

satisfactory results in many circumstances. It is always necessary to determine the signs of the φ_{H_1, H_2} by making use of the redundancy of equations. An error made at this step could introduce serious errors in the individual phase angles. In more complex structures, with correspondingly larger values of σ , this danger becomes greater. On the other hand, one would expect that the use of a larger group of related structure factors would improve matters because of the increased redundancy.

References

- BEEVERS, C. A. & ROBERTSON, J. H. (1950). *Acta Cryst*. 3, 164.
- BERTAUT, E. F. (1955). *Acta Cryst.* 8, 823.
- BUERGER, l~I. J. (1951). *Acta Cryst.* 4, 351.
- CLASTRE, J. & GAY, R. (1950). *J. Phys. Radium,* 11, 75.
- GARI~IDO, J. (1950). *C. R. Acad. Sci. Paris,* 231, 297.
- HA~KER, D. & KASPER, J. S. (1948). *Acta Cryst.* 1, 70.
- HAUPTMAN, H. & KARLE, J. (1953). *The Solution of the Phase Problem. I. The Centrosymmetric Crystal.* A.C.A. Monograph No. 3. New York: Polycrystal Book Service.

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

- McLACHLAN, D. (1951). *Prec. Nat. Acad. Sci. Wash. 37,* 115.
- McLACHLAN, D. & HARKER, D. (1951). *Prec. Nat. Acad. Sci. Wash.* 37, 846.

Acta Cryst. (1958). **11,** 115

The Crystal and Molecular Structure of 2 : 5 - Diamino - 4- Mercapto - 6 - Methyl - Pyrimidine

BY E. N. MASLEN, MISS D. E. JUKES AND C. J. B. CLEWS

Department of Physics, University of Western Australia, Nedlands, Western Australia

(Received 10 *April* 1957)

The crystal and molecular structure of 2:5-diamino-4-mercapto-6-methyl-pyrimidine has been determined from projections about the a and c axes of the crystal. The space group is $P2_1/n$, and the unit cell dimensions are $a = 17.036$, $b = 10.045$, $c = 4.339$ Å, $\beta = 90^{\circ}$ 56'. The *hk*O projection has been solved by locating the heavy atom in a modified Patterson synthesis, and refined by calculation of successive difference syntheses. The z coordinates have been determined from the modified *Okl* Patterson, using steepest-descents methods for the refinement of the atomic coordinates.

1. **Introduction**

Following the investigation of the 4:5- and 4:6 diamino-pyrimidines (Clews & Cochran, 1949; White & Clews, 1956), a study was made of the structure of 2:5-diamino-4-mercapto-6-methyl-pyrimidine to obtain information on the type of hydrogen bonding in the amino groups in this class of diamino pyrimidine, and on the positions of the other hydrogen atoms in the molecule. The results indicated the existence of a mixed structure among these atoms.

The calculated structure gave a markedly nonplanar pyrimidine ring, but this is shown to have resulted from the inability of the method of structure determination to cope adequately with the anisotropy and asymmetry of the thermal motions in the molecule.

2. **Experimental**

The crystals of 2:5-diamino-4-mercapto-6-methyl-

pyrimidine were kindly provided by Dr F. L. Rose of I.C.I. Limited, Manchester, England, who obtained satisfactory crystals, in the form of fine needles, by recrystallization from cellosolves.

These crystals were shown by Weissenberg photographs to be monoclinic, space group $P2_1/n$. Using the θ -method of Weisz, Cochran & Cole (1948), the 100 and 010 spacings were measured with Cu K_{α} radiation, and the 015 spacing with Cu $K\alpha_1$. The β angle was obtained from the distance between two sets of 001 reflexions on Weissenberg photographs of a crystal twinned about the h00 plane. From these measurements the lattice parameters were calculated to be

$$
a = 17.036 \pm 0.006, b = 10.045 \pm 0.004,c = 4.339 \pm 0.002 \text{ Å},\n\beta = 90^{\circ} 56' \pm 6'.
$$

$$
(\lambda \text{ (Cu } K\alpha, \text{ mean value}) = 1.5418 \text{ Å},
$$

$$
\lambda \text{ (Cu } K\alpha_1) = 1.5405 \text{ Å}.
$$

Fig. 1. (a) The complete *hkO* vector synthesis. Chain line: zero contour; broken lines: negative contours; contour interval: 200 units on an arbitrary scale. (b) The same synthesis including only high-order terms. Chain line: zero contour; negative contours omitted; contour interval: 50 units on same scale as (a) .

