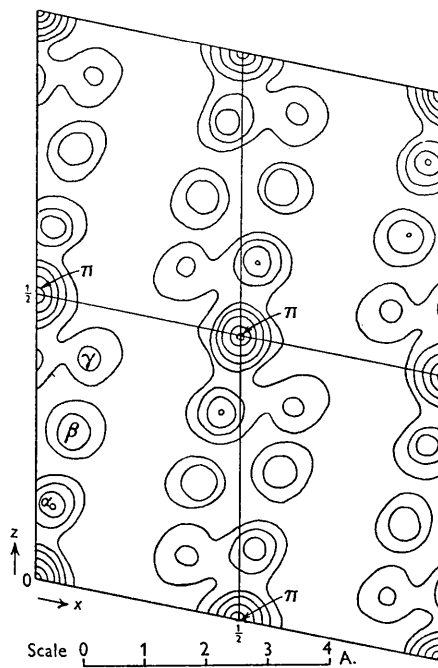


Fig. 4. Patterson projection in direction of b axis.

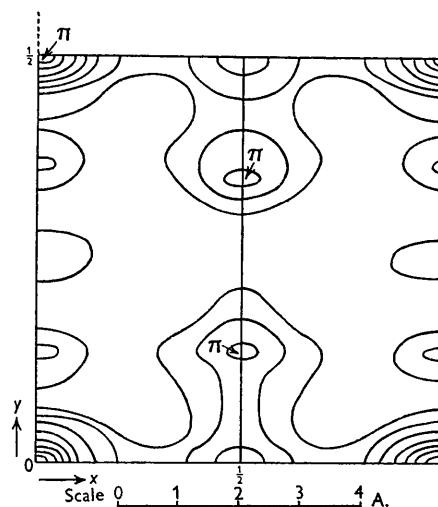


Fig. 5. Patterson projection in direction of c axis.

general pass exactly through the minimum point, so the procedure may have to be repeated a number of times, but a definite improvement can be hoped for at each stage.

In the present instance the method was applied to the y and z co-ordinates derived from the final Fourier a projection; the simple expression

$$R = \sum_{kl} || F_{\text{obs.}} - F_{\text{calc.}} ||$$

(3) Extinction errors have not been allowed for; these will tend to raise the disagreement factor, but will not affect the accuracy of the co-ordinates much, since only low-order terms are affected to any extent.

The variations from the average are barely significant (1.34–1.39 Å.), since these results are based on two-dimensional analysis; a probable error of ± 0.04 Å. is expected. Full three-dimensional analysis may show

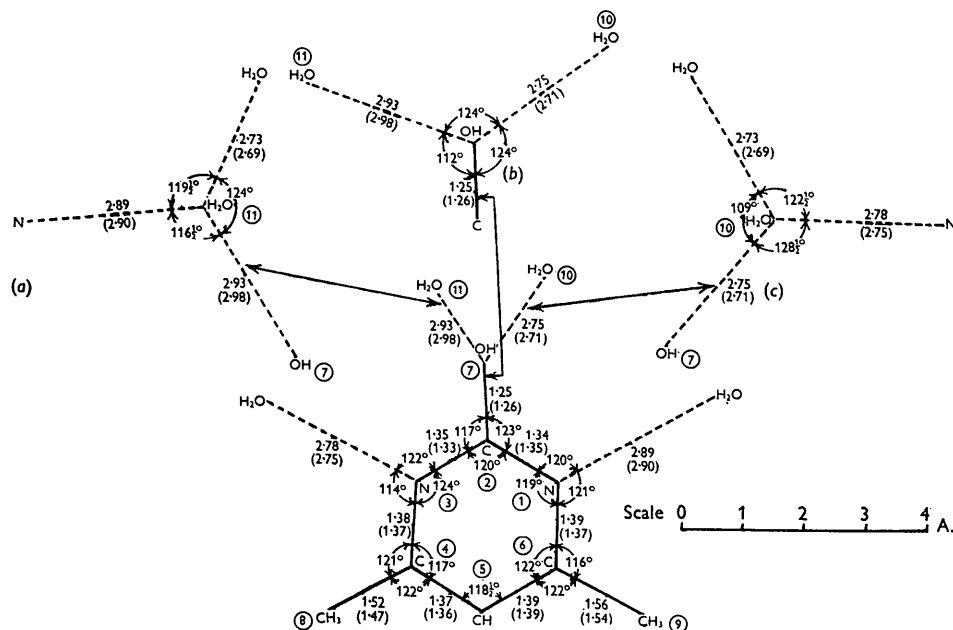


Fig. 8. Pyrimidine molecule viewed in a direction perpendicular to its plane. Bond lengths and bond angles are calculated from co-ordinates derived from steepest descent method, the bond lengths obtained from Fourier projection being in parentheses. (a) Configuration about H_2O molecule. (b) Configuration about the $-\text{OH}$ group, viewed in a direction 70° to the plane of diagram below. (c) Configuration about H_2O molecule.

