Acta Cryst. (1993). C**49**, 1095–1097

Structure of a Dideoxynucleoside Active Against the HIV (AIDS) Virus

BY J. M. GULBIS AND M. F. MACKAY*

Department of Chemistry, La Trobe University, Bundoora, Victoria 3083, *Australia*

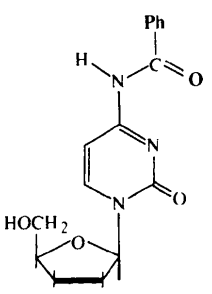
AND G. HOLAN AND S. M. MARCUCCIO

CSIRO, Division of Chemicals and Polymers, Private Bag 10, *Clayton, Victoria* 3168, *Australia*

(*Received* 14 *October* 1992; *accepted* 22 *February* 1993)

Abstract. N^4-Benzoyl-2′,3′-dideoxycytidine, $C_{16}H_{17}N_3O_4.2H_2O$, $M_r = 351.4$, monoclinic, $P2_1$, $a = 4.691$ (1), $b = 14.448$ (2), $c = 12.924$ (3) Å, $\beta = 97.63$ (1)°, $V = 868.2$ (1) Å^3, D_m(flotation) = 1.34 (1), $D_x = 1.344$ Mg m^{-3}, $Z = 2$, $F(000) = 372$, $\lambda = 1.5418$ Å, μ(Cu $K\alpha$) = 0.65 mm^{-1}, $T = 292$ (1) K, final $R = 0.038$ for 1437 observed data. The glycosidic torsion angle C(6)—N(1)—C(1′)—O(4′) is 20.6 (5)° and the pucker of the furanose ring is C(3′) *endo*. Free rotation about the exocyclic C(4′)—C(5′) bond allows the hydroxymethyl substituent to adopt two orientations, *trans* and *gauche*, the latter resulting in a short contact, H(6)···O(5″) of 2.21 (4) Å, indicative of a relatively strong C—H···O intramolecular hydrogen-bonding interaction.

Introduction. The observation that many inhibitors of reverse transcriptase exhibit antiviral activity is of relevance to the design of new compounds for the treatment of human immunosuppressive virus (HIV), the causative agent of AIDS. Foremost among these are the 2′,3′-dideoxynucleosides (Mitsuya & Broder, 1987), a number of which are now undergoing clinical trials in humans (Sarin, 1988), with 2′,3′-dideoxycytidine (DDC), also known as zalcitabine, recently approved for the treatment of AIDS in Austria, Canada and the USA. Reported here is the crystal structure of one such inhibitor, N^4-benzoyl-2′,3′-dideoxycytidine (1).



(1)

* Author to whom correspondence should be addressed.

Experimental. Colourless tabular crystals elongated along **a**, from ethyl acetate containing a small amount of methanol; crystal *ca* 0.13 × 0.16 × 0.53 mm aligned on a Rigaku AFC diffractometer; cell parameters determined by least squares from 2θ values for 25 strong reflections ($41 < 2\theta < 78°$); ω–2θ scans, 2θ scan rate 2° min^{-1}, scan range (1.20 + 0.5tanθ)°, $2\theta_{max} = 130°$, 10 s stationary background counts; three standard reflections monitored every 50 reflections, no significant intensity variation; 1580 data recorded (h −5 to 5, k 0 to 17, l 0 to 15), 1506 independent data ($R_{int} = 0.016$), 1437 for which $I \geq 2\sigma(I)$ used for refinement; intensities corrected for Lorentz and polarization, and for absorption, transmission factors 0.854 to 0.946. Structure solved by direct methods with *SHELXS*86 (Sheldrick, 1985). The —CH_2OH side chain is disordered resulting in two sites for the hydroxyl O(5′) and O(5″) atoms, with site occupancies (constrained to unity) refining to 0.82 (1) and 0.18 (1), respectively. The coordinates of H(4), H(6) and H(O5′) were refined, H(*W*1*a*), H(*W*2*a*) of the water molecules were included at sites located on difference maps. The remainder, apart from H(O5″), H(*W*1*b*), H(*W*2*b*) and those at C(5′) whose sites were not located, and the atoms thus omitted from the analysis, were included at calculated positions. Full-matrix least-squares refinement (*SHELX*76; Sheldrick, 1976) with anisotropic temperature factors given to C, N and O atoms, isotropic for H, converged at $R = 0.038$, $wR = 0.048$, $S = 2.12$ (257 parameters varied); function minimized $\sum w(|F_o| - |F_c|)^2$ with $w = (\sigma|F_o|^2 + 0.0003|F_o|^2)^{-1}$; $(\Delta/\sigma)_{max} = 0.001$; $(\Delta\rho)_{max} = 0.21$, $(\Delta\rho)_{min} = -0.15$ e Å^{-3}. An isotropic extinction correction of the form $F = F_c[1 - (3.82 \times 10^{-6}F^2/\sin\theta)]$ was applied to the calculated structure amplitudes. Atomic scattering factors and anomalous-dispersion corrections from *International Tables for X-ray Crystallography* (1974, Vol. IV, pp. 99, 149). Figures were prepared from the output of *ORTEP*II (Johnson, 1976). Calculations performed on a VAX 11/780 computer.

