Acta Cryst. (1962). **15**, 1174

The Crystal Structures of Salts of Methylated Purines and Pyrimidines. I. 1-Methylcytosine Hydrobromide

By Robert F. Bryan* and Ken-Ichi Tomita

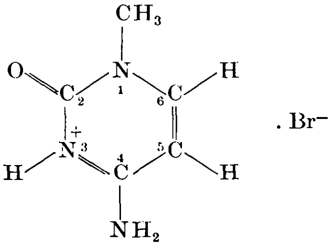
Department of Biology, Massachusetts Institute of Technology, Cambridge 39, Mass., U.S.A.

(*Received 5 September* 1961)

Crystals of 1-methylcytosine HBr are orthorhombic with $a = 12{\cdot}98$, $b = 6{\cdot}80$, and $c = 8{\cdot}83$ Å. The space group is *Pnma* with 4 molecules in the unit cell, all the atoms lying in the mirror planes at $y = \frac{1}{4}$ and $\frac{3}{4}$. Bond lengths and angles have been determined from the (010) projection, $\sigma(x, z) = 0{\cdot}02$ Å, and the dimensions of the pyrimidine ring are in fair agreement with previous predictions. The acidic proton has been identified as being attached to N(3) of the cytosine ring which exists in the keto–amino form in this crystal. No inter-pyrimidine hydrogen bonds are formed, the molecules being held together in layers, in an interlocking double dovetail arrangement, by a planar system of hydrogen bonds of the type N–H · · · Br. The molecular packing is somewhat unusual, the Br^- ions of one layer being sandwiched between the pyrimidine rings of the layers immediately above and below.

Introduction

The first substance whose crystal structure we have determined in this study of the salts of methylated purines and pyrimidines is the hydrobromide of 1-methylcytosine.



Preliminary details of this structure analysis have been given previously (Bryan & Tomita, 1961).

Experimental

Crystals of 1-methylcytosine hydrobromide suitable for X-ray study were prepared by the addition of a slight excess of hydrobromic acid to an aqueous solution of the pyrimidine, which was then allowed to evaporate to dryness at room temperature. The salt crystallises as colourless needles elongated along the *b*-axis. The unit cell dimensions and space group were determined from 30° precession photographs taken with Cu $K\alpha$ radiation.

Crystal data

1-Methylcytosine hydrobromide, $C_5H_7N_3O.HBr$, $M = 216{\cdot}1$.

Orthorhombic prismatic,

$$a = 12{\cdot}98 \pm 0{\cdot}04,\ b = 6{\cdot}80 \pm 0{\cdot}02,\ c = 8{\cdot}83 \pm 0{\cdot}03 \text{ Å}.$$

Volume of the unit cell, 776·7 Å^3, $D_m = 1{\cdot}85$ g.cm.$^{-3}$, $Z = 4$, $D_c = 1{\cdot}87$ g.cm.$^{-3}$, $F(000) = 408$, $\mu = 71{\cdot}4$ cm.$^{-1}$ for Cu $K\alpha$ radiation.

The systematic absences in the X-ray diffraction pattern indicate as possible space groups *Pnma* or $Pn2_1a$. The permitted $0k0$ reflections are uniformly very strong, showing only normal decline, and the diffraction patterns for $h0l$ and $h2l$, and for $h1l$ and $h3l$, are identical in pairs. Thus, it was assumed that the more likely space group was *Pnma* which requires the molecules to be situated in the mirror planes of symmetry at $y = \frac{1}{4}$ and $\frac{3}{4}$. This choice is confirmed by the subsequent analysis.

Intensity data

Intensities were recorded photographically by using the multiple-film technique (Robertson, 1943), with Cu $K\alpha$ radiation, from a needle-shaped crystal of dimensions 0·2 × 0·15 mm. perpendicular to the needle axis. The intensities were measured visually by making use of a scale, prepared from the same crystal, on which a typical reflection was recorded for known lengths of time under the same experimental conditions. The range of intensity recorded on the moving-films was from 1 to 3000. No absorption corrections were applied to the measured intensities, though in some cases these may be sizeable. 134 $h0l$ reflections were recorded, 122 of which had a measurable intensity. For $h1l$ these figures were 132 and 108 respectively, and for $h2l$, 140 and 119. In each case reflections were recorded out to the limit of Cu $K\alpha$ radiation. Lorentz and polarization corrections were applied to the intensity estimates which were then

---

* Present address: Department of Biophysics, The Johns Hopkins University School of Medicine, Baltimore, Maryland, U.S.A.

placed on an approximately absolute scale by making use of Wilson's (1942) method. This same method yielded a preliminary value for the exponent B, in the temperature factor, of 3·0. The corrected intensities were then converted to structure amplitudes by square-rooting.

