

## The Crystal Structure of Isoguanine Sulfate Monohydrate\*

BY E. SUBRAMANIAN† AND RICHARD E. MARSH

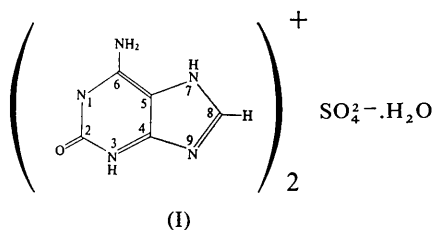
Gates and Crellin Laboratories of Chemistry, California Institute of Technology, Pasadena, California, U.S.A.

(Received 7 May 1969)

The crystal structure of isoguanine sulfate monohydrate,  $(C_5H_6N_5O)_2SO_4 \cdot H_2O$ , has been determined and refined by full-matrix least-squares methods based on 1698 reflections recorded on a diffractometer. The final  $R$  index is 0.037, the goodness of fit 1.6, and the estimated deviations in the positions of the heavier atoms about 0.002 Å. The crystals are monoclinic, space group  $A2/a$ , with cell dimensions  $a = 15.161$  (2),  $b = 14.791$  (2),  $c = 7.080$  (2) Å,  $\beta = 101.04$  (2)°. There are four formula units in the cell; the sulfate ions and water molecules lie on twofold axes. In the isoguanine cation, both nitrogen atoms [N(1) and N(3)] of the six-membered ring are protonated; N(7) of the five-membered ring is also protonated, but N(9) is not. The cations are slightly but significantly nonplanar; they are stacked on top of one another, across centers of symmetry, to form columns parallel to the  $c$  axis, the distances between successive cations within the columns alternating between 3.18 and 3.36 Å. A three-dimensional hydrogen-bond network involves all available protons.

### Introduction

Isoguanine differs from guanine in having the oxygen and amino substituents of the six-membered ring interchanged. Thus, it has many structural features in common with both guanine and cytosine. For this reason we have undertaken the present investigation of the structure of isoguanine sulfate monohydrate (I).



### Experimental

Crystals of isoguanine sulfate monohydrate,  $(C_5H_6N_5O)_2SO_4 \cdot H_2O$ , were grown by slow evaporation of an aqueous solution at room temperature. They grew as thin plates with principal faces  $\{010\}$  and as needles elongated along  $c$ . The space group and preliminary values of the unit-cell dimensions were determined from Weissenberg photographs about  $a$  and  $c$  and a precession photograph of the  $h0l$  zone. More accurate values of the cell dimensions were obtained by measuring  $2\theta$  values for a number of reflections on a General Electric XRD-5 diffractometer, using a fine-slit aperture (0.05°) and a small take-off angle (1°). The

density was measured by flotation in a mixture of benzene and bromoform. The crystal data are summarized in Table 1.

Table 1. Crystal data for isoguanine sulfate monohydrate

$(C_5H_6N_5O)_2SO_4 \cdot H_2O$	F.W. 418.3
$Z = 4$	$F(000)$ 864
Monoclinic	Systematic absences:
$a = 15.161$ (2) Å	$hkl$ : $(k+1)$ odd
$b = 14.791$ (2)	$h0l$ : $h$ odd
$c = 7.080$ (2)	Space group: $Aa$ or $A2/a$
$\beta = 101.04$ (2)°	$A2/a$ assumed
(Cu $K\alpha = 1.5418$ Å)	
$\mu = 26.1$ cm <sup>-1</sup>	
$D_o = 1.75$ g.cm <sup>-3</sup>	$D_c = 1.78$ g.cm <sup>-3</sup>

The space group  $A2/a$  was initially selected over  $Aa$  on the basis of the distribution of intensities, which suggested the presence of a center of symmetry; it was confirmed by the subsequent refinement. Similarly, the presence of four water molecules of crystallization was suggested by the density measurements and confirmed by the structure analysis.

Preliminary three-dimensional intensity data were collected visually from multiple-film equi-inclination Weissenberg photographs of layers 0–5 about  $c$  and of the zero layer about  $a$ . A total of 1324 reflections were recorded in this fashion, of which 1234 were of measurable intensity. These data were used for the derivation and initial refinement of the structure.

