RALPH, R. K., CONNORS, W. J., SCHALLER, H. & KHORANA, H. G. (1963). *J. Amer. Chem. Soc.* **85**, 1983.
SARKAR, P. K. & YANG, J. T. (1965). *Biochem. Biophys. Res. Comm.* **20**, 346.
SHEFTER, E. & TRUEBLOOD, K. N. (1965). *Acta Cryst.* **18**, 1067.
SOLIE, T. N. & SCHELLMAN, J. A. (1968). *J. Mol. Biol.* **33**, 61.
STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175.
SUNDARALINGAM, M. (1965). *J. Amer. Chem. Soc.* **87**, 599.
SUNDARALINGAM, M. & JENSEN, L. H. (1965). *J. Mol. Biol.* **13**, 914.
SUNDARALINGAM, M., RAO, S. T. & BUGG, C. E. (1969). *Abstr. Amer. Cryst. Assoc. Meeting*, Seattle, Wasington, L5.
TS'O, P. O. P. & CHAN, S. I. (1964). *J. Amer. Chem. Soc.* **86**, 4176.
TS'O, P. O. P., MELVIN, I. S. & OLSEN, A. C. (1963). *J. Amer. Chem. Soc.* **85**, 1289.

Acta Cryst. (1970). B**26**, 1101

The Crystal and Molecular Structure of Inosine

BY A. R. I. MUNNS AND P. TOLLIN

Carnegie Laboratory of Physics, The University of Dundee, Dundee, Scotland

(*Received* 14 *July* 1969)

The structure of inosine ($C_{10}N_4O_5H_{12}$) which crystallizes in the space group $P2_1$ with one molecule per asymmetric unit and with unit-cell dimensions: $a=4{\cdot}818\pm0{\cdot}005$, $b=10{\cdot}45\pm0{\cdot}01$, $c=10{\cdot}97\pm0{\cdot}01$ Å and $\beta=90°43'\pm2'$, has been determined from X-ray intensity data collected from linear and four-circle diffractometers. The structure was solved by a Patterson function interpretation method and the positional and thermal parameters were refined by the method of least squares, using anisotropic thermal parameters for the non-hydrogen atoms. The final R value for the 1298 observed reflexions was 0·046 and the standard deviations in the bond lengths and angles are about 0·004 Å and 0·3° respectively. The purine ring in inosine is planar, but both O(10) and C(1′) are significantly displaced from this plane. The dihedral angle between the base and sugar planes is 71·0° and the glycosidic torsion angle, φ_{CN} is $-10{\cdot}6°$. Atom C(3′) of the ribose ring is displaced by 0·63 Å from the plane of the remaining ring atoms and is in the *endo* conformation. The orientation of the C(5′)–O(5′) bond is *gauche* to C(4′)–O(1′) and *trans* to C(4′)–C(3′), the φ_{OO} and φ_{OC} angles being 74·7 and 169·0° respectively. All available groups participate in the hydrogen bonding. There is in addition one particularly short C–O distance involving a hydrogen atom.

Introduction

The determination of the structure of inosine was undertaken as part of a series of structure determinations of nucleic acid components in progress in this laboratory. Inosine is a nucleoside which occurs occasionally in ribonucleic acid (RNA), particularly in molecules of transfer RNA. In transfer RNA inosine appears to form part of a number of anticodons and it has been suggested (Crick, 1966; Woese, 1967) that it is important because it can form a base pair with any of the bases, adenine, uracil or cytidine. Accurate structural information may help towards an understanding of the function of inosine in the anticodon and may be useful if model-building is required for the solution of the structure of crystalline transfer RNA. The determination of this structure provided an opportunity to test the usefulness of incorporating the rotation function of Rossmann & Blow (1962) into the Patterson function interpretation procedure in use in this laboratory. A preliminary account of this work has been given (Tollin & Munns, 1969).

Experimental

Crystals of inosine ($C_{10}N_4O_5H_{12}$), whose structural formula appears in Fig. 1 along with the numbering system used in this paper, were obtained by evaporation from aqueous solutions. Three distinct crystal forms were obtained depending on the rate of evaporation and the temperature at which it took place. The first form, obtained by fast evaporation at 20 °C, consisted of colourless needles showing apparent orthorhombic symmetry and cell dimensions $a=8{\cdot}16\pm0{\cdot}04$, $b=13{\cdot}3\pm0{\cdot}2$, $c=21{\cdot}4\pm0{\cdot}2$ Å. However, these crystals showed a marked tendency to form twinned crystals and were not investigated further. Slow evaporation from partially sealed test tubes at 19 °C produced needle crystals of monoclinic symmetry with cell dimensions $a=6{\cdot}68\pm0{\cdot}05$, $b=11{\cdot}3\pm0{\cdot}1$, $c=17{\cdot}4\pm0{\cdot}1$ Å, $\beta=98{\cdot}3\pm0{\cdot}1°$, belonging to the space group $P2_1$. This structure has been determined by Bugg, Thewalt & Marsh (1968) and independently in this laboratory. A comparison of the results of the two structure determinations is in preparation (Munns, Tollin, Wilson & Young, 1970).

The present structure determination was performed on crystals obtained by slow evaporation at 25 °C. The cell dimensions, obtained from Weissenberg and precession photographs with Cu $K\alpha$ radiation and refined on a linear diffractometer with Mo $K\alpha$ radiation ($\lambda = 0{\cdot}71069$ Å) were $a = 4{\cdot}818 \pm 0{\cdot}005$, $b = 10{\cdot}45 \pm 0{\cdot}01$, $c = 10{\cdot}97 \pm 0{\cdot}01$ Å, $\beta = 90°43' \pm 2'$. The systematic absences, $0k0$ with $k = 2n+1$, determined the space group as $P2_1$. There are two molecules in the unit cell.

All possible reflexions within the sin θ value corresponding to the radius of the limiting sphere for Cu $K\alpha$ radiation were measured on a Hilger & Watts linear diffractometer, with Mo $K\alpha$ radiation and balanced filters, for a crystal mounted on the b axis. For a second crystal mounted on the a axis, data were collected for the five layers up to $(4kl)$. A number of reflexions close to the origin and at the centre of each layer cannot be accurately measured by this technique and the data for these reflexions were collected using a Wooster four-circle diffractometer with Cu $K\alpha$ radiation and b-axis crystal.

The linear diffractometer data for the a- and b-axis

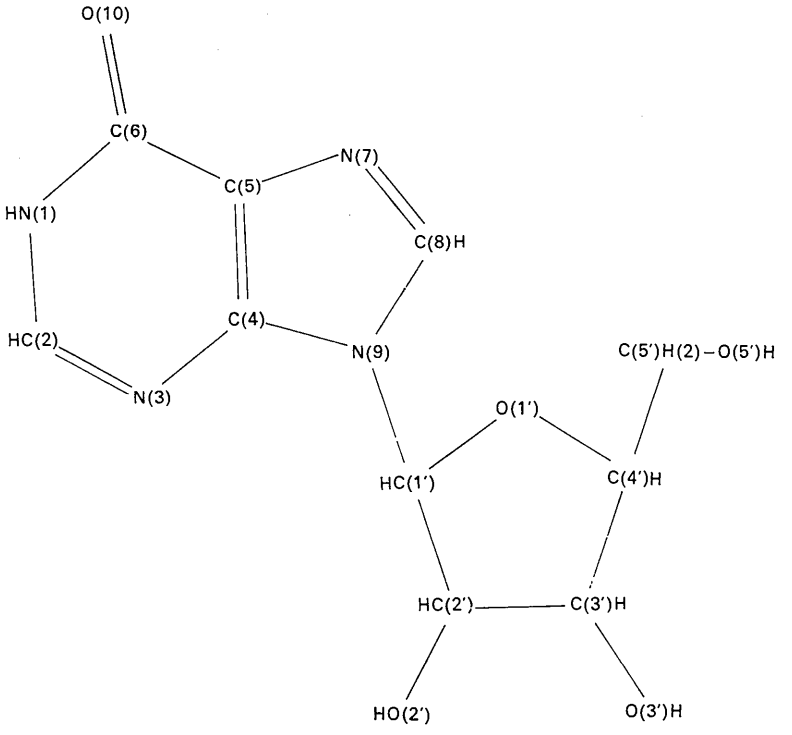

Fig. 1. The structural formula of inosine and the numbering system used in this paper.