0108-2701/93/061095-03$06.00

Table 1. *Final atomic coordinates* ($\times 10^4$) *and isotropic or equivalent isotropic temperature factors* ($Å^2$) *with e.s.d.'s in parentheses*

$$B_{eq} = (8\pi^2/3)\sum_i\sum_j U_{ij} a_i^* a_j^* \mathbf{a}_i.\mathbf{a}_j.$$

	x	y	z	B_{eq}/B_{iso}
N(1)	−3796 (5)	−4231 (3)	−1668 (2)	2.96 (4)
C(2)	−3732 (6)	−3331 (4)	−2049 (2)	2.95 (5)
O(2)	−1916 (5)	−2779 (3)	−1619 (2)	4.02 (4)
N(3)	−5665 (6)	−3080 (3)	−2866 (2)	3.26 (5)
C(4)	−7566 (6)	−3678 (3)	−3299 (2)	2.90 (5)
N(4)	−9460 (6)	−3312 (3)	−4117 (2)	3.38 (5)
C(5)	−7666 (6)	−4601 (3)	−2958 (2)	3.58 (6)
C(6)	−5716 (6)	−4846 (4)	−2132 (3)	3.42 (6)
C(7)	−11461 (8)	−3781 (4)	−4796 (3)	3.82 (6)
O(7)	−11715 (8)	−4620 (3)	−4756 (2)	5.94 (6)
C(8)	−13289 (6)	−3210	−5576 (2)	3.94 (6)
C(9)	−14620 (12)	−3659 (5)	−6462 (3)	7.17 (11)
C(10)	−16413 (15)	−3172 (5)	−7209 (4)	8.92 (14)
C(11)	−16894 (11)	−2249 (5)	−7073 (4)	6.46 (10)
C(12)	−15614 (10)	−1807 (4)	−6205 (3)	5.48 (8)
C(13)	−13793 (8)	−2283 (4)	−5454 (3)	4.40 (7)
C(1′)	−1638 (6)	−4475 (4)	−750 (3)	3.21 (5)
C(2′)	−2651 (8)	−4176 (4)	275 (3)	4.00 (6)
C(3′)	−4029 (8)	−5048 (4)	620 (3)	4.41 (7)
C(4′)	−2146 (6)	−5799 (4)	254 (3)	3.56 (6)
O(4′)	−1285 (5)	−5435 (3)	−716 (2)	3.62 (4)
C(5′)	−3579 (9)	−6724 (4)	61 (4)	5.11 (8)
O(5′)†	−1709 (8)	−7407 (3)	−207 (3)	5.22 (7)
O(5″)‡	−5795 (32)	−6769 (13)	−809 (13)	6.6 (4)
O(*W*1)	−2851 (6)	−1031 (3)	−2697 (2)	4.94 (5)
O(*W*2)	−11700 (6)	−6194 (3)	−6102 (2)	5.04 (6)§

† Atom included with occupancy factor 0.82 (1).
‡ Atom included with occupancy factor 0.18 (1).
§ Isotropic temperature factor.

