

## The Structures of Pyrimidines and Purines.

### V. The Electron Distribution in Adenine Hydrochloride

BY W. COCHRAN

*Crystallographic Laboratory, Cavendish Laboratory, Cambridge, England*

(Received 12 May 1950)

The electron density in crystals of adenine hydrochloride has been measured with sufficient accuracy to distinguish between a number of possible tautomeric forms of the adenine molecule differing only in the positions assigned to hydrogen atoms. The result is presented as the difference between the electron density in the crystal and that appropriate to an assemblage of atoms whose electron distribution is that calculated by Hartree's self-consistent-field method. It is pointed out that this difference is smaller than might have been expected. The electron distribution in N-N and N-O hydrogen bonds supports the view that in a hydrogen bond the proton is at approximately the normal covalent-bond distance from one atom of the pair.

#### 1. Introduction and experimental

The crystal structure of adenine hydrochloride was described in Part II of this series (Broomhead, 1948). In Part IV (Broomhead, 1951) it is concluded that in this compound the adenine molecule is present as a cation, and is to be regarded as one or the other of two tautomers which differ only in the positions assigned to hydrogen atoms. Seven other possible tautomers were eliminated from a consideration of the lengths and dispositions of certain intermolecular bonds in the crystal structure. This paper describes an attempt to confirm and extend these conclusions by direct measurement of the electron density in the crystal. The main object of the work was, in fact, to test the extent to which X-ray diffraction methods can be used to study chemical bonds by measurement of their electron distribution.

Work undertaken with a similar object has been described by Brill, Grimm, Hermann & Peters (1939), and by Brill, Hermann & Peters (1942*a, b, c*). Their results and conclusions, at least in the case of oxalic acid dihydrate, have not been generally accepted (Dunitz & Robertson, 1947). While the measurements made by Brill *et al.* were no doubt accurate, more recent work has shown that better methods for the reduction of the experimental observations are available (van Reijen, 1942; Booth, 1947*a*). In particular, the avoidance of series-termination errors by the introduction of an artificial temperature factor has been shown to be inadvisable, since it may lead to further errors in atomic co-ordinates. Furthermore, by broadening the electron density near the centre of an atom, this method obscures the more interesting details of the electron distribution in the chemical bonds. The procedure adopted in this investigation has been, therefore, to evaluate not the projected electron density

$$\rho_c = \frac{1}{A} \sum_h \sum_l F_o(h0l) \exp \left[ -2\pi i \left( \frac{hx}{a} + \frac{lz}{c} \right) \right],$$

but the function

$$D = \rho_o - \rho_c = \frac{1}{A} \sum_h \sum_l \{F_o(h0l) - F_c(h0l)\} \times \exp \left[ -2\pi i \left( \frac{hx}{a} + \frac{lz}{c} \right) \right],$$

which gives the difference between the observed electron density in the unit cell of the crystal and that calculated for an assemblage of non-bonded atoms in which the electron distribution is known, at least approximately. In this way, systematic errors due to the termination of the Fourier series were virtually eliminated, and by the subtraction of that part of the electron distribution in which we are least interested, the remainder was shown up with greater clarity. The use of this function has been advocated by Booth (1948) among others, and it has been used by Finbak & Norman (1948) in an attempt to improve on the accuracy of the results obtained by Brill *et al.* (1942*c*).

Errors in the experimentally measured  $F_o$ 's naturally set a limit to the accuracy with which the electron density can be determined. An expression for the standard deviation  $\sigma(\rho)$  of the electron density projected on an area  $A$  has been given by the writer (Cochran, 1948). On reasonable assumptions, we find that for  $A = 50 \text{ \AA}^2$ ,  $\sigma(\rho) = 0.3\sigma(F)$ . Thus to measure the electron density to within  $0.1 \text{ e.\AA}^{-2}$ , the structure factors must be accurate to within 0.3 in this case. It is doubtful whether this degree of accuracy could be attained with the photographic techniques usually employed in X-ray analysis.

The results described in this paper are based on measurements made on single crystals of adenine hydrochloride using a Geiger-counter X-ray spectrometer (Cochran, 1950). The relative intensities of 250 ( $h0l$ ) planes were measured using copper and (in part) molybdenum characteristic radiations, the range covered being up to  $\sin \theta/\lambda = 0.74$ . Corrections for absorption in the crystal specimens were made by the

method due to Albrecht (1939), but no corrections for extinction were made since photographic measurements had shown no systematic discrepancies between observed and calculated  $F$ 's for those planes for which  $|F|$  was large.

Since the space group is  $P2/c$  the structure projected on (010) can be referred to a unit cell whose  $c$  axis is one-half that of the true unit cell. Indices, structure factors, etc., in this paper have, for convenience, been referred to this smaller unit cell, for which

$$a=8.77_1, \quad b=4.83_4, \quad c=9.73_0 \text{ \AA.}, \quad \beta=114^\circ 15'$$

and

$$F(000)=186.$$

New techniques used in connexion with the ( $F_o - F_c$ ) synthesis are described in §2, and are summarized in the final paragraph of that section for the benefit of the reader not interested in this purely technical aspect of the work. In §3 certain numerical results and their standard deviations are presented, while these results are discussed in §4.

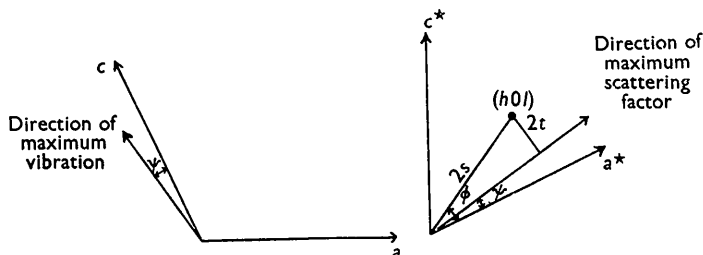


Fig. 1. The relation between the direction of maximum thermal vibration of an atom and its scattering factor.

## 2. Calculation of electron distribution

This section is devoted almost entirely to computational techniques, which it is proposed to describe in some detail, as the accuracy of the final results depends at least as much on their use as on the accuracy of the experimental measurements.

It was pointed out in §1 that the structure factors  $F_o$  were obtained on a purely relative scale. Comparison of observed and calculated  $F$ 's for a number of low-order planes gave an approximate value of the scaling factor. A graph of  $\bar{F}_o/\bar{F}_c$ , where  $\bar{F}_c$  denotes a structure factor calculated for atoms *at rest*, and the average is taken over a number of planes having approximately the same value of  $s = \sin \theta$ , could then be fitted closely by the curve  $\exp[-1.02s^2]$ . Values of  $F_c = \bar{F}_c \exp[-1.02s^2]$  were calculated using the atomic co-ordinates obtained by Broomhead (1948). The value of

$$R = \sum |F_o - F_c| \div \sum |F_o|$$

was 0.178, only slightly less than the final value of 0.20 obtained by Miss Broomhead. Inspection of the electron density calculated from the formula

$$\rho_o(x, z) = \frac{2}{A} \sum_h \sum_l F_o(h0l) \cos 2\pi \left( \frac{hx}{a} + \frac{lz}{c} \right), \quad (1)$$

using the new experimental measurements, showed that 'diffraction effects' caused by the termination of

the series at  $\sin \theta/\lambda = 0.74$  were appreciable and would be certain to obscure much of the detail. This was not unexpected, and in succeeding syntheses the function calculated was

$$D(x, z) = \rho_o - \rho_c \\ = \frac{2}{A} \sum_h \sum_l \{F_o(h0l) - F_c(h0l)\} \cos 2\pi \left( \frac{hx}{a} + \frac{lz}{c} \right). \quad (2)$$

In the first place, however, one refinement could be made, as inspection of the  $\rho_o$  map showed that the electron density in the chlorine atom was not circularly symmetrical. If this asymmetry is due to anisotropic thermal vibration of this atom, it should be possible to represent its atomic scattering factor by

$$f = f_0 \exp[-\{\alpha_C + \beta_C \sin^2(\phi - \psi_C)\} s^2], \quad (3)$$

(Hughes, 1941)

where  $\alpha_C$  and  $\beta_C$  are constants,  $\psi_C$  is the angle between the direction of maximum vibration and the  $c$  axis,

and  $(2s, \phi)$  are the polar co-ordinates of a point in the plane  $k=0$  of the reciprocal lattice (see Fig. 1).

Equation (3) may be written

$$f = f_0 \exp[-(\alpha_C s^2 + \beta_C t^2)],$$

where  $2s, 2t$  are distances in reciprocal space defined in Fig. 1. This form of (3) is convenient for numerical and graphical calculation of  $f$ . It may readily be shown that the electron density of this atom, projected on the line of maximum vibration, is the one-dimensional Fourier transform of

$$f_{\min.} = f_0 \exp[-(\alpha_C + \beta_C) s^2],$$

while that projected on the line of minimum vibration is the transform of  $f_{\max.} = f_0 \exp[-\alpha_C s^2]$  (see, for instance, Wrinch, 1946). Using this result the approximate values of  $\alpha_C$  and  $\beta_C$  were found to be 0.8 and 0.4 respectively, while  $\psi_C$  was  $55^\circ$  by direct measurement. Recalculation of structure factors on this basis reduced  $R$  from 0.178 to 0.164, not a very striking improvement.