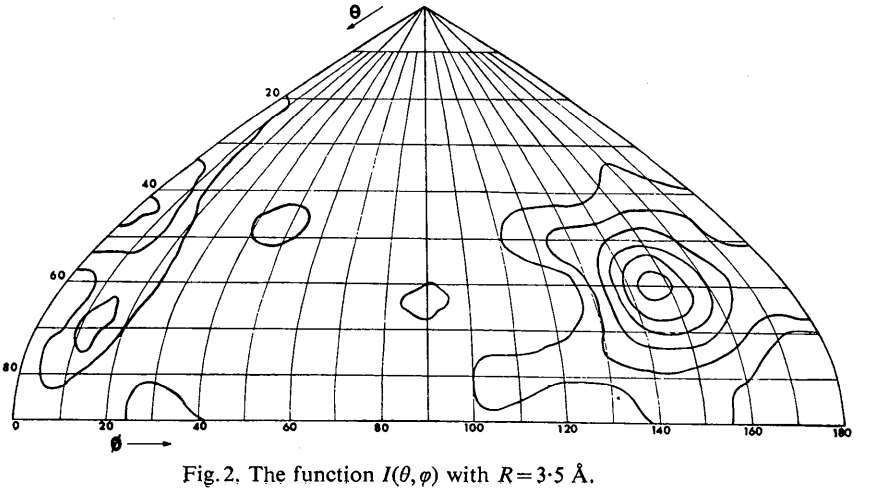

Fig. 2. The function $I(\theta, \varphi)$ with $R = 3{\cdot}5$ Å.

crystals were scaled together to produce a value for each reflexion. The R value between the two independent sets of data for common reflexions was 0·047, where $R=\sum|(F_b-F_a)|/\sum|F_b|$. The four-circle data were scaled to the average values of the linear diffractometer data. Of the 1334 possible observable reflexions, 28 were measured to be effectively zero.

Structure determination

The atoms of the purine residue and atom C(1′) were expected to be planar. Sharpened Patterson coefficients $|F_s(\mathbf{h})|^2$ were calculated using the sharpening function proposed by Wunderlich (1965) with the constants a and p having values of 2·0 and 7·25 respectively. The function $I(\theta,\varphi)$ (Tollin & Cochran, 1964) was computed using all the 1306 $|F_s(\mathbf{h})|^2$ values, and a value of 3·5 Å for the radius R of the disc. The results, plotted on a Sanson–Flamsteed projection, are shown in Fig. 2. Although there will be vectors in the purine residue longer than 3·5 Å, the magnitude of R is restricted by the short a-axis parameter of 4·8 Å. A radius much larger than 3·5 Å would lead to interference with adjacent Patterson function origin peaks. The largest peak on the map indicates that the normal to the purine group has a direction defined by $\theta=60°$ and $\varphi=147°$, where θ and φ are the spherical polar angles with respect to the orthogonal axial set a^*, b, c. The orientation of the purine group in this plane could have been found by calculating the Patterson section as originally suggested (Tollin & Cochran, 1964). However, the opportunity was taken to test the application of the rotation function (Rossmann & Blow, 1962) to an unknown structure. The incorporation of the rotation function into the general Patterson interpretation method in use in this laboratory will be discussed elsewhere (Tollin, Young & Munns, to be published). A model of the purine residue was placed in an orthorhombic unit-cell, with the plane of the purine group in the xy plane, with the C(4)–C(5) bond along the y axis and C(4) at the origin. Sharpened Patterson coefficients were calculated for this model and the 100 largest coefficients were abstracted.

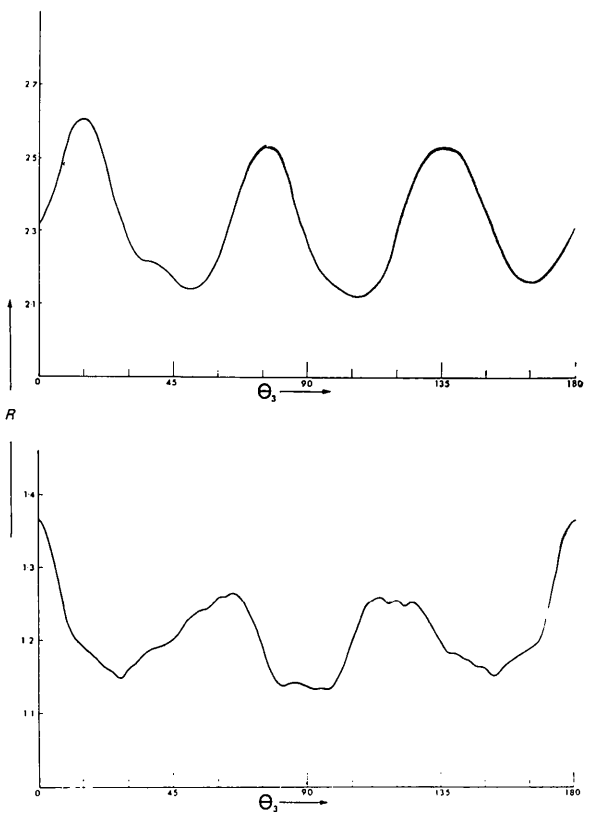

Fig. 3. The rotation functions R for rotations of angle θ_3 in the plane of the purine group, (*a*) for rotation of model with full inosine data and (*b*) for self rotation of the model.

The rotation function was calculated using a program of the type described by Tollin & Rossmann (1966). Two calculations were performed, one using (in the notation of Tollin & Rossmann, 1966) the 100 largest model coefficients as the $|F_h|^2$ and the $|F_s(\mathbf{h})^2|$ for inosine as the $|F_p|^2$ values, the other using all the sharpened model coefficients as the $|F_p|^2$ values. Because of the choice of orientation in the model, a rotation in the plane defined by the spherical polar angles θ and φ corresponds to points in the rotation function having $\theta_1=\varphi+\pi/2$, $\theta_2=\theta$ and θ_3 having all possible values. In the calculation the radius of the sphere used was 3·5 Å and the interpolations were performed over 245 points and 175 points for first and second calculations respectively. The values of the rotation functions for the model against itself, $R(0,0,\theta_3)$, and the model against the inosine structure $R(237°, 60°, \theta_3)$ are shown in Fig. 3. Because of the symmetry of the structure and the model (Tollin, Main & Rossmann, 1966) rotations of θ_3 and $\pi+\theta_3$ cannot be distinguished. The θ_3 angle was determined by comparing these lines through the two rotation functions. It is apparent that a rotation of 15° in θ_3 is required to bring the lower plot into coincidence with the upper.

In addition to the exact symmetry of the structure and the model, the model is almost symmetric about a line perpendicular to the C(4)–C(5) bond in the plane of the molecule. No attempt was made to distinguish between these using the rotation function. The position of the purine group was determined and the correct choice between these alternatives was made using the function $Q(X_0Z_0)$ (Tollin, 1966). The Q functions for the alternatives are shown in Fig. 4. The coordinates used in calculating the Q function were obtained by applying the appropriate rotation matrix and arbitrarily positioning the atom C(4) at the origin. As expected the maps are similar but the largest peak occurs on the map of Fig. 4(*b*) at $X_0=0·1615$ and $Z_0=0·3114$.

The relative coordinates for the correct purine model were referred to the origin of the unit cell by adding on the components indicated by the Q function and were

then used to calculate structure factors. An overall isotropic temperature factor of 4·5 was used. This value, which was felt to be rather high, had been obtained from an analysis of the reflexion data using the method described by Wilson (1942). The R value obtained was 0·50, where $R=\sum(|F_c|-|F_o|)/\sum|F_o|$. An electron density map was calculated using the phases obtained from this structure factor calculation and all the observed structure amplitudes. In addition to peaks at the positions of the purine atoms, peaks of about half the height of these peaks were found. It was possible to interpret these peaks as the missing non-hydrogen atoms of the ribose sugar. The positions of all the non-hydrogen atoms were extracted from this synthesis and used to calculate structure factors and a further electron density map. The R value obtained was 0·31 and the electron density map verified the absence of any water of crystallization.

Refinement of the structure

The structure was refined by the method of least squares using the block-diagonal approximation and a program written by Professor J. Trotter (Toronto Univ.) modified and adapted by the staff of the Computing Laboratory of the University of Dundee for use on an Elliot 4130 computer.