Table 2. *Bond lengths* (Å), *bond angles* (°) *and selected torsion angles* (°) *with e.s.d.'s in parentheses*

N(1)—C(2)	1.392 (7)	C(8)—C(13)	1.373 (6)
N(1)—C(6)	1.348 (6)	C(9)—C(10)	1.385 (8)
N(1)—C(1′)	1.496 (4)	C(10)—C(11)	1.368 (10)
C(2)—O(2)	1.243 (5)	C(11)—C(12)	1.359 (7)
C(2)—N(3)	1.347 (4)	C(12)—C(13)	1.387 (6)
N(3)—C(4)	1.312 (5)	C(1′)—O(4′)	1.397 (7)
C(4)—N(4)	1.392 (4)	C(1′)—C(2′)	1.528 (6)
C(4)—C(5)	1.407 (6)	C(2′)—C(3′)	1.510 (8)
N(4)—C(7)	1.375 (5)	C(3′)—C(4′)	1.514 (7)
C(5)—C(6)	1.358 (5)	C(4′)—C(5′)	1.502 (8)
C(7)—O(7)	1.220 (7)	C(4′)—O(4′)	1.465 (5)
C(7)—C(8)	1.484 (5)	C(5′)—O(5′)	1.394 (7)
C(8)—C(9)	1.390 (5)	C(5′)—O(5″)	1.428 (16)

C(2)—N(1)—C(6)	120.2 (3)	C(8)—C(9)—C(10)	120.1 (4)
C(2)—N(1)—C(1′)	117.2 (3)	C(9)—C(10)—C(11)	120.1 (5)
C(6)—N(1)—C(1′)	122.6 (3)	C(10)—C(11)—C(12)	120.1 (4)
N(1)—C(2)—N(3)	119.0 (3)	C(11)—C(12)—C(13)	120.5 (4)
O(2)—C(2)—N(3)	121.9 (3)	C(8)—C(13)—C(12)	120.2 (3)
C(2)—N(3)—C(4)	120.4 (3)	N(1)—C(1′)—O(4′)	108.9 (3)
N(3)—C(4)—N(4)	113.8 (2)	N(1)—C(1′)—C(2′)	111.4 (3)
N(3)—C(4)—C(5)	122.7 (3)	C(2′)—C(1′)—O(4′)	107.6 (3)
N(4)—C(4)—C(5)	123.6 (3)	C(1′)—C(2′)—C(3′)	102.4 (3)
C(4)—N(4)—C(7)	127.6 (3)	C(2′)—C(3′)—C(4′)	102.4 (3)
C(4)—C(5)—C(6)	116.5 (3)	C(3′)—C(4′)—O(4′)	104.4 (3)
N(1)—C(6)—C(5)	121.2 (3)	C(5′)—C(4′)—O(4′)	110.2 (3)
N(4)—C(7)—O(7)	121.7 (3)	C(3′)—C(4′)—C(5′)	115.0 (3)
N(4)—C(7)—C(8)	116.3 (3)	C(1′)—O(4′)—C(4′)	109.8 (3)
O(7)—C(7)—C(8)	122.0 (3)	C(4′)—C(5′)—O(5′)	112.9 (3)
C(7)—C(8)—C(9)	117.1 (3)	C(4′)—C(5′)—O(5″)	115.7 (7)
C(7)—C(8)—C(13)	123.9 (3)	O(5′)—C(5′)—O(5″)	100.9 (6)
C(9)—C(8)—C(13)	119.0 (3)		

