The Crystal and Molecular Structure of Tubercidin, $C_{11}H_{14}N_4O_4$*

By R. M. Stroud†

Birkbeck College Crystallographic Laboratory, Malet Street, London, W.C.1, England

The crystal structure of tubercidin, $C_{11}H_{14}N_4O_4$, has been determined by application of the symbolic addition technique for obtaining general phases directly from the structure-factor magnitudes. The space group is $P2_1$, and the cell dimensions are $a=9\cdot6752$, $b=9\cdot3038$, $c=6\cdot7166$ Å, and $\beta=94\cdot5536°$. Tubercidin is one of a series of recently discovered antibiotics, all of which are analogues of adenosine, and all of which possess some anti-tumour activity. Tubercidin is the most active compound, in which a carbon atom is substituted for a nitrogen atom at position 7 of the adenine base. The most noticeable induced differences between tubercidin and adenosine amount to an increase in the double-bond character of the glycosidic bond, N(9)–C(1′), and a corresponding decrease in the double-bond character of the adjacent N(9)–C(8) bond in the base, N(9)–C(4) being much less affected.

Introduction

Tubercidin is the most biologically active of a series of recently discovered, modified nucleoside antibiotics, each with selective inhibition of cancer cells (Bloch & Nichol, 1964). It differs from adenosine structurally by the substitution of a C–H group at position 7 of the adenine base, for the normal nitrogen atom (Fig. 1). However, its stability to certain critical enzyme cleavages, and also to acid hydrolysis suggests that the substituted group affects the character of the glycosidic bond directly. This study was carried out as part of an investigation into the series, designed to elucidate some of the principles governing the biological action of these antibiotics, and their individual modifications. The mode of action of tubercidin has been considered by Bloch & Nichol (1963).

Experimental

Materials for the present study were provided by Dr A. J. Williamson of the National Institute for Medical Research, Mill Hill, London.

Crystals of tubercidin were obtained by gradual cooling of a warm aqueous solution of the material. They were tabular with a triangular outline and show the faces (110), (1̄10), (010), and the form {001}.

Unit-cell parameters were measured by comparing high-angle reflexions recorded on zero-layer Weissenberg photographs, with superimposed copper-wire diffraction lines (Stroud, 1968), and by using a 6 cm radius high-angle, precision Weissenberg camera. The standard wavelength of Cu $K\alpha_1$ was taken to be 1·54051 Å. The unit-cell dimensions are:

$$a= 9\cdot6752 \ (3) \ \text{Å}$$
$$b= 9\cdot3038 \ (2)$$
$$c= 6\cdot7166 \ (1)$$
$$\beta=94\cdot5536 \ (1)°.$$

The density calculated for $Z=2$ is $1\cdot443$ g cm^{-3}, whereas the density measured by flotation in a mixture of carbon tetrachloride and benzene was found to be $1\cdot449$ g cm^{-3}. Since tubercidin is optically active the crystals lie in a non-centrosymmetric space group. Systematic absences occurred among the $0k0$ reflexions for k odd only, indicating the space group $P2_1$.

X-ray data for the zero to seventh layers were collected on multiple-film equi-inclination Weissenberg photographs, taken around the b axis, using nickel-filtered Cu $K\alpha$ radiation. These data were scaled together by correlation with a zero-layer photograph taken around the c axis. The intensities were measured by visual comparison with a film strip of standardized exposures, and were corrected for spot size, Lorentz and polarization factors, $\alpha_1\alpha_2$ splitting and absorption, and were scaled together using computer programs written by the author (Stroud, 1968). The absorption correction applied was based on transmission coefficients measured on a Hilger–Watts linear diffractometer by counting on an axial reflexion at different φ angle settings. In total, 1117 independent *hkl* reflexions were measured.

Solution of the structure

The crystal structure was solved using the symbolic addition technique applied to non-centrosymmetric structures (Karle & Karle, 1966).

The data were corrected for vibrational motion and placed on an absolute scale using a K curve (Karle, Hauptman & Christ, 1958), and the normalized structure factors, $|E_h|$, were calculated for the 1117 independent X-ray data measured. A comparison of the quantities $\langle|E_h|^2\rangle$, $\langle|E_h|\rangle$, $\langle|E_h^2-1|\rangle$ and the distribu-

tion of E_h with the theoretical values, conforms well to the expected values for a non-centrosymmetric space group.

