Acta Crystallographica Section C
Crystal Structure Communications
ISSN 0108-2701

2′-Deoxy-5-methylisocytidine

Frank Seela,[a]* Yang He,[a] Hans Reuter[b] and Eva-Maria Heithoff[b]

[a]Laboratorium für Organische und Bioorganische Chemie, Institut für Chemie, Universität Osnabrück, Barbarastrasse 7, D-49069 Osnabrück, Germany, and [b]Anorganische Chemie II, Institut für Chemie, Universität Osnabrück, Barbarastrasse 7, D-49069 Osnabrück, Germany
Correspondence e-mail: fraseela@rz.uni-osnabrueck.de

Received 24 February 2000
Accepted 25 April 2000

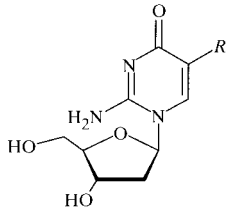
In the title compound, 2-amino-1-(2-deoxy-β-D-*erythro*-pentofuranosyl)-5-methylpyrimidin-4(1*H*)-one, $C_{10}H_{15}N_3O_4$, the conformation of the *N*-glycosidic bond is *syn* and the 2-deoxyribofuranose moiety adopts an unusual ${}^{O}T_1$ sugar pucker. The orientation of the exocyclic C4′—C5′ bond is +*sc* (+*gauche*).

Comment

Transposition of the amino and oxo groups in the canonical nucleic acid base cytosine results in isocytosine. This nucleobase can form a Watson–Crick base pair with isoguanine, forming a DNA with antiparallel (aps) chain orientation (Switzer *et al.*, 1989, 1993; Tor & Dervan, 1993; Roberts *et al.*, 1995; Horn *et al.*, 1995). More interestingly, isocytosine or 5-methylisocytosine can form a reverse Watson–Crick base pair with the canonical DNA base guanine, which then results in parallel (ps) DNA (Sugiyama *et al.*, 1996; Seela *et al.*, 1998; Seela, He & Wei, 1999; Seela, Wei *et al.*, 1999).

2′-Deoxyisocytidine, (I*a*), proved to be problematic in oligonucleotide synthesis. It is deaminated under the normal deprotection conditions (concentrated ammonia, elevated temperature; Switzer *et al.*, 1989), and it was found to be an acid-labile pyrimidine nucleoside (Roberts *et al.*, 1997; Dekker, 1960). Only when a rather labile protecting group is chosen for compound (I*a*), as well as for the other nucleosides used in the solid-phase synthesis, can intact oligonucleotides be isolated. Another approach that circumvents this problem was the use of the 5-methyl derivative of 2′-deoxyisocytosine [5miC_d, (I*b*)] as a substitute (Tor & Dervan, 1993; Jurczyk *et al.*, 1998; Bukowska & Kusmierek, 1996). Many oligonucleotides containing compound (I*b*) have been prepared by our group and the base-pairing properties were studied in parallel and in antiparallel DNA (Seela, He & Wei, 1999). As it has been recently shown that the base-pairing properties of (I*b*) are similar to those of (Ia), it was decided to elucidate the molecular structure of the title compound (I*b*).

The structure of (I*b*) is shown in Fig. 1. Similar to the significant nonplanarity of the pyrimidine ring of 2′-deoxy-5-methylcytidine [5mdC, (II); Sato, 1988], the nucleobase of (I*b*) is somewhat nonplanar. The deviations of its C and N atoms from the least-squares plane are N1 = −0.040 (3), C2 = 0.035 (3), N3 = 0.009 (3), C4 = −0.043 (3), C5 = 0.036 (3) and C6 = 0.005 (3) Å.



