Acta Cryst. (1974). B30. 1651

Crystal and Molecular Structure of 9-β-D-Arabinofuranosyladenine*

BY GERARD BUNICK AND DONALD VOET

Department of Chemistry and the Laboratory for Research on the Structure of Matter, University of Pennsylvania, Philadelphia, Pennsylvania 19174, *U.S.A.*

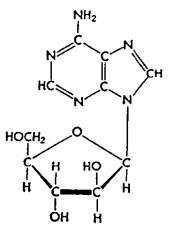
(*Received* 27 *December* 1973; *accepted* 26 *February* 1974)

Crystals of 9-β-D-arabinofuranosyladenine, $C_{10}H_{13}N_5O_4$, have the space group $P2_12_12_1$ with unit-cell dimensions $a=5{\cdot}038$ (1), $b=10{\cdot}435$ (7) and $c=21{\cdot}419$ (4) Å and with four molecules per unit cell. Intensity data were collected with an automated diffractometer using graphite monochromated Cu $K\alpha$ radiation. The structure was solved by direct methods and refined by least-squares calculations to a discrepancy index of $R=0{\cdot}032$ based on 1085 unique reflections. The arabinose ring is puckered in the C(3′) *endo* conformation. The conformation about the C(4′)–C(5′) bond is *gauche–trans* and that about the glycosidic bond is *anti*. The puckering conformation of the arabinose ring in this as well as in other arabinonucleoside structures is due to the strain caused by the rather short N(9)···O(2′) contact. The conformation of 9-β-D-arabinofuranosyladenine is similar to that of adenosine.

Introduction

Several β-arabinonucleosides exhibit potent antibiotic activity in the treatment of viral and bacterial infections and show great promise in the chemotherapeutic treatment of malignant tumors (Cohen, 1966; Schabel, 1968; Suhadolnik, 1970). The first known representatives of this class of compounds, 1-β-D-arabinosylthymine (spongothymine) and 1-β-D-arabinosyluracil (spongouracil), were discovered in extracts of the sponge *Cryptotethya crypta* (Bergman & Feeney, 1950; Burke, 1955).

The synthesis of 9-β-D-arabinofuranosyladenine (Ara-A) (I) (Lee, Benitez, Goodman & Baker, 1960) preceded its discovery in the culture filtrates of *Streptomyces antibioticus* (Parke, Davis & Co., 1967). Ara-A is cytocidal towards bacterial and tumor cells and shows broad-spectrum activity against DNA viruses both in cell cultures and in animals (Cohen, 1966; Schabel, 1968). Ara-A is an effective chemotherapeutic agent in the treatment of several types of mammalian tumors (Cohen, 1966; Suhadolnik, 1970). It exhibits a high activity in producing chromosomal breaks, inhibits DNA-polymerase and, in *in vitro* studies, has been shown to be incorporated into DNA as the terminal residue (Cohen, 1966; Suhadolnik, 1970).



(I)

The low toxicity of Ara-A to normal tissues together with its potent antibiotic and chemotherapeutic activities has stimulated continued interest in the pharmacological properties of this compound. The possibility that the conformation of Ara-A might be correlated with its antibiotic and chemotherapeutic properties prompted the determination of its crystal structure.

Experimental

Crystals of Ara-A ($C_{10}H_{13}N_5O_4$) were grown at room temperature by the vapor diffusion of ethanol into an aqueous solution of Ara-A (Nutritional Biochemicals Corporation) over a two to four week period. Use of this technique usually yielded crystals in the form of very thin needles that were unsuitable for X-ray analyses. However, on one occasion, a larger crystal with dimensions $0{\cdot}28\times0{\cdot}28\times0{\cdot}06$ mm was isolated.

Preliminary Weissenberg and precession photographs of the crystal revealed that it had orthorhombic lattice symmetry. The systematic absences of the $h00$ reflections for h odd, the $0k0$ reflections for k odd and the $00l$ reflections for l odd indicated that the space group of the crystal was $P2_12_12_1$. All subsequent X-ray measurements were made on a Picker FACS-1 diffractometer equipped with a pyrolytic graphite monochromator and employing Cu $K\alpha$ radiation ($\lambda=1{\cdot}5418$ Å). The unit-cell parameters of the crystal, as determined from the least-squares analysis of the angular positions of 12 reflections, are presented in Table 1. The density of the crystal could not be determined as only one crystal of Ara-A that was large enough to work with was available. However the calculated density, assuming one molecule per asymmetric unit of the unit cell, of 1·576 g cm^{-3} is in reasonable agreement with corresponding values found for similar substances [for example the experimentally determined density of adenosine is 1·54 g cm^{-3} (Lai & Marsh, 1972)].

* This paper was presented at the ACA meeting, Gainsville, Fla., January, 1973.

Table 1. *Crystal data for Ara-A*

9-β-D-Arabinofuranosyladenine	$C_{10}H_{13}N_5O_4$
Space group: $P2_12_12_1$	F.W. 267·25
a = 5·038 (1) Å	$Z=4$
b = 10·435 (7)	$D_x=1·576$ g cm^{-3}
c = 21·419 (4)	$F(000)=560$
$V_c=1126$ Å^3	

The X-ray diffraction data were measured with the θ–2θ scan mode, a scan rate of 1° min^{-1}, a scan range of 1·5° and a take-off angle of 2·5°. Stationary background counts of 20 s duration each were taken at both limits of each scan. The 234, $02\bar{7}$ and $\bar{2}27$ reflections were monitored after every fiftieth measurement. They showed no significant change during the data collection. 1085 reflections were measured to the limit $2\theta=125°$.

Structure determination and refinement

The measured intensities of the reflections, I, were corrected for Lorentz and polarization effects. Standard deviations, $\sigma(I)$, were calculated (Stout & Jensen, 1968) with an instrumental instability factor of 0·03. Of the 1085 measured unique reflections 64 had $I<2·33\sigma(I)$. The remainder are observed at the 98% confidence level. No absorption corrections were made due to the small size of the crystal and its low linear absorption coefficient for Cu $K\alpha$ radiation ($\mu=10·7$ cm^{-1}). The normalized structure factors were calculated using the method of polynomial regression to determine $\langle I\rangle$ as a function of $\sin\theta/\lambda$ (Voet, 1972).