The present values of bond lengths must therefore be regarded as tentative and liable to correction by refined analysis. The nature of the structure can, however, be described in general terms as follows.

The pyrimidine ring (Fig. 8) has resemblances to the benzene ring in that it is planar and the bonds are intermediate between double and single. Taking the co-ordinates of atoms 2, 4, 6 in the ring as defining a plane, the displacements of the other atoms are as follows, positive and negative signs indicating displacements on opposite sides of the reference plane:

- (1) -0.00 Å., (3) -0.01 Å., (5) -0.05 Å.,
 (7) $+0.06$ Å., (8) $+0.10$ Å., (9) -0.06 Å.,
 (10') -0.20 Å., (11') -0.12 Å.

The lengths of the $\text{C}=\text{C}$ bonds in the ring (1.37, 1.39 Å.) correspond with those found in benzene compounds (1.38–1.39 Å.). The average of the four $\text{C}=\text{N}$ bond lengths is 1.36–1.37 Å., which compares with the averages in other compounds:

Melamine (Hughes, 1941)	1.34 ₆ Å.
Dicyandiamide (Hughes, 1940)	1.35 ₇ Å.
Diketopiperazine (Corey, 1938)	1.37 Å.
Cyanuric triazide (Knaggs, 1935)	1.34 ₅ Å.

a tendency towards closer equality as in melamine or dicyandiamide, or divergence as in diketopiperazine or cyanuric triazide (Fig. 9).

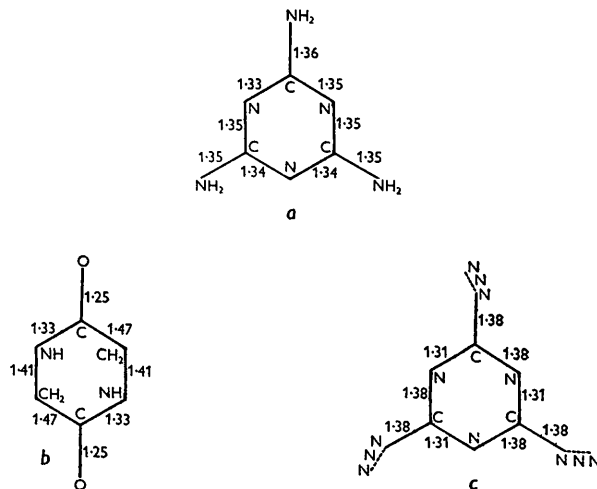


Fig. 9. Bond lengths in (a) melamine, (b) diketopiperazine, (c) cyanuric triazide.

The bond between the hydroxyl group and the 2-carbon is of length 1.25 ± 0.04 Å. and resembles that found in several compounds showing resonance (e.g.

oxalic acid dihydrate and other carboxylic compounds); there is a clear difference from resorcinol
($C\equiv OH = 1.36-1.37 \text{ \AA}$.)

water molecules are linked to each hydroxyl group at distances of 2.75 and 2.93 \AA ., while each water is further linked to another water molecule next above in

