The Crystal and Molecular Structure of 6-Thioguanine[1]

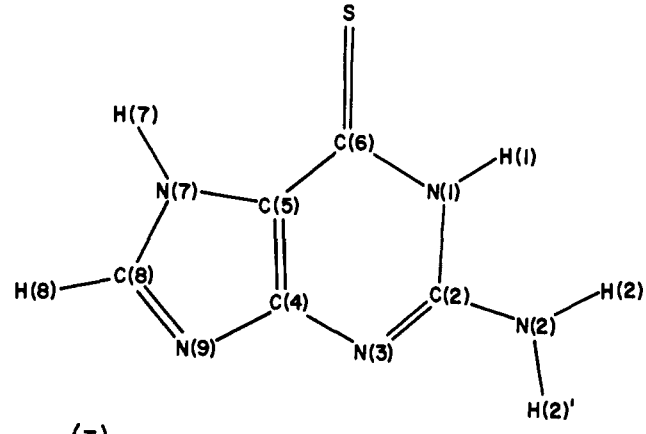
Charles E. Bugg* and Ulf Thewalt

Contribution from the Institute of Dental Research and Department of Chemistry, University of Alabama in Birmingham, Birmingham, Alabama 35233, and Mineralogisches Institut, Universität Marburg, 355 Marburg/Lahn, West Germany. *Received April 4, 1970*

Abstract: The crystal structure of the antitumor guanine analog, 6-thioguanine ($C_5H_5N_5S$), was determined from three-dimensional diffractometer data and refined by block-diagonal least squares to an R index of 0.041. In contrast to the crystal structure of guanine, the thioguanine molecule assumes the tautomer form with a hydrogen atom bonded to atom N-7, rather than N-9, of the purine ring. The base stacking is different from that found in the crystal structure of guanine, but is similar to the purine stacking in crystals of guanine hydrochloride and 9-methylguanine hydrobromide, two protonated derivatives which also have hydrogen atoms at the N-7 position. The sulfur substituent is an acceptor in the formation of hydrogen bonds with atoms N-7 and N-2 of symmetry related molecules; both of these hydrogen bonds are approximately 3.3 Å in length. This is considerably different from the lengths of 2.8–3.0 Å for hydrogen bonds between carbonyl oxygen atoms and nitrogen donors as found in the crystal structures of guanine and related compounds, as well as in the guanine–cytosine base pairs of nucleic acids. It is suggested that guanine–cytosine and thioguanine–cytosine hydrogen bonded base pairs might have different dimensions, and that distortions in hydrogen bonding caused by the substitution of thioguanine for guanine in nucleic acids may account, in part, for the biological properties of thioguanine.

Thioguanine (2-amino-6-mercaptopurine, $C_5H_5N_5S$; I) is a metabolic inhibitor[2] with antitumor activity.[3-14] Several studies have indicated that the biological activity of thioguanine is due, in part, to its incorporation in nucleic acids,[15-19] probably by replacement of guanine. Very little is known about the factors which account for the action of thioguanine once it enters nucleic acids; however, it seems likely that this purine analog might interfere with the normal interactions among the bases of nucleic acids. Interactions among purine and pyrimidine bases are apparently of two principal types: hydrogen bonding[20] and base stacking.[21] Both types of interactions appear to play major roles in establishing and maintaining the physical and biological properties of nucleic acids. Alterations in these interactions, resulting from replacement of guanine by thioguanine, could possibly produce faulty nucleic acids, thus accounting for the biological activity of thioguanine.

We have determined the crystal structure of thioguanine, in order to compare the solid state hydrogen bonding and base stacking interactions with those found for guanine.[22,23]



(I)

Experimental Section

Large, pale yellow needles of thioguanine were obtained by slowly cooling a hot, saturated aqueous solution. Weissenberg and oscillation photographs showed the Laue symmetry to be mmm (D_{2h}). The needle axis was chosen as the c axis. The space group is $P2_12_12_1$ as indicated by the systematic absence of reflections $h00$ with h odd, $0k0$ with k odd, and $00l$ with l odd.