(I*a*) *R* = H
(I*b*) *R* = Me

(II)

Bond lengths and angles are summarized in Table 1. The glycosidic bond length (N1—C1′) of 1.478 (5) Å is 0.017 (8) Å longer than the corresponding bond length in (II). The most distinct feature of the conformation of (I*b*), different from other pyrimidine nucleosides, is the *syn* conformation of its glycosidic bond, which corresponds to a torsion angle χ_{CN} (O4′—C1′—N1—C2) = 58.2 (5)° (Table 1). The glycosidic torsion angle in pyrimidine nucleosides is generally found to be *anti*; very few *syn* structures are known (Saenger, 1989). For example, (II) has an *anti* glycosidic bond with χ_{CN} = −131.7°. 4-Thiouridine hydrate (Saenger & Scheit, 1970) was the first pyrimidine nucleoside found to exist in the *syn* conformation in the crystalline state, with χ_{CN} = 76.8°. 2′-Deoxy-2′-(*R*)-phenylsulfinyluridine (Hata *et al.*, 1991) also adopts a *syn* conformation, with χ_{CN} = 61.6°. The dihydro analogue of (I*b*), the ribonucleoside 5,6-dihydroisocytidine monohydrate, which also has the amino group at the 2-position, adopts the usual *anti* conformation, with χ_{CN} = 107.6° (Kojic-Prodic *et al.*, 1976). The reason for the preference of a *syn* conformer of (I*b*) is a bifurcated hydrogen bond formed between the 2-amino group of the base and atoms O4′ and O5′ of the sugar moiety. The 2-amino group is located above the 2′-deoxyribofuranose moiety, as shown in Fig. 1.

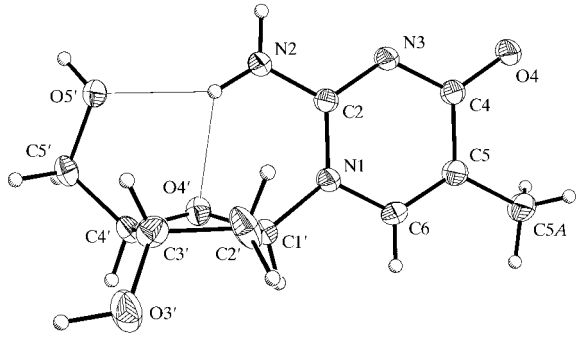


Figure 1
A perspective view of (I*b*) showing the bifurcated intramolecular hydrogen bond, which stabilizes the *syn* conformation of the glycosidic bond. Displacement ellipsoids are drawn at the 25% probability level and H atoms are shown as spheres of an arbitrary size.

The next major conformational parameter of interest is the pucker of the deoxyribofuranose moiety of (I*b*). Its phase angle of pseudorotation, *P*, is 102.4 (6)°. The following two points are noteworthy. Firstly, the *P* value of (I*b*) lies outside the preferred pseudorotation range of nucleosides: the C3′-*endo* domain is in the region *P* = 340–40° (North) and the C2′-*endo* domain is in the region *P* = 140–200° (South); thus, the *P* value of (I*b*) corresponds to a sugar pucker mode of ${}^{O}T_{1}$, which is also shown in Fig. 1. Secondly, in contrast with the standard sugars, where the average value of Ψ_m is 38.6 (3)° (Saenger, 1984), the maximum puckering amplitude in (I*b*) of only 26.1 (1)° is significantly smaller. This indicates that the sugar moiety of (I*b*) is flattened compared with those of the normal nucleoside sugar moieties. Such cases are also found in cyclic nucleosides, where the sugar conformation and the puckering amplitude are often constrained (Saenger, 1989). The conformation about the C5′—C4′ bond is +*sc* (+*gauche*) (Saenger, 1984), with a torsion angle γ (O5′—C5′—C4′—C3′) of 62.7 (4)°. For comparison, the sugar ring of (II) is puckered in a typical C2′-*endo* envelope form (*P* = 161.5°), with a normal maximum amplitude Ψ_m = 37.9°. Furthermore, the oxygen O5′ is located out of the sugar ring, with γ = 178.6° (Sato, 1988).