All 19 non-hydrogen atoms of Ara-A were located in an E map in which the phases were calculated according to the program *MULTAN* (Main, Woolfson & Germain, 1971) using the 126 reflections with $|E|\geq 1·50$. The structure was refined by the full-matrix least-squares method. The atomic scattering factors for non-hydrogen atoms were taken from Cromer & Waber (1965) and those for hydrogen atoms were taken from Stewart, Davidson & Simpson (1965). A difference Fourier map calculated after the structure had been partially refined revealed the positions of all 13 hydrogen atoms in the structure. Final refinement, in which the overall scale factor, all atomic positional parameters, the anisotropic temperature parameters of the non-hydrogen atoms and the isotropic temperature parameters of the hydrogen atoms were simultaneously varied, caused the discrepancy index to converge to the value $R=0·032$ based on the 1085 observed unique reflections. The final parameter shifts were all less than their respective estimated standard deviations. The largest peak in the difference Fourier map based on the final structural parameters corresponded to an electron density of 0·14 eÅ^{-3}.

Results

The molecular configuration of Ara-A in the crystal structure is illustrated in Fig. 1. This *ORTEP* representation (Johnson, 1965) of the molecule also defines the atomic numbering system used in this report. Table 2 contains the final fractional coordinates and thermal parameters for all atoms in the asymmetric unit of the unit cell together with their standard deviations as estimated from the variance-covariance matrix of the final cycle of the least-squares refinement. Table 3 lists the observed and the calculated structure factors.

Table 2. *Positional and thermal parameters of Ara-A*

The anisotropic temperature factors have the functional form $T=\exp[-(h^2\beta_{11}+k^2\beta_{22}+l^2\beta_{33}+2hk\beta_{12}+2hl\beta_{13}+2kl\beta_{23})\times 10^{-4}]$. Isotropic temperature factors have the functional form $T=\exp(-B\sin^2\theta/\lambda^2)$. Standard deviations, as determined from the variance-covariance matrix of the final cycle of least-squares refinement, are given in parentheses and refer to the least significant digits of their corresponding parameters.

	x	y	z	β_{11}	β_{22}	β_{33}	β_{12}	β_{13}	β_{23}
N(1)	0·4219 (6)	0·9732 (2)	0·4969 (1)	349 (12)	84 (3)	10 (0)	14 (5)	−12 (2)	1 (1)
C(2)	0·2982 (7)	0·8790 (3)	0·5285 (1)	348 (15)	91 (3)	12 (1)	−7 (7)	−21 (3)	−5 (1)
N(3)	0·3311 (5)	0·8411 (2)	0·5871 (1)	267 (11)	74 (2)	11 (0)	−19 (5)	−18 (2)	−3 (1)
C(4)	0·5231 (5)	0·9097 (2)	0·6149 (1)	217 (11)	46 (2)	8 (0)	5 (5)	−1 (2)	−4 (1)
C(5)	0·6704 (5)	1·0071 (2)	0·5885 (1)	232 (11)	51 (2)	7 (0)	13 (5)	0 (2)	0 (1)
C(6)	0·6098 (6)	1·0419 (3)	0·5271 (1)	285 (12)	60 (3)	9 (1)	30 (5)	4 (2)	0 (1)
N(7)	0·8505 (5)	1·0555 (2)	0·6313 (1)	261 (10)	56 (2)	10 (0)	−21 (4)	−1 (2)	1 (1)
C(8)	0·8089 (6)	0·9882 (2)	0·6818 (1)	230 (12)	52 (2)	8 (1)	−21 (5)	−1 (2)	0 (1)
N(9)	0·6122 (4)	0·8987 (2)	0·6754 (1)	196 (9)	49 (2)	7 (0)	−11 (4)	−6 (2)	1 (1)
N(6)	0·7284 (7)	1·1421 (3)	0·4984 (1)	514 (16)	83 (3)	10 (1)	−6 (6)	5 (3)	9 (1)
C(1′)	0·4850 (5)	0·8257 (2)	0·7250 (1)	183 (10)	45 (2)	7 (0)	−8 (5)	1 (2)	−2 (1)
C(2′)	0·6736 (5)	0·7399 (2)	0·7635 (1)	171 (10)	41 (2)	9 (1)	−3 (4)	5 (2)	1 (1)
C(3′)	0·7236 (5)	0·8190 (2)	0·8219 (1)	198 (11)	48 (2)	6 (0)	7 (5)	1 (2)	2 (1)
C(4′)	0·4593 (6)	0·8874 (3)	0·8303 (1)	203 (11)	50 (3)	8 (1)	8 (5)	5 (2)	4 (1)
C(5′)	0·4710 (6)	1·0108 (3)	0·8664 (1)	320 (14)	60 (3)	11 (1)	53 (6)	−3 (3)	−4 (1)
O(1′)	0·3727 (4)	0·9158 (2)	0·7677 (1)	234 (8)	61 (2)	9 (0)	33 (3)	−6 (2)	−1 (1)
O(2′)	0·9133 (4)	0·7084 (2)	0·7332 (1)	182 (7)	63 (2)	10 (0)	8 (3)	3 (2)	−6 (1)
O(3′)	0·7808 (4)	0·7482 (2)	0·8765 (1)	364 (10)	64 (2)	7 (0)	67 (4)	−3 (2)	1 (1)
O(5′)	0·2157 (5)	1·0688 (2)	0·8707 (1)	414 (11)	89 (2)	9 (0)	110 (5)	17 (2)	4 (1)

Table 2 (*cont.*)

	x	*y*	*z*	*B*
H(1)	0·164 (8)	0·833 (3)	0·503 (1)	4·1 (0·9)
H(2)	0·907 (6)	1·000 (2)	0·722 (1)	1·3 (0·5)
H(3)	0·701 (7)	1·156 (3)	0·455 (2)	4·2 (0·9)
H(4)	0·843 (8)	1·196 (3)	0·520 (2)	5·6 (1·0)
H(5)	0·337 (6)	0·779 (2)	0·703 (1)	1·5 (0·5)
H(6)	0·567 (6)	0·657 (3)	0·777 (1)	1·5 (0·6)
H(7)	0·880 (8)	0·667 (4)	0·697 (1)	4·7 (0·9)
H(8)	0·862 (6)	0·882 (3)	0·813 (1)	1·3 (0·6)
H(9)	0·877 (9)	0·689 (4)	0·870 (2)	5·2 (1·1)
H(10)	0·329 (7)	0·824 (3)	0·850 (1)	2·6 (0·7)
H(11)	0·536 (6)	0·996 (3)	0·911 (1)	2·3 (0·6)
H(12)	0·589 (7)	1·071 (3)	0·840 (1)	3·6 (0·8)
H(13)	0·173 (7)	1·061 (3)	0·906 (2)	3·4 (0·8)

Fig. 2 presents the covalent bond angles and distances of the structure. Standard deviations of bond lengths and bond angles not involving hydrogen atoms average 0·003 Å and 0·2°, respectively. The corresponding quantities for bonds involving hydrogen atoms are 0·03 Å and 2°, respectively.