The first  $D$  synthesis (equation (2)) was now calculated, and showed that the co-ordinates of several atoms were considerably in error. Since it will be necessary to refer to a number of such syntheses, they will be distinguished as  $D_1, D_2$ , etc. Corrections to atomic co-ordinates were made as follows. Inspection of the  $\rho_o$  map (equation (1)) showed that the electron

density within 0.4 Å. of the centre of an atom could be represented by  $\rho_o(r) = \rho_o(0) \exp[-pr^2]$ , (4)

(Costain, 1941)

with  $p = 4.9$ ,  $\rho_o(0) = 10.2$  and  $11.2 \text{ e.Å.}^{-2}$  for carbon and nitrogen atoms respectively. If the origin is taken at the point assumed to be the atomic centre in calculating structure factors, we can write

$$D(r) = \rho_o(r) - \rho_c(r) \\ = \rho_o(0) \exp[-p(r-\Delta)^2] - \rho_c(0) \exp[-p'r^2],$$

where  $\Delta$  is the required correction. If the atomic scattering factor of this atom has been correctly chosen,  $p = p'$  and  $\rho_o(0) = \rho_c(0)$ , but this cannot be assumed in advance. For small values of  $r$ ,

$$D(r) = \rho_o(0)(1 - pr^2 + 2pr\Delta) - \rho_c(0)(1 - p'r^2),$$

and therefore

$$\Delta = \frac{(dD/dr)_{r=0}}{2p\rho_o(0)} \quad (5)$$

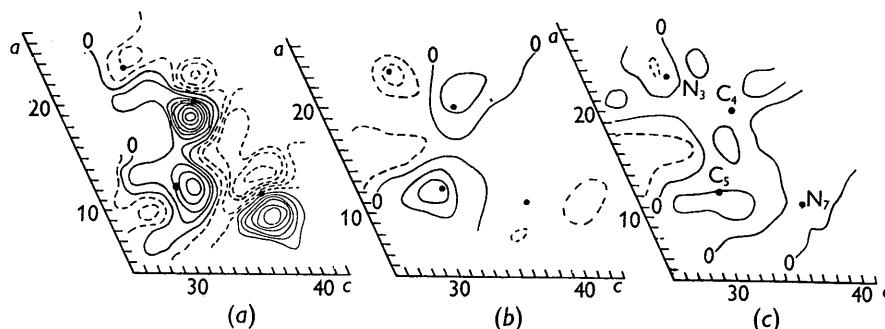


Fig. 2. (a) Variation of  $D_1$  in the neighbourhood of atoms  $N_3$ ,  $C_4$ ,  $C_5$  and  $N_7$ , showing the uniform slope of this function near atomic co-ordinates (indicated by dots). (b) Same for  $D_4$ . The co-ordinates have been corrected to remove the slope. (c) Same for  $D_8$ . The temperature-factor parameters  $\alpha_i$  have been adjusted to make  $D$  approximately zero at the centre of each atom. Contours at an interval of  $0.25 \text{ e.Å.}^{-2}$ , zero contour indicated by 0, negative contours broken. Lines parallel to  $a$  and to  $c$  are subdivided into  $\frac{1}{10}$ ths of  $a$  and  $c$  respectively.

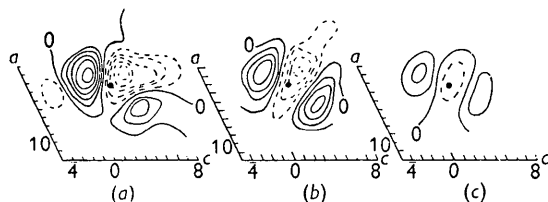


Fig. 3. Variation of (a)  $D_1$ , (b)  $D_2$  and (c)  $D_3$  in the neighbourhood of the co-ordinates of the chlorine atom, showing the gradual reduction in the value of  $D$  as the co-ordinates and temperature-factor parameters are corrected. Contours at an interval of  $\frac{1}{2} \text{ e.Å.}^{-2}$ , otherwise as in Fig. 2.

A section from the map of the function  $D_1$  is shown in Fig. 2(a). It will be noticed that the function has a considerable slope in the neighbourhood of the co-ordinates of atoms  $C_4$  and  $N_7$  in particular. The corrections for each atom were calculated from (5), the maximum value of  $\Delta$  being 0.08 Å. Fig. 3(a) shows the variation of  $D_1$  in the neighbourhood of the co-ordinates of the chlorine atom. It is clear from this figure that the  $z$  co-ordinate of this atom is smaller than that assumed, and that one or more of the constants of (3) have been wrongly estimated. Structure factors based on the

corrected co-ordinates were now calculated, and the value of  $R$  was reduced to 0.14. Calculation of  $D_2$  showed that the slope of this function near atomic centres was now small in all cases, and the form of  $D_2$  in the neighbourhood of the chlorine co-ordinates was more symmetrical (Fig. 3(b)). Calculations based on the form of  $D_2$  in this region gave new values  $\alpha_C = 0.7$ ,  $\beta_C = 0.7$  and  $\psi_C = 50^\circ$  for the temperature-factor parameters of this atom. Certain maxima of  $D_2$  could reasonably be ascribed to the presence of hydrogen atoms, and the contributions of five of these were included when the structure factors were recalculated. The atomic scattering factor of hydrogen was taken to be

$$f = f_0 \exp[-1.00s^2],$$

where  $f_0$  is the atomic scattering factor of a hydrogen atom in its ground state. These changes reduced  $R$  to 0.10.  $D_3$  was now calculated; a section from it in the

region of the chlorine co-ordinates is shown in Fig. 3(c). At this stage a change of 5% in the value of the scaling factor relating  $F_o$  to  $F_c$  was made on the basis of a comparison of observed and calculated  $F$ 's in the range  $0.3 < \sin \theta / \lambda < 0.4$ . This change was necessitated by the previous changes in the temperature-factor parameters of the chlorine atom.

The value of these parameters for the chlorine atom and for the adenine molecule as a whole were now more accurately determined as follows.

Let  $C$ ,  $M$ ,  $W$  and  $H$  be the contributions to  $F_c$  of the chlorine atom, the adenine molecule (excluding hydrogen), the oxygen atom of the water molecule, and the hydrogen atoms respectively, so that

$$F_c = C + M + W + H.$$

Assuming for the present that the temperature-factor parameters of the water molecule and the hydrogen atoms have been correctly chosen, and that only the 'acoustic' modes of vibration of the adenine molecule need be considered, we may write

$$F_o = C \cdot \exp[\Delta\gamma_C \cdot s^2] + M \exp[\Delta\gamma_M \cdot s^2] + W + H,$$

where  $\gamma = \alpha + \beta \sin^2(\phi - \psi)$ .

Hence, assuming  $\Delta\alpha_C$ , etc., small compared with unity,

$$F_o - F_c = s^2(C \cdot \Delta\gamma_C + M \cdot \Delta\gamma_M). \quad (6)$$

All structure factors with  $\sin\theta/\lambda > 0.25$  were now grouped into nine sets, each centred around one of  $\phi = 10^\circ, 30^\circ, \dots, 170^\circ$ . Each set gave 20–25 equations of the form of (6), and within each set the value of  $\phi$ , and therefore of  $\Delta\gamma_C$  and  $\Delta\gamma_M$ , was assumed constant. These equations were solved by the method of least squares, and the values obtained for  $\Delta\gamma_C$ , etc., were added as corrections to give new values of  $\alpha_C$ , etc. The corrected values for the chlorine atom were, for instance,  $\alpha_C = 0.80$ ,  $\beta_C = 0.90$ ,  $\psi_C = 43\frac{1}{2}^\circ$ . Values of  $\gamma_C$  and  $\gamma_M$  are plotted against  $\phi$  in Fig. 4. It will be observed that in both cases, particularly the former, the experimental values are closely fitted by a curve of the form  $\alpha + \beta \sin^2(\phi - \psi)$ . This may be regarded as experimental confirmation of the result given as equation (3).

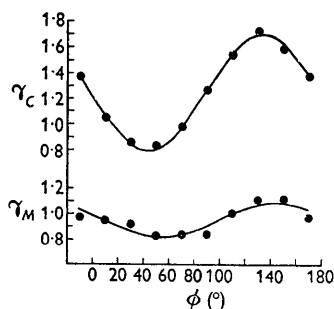


Fig. 4. Comparison of measured values of the factor  $\gamma$  (dots) with values of  $\alpha + \beta \sin^2(\phi - \psi)$  (full line). The upper graph refers to the chlorine atom, for which  $\alpha = 0.80$ ,  $\beta = 0.90$ ,  $\psi = 43\frac{1}{2}^\circ$ ; the lower to the adenine molecule for which  $\alpha = 0.80$ ,  $\beta = 0.28$ ,  $\psi = 55^\circ$ .