The hydrogen-bond pattern of (I*b*) is summarized in Table 2. The intramolecular bifurcated hydrogen bond (N2—H22···O5′ and N2—H22···O4′), which does not exist in molecule (II), is particularly important. This hydrogen bond stabilizes the *syn* conformation of the pyrimidine moiety and restricts the conformation of the sugar ring in a way similar to the cyclic nucleosides.

Beside this important intramolecular hydrogen bond, there are three intermolecular hydrogen bonds responsible for the packing of the molecules in (I*b*) (Table 2). The O3′—H3′···O4^{ii} bond connects molecules parallel to the *c* axis to form infinite chains [symmetry code: (ii) *x*, *y*, *z* − 1]. Three of these chains placed around the threefold axis form triple strands. Each molecule of the first chain of a strand forms O5′—H5′···O4^{iii} and N2—H21···O5′^{i} hydrogen bonds with molecules of the second chain [symmetry codes: (i) $\frac{2}{3} - y, \frac{4}{3} + x - y, \frac{1}{3} + z$; (iii) $\frac{2}{3} - y, \frac{4}{3} + x - y, z - \frac{2}{3}$]. At the same time, atoms O5′ and O4 accept hydrogen bonds from two molecules of the third chain. The second chain donates hydrogen bonds to the third chain. The triple strands are not connected to each other.

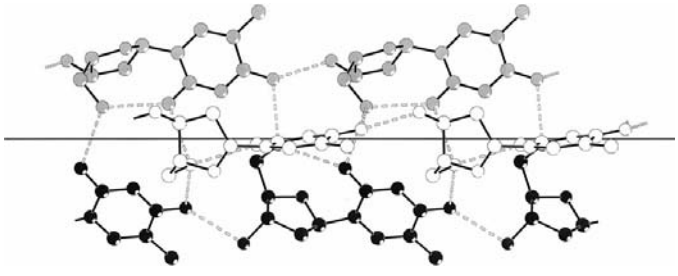


Figure 2
The arrangement of the molecules in (I*b*) around a crystallographic 3_1 axis. Intermolecular hydrogen bonds are represented by dashed lines. H atoms and intramolecular hydrogen bonds have been omitted for clarity.

Fig. 2 shows the crystal packing of (I*b*) along the threefold *c* axis.

Experimental

Compound (I*b*) was synthesized from 2,5′-anhydrothymidine (Watanabe *et al.*, 1978), according to the procedure reported by Kowollik & Langen (1968), and was crystallized from acetone/MeOH (8:2).

Crystal data

$C_{10}H_{15}N_3O_4$
M_r = 241.25
Trigonal, *R*3
a = 16.611 (3) Å
c = 10.6471 (13) Å
V = 2544.4 (7) Å^3
Z = 9
D_x = 1.417 Mg m^{-3}
Mo *K*α radiation
Cell parameters from 41 reflections
θ = 2.44–12.34°
μ = 0.111 mm^{-1}
T = 293 (2) K
Prism, colourless
0.51 × 0.17 × 0.11 mm

Data collection

Siemens *P*4 diffractometer
2θ/ω scans
Absorption correction: empirical (*SHELXTL*; Sheldrick, 1997)
T_{min} = 0.914, T_{max} = 0.988
2165 measured reflections
975 independent reflections
892 reflections with $I > 2\sigma(I)$
R_{int} = 0.037
θ_{max} = 24.96°
h = −17 → 17
k = −19 → 19
l = −12 → 12
3 standard reflections
every 97 reflections
intensity decay: none

Refinement

Refinement on F^2
$R[F^2 > 2\sigma(F^2)]$ = 0.039
$wR(F^2)$ = 0.096
S = 1.053
975 reflections
163 parameters
H atoms treated by a mixture of independent and constrained refinement
$w = 1/[\sigma^2(F_o^2) + (0.0489P)^2 + 2.6081P]$ where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max} < 0.001$
$\Delta\rho_{max}$ = 0.46 e Å^{-3}
$\Delta\rho_{min}$ = −0.21 e Å^{-3}
Extinction correction: *SHELXL*93 (Sheldrick, 1993)
Extinction coefficient: 0.0018 (5)

Table 1
Selected geometric parameters (Å, °).