A fragment of length 0.15 mm and cross section dimensions 0.08 × 0.06 mm was sliced from a needle crystal and was mounted on a Picker FACS-1 diffractometer with the c axis parallel to the ϕ axis of the diffractometer. The 2θ values for a number of reflections were measured; the unit-cell parameters obtained from a least-squares analysis of these measurements are a = 16.313 (2), b = 9.850 (1), and c = 4.239 (1) Å. The density calculated by

* To whom inquiries should be addressed: Institute of Dental Research, University of Alabama in Birmingham, Birmingham, Alabama 35233.

(1) Supported, in part, by USPHS Research Grant No. DE-02670 from the National Institute of Dental Research.

(2) G. H. Hitchings and G. B. Elion, "Metabolic Inhibitors," Vol. 1, J. H. Quastel and R. M. Hochster, Ed., Academic Press, New York, N. Y., 1963, pp 215–237.
(3) A. M. Williams and E. Joranger, *Brit. J. Cancer*, **15**, 342 (1961).
(4) J. F. Henderson and I. G. Junga, *Cancer Res.*, **21**, 7 (1961).
(5) A. C. Sartorelli, E. J. Schoolar, Jr., and P. F. Kruse, Jr., *Proc. Soc. Exp. Biol. Med.*, **104**, 266 (1960).
(6) A. C. Sartorelli and G. A. LePage, *Cancer Res.*, **18**, 938 (1958).
(7) A. C. Sartorelli and B. A. Booth, *ibid.*, **27**, 1614 (1967).
(8) D. H. Adams, *ibid.*, **24**, 250 (1964).
(9) E. R. Lefkowitz, W. A. Creasey, P. Calabresi, and A. C. Sartorelli, *ibid.*, **25**, 1207 (1965).
(10) R. W. Sidwell, S. M. Sellers, G. J. Dixon, and F. M. Schabel, Jr., *ibid.*, **28**, 35 (1968).
(11) G. A. LePage and M. Jones, *ibid.*, **21**, 642 (1961).
(12) C. W. Noell and R. K. Robins, *J. Med. Pharm. Chem.*, **5**, 1074 (1962).
(13) B. A. Booth and A. C. Sartorelli, *Science*, **138**, 518 (1962).
(14) G. A. LePage and N. Howard, *Cancer Res.*, **23**, 622 (1963).
(15) G. A. LePage and I. G. Junga, *ibid.*, **23**, 739 (1963).
(16) G. A. LePage, *ibid.*, **20**, 403 (1960).
(17) G. A. LePage and M. Jones, *ibid.*, **21**, 1590 (1961).
(18) D. H. Adams, *Biochem. J.*, **89**, 240 (1963).
(19) H. S. Kaplan, D. C. Smith, and P. Tomlin, *Nature (London)*, **190**, 794 (1961).
(20) K. Hoogsteen, "Molecular Associations in Biology," B. Pullman, Ed., Academic Press, New York, N. Y., 1968, pp 21–38.
(21) See P. O. P. Ts'o, in ref 20, pp 39–75.
(22) C. E. Bugg, U. Thewalt, and R. E. Marsh, *Biochem. Biophys. Res. Commun.*, **33**, 436 (1968).
(23) U. Thewalt, C. E. Bugg, and R. E. Marsh, submitted for publication.