By the use of the coordinates of the 19 heavy atoms obtained from the electron density map, the structure was refined using individual isotropic temperature factors. At this stage all but the very largest reflexions were given equal weight. In three cycles the R value

Table 1. *Positional parameters* ($\times 10^4$) *as fractions of the cell edges and standard deviations for the non-hydrogen atoms*

	x/a	*y/b*	*z/c*
N(1)	4943 (6)	1879 (3)	3715 (3)
C(2)	3372 (8)	1952 (4)	2684 (3)
N(3)	1625 (6)	1070 (3)	2317 (3)
C(4)	1509 (6)	0072 (3)	3114 (3)
C(5)	2965 (6)	−0095 (3)	4187 (3)
C(6)	4995 (7)	0847 (3)	4525 (3)
N(7)	2265 (6)	−1219 (3)	4749 (2)
C(8)	0411 (7)	−1724 (3)	4012 (3)
N(9)	−0148 (5)	−0990 (3)	3001 (2)
O(10)	6676 (5)	0819 (3)	5381 (2)
C(1′)	−1932 (6)	−1311 (3)	1940 (3)
O(1′)	−3507 (4)	−2412 (2)	2230 (2)
C(2′)	−0206 (6)	−1650 (3)	0824 (3)
O(2′)	−1851 (5)	−1506 (2)	−0250 (2)
C(3′)	0225 (6)	−3080 (3)	1029 (3)
O(3′)	1166 (4)	−3781 (2)	0013 (2)
C(4′)	−2618 (6)	−3485 (3)	1479 (3)
C(5′)	−2595 (7)	−4696 (3)	2221 (3)
O(5′)	−5333 (5)	−5168 (3)	2426 (2)

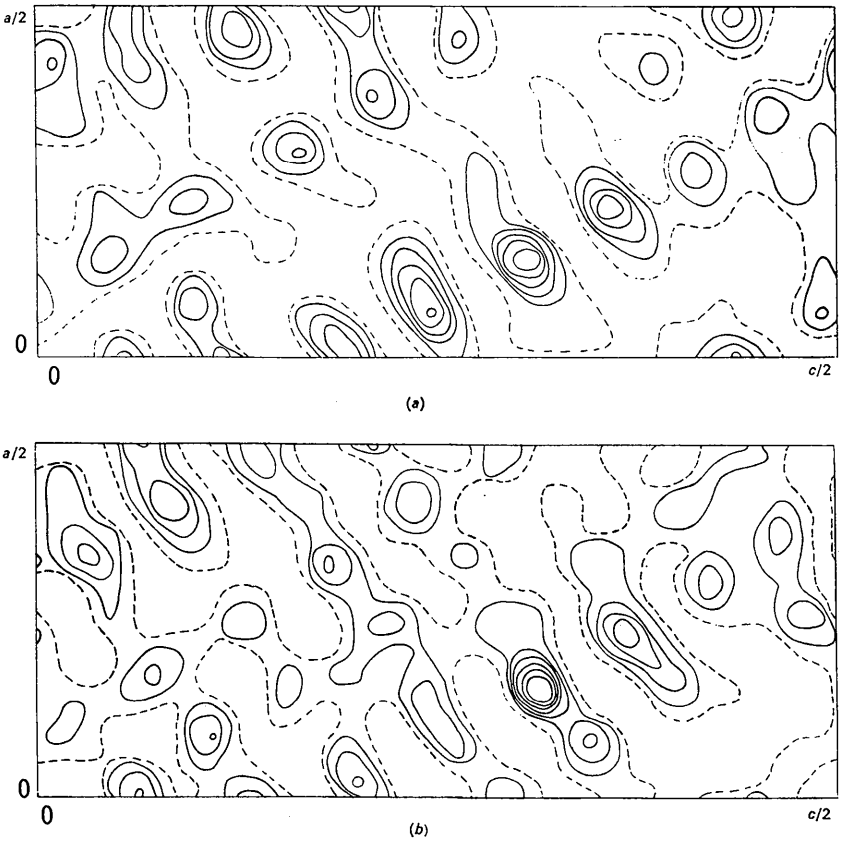

Fig. 4. The $Q(X_0, Z_0)$ functions for the alternative purine model orientations.

dropped to 0·10, and the temperature factors became more reasonable, all lying in the range 1·7 to 2·7. An electron density map and a difference Fourier synthesis were then calculated.

The possible positions of the 12 hydrogen atoms were calculated assuming normal bond lengths, angles and possible hydrogen bond schemes suggested by close contacts of non-hydrogen atoms. All the hydrogen atoms, apart from H(5′) and H(5″), could be identified easily, with peaks in the Fourier and in the difference Fourier syntheses of height about 0·4 e.Å^{-3}. The failure to observe the hydrogen atoms H(5′) and H(5″) was due in part to the large anisotropy of the close-by atoms C(5′) and O(5′). Since the missing hydrogen atoms are unimportant in the bonding scheme, their coordinates were taken as those calculated. The 12 hydrogen atoms were introduced into the refinement with fixed coordinates and an overall temperature factor of 2·5. The 19 non-hydrogen atoms were refined using individual anisotropic temperature factors of the form

$$T=\exp\left[-(\beta_{11}h^2+\beta_{12}kl+\beta_{13}hl+\beta_{22}k^2+\beta_{23}kl+\beta_{33}l^2)\right].$$

All the observed reflexions were used and the weighting scheme was unchanged. An R value of 0·062 was reached before the hydrogen coordinates and temperature factors were refined by least squares. Shifts in coordinates and temperature factors were severely curtailed by fudge factors to avoid excessive movements. The R value fell to 0·054. It was clear that eight low-order reflexions, 004, 011, 012, 10−3, 10−1, 102, 11−2 and 11−1, were calculating high. All of these reflexions had been measured with high counts on the linear diffractometer and although precautions had been taken to minimize lost counts by keeping to low counting rates it appeared that these reflexions may have had a peak counting rate high enough to cause error. These eight reflexions were removed from the further refinements.

For the final stage of refinement each reflexion was given a weight ω given by

$$\omega=\frac{1}{1+[(|F_o|-a)/b]^2},$$

where a and b are constants. Values of $a=8$ and $b=6$ were chosen so that the average value of ω $(|F_o|-|F_c|)^2$ (*i.e.* $\omega\Delta^2$) as a function of structure amplitude was approximately constant. Initially the coordinates and thermal parameters of the 12 hydrogen atoms were fixed. The thermal parameters were chosen in each case to be the isotropic temperature factors of the heavy atom to which the hydrogen atom was covalently bonded. Only at the end of the refinement were the parameters of hydrogen atoms allowed to change. The

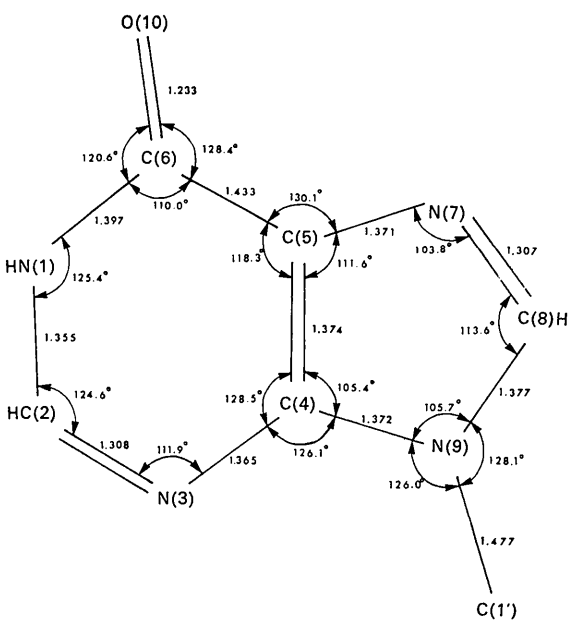

Fig. 5. The bond lengths and valence bond angles for the non-hydrogen atoms of the purine base.