Subsequently, complete intensity data were collected on a Datex-automated General Electric XRD-5 diffractometer, equipped with a scintillation counter and a frequency discriminator; nickel-filtered Cu  $K\alpha$  radiation was used. The crystal was a square prism of dimensions 0.10 × 0.13 × 0.13 mm. Counts were recorded during a  $\theta$ – $2\theta$  scan, at a scanning speed of 2° (in  $2\theta$ ) min<sup>-1</sup>, and for 20 seconds at the two extrema of each scan.

\* Contribution No. 3841 from the Gates and Crellin Laboratories of Chemistry. This investigation was supported in part by a Research Grant No. HE-02143 from the National Heart Institute, U.S. Public Health Service, to the California Institute of Technology.

† Present address: Department of Physics, Indian Institute of Science, Bangalore-12, India.

In order to insure capturing the entire  $\alpha_1$ - $\alpha_2$  doublet, the scan range increased linearly with  $2\theta$ , being  $1.6^\circ$  at  $2\theta=0^\circ$  and  $4.6^\circ$  at  $2\theta=150^\circ$ . The 10,2,0 reflection was monitored every 15 reflections; its intensity showed no significant change with time. All reflections (1698 in number) with  $2\theta$  values between  $4\frac{1}{2}$  and  $153^\circ$  were recorded in this manner.

Each intensity  $I$  was assigned a variance  $\sigma^2(I)$  based on the statistics of the scan and background counts, but including an additional term,  $(0.02I)^2$ , to take account of unknown experimental errors. The intensities and their standard deviations were corrected for Lorentz and polarization factors, but not for absorption ( $\mu R \approx 0.15$ ).

#### Determination and refinement of the structure

A sharpened three-dimensional Patterson map, calculated from the visually estimated intensities, quickly led to the approximate coordinates of the atoms of the sulfate group – which, perforce, lies on a twofold rotation axis. The positions of the remaining heavy atoms were recovered from a three-dimensional electron density map phased by the sulfate group. Full-matrix least-squares refinement, including individual isotropic temperature factors, led to an  $R$  index of 0.17. At this stage a difference map evaluated in the plane of the isoguanine ion indicated the positions of its six hydrogen atoms; the water hydrogen atoms were positioned so as to form O(15)-H...O(13) hydrogen bonds. Further refinement, including anisotropic temperature factors for the heavier atoms, decreased the  $R$  index to 0.11 but resulted in negative values for the  $B(33)$  terms of some of the atoms (presumably reflecting incorrect scaling of the upper-layer Weissenberg photographs). At this stage the intensities were re-collected on the diffractometer.

Final refinement was by full-matrix least-squares, carried out (as were all other computations) on an IBM 7094 using the CRYRM system (Duchamp, 1964). A total of 156 parameters were adjusted; these were the coordinates for 22 atoms in the asymmetric unit (two of them, the sulfur atom and water oxygen atom, lying in special positions), isotropic temperature factors for the seven hydrogen atoms, anisotropic temperature factors for the 15 heavier atoms, and a scale factor. The quantity minimized was  $\sum w(F_o^2 - F_c^2/k^2)^2$ , where  $k$  is the scale factor and weights  $w$  were taken equal to the reciprocals of the variances of  $F_o^2$  evaluated during the data reduction process (see *Experimental*). Form factors for C, N, O, and S were taken from *International Tables for X-ray Crystallography* (1962) and for hydrogen from Table II of Stewart, Davidson & Simpson (1965).

The refinement converged with an  $R$  index of 0.037 and a goodness of fit,  $\left(\frac{\sum w(F_o^2 - F_c^2/k^2)^2}{m-s}\right)^{1/2}$ , of 1.6. During the final cycle no parameter shift exceeded 10% of its standard deviation. The final parameters and their

estimated standard deviations are given in Tables 2 and 3; observed and calculated structure factors are given in Table 4. The final electron density evaluated in the plane of the isoguanine ion is shown in Fig. 1

Table 2. Final positional parameters and their estimated standard deviations

Values for the heavier atoms have been multiplied by  $10^5$ ; for the hydrogen atoms, by  $10^4$ .