Table I. The Final Heavy Atom Parameters and Their Estimated Standard Deviations[a]

Atom	x	y	z	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
N-1	5407 (2)	5401 (3)	4103 (8)	29 (1)	51 (3)	484 (21)	5 (3)	30 (10)	−21 (14)
C-2	5096 (2)	6541 (3)	2734 (12)	26 (1)	67 (3)	410 (22)	4 (4)	13 (12)	−42 (18)
N-2	4442 (2)	6382 (3)	940 (10)	41 (1)	64 (3)	739 (26)	1 (4)	−102 (12)	31 (16)
N-3	5416 (2)	7773 (3)	3033 (8)	27 (1)	54 (3)	523 (21)	2 (3)	16 (10)	3 (16)
C-4	6088 (2)	7804 (3)	4909 (10)	28 (1)	63 (3)	418 (23)	25 (4)	36 (10)	51 (19)
C-5	6429 (2)	6674 (4)	6396 (9)	28 (1)	66 (3)	472 (24)	5 (4)	8 (12)	−51 (17)
C-6	6104 (2)	5392 (3)	6072 (10)	28 (1)	70 (4)	380 (23)	19 (4)	67 (11)	−10 (17)
N-7	7091 (2)	7134 (3)	8082 (9)	25 (1)	84 (3)	509 (21)	10 (3)	−34 (9)	−49 (18)
C-8	7130 (2)	8477 (4)	7550 (12)	36 (2)	78 (4)	651 (30)	7 (4)	−32 (16)	−26 (24)
N-9	6536 (2)	8929 (3)	5654 (9)	32 (1)	66 (3)	644 (22)	4 (4)	−25 (10)	17 (18)
S	6444 (1)	3957 (1)	7825 (2)	29 (0)	54 (1)	406 (5)	10 (1)	−9 (3)	27 (4)

[a] The values have been multiplied by 10^4. The temperature factors are in the form $T = \exp(-\beta_{11}h^2 - \beta_{22}k^2 - \beta_{33}l^2 - \beta_{12}hk - \beta_{13}hl - \beta_{23}kl)$.

assuming that there are four thioguanine molecules per unit cell is 1.630 g cm^{-3}; the density measured by flotation is 1.64 g cm^{-3}.

Intensity data were collected with the diffractometer, using a scintillation counter, nickel-filtered copper radiation, and a θ–2θ scanning technique. Measurements were made for the 690 reflections in the range $4° \leq 2\theta \leq 128°$; this represents approximately 71% of the unique reflections in the copper sphere of reflection. The intensity values were assigned variances, $\sigma^2(I)$, according to the statistics of the scan and background counts plus an additional term $(0.03S)^2$, S being the scan counts. The intensities and their standard deviations were corrected for Lorentz and polarization factors and placed on an approximately absolute scale by means of a Wilson plot. No reflections were discarded or considered to be unobservable. No corrections for absorption effects were applied; however, considering the small size of the crystal used and the magnitude of the linear absorption coefficient (μ = 36 cm^{-1}), it is expected that absorption errors should have little effect on any parameters other than the temperature factors.

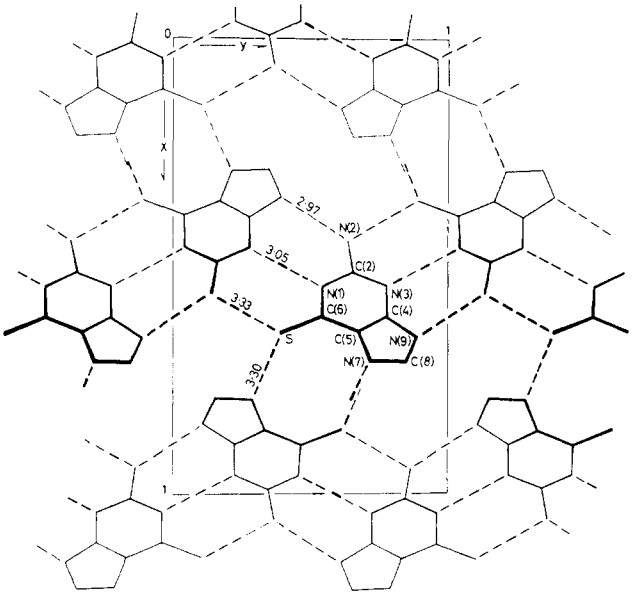


Figure 1. The structure of thioguanine viewed down the c axis (dashed lines represent hydrogen bonds; the hydrogen bond distances (Å) are shown).