Table 2. *Anisotropic thermal parameters* ($\times 10^5$) *and standard deviations for the non-hydrogen atoms*

	B_{11}	B_{12}	B_{13}	B_{22}	B_{23}	B_{33}
N(1)	3114 (109)	−758 (91)	−555 (86)	516 (23)	−60 (40)	543 (21)
C(2)	4032 (149)	−814 (120)	−608 (111)	521 (28)	143 (50)	584 (21)
N(3)	3489 (113)	−545 (93)	−899 (82)	446 (22)	183 (37)	496 (20)
C(4)	2129 (102)	30 (91)	−198 (81)	407 (25)	−101 (41)	441 (21)
C(5)	2424 (108)	−85 (96)	−240 (82)	395 (22)	38 (41)	414 (20)
C(6)	2544 (109)	56 (102)	−188 (85)	483 (26)	−70 (43)	399 (20)
N(7)	2775 (100)	−154 (89)	−251 (77)	463 (22)	109 (39)	481 (19)
C(8)	2800 (119)	−56 (101)	−536 (93)	472 (26)	95 (46)	485 (23)
N(9)	2398 (95)	−9 (77)	−306 (72)	362 (20)	10 (35)	407 (17)
O(10)	3188 (91)	−462 (82)	−914 (69)	656 (21)	30 (36)	544 (17)
C(1′)	1869 (100)	−202 (92)	−405 (81)	441 (25)	35 (41)	450 (20)
O(1′)	2024 (74)	−310 (68)	109 (62)	450 (17)	−224 (31)	549 (17)
C(2′)	2583 (111)	−281 (93)	−348 (85)	369 (23)	−13 (40)	409 (20)
O(2′)	3991 (99)	−104 (79)	−723 (69)	469 (19)	71 (32)	415 (16)
C(3′)	1981 (99)	−23 (88)	−182 (78)	377 (22)	−16 (39)	403 (20)
O(3′)	2445 (80)	335 (70)	−323 (61)	498 (19)	−163 (32)	433 (15)
C(4′)	2034 (103)	−167 (87)	−167 (85)	354 (23)	−33 (42)	501 (22)
C(5′)	2742 (117)	−356 (100)	11 (99)	386 (25)	126 (46)	667 (21)
O(5′)	3477 (96)	−1177 (79)	504 (74)	550 (19)	119 (36)	639 (19)

final R value was 0·046. Inclusion of the eight low order reflexions and the unobserved reflexions with a value of zero, gave an R value of 0·051. A difference synthesis was calculated to check that there were no unexplained high regions. Finally a plot of the average value of $\omega\Delta^2$ against structure amplitude was obtained to check that it was reasonably constant.

The maximum heavy atom and hydrogen atom coordinate shift were 0·0007 and 0·02 Å respectively, correponding to $0{\cdot}2\sigma$ and $0{\cdot}5\sigma$ respectively. The final coordinates for the non-hydrogen atoms and their estimated standard deviations appear in Table 1 and the anisotropic temperature factors appear in Table 2. The coordinates of the hydrogen atoms and their isotropic temperature factors appear in Table 3. Observed and calculated structure amplitudes are listed in Table 4. The intramolecular bond lengths and angles for the non-hydrogen atoms are given in Table 5 and the covalent bonds involving the hydrogen atoms are listed in Table 6. The standard deviations quoted for the bond lengths have not been modified to include the error in the cell dimensions.

At the 99·5% level of confidence the error in the bond lengths is an order of magnitude greater than the errors in the individual cell parameters.

Table 3. *Positional parameters* ($\times 10^3$) *as fractions of the cell edges, the isotropic temperature factor and standard deviations for the hydrogen atoms*

	x/a	y/b	z/c	β
H(1)	605 (9)	247 (6)	393 (4)	2·1 (1·0)
H(2)	359 (9)	267 (5)	213 (4)	1·9 (0·9)
H(8)	−055 (9)	−254 (5)	418 (4)	1·4 (0·9)
H(1′)	−339 (8)	−061 (4)	173 (4)	0·9 (0·8)
H(2′)	154 (8)	−112 (4)	083 (3)	0·5 (0·8)
H(3′)	153 (8)	−322 (4)	167 (4)	0·7 (0·8)
H(4′)	−388 (8)	−356 (4)	079 (4)	0·5 (0·7)
H(5′)	−160 (9)	−461 (4)	299 (4)	1·2 (0·9)
H(5″)	−175 (9)	−534 (5)	167 (4)	1·7 (0·9)
H(2″)	−204 (10)	−069 (5)	−044 (4)	2·0 (1·0)
H(3″)	000 (10)	−358 (5)	−060 (4)	2·2 (1·0)
H(5″)	−594 (10)	−485 (5)	311 (4)	2·3 (1·0)

Discussion of the molecular and crystal structure

The conformation and molecular dimensions of the purine and sugar units of inosine are considered and compared with similar structures. The intermolecular contacts and in particular the hydrogen bonding scheme are discussed.

The purine base

The least-squares plane through the nine non-hydrogen atoms of the purine residue is given in Table 7 along with the deviations of the most relevant atoms from this plane. In terms of the θ,φ angles previously derived by integrating the Patterson function, this plane corresponds to $\theta = 59{\cdot}5°$ and $\varphi = 147°$.

The large displacements of the atoms O(10) and C(1′) correspond to angles of 7·5 and 5·5° respectively between the connecting bond and its projection in the base plane. Variation in the displacement of the atom C(1′) from the base plane in nucleoside structures is generally acknowledged to be due to packing forces since different deviations of this atom have been noted for different crystal forms of the same nucleoside (Iball, Morgan & Wilson, 1968). Large displacements of this atom have been observed in the related compounds, for example, displacements of 0·22 Å in deoxyadenosine (Watson, Sutor & Tollin, 1965), 0·42 Å in vitamin B_{12} (Hodgkin, Lindsey, Spark, Trueblood & White, 1962), 0·15 Å in 5-bromo-2′-deoxyuridine (Harris & McIntyre, 1964) and 0·15 Å in 5′-bromo-uridine (Iball, Morgan & Wilson, 1966).

Intermolecular forces may also be the reason for the deviation of atom O(10) from the base plane. The bond lengths and angles of the purine base appear in Fig. 5 and are in close agreement with those previously determined purine based structures. A comparison of the bond lengths and angles of inosine with guanine hydrochloride dihydrate (Iball & Wilson, 1965), 8-azaguanine monohydrate (Sletten, Sletten & Jensen, 1968), deoxyadenosine (Watson, Sutor & Tollin, 1965), and purine (Watson, Sweet & Marsh, 1965), appears in Table 8.

The glycosidic bond length of 1·477 Å is in good agreement with other nucleoside structures.

The ribose ring

In the crystal structure of nucleosides and nucleotides, strain in the five-membered ring of the sugar, close intermolecular contact of exocyclic substituents and special packing considerations such as intermolecular hydrogen bonds cause a puckering of the ring such that the C(2′) or C(3′) atom is displaced by about 0·5 Å from the plane of the other four atoms (Spencer, 1959). This displacement may be *endo*, *i.e.* lying on the same side of the sugar plane as C(5′), or *exo*, *i.e.* lying on the opposite side.

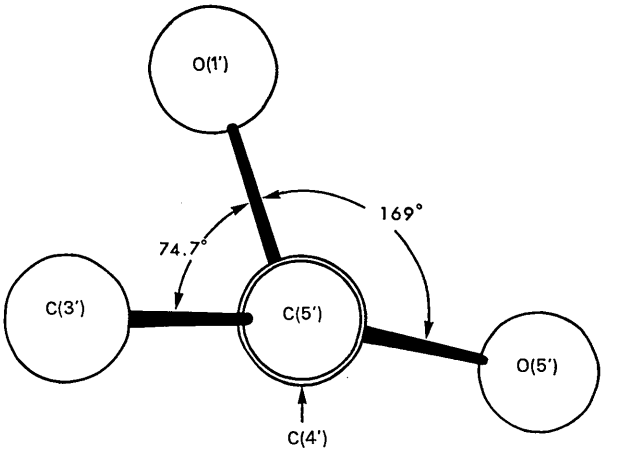

Fig. 6. The conformation of the C(5′) atom of the sugar viewed down the C(5′)–C(4′) bond.