The related S-S vectors at $(2x, 2y)$, $(2x, 0)$, $(0, 2y)$ are indicated by dots.

3. The x and y coordinates of the S, N and G atoms

The *hkO* data were collected using a crystal approximately 1 mm. in length and 0.08 mm. in diameter of cross-section. The intensities were corrected for Lorentz and polarziation factors, but no allowance was made for absorption as the small size and regular shape of the crystal made this negligible.

The Patterson like function

$$
\frac{1}{A}\sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty}|F_{hk}|^2\cos 2\pi\left(hu+\frac{h+k}{4}\right)\cos 2\pi\left(kv-\frac{h+k}{4}\right)
$$

was calculated, but the S-S vectors were not obvious when the complete data were used (Fig. $1(a)$). However, by repeating the synthesis using only those reflexions with $\sin \theta > 0.5$, and thereby making use of the proportionally higher reflecting power of the S atom at large Bragg angles (Fig. $1(b)$), a satisfactory estimate of the S coordinates was obtained and the solution of the complete function in terms of S-S and S-light-atom vectors became possible. A Fourier projection was calculated using signs based on this estimate of the sulphur position, and resolution was obtained immediately, giving a molecule with the anticipated orientation. Structure factors were calculated using the scattering factor curves of McWeeny (1951) for the C and N atoms, and that of Viervoll & \ddot{O} grim (1949) for the S atom. The temperature factor was determined by plotting $\sum I_o/\sum F_c^2$ against sin² θ/λ^2 over ranges of 0.1 in sin² θ . Making allowances for this, the scaling factor was then obtained from the equality $\sum I_c = \sum F_c^2$. The first structure had an R-factor* of 0.285 . By a series of difference syntheses the S, N and C atom coordinates were adjusted till the R -factor was 0.103 , when refinement by this method was complete. A Fourier synthesis, using the final signs of the structure factors (Fig. $2(a)$), is well resolved, with the heights of the C and N peaks significantly different.

4. The x and y **coordinates of the H atoms**

In the difference synthesis calculated using the structure factors from the final set of S, N and C coordinates (Fig. $2(b)$), the positions of all the hydrogen atoms in the structure are not uniquely determined, as there are three peaks of significant height in the neighbourhood of the N_9 atom, and two near the S_8 atom. (For the numbering of the positions in the molecule see Fig. 5.)

* In all R-factor calculations all the terms to the Cu K_{α} limit were included, with the unobserved reflexions set equal to half their maximum possible value.

Fig. 2. (a) The *hk*O Fourier synthesis. Chain line: zero contour; contour interval: 1 e. Å⁻². Five of the contours around the sulphur atom are omitted. (b) The hk0 difference synthesis. Broken lines: negative contours; zero contour omitted; contour interval: 0.2 e.Å⁻². The molecule is drawn in from the corresponding peaks in the Fourier synthesis. Alternative positions for the hydrogen atoms are given by bonds shown as chain lines.

The peaks near the N_7 position are well resolved, though they are not sufficiently clear for the accurate determination of the projected bond lengths. They occur in positions which indicate that $N-H \cdots N$ and $N-H \cdots S$ bonding is possible.

The difference synthesis in the region of the methyl group is not well resolved. However, by assuming that the value of the difference synthesis in this region represents a projection of the electron density of the hydrogen atoms, and by rotating a tetrahedral model till the projected electron density of the model corresponded with the difference synthesis, a satisfactory structure was obtained.

The solution of the remainder of the difference synthesis was possible only after the third coordinates had been determined approximately, at which stage model structures and stacking considerations could be used in the analysis. The significant features could be explained only by assuming the existence of alternative positions for the hydrogen atoms, neither of which involves an imino structure.

In the first alternative the amino hydrogens are at $(0.32, 0.34)$ and $(0.25, 0.28)$. From considerations of the van der Waals radii of the atoms involved it is then expected that the hydrogen attached to the sulphur would be in the $(0.17, 0.06)$ position.