The bond angles and bond distances of the adenine moiety are in excellent agreement with the corresponding quantities found in previously published analyses of crystal structures containing neutral adenine residues (Voet & Rich, 1970). Table 4 indicates that the only ring atoms of the adenine group to deviate from coplanarity by more than three times the standard deviations of their positional parameters are atoms C(6) and N(9). However, a comparison of a number of crystal structures containing adenine residues (Voet & Rich, 1970; Kennard, Motherwell, Coppola, Griffin, Reese & Larson, 1971) revealed no systematic mode by which the adenine ring atoms deviated from coplanarity. Therefore it is concluded that the deviations of the adenine ring atoms of Ara-A from coplanarity are due to the distorting action of packing forces in the crystal.

Table 3. *Observed and calculated structure factors for 9-β-*D*-arabinofuranosyladenine*

Table 4. *Deviations of atoms from the least-squares planes through the adenine and the arabinofuranose rings*

Adenine atom	Deviation†	Arabinofuranose atom	Deviation‡
N(1)	0·001 Å	C(1′)	−0·023 Å
C(2)	0·014	C(2′)	0·014
N(3)	−0·001	O(1′)	0·025
C(4)	0·002	C(4′)	−0·015
C(5)	0·013	C(3′)	0·563*
C(6)	−0·027	C(5′)	0·747*
N(7)	0·015		
C(8)	0·000		
N(9)	−0·018		
N(6)	−0·118*		
C(1′)	−0·282*		
O(2′)	1·947*		
r.m.s. deviation	0·014 Å	r.m.s. deviation	0·020 Å

* Atoms not used in calculating the least-squares plane.
† The equation of the least-squares plane of the adenine ring is $0{\cdot}6835x+0{\cdot}6606y+0{\cdot}3104z=10{\cdot}1504$ Å.
‡ The equation of the least-squares plane of the four most coplanar atoms of the arabinofuranose ring is $0{\cdot}7578x-0{\cdot}6383y+0{\cdot}1351z=0{\cdot}0880$ Å.

Although atoms directly substituent to an adenine ring are often observed to be significantly deviated from the adenine plane, this deviation for atom C(1′) of Ara-A is far more than is usually observed (Voet & Rich, 1970). Table 4 indicates that atom C(1′) is displaced 0·282 Å from the least-squares plane of the adenine ring to the opposite side of the adenine ring from atom O(2′). This is a reflection of the close contact between atoms O(2′) and N(9) (Sherfinski & Marsh, 1973). This close contact is further discussed below. Atom N(6), which is also directly substituent to the adenine ring, deviates 0·118 Å from the plane of the adenine ring to the same side of the ring as atom C(1′). This mode of deviation of atoms N(6) and C(1′) from coplanarity with the adenine ring is the most common one observed in the various adenine-containing crystal structures.

It can be seen from Table 4 that the puckering conformation of the arabinofuranosyl ring is C(3′) *endo* (^3E). Table 5, which contains a list of the important torsion angles in Ara-A, indicates that the adenine residue is in the *anti* conformation with respect to the arabinose ring, that the conformation about the exocyclic bond C(4′)–C(5′) is *gauche–trans* and that the torsion angles about the arabinose residue lie within the usual ranges found in furanose rings with the C(3′) *endo* puckering conformation (Sundaralingam, 1969).

Comparison of Figs. 2 and 3 indicates that the bond angles and bond distances of the arabinose residue are in good agreement with the corresponding quantities of the more accurate published arabinonucleoside structures (Sherfinski & Marsh, 1973; Tollin, Wilson & Young, 1973; Saenger, 1972; Tougard, 1973*a*). Likewise, there is good agreement between the arabinose bond parameters of the present structure and the corresponding parameters derived from averaging

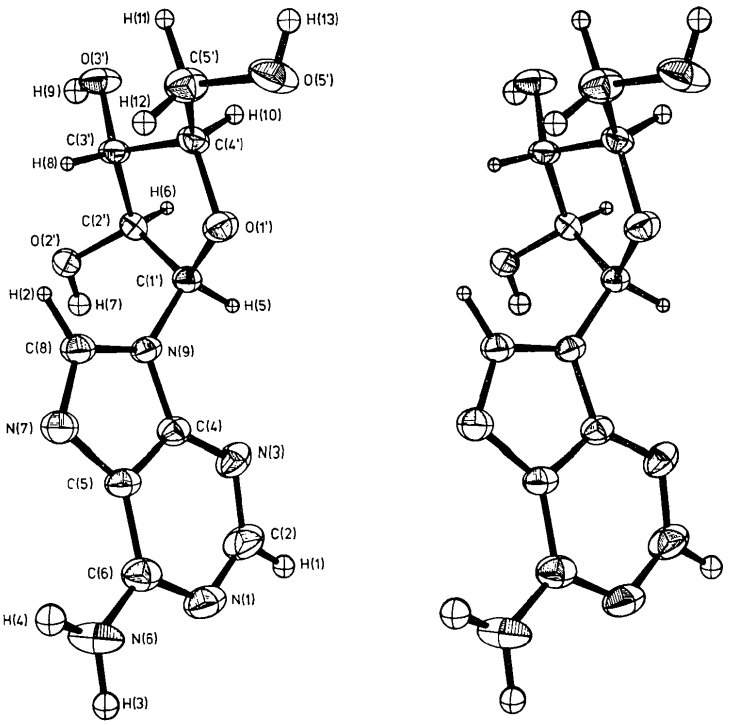
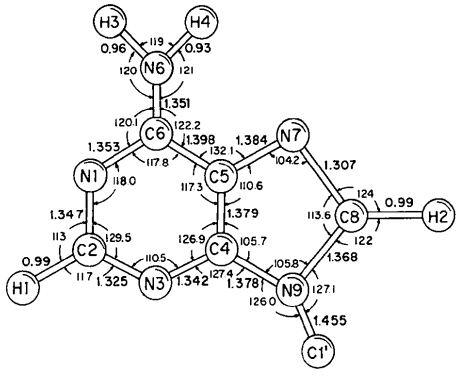


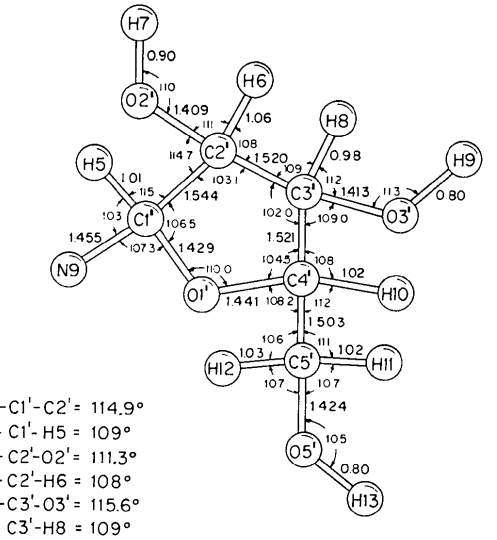
Fig. 1. A stereographic view (Johnson, 1965) of a molecule of 9-β-D-arabinofuranosyladenine. Thermal ellipsoids are drawn at the 50% probability level. Hydrogen atoms are represented as spheres at the 25% probability level.

those of the various published ribofuranosyl rings (Saenger & Eckstein, 1970).