C(6)—N(1)—C(2)—O(2)	−178.9 (4)	C(6)—N(1)—C(1′)—C(2′)	−98.0 (5)
C(6)—N(1)—C(2)—N(3)	1.9 (6)	C(2)—N(1)—C(1′)—C(2′)	83.0 (4)
N(1)—C(2)—N(3)—C(4)	−0.4 (6)	O(2)—C(2)—N(1)—C(1′)	0.1 (5)
C(2)—N(3)—C(4)—C(5)	−1.3 (5)	C(1′)—C(2′)—C(3′)—C(4′)	35.0 (4)
N(3)—C(4)—C(5)—C(6)	1.5 (5)	C(2′)—C(3′)—C(4′)—O(4′)	−34.1 (4)
C(4)—C(5)—C(6)—N(1)	0.0 (6)	C(3′)—C(4′)—O(4′)—C(1′)	19.8 (4)
C(5)—C(6)—N(1)—C(2)	−1.6 (6)	C(4′)—O(4′)—C(1′)—C(2′)	2.8 (5)
C(5)—C(4)—N(4)—C(7)	−8.2 (6)	O(4′)—C(1′)—C(2′)—C(3′)	−24.2 (4)
C(4)—N(4)—C(7)—O(7)	−2.2 (7)	C(3′)—C(4′)—C(5′)—O(5′)	−176.0 (4)
C(4)—N(4)—C(7)—C(8)	177.7 (3)	C(3′)—C(4′)—C(5′)—O(5″)	68.5 (9)
N(4)—C(7)—C(8)—C(13)	−22.5 (5)	O(4′)—C(4′)—C(5′)—O(5′)	66.4 (5)
C(5)—C(6)—N(1)—C(1′)	179.4 (4)	O(4′)—C(4′)—C(5′)—O(5″)	−49.1 (9)

Discussion. Final atomic coordinates for non-H atoms are given in Table 1,* bond lengths, bond angles and selected torsion angles are listed in Table 2 and the molecular conformation and atom-numbering scheme is illustrated in Fig. 1. The conformational parameters used hereafter follow the guidelines of the IUPAC–IUB Joint Commission on Biochemical Nomenclature (1983). The glycosidic torsion angle χ_{CN} [C(2)—N(1)—C(1′)—O(4′)] is −158.4 (3)° and the pucker of the furanose ring is C(3′) *endo* (^3E, $P = 14.0°$, $\psi_m = 36.1°$). A disordering of the —CH_2OH side chain gives rise to two rotamers of approximate populations 80 and 20%. The two C(4′)—C(5′) conformations are +*ap*(*gauche*$^-$ *trans*) with $\gamma = 176.0\ (4)°$, and +*sc*(*gauche*$^-$ *gauche*) with $\gamma = 68.5\ (9)°$. A similar disorder has been noted in crystals of the antiviral nucleoside, 2′-fluoro-5-iodoarabinoylcytosine (FIAC) (Birnbaum, Cygler, Watanabe & Fox, 1982), but in this case the populations of the two rotamers were equal. Interestestingly, in crystals of 3′-azido-3′-deoxythymidine (AZT) (Gurskaya, Tsapkina, Skaptsova, Kraevskil, Lindeman & Struchkov, 1986; Camerman, Mastropaolo & Camerman, 1987; Birnbaum, Giziewicz, Gabe, Lin & Prusoff, 1987; Dyer, Low, Tollin, Wilson & Howie, 1988) there are two crystallographically independent molecules of the nucleoside, one with the *trans* and one with the *gauche*$^+$ conformation, whereas in crystals of 2′,3′-dideoxycytidine (DDC) (Birnbaum, Lin & Prusoff, 1988; Silverton, Quinn, Haugwitz & Todara, 1988) only the *trans* rotamer is observed. As in AZT, FIAC and other comparable nucleosides (Birnbaum, Blonski & Hruska, 1983; Birnbaum, Sadana, Blonski & Hruska, 1986) the *gauche* rotamer in the title compound is stabilized by an intramolecular C(6)—H(6)···O(5″) hydrogen bond (Taylor & Kennard, 1982). The H(6)···O(5″) distance of 2.21 (4) Å is very similar to the values reported for this interaction in the other nucleosides. Although the dimensions are less favourable than for the intramolecular interaction in the *gauche*$^+$ rotamer, there is indication of