The second structure has the amino hydrogens at $(0.32, 0.34)$ and $(0.29, 0.42)$ with the hydrogen of the mercapto group at $(0.18, 0.12)$. The heights of the peaks in the difference synthesis indicate the occurrence of the alternative forms in about equal numbers.

When allowance had been made in the calculated structure factors on the basis of this mixed structure, using the scattering factors by McWeeny for the hydrogen atoms, the R -factor was reduced to 0.080.

5. **Other features in the** *hkO* **difference** synthesis.

The atomic positions have been adjusted to occur at regions of zero slope in the difference synthesis. No significant features remain in the synthesis except those already ascribed to the hydrogen atom contribution, and one at the sulphur position, 0.6 e. \AA^{-2} in height, which presumably indicates that the sulphur atom has a smaller vibration than that assumed in the temperature-factor expression. This is an acceptable supposition since the sulphur atom is of greater mass than the carbon and nitrogen atoms.

6. Determination of the z parameters

Stacking considerations indicated that some resolution would be possible in the *Okl* plane. Data were collected using a specimen whose depth was approximately 0.08 mm., and cross-section 0.65×0.30 mm. Lorentz and polarization corrections, but no absorption correction, were applied to the observed intensities.

Syntheses similar to those applied to the solution of the *hkO* projection were calculated (Fig. 3(a) and $3(b)$), and that including only the high-angle terms

Fig. 3. (a) The complete *Okl* vector synthesis. Chain line: zero contour; broken lines: negative contours; contour interval: 100 units on an arbitrary scale. (b) The same synthesis including only high-angle terms. Chain line: zero contour; negative contours omitted; contour interval: 20 units on same scale as (a) .

The related S-S vectors are indicated by dots.

had three peaks along the ordinate $y = 0.200$ which could correspond to the S-S vector. Having estimated the differences in the z coordinates of each of the light atoms and the sulphur from their length in projection and the known bond lengths in the pyrimidine ring (Clews & Cochran, 1949), structure factors were evaluated for each possible structure. Error syntheses were calculated for each set and enabled two of the possibilities to be eliminated. The third structure, using an isotropic temperature factor, had an R-factor of 0.44, which was reduced to 0.27 by making adjustments indicated in the error synthesis. A temperature factor expression of the form

$$
F_o = F_c \exp -\{(B_1(kv)^2/\lambda^2 + B_2(l\omega)^2/\lambda^2)\}\
$$

was employed for each $0kl$ reflexion. The value of B_1 was assumed to be equal to the B value in the *hkO* projection, and $B₂$ was then calculated by plotting

$$
\ln \left[\Sigma I_o / \Sigma \{ F_c^2 \exp -2B_1 (k\nu)^2 / \lambda^2 \} \right]
$$

against $(l\omega)^2/\lambda^2$, the summation being made over terms of constant l . The use of this anisotropic temperature factor reduced the R -factor to 0.22.

The refinement of the coordinates was then effected by the use of the steepest-descents formulae of Qurashi (1949), in the course of which it was noticed that the values of the z coordinates, especially of similar atoms close together in the projection, were strongly interdependent. The refinement was continued till no further shifts were indicated, the R-factor at this stage being 0.120.

It was thought that the z coordinates determined by this method could still be improved, as earlier work, including the three-dimensional analysis of Clews & Cochran (1949), had shown the pyrimidine ring to be planar, within the limits of experimental error,

Fig. 4. (a) The 0kl sharpened Fourier synthesis. Chain line: zero contour; broken lines: contours below 10 and above zero; contour interval: $2 e.A^{-2}$. In this synthesis the terms have been sharpened so that the temperature factor is removed. It shows considerably more resolution than the normal Fourier, without radical alteration to any of the peak positions. (b) The 0kl difference synthesis, showing concurrence with expected hydrogen atom positions. Chain line: zero contour; broken lines: negative contours; contour interval: 0.2 e.Å⁻². The bonds in one molecule are drawn in from the Fourier. Bonds to alternative positions for the hydrogen atoms are shown as broken lines.