Further small corrections ( $\sim 0.01$  A.) to atomic co-ordinates were now made in order to eliminate the slope of  $D_3$  in the neighbourhood of atomic centres, and a fourth set of structure factors was calculated. The value of  $R$  was 0.077, and was reduced to 0.071 when the contributions from a further two hydrogen atoms were allowed for. The variation of  $D_4$  in the neighbourhood of the co-ordinates of the water molecule showed the assumption of an isotropic temperature factor for this atom to be incorrect. This, and the possibility of an error in the scaling factor  $K$  relating  $F_o$  to  $F_c$ , was allowed for by taking, as before,

$$F_c = C + M + W + H,$$

and

$$KF_o = C \cdot \exp[\Delta\gamma_C \cdot s^2] + M \cdot \exp[\Delta\gamma_M \cdot s^2] + W \exp[\Delta\gamma_W \cdot s^2] + H.$$

Therefore,

$$F_o - F_c = F_o \cdot \Delta K + s^2(C \cdot \Delta\gamma_C + M \cdot \Delta\gamma_M + W \cdot \Delta\gamma_W), \quad (7)$$

where  $\Delta K = 1 - K$ .

Equations were grouped into nine sets as before, and solved by the method of least squares for the four unknowns.  $K$  should, of course, be the same in each set;

values found ranged from  $-0.008$  to  $+0.030$ , the mean value being  $+0.01 \pm 0.005$ . Negligible values of  $\Delta\gamma_C$  were found, while values of  $\Delta\gamma_M$  indicated a change of 0.05 in the value of  $\alpha_M$ . Values of  $\Delta\gamma_W$  led to  $\alpha_W = 0.8$ ,  $\beta_W = 0.5$ ,  $\psi_W = 33^\circ$ . In the  $D_4$  map the variation of this function in the immediate neighbourhood of the co-ordinates of the amino nitrogen ( $N_{10}$ ) had been closely similar to its variation in the neighbourhood of the water molecule. The temperature-factor parameters of  $N_{10}$  were therefore adjusted accordingly.

The possibility that all atoms of the adenine molecule did not have the same temperature factor was next considered. The calculations referred to above showed that the thermal vibration of the molecule as a whole was anisotropic, and in order to avoid an excessive amount of calculation it was assumed that  $\beta_M$  was the same for all atoms, except  $N_{10}$ , for which this assumption was clearly incorrect. Two ways of choosing the constant  $\alpha_r$  for each atom of the molecule suggest themselves:

(i) We may choose the value of  $\alpha_r$  to make the calculated and observed electron densities at the centre of the  $r$ th atom agree exactly. This is perhaps the better method to adopt, as  $\rho_o$  at this point should be nearly independent of any redistribution of electrons that has taken place on bonding. A positive value of  $\rho_o - \rho_c$  at an atomic centre would correspond to too great a value of  $\alpha_r$ . A quantitative relation may be established as follows:

$$\text{Writing} \quad 2\pi(hx/a + lz/c) = \Theta,$$

$$F_c = C + W + H + 2 \sum_r f_r \cos \Theta_r,$$

where  $f_r = f_o \exp[-(\alpha_M + \beta_M \sin^2(\phi - \psi_M))s^2]$ ,

and  $f_o$  denotes, as before, the scattering factor of a carbon or nitrogen atom at rest.

Now assume

$$F_o = C + W + H + 2 \sum_r \{f_r \exp[\Delta\alpha_r \cdot s^2]\} \cos \Theta_r.$$

$$\text{Then } \rho_o - \rho_c = \frac{2}{A} \sum_h \sum_l 2 \left\{ \sum_r f_r \cdot \Delta\alpha_r \cdot s^2 \cos \Theta_r \right\} \cos \Theta, \quad (8)$$

and, since  $\sum_h \sum_l \cos \Theta_r \cos \Theta_{r'} = 0$  unless  $r = r'$ ,

$$D(x_r, z_r) \doteq \frac{2}{A} \Delta\alpha_r \cdot \sum_h \sum_l f_r \cdot s^2. \quad (9)$$

Equation (8) was used to evaluate  $\rho_o - \rho_c$  for an atom (average of carbon and nitrogen) with  $\Delta\alpha_r = 0.1$ . The result is shown in Fig. 5(a), from which it will be observed that this value of  $\Delta\alpha_r$  leads to a value  $\rho_o - \rho_c = 0.35 \text{ e.A.}^{-2}$  at the centre of the atom, but values at distances greater than about 1 A. from the centre are very small. Values of  $D_5$  at the centres of atoms were used to find values of  $\Delta\alpha_r$  from (9). The results are given in Table 1.

(ii) Alternatively, we may choose the value of  $\alpha_r$  for each atom to give the best possible agreement between

$F_o$  and  $F_c$ . This will not necessarily lead to the same values for the  $\alpha_r$ 's as the procedure outlined in (i) above. We have, as before,

$$F_o - F_c = 2 \sum_r f_r \cdot \Delta\alpha_r \cdot s^2 \cos \Theta_r,$$

the sum being taken over the atoms of the adenine molecule. This set of 250 equations was solved for the ten  $\Delta\alpha_r$ 's by a modification of the least-squares method. The results are shown in Table 1, and the agreement with values obtained by the first method is satisfactory.

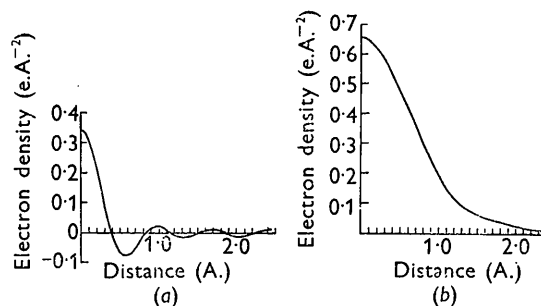


Fig. 5. (a) Difference between the electron densities of two atoms whose temperature-factor parameters,  $\alpha$  and  $\alpha + \Delta\alpha$ , differ by 0.1. (b) Difference between the electron densities of a chlorine ion and a chlorine atom, calculated from structure factors given by James & Brindley.

Table 1. Comparison of corrections to temperature-factor parameters obtained by two methods

Atom I	Value of $\Delta\alpha$ from $D_5$	Value of $\Delta\alpha$ by least squares
N <sub>1</sub>	-0.095	-0.118
N <sub>3</sub>	-0.128	-0.138
N <sub>7</sub>	+0.007	+0.032
N <sub>9</sub>	-0.080	-0.096
N <sub>10</sub>	-0.108	-0.150
C <sub>2</sub>	-0.108	-0.075
C <sub>4</sub>	+0.095	+0.106
C <sub>5</sub>	+0.162	+0.170
C <sub>6</sub>	+0.013	+0.075
C <sub>8</sub>	-0.034	-0.032

The values of the temperature-factor parameters used for the final calculation of structure factors are shown in Table 2. Some of them were rounded off from the exact values derived from Table 1 in order to simplify subsequent calculation. The effect on the  $D$

Table 2. Temperature-factor parameters of all atoms

Atom	$\alpha$	$\beta$	$\psi$	$\gamma = \alpha + \frac{1}{2}\beta$
Cl	0.8	0.9	43½°	1.25
W	0.8	0.5	33°	1.05
N <sub>10</sub>	0.7	0.8	38°	1.10
N <sub>1</sub>	0.9	0.28	55°	1.04
N <sub>3</sub>	0.9	0.28	55°	1.04
N <sub>7</sub>	0.8	0.28	55°	0.94
N <sub>9</sub>	0.9	0.28	55°	1.04
C <sub>2</sub>	0.9	0.28	55°	1.04
C <sub>4</sub>	0.7	0.28	55°	0.84
C <sub>5</sub>	0.7	0.28	55°	0.84
C <sub>6</sub>	0.8	0.28	55°	0.94
C <sub>8</sub>	0.8	0.28	55°	0.94

map of the correction of the temperature-factor parameters can be seen by comparing a section from  $D_6$  (Fig. 2(c)) with a corresponding section from  $D_4$  (Fig. 2(b)).