Table 4. *Observed and calculated structure factors*

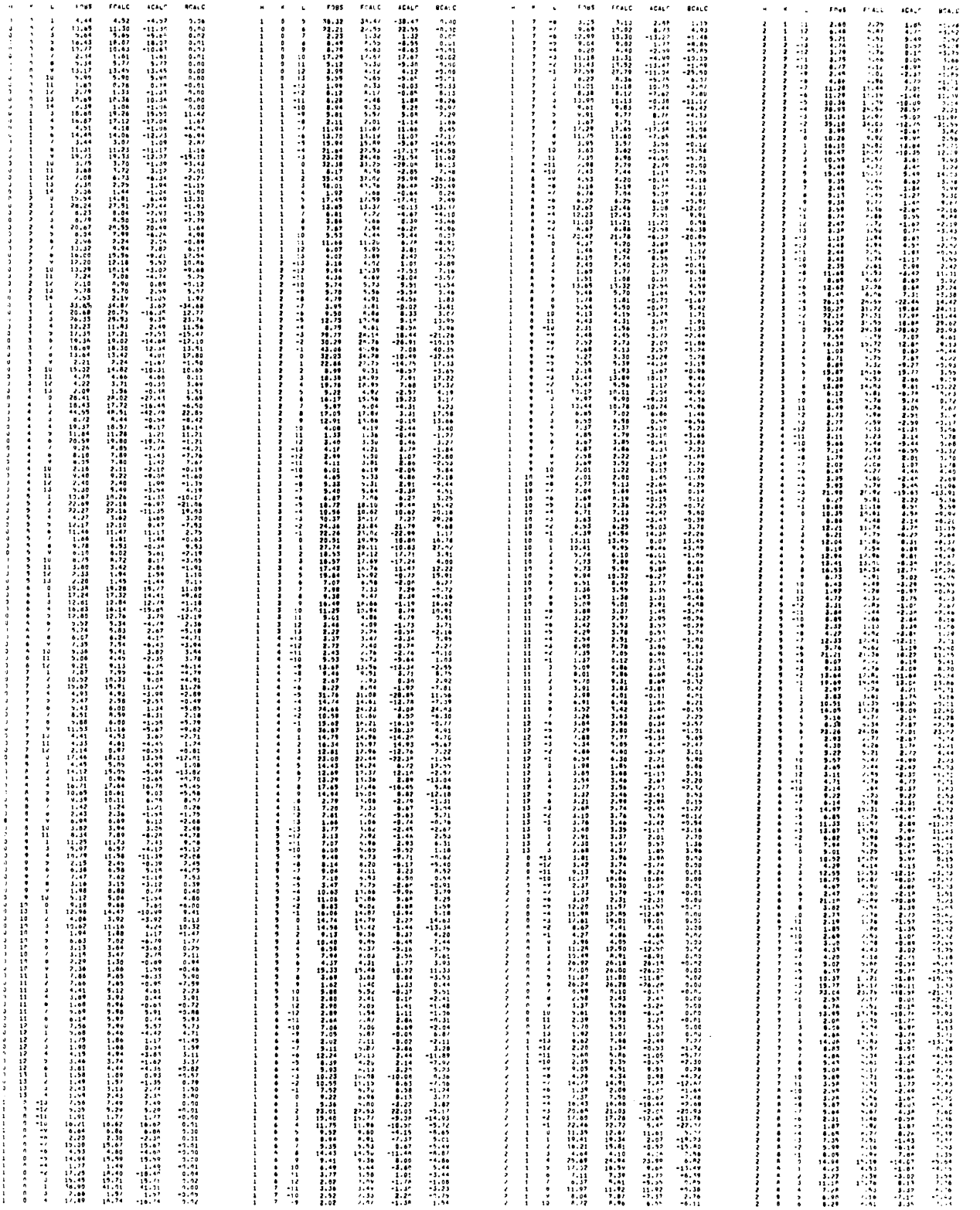

Table 4 (*cont.*)

H	K	L	FOBS	FCALC	ACALC	BCALC
2	4	7	3.28	3.03	1.09	2.49
2	8	8	2.66	2.94	2.67	1.22
2	8	9	4.65	4.49	-4.49	-0.09
2	8	10	7.27	6.91	-4.13	5.40
2	9	-9	2.49	2.93	1.75	2.74
2	9	-8	2.36	1.49	0.73	-1.19
2	9	-7	9.67	8.75	-6.15	-6.91
2	9	-6	1.09	1.02	0.08	1.02
2	9	-5	3.11	3.14	3.14	0.43
2	9	-4	2.45	2.28	2.12	-0.83
2	9	-3	11.73	10.54	-10.12	2.93
2	9	-2	9.74	9.49	7.61	-5.63
2	9	-1	12.49	12.55	9.19	8.55
2	9	0	4.44	4.52	1.04	4.40
2	9	1	12.67	13.26	-11.42	-6.74
2	9	2	9.47	8.88	-5.29	-7.16
2	9	3	8.40	8.32	3.62	-7.49
2	9	4	9.70	9.88	5.93	7.91
2	9	5	7.25	7.31	3.11	-7.40
2	9	6	4.60	4.99	1.25	-4.84
2	9	7	4.95	3.84	2.75	-2.60
2	9	8	1.85	2.33	2.29	0.26
2	9	9	7.48	7.29	7.27	0.59
2	10	-9	1.90	1.37	-0.21	-1.35
2	10	-7	2.78	2.74	2.54	1.33
2	10	-6	2.34	2.39	2.03	1.25
2	10	-5	4.18	4.90	-3.74	1.63
2	10	-4	3.78	3.80	1.56	-3.47
2	10	-3	2.67	2.22	0.47	-2.19

EXTINCTIONS AND UNOBSERVED REFLECTIONS.

In inosine, the sugar ring has C(3′) *endo* with a displacement of 0·63 Å from the plane through C(1′), C(2′), C(4′) and O(1′). The equation of this plane and the deviations from it are given in Table 7. The large displacement of C(3′) brings its attached oxygen atom O(3′) into a position 0·44 Å above [that is on the same side as C(5′)] the four atom sugar plane.

Shefter & Trueblood (1965) have defined the angles φ_{OO} and φ_{OC} which described the variable conformation of the C(5′)–O(5′) bond. The angle between the projected C(5′)–O(5′) and C(4′)–O(1′) bonds is φ_{OO} and the angle between the projected C(5′)–O(5′) and C(4′)–C(3′) bonds, φ_{OC}.

In inosine these angles are $\varphi_{OO} = 74{\cdot}7°$ and $\varphi_{OC} = 169{\cdot}0°$ so that C(5′)–O(5′) is *gauche* to C(4′)–O(1′) and *trans* to C(4′)–C(3′). This information is shown in Fig. 6. In the majority of nucleosides the conformation is *gauche–gauche*. The barium and disodium inosine 5′-monophosphate structures (Nagashima & Iitaka, 1968) both have a *gauche–gauche* conformation but with C(2′) *endo*. Other structures with the *gauche–trans* conformation do exist but with varying sugar-ring pucker. Some of these are: adenosine 3′-phosphate (Sundaralingam, 1966) which has C(3′) *endo*, 5-bromodeoxyuridine (Iball, Morgan & Wilson, 1966) with C(2′) *endo*, and thymidine (Young, Tollin & Wilson, 1969) and deoxyadenosine (Watson, Sutor & Tollin, 1965) both of which have C(3′) in the *exo* conformation.

The bond lengths and angles of the sugar moiety of inosine are included in Table 5 and appear in Fig. 7(*a*). The details of the conformation of the sugar residues in various well refined nucleoside and nucleotide struc-

Table 5. *Bond lengths, valence bond angles and standard deviations involving only non-hydrogen atoms*

N(1) –C(2)	1·355 (5) Å	C(2) –N(3)	1·308 (5) Å
N(3) –C(4)	1·365 (4)	C(4) –C(5)	1·374 (4)
C(4) –N(9)	1·372 (4)	C(5) –C(6)	1·433 (5)
C(6) –N(1)	1·397 (4)	C(6) –O(10)	1·233 (4)
C(5) –N(7)	1·371 (4)	N(7) –C(8)	1·307 (4)
C(8) –N(9)	1·372 (4)	N(9) –C(1′)	1·477 (4)
C(1′)–O(1′)	1·417 (4)	C(1′)–C(2′)	1·530 (4)
O(1′)–C(4′)	1·459 (4)	C(2′)–O(2′)	1·420 (4)
C(2′)–C(3′)	1·525 (4)	C(3′)–O(3′)	1·413 (4)
C(3′)–C(4′)	1·522 (4)	C(4′)–C(5′)	1·506 (4)
C(5′)–O(5′)	1·428 (4)		
C(5) –N(7) –C(8)	103·8 (0·3)°		
N(1) –C(2) –N(3)	124·6 (0·3)	C(2) –N(3) –C(4)	111·9 (0·3)°
N(3) –C(4) –C(5)	128·5 (0·3)	C(4) –C(5) –C(6)	118·3 (0·3)
C(5) –C(6) –N(1)	111·0 (0·3)	C(6) –N(1) –C(2)	125·4 (0·3)
N(1) –C(6) –O(10)	120·6 (0·3)	C(5) –C(6) –O(10)	128·4 (0·3)
C(6) –C(5) –N(7)	130·1 (0·3)	C(4) –C(5) –N(7)	111·6 (0·3)
N(7) –C(8) –N(9)	113·6 (0·3)	C(8) –N(9) –C(4)	105·7 (0·3)
N(9) –C(4) –C(5)	105·4 (0·3)	N(9) –C(4) –N(3)	126·1 (0·3)
C(8) –N(9) –C(1′)	128·1 (0·3)	C(4) –N(9) –C(1′)	126·0 (0·3)
N(9) –C(1′)–O(1′)	108·4 (0·2)	N(9) –C(1′)–C(2′)	111·5 (0·2)
O(1′)–C(1′)–C(2′)	106·8 (0·2)	C(1′)–O(1′)–C(4′)	109·6 (0·2)
C(1′)–C(2′)–O(2′)	109·6 (0·2)	C(1′)–C(2′)–C(3′)	100·6 (0·2)
O(2′)–C(2′)–C(3′)	107·5 (0·2)	C(2′)–C(3′)–C(4′)	101·5 (0·2)
C(2′)–C(3′)–O(3′)	115·9 (0·2)	O(3′)–C(3′)–C(4′)	114·2 (0·2)
C(3′)–C(4′)–O(1′)	104·0 (0·2)	C(3′)–C(4′)–C(5′)	114·1 (0·3)
O(1′)–C(4′)–C(5′)	110·0 (0·3)	C(4′)–C(5′)–O(5′)	112·0 (0·3)