Structure determination

The crystallographic requirement, embodied in the choice of space group *Pnma*, that the molecules lie in the mirror planes of symmetry at $y=\frac{1}{4}$ and $\frac{3}{4}$, means that only two positional parameters remain to be determined for each atom in the asymmetric unit to specify the structure completely. In the present case, the asymmetric unit consists, stoichiometrically, of one molecule of 1-methylcytosine and one molecule of hydrogen bromide. It was decided to try to determine the necessary positional parameters from the projection of the structure down [010] and, if this attempt should fail, to make use of the partial three-dimensional intensity data which had already been collected. This attempt to solve the structure from a single projection was successful and the three-dimensional intensity data were only used to check the correctness of the proposed structure in the final stages of the analysis.

To solve the phase problem, the heavy atom method was chosen with the Br^- ion being used as the phase determining agent. To find the x- and z-coordinates of the bromine ion the Patterson function $P(u,w)$ was calculated, with quantities $I.Lp^{-2}$ being used as coefficients, where I represents a particular intensity and Lp the appropriate Lorentz and polarization factor. This double application of the correction produces a useful sharpening effect on the peaks in the Patterson. The heavy atom was located then at the position $x=0·113$, $z=0·208$, with respect to the centre of symmetry of the projection as origin. These coordinates were used to calculate the contribution of the bromine ion to the $h0l$ structure factors and in this way about 100 of these reflections seemed to be phase determined by the heavy atom. The value of $R=\Sigma(|F_o|-|F_{\text{Br}}|)/\Sigma|F_o|$ was 0·32 for these reflections.

The terms thus phased were included in a Fourier synthesis to give a preliminary electron-density distribution $\varrho(x,z)$ which proved difficult to interpret in terms of our preconceptions as to what type of structure was likely to be found. The bromine ion was surrounded by what appeared to be a substantial diffraction maximum and comparatively few other peaks of high density were to be found. A trial structure was first sought on the assumption that, although there might be considerable overlap in projection between the pyrimidine rings, there was unlikely to be any great overlap between the rings and the bromine ions. No satisfactory structure could be found on this basis, however, and it became evident that what we had designated a diffraction ripple around the bromine ion was, in fact, electron density resulting from the presence of several light atoms. There was, indeed, no overlap between the cytosine rings and each bromine lay, in projection, almost at the centre of one such ring. Once this was realized, the correct structure was soon found and was rapidly refined by means of F_o and (F_o-F_{Br}) syntheses. One such (F_o-F_{Br}) synthesis is shown in Fig. 1 which also shows the relationship between the molecules in the different layers. The value of the conventional index R at this stage was 0·18.

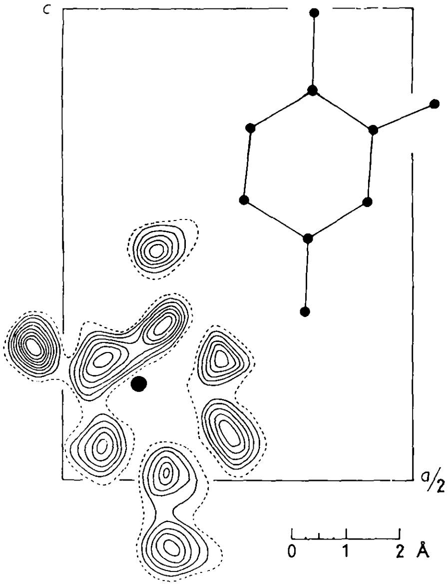


Fig. 1. Electron-density projection $\varrho(x,z)$. The contribution of the bromine atom has been subtracted from the structure factors. The contoured atoms are at $y=\frac{1}{4}$, the position of the molecule at $y=\frac{3}{4}$ is indicated by the solid lines. The bromine atom at $y=\frac{3}{4}$ is represented by the large black circle.