* Lists of structure amplitudes, anisotropic thermal parameters, H-atom coordinates and short intermolecular contacts have been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 55947 (19 pp.). Copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England. [CIF reference: HL1026]

a weak interaction in both rotamers between the C(6) proton and the furanose ring O atom, the C(6)···O(4′) and H(6)···O(4′) distances being 2.717 (5) and 2.31 (3) Å, respectively, with the angle at H(6) 101 (2)° [*cf.* the similar interaction in AZT (Birnbaum, Giziewicz, Gabe, Lin & Prusoff, 1987)]. This hydrogen bond must therefore be considered bifurcated in the *gauche*$^+$ rotamer with H(6) being donated to both O(5″) and O(4′). The pyrimidine ring atoms are coplanar to within 0.01 (1) Å and the interplanar angle between the pyrimidine and phenyl rings is 31 (1)°. The peptide bond adopts the usual *trans* planar conformation. The bond lengths and angles are in good agreement with comparable structures.

The crystal packing is illustrated in Fig. 2. The nucleoside molecules are stacked in layers approximately parallel to the (101) planes. Intermolecular hydrogen bonds link the nucleoside and water molecules into a three-dimensional network (Table 3).

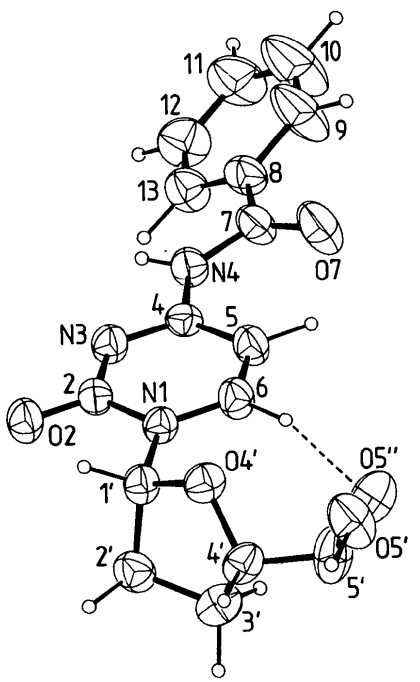


Fig. 1. A perspective view of the molecule with thermal ellipsoids scaled to 50% probability. The C symbol is omitted for C atoms and the H atoms are denoted by spheres of arbitrary radii.

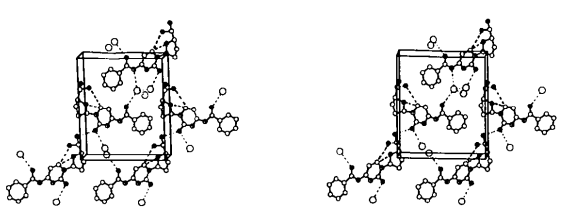


Fig. 2. Stereoview of the crystal packing along **a** with the *b* axis vertical.

Table 3. *Hydrogen-bonding geometry* (Å, °) *with e.s.d.'s in parentheses*

D—H···*X*	*D*···*X*	H···*X*	*D*—H	*D*—H···*X*
C(6)—H(6)···O(5″)*	3.26 (2)	2.21 (4)	1.05 (4)	178 (3)
C(6)—H(6)···O(4′)*	2.717 (5)	2.31 (3)	1.05 (4)	101 (2)
O(5′)—H(O5′)···O(2)^{i}	2.772 (4)	1.81 (6)	0.98 (6)	166 (6)
N(4)···H(4)···O(*W*2)^{ii}	3.114 (6)	2.31 (4)	0.85 (4)	158 (6)
O(*W*1)—H(*W*1*a*)···O(2)	2.890 (6)	1.87	1.03	174
O(*W*2)—H(*W*2*a*)···O(7)	2.864 (5)	1.80	1.07	171
O(*W*2)···O(*W*1)^{iii}	2.811 (4)			
O(*W*2)···O(*W*1)^{iv}	2.819 (4)			
O(5′)···O(5″)^{v}	3.12 (2)			