Ara-A and 9-β-D-arabinosyl-4-thiouracil (Ara-4SU) (Saenger, 1972) are the only reported arabinonucleosides in the C(3′) *endo* conformation. The corresponding exocyclic bond angles about atom C(2′) in these structures are closely comparable but those about atom C(3′) are in poor agreement [the C(1′)–C(2′)–O(2′), C(3′)–C(2′)–O(2′), C(2′)–C(3′)–O(3′) and C(4′)–C(3′)–O(3′) angles in Ara-4SU are 114·3, 111·8, 110·2 and 113·3°, respectively (Saenger, 1972)]. The magnitudes of the exocyclic bond angles about atom C(3′) of Ara-A are in good agreement with the corresponding average angles for ribonucleosides shown in Fig. 3 whereas this is not the case with Ara-4SU. Hence, in the absence of further structural data concerning C(3′) *endo* arabinonucleosides, it must be concluded that atom O(3′) of Ara-4SU is in a distorted position.



(*a*)



∠ N9-C1'-C2' = 114.9°
∠ O1'-C1'-H5 = 109°
∠ C3'-C2'-O2' = 111.3°
∠ C1'-C2'-H6 = 108°
∠ C2'-C3'-O3' = 115.6°
∠ C4'-C3'-H8 = 109°
∠ O1'-C4'-H10 = 109°
∠ C3'-C4'-C5' = 115.4°
∠ H11-C5'-H12 = 114°
∠ C4'-C5'-O5' = 111.2°

(*b*)

Fig. 2. The covalent bond distances (Å) and angles (°) of (*a*) the adenine moiety and (*b*) the arabinofuranose moiety of 9-β-D-arabinofuranosyladenine. The estimated standard deviations of these quantities are approximately 0·003 Å and 0·2°, respectively, for bonds not involving hydrogen atoms. For bonds involving hydrogen atoms these quantities are about 0·03 Å and 2°, respectively.

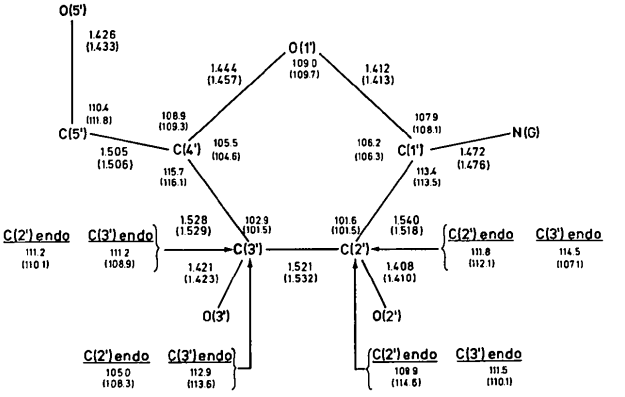


Fig. 3. The average bond distances and angles of the arabinose moiety in the more accurately determined crystal structures containing arabinonucleosides. These include the structures of Ara-A, 1-β-D-arabinofuranosylcytosine hydrochloride (Sherfinski & Marsh, 1973), 1-β-D-arabinofuranosyluracil (Tollin, Wilson & Young, 1973), 1-β-D-arabinofuranosyl-4-thiouracil monohydrate (Saenger, 1972) and 1-β-D-arabinofuranosylthymine (Tougard, 1973*a*). The average values of the bond distances and angles of ribose, as determined from the crystal structures of various nucleosides and nucleotides (Saenger & Eckstein, 1970), are presented in parentheses below the corresponding quantity of the arabinose residue.

As can be seen in Fig. 3, the average magnitude of the C(1′)–C(2′)–O(2′) bond angle for C(3′) *endo* arabinonucleosides differs from the corresponding quantity in ribonucleosides by 7°. However atom O(2′) is on differing sides of the furanose ring in arabinonucleosides and ribonucleosides. This suggests that arabinonucleosides in the C(2′) *endo* and C(3′) *endo* conformations should be correlated with ribonucleosides in the C(2′) *exo* and C(3′) *exo* conformations respectively, when comparing the exocyclic bond angles about atom C(2′) between these two types of nucleosides. The C(2′) *exo* and C(3′) *exo* conformations are rarely observed in ribonucleoside structures. However they closely resemble the C(3′) *endo* and C(2′) *endo* conformations, respectively. Hence in comparing exocyclic bond angles about atom C(2′), the C(2′) *endo* and C(3′) *endo* conformations in arabinonucleosides should be compared with the C(3′) *endo* and C(2′) *endo* conformations, respectively, in ribonucleosides. It can be seen from Fig. 3 that if such comparisons are made, the overall agreement between bond angles C(1′)–C(2′)–O(2′) and C(3′)–C(2′)–O(2′) in arabinonucleosides and ribonucleosides is significantly improved.

Figs. 4 and 5 are two different views of the crystal structure of Ara-A. The hydrogen-bonding pattern present in the crystal structure can be seen in these figures. The hydrogen-bond length and angle data are presented in Table 6. Each molecule of Ara-A is

Table 5. *Arabinose conformational parameters in β-arabinonucleosides*

Pucker conformation	Ara-A C(3′)*endo*	Ara-C.HCl C(2′)*endo*	Ara-4SU C(3′)*endo*	Ara-U C(2′)*endo*	Ara-5 BrU C(1′)*exo*–O(1′)*endo*	Ara-T C(1′)*exo*–O(1′)*endo*	Ara-C C(2′)–*endo*
Glycosidic torsion angle, χ	57·8°, *anti*	26·7°, *anti*	36·0°, *anti*	34·0°, *anti*	25°, *anti*	24·1°, *anti*	28°, *anti*
Furanose ring torsion angles							
O(1′)–C(1′)–C(2′)–C(3′)	−19·0°	25·2°	−27·5°	38·1°	27°	32·0°	35°
C(1′)–C(2′)–C(3′)–C(4′)	33·5	−34·3	39·0	−34·0	−10	−10·5	−34
C(2′)–C(3′)–C(4′)–O(1′)	−36·9	32·8	−38·3	19·4	−12	−13·4	22
C(3′)–C(4′)–O(1′)–C(1′)	26·2	−4·6	22·1	5·0	31	34·6	1
C(4′)–O(1′)–C(1′)–C(2′)	−4·4	−17·9	3·5	−27·5	−37	−42·1	−23
Exocyclic torsion angles							
O(5′)–C(5′)–C(4′)–O(1′)	62·1°, *g*	68·7°, *g*	−55·5°, *g*	−63·3°, *g*	−60°, *g*	−59·5°, *g*	−68°, *g*
O(5′)–C(5′)–C(4′)–C(3′)	178·8, *t*	−171·5, *t*	62·6, *g*	55·8, *g*	61, *g*	60·2, *g*	51, *g*
N(*G*)···O(2′) contact distance	2·788 Å	2·754 Å	2·744 Å	2·677 Å	2·7 Å	2·786 Å	2·85Å