Table 1. *Atomic coordinates expressed as fractions of the unit cell edges*

Atom	x	y	z
Br	−0·1088	0·2500	−0·2034
N(1)	0·1367	0·2500	0·3240
N(3)	0·0596	0·2500	0·0850
N(7)	0·1503	0·2500	−0·1462
C(2)	0·0498	0·2500	0·2365
C(4)	0·1471	0·2500	0·0068
C(5)	0·2377	0·2500	0·0906
C(6)	0·2257	0·2500	0·2434
C(8)	0·1319	0·2500	0·4912
O	−0·0373	0·2500	0·2912
H, N(3)	−0·0084	0·2500	0·0272
H, N(7)	0·0862	0·2500	−0·2134
H, N(7)	0·2187	0·2500	−0·2020
H, C(5)	0·3095	0·2500	0·0272
H, C(6)	0·2926	0·2500	0·3201
H, C(8)	0·0524	0·2500	0·5312
H, C(8)	0·1802	0·1150	0·5357
H, C(8)	0·1802	0·3850	0·5357

Refinement

The structure was refined by means of $(F_o - F_c)$ syntheses making use of the $h0l$ reflections. The first such synthesis gave good indications of the presence of all of the hydrogen atoms in the structure though it was not possible to decide upon the proper orientation of the methyl group. The acidic hydrogen was readily identified as being attached to N(3). Contributions from all atoms save the methyl hydrogens were then included in a structure factor calculation and a second difference Fourier calculated. From this, and