Phase determination

Arbitrary phases were assigned to a linearly independent set of three E_h which are not structure invariant, in such a way that they restrict the origin to one particular position, for one of the enantiomorphic
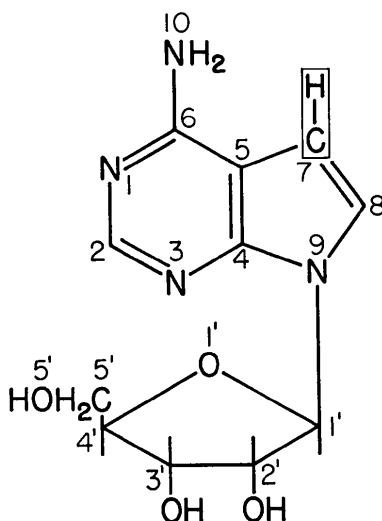

Fig. 1. Chemical composition of tubercidin. The rectangle indicates the substituted C-H group at position 7 of the base. In adenosine N occupies position 7.
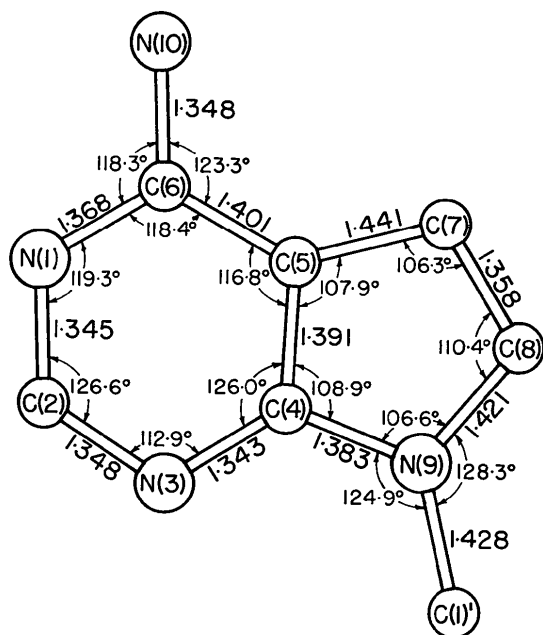

Fig. 2. Bond distances and angles in the base of tubercidin.

forms of the solution, and new phases were calculated by application of equation (1):

$$\varphi_h = \langle \varphi_k + \varphi_{h-k} \rangle_{kr} . \tag{1}$$

The terms used in implementing this relationship at the start were those for which $|E_h E_k E_{h-k}| > 8\cdot5$, corresponding to an expected variance of angle of about $\pm 0\cdot5$ square radians. This criterion was established by trial. Unfortunately, it was not possible to choose the highest $|E_h|$ reflexions of the $h0l$ set to specify the origin in two dimensions since two of them were structure invariants, while four more had no interactions which satisfied the variance condition chosen for new phase indications.

The assignments made to specify the origin were:

hkl	$|E_h|$	Phase
$10\bar{6}$	1·95	0
$40\bar{1}$	2·09	0
$71\bar{4}$	2·45	0 .

Symbolic phases were assigned to other E_h sequentially to enable the phase expansion to continue, and they are listed with the total number of symbolic phases determined (including the origin assignments) at each stage.

h Phase	$|E_h|$	Total phases
Origin set		5
$13\bar{8}=a$	2·99	11
$206=b$	2·20	20
$790=c$	2·76	47

The phase-averaging process equation (1) was then reiterated through all the E_h down to $|E_h|=1\cdot59$, constituting the highest 104 $|E_h|$ of this limited data set. Two cycles of reiteration were carried out, giving a total of 72 acceptable symbolic phases for hkl's with $|E_h| > 1\cdot59$. It soon became clear that b must have the value $b=0$, to avoid a large number of inconsistent indications for the terms in equation (1).

During the course of reiteration several different symbolic expressions arose among terms $(\varphi_k + \varphi_{h-k})$ for given E_h. Having taken b as $b=0$, the total number of these indications involving c and a, through the highest 104 $|E_h|$ and their type were collected and appeared to fall into three distinct categories.

1. Those containing the largest number of indications, generally relating c to a, in the form $c=\pi+na$, where n is an integer.
2. Those relating c to a in the form $c=na$.
3. Those relating a directly to 0 or π.

It was felt that the type-1 indications, accounting for 73% of the total number, gave a clue as to a possible relationship between the symbols a and c.

The trial relationship

$$c=\pi+3\cdot56a \tag{2}$$

was inferred from the type-1 indications. It was derived as a weighted average and used in the substitution of values for the symbolic phases a and c.

The enantiomorph was specified by restricting a to the region between 0 and π. Values of a, from $\pi/8$, to π radians, were substituted into the basic set of 72 phases in steps of $\pi/8$ using equation (2) to calculate the value of c, and the eight sets of phases so generated expanded using the tangent formula (Karle & Hauptman, 1956). In this way the expansion was carried down to $|E_h| = 1\cdot0$, through the 419 highest $|E_h|$ values, in four stages. Finally, two refinement cycles using the tangent formula proved sufficient for good convergence. From 10 to 25% of the phases were rejected at this stage because of inadequate statistical indications, and E maps were calculated using each of the eight expanded phase sets. The map corresponding to

$a = 3\pi/4$ alone revealed the structure of the molecule. The trial adenine ring was planar, and all the supposed bonded distances between the remaining atoms were about 1·4 to 1·5 Å.