Abbreviations and symbols

Ara-A:	9-β-D-Arabinofuranosyladenine (this study).
Ara-C.HCl:	1-β-D-Arabinofuranosylcytosine hydrochloride (Sherfinski & Marsh, 1973).
Ara-4SU:	1-β-D-Arabinofuranosyl-4-thiouracil monohydrate (Saenger, 1972).
Ara-U:	1-β-D-Arabinofuranosyluracil (Tollin, Wilson & Young, 1973).
Ara-5BrU:	1-β-D-Arabinofuranosyl-5-bromouracil (Tougard, 1969; 1973*b*).
Ara-T:	1-β-D-Arabinofuranosylthymine (Tougard, 1973*a*).
Ara-C:	1-β-D-Arabinofuranosylcytosine (Lefébvre-Soubeyran & Tougard, 1973; Chwang & Sundaralingam, 1973).
χ:	The torsion angle C(8)–N(9)–C(1′)–O(1′) for purines and C(6)–N(1)–C(1′)–O(1′) for pyrimidines as defined by Sundaralingam (1969).
N(*G*):	The glycosidic nitrogen atom: N(9) in purines, N(1) in pyrimidines.
g:	*gauche*
t:	*trans*

Torsion angles are defined according to the right-hand rule (Klyne & Prelog, 1960).

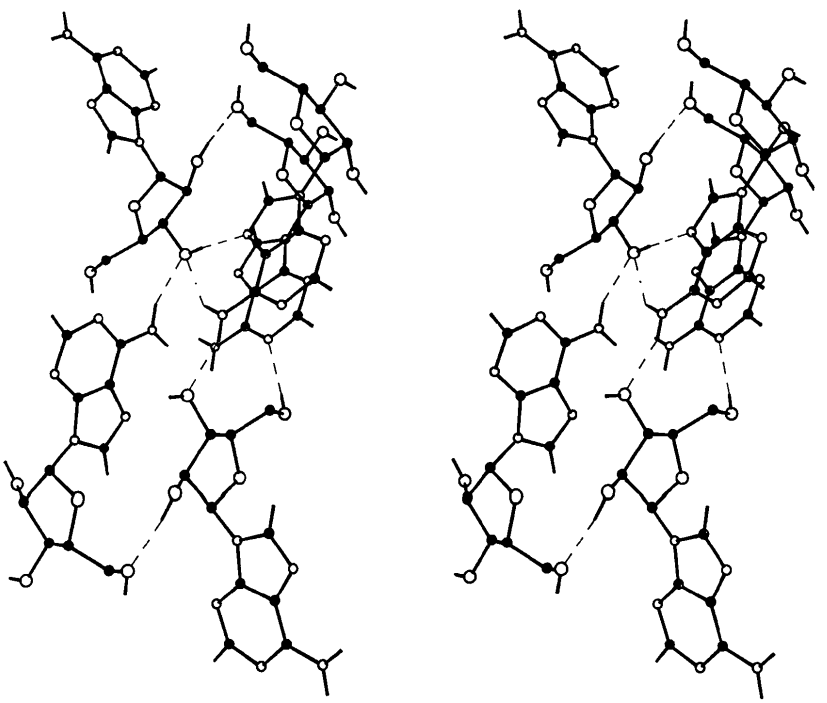


Fig. 4. A stereographic view, roughly along the *a* axis, of the crystal structure of Ara-A. This view illustrates all of the unique hydrogen bonds in the crystal structure. Hydrogen bonds are represented as dashed lines and the overly long hydrogen bond is a dot-dashed line. Non-hydrogen atoms are represented as spheres of arbitrary radii with carbon atoms as filled circles, nitrogen atoms as unfilled circles and oxygen atoms as large unfilled circles. Those hydrogen atoms on the arabinofuranose residue not participating in hydrogen bonding have been deleted from the illustration for the sake of clarity.

capable of forming five hydrogen bonds. However, as can be seen from Table 6, only four of them are formed. The atomic grouping N(6)–H(4)···O(3′*e*) has the proper directionality to form a hydrogen bond but the N(6)···O(3′*e*) distance of 3·809 Å is too long for such an interaction (Hamilton & Ibers, 1968). It is most often the case that molecules arrange themselves in a crystalline lattice so as to form the maximum number of hydrogen bonds of which they are capable. However there are several known exceptions to this rule. In the case of Ara-A the N(6)···O(3′*e*) hydrogen bond cannot form due to the close packing of hydrogen atoms about atom O(3′). It can be seen from Figs. 4 and 5 and from Table 6 that the N(6)–H(4) bond is wedged between the N(6)···O(3′*a*) and the O(3′)···N(7*c*) hydrogen bonds. This results in H(3)···H(4) and H(4)···H(9) contacts of 2·43 and 2·75 Å respectively. These are quite close to the minimal H···H van der

Table 6. *Hydrogen-bond data*

Hydrogen bond *D*–H···*A*	Distance *D*···*A*	Distance H···*A*	Angle *D*–H···*A*
(1) Normal hydrogen bonds			
N(6)–H(3)···O(3′*a*)	2·852 (3) Å	1·95 (3) Å	155 (3)°
O(2′)–H(7)···O(5′*b*)	2·737 (3)	1·84 (4)	172 (3)
O(3′)–H(9)···N(7*c*)	2·742 (3)	1·95 (4)	169 (4)
O(5′)–H(13)...N(1*d*)	2·825 (3)	2·03 (3)	175 (3)
(2) Hydrogen-bond candidate with large *D*···*A* distance			
N(6)–H(4)···O(3′*e*)	3·809 (4)	2·96 (4)	151 (3)

Lower-case letters accompanying atom numbers refer to atoms related to those of Table 3 by the following symmetry operations:

(*a*) $\frac{3}{2}-x,\ 2-y,\ -\frac{1}{2}+z$; (*d*) $\frac{1}{2}-x,\ 2-y,\ \frac{1}{2}+z$
(*b*) $1-x,\ -\frac{1}{2}+y,\ \frac{3}{2}-z$; (*e*) $2-x,\ \frac{1}{2}+y,\ \frac{3}{2}-z$
(*c*) $2-x,\ -\frac{1}{2}+y,\ \frac{3}{2}-z$

Standard deviations as determined from the final cycle of least-squares refinement are given in parentheses and refer to the least significant digit of their corresponding parameter.