N1—C2	1.378 (5)	C5—C5*A*	1.504 (5)
N1—C6	1.384 (5)	C1′—O4′	1.413 (5)
N1—C1′	1.478 (5)	C1′—C2′	1.523 (7)
C2—N3	1.316 (4)	C2′—C3′	1.496 (6)
C2—N2	1.328 (5)	C3′—O3′	1.412 (5)
N3—C4	1.358 (5)	C3′—C4′	1.539 (6)
C4—O4	1.249 (4)	C4′—O4′	1.444 (4)
C4—C5	1.445 (5)	C4′—C5′	1.505 (6)
C5—C6	1.339 (6)	C5′—O5′	1.420 (5)
C2—N1—C6	117.7 (3)	O4′—C1′—C2′	106.5 (3)
N3—C2—N1	122.3 (3)	C3′—C2′—C1′	106.2 (3)
C2—N3—C4	120.5 (3)	C2′—C3′—C4′	105.2 (3)
N3—C4—C5	119.4 (3)	O4′—C4′—C3′	106.5 (3)
C6—C5—C4	117.3 (3)	C1′—O4′—C4′	109.2 (3)
C5—C6—N1	122.2 (4)	O5′—C5′—C4′	108.9 (3)
C6—N1—C2—N3	−7.1 (6)	O4′—C1′—C2′—C3′	19.4 (5)
N1—C2—N3—C4	2.6 (6)	C1′—C2′—C3′—C4′	−5.6 (5)
C2—N3—C4—C5	4.8 (6)	C2′—C3′—C4′—O4′	−9.7 (5)
N3—C4—C5—C6	−7.4 (6)	C5′—C4′—C3′—O3′	112.3 (4)
C4—C5—C6—N1	2.8 (6)	C4′—O4′—C1′—C2′	−26.4 (4)
C5—C6—N1—C2	4.1 (6)	C3′—C4′—O4′—C1′	22.8 (4)
O4′—C1′—N1—C2	58.2 (5)	O5′—C5′—C4′—C3′	62.7 (4)

Table 2
Hydrogen-bonding geometry (Å, °).

D—H···A	D—H	H···A	D···A	D—H···A
N2—H21···O5′^{i}	0.88	2.12	2.977 (4)	167
N2—H22···O5′	0.90	2.06	2.864 (4)	149
N2—H22···O4′	0.90	2.18	2.814 (4)	127
O3′—H3′···O4^{ii}	1.09	1.75	2.744 (4)	149
O5′—H5′···O4^{iii}	0.95	1.78	2.694 (4)	161

Symmetry codes: (i) $\frac{2}{3}-y,\frac{4}{3}+x-y,\frac{1}{3}+z$; (ii) $x,y,z-1$; (iii) $\frac{2}{3}-y,\frac{4}{3}+x-y,z-\frac{2}{3}$.

In the absence of suitable anomalous scatterers, the measured Friedel data could not be used to determine the absolute structure. Therefore, Friedel reflections were merged. However, comparison with the known configuration of the parent molecule indicates that the proposed configuration is correct. All H atoms were located in difference Fourier syntheses. Nevertheless, all H atoms, except hydroxyl- and amino-H, were generated in idealized positions and refined using a riding model, with displacement parameters fixed at 1.5 times the (equivalent) isotropic displacement parameters of the parent atoms. The amino- and hydroxyl-H atom coordinates (as found in the difference Fourier map) were refined as riding on the parent N and O atoms, respectively. The N—H and O—H bond lengths were refined using an AFIX4 instruction.

Data collection: *XSCANS* (Siemens, 1996); cell refinement: *XSCANS*; data reduction: *SHELXTL* (Sheldrick, 1997); program(s) used to solve structure: *SHELXS*86 (Sheldrick, 1990); program(s) used to refine structure: *SHELXL*93 (Sheldrick, 1993); molecular graphics: *SHELXTL* and *DIAMOND* (Brandenburg, 1999); software used to prepare material for publication: *SHELXTL*.