A suitable trial structure was readily obtained by the heavy atom method: coordinates for the sulfur atom were determined from a sharpened Patterson map, and the other nonhydrogen atoms were located in a Fourier map calculated with phases based on the sulfur atom. Three-dimensional refinement of the trial structure was carried out by block-diagonal least squares. The quantity minimized was $\Sigma w(F_o - (1/k)F_c)^2$, where k is a scale factor and the weight w is equal to $(2F_o/\sigma(F_o^2))^2$. Initially, the heavy atom positional and anisotropic temperature factors refined to a R index $(\Sigma||F_o| - |F_c||/\Sigma|F_o|)$ of 0.06. At this stage a difference Fourier map clearly revealed the positions of the five hydrogen atoms. Finally, all positional parameters, along with anisotropic temperature factors for the heavy atoms and isotropic temperature factors for the hydrogen atoms, were refined. Atomic scattering factors for the nonhydrogen atoms were obtained from the "International Tables for X-Ray Crystallography,"[24] and those for the hydrogen atoms were from Stewart, Davidson, and Simpson.[25]

Table II. The Final Hydrogen Atom Parameters and Their Estimated Standard Deviations[a]

Atom	x	y	z	B, Å^2
H-1	516 (2)	454 (3)	363 (9)	3.4 (0.9)
H-2	411 (3)	543 (5)	081 (14)	8.9 (1.5)
H-2′	421 (2)	711 (4)	017 (11)	5.4 (1.1)
H-7	751 (3)	650 (5)	946 (13)	9.5 (1.6)
H-8	758 (2)	900 (3)	821 (9)	4.1 (0.9)

[a] The positional parameters have been multiplied by 10^3.

The final R index is 0.041. In the last cycle of refinement, no parameter shift exceeded one-fifth of its indicated standard deviation. The goodness-of-fit, $(\Sigma(1/\sigma^2(F_o^2))(F_o^2 - F_c^2/k^2)^2/(m - s))^{1/2}$, where m is the number of reflections used and s is the number of parameters refined, is 1.48. The average estimated standard deviations in the positional coordinates of the heavy atoms are 0.001–0.004 Å and those for the hydrogen atoms are 0.04–0.05 Å; this corresponds to esd's of about 0.005 Å for bond lengths involving only heavy atoms and 0.04 Å for bond lengths involving hydrogen atoms. The esd's in bond angles are about 0.3° for angles involving only heavy atoms and 2° for angles involving hydrogen atoms. At the conclusion of the refinement, a three-dimensional electron-density difference map was calculated with only the heavy atom contributions included in the values for the calculated structure factors. This map showed regions of electron density in excess of 0.4 electron/Å^3 at all calculated hydrogen atom positions; no other peaks or troughs exceeding 0.3 electron/Å^3 were apparent in this map.

Results

The final heavy atom parameters and their standard deviations are listed in Table I; the final hydrogen atom parameters and their standard deviations are listed in Table II.

In the crystal structure, thioguanine exists as the tautomer represented by structural formula I, with a hydrogen atom bonded to atom N-7, rather than to atom N-9 as is found in the crystal structure of guanine.[22,23]

Figure 1 shows the structure viewed down the c axis; Figure 2 shows the structure projected down the b axis. The hydrogen bond lengths are included in Figure 1, and additional hydrogen bond data are compiled in

(24) "International Tables for X-Ray Crystallography," Vol. III, Kynoch Press, Birmingham, England, 1962, p 202.

(25) R. F. Stewart, E. R. Davidson, and W. T. Simpson, *J. Chem. Phys.*, 42, 3175 (1965).