Table 2. *Table of observed and calculated structure factors*

Accidentally absent spectra are omitted for brevity

h 0 0	F_o	F_c
2	24.1	22.6
6	49.5	46.2 -
8	56.8	57.0
10	10.2	7.6
12	31.0	26.7 -
14	22.3	22.4 -
h 0 1		
1	44.8	43.0
2	143.2	131.6 -
3	21.0	14.9 -
4	43.9	41.8 -
6	66.8	66.5
8	65.0	64.2
10	41.0	31.4 -
12	26.8	24.3 -
16	21.2	22.3
h 0 2		
1	37.4	38.9 -
2	36.7	36.4 -
3	42.8	50.1 -
4	45.1	46.1
5	44.6	43.1 -
6	51.4	47.2
7	35.9	34.7
8	25.4	30.2 -
10	34.8	33.5 -
11	21.9	22.0 -
12	23.7	18.5
14	29.3	27.7
15	23.2	16.1
h 0 3		
1	56.6	54.7 -
2	56.8	60.8
3	51.2	49.7
5	64.1	61.5
6	49.8	47.9 -
8	38.6	30.9
9	50.1	43.3 -
10	16.4	16.6
11	14.1	13.6 -
12	25.2	20.8
13	23.5	18.4
15	15.6	11.0
16	8.7	11.4 -
h 0 4		
0	71.5	62.7
1	23.6	23.4
2	29.4	14.6
3	63.0	69.2
4	26.8	24.3 -
5	11.5	4.4 -
6	25.0	12.6 -
7	47.2	46.6 -
10	15.0	11.9
11	46.2	42.4
13	20.3	17.5
15	18.3	15.7 -
h 0 5		
1	35.2	33.5
2	28.0	20.4 -
3	31.2	36.6 -
5	55.6	58.4 -
7	7.2	5.1
9	46.3	43.9
10	6.3	3.5 -
11	19.2	20.6
13	20.4	18.9 -
15	13.9	13.1 -
h 0 6		
0	8.0	13.6
1	41.3	45.0 -
3	53.5	60.2 -
4	40.0	41.8 -
6	21.2	17.9 -
7	23.7	32.4
8	25.3	18.3
11	22.6	26.8 -
h 0 7		
1	32.0	33.8 -
2	29.4	22.1 -
5	36.5	41.8
6	32.4	34.0
7	16.6	7.8
9	29.0	31.9 -
10	9.0	6.3 -
11	13.7	6.0 -
12	10.8	12.5 -
13	14.0	14.8
h 0 8		
0	21.0	22.2 -
1	28.3	32.6
3	33.4	32.1
4	23.6	21.4
7	16.6	19.2 -
8	13.6	11.2 -
10	11.0	16.6 -
11	11.0	10.7
h 0 9		
1	8.5	13.3
2	29.6	29.9
4	9.0	8.7
5	11.0	13.5 -
6	22.4	23.5 -
h 0 10		
0	24.4	30.5
4	24.8	22.2 -
h 1 0		
2	10.4	8.0
4	59.0	64.4
6	54.5	50.9
8	51.7	56.4 -
10	18.5	18.2
12	61.2	66.9
14	5.7	3.3 -
h 1 1		
0	21.1	29.2
1	6.9	7.2 -
2	3.5	2.7
3	32.2	29.1
4	109.9	123.6 -
5	32.2	29.9 -
6	28.5	29.6 -
7	42.7	31.8 -
8	50.0	47.5
10	59.3	56.9
11	32.7	24.8
14	26.5	24.9 -
h 1 2		
1	54.0	62.2
2	102.9	107.2 -
3	61.2	52.1 -
4	53.2	43.1 -
5	52.9	50.8 -
6	51.8	58.5
7	39.8	31.7
8	32.5	24.3
9	31.7	29.4
10	16.7	23.7 -
12	33.2	30.7 -
13	21.0	16.8 -
15	8.0	7.7 -
h 1 3		
0	43.7	49.9 -
1	42.2	40.5 -
2	4.6	3.4 -
3	52.2	53.4 -
4	53.8	54.9
5	38.8	33.7
6	39.2	38.2
7	55.2	66.6
8	26.0	28.0 -
10	33.5	27.7
11	29.3	28.1 -
13	8.5	11.1 -
14	19.0	15.0
15	10.3	10.3
h 1 4		
1	83.9	99.1 -
2	27.0	26.2
3	40.5	37.9
4	26.9	17.4
5	62.5	63.3
6	13.0	17.4 -
7	7.3	8.1 -
8	8.9	10.6 -
9	43.9	37.1 -
10	3.2	3.7
12	19.7	14.6
13	21.2	20.1
h 1 5		
0	40.5	23.6 -
1	56.3	57.7
3	55.8	59.0
4	12.2	14.2 -
5	25.7	21.6 -
7	65.3	65.2 -
10	15.5	13.2
11	30.0	26.9
13	7.8	9.0
h 1 6		
1	53.5	52.3
3	16.9	18.0 -
5	50.0	48.4 -
6	14.5	22.0
8	14.8	19.2 -
9	40.3	36.9
13	15.0	22.3 -
h 1 7		
0	26.7	29.8
1	28.5	23.4 -
3	29.8	31.0 -
5	38.3	30.8
6	15.8	21.4 -
7	39.7	34.6
8	8.6	5.4
9	7.3	6.3
11	17.3	17.4 -
12	5.1	1.9 -
h 1 8		
1	16.7	16.9 -
2	25.8	36.1 -
5	23.2	21.9
6	23.0	16.4
9	15.3	18.7 -
10	10.6	12.8 -
11	9.5	10.1 -
h 1 9		
0	17.5	22.1 -
3	14.7	11.1
4	19.7	23.7
6	8.8	13.7
7	17.2	14.6 -
8	16.7	14.1 -
h 1 10		
1	5.1	4.5
2	17.8	23.9
h 1 11		
0	12.8	19.2
h 2 0		
2	24.2	24.3 -
6	48.1	43.2
8	50.4	53.0 -
10	12.1	7.5 -
12	25.2	24.5
14	22.3	20.3
h 2 1		
1	40.0	36.4 -
2	108.2	108.6
3	19.3	17.2
4	37.5	34.5
5	14.8	17.4
6	57.1	58.6 -
8	54.3	56.8 -
10	37.0	28.4
12	24.3	22.3
h 2 2		
1	33.9	31.0
2	30.3	27.8
3	39.3	43.7
4	46.1	43.0 -
5	30.7	34.6
6	38.0	39.4 -
7	29.7	32.0 -
8	30.9	26.3
9	8.7	7.4 -
10	30.8	30.7
11	19.9	20.1
12	21.6	16.5 -
13	9.6	10.7
14	32.3	25.6 -
15	22.4	14.7 -
h 2 3		
1	51.0	45.9
2	54.0	51.3 -
3	37.5	39.7 -
5	56.2	55.9 -
6	42.3	42.6
8	31.5	28.1
9	8.7	7.4 -
10	16.0	15.5 -
11	12.4	12.0
12	22.8	18.9
13	20.0	16.9
15	14.6	10.1
h 2 4		
0	57.7	51.1 -
1	23.0	24.7 -
2	23.3	13.2 -
3	61.2	61.0 -
4	23.1	21.6
5	12.7	5.3
6	20.4	12.1
7	39.5	41.5
10	16.6	10.7 -
11	39.4	38.5 -
12	10.4	12.2
13	18.3	16.2 -
14	14.4	8.4
h 2 5		
1	36.8	31.5 -
2	25.0	18.8
3	28.8	33.2
5	44.0	52.2
9	36.6	39.6 -
11	21.8	19.0 -
13	21.7	17.4
14	14.5	12.2
h 2 6		
0	9.0	11.7 -
1	31.3	39.5
3	51.9	54.3
4	33.9	37.1
6	21.9	15.9 -
7	23.4	30.3 -
8	22.1	16.7 -
11	22.9	24.7
13	9.5	9.9
h 2 7		
1	27.6	30.9
2	25.1	20.7
5	34.4	31.0 -
6	25.9	30.6 -
7	14.9	6.9 -
9	30.0	29.3
10	11.6	5.6
11	10.7	5.5
12	10.5	11.6
h 2 8		
0	19.2	20.0
1	24.9	30.0 -
3	25.4	29.1 -
4	21.9	19.5 -
6	12.5	13.9 -
7	17.1	17.5
8	16.1	10.2
10	13.2	15.3
11	8.8	9.9 -
h 2 9		
1	15.0	12.3 -
2	26.6	27.2 -
5	13.4	12.4
6	20.6	21.7
8	12.3	13.5
9	12.5	9.1 -
h 2 10		
0	23.9	28.0 -
3	9.0	5.9
4	19.9	20.5
6	9.8	11.5