Computer programs to perform the correction of data, computation of $|E_h|$ values, symbolic addition process, and the tangent formula expansion were written by the author for the London University Atlas computer (Stroud, 1968).

Refinement of the structure

The trial positions obtained from the E map were refined using a block-diagonal approximation of the least-squares process together with Hughes's (1941)

Table 1. *Tubercidin. The final atomic positions of one molecule in fractional coordinates, and their standard deviations*

|	x	y	z
N(1)	0·6002 (5)	0·9596 (13)	0·4502 (13)
C(2)	0·4609 (9)	0·9660 (25)	0·4363 (28)
N(3)	0·3763 (5)	0·9149 (12)	0·5698 (12)
C(4)	0·4451 (7)	0,8470 (13)	0·7238 (13)
C(5)	0·5883 (7)	0·8323 (13)	0·7544 (16)
C(6)	0·6675 (7)	0·8930 (12)	0·6102 (17)
N(10)	0·8070 (5)	0·8860 (12)	0·6196 (12)
N(9)	0·3842 (5)	0·7821 (12)	0·8802 (9)
C(8)	0·4941 (11)	0·7276 (19)	1·0117 (20)
C(7)	0·6184 (8)	0·7564 (20)	0·9397 (13)
C(1')	0·2384 (7)	0·7657 (15)	0·8924 (16)
O(1')	0·1943 (5)	0·8507 (8)	1·0531 (10)
O(2')	0·1916 (4)	0·5162 (9)	0·7857 (9)
C(2')	0·1966 (7)	0·6113 (12)	0·9470 (16)
C(3')	0·0502 (6)	0·6422 (12)	1·0116 (12)
O(3')	−0·0471 (3)	0·6575 (8)	0·8426 (7)
C(4')	0·0672 (7)	0·7872 (12)	1·1156 (13)
C(5')	0·0811 (10)	0·7756 (18)	1·3432 (15)
O(5')	0·0968 (4)	0·9179 (8)	1·4284 (8)
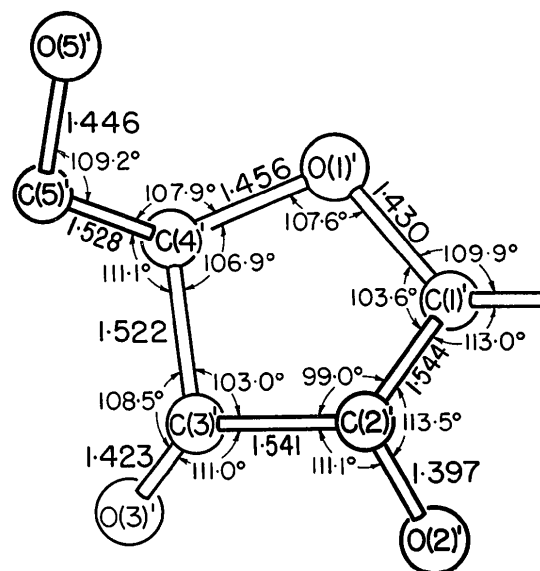
Fig. 3. Bond distances and angles of the ribose moiety in tubercidin.

Table 2. *Tubercidin. Final anisotropic thermal vibration parameters and their standard deviations* ($\times 10^4$)

The form of the temperature correction is $T = \exp\left[-(\beta_{11}h^2 + \beta_{22}k^2 + \beta_{33}l^2 + \beta_{23}kl + \beta_{31}hl + \beta_{12}hk)\right]$.

	β_{11}	β_{22}	β_{33}	β_{23}	β_{31}	β_{12}
N(1)	40 (5)	92 (23)	157 (30)	106 (61)	61 (29)	− 9 (24)
C(2)	42 (7)	117 (43)	224 (62)	158 (150)	55 (50)	− 4 (41)
N(3)	48 (5)	61 (17)	136 (26)	83 (56)	28 (28)	14 (22)
C(4)	44 (6)	58 (22)	83 (23)	6 (57)	40 (29)	− 7 (27)
C(5)	35 (5)	56 (24)	116 (29)	5 (62)	−10 (30)	2 (23)
C(6)	36 (5)	30 (22)	145 (32)	− 20 (64)	29 (31)	−12 (22)
N(10)	41 (4)	86 (22)	123 (24)	42 (58)	64 (26)	− 8 (23)
N(9)	37 (4)	104 (22)	74 (17)	21 (50)	5 (20)	− 5 (24)
C(8)	60 (6)	72 (34)	105 (25)	67 (67)	10 (32)	0 (31)
C(7)	49 (6)	110 (35)	88 (25)	− 75 (80)	44 (33)	−26 (36)
C(1')	41 (5)	52 (23)	113 (26)	9 (68)	65 (30)	−12 (26)
O(1')	65 (4)	45 (11)	193 (25)	− 35 (45)	137 (29)	−64 (17)
O(2')	41 (3)	84 (15)	178 (24)	−137 (53)	67 (23)	−28 (17)
C(2')	41 (6)	42 (23)	130 (29)	− 63 (62)	49 (33)	−46 (25)
C(3')	28 (4)	55 (21)	72 (20)	−6 (50)	30 (22)	41 (22)
O(3')	29 (2)	46 (10)	142 (18)	27 (35)	−45 (16)	3 (12)
C(4')	52 (6)	43 (21)	81 (22)	− 12 (56)	69 (29)	−28 (27)
C(5')	76 (10)	74 (28)	86 (25)	− 78 (75)	17 (38)	− 2 (39)
O(5')	71 (5)	52 (11)	113 (16)	− 83 (34)	10 (21)	32 (18)