Values of  $F_c$  calculated on the basis of these constants and the atomic co-ordinates set out in Table 3 are compared with values of  $F_o$  in Table 4. The value of  $R$  is 0.061, and the mean value of  $|F_o - F_c|$  is 0.44. These values of  $F_c$  were used to calculate  $D_6$ . Structure

Table 3. Atomic co-ordinates

Atom	$x$	$z$	$y$
Cl	0.2812	0.0482	0.178 obs.
O <sub>W</sub>	0	0	0.745 obs.
N <sub>1</sub>	0.1762	0.3043	0.5998 calc.
N <sub>3</sub>	0.3960	0.5290	0.7553 calc.
N <sub>7</sub>	0.1883	0.6326	0.2090 calc.
N <sub>9</sub>	0.4060	0.7458	0.5013 calc.
N <sub>10</sub>	-0.0093	0.2890	0.2481 calc.
C <sub>2</sub>	0.3098	0.3838	0.7665 calc.
C <sub>4</sub>	0.3395	0.5958	0.5579 calc.
C <sub>5</sub>	0.2078	0.5283	0.3798 calc.
C <sub>6</sub>	0.1180	0.3717	0.3982 calc.
C <sub>8</sub>	0.3136	0.7618	0.2961 calc.
H <sub>0</sub>	0.087	0.025	—
H <sub>1</sub>	0.120	0.205	0.613 calc.
H <sub>2</sub>	0.335	0.318	0.900 calc.
H <sub>10</sub>	-0.048	0.200	0.283 calc.
H <sub>10</sub>	-0.058	0.310	0.124 calc.
H <sub>8</sub>	0.338	0.850	0.234 calc.
H <sub>9</sub>	0.513	0.825	0.615 calc.

factors  $F'_c$  which did not include the contributions of the hydrogen atoms, were also calculated.  $F_c$  differs from  $F'_c$  only in the range  $\sin \theta/\lambda < 0.45$ . Within this range,  $R = 0.046$  and  $|F_o - F'_c| = 0.50$ , while  $R' = 0.08$  and  $|F_o - F'_c| = 0.88$ .  $D_7$  was calculated according to equation (2) with values of  $F_o - F'_c$  as coefficients. The result is shown in Fig. 6.

It may be as well to summarize at this point the procedure outlined in this section, and the assumptions on which the electron distribution shown in Fig. 6 was obtained. It was assumed that the observed structure amplitudes could best be represented by

$$KF_o = 2 \sum_r f_{0r} \exp [ -(\alpha_r + \beta_r \sin^2(\phi - \psi_r)) s^2 ] \times \cos 2\pi \left( \frac{hx_r}{a} + \frac{lz_r}{c} \right).$$

The constant  $K$  was then adjusted to unity, and the best values of the atomic co-ordinates ( $x_r, z_r$ ) and the temperature-factor parameters  $\alpha_r, \beta_r$  and  $\psi_r$  was found by a method of successive approximations.  $f_0$  is the atomic scattering factor of an atom whose electron distribution is that calculated by the method of the self-consistent field (Hartree, 1928). The electron distribution shown in Fig. 6 is therefore the difference between that existing in the crystal and that obtained by superposing 'Hartree' chlorine, oxygen, nitrogen and carbon atoms, with appropriate amplitudes of thermal vibration, at the appropriate points in the unit cell.

Table 4. Comparison of observed and calculated structure factors

\* Terms marked with an asterisk were omitted from the final synthesis because of uncertain sign.

<i>h</i>	<i>l</i>	<i>F<sub>o</sub></i>	<i>F<sub>c</sub></i>	<i>h</i>	<i>l</i>	<i>F<sub>o</sub></i>	<i>F<sub>c</sub></i>	<i>h</i>	<i>l</i>	<i>F<sub>o</sub></i>	<i>F<sub>c</sub></i>
0	0		186	2		4.9	4.0	8		3.4	- 3.0
1		3.2	3.7	3		10.6	10.6		7	5.9	5.4
2		33.8	- 32.9	4		7.3	- 7.0	II		1.5	- 1.5
3		25.4	24.1	5		12.2	- 12.3	IO		6.3	- 6.2
4		4.9	5.3	6		18.2	19.2	9		1.5	1.5
5		4.4	4.2	7		12.2	11.4	8		8.1	9.1
6		7.0	7.2	8		7.8	- 7.2	7		3.0	2.9
7		11.9	12.3	9		0	0.2	6		0.3	- 0.2
8		2.9	- 3.0	II	4	1.8	1.7	5		9.7	9.7
9		2.5	- 2.0	II		0	0.4	4		7.3	7.1
10		9.2	9.0	IO		0	0.2	3		3.2	- 3.1
11		2.2	1.9	9		6.6	- 6.5	2		7.8	8.0
II	1	3.7	3.7	8		8.0	7.8	I		17.2	17.5
IO		3.7	- 4.4	7		7.5	7.8	0		6.1	- 6.0
9		4.4	- 4.1	6		11.8	- 11.8	1		7.4	- 6.2
8		4.2	4.6	5		13.2	12.2	2		17.0	17.0
7		10.5	10.1	4		13.9	13.8	3		8.3	8.6
6		17.8	- 16.7	3		1.0	1.2	4		3.7	- 3.9
5		19.8	- 18.2	2		0.8*	0	5		5.1	5.2
4		26.8	26.0	I		28.3	28.0	6		7.3	7.6
3		6.6	6.2	0		16.0	16.4	7		2.4	- 2.2
2		21.5	21.5	1		33.0	- 32.9	II	8	1.2	- 0.7
I		0.7	0.9	2		13.0	12.7	II		0	- 0.4
0		19.8	- 20.8	3		21.2	21.1	IO		0	0.7
1		5.9	5.4	4		2.2	2.7	9		1.4*	0
2		20.4	- 19.0	5		5.1	5.9	8		3.6	3.1
3		25.8	25.4	6		1.7	1.4	7		2.5	- 2.5
4		18.3	17.3	7		5.1	5.0	6		0	- 0.2
5		14.6	- 14.4	8		3.0	- 2.1	5		7.3	7.6
6		3.5	3.5	9		0	0	4		10.2	- 9.8
7		13.2	13.0	II	5	3.7	3.8	3		8.0	- 7.7
8		3.5	1.8	II		2.2	- 2.2	2		3.4	3.1
9		7.3	- 6.7	IO		8.1	- 8.7	I		10.4	11.2
10		2.2	1.7	9		5.9	6.3	0		2.2	3.0
II	2	1.2	- 1.0	8		3.4	3.0	1		10.7	- 11.1
II		2.6	2.2	7		2.7	2.4	2		5.2	5.6
IO		2.2	- 2.1	6		4.6	- 4.4	3		5.1	5.4
9		3.4	- 3.4	5		8.1	- 7.5	4		5.6	- 5.0
8		4.7	5.2	4		5.8	6.8	5		0	0.4
7		2.7	2.7	3		12.9	- 14.0	6		2.9	2.8
6		1.7	- 2.9	2		2.2	1.8	II	9	2.4	2.5
5		0	- 1.0	I		1.7	1.0	II		2.5	- 2.7
4		9.5	9.5	0		2.4	1.9	IO		0	- 0.1
3		2.9	- 2.1	1		4.9	- 5.6	9		7.3	7.8
2		31.6	- 33.4	2		7.1	- 7.5	8		1.4	- 1.7
I		8.5	8.3	3		20.2	21.0	7		6.6	- 6.6
0		10.0	9.4	4		8.1	- 8.0	6		8.8	- 8.3
1		20.7	- 21.0	5		14.4	- 14.5	5		8.0	7.2
2		8.5	- 9.0	6		7.6	7.9	4		9.7	9.6
3		16.1	16.4	7		0	- 0.9	3		7.0	- 7.0
4		9.5	9.2	8		4.2	- 4.6	2		7.6	7.5
5		12.0	- 11.5	II	6	2.0	- 1.2	I		3.2	- 2.9
6		9.3	- 9.2	II		7.8	8.1	0		11.6	- 12.4
7		1.2	0.9	IO		1.8	1.6	1		4.2	4.2
8		1.0*	- 0.1	9		1.2*	- 0.4	2		6.3	6.8
9		2.2	- 2.3	8		8.8	9.0	3		2.7	- 2.7
10		3.5	3.2	7		12.0	- 11.5	4		10.2	- 10.0
II	3	0	0	6		8.6	- 8.3	5		1.9	2.6
II		8.1	8.3	5		19.3	19.3	II	10	1.0	1.8
IO		1.7	2.4	4		16.0	16.1	IO		5.1	- 5.1
9		3.2	- 3.4	3		15.0	- 13.8	9		2.7	3.2
8		9.6	9.4	2		4.6	- 4.1	8		2.7	2.7
7		6.8	6.7	I		26.6	25.5	7		3.4	- 4.2
6		8.5	- 7.3	0		3.4	- 3.5	6		15.8	16.0
5		6.3	6.7	1		3.7	- 4.0	5		6.8	7.2
4		33.2	33.7	2		11.8	11.7	4		9.1	- 10.3
3		2.6	2.2	3		4.1	- 4.1	3		1.3	- 1.3
2		25.7	- 26.8	4		7.0	- 7.1	2		8.5	8.3
I		35.8	40.0	5		0	- 0.7	I		7.3	7.1
0		33.6	36.0	6		7.6	6.6	0		4.9	- 4.6
1		5.7	5.0	7		3.0	3.2				

Table 4 (cont.)