three further such syntheses, minor shifts were made in the positions of the atoms to give the coordinates listed in Table 1. In the case of the hydrogen atoms the coordinates listed correspond closely to the centres of peaks of positive electron-density, of heights ranging from 0·3 to 1·2 e.Å^{-2}, occurring in the first $(F_o - F_c)$ synthesis. No attempt has been made to refine these positions save in the case of the methyl–hydrogen atoms which have been placed in calculated positions corresponding to what is believed to be the correct orientation of this group. The reasons for this supposition are given in the subsequent section.

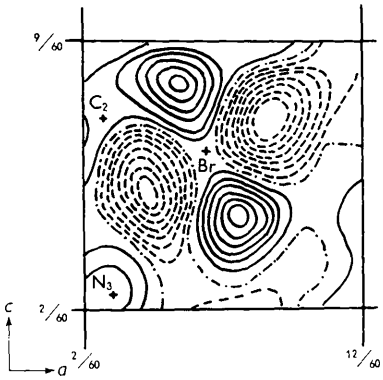


Fig. 2. Illustration of the anisotropic thermal motion of the bromine ion as illustrated by the fourth $(F_o - F_c)$ synthesis. Contours approximately every 0·2 e.Å^2.

The final difference Fourier showed quite clearly the presence of anisotropy in the thermal vibration of the bromine ion. The appearance of this feature is shown in Fig. 2. No allowance was made for this anisotropy in calculating structure factors, a value of B of 2·75 being applied in calculating the isotropic temperature factors for all the atoms. The scattering functions used were: for Br, the curve of Thomas & Umeda (1957); for C, N, and O, the functions of Berghuis *et al.* (1955); and for hydrogen, the curve due to McWeeny (1951). Structure factors were calculated on an IBM 709 data processing machine using the program of A.C.T. North; Fourier syntheses were computed on the same machine with the program MIFR 1 (Shoemaker & Sly, 1961). Calculations of bond distances and angles were made by using a program due to R. Langridge.