Table 6. *Bond lengths, valence bond angles and standard deviations involving hydrogen atoms*

N(1) –H(1)	0·84 (5) Å	C(2) –H(2)	0·98 (5) Å
C(8) –H(8)	0·99 (5)	C(1′)–H(1′)	1·04 (4)
C(2′)–H(2′)	1·01 (4)	C(3′)–H(3′)	0·95 (4)
C(4′)–H(4′)	0·96 (4)	C(5′)–H(5′)	0·97 (4)
C(5′)–H(5″)	0·99 (5)	O(2′)–H(2″)	0·88 (5)
O(3′)–H(3″)	0·90 (5)	O(5′)–H(5‴)	0·88 (5)
H(1)—N(1) –C(6)	112·3 (3·3)°	H(1) —N(1) –C(2)	122·3 (3·3)°
H(2)—C(2) –N(1)	120·3 (3·0)	H(2) —C(2) –N(3)	115·0 (3·0)
H(8)—C(8) –N(7)	123·4 (2·6)	H(8) —C(8) –N(9)	123·0 (2·6)
H(1′) –C(1′)–N(9)	113·2 (2·4)	H(1′)—C(1′)–O(1′)	105·3 (2·4)
H(1′) –C(1′)–C(2′)	111·2 (2·4)	N(2′)—C(2′)–C(1′)	109·4 (2·4)
H(2′) –C(2′)–C(3′)	115·0 (2·4)	H(2′)—C(2′)–O(2′)	114·0 (2·4)
H(3′) –C(3′)–C(2′)	110·4 (2·5)	H(3′)—C(3′)–C(4′)	107·7 (2·5)
H(3′) –C(3′)–O(3′)	106·9 (2·5)	H(4′)—C(4′)–C(3′)	109·3 (2·5)
H(4′) –C(4′)–C(5′)	110·7 (2·5)	H(4′)—C(4′)–O(1′)	108·4 (2·5)
H(5′) –C(5′)–C(4′)	112·9 (2·6)	H(5′)—C(5′)–O(5′)	110·2 (2·6)
H(5′) –C(5′)–H(5″)	113·0 (3·7)	H(5″) –C(5′)–C(4′)	104·0 (2·7)
H(5″)–C(5′)–O(5′)	104·4 (2·7)	H(2″) –O(2′)–C(2′)	110·9 (3·2)
H(3″)–O(3′)–C(3′)	105·5 (3·2)	H(5‴)–O(5′)–C(5′)	108·8 (3·3)

tures have been examined by Sundaralingam & Jensen (1965). The expected bond lengths and angles predicted by these authors for a C(3′) *endo* configuration appear in Fig. 7(*b*) for comparison with the inosine structure.

Only the C(1′)–C(2′) bond distance is significantly different from the predicted values of Sundaralingam & Jensen. As in other structures the C(1′)–O(1′) bond is significantly less than the C(4′)–O(1′) bond, the difference being 0·042 Å, but the C(3′)–O(3′) bond involving the out of plane C(3′) atom is not significantly less than the C(2′)–O(2′) bond involving the in-plane C(2′) atom.

Sundaralingam & Jensen analysed the three types of angles in the sugar ring, CCC, CCO and COC, and produced average values of 101·7, 105·8 and 109·3° respectively. In inosine the values are 101·1, 105·4 and 109·6° respectively. The angles around the in-plane C(2′) atom are all significantly different from their predicted values, in particular the angle C(2′)–C(3′)–C(4′) at the out of plane C(3′) atom is 0·9° greater than that at the in-plane C(2′) atom, C(1′)–C(2′)–C(3′). Sundaralingam & Jensen suggest that the former angle should invariably be about 1·6° less than the latter. It seems probable that strong intermolecular packing forces involving both O(2′) and O(3′) cause inosine to differ.

In agreement with other structures the exocyclic angles N(9)–C(1′)–O(1′) and C(5′)–C(4′)–O(1′) are both significantly greater than the angles N(9)–C(1′)–C(2′) and C(5′)–C(4′)–C(3′) respectively.

Conformation of the molecule

The dihedral angle between the sugar base plane and that of the purine is 71°. The relative orientation of the sugar and purine groups can be described by the torsion angle φ_{CN} defined by Donohue & Trueblood (1960) as the angle formed by the trace of the plane of the base with the projection of the C(1′)–O(1′) bond of the sugar ring when viewed along the C(1′)–N(9) glycosidic bond. If C(1′) is displaced from the base plane then the projection of the C(1′)–N(9) bond in the plane is used (Haschemeyer & Rich, 1967). In inosine the conformation is *anti* with $\varphi_{CN} = -10{\cdot}6°$. This value agrees with nucleotide and nucleosides previously studied, all of which with the exception of deoxyguanosine (Haschemeyer & Sobell, 1965), are in the *anti* conformation but with large variation in the φ_{CN} angle. A view of the conformation of inosine is shown in Fig. 8.

Molecular packing and hydrogen bonds

The molecular packing of purine and pyrimidine nucleotides and nucleosides in their crystal structures is of importance in the understanding of the structure of nucleic acids. It has been observed that purine-based nucleosides tend to pack in parallel planes of separation of about 3·5 Å. In inosine the perpendicular distance between the purine planes of molecules separated by an *a* axis lattice translation is 3·5 Å. The

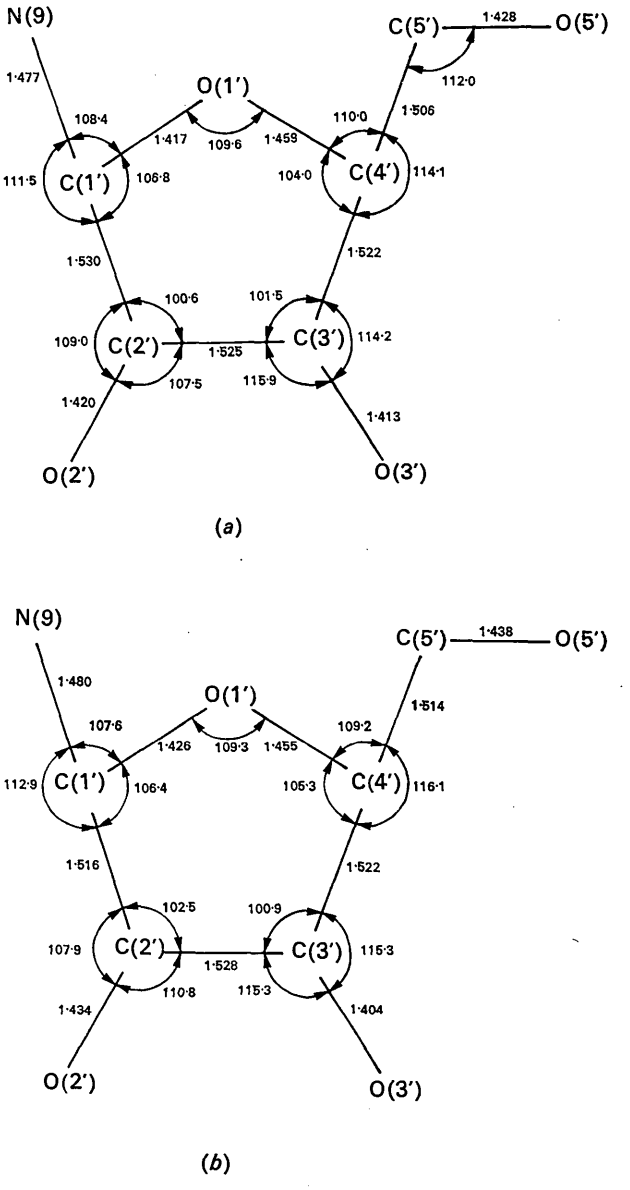

Fig. 7. The bond lengths and valence bond angles of the ribose sugar (*a*) of inosine and (*b*) of the model structure proposed by Sundaralingam & Jensen (1965) for C(3′) *endo* conformation.