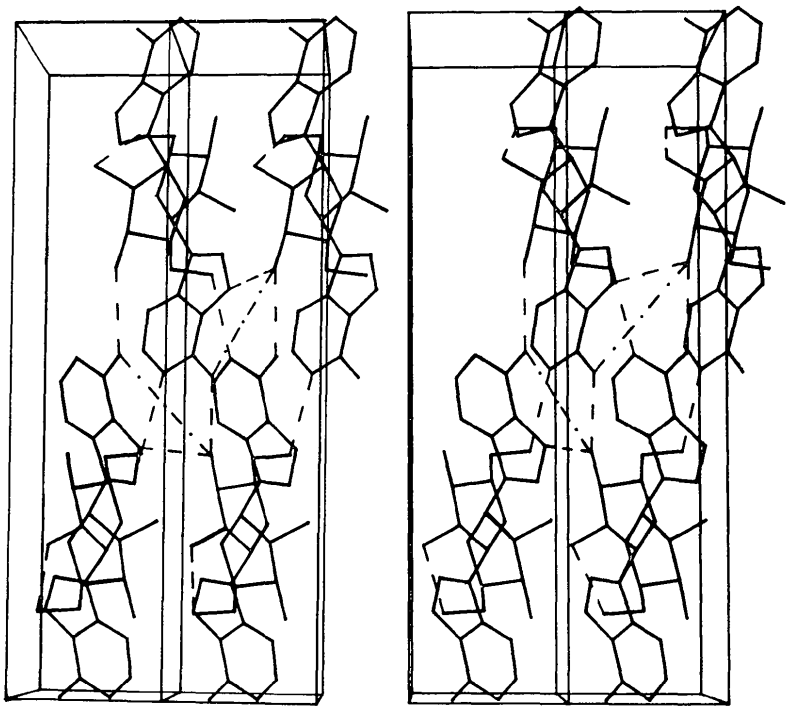


Fig. 5. A stereographic view of a portion of the crystal structure of Ara-A illustrating the packing of its molecules in two adjacent unit cells. The view is roughly along the *b* axis with the *a* and *c* axes extending horizontally and vertically, respectively. Hydrogen bonds are represented as dashed lines and the overly long hydrogen bonds are dot-dashed lines. Hydrogen atoms have been omitted for the sake of clarity.

Waals distance of 2·4 Å (Pauling, 1960). Therefore the formation of the N(6)···O(3′*e*) hydrogen bond by a movement of the N(6)–H(4) bond towards atom O(3′*e*) would result in unacceptable crowding of hydrogen atoms about this atom.

The marginally close contacts of 3·249 Å between atoms C(8) and O(2′*e*) and 2·54 Å between atoms H(2) and O(2′*e*) together with the C(8)–H(2)–O(2′*e*) and H(2)–C(8)–O(2′*e*) angles of 128° and 38°, respectively, is suggestive of a rather weak C–H···O hydrogen bond-like interaction. There is ample structural precedent for the participation of atom C(8) of adenine in such an interaction (Voet, 1972).

The adenine residues form infinite stacks of parallel planar molecules. This is illustrated in Fig. 6. Adjacent molecules in these stacks, which are related by the *a*-axis translation, have an interplanar spacing of 3·44 Å. The closest interatomic distance within a stack, that between atoms N(7) and N(3), is 3·430 Å. This contact is within the usual range for normal stacking interactions (Pauling, 1960). It will be seen from Fig. 6 that the stacking interactions between adenine rings are characterized by the lack of overlap between neighboring molecules. Such a stacking pattern is characteristic of purine and pyrimidine interactions (Bugg, 1972; Bugg, Thomas, Rao & Sundaralingam, 1971).

The non-bonding intramolecular contact between atoms N(9) and O(2′) of 2·788 Å is less than the minimum van der Waals N···O contact of 2·9 Å (Pauling, 1960). There are no unusually short intermolecular contacts in the crystal structure of Ara-A other than those discussed above.

Discussion

Model-building studies indicate that the O(2′) atom of arabinonucleosides greatly hinders the free rotation of both purines and pyrimidines about the glycosidic bond. Hence the allowed rotational range of arabinonucleosides about the glycosidic bond must be considerably smaller than that of the corresponding ribonucleosides. This was confirmed by optical rotatory dispersion (ORD) and circular dichroism (CD) studies of solutions of arabinonucleosides (Guschlbauer & Private de Garilhe, 1969). Crystallographic studies concerning the conformations of arabinonucleosides are summarized by the structural data in Table 5. Here it is seen that for the six published structures of arabinosylpyrimidines the torsion angles of the bases about their glycosidic bonds, χ (Sundaralingam, 1969), lie within 12° of each other. In contrast, this torsion angle for the *anti* conformation spans a range of over 60° in crystal structures of various β-ribosylpyrimidines (Sundaralingam, 1969).

The torsion angle of Ara-A about the glycosidic bond is $\chi = 57{\cdot}8°$ which is an *anti* conformation. The fact that Ara-A is the first β-arabinosylpurine whose structure has been reported of course limits comparative studies. Nevertheless the observation that the value of χ observed for Ara-A lies well outside the range of χ observed for the β-arabinosylpyrimidines when taken together with similar observations for the analogous ribonucleosides (Sundaralingam, 1969) suggests that β-arabinosylpurines have greater rotational freedom about their glycosidic bonds than do β-arabinosylpyrimidines.

It can be seen in Table 5 that the nonbonding interatomic contact N(*G*)···O(2′), in all published arabinonucleoside structures, is less than the minimal N···O van der Waals contact of 2·9 Å (Pauling, 1960). This indicates that the arabinose sugar in β-D-arabinonucleosides is strained in a manner not found in ribonucleosides. However the close O(2′)···O(3′) nonbonding intramolecular contact of approximately 2·7 Å in ribonucleosides is not present in arabinonucleosides. The foregoing is corroborated by Fig. 3 which shows that the C(1′)–C(2′) bond of arabinonucleosides is 0·02 Å longer, on the average, than the corresponding bond in ribonucleosides. Fig. 3 further reveals that the bond angles involving the C(2′)–O(2′) and the C(3′)–O(3′) bonds of arabinonucleosides are distorted in the expected manner relative to the corresponding bond angles of ribonucleosides. There are no other significant differences between corresponding bond parameters of arabino- and ribonucleosides.