$h$	$l$	$F_o$	$F_c$	$h$	$l$	$F_o$	$F_c$	$h$	$l$	$F_o$	$F_c$
1		2.4	2.4	$\bar{1}$		4.9	4.6	$\bar{2}$		3.2	3.2
2		5.4	5.5	0		2.7	2.9	$\bar{1}$		4.0	3.2
3		0	0.1	1		3.2	3.6	0		7.3	6.9
$\bar{11}$	11	0	1.3	2		6.4	6.1	1		2.2	1.6
$\bar{10}$		3.0	2.9	3		5.6	4.9	$\bar{8}$	13	1.9	2.5
$\bar{9}$		1.5	1.1	$\bar{10}$	12	0	0.4	$\bar{7}$		0	1.2
$\bar{8}$		3.2	3.3	9		2.4	2.2	$\bar{6}$		5.6	5.3
$\bar{7}$		0	0.9	$\bar{8}$		3.2	2.9	$\bar{5}$		2.7	2.5
$\bar{6}$		4.2	4.4	$\bar{7}$		4.4	3.5	$\bar{4}$		3.7	4.0
$\bar{5}$		5.1	5.6	$\bar{6}$		2.0	1.9	$\bar{3}$		0	1.1
$\bar{4}$		1.0	1.0	$\bar{5}$		0	0.2	$\bar{2}$		3.4	3.7
$\bar{3}$		3.7	4.1	$\bar{4}$		2.5	2.6	$\bar{1}$		4.2	4.3
$\bar{2}$		7.0	6.9	$\bar{3}$		4.1	3.6				

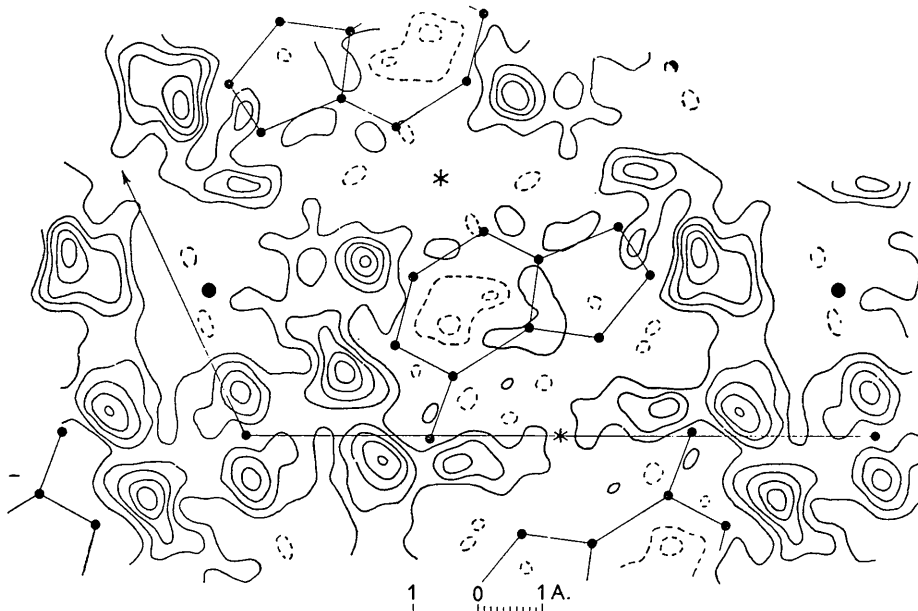


Fig. 6. Difference between the electron density projected on (010) and that calculated for isolated Cl, O, N and C atoms whose centres are indicated by dots. Contours at 0.2, 0.4, ... 1.0 e.A.<sup>-2</sup>, negative contours broken.

### 3. Numerical results

Bond lengths and bond angles calculated from the co-ordinates given in Table 3 are shown in Fig. 7. Since the  $y$  co-ordinates have not been determined in this investigation, it has been assumed that the molecule is planar. Support for this assumption comes from accurate three-dimensional measurements on naphthalene (Abrahams, Robertson & White, 1949), anthracene (Mathieson, Robertson & Sinclair, 1950) and a substituted pyrimidine (Clews & Cochran, 1949). If the molecule is planar, we have  $y = Ax + Bz + C$  ( $x, y, z$  fractional co-ordinates, diad axis taken as origin in  $xz$  plane). The values of  $A, B$  and  $C$  in best agreement with the  $y$  co-ordinates found by Broomhead (1948) were  $A = 2.0081, B = -1.2742, C = 0.6301$ . If the neighbouring molecules related by the operation of the centres of symmetry at  $(0, 0, \frac{1}{2})$  and  $(\frac{1}{2}, 0, \frac{1}{2})$  are accurately coplanar, as a model of the crystal structure and a con-

sideration of intermolecular forces suggest, we should find  $A = 2.00, B = -2C$ . The value of  $C$  fulfilling these conditions and in best agreement with previous measurements was 0.6320. Bond lengths were calculated using both sets of constants, but in no case did they differ by as much as 0.005 Å. The standard deviations of bond lengths and other measurements are derived later in this section.

The number of electrons associated with each atom was evaluated as follows. If  $n_o$  and  $n_c$  are the numbers of electrons in an area  $S$ , corresponding to electron densities  $\rho_o$  and  $\rho_c$ ,

$$n_o - n_c = \int_S (\rho_o - \rho_c) dS = \int_S D dS;$$

or approximately  $n_o = n_c + \delta S \cdot \Sigma D$ , where the sum is taken over the points inside  $S$  at which  $D$  was evaluated. Each Fourier synthesis was evaluated at intervals of  $a/60$  and  $c/60$ , so that  $\delta S = A/3600$ . The choice of area

associated with each atom, or group of atoms, was necessarily somewhat arbitrary and is shown in Fig. 8. Values of  $n_o$  derived in this way from  $D_6$  (hydrogen atoms subtracted) and from  $D_7$  (hydrogen atoms not subtracted) are given in Table 5.

oxygen and hydrogen atoms, and of  $f_0 \exp[-1.25s^2]$  for a chlorine atom (Fig. 9). If the electron spread in a bonded hydrogen atom is even qualitatively similar to that in an isolated one, a considerable number of electrons will lie outside the boundaries of the areas

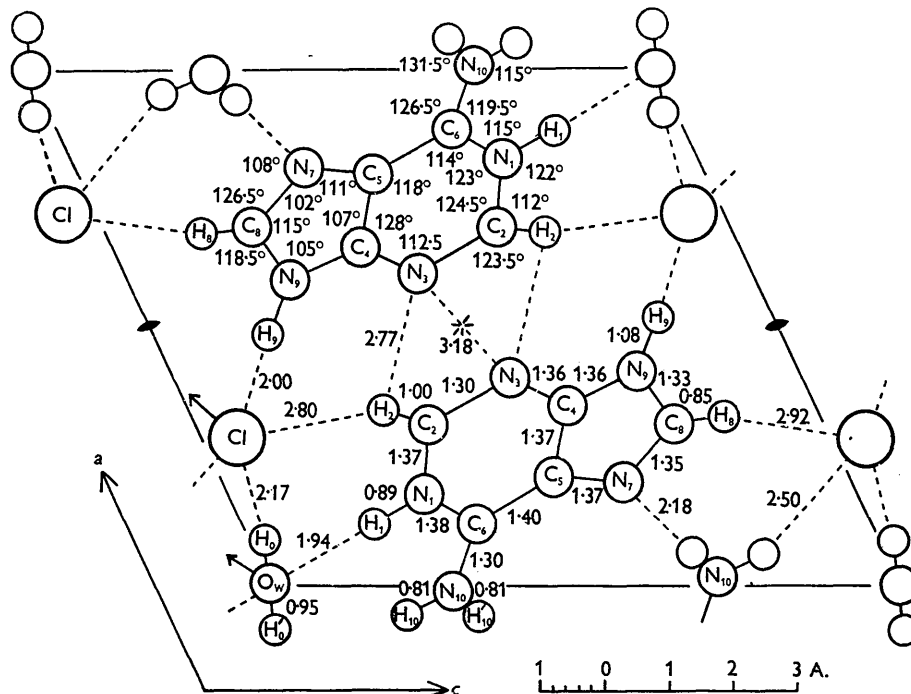


Fig. 7. Bond lengths and bond angles in adenine hydrochloride. Arrows indicate the directions of maximum thermal vibration of the chlorine atom and the water molecule.