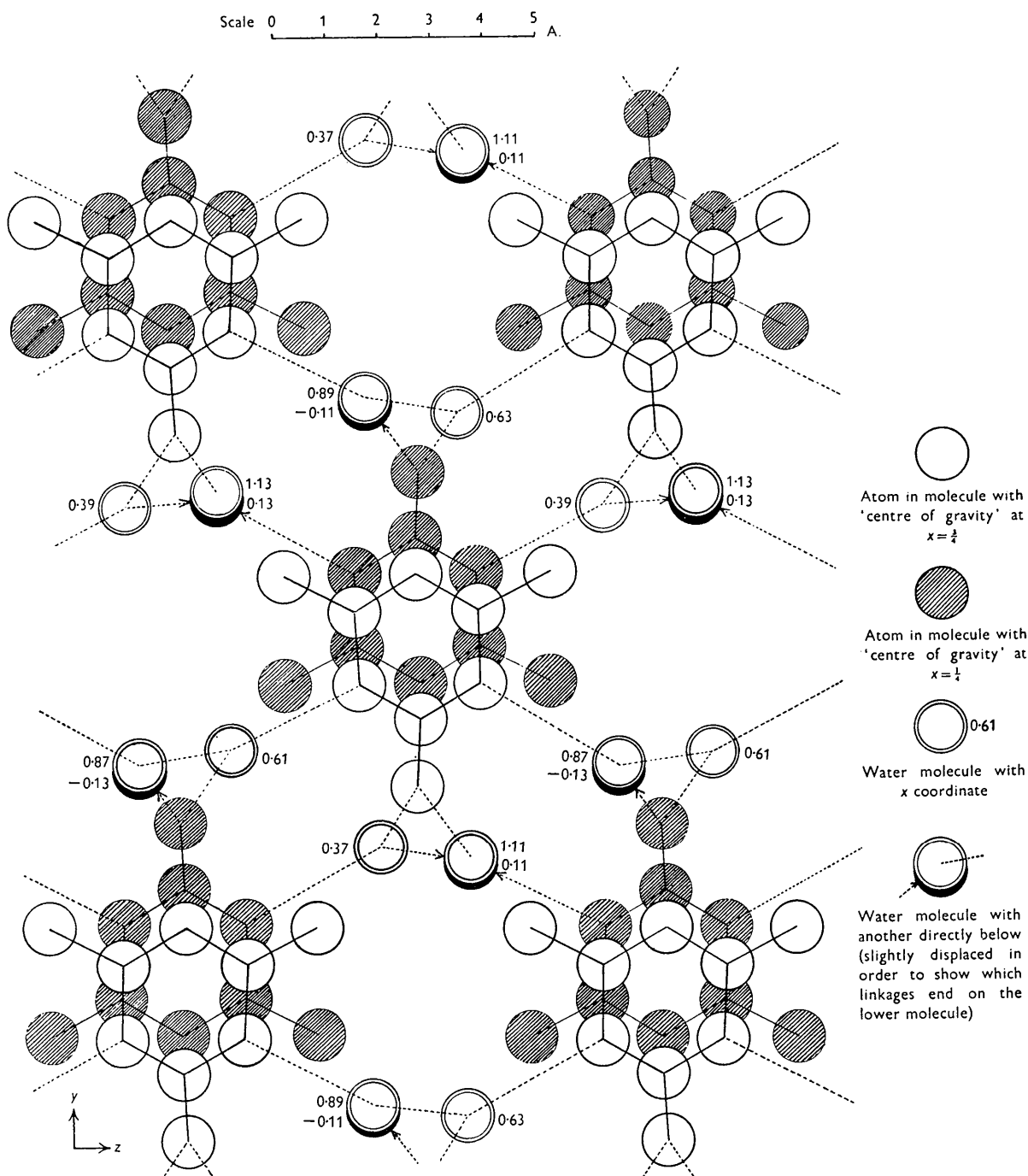


Fig. 10. The structure viewed in the direction of the a axis, showing intermolecular linkages - - - -. All molecules are tilted downwards to the left by 10° .

(Robertson, 1936, 1938) in which the resonance cannot so readily extend from the ring to the hydroxyl groups.

The molecules are linked to one another by hydrogen bridges as shown by dotted lines in Figs. 6-8. Two

the a direction at 2.73 \AA . and to the nitrogen atom of the pyrimidine molecule related to the original one by the glide plane (2.78, 2.89 \AA .), thus knitting the structure together firmly (Fig. 10).

The intermolecular linkages seem to provide good evidence in favour of the theory of mesohydric tautomerism given by Hunter (1946), but a full discussion of the resonance phenomena cannot be attempted here. It is hoped to present this, together with the results of further work on the anhydrous form of this compound, at a later date.

The author is indebted to Dr C. H. Carlisle for much helpful advice and encouragement throughout the course of this work. Acknowledgement is also due to the Department of Scientific and Industrial Research for a maintenance grant.

References

- BOOTH, A. D. (1947). *Nature, Lond.*, **160**, 196.
 CLEWS, C. J. B. & COCHRAN, W. (1947). *Nature, Lond.*, **159**, 264.
 CLEWS, C. J. B. & COCHRAN, W. (1948). *Acta Cryst.* **1**, 4.
 COREY, R. B. (1938). *J. Amer. Chem. Soc.* **60**, 1598.
 DE HAAN, T. (1908). *Rec. Trav. chim. Pays-Bas*, **27**, 162.
 HUGHES, E. W. (1940). *J. Amer. Chem. Soc.* **62**, 1258.
 HUGHES, E. W. (1941). *J. Amer. Chem. Soc.* **63**, 1737.
 HUNTER, L. (1946). *Rep. Progr. Chem.* **141**.
 JAEGER, F. M. (1908). *Z. Krystallogr.* **44**, 569.
 KNAGGS, I. E. (1935). *Proc. Roy. Soc. A*, **150**, 576.
 ROBERTSON, J. M. (1936). *Proc. Roy. Soc. A*, **157**, 79.
 ROBERTSON, J. M. (1938). *Proc. Roy. Soc. A*, **167**, 122.