The atoms of the furanose ring deviate from coplanarity in order to reduce the steric interference between their otherwise eclipsed substituent atoms (Spencer, 1959). The most serious short contact in β-ribofuranose residues, that between atoms O(2′) and O(3′), is best relieved by a twist about the C(2′)–C(3′) bond. This brings either atom C(2′) or atom C(3′) out of coplanarity with the other four atoms of the furanose ring [see Fig. 3 of Spencer (1959)] and,

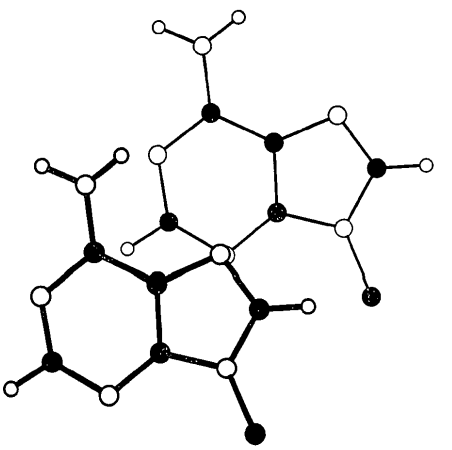


Fig. 6. A projection of two consecutive adenine residues onto the plane of one of the adenine rings. This figure illustrates the base stacking pattern present in the structure of Ara-A. Atoms are represented as circles of arbitrary radii with carbon atoms as filled circles, nitrogen atoms as unfilled circles and hydrogen atoms as smaller, unfilled circles.

hence, accounts for the observation that the sugar puckering conformation in all reported β-ribonucleoside structures is either C(2′) *endo* (^2E), C(3′) *endo* (^3E), C(2′) *exo* (E_2) or C(3′) *exo* (E_3) (Sundaralingam, 1965; 1969). However the two *exo* conformers are less favored, as is observed crystallographically (Sundaralingam, 1969), because they would cause repulsive interactions between atom O(1′) and the oxygen atom substituent to the *exo* carbon atom (Davis, 1973).

The short O(2′)···O(3′) contact present in the ribofuranose residue has been replaced in the arabinofuranose residue by a short O(2′)···N(*G*) contact. Thus the arabinofuranose ring must assume a conformation that results in a twist about its C(1′)–C(2′) bond. This can be achieved with the arabinofuranose ring in any one of the four conformations normally assumed by the ribofuranose ring although the C(3′) *exo* conformer would be less favored for the previously stated reasons. However, the desired twist about the C(1′)–C(2′) bond can also be achieved if the arabinofuranose ring assumes a C(1′) *exo* conformation. [The C(1′) *endo* conformation would also relieve the close O(2′)···N(*G*) contact but this would result in unacceptably close contacts between the atoms substituent to atom C(5′) and the atoms of the purine or pyrimidine base.] However this C(1′) *exo* conformation would have to be supplemented by a twist about the C(3′)–C(4′) bond in order to relieve the otherwise eclipsed conformation about this bond. This occurs when the arabinofuranose ring assumes either the C(1′) *exo*–O(1′) *endo* ($^0T_1$) or the C(1′) *exo*–C(2′) *endo* ($^2T_1$) conformations.

The data in Table 5 corroborate the foregoing arguments. Of the seven reported β-arabinonucleosides, three are in the C(2′) *endo* conformation, two are in the C(3′) *endo* conformation and two (which are closely isomorphous) are in the C(1′) *exo*–O(1′) *endo* conformation. Hence it can be concluded that there is a wider range of sugar pucker conformations available to β-arabinonucleosides as compared with that of β-ribonucleosides.

It is of interest to compare the structure of Ara-A with that of adenosine because Ara-A presumably acts as an adenosine antagonist in biological processes. The crystal structure of adenosine (Lai & Marsh, 1972) has been determined to an accuracy comparable with that of Ara-A (R=0·024 based on 1333 independent reflections). The sugar conformation of Ara-A is quite similar to that of adenosine. Both molecules are in the C(3′) *endo* conformation and maintain a *gauche–trans* conformation about their C(4′)–C(5′) exocyclic bond. The conformation about the glycosidic bond in both molecules is in the *anti* range. However the glycosidic torsional angle, χ, of adenosine is 9·9° whereas the corresponding quantity of Ara-A is 57·8°.

Adenosine crystallizes in the space group $P2_1$. This is different from the space group in which Ara-A crystallizes and therefore the molecular packing within the two structures is quite different. Adenosine forms its full complement of hydrogen bonds but none of these are the same as the hydrogen bonds that are observed in the structure of Ara-A. Both structures form a weak C–H···O hydrogen bond-like interaction. However in adenosine this involves atom C(2) whereas in Ara-A it involves atom C(8). It is of interest that the base-stacking interactions in the crystal structures of adenosine (Bugg, 1972) and Ara-A show some similarity. In both structures there are infinite stacks of adenine residues in which neighboring rings are related by the *a* (short) axis translation.

The structure of Ara-A was elucidated in order to assess the conformational similarities and differences between adenosine and Ara-A. This information could quite possibly lead to a greater understanding of the mechanism through which Ara-A acts as an adenosine antagonist. The conformational similarities between Ara-A and adenosine are such that Ara-A is quite likely to be accepted into many biosynthetic pathways that normally incorporate adenosine. However the biochemical evidence (Cohen, 1966; Schabel, 1968; Suhadolnik, 1970) suggests that once Ara-A is incorporated into various biological molecules, the configurational difference between adenosine and Ara-A renders these molecules partially or even totally biochemically inactive. A less likely reason for the anomalous biochemical behavior of Ara-A-containing molecules is that their biological activity may require a conformational freedom about the glycosidic bond that is not possible in Ara-A.

In order to understand further the biochemical similarities and differences between adenosine and Ara-A it will be necessary to investigate the binding-site properties of the enzymes that can substitute Ara-A in place of adenosine in biosynthetic reactions and to elucidate the nature of the structural and chemical perturbations that Ara-A confers on the molecules in which it is incorporated.

This work was supported, in part, by NIH Grant No. GM 18632 and by the National Science Foundation.

References

BERGMAN, W. & FEENEY, R. J. (1950). *J. Amer. Chem. Soc.* **72**, 2809–2810.

BUGG, C. E. (1972). In *The Purines–Theory and Experiment*, edited by E. D. BERGMAN & B. PULLMAN, pp. 178–204. New York: Academic Press.