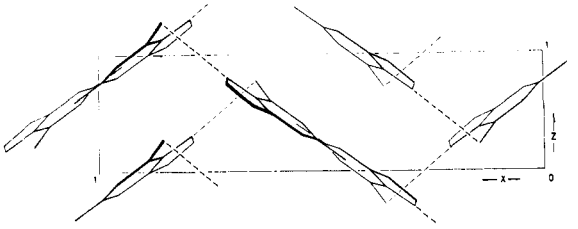


Figure 2. The structure of thioguanine viewed down the *b* axis (dashed lines represent hydrogen bonds between the sulfur substituent and atom N-7).

rection. Within these ribbons each base is joined to the two adjacent bases by a total of six hydrogen bonds. The dihedral angle between the planes of adjacent molecules in these ribbons is 19°. The ribbons are stacked in the *c* direction, with adjacent molecules along *c* being separated by an interplanar spacing of 3.37 Å; the molecular stacking along *c*, as viewed perpendicular to the thioguanine plane, is shown in Figure 3a. As viewed down the *b* axis (Figure 2) the ribbons form a herringbone pattern, with adjacent ribbons joined by hydrogen bonds between atom N-7 and the sulfur substituent.

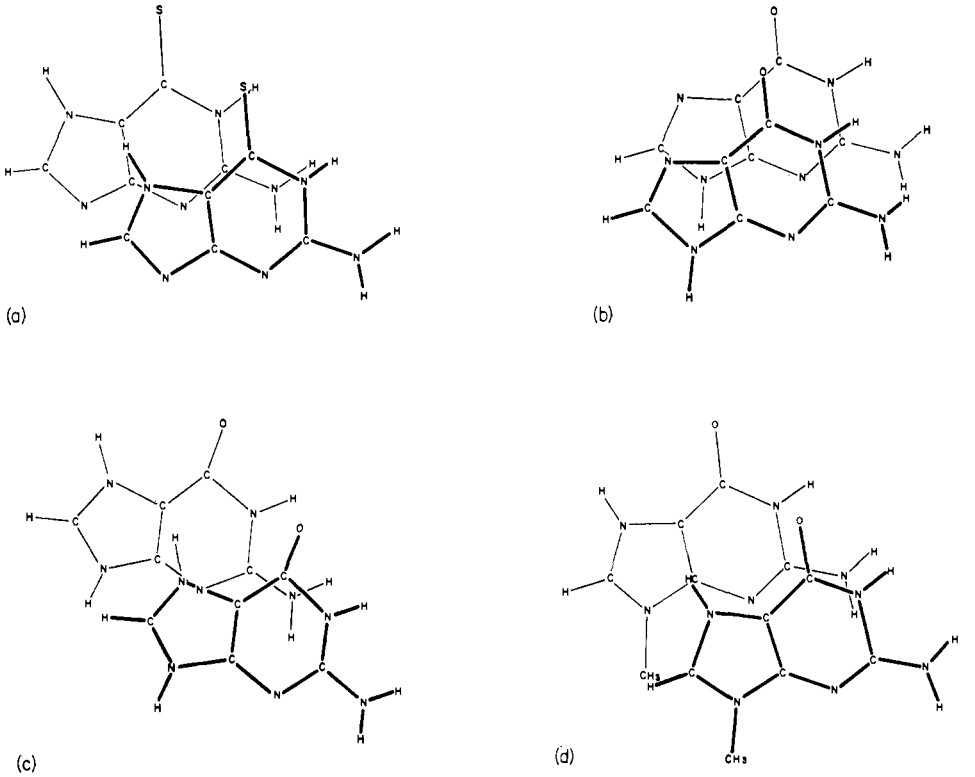


Figure 3. Base stacking in thioguanine and related compounds (all views are perpendicular to the least-squares planes of the purine rings): (a) thioguanine (interplanar spacing = 3.37 Å); (b) guanine[22,23] (interplanar spacing = 3.30 Å); (c) guanine hydrochloride monohydrate[37] (interplanar spacing = 3.30 Å; hydrogen atoms were added to the published atomic parameters on the basis of the expected molecular stereochemistry); (d) 9-methylguanine hydrobromide[38] (interplanar spacing = 3.39 Å; hydrogen atoms were added to the published atomic parameters on the basis of the expected molecular stereochemistry; this stacking pattern is also found in crystals of guanine hydrochloride dihydrate[39]).