weighting scheme. The function minimized was $R_1 = \sum_{hkl} w(|F_o| - |F_c|)^2$. Several cycles reduced the R index, $R = \sum |F_o - F_c|/\sum|F_o|$, from 0·30 for the trial structure to 0·11, when refinement was terminated. Shifts at this point were all less than 0·1 standard deviations. The final positional parameters are shown in Table 1 together with their standard deviations while, Table 2 shows the final anisotropic temperature parameters and their standard deviations. The mean values of the standard deviations as derived from the least-squares refinement are 0·013 Å for the bonds, and 0·9° for the angles.

Table 3. Intramolecular bond distances

	Tubercidin		Adenosine*
	Distance (Å)	Distance (Å)‡	Distance (Å)
N(1) –C(2)	1·345 (10)	1·350 (4)	1·340
C(2) –N(3)	1·348 (18)	1·329 (4)	1·330
N(3) –C(4)	1·343 (13)	1·346 (4)	1·349
C(4) –C(5)	1·391 (10)	1·398 (4)	1·381
C(5) –C(6)	1·401 (14)	1·406 (4)	1·415
C(6) –N(1)	1·361 (14)	1·347 (4)	1·351
C(6) –N(10)	1·348 (8)	1·341 (4)	1·332
C(5) –C(7)†	1·441 (16)	1·433 (4)	1·385
C(7)† –C(8)	1·358 (14)	1·359 (4)	1·308
C(8) –N(9)	1·421 (14)	1·400 (4)	1·362
N(9) –C(4)	1·383 (12)	1·370 (4)	1·374
N(9) –C(1′)	1·428 (8)	1·438 (4)	1·466
C(1′)–C(2′)	1·544 (11)	1·520 (4)	1·530
C(2′)–C(3′)	1·541 (9)	1·526 (4)	1·528
C(2′)–O(2′)	1·397 (13)	1·401 (4)	1·411
C(3′)–C(4′)	1·522 (15)	1·527 (4)	1·522
C(3′)–O(3′)	1·423 (8)	1·423 (4)	1·417
C(4′)–O(1′)	1·456 (9)	1·451 (4)	1·450
O(1′)–C(1′)	1·430 (13)	1·422 (4)	1·411
C(4′)–C(5′)	1·528 (13)	1·506 (4)	1·509
C(5′)–O(5′)	1·446 (17)	1·452 (4)	1·420

Overall r.m.s. difference = 0·013 Å between Stroud bond lengths and Abola & Sundaralingam bond lengths.

* From Lai & Marsh (1972).
† N(7) in adenosine.
‡ After further refinement of Abola & Sundaralingam (1973).

Molecular parameters

The final interatomic bond distances and angles and the hydrogen bond lengths, together with their standard deviations, are shown in Tables 3, 4, and 5. Table 3 also includes the corresponding bond lengths obtained from the crystallographic study of adenosine by Lai & Marsh (1972) for comparison. The bond lengths and angles in the base and ribose moieties of tubercidin are shown diagrammatically in Figs. 2 and 3. A three-dimensional electron-density synthesis was calculated using the observed structures amplitudes, F_{obs} and the phases φ_h, calculated from the refined set of parameters. This map showed the complete molecule and no other significant peaks above the background. A complete list of the observed and calculated structure factors, and their phase angles φ_h are listed in Table 6.

The purine residue

The base residue of tubercidin is planar, and the equation of the least-squares plane through the nine atoms of the purine ring is

$$0·0933x + 8·0934y + 3·3678z = 9·291$$
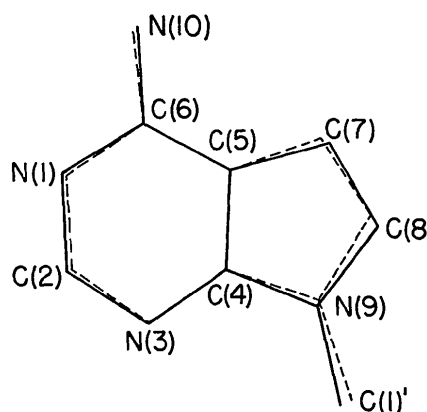
Fig. 4. Diagram of the planar purine residue (solid line) in tubercidin, compared with the base in deoxyadenosine (dotted line).