N(1)	7682 (8)	-11933 (8)	66204 (19)
C(2)	-736 (10)	-15982 (10)	60623 (22)
N(3)	-7946 (8)	-10971 (8)	63545 (20)
C(4)	-6742 (9)	-2476 (9)	71157 (21)
C(5)	1699 (9)	1312 (9)	76358 (21)
C(6)	9400 (9)	-3463 (9)	73575 (21)
N(7)	238 (8)	9907 (8)	82728 (19)
C(8)	-8682 (9)	10711 (11)	80999 (23)
N(9)	-13249 (8)	3341 (8)	74038 (19)
N(10)	17626 (8)	-298 (10)	77356 (21)
O(11)	-1447 (8)	-23448 (7)	53381 (18)
O(12)	32762 (7)	17168 (8)	7806 (21)
O(13)	22933 (7)	28650 (8)	15604 (17)
S(14)	$\frac{1}{2}$	22858 (3)	0
O(15)	$\frac{1}{2}$	19002 (12)	$\frac{1}{2}$
H(16)	1843 (14)	535 (16)	8255 (33)
H(17)	2212 (16)	-333 (16)	7384 (34)
H(18)	1241 (14)	-1549 (14)	6484 (32)
H(19)	-1334 (14)	-1337 (13)	5937 (28)
H(20)	425 (14)	1412 (14)	8725 (30)
H(21)	-1124 (11)	1599 (12)	8454 (26)
H(22)	2516 (18)	2335 (17)	6075 (35)

#### Description of the structure

An interesting feature of the structure is the location of the extra proton in the isoguanine cation. We find both nitrogen atoms of the six-membered ring, N(1) and N(3), to be protonated, along with N(7) of the five-membered ring. In direct contrast, the cations in gua-

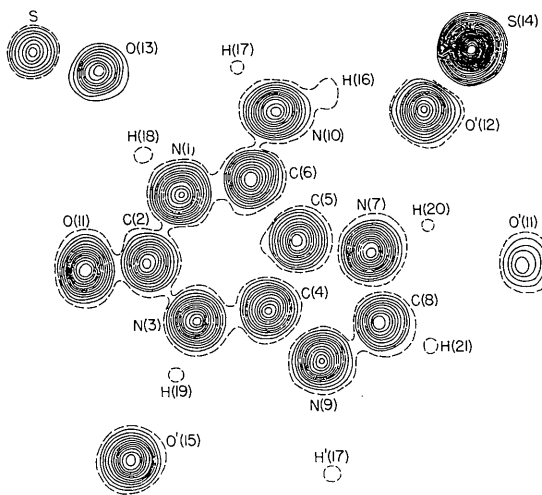


Fig. 1. The electron density in the plane of the isoguanine cation, calculated at the conclusion of the refinement. Contours are at 1 (dashed), 2, 3 ... e.Å<sup>-3</sup>.

nine hydrochloride (Iball & Wilson, 1965) and hypoxanthine hydrochloride (Sletten & Jensen, 1969) are protonated at the two nitrogen atoms [N(7) and N(9)] of the five-membered ring and at N(1) of the six-membered ring. In the neutral molecules guanine, guanosine, and

inosine (Bugg, Thewalt & Marsh, 1968) and in 9-ethylhypoxanthine (Kim & Rich, 1967), which are substituted on N(7) or N(9), the remaining proton is on N(1), while purine itself is protonated only at N(7) (Watson, Sweet & Marsh, 1965). It appears that rather subtle effects, such as intermolecular forces, may be responsible for the particular tautomeric form that these bases choose.

The isoguanine cation is significantly non-planar. Least-squares planes through various groups of atoms are described in Table 5 and shown schematically in Fig. 2. The six-membered ring is puckered in such a way that the *exo* atoms N(10) and O(11) are appreciably out of plane; in addition, there is a folding along the C(4)–C(5) bond, the dihedral angle between the planes of the five- and six-membered rings being 2.5°. Very similar, but less pronounced, patterns of nonplanarity have been noted by Sletten & Jensen (1969) in a number of other purine derivatives.

The bond distances and angles are shown in Fig. 3. The formal standard deviations in these values are about 0.002 Å and 0.1° where only heavier atoms are involved, and about 0.02 Å and 1° where hydrogen atoms are involved. The values have not been corrected for the effects of thermal motion; as discussed later, these effects appear to be small.