 δ is the displacement from the calculated mean plane of the molecule: $0.113x(A) - 0.551y(A) + 0.827z(A) + 0.704 = 0.$

	hk0 projection		Okl projection	
	x/a	y/b	y/b	z/c
H_{11}	0.46	-0.09	-0.11	-0.65
H_{12}	0.39	-0.15	-0.15	-0.48
H'_{13}	0.17	0.06	0.06	0.22
H_{13}	0.18	0.12	0.15	0.16
H_{14}	0.32	0.34	0.33	0.40
H'_{15}	0.25	0.28	0.29	0.40
H_{15}	$0 - 29$	0.42	0.42	0.18
\rm{H}_{16}	0.50	0.30	0.27	-0.10
H_{17}	0.44	0.36	0.31	0.21
H_{18}	0.46	0.37	0.37	-0.15

Table 2. *Hydrogen coordinates*

Table 3. *Temperature-factor coefficients* $hk0$ projection

Average temperature factor for all atoms: 3.0_{9} Å² *Okl* projectiom

whereas the present structure, as shown in Table l, deviates significantly from planarity, and some of the calculated bond lengths, involving the use of the coordinates of the atoms where considerable overlap occurs, differ from generally accepted values.

The difference synthesis at this stage $(Fig. 4(b))$ clearly indicated that the discrepancies were due to variation of the temperature factor in various parts of the molecule. It was expected that the direction of maximum vibration of the atoms would be perpendicular to the plane of the pyrimidine ring. As the molecule is tilted out of the $hk0$ plane by about 30° , this would imply that the temperature-factor ellipse would be asymmetric to the symmetry elements of the cell. The result of this would be that an atom at x, y, z, would, for any reflection *Okl,* have a different scattering factor from two of the atoms related to it by symmetry. It was thought unlikely, however, that this would have any effect on the z coordinates in the cell. To verify this an allowance was made for the effect by the method of Rollett (1955). Although it proved possible to alter the coordinates somewhat by varying the two temperature-factor parameters it was possible to effect no improvement on the structure by this means, that with the minimum R value being the same as that previously determined. The difference synthesis on the structure factors including this allowance had approximately the same significant features as that calculated with a symmetrical temperature factor. The lack of improvement in the R factor indicated that this type of thermal motion was not responsible for the discrepancies in the structure.

7. Verification of the postulated hydrogen-atom coordinates, and analysis of the *Okl* **difference synthesis of the complete structure**

Although the light-atom coordinates were still slightly in error at this stage, the resulting discrepancies in the structure factors would vary considerably with the indices of the reflexion. As the values of the structure factors with low l index are not much affected by small errors in the z coordinates, the standard deviation in these would be only slightly greater than that in the corresponding *hkO* structure factors. Owing to the small c spacing, and the rapid drop in its scattering factor with increasing Bragg angle, the hydrogenatom contribution to the structure factors is confined almost entirely to the terms with low l value, so that in a difference synthesis these terms should form significant peaks at the hydrogen positions. The remaining terms due to errors in light-atom coordinates have a much smaller temperature factor, so it is unlikely that they would obscure the effect of the hydrogen atoms. Peaks do, in fact, occur at approximately the expected positions. Slight shifts appear to result from overlap with other features in the synthesis, and a small rotational shift of the methyl group from the orientation previously assumed is necessary. The

inclusion in the structure factors of the hydrogen-atom contribution, using an average isotropic temperature factor for these atoms, reduces the R -factor to 0.106. As expected, this allowance has little effect on the terms with high 1 index, and therefore on the z coordinates of the light atoms.

As has been observed earlier, the sulphur atom, being heavier than the nitrogen and carbon atoms, has a smaller amplitude of thermal vibration. Decreasing B by 0.15 \tilde{A}^2 eliminated the peak at the sulphur position in the $hk0$ difference synthesis, so that, as an approximate allowance for the heavier mass of the sulphur atom, the same correction was applied to each B value in the *Okl* structure factors. This reduced the R-factor to 0.104, and the standard deviation of the electron density to 0.38 e. \AA^{-2} .

In a difference synthesis the atoms C_2 , N_3 , and N_7 (shown on the left-hand side of the molecule in Fig. 5)

Fig. 5. Diagram showing the *nomenclature* used in describing the structure, bond lengths and interatomic distances. Bonds are shown as unbroken straight lines, except that those to hydrogen atoms in alternative positions are shown as light broken lines. Hydrogen-bonded vectors are shown as heavy broken lines and unbonded interatomie distances as chain lines.