Table 4. Intramolecular bond angles and their standard deviations

C(6)–N(1)–C(2)	119·3 (1·1)°	C(8) –N(9) –C(1′)	128·3 (0·8)°
N(1)–C(2)–N(3)	126·6 (1·6)	N(9) –C(1′)–C(2′)	113·0 (0·8)
C(2)–N(3)–C(4)	112·9 (1·3)	N(9) –C(1′)–O(1′)	109·9 (1·0)
N(3)–C(4)–C(5)	126·0 (1·1)	O(1′)–C(1′)–C(2′)	103·6 (0·6)
C(4)–C(5)–C(6)	116·8 (1·0)	C(1′)–C(2′)–C(3′)	99·0 (0·8)
C(5)–C(6)–N(1)	118·4 (0·8)	C(2′)–C(3′)–C(4′)	103·0 (0·8)
N(1)–C(6)–N(10)	118·3 (1·0)	C(3′)–C(4′)–C(5′)	111·1 (0·7)
C(5)–C(6)–N(10)	123·3 (1·0)	C(3′)–C(4′)–O(1′)	106·9 (0·7)
N(3)–C(4)–N(9)	125·1 (0·9)	O(1′)–C(4′)–C(5′)	107·9 (0·6)
C(6)–C(5)–C(7)	135·3 (0·7)	C(4′)–O(1′)–C(1′)	107·6 (0·8)
C(4)–C(5)–C(7)	107·9 (0·7)	C(1′)–C(2′)–O(2′)	113·5 (0·5)
C(5)–C(7)–C(8)	106·3 (1·3)	C(3′)–C(2′)–O(2′)	111·1 (0·8)
C(7)–C(8)–N(9)	110·4 (1·4)	C(2′)–C(3′)–O(3′)	111·0 (0·9)
C(8)–N(9)–C(4)	106·6 (0·8)	C(4′)–C(3′)–O(3′)	108·5 (0·7)
N(9)–C(4)–C(5)	108·9 (1·0)	C(4′)–C(5′)–O(5′)	109·2 (0·7)
C(4)–N(9)–C(1′)	124·9 (0·7)		

Table 5. *Hydrogen-bond lengths*

The numbers in parentheses indicate the molecule to which the second atom belongs; a 1 or 0 before the / indicates whether the atom is related by a 2_1 axis to the first listed molecule, and the triple following / gives the associated translation involved in units of **a**, **b**, and **c**.

N(1)—O(2')	(1/1 0 1)	2.712 ± 0.008 Å
N(3)—O(5')	(0/0 0-1)	2.794 ± 0.007
N(10)–O(2')	(1/1 0 1)	2.981 ± 0.011
N(10)–O(3')	(0/1 0 0)	2.902 ± 0.011
N(10)–O(5')	(0/1 0-1)	3.189 ± 0.007
O(3') –O(5')	(1/0 0-1)	2.894 ± 0.008

with respect to crystal axes *a*, *b*, and *c*. The distances of these atoms and C(1') and N(10) from this plane

are shown in Table 7. The atoms in the rings themselves are insignificantly displaced from the plane, with a mean deviation of about 0·007 Å. C(1') of the glycosidic bond is displaced by 0·115 Å from the base plane; however, such a displacement is not uncommon in purine nucleosides. The glycosidic bond itself, C(1')–N(9), makes an angle of 4°35' with the base plane. It is of interest to note that in the study of adenosine (Lai & Marsh, 1972) C(1') does not deviate significantly from the base plane.

The bond lengths and angles in the base appear to be significantly changed by the presence of C–H at position 7 of the base, and they are compared with those of the adenine base in deoxyadenosine in Fig. 4.

Table 6. *List of observed and calculated structure factors for tubercidin*

The format is *hk*, followed by columns of *l*, $100F_{obs}$, $100F_{calc}$, φ_{calc}.

[Table of structure factors — illegible]

Table 6 (*cont.*)