The bond distances are in good general agreement with the values tabulated by Sletten & Jensen (1969) for various purine derivatives, although a few differences are noteworthy: (1) The N(1)–C(2) distance is longer, and the N(1)–C(6) distance shorter, in the isoguanine cation than in guanine and hypoxanthine cations. This difference is undoubtedly an effect of the different position of the carbonyl oxygen atom O(11);

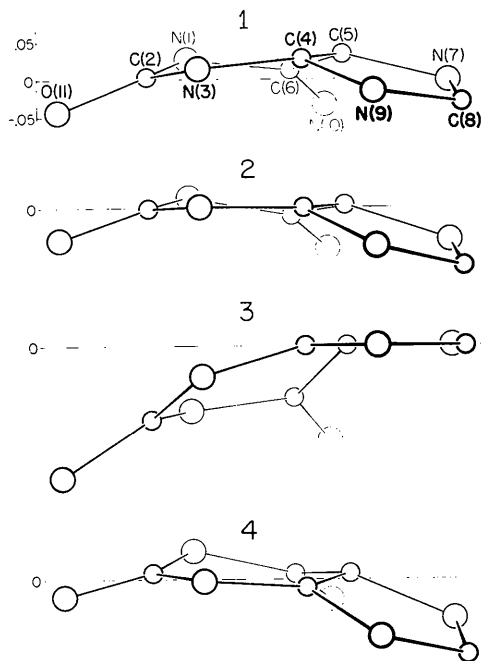


Fig. 2. Schematic representations showing the deviations of the atoms of the isoguanine cation from various least-squares planes. In each case the horizontal line is the trace of the best plane through the specified atoms (see Table 5).

Table 3. Final temperature parameters and their estimated standard deviations

The anisotropic coefficients for the heavier atoms are of the form

$$\exp [-(b_{11}h^2 - b_{22}k^2 - b_{33}l^2 - b_{12}hk - b_{13}hl - b_{23}kl)],$$

and have been multiplied by 10<sup>5</sup>. For the hydrogen atoms, isotropic *B*'s are given.

	<i>b</i> <sub>11</sub>	<i>b</i> <sub>22</sub>	<i>b</i> <sub>33</sub>	<i>b</i> <sub>12</sub>	<i>b</i> <sub>13</sub>	<i>b</i> <sub>23</sub>
N(1)	179 (5)	177 (6)	1164 (29)	56 (8)	225 (19)	-83 (19)
C(2)	240 (6)	167 (6)	930 (31)	-8 (10)	152 (22)	-13 (21)
N(3)	164 (5)	167 (5)	1314 (31)	-56 (8)	134 (19)	-113 (20)
C(4)	156 (5)	164 (6)	910 (29)	-17 (9)	172 (20)	20 (20)
C(5)	148 (6)	155 (6)	955 (29)	-4 (9)	141 (20)	-36 (20)
C(6)	160 (5)	187 (6)	844 (29)	20 (9)	141 (19)	42 (21)
N(7)	162 (5)	151 (5)	1107 (27)	-29 (8)	165 (18)	-101 (20)
C(8)	179 (6)	195 (6)	1235 (34)	31 (10)	205 (22)	-112 (23)
N(9)	144 (5)	208 (5)	1294 (30)	7 (8)	219 (18)	-84 (21)
N(10)	130 (5)	275 (6)	1542 (32)	16 (9)	205 (20)	-200 (23)
O(11)	341 (6)	184 (5)	1519 (28)	-11 (8)	230 (19)	-295 (19)
O(12)	223 (5)	309 (6)	2371 (35)	238 (9)	419 (21)	529 (23)
O(13)	242 (5)	267 (5)	1243 (26)	74 (8)	210 (17)	-233 (18)
S(14)	110 (2)	133 (2)	1094 (11)	0	159 (7)	0
O(15)	229 (7)	266 (7)	1229 (37)	0	182 (25)	0
H(16)		3.7 (5)				
H(17)		4.7 (6)				
H(18)		3.4 (4)				
H(19)		2.7 (4)				
H(20)		3.3 (4)				
H(21)		1.9 (3)				
H(22)		8.1 (6)				

the C=O bond is in conjugation with the N(1)-C(6) bond in isoguanine and with the N(1)-C(2) bond in guanine and hypoxanthine. The C(2)-N(3) bond is longer, and the C(4)-N(9) bond shorter, in isoguanine, reflecting the different localities of the extra proton. The

corresponding effect on the N(3)-C(4) and C(8)-N(9) distances appears to be small.