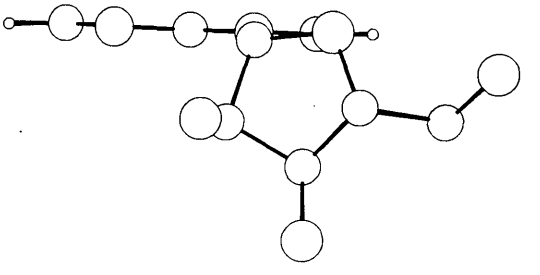

Fig. 8. The conformation of the inosine molecule viewed along the projection of the C(1′)–N(9) bond in the plane of the purine base.

arrangement of the molecules in the unit cell in projection down the *a* axis is shown in Fig. 9.

This structure of inosine is very tightly bonded, and this fact is reflected in the rather high calculated density of 1·61 g.cm^{-3}. All possible hydrogen bonds are formed and there are a number of close intermolecular contacts, some involving hydrogen atoms attached to carbon atoms of the purine residue. Table 9 lists the close contacts involving hydrogen atoms. All other intermolecular contacts, with two exceptions, are greater than 3·35 Å. The two exceptions are the contact between N(7) and O(10) at $1-x$, $-\frac{1}{2}+y$, $1-z$, and that between O(3′) and C(2) at $-x$, $-\frac{1}{2}+y$, $-z$, which have distances of 3·14 and 3·17 Å respectively. How-

Table 7. *Deviations of non-hydrogen atoms from the least-squares plane through the purine base, and through the sugar*

(I) Purine plane: $0{\cdot}7237x-0{\cdot}4684y-0{\cdot}5068z+1{\cdot}2786=0$
(II) Sugar plane: $0{\cdot}7363x-0{\cdot}2892y+0{\cdot}6329z-0{\cdot}9278=0$

	Deviations from (I)	Δ/σ		Deviations from (II)	Δ/σ
N(1)	0·020 Å	6·4	C(1′)	0·042 Å	13·5
C(2)	0·020	5·2	C(2′)	−0·025	8·0
N(3)	−0·010	3·2	C(4′)	0·026	8·4
C(4)	−0·008	2·6	O(1′)	−0·043	20·7
C(5)	−0·010	3·2	C(3′)*	0·626	
C(6)	−0·045	14·0	O(3′)*	0·440	
N(7)	0·023	7·4	O(2′)*	−1·379	
C(8)	0·004	1·2	C(5′)*	0·845	
N(9)	−0·014	5·4	N(9)*	1·320	
O(10)*	−0·161	62·3			
C(1′)*	−0·149	47·8			
H(1)*	0·040	0·8			
H(2)*	0·018	0·4			
H(8)*	0·032	0·6			

* Omitted from least-squares plane calculation.

Table 8. *Comparison of bond lengths and angles of the purine base in related compounds*

	MT	IW	SSJ	WST	WSM
N(1)-C(2)	1·355 Å	1·374 Å	1·379 Å	1·317 Å	1·349 Å
C(2)-N(3)	1·308	1·318	1·334	1·326	1·332
N(3)-C(4)	1·365	1·345	1·351	1·346	1·336
C(4)-C(5)	1·374	1·377	1·383	1·392	1·402
C(5)-C(6)	1·433	1·414	1·430	1·414	1·389
C(6)-N(1)	1·397	1·390	1·385	1·336	1·330
C(6)-O(10)	1·233	1·237	1·224	—	—
C(5)-N(7)	1·371	1·378	1·361	1·375	1·374
N(7)-C(8)	1·307	1·322	1·303	1·307	1·332
C(8)-N(9)	1·372	1·335	1·359	1·361	1·312
N(9)-C(4)	1·372	1·375	1·350	1·369	1·374
N(9)-C(1′)	1·477	—	—	1·472	—
N(1)-C(2)-N(3)	124·6°	123·4°	123·2°	128·8°	128·2°
C(2)-N(3)-C(4)	111·9	112·8	112·4	111·0	113·2
N(3)-C(4)-C(5)	128·5	127·6	128·0	126·9	123·5
C(4)-C(5)-C(6)	118·3	119·9	119·6	115·4	118·2
C(5)-C(6)-N(1)	110·0	110·8	110·7	118·1	119·0
C(6)-N(1)-C(2)	125·4	125·6	126·2	119·8	118·0
C(5)-C(6)-O(10)	128·4	128·9	127·8	—	—
N(1)-C(6)-O(10)	120·6	120·3	121·6	—	—
C(6)-C(5)-N(7)	130·1	132·7	131·1	133·9	136·6
C(4)-C(5)-N(7)	111·6	107·4	109·3	110·7	105·3
C(5)-N(7)-C(8)	103·8	108·2	108·1	104·4	106·4
N(7)-C(8)-N(9)	113·6	109·6	108·3	113·2	114·6
C(8)-N(9)-C(4)	105·7	108·6	110·4	106·8	104·2
C(8)-N(9)-C(1′)	128·1	—	—	129·6	—
N(3)-C(4)-N(9)	126·1	126·3	128·0	128·3	126·9
C(4)-C(5)-N(9)	105·4	106·2	104·0	104·8	109·6
C(4)-N(9)-C(1′)	126·0	—	—	123·2	—

MT, inosine (this paper).
IW, guanine hydrochloride dihydrate (Iball & Wilson, 1965).
SSJ, 8-azaguanine monohydrate (Sletten, Sletten & Jensen, 1968).
WST, deoxyadenosine (Watson, Sutor & Tollin, 1965).
WSM, purine (Watson, Sweet & Marsh, 1965).

ever, atom N(7) is next to atom C(8) and C(2) is next to N(3) and these atoms, C(8) and N(3), have close contacts involving hydrogen atoms with O(10) and O(3′) respectively.

Of the eight close contacts involving hydrogen atoms listed in Table 9, the first four are normal hydrogen bonds. Of these the O(5′)–O(10) bond is almost linear while the other three have hydrogen–donor–acceptor angles of about 21°. Such non-linearity of hydrogen bonds is common in structures of nucleosides (Donohue, 1968).

Three of the remaining close contacts of Table 9 form a particularly interesting system involving the atoms O(2′), C(2), O(3′) and O(5′). The positions of these atoms and the hydrogen atoms H(2″) and H(2) are shown in Fig. 10. The possibility of the existence of C–H···O hydrogen bonds has received some discussion (Sutor, 1962; Donohue, 1968; Hamilton & Ibers, 1968). Despite the short C···O distances of 3·196 and 3·087 Å for the contacts C(2)–H(2)···O(2′) and C(2)–H(2)···O(5′), the H···O distances of 2·37 and 2·34 Å may well exclude the possibility of these contacts representing C–H···O hydrogen bonds. However, it is tempting to suggest that the H···O distances are rather long because this hydrogen atom is forming a bifurcated hydrogen bond with the atoms O(2′) and O(5′). There is some further evidence to support the existence of these bonds. In inosine there are four hydrogen donor sites, N(1), O(2′), O(3′) and O(5′), but there are six possible acceptor sites N(3), N(7), O(10),

Table 9. *Close intermolecular contacts involving hydrogen and non-hydrogen atoms*

A–H···*B*	*B* equipoint	*A*–H (±0·05)	*A*–*B* (±0·004)	H···*B* (±0·05)	H–*A*–*B* (±3°)
N(1)–H(1) ····N(7)	($1-x$ $\frac{1}{2}+y$ $1-z$)	0·84 Å	2·923 Å	2·15 Å	20°
O(3′)–H(3′) ···N(3)	($-x$ $-\frac{1}{2}+y$ $-z$)	0·90	2·877	2·06	21
O(5′)–H(5‴)···O(10)	($-x$ $-\frac{1}{2}+y$ $1-z$)	0·88	2·704	1·83	5½
O(2′)–H(2″) ···O(3′)	($-x$ $\frac{1}{2}+y$ $-z$)	0·88	2·878	2·10	23
O(2′)–H(2″) ···O(5′)	($-1-x$ $\frac{1}{2}+y$ $-z$)	0·88	3·068	2·56	48
C(2) –H(2) ····O(2′)	($-x$ $\frac{1}{2}+y$ $-z$)	0·98	3·196	2·37	27½
C(2) –H(2) ····O(5)	($1+x$ $1+y$ z)	0·98	3·087	2·34	34
C(8) –H(8) ····O(10)	($1-x$ $-\frac{1}{2}+y$ $1-z$)	0·99	2·997	2·57	56