BUGG, C. E., THOMAS, J. M., RAO, S. T. & SUNDARALINGAM, M. (1971). *Biopolymers*, **10**, 175–219.

BURKE, D. C. (1955). *J. Org. Chem.* **20**, 643–645.

CHWANG, A. K. & SUNDARALINGAM, M. (1973). *Nature New Biol.* **243**, 78–80.

COHEN, S. S. (1966). *Progr. Nucl. Acid Res. Mol. Biol.* **5**, 1–82.

CROMER, D. T. & WABER, J. T. (1965). *Acta Cryst.* **18**, 104–109.

DAVIS, R. C. (1973). Private communication.

GUSCHLBAUER, W. & PRIVATE DE GARILHE, M. (1969). *Bull. Soc. Chim. Biol.* **51**, 1511–1519.
HAMILTON, W. C. & IBERS, J. A. (1968). *Hydrogen Bonding in Solids*, pp. 15–16. New York: Benjamin.
JOHNSON, C. K. (1965). *ORTEP*. Report ORNL-3794, Oak Ridge National Laboratory, Oak Ridge, Tennessee.
KENNARD, O., MOTHERWELL, W. D. S., COPPOLA, J. C., GRIFFIN, B. E., REESE, C. B. & LARSON, A. C. (1971). *J. Chem. Soc.* (*B*), pp. 1940–1946.
KLYNE, W. & PRELOG, V. (1960). *Experientia*, **16**, 521–523.
LAI, T. F. & MARSH, R. E. (1972). *Acta Cryst.* B**28**, 1982–1989.
LEE, W. W., BENITEZ, A., GOODMAN, L. & BAKER, B. R. (1960). *J. Amer. Chem. Soc.* **82**, 2648–2649.
LEFÈBVRE-SOUBREYAN, O. & TOUGARD, P. (1973). *C. R. Acad. Sci. Paris*, **276**, 403–406.
MAIN, P., WOOLFSON, M. M. & GERMAIN, G. (1971). *MULTAN, A Computer Program for the Automatic Solution of Crystal Structures*. Department of Physics, Univ. of York, York, England.
PARKE, DAVIS & CO. (1967). Patent No. 671,557.
PAULING, L. (1960). *The Nature of the Chemical Bond*, 3rd ed. Ithaca: Cornell Univ. Press.
SAENGER, W. (1972). *J. Amer. Chem. Soc.* **94**, 621–626.
SAENGER, W. & ECKSTEIN, F. (1970). *J. Amer. Chem. Soc.* **92**, 4712–4718.
SCHABEL, F. M. JR (1968). *Chemotherapy*, **13**, 321–338.
SHERFINSKI, J. S. & MARSH, R. E. (1973). *Acta Cryst.* B**29**, 192–198.
SPENCER, M. (1959). *Acta Cryst.* **12**, 59–65.
STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* **42**, 3175–3187.
STOUT, G. H. & JENSEN, L. H. (1968). *X-ray Structure Determination*, pp. 456–458. New York: Macmillan.
SUHADOLNIK, R. J. (1970). *Nucleoside Antibiotics*, chap. 3. New York: Wiley–Interscience.
SUNDARALINGAM, M. (1965). *J. Amer. Chem. Soc.* **87**, 599–606.
SUNDARALINGAM, M. (1969). *Biopolymers*, **7**, 821–860.
TOLLIN, P., WILSON, H. R. & YOUNG, D. W. (1973). *Acta Cryst.* B**29**, 1641–1647.
TOUGARD, P. (1969). *Biochem. Biophys. Res. Commun.* **37**, 961–964.
TOUGARD, P. (1973*a*). *Acta Cryst.* B**29**, 2227–2232.
TOUGARD, P. (1973*b*). *Biochem. Biophys. Acta*, **319**, 116–121.
VOET, D. (1972). *J. Amer. Chem. Soc.* **94**, 8213–8222.
VOET, D. & RICH, A. (1970). *Prog. Nucl. Acid Res. Mol. Biol.* **10**, 183–265.

Acta Cryst. (1974). B**30**, 1660

The Crystal Structure of the Thiamine Hydrochloride Copper(II) Complex

BY M. R. CAIRA, G. V. FAZAKERLEY, P. W. LINDER AND L. R. NASSIMBENI

Department of Chemistry, University of Cape Town, South Africa

(*Received* 7 *March* 1974; *accepted* 11 *March* 1974)

The structure of the 1:1 thiamine hydrochloride–$CuCl_2$ complex, $C_{12}H_{18}N_4OSCl_2.CuCl_2$, has been elucidated by Patterson and Fourier methods and refined by full-matrix least-squares computations to $R=0{\cdot}029$ for 1607 independent reflexions. The monoclinic unit cell, space group $P2_1/c$, with $a=9{\cdot}488$ (5), $b=16{\cdot}871$ (7), $c=12{\cdot}940$ (5) Å and $\beta=117{\cdot}2$ (2)°, contains four complex units. There is no direct bonding between the metal atom and the organic molecule. Instead each thiamine cation is associated with a tetrachlorocuprate anion. The four Cu–Cl bond lengths average 2·25 Å, but the Cl–Cu–Cl angles deviate significantly from ideal tetrahedral geometry. All the hydrogen atoms were located in difference syntheses. In addition to an N–H···O hydrogen bond, the presence of short intermolecular H···Cl distances is taken as evidence of N–H···Cl and C–H···Cl interactions.

Introduction

The crystal structures of several molecules containing the thiamine entity have been reported. Among these are thiamine hydrochloride (vitamin B_1: Kraut & Reed, 1962), thiamine monophosphate (Karle & Britts, 1966), thiamine pyrophosphate (Carlisle & Cook, 1969), thiamine pyrophosphate hydrochloride (Pletcher & Sax, 1971) and bis(protonated thiamine)-tetrachlorodioxouranium(VI) (Marzotto, Bandoli, Clemente, Benetollo & Galzigna, 1973).

As part of a programme investigating the interaction of metal ions and biological molecules, we have elucidated the crystal structure of the thiamine hydrochloride Cu(II) complex (Marzotto, Nicolini, Signor & Galzigna, 1970). Interactions of copper with thiamine are important because it has been shown (Kobayashi, 1972) that copper specifically promotes reactions involving the thiazole ring either by breaking the ring to form thiamine disulphide or by linking with the amino group to form thiochrome. This analysis was undertaken in order to establish whether or not there is direct bonding between the copper atom and thiamine in the solid state of the complex, and if so, which of the rings, thiazolium or pyrimidine, is involved. It has been suggested (Marzotto *et al.*, 1970) that in solution the thiamine molecule is bonded to the metal by a pyrimidine nitrogen atom.