• 5	5		
-6	202	315	-0.41
-5	558	453	-1.77
-4	1075	1145	1.76
-3	272	335	-0.25
-2	1311	1307	1.23
-1	961	927	0.28
0	928	855	0.77
1	333	359	1.71
2	820	760	-2.02
3	1403	1526	-2.01
4	1317	1385	-2.41
5	220	154	-3.00
• 6	5		
-6	182	210	-2.73
-5	450	359	-1.43
-4	887	[illegible]	-3.02
-3	1155	1155	2.91
-2	1071	1173	2.72
-1	500	477	0.63
0	576	541	1.70
1	1322	1318	2.11
2	1157	1053	-3.03
3	523	565	2.26
4	1074	1063	-1.76
5	477	473	-3.12
• 7	5		
-5	405	438	2.03
-4	374	481	-2.54
-3	572	531	2.73
-2	542	515	2.14
-1	936	955	-2.15
0	[illegible]	903	1.81
1	1349	1470	-1.38
2	690	684	1.18
3	715	713	2.81
4	210	133	1.57
5	1074	948	0.57
• 8	5		
-5	943	717	-2.57
-4	208	116	0.19
-3	227	176	0.24
-2	722	706	2.71
-1	1411	1458	0.74
0	571	561	-2.25
1	564	556	-2.05
2	1031	980	-2.81
3	749	766	-2.52
4	177	233	0.84
• 9	5		
-5	982	970	2.86
-4	550	500	2.52
-3	166	[illegible]	-1.74
-2	1113	1002	-0.21
-1	222	221	0.23
0	222	200	0.20
1	329	325	-0.03
2	201	311	-1.13
3	480	530	-1.37
4	573	437	1.54
• 10	5		
-4	426	426	0.77
-3	500	547	1.02
-2	611	649	0.73
-1	552	635	-0.15
0	157	137	1.74
1	239	293	-1.16
2	237	270	2.09
3	371	345	-0.07
• 0	6		
1	1315	1457	-0.65
2	1709	1832	-2.77
3	1038	1023	[illegible]
4	2468	2497	0.06
5	741	725	1.86
6	421	459	-0.02
7	432	339	-1.94
• 1	6		
-6	252	356	-1.23
-5	704	710	0.20
-4	281	215	0.81
-3	675	612	1.06
-2	1637	1671	0.08
-1	1403	1415	-0.48
0	540	720	-0.72
1	1912	1502	-0.63
2	366	307	2.67
3	2276	2076	-1.73
4	818	760	-2.70
5	416	333	-0.71
6	1022	1011	-0.72
7	178	150	-0.43
• 2	6		
-7	500	395	0.39
-6	255	444	0.90
-5	547	437	2.04
-4	1360	1285	0.73
-3	426	384	-2.70
-2	822	713	-0.14
-1	563	526	-2.77
0	1400	1414	-1.19
1	1470	1339	1.15
2	1373	1438	3.13
3	478	355	2.24
4	1281	1213	3.11
5	1004	841	-1.78
6	307	345	-2.32
7	159	160	-1.73
• 3	6		
-6	452	688	-0.92
-5	365	260	2.59
-4	1354	1286	-2.33
-3	1057	1004	-2.37
-2	478	340	-0.80
-1	1079	1164	-1.30
0	1190	1222	-2.03
1	852	812	-0.35
2	255	214	2.10
3	648	656	0.36
4	656	774	0.36
5	526	434	-0.16
6	324	415	-0.16
7	261	257	-0.69
• 4	6		
-6	824	358	-2.06
-5	276	210	2.55
-4	1025	963	-2.35
-3	1400	1517	-2.16
-2	1076	1102	-0.54
-1	906	679	-2.15
0	1210	1129	0.71
1	1164	1160	2.70
2	751	635	1.67
3	1013	1155	1.07
4	548	457	1.17
5	770	740	-0.43
• 5	6		
-7	274	109	-1.76
-6	214	166	0.02
-5	305	200	0.06
-4	467	480	-0.61
-3	371	411	-0.03
-2	755	761	-0.93
-1	1231	1147	-1.27
0	919	926	-0.12
1	592	562	2.45
2	910	1000	[illegible]
3	758	755	2.02
4	452	470	-0.11
5	665	511	-0.04
• 6	6		
-5	528	449	1.37
-4	712	641	-2.33
-3	902	907	1.54
-2	470	457	0.73
-1	830	715	0.02
0	1510	1560	1.22
1	603	706	-2.85
2	660	527	1.41
3	828	915	-1.64
4	612	505	-2.09
5	318	276	-3.05
• 7	6		
-5	721	747	-1.49
-4	855	243	1.02
-3	815	729	0.11
-2	[illegible]	[illegible]	[illegible]
-1	1064	884	2.70
0	831	706	0.31
1	1002	770	1.73
2	321	269	2.30
3	357	241	-1.19
4	655	612	-1.64
5	316	334	-1.61
• 8	6		
-4	945	858	3.14
-3	192	169	1.03
-2	550	601	3.01
-1	550	493	1.43
0	1310	1479	-2.05
1	1311	1343	2.37
2	657	573	1.89
3	226	207	0.87
4	507	437	1.12
• 9	6		
-4	541	480	-2.74
-3	900	612	-0.07
-2	373	373	2.13
-1	261	257	0.17
0	655	615	1.74
3	175	205	1.82
• 10	6		
-3	578	524	2.58
-2	931	107	-0.29
-1	558	561	1.17
0	414	478	2.61
• 0	7		
1	166	160	1.50
2	1327	1610	1.89
3	1151	1169	0.29
4	271	221	-0.63
5	370	255	-0.18
6	173	203	1.00
• 1	7		
-6	358	317	-3.07
-5	544	522	-0.45
-4	1130	1071	2.67
-3	1809	2022	0.81
-2	806	725	-2.71
-1	413	957	1.29
0	1179	953	-1.63
1	722	771	2.33
2	624	561	1.67
3	716	755	-2.00
4	646	731	0.64
5	422	372	-0.82
6	459	381	1.08
• 2	7		
-6	249	159	0.18
-5	305	205	1.64
-4	966	1004	0.45
-3	423	820	-0.17
-2	366	273	1.95
-1	1096	1157	1.27
0	463	463	-1.84
1	1150	1143	-2.86
2	575	664	-2.03
3	1000	1075	-2.37
4	443	476	-1.51
5	631	446	-1.32
6	327	310	3.12
• 3	7		
-6	415	456	1.63
-5	655	615	2.02
-4	700	667	1.43
-3	756	715	-0.03
-2	533	496	-0.90
-1	773	670	-0.09
0	1126	1150	-2.31
1	845	915	0.14
2	1011	1172	-1.67
3	730	473	-1.10
4	346	335	-2.44
5	368	402	-1.76
6	405	308	-1.71
• 4	7		
-5	243	264	1.49
-4	288	160	-1.47
-3	1000	1014	-1.80
-2	725	776	-1.38
-1	437	508	-1.10
0	876	796	-2.05
1	209	265	0.11
2	820	896	-1.06
3	659	727	-1.66
4	368	430	-0.26
• 5	7		
-5	370	454	-3.10
-4	259	151	0.25
-3	1663	1620	-2.34
-2	962	917	-1.35
-1	1375	1301	-1.26
0	568	571	3.05
1	436	476	2.03
2	222	166	-0.54
3	1258	1191	0.33
4	809	609	-0.01
• 6	7		
-4	247	269	-2.38
-3	451	363	-0.48
-2	441	403	0.90
-1	754	667	-1.81
0	844	711	-2.90
1	730	713	0.73
2	737	736	-2.02
3	410	363	-0.03
• 7	7		
-4	413	351	-1.00
-3	423	304	-1.41
-2	311	293	1.54
-1	777	731	2.75
0	1426	1395	0.70
1	1645	1301	-2.13
2	320	283	-1.09
3	341	363	-2.73
• 8	7		
-3	764	639	-0.06
-2	824	870	3.01
-1	613	556	-1.43
0	[illegible]	615	2.64
1	674	592	0.00
2	790	405	-0.55
• 9	7		
-3	264	226	-2.06
-2	490	456	-1.98
-1	328	369	2.87
0	307	301	1.33
1	566	[illegible]	-1.37
2	309	308	-1.71
• 10	7		
-1	405	911	-2.64
0	413	410	1.79
• 1	8		
0	1619	1724	1.44
• 2	8		
0	756	790	3.02
• 3	8		
0	1411	1431	1.93
• 4	8		
0	886	869	-0.56
• 6	8		
0	777	791	-0.56
• 7	8		
0	921	1123	-2.04
• 3	9		
0	964	1035	2.43
• 4	9		
0	773	[illegible]	2.53
• 6	9		
0	216	228	-2.54
• 7	9		
0	1303	1317	-0.85
• 6	10		
0	234	454	0.57
• 0	2		
0	6374	6745	-1.98
• 0	4		
0	1008	1157	-2.60
• 0	6		
0	1311	1160	-1.52
• 0	8		
0	1519	1524	0.87