The conventional residuals, R, were calculated, excluding unobserved reflections, as 0·117 for $h0l$, 0·123 for $h1l$, and 0·121 for $h2l$. Lists of structure factors are given in Table 2, based on the coordinates given in Table 1. Unobserved reflections are omitted for the sake of brevity, no anomalously high values of F_c were found for any of these.

Description and discussion of the structure

The bond lengths and angles found in the present analysis for the cation of 1-methylcytosine are shown in Fig. 3, which also shows the models proposed by Pauling (1960) and by Spencer (1959) for the molecule of cytosine on the basis of previous X-ray analyses of related compounds.

The standard deviation in the position of a light atom as calculated by Cruickshank's (1949) method is $\sigma(x, z) = 0{\cdot}02$ Å. Since there is overlap in projection between some of the atoms of the pyrimidine ring and the bromine ion this figure probably represents an underestimate and we would suggest that the standard deviation in a bond length involving light atoms is

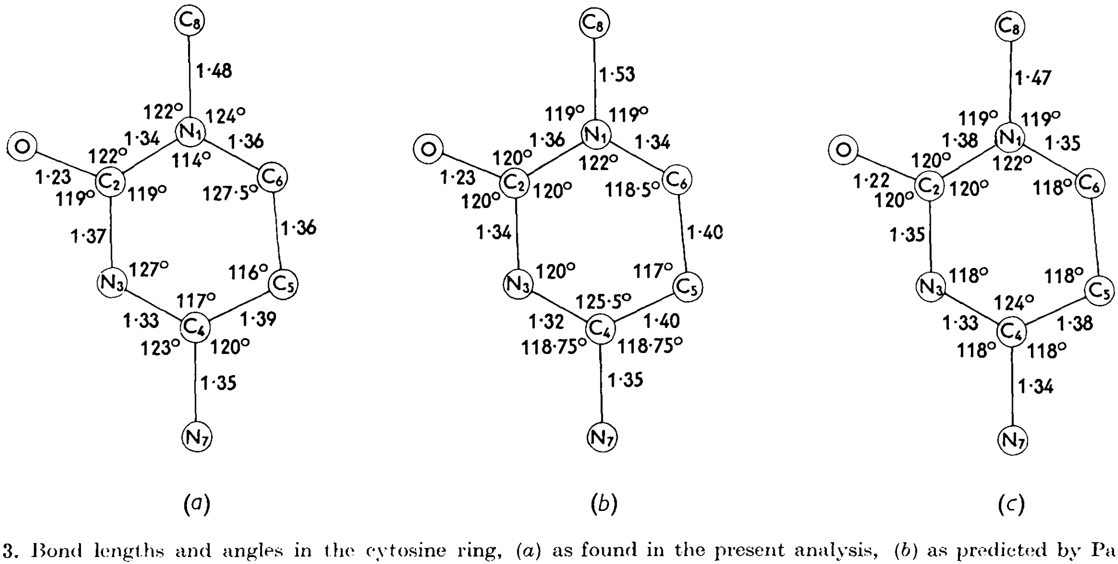


Fig. 3. Bond lengths and angles in the cytosine ring, (*a*) as found in the present analysis, (*b*) as predicted by Pauling, (*c*) as predicted by Spencer.

probably around 0·04 Å, and in a bond angle is correspondingly about 2°. The standard deviation in the position of the bromine ion is 0·004 Å.

The bond distances which we find for the 1-methylcytosine cation are then in good agreement with those proposed by both Pauling and Spencer, though the size of our standard deviation does not allow us to prefer one of these models over the other. From the bond lengths, and from the hydrogen bonding scheme, we would deduce that the cytosine cation exists in the keto-amino form with the acidic proton attached to N(3). There are some sizeable differences between the bond angles which we measure for the cation and those angles embodied in either of the model structures. In particular, the angle at N(1) is 114° rather than the 122° suggested by both Pauling and Spencer and other differences of equal magnitude occur and would appear to be significant, though in view of the overlap mentioned we would hesitate to be overconfident on this point.

The molecules have been assumed to lie in the space group mirror planes and are thus, by definition, planar. A packing drawing of the molecules within a given layer is shown in Fig. 4. The van der Waals' radii are inscribed, the following values being used: C, 1·7 Å; N, 1·5 Å; O, 1·4 Å; H, 1·0 Å; and Br^-, 2·00 Å. The molecules are packed together in an interlocking double dovetail arrangement and it is obvious that this packing is very efficient. Various intermolecular distances are given in Fig. 4.