Financial support by the Deutsche Forschungsgemeinschaft, Roche Diagnostics GmbH and the Fonds der Chemischen Industrie is gratefully acknowledged.

Supplementary data for this paper are available from the IUCr electronic archives (Reference: JZ1402). Services for accessing these data are described at the back of the journal.

References

Brandenburg, K. (1999). *DIAMOND*. Release 2.1*c*. Crystal Impact GbR, Bonn, Germany.
Bukowska, A. M. & Kusmierek, J. T. (1996). *Acta Biochim. Pol.* **43**, 247–254.
Dekker, C. A. (1960). *Ann. Rev. Biochem.* **29**, 453–474.
Hata, T., Matsuda, A. & Miyasaka, T. (1991). *Anal. Sci.* **6**, 987–988.
Horn, T., Chang, C. A. & Collins, M. L. (1995). *Tetrahedron Lett.* **36**, 2033–2036.
Jurczyk, S. C., Kodra, J. T., Rozzell, J. D., Benner, S. A. & Battersby, T. R. (1998). *Helv. Chim. Acta*, **81**, 793–811.
Kojic-Prodic, B., Ruzic Toros, Z. & Coffou, E. (1976). *Acta Cryst.* B**32**, 1103–1107.
Kowollik, G. & Langen, P. (1968). *J. Prakt. Chem.* **37**, 311–318.
Roberts, C., Bandaru, R. & Switzer, C. (1995). *Tetrahedron Lett.* **36**, 3601–3604.
Roberts, C., Bandaru, R. & Switzer, C. (1997). *J. Am. Chem. Soc.* **119**, 4640–4649.
Saenger, W. (1984). *Principles of Nucleic Acid Structure*, edited by C. R. Cantor, p. 55. New York: Springer-Verlag.
Saenger, W. (1989). *Landolt-Börnstein New Series*, Vol. 1, *Nucleic Acids*, Subvolume a, edited by O. Madelung, pp. 24–25. Berlin: Springer-Verlag.
Saenger, W. & Scheit, K. H. (1970). *J. Mol. Biol.* **50**, 153–169.
Sato, T. (1988). *Acta Cryst.* C**44**, 870–872.
Seela, F., He, Y. & Wei, C. (1999). *Tetrahedron*, **55**, 9481–9500.
Seela, F., Wei, C., Melenewski, A. & Feiling, E. (1998). *Nucleosides Nucleotides*, **17**, 2045–2052.
Seela, F., Wei, C., Melenewski, A., He, Y., Kröschel, R. & Feiling, E. (1999). *Nucleosides Nucleotides*, **18**, 1543–1548.
Sheldrick, G. M. (1990). *Acta Cryst.* A**46**, 467–473.
Sheldrick, G. M. (1993). *SHELXL*93. University of Göttingen, Germany.
Sheldrick, G. M. (1997). *SHELXTL*. Release 5.10. Bruker AXS Inc., Madison, Wisconsin, USA.
Siemens (1996). *XSCANS*. Release 2.2. Siemens Analytical X-ray Instruments Inc., Madison, Wisconsin, USA.
Sugiyama, H., Ikeda, S. & Saito, I. (1996). *J. Am. Chem. Soc.* **118**, 9994–9995.
Switzer, C., Moroney, S. E. & Benner, S. A. (1989). *J. Am. Chem. Soc.* **111**, 8322–8323.
Switzer, C. Y., Moroney, S. E. & Benner, S. A. (1993). *Biochemistry*, **32**, 10489–10496.
Tor, Y. & Dervan, P. B. (1993). *J. Am. Chem. Soc.* **115**, 4461–4467.
Watanabe, K. A., Reichman, U., Chu, C.-K. & Fox, J. J. (1978). *Nucleic Acid Chemistry*, Part 1, edited by R. S. Tipson & L. B. Townsend, pp. 273–277. New York: John Wiley and Sons.