The large difference between the C-N-C angles at N(1) and N(3) is surprising. In guanine derivatives in which N(1) is protonated but N(3) is not, the angle at

Table 4. Observed and calculated structure factors ( $\times 10$ )

Reflections indicated with an asterisk (\*) were given zero weight in the least-squares calculations.

Table with columns for h, k, l, observed structure factors, and calculated structure factors. The table contains a large grid of numerical data points for various reflections.

N(1) is about  $125^\circ$  and that at N(3) about  $112^\circ$ ; in the isoguanine cation, with both nitrogen atoms protonated, the angles, though more nearly equal, still differ by  $6^\circ$ . Enlarging the angle at N(1) appears to be the most favorable way of accommodating the smaller internal angles at the neighboring carbon atoms C(2) and C(6), which in turn result from the double-bond character in the C(2)–O(11) and C(6)–N(10) bonds.

The angles at N(7) and N(9) are similar to those observed in other guanine derivatives, the angle at the three-coordinate atom [in this case, N(7)] being slightly larger than that at the two-coordinate atom.

### The temperature factors

The temperature factors for most of the atoms are small; they are represented as ellipsoids in Fig. 4. The largest amplitudes of motion for the atoms of the cation tend to be out-of-plane, ranging up to  $0.20 \text{ \AA}$  (r.m.s.) for N(10) and O(11), but there is no clear pattern of rigid-body motion. No correction to the bond distances within the five- or six-membered rings appears to be necessary; however, a riding-model treatment (Busing & Levy, 1965) indicates that the C(2)–O(11) bond distance should be increased by about  $0.012 \text{ \AA}$  and the C(6)–N(10) distance by about  $0.008 \text{ \AA}$ .

The motion of oxygen atom O(12) (maximum r.m.s. amplitude,  $0.25 \text{ \AA}$ ) of the sulfate group is considerably larger than that of O(13) ( $0.19 \text{ \AA}$ ). The indicated corrections to the S–O(12) distance is about  $0.02 \text{ \AA}$  and to the S–O(13) distance about  $0.01 \text{ \AA}$ , bringing both values up to about  $1.49 \text{ \AA}$ .

### Packing of the molecules

A stereoscopic view of a portion of the structure is shown in Fig. 4; views down the *b* and *c* axes are shown in Figs. 5 and 6. The two most interesting features of the intermolecular arrangement are the hydrogen bonding and the stacking of isoguanine cations.

### (i) The hydrogen bonding

There are six crystallographically distinct hydrogen atoms available for forming hydrogen bonds: five on the isoguanine cation, and one on the water molecule. All six are involved in hydrogen bonding. Similarly, all