Table III. The molecules form a tightly hydrogen bonded network with all eligible oxygen, nitrogen, and sulfur hydrogen bond donors and acceptors participating in the formation of hydrogen bonds. The bases are hydrogen bonded around screw axes to form approximately planar ribbons running in the *b* direction.

Table III. Hydrogen Bond Distances and Angles

Donor atom	Hydrogen atom	Acceptor atom	Distances, Å: Donor–acceptor	Distances, Å: Hydrogen–acceptor	Donor–hydrogen–acceptor angle, deg
N-1	H-1	N-3	3.053	2.10	172
N-2	H-2	N-9	2.973	1.92	163
N-2	H-2′	S	3.327	2.46	171
N-7	H-7	S	3.303	2.27	157

The bond distances and angles are shown in Figure 4. In Table IV the bond distances within the purine ring are compared with those found for guanine; several large differences, primarily in the imidazole

Table IV. Comparison of Bond Lengths within the Purine Rings of Thioguanine and Guanine

Bond	Thioguanine, Å	Guanine, Å
N(1)–C(2)	1.363	1.371
N(1)–C(6)	1.411	1.398
C(2)–N(3)	1.327	1.315
N(3)–C(4)	1.355	1.364
C(4)–C(5)	1.395	1.392
C(4)–N(9)	1.364	1.364
C(5)–C(6)	1.376	1.405
C(5)–N(7)	1.372	1.405
N(7)–C(8)	1.344	1.319
C(8)–N(9)	1.335	1.369

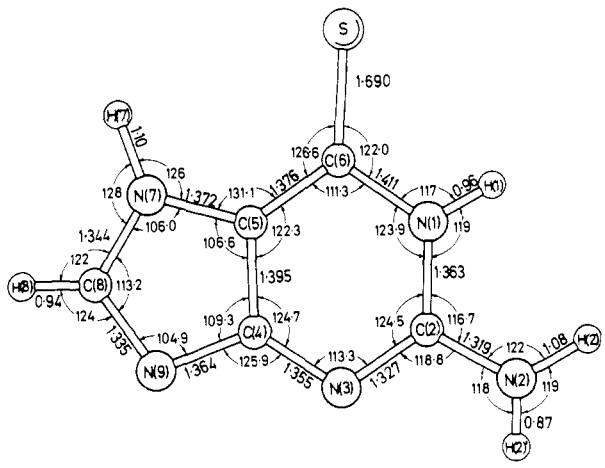


Figure 4. Bond distances (Å) and angles (degrees) within the thioguanine molecule.

rings, are apparent. Table V lists the atomic deviations from least-squares planes through the molecule.

Table V. Deviations from the Least-Squares Planes through Thioguanine

	Plane A deviation, Å	Plane B deviation, Å
N-1	0.012	0.003
C-2	−0.009	−0.010
N-2	−0.003	−0.002[a]
N-3	−0.003	0.002
C-4	0.002	0.004
C-5	0.009	0.003
C-6	0.016	0.004
S	−0.017	−0.040[a]
N-7	−0.001	−0.008
C-8	−0.002	−0.001
N-9	−0.004	0.003
H-1	0.08[a]	0.07[a]
H-2	−0.12[a]	−0.12[a]
H-2′	−0.08[a]	−0.07[a]
H-7	0.04[a]	0.03[a]
H-8	0.12[a]	0.12[a]

[a] Atoms excluded from the calculation of the least-squares plane. The equations of the least-squares planes, with the coefficients of X (ax), Y (by), and Z (cz) equal to the direction cosines with respect to the crystallographic axes, are: plane A, $0.5851X - 0.1643Y - 0.7942Z = 2.893$ Å; plane B, $0.5838X - 0.1588Y - 0.7962Z = 2.917$ Å.

Copies of the structure factor tables will be furnished upon request.