Table 5. Numbers of electrons in certain atoms or groups of atoms

Atom group	$n_o$	$n_o$ from $D_6$	$n_o$ from $D_7$	Average $n_o$
(NH) <sub>1</sub>	8	7.96	7.86	7.91
(CH) <sub>2</sub>	7	6.92	6.77	6.85
N <sub>3</sub>	7	7.00	6.97	6.98
C <sub>4</sub>	6	6.10	6.10	6.10
C <sub>5</sub>	6	6.22	6.20	6.21
C <sub>6</sub>	6	5.87	5.93	5.90
N <sub>7</sub>	7	6.89	6.98	6.94
(CH) <sub>2</sub>	7	7.23	7.28	7.25
(NH) <sub>2</sub>	8	7.80	7.56	7.68
(NH <sub>2</sub> ) <sub>10</sub>	9	8.78	8.35	8.57
$\frac{1}{2}$ H <sub>2</sub> O	$\frac{1}{2} \times 10$	$\frac{1}{2} \times 9.96$	$\frac{1}{2} \times 9.80$	$\frac{1}{2} \times 9.88$
Cl	17	16.85	17.26	17.05
Total	93	92.6	92.16	92.38

The standard deviation of the number of electrons in one of the areas considered is shown later to be less than 0.1 e. Values of  $n_o$  derived from  $D_6$  and from  $D_7$  often differ by considerably more than this, particularly when a hydrogen atom is involved. The explanation of this is to be sought in the spreading of the electron density outside the area in which an electron count was made. The electron distribution in isolated atoms has been calculated by forming the two-dimensional Fourier transforms of  $f_0 \exp[-1.00s^2]$  for carbon, nitrogen,

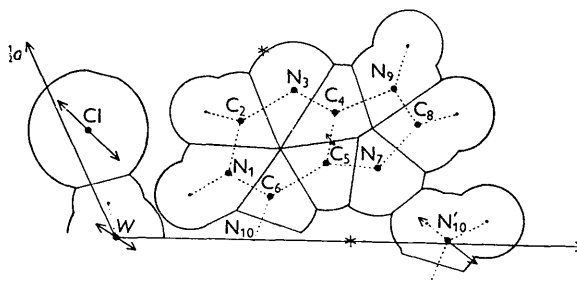


Fig. 8. Areas associated with each atom or group of atoms in which electron counts were made. The arrows show the direction of maximum thermal vibration of each atom or group of atoms.

shown in Fig. 8, and values of  $n_o$  from  $D_7$  will be too low. Values taken from  $D_6$  tend to be too high for a similar reason. This, and not the random experimental error, constitutes the main error in the measurement of  $n_o$ . The average of the values from  $D_6$  and  $D_7$  will be taken as the best estimate in each case, and should not be in error by more than 0.2 e.

Before discussing the results, an attempt will be made to estimate their standard deviations, so that conclusions are not drawn from results which are not significantly outside the range of experimental error.



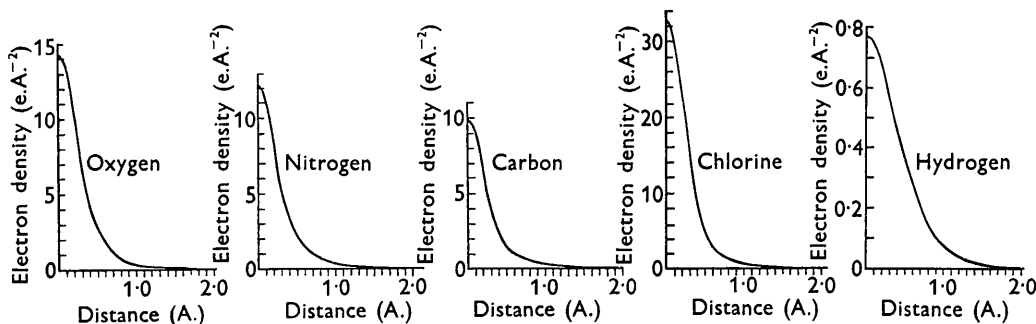


Fig. 9. The electron distribution in isolated atoms.

(a) *Standard deviation of the electron density*

$$\text{From } \rho_o = \frac{2}{A} \sum_h \sum_l F_o \cos \Theta,$$

we obtain immediately

$$\Delta \rho_o = \frac{2}{A} \sum_h \sum_l \Delta F_o \cos \Theta,$$

and therefore

$$\sigma(\rho_o) = \{(\overline{\Delta \rho_o^2})\}^{\frac{1}{2}} = \frac{2}{A} \sigma(F_o) (\frac{1}{2}N)^{\frac{1}{2}}, \quad (10)$$

where  $\sigma(F_o) = \{(\overline{\Delta F_o^2})\}^{\frac{1}{2}}$  and  $N$  is the number of terms in the series. The quantity  $\sigma(\rho_o - \rho_c)$  will be equal to  $\sigma(\rho_o)$ , provided rounding-off errors in the calculation of the  $F_c$ 's are small. This was the case in the present work. A lower limit to the value of  $\sigma(F_o)$  can be got by comparing independent observations. This gave  $\sigma(F_o) > 0.3$ . An upper limit is got by taking

$$\sigma(F_o) = 1.25 |F_o - F_c|,$$

which gives  $\sigma(F_o) < 0.55$ . The former value is an underestimate if systematic errors were present in the observations, the latter must be an over-estimate, since the atomic scattering factors used in calculating  $F_c$ 's cannot be exactly correct. Accordingly we take  $\sigma(F_o) = 0.4$ , and hence  $\sigma(\rho_o) = \sigma(\rho_o - \rho_c) = 0.1 \text{ e.A.}^{-2}$ , since  $A = 77.8 \text{ A.}^2$ ,  $N = 250$ .

(b) *Standard deviation of bond length*

We make use of a result given by Booth (1947b)

$$\sigma(x_r) = \frac{2}{N_r} \left(\frac{\pi}{\lambda p}\right)^2 \left(\frac{\pi}{A}\right)^{\frac{1}{2}} \sigma(F_o).$$

$\sigma(x_r)$  is the standard deviation of the  $x$  co-ordinate of the  $r$ th atom,  $N_r$  its atomic number and  $p = 4.9$  in this case. Hence for a carbon atom,  $\sigma(x_r) = 0.0056 \text{ A}$ . This does not take into account errors caused by possibly incomplete correction for the termination of the series, which is, of course, automatically accomplished, as far as is possible, by the technique of the  $(F_o - F_c)$  synthesis. A result given by Cruickshank (1949) allows one to estimate the standard deviation resulting from all sources of error. It may readily be shown that Cruickshank's result is equivalent to

$$\sigma(x_r) = \left\{ \frac{1}{A} \int_A \left( \frac{\partial}{\partial x} (\rho_o - \rho_c) \right)^2 dA \right\}^{\frac{1}{2}} \div \left( \frac{\partial^2 \rho_o}{\partial x^2} \right)_{x=x_r, z=z_r}$$

The average value of  $\left( \frac{\partial}{\partial x} (\rho_o - \rho_c) \right)^2$  was evaluated over 86 points of the function  $D_r$ , at which one might reasonably expect the true value of  $\rho_o$  and therefore of  $\frac{\partial}{\partial x} (\rho_o - \rho_c)$  to be zero. This gave  $\sigma(x_r) = 0.007 \text{ A}$ . We may therefore conclude that the standard deviation of a CC or CN bond length does not exceed  $0.01 \text{ A}$ .

(c) *Standard deviation of number of electrons in a particular area*

$$\text{We have } n_o = \int \rho_o dS,$$

where  $dS$  is an element of area  $= A dx dz / ac$ . If the area is bounded by the lines  $x = x_1, x_2$  and  $z = z_1, z_2$ , then

$$n_o = \frac{2}{\pi^2} \sum_h \sum_l \frac{F_o}{hl} \sin \pi h \left( \frac{x_2 - x_1}{a} \right) \sin \pi l \left( \frac{z_2 - z_1}{c} \right) \times \cos 2\pi \left( \frac{h(x_1 + x_2)}{2a} + \frac{l(z_1 + z_2)}{2c} \right).$$

Hence we have, fairly directly,

$$\sigma(n_o) = (\sqrt{2}) \sigma(F_o) \frac{x_2 - x_1}{a} \frac{z_2 - z_1}{c} \left\{ \sum_h \sum_l \left( \frac{\sin X}{X} \right)^2 \left( \frac{\sin Z}{Z} \right)^2 \right\}^{\frac{1}{2}},$$

where  $X = \pi h(x_2 - x_1)/a$ , etc.

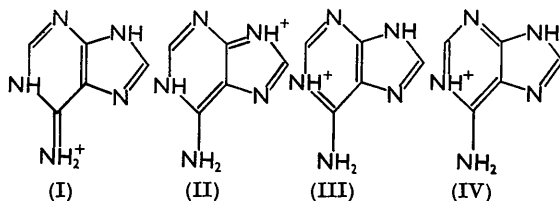
The summation was evaluated numerically for the case  $(x_2 - x_1)/a = (z_2 - z_1)/c = 1/6$ , which corresponds to an area of  $2.16 \text{ A.}^2$ , about equal to the average area over which electron counts were made, and the value of  $\sum \sum \left( \frac{\sin X}{X} \right)^2 \left( \frac{\sin Z}{Z} \right)^2$  was found to be 14.4. Hence

$\sigma(n_o) = 0.06 \text{ e}$ . This value is small compared with the uncertainty introduced by the spreading of the electrons for a particular atom outside the area which can be associated with this atom, as has already been pointed out.