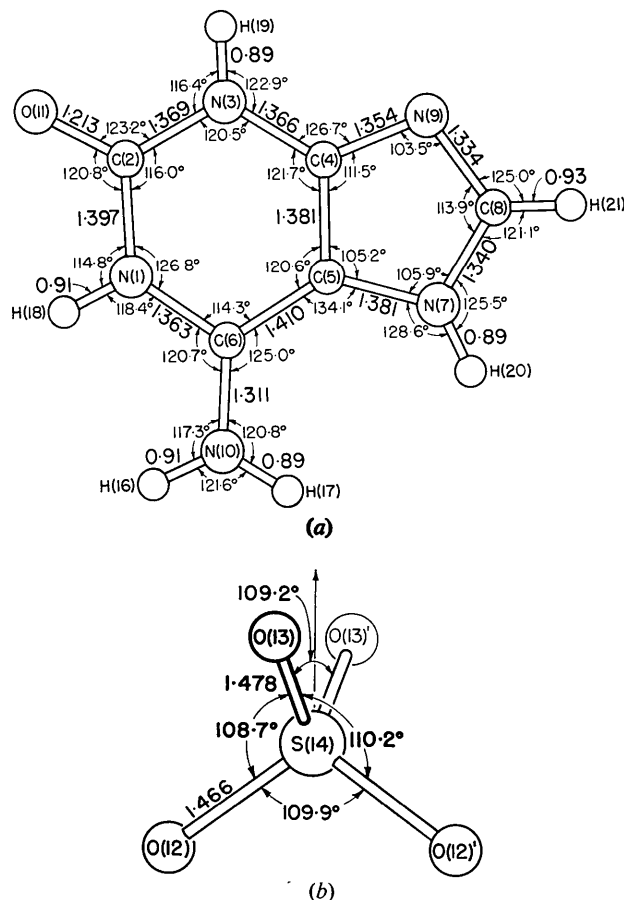


Fig. 3. Bond distances ( $\text{\AA}$ ) and angles ( $^\circ$ ): (a) the cation, (b) the anion.

Table 5. Equations of, and deviations from, various least-squares planes through atoms of the cation

The direction cosines *q* are with respect to *a*, *b* and *c*\*; *D* is the origin-to-plane distance.

	Plane 1	Plane 2	Plane 3	Plane 4
$q_a$	-0.0932	-0.0914	-0.1144	-0.0798
$q_b$	-0.3751	-0.3825	-0.3512	-0.3878
$q_c^*$	0.9222	0.9193	0.9292	0.9182
<i>D</i>	4.8533 $\text{\AA}$	4.8691 $\text{\AA}$	4.9516 $\text{\AA}$	4.8511 $\text{\AA}$
N(1)	0.025 $\text{\AA}$	0.012 $\text{\AA}$ *	-0.087 $\text{\AA}$ *	0.037 $\text{\AA}$
C(2)	0.004	-0.006	-0.099*	0.009
N(3)	0.019	0.001	-0.042*	-0.002
C(4)	0.028	-0.001	-0.001	-0.010
C(5)	0.038	0.006	0.000	0.008
C(6)	0.012	-0.011	-0.069*	0.008
N(7)	-0.002	-0.045*	0.000	-0.054*
C(8)	-0.033	-0.079*	-0.001	-0.104*
N(9)	-0.015	-0.052*	0.001	-0.078*
N(10)	-0.033	-0.058*	-0.126*	-0.027
O(11)	-0.044	-0.044*	-0.177*	-0.023

\* Indicates atom assigned zero weight in calculating the plane; all other atoms were weighted equally.