Discussion

It is interesting that thioguanine crystallizes in a tautomer form different from that found for guanine. Since the sugar is substituted at the nine position of purine nucleosides and nucleotides, it is usually assumed that the free bases have hydrogen atoms bonded to this position. However, as pointed out by Marsh[26] and by Donohue,[27] the bases might exist in a number of different tautomer forms. There is no clear evidence that, in the solid state, atom N-9 of purines has a greater affinity for the hydrogen atom than does atom N-7; in the crystal structures of purine,[28] 6-mercaptopurine,[29,30] theophylline,[31] and 2-mercapto-6-methylpurine,[27] the hydrogen atom is also in the N-7, rather than the N-9, position.

Since guanine and thioguanine crystallize as different tautomers, it is not possible to determine the specific effects which the sulfur substituent has on the bond lengths and angles within the purine ring. However, comparison of the bond distances and angles in inosine[32] and 6-thiopurine riboside[33] reveals that replacement of the carbonyl oxygen atom of inosine by a sulfur atom has little effect on the bond lengths and angles within the purine moiety. It is likely that the differences in the bond lengths within the purine rings of guanine and thioguanine (Table III) may be attributed primarily to the difference in tautomer form; this possibility is further supported by the finding that the bond lengths and angles within the imidazole ring of thioguanine are only slightly different from the corresponding values found for purine.[28] The length of the C(6)–S bond (1.69 Å) is in good agreement with the carbon–sulfur bonds in 6-mercaptopurine[29,30] and 6-mercaptopurine riboside.[33]

An important type of interaction between purines in aqueous solution[21] and in the solid state[34–36] involves the vertical stacking of parallel bases. In Figure 3, the base stacking found in thioguanine is compared with that found in the crystal structures of guanine and two protonated guanine derivatives. Whereas considerable overlap of bases is found for guanine,[22,23] base stacking in thioguanine involves little base overlap. However, it is especially noteworthy that the base stacking pattern in thioguanine is practically identical with that found in the crystal structures of guanine hydrochloride monohydrate[37] (Figure 3c), 9-methylguanine hydrobromide[38] (Figure 3d), and guanine hydrochloride dihydrate[39] (Figure 3d), derivatives which also have a hydrogen atom at the N-7 position; this similarity is somewhat surprising since these protonated guanine derivatives possess formal positive charges which usually would be repulsive. In the solid state base stacking of other nucleic acid constituents, it is generally found that closely related purines and pyrimidines show similar stacking patterns.[35,36] This is another striking example of a specific base stacking pattern which persists in several different crystalline environments. Considering the similarity between stacking in guanine derivatives and in thioguanine, it is possible that replacement of guanine by thioguanine in nucleic acids might have little direct effect on base stacking interactions.

(26) R. E. Marsh, "Structural Chemistry and Molecular Biology," A. Rich and N. Davidson, Ed., W. H. Freeman, San Francisco, Calif., 1968, pp 484–489.
(27) J. Donohue, *Acta Crystallogr., Sect. B*, **25**, 2418 (1969).
(28) D. G. Watson, R. M. Sweet, and R. E. Marsh, *ibid.*, **19**, 573 (1965).
(29) E. Sletten, J. Sletten, and L. H. Jensen, *ibid., Sect. B*, **25**, 1330 (1969).
(30) G. M. Brown, *ibid., Sect. B*, **25**, 1338 (1969).
(31) D. J. Sutor, *ibid.*, **11**, 83 (1958).
(32) U. Thewalt, C. E. Bugg, and R. E. Marsh, *ibid.*, in press.
(33) E. Shefter, *J. Pharm. Sci.*, **57**, 1157 (1968).
(34) C. E. Bugg and U. Thewalt, *Biochem. Biophys. Res. Commun.*, **37**, 623 (1969).
(35) M. Sundaralingam, S. T. Rao, C. E. Bugg, and J. Thomas, the American Crystallographic Association Meeting, March 1969, Paper L-5.
(36) C. E. Bugg, J. Thomas, M. Sundaralingam, and S. T. Rao, *Biopolymers*, in press.
(37) J. M. Broomhead, *Acta Crystallogr.*, **4**, 92 (1951).
(38) H. M. Sobell and K. Tomita, *ibid.*, **17**, 126 (1964).
(39) J. Iball and H. R. Wilson, *Proc. Roy. Soc., Ser. A*, **288**, 418 (1965).