## 4. Discussion

The bond lengths and bond angles shown in Fig. 7 differ considerably from those given previously (Broomhead, 1948), which were derived from photographic measurements. The mean deviations between the two sets of results are  $0.045 \text{ A}$ . for bond lengths and  $3\frac{1}{2}^\circ$  for bond angles. Bond lengths reported here do not appear to support suggestions made by Taylor (1949) regarding

the molecular structure of adenine. Later discussion suggests that if we attempt to describe the molecular structure of adenine in terms of resonance between valence bond structures, (I) and (II) must make important contributions. In fact, by suitably weighting



the contributions from the four structures shown above, one already obtains bond lengths in the pyrimidine ring in good agreement with those observed, while the predicted electron distribution is in qualitative agreement with that observed. These structures consistently make  $N_9C_8$  and  $N_7C_5$  single bonds, and  $N_7C_8$  a double bond, however. In fact all bonds of the iminazole ring are found to have 20–50 % double-bond characters, and it is not possible to account for this without postulating contributions from structures which do not appear plausible, and which find no support from the experimentally determined electron distribution. It must be concluded that this type of approach to the problem is too naïve to predict the bond lengths in a relatively complex molecule where the interaction with the chlorine atom should also be taken into account.

The electron distribution confirms and extends the conclusions reached in Part IV. One-half of the unit cell we are considering contains one molecule of adenine ( $C_5H_5N_6$ ), one of HCl and half of the water molecule which is in a special position, that is, 93 electrons in all. Of the seven hydrogen atoms, it was known in advance that one must be bound to each of  $C_2$  and  $C_8$ . Electron-density maxima occur at distances of 1.00 and 0.85 Å. from these atoms, and in directions fairly symmetrically disposed in relation to neighbouring covalent bonds (Fig. 7). We have assumed, of course, that in space these maxima lie in the molecular plane. The CH distances given above have a standard deviation which must be about 0.1 Å., and there is no reason to suppose that the proton coincides exactly with the point of maximum electron density to which measurements were made.

In the pyrimidine ring, one can say with confidence that there is a hydrogen atom covalently bound to  $N_1$ , but not to  $N_3$ . This could not have been predicted in advance, although it was strongly suggested by packing considerations. The height of the electron-density maximum adjacent to  $N_1$  is  $1.07 e.A.^{-2}$ , and it occurs at a distance of 0.89 Å. from  $N_1$ , almost exactly on the line joining  $N_1$  to the centre of the water molecule. The distance of the hydrogen atom (always with the reservation that this refers to the point of maximum electron density) from the centre of the water molecule ( $O_W$ ) is 1.94 Å., while  $N_1O_W = 2.81$  Å. The lengths and

dispositions of these bonds suggest that this is a typical N—H—O hydrogen bond, and the distance  $H_1O_W = 1.94$  Å. is significantly less than the sum of the van der Waals radii of hydrogen and oxygen, which is  $1.2 + 1.4 = 2.6$  Å.

From the position of the electron-density maximum, it would be unreasonable to assume that the proton is situated elsewhere than at a point distant  $1.0 \pm 0.2$  Å. from the centre of  $N_1$ . This agrees with infra-red spectroscopic evidence on the position of the proton in at least the majority of hydrogen bonds (see, for instance, Pauling, 1945). The electron density in the region between  $H_1$  and  $O_W$  is never less than  $0.2 e.A.^{-2}$ , but this is not sufficiently great compared with the standard deviation of  $0.1 e.A.^{-2}$  for any conclusions to be drawn.

Turning to the distribution in the neighbourhood of the extra-ring nitrogen ( $N_{10}$ ), it is reasonable to assume from the presence of two electron-density maxima, both at distances of 0.81 Å. from the centre of this atom, that the adenine molecule is in the ( $-NH_2$ ), and not the ( $=NH$ ) form. The heights of these maxima are 0.72 and  $1.00 e.A.^{-2}$ , the smaller ( $H'_{10}$ ) occurring along the line joining  $N_{10}$  to  $N'_7$  ( $N_7$  of an adjacent molecule), and the larger ( $H_{10}$ ) approximately on the line  $N_{10}Cl$ . The distances  $H_{10}Cl$  and  $H'_{10}N'_7$  are 2.50 and 2.18 Å. respectively, again significantly less than the distances obtained by adding corresponding van der Waals radii, which are 3.0 and 2.7 Å. respectively. The electron distribution in  $H'_{10}$  is the more diffuse of the two, and again there are indications of a 'bridge' extending towards (or from)  $N'_7$ . The number of electrons associated with the amino group is definitely less than 9, the deficiency being about 0.4 e. (Table 5). This is in qualitative agreement with the pronounced shortening of the bond  $C_6N_{10}$  from the single-bond value of 1.47 Å., since both effects can be explained in terms of contributions to the molecular state from resonance structures such as (I) above. Indeed, the quantitative agreement is good. A deficiency of  $0.4 \pm 0.1 e.$  corresponds to a double-bond character of  $40 \pm 10$  %, and Pauling's curve (1945) relating bond length to double-bond character shows this to correspond to a bond length of  $1.33 \pm 0.02$  Å., while the measured value is  $1.30 \pm 0.01$  Å.

We have so far located five of the seven hydrogen atoms. A sixth, forming a part of the water molecule and related to the other hydrogen atom of this molecule by the operation of a diad axis, is also clearly resolved in Fig. 6. The maximum electron density is  $0.87 e.A.^{-2}$ , and if we assume that the hydrogen atom is directed towards a neighbouring chlorine atom, in space as well as in projection, the distance  $O_W H_6$  is 0.95 Å., and the angle HOH is  $96^\circ$ . The distance  $H_6Cl$  is then 2.17 Å., much shorter than the 3.0 Å. obtained by adding the van der Waals radii. In contrast, the distances  $H_2Cl$  and  $H_3Cl$  are 2.80 and 2.92 Å. respectively. These may be regarded as normal van der Waals contacts, and the shortening of 0.5 Å. observed in the other cases is to be regarded as characteristic of hydrogen bonding.

The seventh hydrogen atom appears to be covalently bound to  $N_9$ , although the evidence is not quite so conclusive in this case. A peak of height  $0.7 \text{ e.}\text{\AA}^{-2}$  occurs  $1.08 \text{ \AA}$ . from  $N_9$  along the line joining this atom to a chlorine atom, and although it contains considerably less than one electron, it is the only large peak in Fig. 6 still not accounted for. The number of electrons in the group  $(NH)_9$  is  $7.7 \pm 0.2$  (Table 5).

We have now associated an additional proton with the adenine molecule, but less than one extra electron since both  $(NH_2)_{10}$  and  $(NH)_9$  are electron-deficient. This is in qualitative agreement with the assumption, implicit in the above discussion of bond lengths, that adenine hydrochloride is to be regarded as  $(AH)^+Cl^-$ . One may therefore ask whether there is any evidence from the electron distribution that the chlorine atom is ionized. At this point the question of the accuracy of the atomic scattering factors used in calculating  $\rho_c$  must be raised. The atomic scattering factor for oxygen comes directly from the electron distribution calculated by the self-consistent-field method, while those for carbon and nitrogen were obtained by an interpolation method based on self-consistent-field calculations for other atoms of low atomic number (James & Brindley, 1931). The atomic scattering factor of Cl is based on the calculated distribution in  $Cl^-$ , and James & Brindley state that to obtain  $f(Cl)$  they subtracted one-sixth of the contribution to  $f(Cl^-)$  of the six ( $3p$ ) electrons. The difference between the projected electron densities in  $Cl^-$  and in Cl should be obtained on calculating the two-dimensional Fourier transform of  $f(Cl^-) - f(Cl)$ . When this is done, using the values published by James & Brindley, the distribution shown in Fig. 5 (*b*) is obtained. This is obviously incorrect, since it predicts a maximum difference of electron density at the nucleus, while we should in fact expect the difference there to be almost zero and the maximum difference to occur at a distance of perhaps  $0.7 \text{ \AA}$ . from the centre. The explanation would appear to be that  $f(Cl^-) - f(Cl)$  is alternately positive and negative as  $\sin \theta/\lambda$  increases from zero, but that only in the range  $\sin \theta/\lambda < 0.4$ , where they considered the contribution of the ( $3p$ ) electrons to be appreciable, did James & Brindley make any change from  $f(Cl^-)$  in tabulating  $f(Cl)$ . The result given by calculation is in any case only an approximation to the actual distribution in an isolated atom or ion, and the effect on  $\rho_o - \rho_c$  of using an incorrect atomic scattering factor is merely to change the distribution of electrons in an area surrounding the nucleus. The number of electrons in this area is not changed, although the area must be increasingly extended the more  $f$  deviates from the true value, particularly at large values of  $\sin \theta/\lambda$ . Values of  $D_7$  in the immediate neighbourhood of the chlorine atom show no systematic variation from zero, but a circle of radius  $1.6 \text{ \AA}$ . drawn about the centre of this atom passes through a region in which  $D$  is consistently positive. The number of electrons lying between circles of radii  $1.3$  and  $2.0 \text{ \AA}$ ., excluding areas

already associated with other atoms, is about  $0.4$ . The direct evidence as to the state of ionization of the chlorine atom is therefore inconclusive, but it is improbable that it is fully ionized. Any interaction between the chlorine atom and the adenine molecule must take place mainly through the hydrogen bonds linking Cl to  $N_{10}$  and to  $N_9$ .