Table 7. *Distances of the atoms from the least-squares base plane*

N(1)	−0·001 Å	N(9)	−0·009 Å
C(2)	−0·009	C(8)	0·003
N(3)	0·019	C(7)	0·005
C(4)	−0·005	N(10)	−0·006
C(5)	−0·007	C(1′)	−0·115
C(6)	0·005	O(1′)	1·110

The bonds involving C(7) are both longer than they are in adenosine (Table 3) which has a nitrogen atom at position 7 of the base. However, these increases would be expected in the immediate vicinity of the substituted atom.

½
Z
0
−½
0
X
½
1

Fig. 5. Packing diagram of the molecules in a unit cell. The dotted lines show hydrogen bonds present, and the broken hydrogen bond aligns with a molecule translated by *b* from the one shown.

The glycosidic bond N(9)–C(1′) observed as 1·428 (8) Å is significantly shorter than that observed in other nucleoside structures, where the bond length is about 1·47 Å, very close to that of a single, paraffinic bond length, C–N (see, for example, Kraut & Jensen, 1963; Sundaralingam, 1966; Lai & Marsh, 1972). The short length of the glycosidic bond indicates an unusual amount of double-bond character, which might account for the increased stability of tubercidin to acid hydrolysis, and to enzymic or phosphorolytic cleavage. The further refinement of this structure using diffractometer collected data (Abola & Sundaralingam, 1973) confirms the relative shortening of this bond (1·438 Å). The C(8)–N(9) bond is more affected by the substitution at position 7 than is the C(4)–N(9) bond, and the refinement of Abola & Sundaralingam indicates that this latter bond length is insignificantly different from C(4)–N(9) in adenosine (Lai & Marsh, 1972) while the N(9)–C(8) bond, which is adjacent to the glycosidic bond, is increased by about 0·037 Å. It therefore has a smaller proportion of double-bond character in tubercidin than it does in adenosine. The C(4)–C(5) bond length is 0·17 Å longer in tubercidin also.