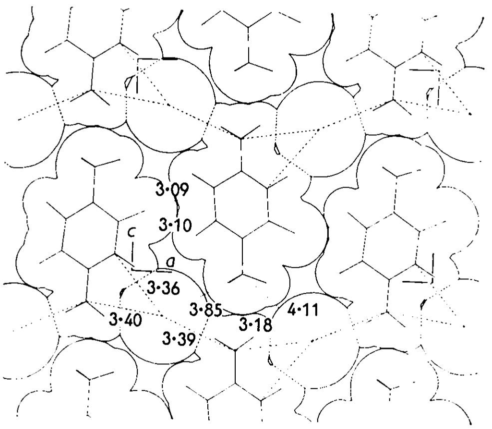


Fig. 4. Packing drawing showing the arrangement of molecules within a given layer of the structure. Conventional van der Waals' radii are inscribed. One unit cell is outlined and certain intermolecular approaches are given. –·–·–· hydrogen bonds, ······ normal van der Waals' interactions.

There are no inter-pyrimidine hydrogen bonds in the structure, the principal forces operating to hold the molecules together being the array of N–H · · · Br hydrogen bonds running in the direction of the a-axis. The lengths of these bonds, 3·36, 3·39, and 3·40 Å are unremarkable; more worthy of comment is the way in which the bromine ion is arranged so as to participate in two hydrogen bonds N(3)–H · · · Br and N(7)–H · · · Br. The deviation from linearity in each of these bonds is over 25° and it would seem that there is a very strong electrostatic contribution to the energy of both.

From the distances of the methyl-carbon atom from each of the two neighbouring bromine ions it was decided that the orientation of the methyl-group was such that the in-plane hydrogen atom points towards the bromine ion 4·11 Å distant and that the out of plane hydrogens are directed towards the corresponding ion 3·85 Å distant. Free rotation of the methyl-group about the C–N bond is probably prevented in the crystal by the shortness of this latter approach.

The oxygen atom of one molecule in a given layer is equidistant from the atoms C(5) and C(6) of an adjacent molecule. To accomodate this the hydrogen atom attached to C(6) must be bent somewhat from the line bisecting the ring angle at C(6). This distortion would require but little energy to accomplish and the position of the electron-density peak assigned to this hydrogen atom indicates that it does indeed take place.

The shortest contacts between the layers are between the bromine ions of one layer and the atoms N(1), N(3), and C(2), in the layers above and below. The distances N(1) · · · Br and N(3) · · · Br are 3·58 and 3·61 Å respectively and are unexceptional. The distance C(2) · · · Br of 3·50 Å, on the other hand, is somewhat shorter than the sum of the conventional van der Waals' radii 3·6 to 3·7 Å. Whilst this difference is quite small it does indicate that the 'radius' is declining in the direction of the centre of the pyrimidine ring.

This work was supported by grants from the National Cancer Institute of the U.S. Public Health Service and the American Cancer Society. One of us (R. F. B.) was in receipt of a U.S. Public Health Service Postdoctoral Fellowship.

We are grateful to the Computation Centre M.I.T. for the use of calculating facilities.

References

Berghuis, J., Haanappel, IJ. M., Potters, M., Loopstra, B. O., MacGillavry, C. H. & Veenendaal, A. L. (1955). *Acta Cryst.* **8**, 478.

Bryan, R. F. & Tomita, K. (1961). *Nature, Lond.* **192**, 812.

Cruickshank, D. W. J. (1949). *Acta Cryst.* **2**, 65.

McWeeny, R. (1951). *Acta Cryst.* **4**, 513.

Pauling, L. (1960). *The Nature of the Chemical Bond.* 3rd edn. Ithaca: Cornell University Press.

Robertson, J. M. (1943). *J. Sci. Instrum.* **20**, 175.

Spencer, M. (1959). *Acta Cryst.* **12**, 59.

Thomas, L. H. & Umeda, K. (1957). *J. Chem. Phys.* **26**, 293.

Wilson, A. J. C. (1942). *Nature, Lond.* **150**, 152.