One unexpected feature of the electron distribution as a whole is its similarity to the 'calculated' distribution obtained by superposing isolated atoms. This is shown by the absence of the discrete maxima which one might have expected to find at the centres of covalent bonds. These maxima could be obscured by wrong choice of temperature-factor parameters of neighbouring atoms, but quite elaborate precautions were taken to ensure that the constants  $\alpha_r$  were as nearly correct as possible (see §2). The value of  $D_7$  is positive at the centres of CC and CN bonds (except in one case), the average value being  $+0.2 \text{ e.}\text{\AA}^{-2}$ . Since this is an average over ten bonds, it seems to differ from zero by an amount outside the range of random experimental error. Only in the case of the  $N_9C_8$  bond is there a pronounced maximum, which, however, contains only  $0.1 \text{ e}$ . The electron density in a covalent bond is therefore very little more than is obtained by placing two 'non-interacting' atoms one bond-length apart. This conclusion finds support from results published by Robertson (1945), who compared the projected electron densities in single, conjugated, double and triple CC bonds, and remarked that the apparent increase in electron density at the centre of a bond in at least the first three cases was about the same as resulted from bringing pairs of average carbon atoms to within the specified distances. Brill *et al.* (1939) conclude from their measurements on diamond that the electron density at the centre of a bond is higher than that at other points at the same distance from the nucleus, but it is not clear whether they have taken overlapping into account. Franklin (1950), however, finds it necessary to postulate a relatively large concentration of electrons around the centres of the bonds in graphite in order to explain certain features of the X-ray scattering of this substance. Unfortunately, it does not appear to be possible at the present time to predict theoretically even an approximate value of the electron density in a covalent bond, and clearly more results are required before definite conclusions can be drawn.

Even the hydrogen atoms have an electron distribution which is very similar to that of an isolated atom, the observed maximum electron densities being  $1.08, 1.07, 1.00, 0.87, 0.72$  and  $0.70 \text{ e.}\text{\AA}^{-2}$ , as compared with the calculated value of  $0.77 \text{ e.}\text{\AA}^{-2}$ . The observed electron distribution has a tendency to be elongated in a direction perpendicular to the covalent bond in the case of the hydrogen atoms, and the direction of maximum thermal vibration of  $N_{10}$  is also nearly perpendicular to the covalent bond  $C_6N_{10}$ . This is an indication of the smaller value of the force constant for the bending

rather than for the stretching of a bond. The directions of maximum thermal vibration of the chlorine atom and the water molecule can also be understood when one considers their environment, the direction in both cases being that in which they make no bonds (Fig. 7).

Finally, we may comment on the very close agreement obtained between calculated and observed structure factors. Over the range  $\sin \theta/\lambda < 0.5$ , two sets of independent measurements on two crystal specimens were made, and the value of  $\Sigma |F_1 - F_2| \div \Sigma |F_1|$  was 0.026. Over the same range  $\Sigma |F_o - F_c| \div \Sigma |F_o| = 0.046$ . This is simply another indication of the fact that the actual electron distribution differs very little from that in a set of 'Hartree' atoms, so that the corresponding atomic scattering factors, including that of hydrogen, are nearly correct. Nevertheless, it is surprising to find that the average atomic scattering factor of an isolated atom is correctly represented by the values given by James & Brindley, to within 4% over this range, as these results suggest. However, it must be remembered that any error in  $f_0$ , which corresponded to a general spreading out of the calculated electron distribution such as results from a neglect of the exchange principle (Hartree & Hartree, 1936), would be at least partially compensated by our method of estimating the temperature-factor parameter.

I should like to conclude by thanking Prof. Sir Lawrence Bragg and Dr W. H. Taylor for their continued support and encouragement. I am indebted to Dr M. V. Wilkes for permission to use the Hollerith tabulator on which much of the computing work was done, and to Mrs Gill for her assistance with the latter.

### References

- ABRAHAM, S. C., ROBERTSON, J. M. & WHITE, J. G. (1949). *Acta Cryst.* **2**, 233.
- ALBRECHT, G. (1939). *Rev. Sci. Instrum.* **10**, 221.
- BOOTH, A. D. (1947a). *Proc. Roy. Soc. A*, **190**, 490.
- BOOTH, A. D. (1947b). *Proc. Roy. Soc. A*, **190**, 482.
- BOOTH, A. D. (1948). *Nature, Lond.*, **161**, 765.
- BRILL, R., GRIMM, H. G., HERMANN, C. & PETERS, C. (1939). *Ann. Phys., Lpz.*, **34**, 393.
- BRILL, R., HERMANN, C. & PETERS, C. (1942a). *Ann. Phys., Lpz.*, **41**, 37.
- BRILL, R., HERMANN, C. & PETERS, C. (1942b). *Ann. Phys., Lpz.*, **41**, 233.
- BRILL, R., HERMANN, C. & PETERS, C. (1942c). *Ann. Phys., Lpz.*, **42**, 357.
- BROOMHEAD, J. M. (1948). *Acta Cryst.* **1**, 324.
- BROOMHEAD, J. M. (1951). *Acta Cryst.* **4**, 92.
- CLEWS, C. J. B. & COCHRAN, W. (1949). *Acta Cryst.* **2**, 46.
- COCHRAN, W. (1948). *Acta Cryst.* **1**, 54.
- COCHRAN, W. (1950). *Acta Cryst.* **3**, 268.
- COSTAIN, W. (1941). Ph.D. Thesis, University of Birmingham.
- CRUCKSHANK, D. W. J. (1949). *Acta Cryst.* **2**, 154.
- DUNITZ, J. D. & ROBERTSON, J. M. (1947). *J. Chem. Soc.* p. 142.
- FINBAK, C. & NORMAN, N. (1948). *Acta Chem. Scand.* **2**, 813.
- FRANKLIN, R. E. (1950). *Nature, Lond.*, **165**, 71.
- HARTREE, D. R. (1928). *Proc. Camb. Phil. Soc.* **24**, 89, 111.
- HARTREE, D. R. & HARTREE, W. (1936). *Proc. Roy. Soc. A*, **156**, 45.
- HUGHES, E. W. (1941). *J. Amer. Chem. Soc.* **63**, 1737.
- JAMES, R. W. & BRINDLEY, G. W. (1931). *Phil. Mag.* **12**, 81.
- MATHIESON, A. McL., ROBERTSON, J. M. & SINCLAIR, V. C. (1950). *Acta Cryst.* **3**, 251.
- PAULING, L. (1945). *The Nature of the Chemical Bond*. Cornell University Press.
- REIJEN L. J. VAN (1942). *Physica, 's Grav.*, **5**, 461.
- ROBERTSON, J. M. (1945). *J. Chem. Soc.* p. 249.
- TAYLOR, H. F. W. (1949). *Nature, Lond.*, **164**, 750.
- WRINCH, D. (1946). *Fourier Transforms and Structure Factors*. ASXRED Monograph, no. 2.

*Acta Cryst.* (1951). **4**, 92

## The Structures of Pyrimidines and Purines. IV. The Crystal Structure of Guanine Hydrochloride and its Relation to that of Adenine Hydrochloride

BY JUNE M. BROOMHEAD\*

*Crystallographic Laboratory, Cavendish Laboratory, Cambridge, England*

(Received 23 May 1950)

The crystal structure of guanine hydrochloride monohydrate has been established by X-ray analysis, employing two-dimensional Fourier methods. The cell dimensions and space group are different from those of adenine hydrochloride hemihydrate, but, in spite of this, the two structures are strikingly similar. From the disposition of intermolecular hydrogen bonds an attempt is made to deduce the positions of the hydrogen atoms covalently bound to nitrogen atoms of the purine molecules.

### 1. Introduction

An X-ray study of the hydrogen chloride salts of the purines adenine and guanine was undertaken as part of

\* Now at the Department of Chemical Crystallography, University Museum, Oxford, England.

a programme of investigation of a group of comparatively simple molecules which are constituents of nucleic acids. In this particular case the main object of the research has been to establish, if possible, which tautomeric forms of adenine and guanine exist in the solid