The ribose moiety

The ribose ring is puckered, with C(2′) displaced by 0·665 Å from the best least-squares plane through the other four atoms of the furanose ring, and to the same side of the ring as C(5′). The plane is described with respect to crystallographic axes by the equation

$$5{\cdot}2799x - 4{\cdot}4067y + 4{\cdot}3852z = 1{\cdot}829$$

and displacements of the atoms from this plane are given in Table 8. The puckering is thus described as C(2′) *endo* puckering.

Table 8. *Distances of atoms from the sugar plane*

C(1′)	−0·038 Å
O(1′)	0·060
C(3′)	0·035
C(4′)	−0·058
C(1′)	0·661

The average length of the C–C bonds in the sugar ring system is 1·54 Å, while the average C–O bond length is 1·430 Å, in agreement with the accepted single bond length of about 1·43 Å.

The torsion angle φ_{CN}, defined by Donohue & Trueblood (1960) as the dihedral angle between the base plane, and the plane containing N(9)–C(1')–O(1'), is −65° in tubercidin. This is another example of an adenosine molecule showing 2' puckering of the sugar ring with *endo* displacement of C(2'); *i.e.*, lying on the same side of the sugar ring as C(5'), and the large value of φ_{CN} here observed is not unusual in this conformation. For adenosine in β-adenosine-2'-β-uridine-5'-phosphoric acid (Shefter, Barlow, Sparks & Trueblood, 1964; 1969) the angle is $\varphi_{CN} = -54°$. Molecules showing 3' puckering generally show smaller rotation angles (Haschemeyer & Rich, 1967). The angle for tubercidin falls into the category of 'allowed angles' in purine nucleosides, based on normal and rare intramolecular van der Waals radii for C(2') *endo* puckering, as discussed by Haschemeyer & Rich.

Hydrogen bonding and packing of the molecules

The scheme of hydrogen bonds holding the molecules together in the crystal lattice is shown in Fig. 5 while the hydrogen-bond lengths are presented in Table 5. Each molecule is linked to six others. The nitrogen atom N(10) and oxygen atom O(5') are both involved in three hydrogen bonds. Two of the N(10) bonds lie almost in the (010) plane. O(2') and O(3') are both involved in two bonds, while N(3) and N(1) are involved in only one bond. No attempt was made to locate the hydrogen atoms.

Discussion

It appears that, apart from the expected changes in the bonds to C(7), the main difference between adenosine and tubercidin amounts to a decrease in the double-bond character of C(8)–N(9), and an increase in the double-bond character of the adjacent glycosidic bond. These differences are possibly accompanied by relatively minor changes in other parts of the molecule.

The structure obtained through this application of the symbolic addition technique was in the correct absolute configuration for tubercidin.

References

ABOLA, J. & SUNDARALINGAM, M. (1973). *Acta Cryst.* B29, 697–703.
BLOCH, A. & NICHOL, C. A. (1963). Studies on the mode of action of tubercidin. *Abstracts 145th Meeting of American Chemical Society*, p. 35 ff.
BLOCH, A. & NICHOL, C. A. (1964). *Antimicrobial Agents and Chemotherapy*, p. 530. American Society for Microbiology.
DONOHUE, J. & TRUEBLOOD, K. N. (1960). *J. Mol. Biol.* 2, 363–371.
HASCHEMEYER, A. E. V. & RICH, A. (1967). *J. Mol. Biol.* 27, 369–384.
HUGHES, E. W. (1941). *J. Amer. Chem. Soc.* 63, 1737–1751.
KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* 9, 635–651.
KARLE, J., HAUPTMAN, H. & CHRIST, C. L. (1958). *Acta Cryst.* 11, 757–761.
KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* 21, 849–859.
KRAUT, J. & JENSEN, L. H. (1963). *Acta Cryst.* 16, 79–88.
LAI, T. F. & MARSH, R. E. (1972). *Acta Cryst.* B28, 1982–1989.
SHEFTER, E., BARLOW, M., SPARKS, R. & TRUEBLOOD, K. N. (1964). *J. Amer. Chem. Soc.* 86, 1872–1874.
SHEFTER, E., BARLOW, M., SPARKS, R. & TRUEBLOOD, K. N. (1969). *Acta Cryst.* B25, 895–909.
STROUD, R. M. (1968). Ph.D. Thesis, Univ. of London.
SUNDARALINGAM, M. (1966). *Acta Cryst.* 21, 495–506.