Symmetry code: (i) $-x, -\frac{1}{2}+y, -z$; (ii) $-2-x, \frac{1}{2}+y, -1-z$; (iii) $-1-x, -\frac{1}{2}+y, -1-z$; (iv) $-2-x, -\frac{1}{2}+y, -1-z$; (v) $1+x, y, z$.

* Intramolecular interaction.

The interaction between adjacent nucleoside molecules related by the twofold screw axis, in which O(5′) of the *trans* rotamer donates its proton to the exocyclic O(2) atom ($-x, -\frac{1}{2}+y, -z$) of the pyrimidine ring, link the molecules into infinite spirals along the [010] direction. This relatively strong interaction stabilizes the *trans* conformation of the —CH_2OH side chain of the more highly populated *trans* rotamer. Moreover, this interaction provides the only direct linkage between the nucleoside molecules, the other linkages involving bridging *via* the water molecules. There is no close overlap of the aromatic systems in the crystal.

References

BIRNBAUM, G. I., BLONSKI, W. J. P. & HRUSKA, F. E. (1983). *Can. J. Chem.* **61**, 2299–2304.

BIRNBAUM, G. I., CYGLER, M., WATANABE, K. A. & FOX, J. J. (1982). *J. Am. Chem. Soc.* **104**, 7626–7630.

BIRNBAUM, G. I., GIZIEWICZ, J., GABE, E. J., LIN, T.-S. & PRUSOFF, W. H. (1987). *Can. J. Chem.* **65**, 2135–2139.

BIRNBAUM, G. I., LIN, T.-S. & PRUSOFF, W. H. (1988). *Biochem. Biophys. Res. Commun.* **151**, 608–614.

BIRNBAUM, G. I., SADANA, K. L., BLONSKI, W. J. P. & HRUSKA, F. E. (1986). *J. Am. Chem. Soc.* **108**, 1671–1675.

CAMERMAN, A., MASTROPAOLO, D. & CAMERMAN, N. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 8239–8242.

DYER, I., LOW, J. N., TOLLIN, P., WILSON, H. R. & HOWIE, R. A. (1988). *Acta Cryst.* C**44**, 767–769.

GURSKAYA, G. V., TSAPKINA, E. N., SKAPTSOVA, N. V., KRAEVSKIL, A. A., LINDEMAN, S. V. & STRUCHKOV, YU. T. (1986). *Dokl. Akad. Nauk. SSSR*, **291**, 854–859; (1987). *Chem. Abstr.* **106**, 111695*g*.

IUPAC–IUB JOINT COMMISSION ON BIOCHEMICAL NOMENCLATURE (1983). *Pure Appl. Chem.* **55**, 1273–1280.

JOHNSON, C. K. (1976). *ORTEP*II. Report ORNL-5138. Oak Ridge National Laboratory, Tennessee, USA.

MITSUYA, H. & BRODER, S. (1987). *Nature (London)*, **325**, 773–778.

SARIN, P. S. (1988). *Annu. Rev. Pharmacol.* **28**, 411–428.

SHELDRICK, G. M. (1976). *SHELX*76. Program for crystal structure determination. Univ. of Cambridge, England.

SHELDRICK, G. M. (1985). *SHELXS*86. In *Crystallographic Computing* 3, edited by G. M. Sheldrick, C. Krüger & R. Goddard, pp. 175–189. Oxford Univ. Press.

SILVERTON, J. V., QUINN, F. R., HAUGWITZ, R. D. & TODARA, L. J. (1988). *Acta Cryst.* C**44**, 321–324.

TAYLOR, R. & KENNARD, O. (1982). *J. Am. Chem. Soc.* **104**, 5063–5070.