One important difference between guanine and thioguanine is in the hydrogen bonding capabilities of the sulfur substituent of thioguanine as compared to the carbonyl oxygen atom of guanine. As can be seen in Figure 1 and Table III, the sulfur atom of thioguanine accepts hydrogen bonds from atoms N-7 and N-2; the lengths of these hydrogen bonds are 3.30 and 3.33 Å, respectively, and the distances and angles involved (Table III) are in good agreement with those found for nitrogen–sulfur hydrogen bonds in a number of other crystal structures.[40] In the crystal structure of guanine,[22,23] the carbonyl oxygen atom also accepts a hydrogen bond from an amino nitrogen atom; however, in this case, the length of the hydrogen bond is 2.93 Å.

Considering specific differences between the crystal structures of guanine[22,23] and thioguanine, it might be expected that substitution of thioguanine for guanine would affect base pairing in double helical nucleic acids. Of particular significance are the large differences between the C(6)–S and C(6)–O(6) covalent bond lengths (1.69 *vs.* 1.24 Å), and between the N–H---S and N–H---O hydrogen bond lengths (3.3 *vs.* 2.9 Å). In the Watson–Crick scheme for base pairing in double helical nucleic acids, the carbonyl oxygen atom of guanine forms a hydrogen bond with the amino nitrogen atom of cytosine.[41] The generally accepted length of this N–H---O hydrogen bond is 2.8–3.0 Å, in agreement with the crystal structure of guanine. On the other hand, if guanine were replaced by thioguanine in nucleic acids, the N–H---S hydrogen bond in thioguanine–cytosine base pairs would probably assume a value close to 3.3 Å, the same as that found in the crystal structure of thioguanine. Coupled with the difference in the C–S and C–O covalent bond lengths, this could result in the formation of thioguanine–cytosine base pairs with dimensions which are considerably different from those of guanine–cytosine pairs.

(40) J. Donohue, *J. Mol. Biol.*, **45**, 231 (1969).
(41) J. D. Watson and F. H. C. Crick, *Nature (London)*, **171**, 737 (1953).

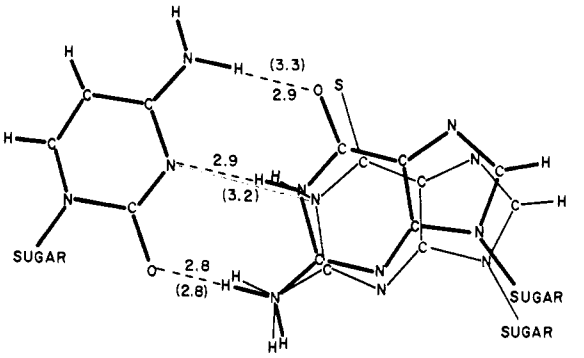


Figure 5. Possible differences between hydrogen bonding in guanine–cytosine and thioguanine–cytosine base pairs. The numbers represent hydrogen bond lengths (Å), with the lengths in parentheses corresponding to thioguanine.

Figure 5 shows the likely differences between thioguanine–cytosine and guanine–cytosine base pairs. The glycosidic carbon atom bonded to atom N-9 of thioguanine could be displaced by more than three-fourths of an ångström from the position which this carbon atom would normally occupy in guanine–cytosine base pairs; this might result in a great deal of distortion in the sugar–phosphate backbone of double helical nucleic acids. Concomitantly, the other two hydrogen bonds to cytosine would be weakened. It is conceivable that such structural differences between thioguanine–cytosine and guanine–cytosine base pairs might be adequately large to disrupt the normal biological functioning of nucleic acids containing thioguanine, thus accounting for the antimetabolite[2] and antitumor activity[3–14] of this purine analog.