The diagram is a projection of the bond lengths on to the $hk0$ plane, and is drawn to scale. The bond lengths calculated from the observed structure are shown above the bond, with the standard deviation calculated as indicated in the text. The values in parenthesis are the calculated bond lengths if the atoms in the ring are assumed to have the z coordinate of the calculated mean plane.

are at heights of about -0.3 e. \AA^{-2} . Those on the righthand side $(C_5, C_6, N_9 \text{ and } C_{10})$ are at approximately zero, while the atoms along a central line in the molecule $(N_1, C_4 \text{ and } S_8)$ are at 1.0, 0.0, and 0.5 e. \AA^{-2} , though the C_4 height is expected to have been reduced somewhat because of the proximity of the S_8 and N_1 peaks. That the peak heights occur in this order is quite significant. For example, if the errors in the synthesis were randomly distributed the probability of the peak heights at the C_2 , N_3 and N_7 positions being simultaneously < 0.3 e. Å⁻² is approximately 0.020. The most probable explanation of the difference synthesis is that the molecule oscillates about an axis approximately through the centre of gravity of the molecule and the sulphur atom, so that the left-hand side of the molecule would vibrate more than the righthand side, with the vibration of the central atoms less than that of either side. This would account satisfactorily for all the discrepancies in the structure. To allow for this would involve, in the determination of the z coordinates, the use of one positional and two vibrational parameters per atom, totalling thirty parameters in all. However there are only 46 non-zero reflexions available for the solution of the equations involving these parameters, so that the structure is not sufficiently over-determined for the results of a more accurate refinement to be significant. It would be necessary, therefore, if greater accuracy were to be attained, to use more data than are available in this projection, preferably a full three-dimensional calculation; it is hoped to undertake this work at a later stage to solve fully several further problems in the chemistry of this pyrimidine derivative which have been suggested by the present structure determination.

8. Accuracy

In the *hkO* projection the formulae of Cruickshank (1949) for the standard deviation of the electron density and atomic coordinates are applicable. Putting $\delta F = F_o - F_c$ in the expression $\sigma(\rho) = (\Sigma \delta F^2)^{\frac{1}{2}}/A$, the standard deviation of the electron density is estimated to be 0.20 e. \AA^{-2} , which justifies the earlier assumption that the hydrogen atoms would show significant peaks in the difference synthesis.

The standard deviation of the atomic coordinates is given by

$$
\sigma(x) = \frac{2\pi}{Aa} \frac{(\sum h^2 \delta F^2)^{\frac{1}{2}}}{|\delta^2 \varrho/\delta x|}.
$$

For the S, N and C atoms $\frac{\partial^2 \rho}{\partial x^2}$ was obtained from the Fourier projection by graphical methods, assuming an electron-density distribution of the form

$$
\varrho(r) = \varrho(0) \exp -(pr^2).
$$

In the case of the hydrogen atoms $\frac{\partial^2 \rho}{\partial x^2}$ was estimated from the well resolved H_{14} peak in the difference synthesis. This clearly underestimates the

error for those atoms in the hybrid system, and for the partly resolved peaks in the methyl group.

The standard deviations obtained by this method for the *hkO* projection, were

For the *Okl* projection Cruickshank's formula gives 0.38 e.Å^{-2} for the standard deviation of the electron density. This includes errors due to temperature-factor parameters and series termination, and random errors. Owing to the high degree of overlap in this projection, Cruickshank's formula for the standard deviation of the z coordinate is inapplicable; furthermore it was not found possible to calculate the accuracy of the z-coordinate determination directly from the data.

An approximate average value of $\sigma(z)$ is given by the r.m.s, deviation of atoms in the pyrimidine ring from the calculated mean plane of the molecule. This estimate is ± 0.13 Å⁻². This necessarily overestimates the average standard deviation for it includes the error of the mean plane in addition to errors in individual coordinates, and the calculation gives higher weight to the atoms most subject to high overlap and temperature-factor errors. The displacements of substituent atoms from the mean plane have been omitted in the calculation, as earlier work has shown that the bonds to substituent atoms are frequently tilted out of the plane of the pyrimidine ring. The errors in the calculated bond lengths (Fig. 5) will therefore also be overestimated, particularly for bonds involving substituent atoms.