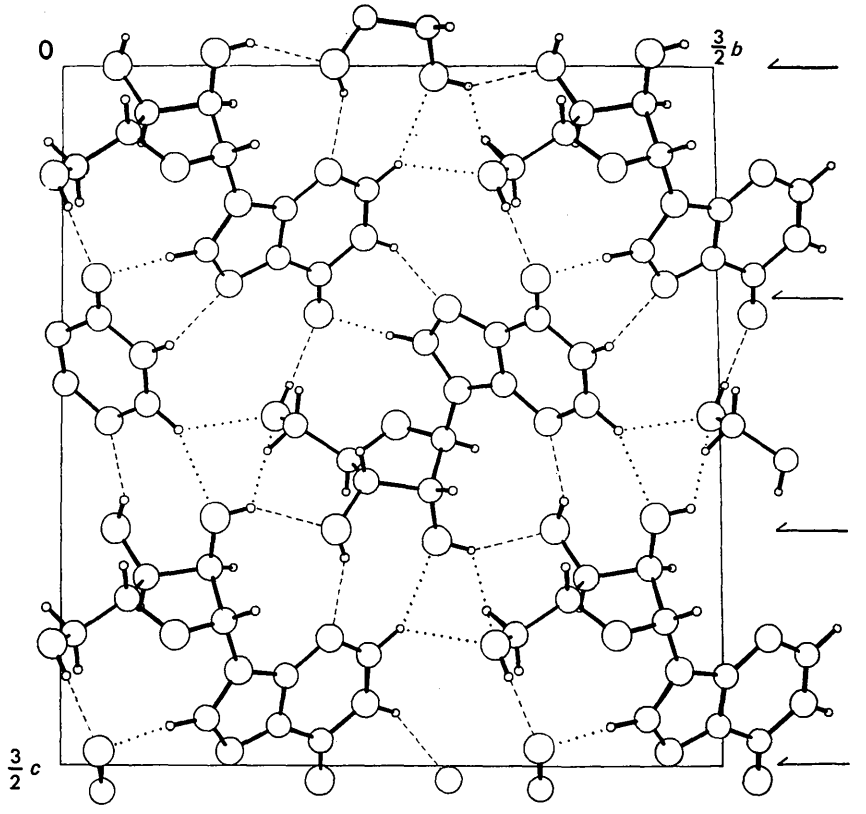

Fig. 9. The arrangement of the molecules of inosine in the unit cell viewed as a projection onto the b,c^* plane. Hydrogen bonds are shown as broken lines and the other close contacts listed in Table 8 are shown as dotted lines.

O(2′), O(3′) and O(5′). Thus the inclusion of the hydrogen atom H(2) completes the hydrogen bonding scheme. The O(2′) to O(5′) distance is short, presumably because of the very strong bonding between the molecules and not because a hydrogen bond is being formed. The remaining close contact, that between C(8) and O(10) is clearly not a C–H···O hydrogen bond since both the H···O distance and the bond angle are large.

The authors wish to thank the Computing Laboratories of the Universities of St Andrews and Dundee for making computer facilities available and Mr P. E. Adamson for the use of the molecular drawing program. Thanks are also due to Dr J. Iball for assistance with the collection of the data from the four-circle diffractometer, to Mr J. Low for assistance with the linear diffractometer, to Dr H. R. Wilson and Dr D. W. Young for helpful discussions, and to Miss P. Hendry and Mrs E. M. Millar for technical and clerical assistance. One of us (A.R.I.M.) wishes to acknowledge the receipt of a Science Research Council Studentship.

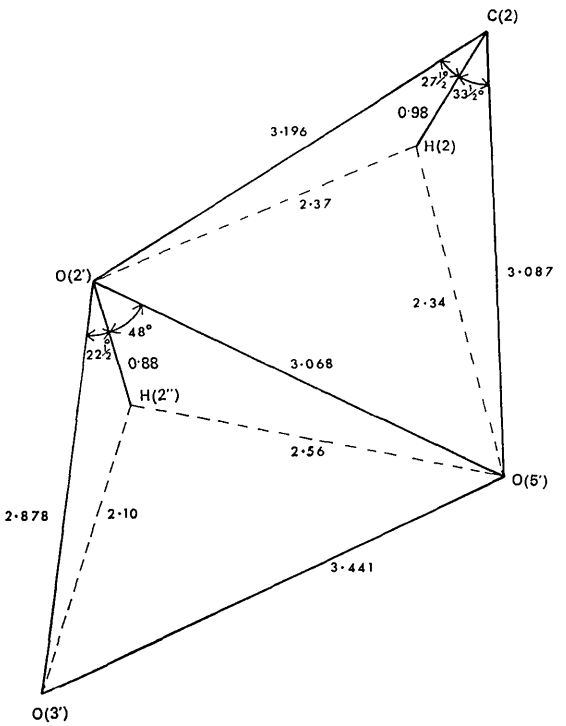

Fig. 10. The conformation of the close contacts involving atoms C(2) and C(2′).

References

BUGG, C. E., THEWALT, U. & MARSH, R. E. (1968). *Biochem. Biophys. Res. Comm.* **33**, 436.

CRICK, F. H. C. (1966). *J. Mol. Biol.* **19**, 548.

DONOHUE, J. (1968). *Structural Chemistry and Molecular Biology*. San Francisco: Freeman.

DONOHUE, J. & TRUEBLOOD, K. N. (1960). *J. Mol. Biol.* **2**, 363.

HAMILTON, W. C. & IBERS, J. A. (1968). *Hydrogen Bonding in Solids*. New York: Benjamin.

HARRIS, D. R. & MACINTYRE, W. M. (1964). *Biophys. J.* **4**, 203.

HASCHEMEYER, A. E. V. & RICH, A. (1967). *J. Mol. Biol.* **27**, 369.

HASCHEMEYER, A. E. V. & SOBELL, H. M. (1965). *Acta Cryst.* **19**, 125.

HODGKIN, D. C., LINDSEY, J., SPARKS, R. A., TRUEBLOOD, K. N. & WHITE, J. G. (1962). *Proc. Roy. Soc.* A**266**, 494.

IBALL, J., MORGAN, C. H. & WILSON, H. R. (1966). *Proc. Roy. Soc.* A**295**, 320.

IBALL, J., MORGAN, C. H. & WILSON, H. R. (1968). *Proc. Roy. Soc.* A**302**, 225.

IBALL, J. & WILSON, H. R. (1965). *Proc. Roy. Soc.* A**288**, 418.

NAGASHIMA, N. & IITAKA, Y. (1968). *Acta Cryst.* B**24**, 1136.

MUNNS, A. R. I., TOLLIN, P., WILSON, A. J. C. & YOUNG D. W. (1970). *Acta Cryst.* B**26**, 1114.

ROSSMANN, M. G. & BLOW, D. M. (1962). *Acta Cryst.* **15**, 24.

SHEFTER, E. & TRUEBLOOD, K. N. (1965). *Acta Cryst.* **18**, 1067.

SLETTEN, J., SLETTEN, E. & JENSEN, L. H. (1968). *Acta Cryst.* B**24**, 1692.

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

SUNDARALINGAM, M. (1966).*Acta Cryst.* **21**, 495.

SUNDARALINGAM, M. & JENSEN, L. H. (1965). *J. Mol. Biol.* **13**, 930.

SUTOR, D. J. (1962). *Nature, Lond.* **195**, 68.

TOLLIN, P. (1966). *Acta Cryst.* **21**, 613.

TOLLIN, P. & COCHRAN, W. (1964). *Acta Cryst.* **17**, 1332.

TOLLIN, P., MAIN, P. & ROSSMANN, M. G. (1966). *Acta Cryst.* **20**, 404.

TOLLIN, P. & MUNNS, A. R. I. (1969). *Nature, Lond.* **222**, 1170.

TOLLIN, P & ROSSMANN, M. G. (1966). *Acta Cryst.* **21**, 872.

WATSON, D. C., SUTOR, D. J. & TOLLIN, P. (1965). *Acta Cryst.* **19**, 111.

WATSON, D. G., SWEET, R. M. & MARSH, R. E. (1965). *Acta Cryst.* **19**, 573.

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

WOESE, C. R. (1967). *The Genetic Code*. New York: Harper & Row.

WUNDERLICH, J. A. (1965). *Acta Cryst.* **19**, 200.

YOUNG, D. W., TOLLIN, P. & WILSON, H. R. (1969). *Acta Cryst.* B**25**, 1423.