Acta Cryst. (1948). **1**, 174

The Application of the Harker-Kasper Method of Phase Determination

BY J. GILLIS

Department of Applied Mathematics, Weizmann Research Institute, Rehovot, Palestine

(Received 4 October 1947 and in revised form 19 April 1948)

The Harker-Kasper inequalities and related ones are applied to the data on oxalic acid dihydrate given by Robertson and Woodward. The technique of obtaining some forty signs of the structure amplitudes is developed in detail. Some general remarks on the method conclude the paper.

1. Introduction

1.1. The main purpose of this paper is the application of the method of inequalities, due to Harker & Kasper (1948), to the determination of the phases of some of the structure factors in an actual case. We discuss in § 2 the data on oxalic acid dihydrate given by Robertson & Woodward (1936). The working will be shown in some detail to bring out the various technical ideas which have had to be used in the process. In § 3 we discuss some general considerations which emerge from the work of § 2, and also some of the wider possibilities of the method as well as its limitations. These points are further illustrated by another example in § 4.

1.2. For the sake of completeness we recall here the necessary definitions. If F_{hkl} represents a general structure factor, we define the *unitary structure amplitudes*, U_{hkl} , by

$$U_{hkl} = F_{hkl} / f_{hkl}. \quad (1.1)$$

Here f_{hkl} is a suitable mean atomic scattering factor for the hkl direction normalized to make

$$U_{000} = 1. \quad (1.2)$$

U_{hkl} is in fact no other than the quantity ${}^a\hat{F}_{hkl}$ introduced by Harker & Kasper and also called \hat{F}_{hkl} by the present writer (Gillis, 1948). The new symbol has been introduced for reasons of typographic simplicity.

If we know the values of $|F_{hkl}|$, we can easily calculate those of $|U_{hkl}|$. Our purpose will be to use this latter information to deduce also the *phases* of the U_{hkl} 's, and these are, of course, the same as the phases

of the corresponding F_{hkl} 's. In this paper we shall be concerned only with real F_{hkl} 's, so that all we have to determine is their signs.

We shall often use a single capital letter to denote a triplet; e.g. U_H for U_{hkl} . Then U_{2H} will represent $U_{2h, 2k, 2l}$, $U_{H+H'}$ will stand for $U_{h+h', k+k', l+l'}$, etc. We shall use the symbol $[hkl]$ to denote the sign of F_{hkl} , i.e. if the phase of F_{hkl} is α

$$[hkl] = e^{i\alpha}. \quad (1.3)$$

As we have already remarked, we shall be dealing in this paper only with cases in which $[hkl] = \pm 1$.

1.3. Our argument will depend on the following inequalities. They are asserted here for the $h0l$ terms of a crystal of space group $P2_1/n$, except for (d) which is asserted for the $0kl$ terms. (a), (b), (c), (e) and (f) are, however, in fact true of *all* the terms of a centrosymmetric crystal, and this fact will be used in § 3. Proofs of (a), (b), (c) and (d) have been given by Harker & Kasper (1948), and of (e) and (f) by the writer (Gillis, 1948):

$$(a) \quad U_H^2 \leq \frac{1}{2}(1 + U_{2H}).$$

$$(b) \quad (U_H + U_{H'})^2 \leq (1 + U_{H+H'}) (1 + U_{H-H'}).$$

$$(c) \quad (U_H - U_{H'})^2 \leq (1 - U_{H+H'}) (1 - U_{H-H'}).$$

$$(d) \quad U_{0kl}^2 \leq \frac{1}{4}(1 \pm U_{0,2k,0} \pm U_{0,0,2l} + U_{0,2k,2l}),$$

the \pm sign being taken according as $k+l$ is even or odd.

$$(e) \quad |U_{3H} + 3U_H|^3 \leq \frac{1}{2}(1 + U_{2H})(3 + 4U_{2H} + U_{4H})^2.$$

$$(f) \quad |U_{3H} + 3U_H| \leq 2(1 + U_{2H}).$$