9. Discussion

In simple pyrimidines the 2-, 4-, and 6-positions are highly electron deficient, with the 5-position less so, but still not as fully saturated as an aromatic nucleus. The presence of electron-releasing substituents in the ring tends to reduce this deficiency. In a review of the available chemical data, Brown (1953) showed that there was some evidence that this would result in the presence of tautomeric forms if the electron-releasing substituent were on the 2-, 4-, or 6-position. In the present molecule tautomerism could conceivably exist between the forms (I) – (IV) in addition to the normal form (V).

The structure determination shows clearly that if any of the tautomeric forms do exist in the crystalline state they do so in proportions which cannot be detected crystallographically. It is apparent, though, that there is some interaction between the substituent groups and the ring at the 2- and 4-positions, for the N_7-C_2 bond length (1.33 Å) and the S_8-C_4 bond length (1.71 Å) are shorter than the corresponding normal single-bond lengths, 1.47 and 1.81 Å respectively (Pauling, 1948). As expected, the amino group on the 5-position does not show this interaction, the N_9-C_5 bond length (1.47 Å) agreeing closely with the normal single-bond length.

There is strong evidence that in the case of the $N₇$ atom the electron releasing is effected, at least in part, by hydrogen bonding. It has already been shown that the hydrogen atoms are directed favourably for bonding to the neighbouring S_8 and N_1 atoms. Although it is rather surprising that such bonding would connect to SH, itself an electron-releasing group, confirmation of the bonding to the sulphur is readily obtained from the interatomic distance between these two atoms. The distance of closest approach in van der Waals contact was obtained by modifying the value obtained by Wright (1956) for the distance of closest approach of an oxygen to a sulphur atom in glutathione (3.47 Å) , to take account of the difference in van der Waals radius of the amino group in this orientation and of the oxygen atom, giving a net value

of 4.10 A. That actually occurring in the crystal is 3.40 A, indicating the existence of strong hydrogen bonding between these two atoms. Similar calculations have been made for the N_1-N_7 distance, confirming the bonding between these two atoms.

The N_g-N_3 and N_g-S_8 vectors form an approximately tetrahedral system similar to that around the N_7 atom. The lengths of these vectors (2.96 and 3.45 \AA) respectively) are not, however, significantly shorter than the normal N-N and N-S van der Waals contacts $(3.00 \text{ and } 3.35 \text{ Å}$ (Pauling, 1948)) unless a hydrogen atom is interposed. This does not appear to be the case, for the hydrogens tend to occupy positions determined by stacking considerations rather than hydrogen-bonded sites.

In general the structure based on chemical evidence has been well substantiated, except that it appears, at least in the crystalline state, that the electronreleasing groups tend preferentially to form hydrogen bonds rather than tautomeric states. This is perhaps due to the high energy barriers in a crystalline structure, which prevent tautomeric transitions.

The authors wish to acknowledge the generous financial assistance of the Research Grants Committee of the University of Western Australia. Sincere thanks are also due to Messrs D. E. Henshaw and G. B. Robertson for their assistance in the early stages of this work, and to Dr D. M. Hodgkin, F. R. S., for her advice on some of the methods in the later stages of the refinement of the structure.

References

- BROWN, D.J. (1953). *Rev. Pure Appl. Chem. 3,* 116.
- CLEWS, C. J. B. & COCHRAN, W. (1949). *Acta Cryst.* 2, 46.
- CRUICKSHAHK, D. W. J. (1949). *Acta Cryst. 2,* 65.
- MCWEEHY, R. (1951). *Acta Cryst.* 4, 513.
- PAULING, L. (1948). *The Nature of the Chemical Bond*, 2nd ed. Ithaca: Cornell University Press.
- QURASHI, M. M. (1949). *Acta Cryst.* 2, 404.
- ROLLETT, J. S. (1955). *Acta Cryst.* 8, 487.
- VIERVOLL, H. & ()GRIM, O. (1949). *Acta Cryst.* 2, 277.
- WEISZ, O., COCHRAN, W. & COLE, W. F. (1948). *Acta Cryst.* 1, 83.
- WHIWE, N.E. & CLEWS, C. J. B. (1956). *Acta Cryst. 9,* 586.
- WRIGHT, W. B. (1956). *Chem. Ind. Rev.* No. 20, 437.