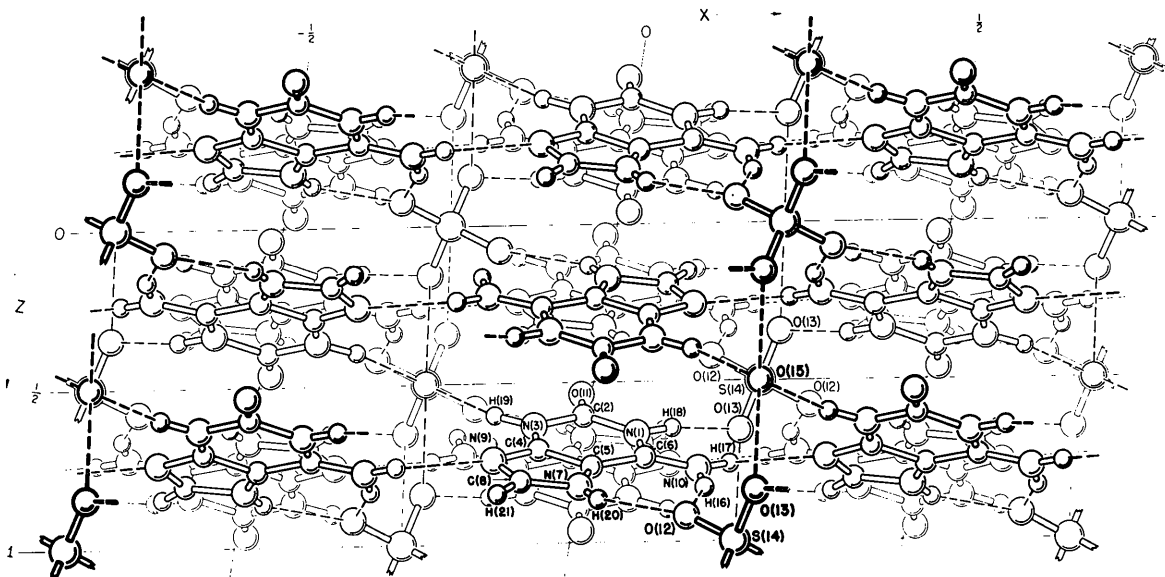


Fig. 5. A drawing of the structure viewed down the  $b$  axis. The hydrogen atoms of the water molecules are omitted.

tween atoms of different electronegativities, suggesting that dipole attractions are at least partly responsible for the stacking. It is somewhat surprising that these attractions are sufficient to overcome the repulsion due to the positive charges on the isoguanine cations; however, these positive charges are distributed throughout the cations and probably transmitted, in part at least, to the neighboring hydrogen-bond acceptors.

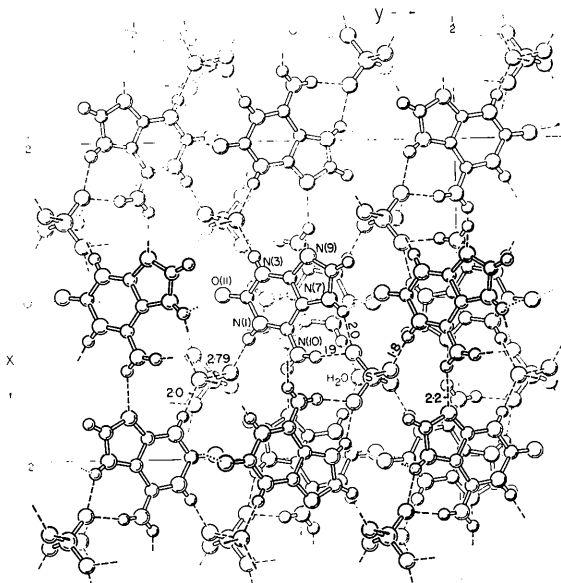


Fig. 6. A drawing of the structure viewed along the  $c$  axis. The hydrogen atoms of the water molecules are omitted.

Table 7. Shortest interatomic distances (less than 3.4 Å) between cations within a stack

Relative to the parent cation at  $x, y, z$  (Table 2), distances of type  $a$  involve the cation at  $-x, -y, 1-z$  and that of type  $b$  the cation at  $-x, -y, 2-z$ .

	Type	$D$
N(1)···C(4)	$a$	3.379 Å
N(1)···C(8)	$a$	3.379
N(1)···N(9)	$a$	3.371
C(2)···C(5)	$a$	3.382
C(2)···N(7)	$a$	3.213
N(3)···C(6)	$a$	3.360
C(4)···C(6)	$a$	3.236
N(7)···O(11)	$a$	3.280
C(5)···N(7)	$b$	3.399

#### References

- BUGG, C. E., THEWALT, U. T. & MARSH, R. E. (1968). *Biochem. Biophys. Res. Comm.* **33**, 436.
- BUSING, W. R. & LEVY, H. A. (1964). *Acta Cryst.* **17**, 142.
- CHAN, S. I., SCHWEIZER, M. P., TS'O, P. O. P. & HELMKAMP, G. K. (1964). *J. Amer. Chem. Soc.* **86**, 4182.
- DUCHAMP, D. J. (1964). *Program and Abstracts*, A. C. A. meeting, Bozeman, Montana; paper B-14, p. 29.
- IBALL, J. & WILSON, H. R. (1965). *Proc. Roy. Soc. A* **288**, 418.
- International Tables for X-Ray Crystallography* (1962). Vol. III. Table 3.3.1 A. Birmingham: Kynoch Press.
- KIM, S.-H. & RICH, A. (1967) *Science*, **158**, 1946.
- SLETTEN, E. & JENSEN, L. H. (1969). *Acta Cryst.* **B25**, 1608.
- STEWART, R. F., DAVIDSON, F. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175.
- WATSON, D. G., SWEET, R. M. & MARSH, R. E. (1965). *Acta Cryst.* **19**, 573.