Acta Crystallographica Section C
Crystal Structure Communications
ISSN 0108-2701

1-(β-D-Erythrofuranosyl)cytidine (β-erythrocytidine)

Paul C. Kline,[a] Bruce C. Noll,[b] Hongqiu Zhao[b] and Anthony S. Serianni[b]*

[a]Department of Chemistry, Middle Tennessee State University, Murfreesboro, TN 37132, USA, and [b]Department of Chemistry and Biochemistry, 251 Nieuwland Science Hall, University of Notre Dame, Notre Dame, IN 46556-5670, USA
Correspondence e-mail: serianni.1@nd.edu

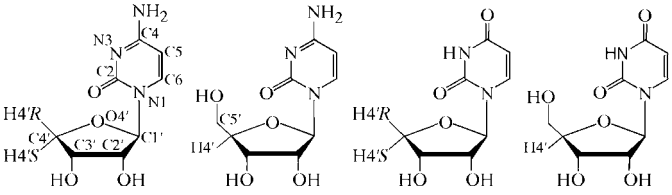
Received 16 October 2006
Accepted 18 December 2006
Online 23 January 2007

1-(β-D-Erythrofuranosyl)cytidine, $C_8H_{11}N_3O_4$, (I), a derivative of β-cytidine, (II), lacks an exocyclic hydroxymethyl ($-CH_2OH$) substituent at C4′ and crystallizes in a global conformation different from that observed for (II). In (I), the β-D-erythrofuranosyl ring assumes an E_3 conformation (C3′-*exo*; *S*, *i.e.* south), and the N-glycoside bond conformation is *syn*. In contrast, (II) contains a β-D-ribofuranosyl ring in a $^3T_2$ conformation (*N*, *i.e.* north) and an *anti*-N-glycoside linkage. These crystallographic properties mimic those found in aqueous solution by NMR with respect to furanose conformation. Removal of the $-CH_2OH$ group thus affects the global conformation of the aldofuranosyl ring. These results provide further support for *S*/*syn*–*anti* and *N*/*anti* correlations in pyrimidine nucleosides. The crystal structure of (I) was determined at 200 K.

Comment

Tetrofuranose-containing nucleosides and their phosphate esters have attracted attention recently as building blocks in the preparation of novel oligonucleotides (Schoning *et al.*, 2000; Kempeneers *et al.*, 2004). Oligonucleotides containing β-tetrofuranosyl rings, which lack the exocyclic C5′ hydroxymethyl group found in β-ribofuranosyl rings, are necessarily assembled *via* 2′→3′-phosphodiester linkages. The *trans*-O2′,O3′ configuration found in L-*threo* derivatives (TNA) mimics backbones formed from conventional 3′→5′ linkages (Wilds *et al.*, 2002), whereas the *cis*-O2′,O3′ configuration found in *erythro* building blocks presumably results in oligonucleotides with considerably different structures and topologies.

Nucleoside conformation in solution is affected significantly by removal of the furanose exocyclic hydroxymethyl functionality. NMR investigations of β-erythronucleosides have revealed very different J_{HH}, J_{CH} and J_{CC} values, indicative of a major shift in the preferred conformation of the furanose ring relative to that observed in the corresponding β-ribonucleosides (Kline & Serianni, 1992). For example, the endocyclic $^3J_{H1',H2'}$, $^3J_{H2',H3'}$ and $^3J_{H3',H4'}$ values are 5.6, 4.6 and 3.3 Hz (to H4′*S*) in β-erythrocytidine, (I), whereas the corresponding values in β-cytidine, (II), are 3.9, 5.3 and 6.0 Hz, respectively. Pseudorotational analysis (Rao *et al.*, 1981) of these *J* couplings reveals a highly preferred *S* (south) furanose conformation in (I) (79%; $P = 186.4°$ and $\tau_m = 40°$), whereas a similar analysis for (II) reveals a preferred *N* (north) conformation (58%; $P = 21.8°$ and $\tau_m = 38°$) (Fig. 1). The different preferred furanose conformations were expected to affect N-glycoside conformation in (I) and (II), but NMR data were unavailable to address this question.



A comparison of the low-temperature crystal structure of (I) determined here (Fig. 2) with that of (II) (Ward, 1993) shows differences in furanose conformation nearly identical to those observed in solution. Removal of the exocyclic $-CH_2OH$ group stabilizes *S* conformations, with E_3 (C3′-*exo*) preferred in (I) ($P = 197.8°$ and $\tau_m = 44°$) and $^3T_2$ preferred in (II) ($P = 7.7°$ and $\tau_m = 39°$) (Fig. 1 and Table 1). These results can be explained based on a preferred quasi-equatorial orientation of the bulkier $-CH_2OH$ substituent in (II), which is achievable in *N* conformations like $^3T_2$. While the pseudorotational phase angles, *P*, differ significantly, the puckering amplitudes, τ_m, are similar in (I) and (II) (Table 1). Endocyclic C1—C2 and C2—C3 bond lengths in (I) are larger by ∼0.01 Å, whereas $r_{C3,C4}$ in (I) and (II) are more similar (Δ = 0.004 Å) (Table 1).

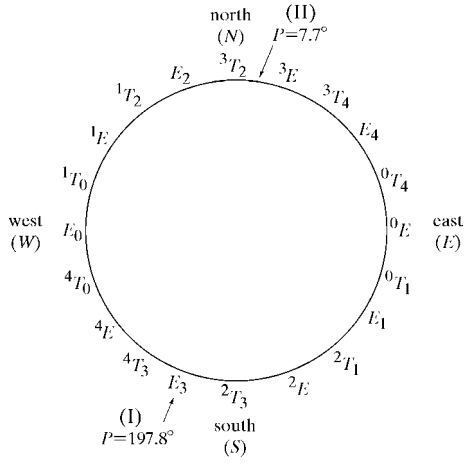


Figure 1
The pseudorotational itinerary of an aldofuranose ring. *E* and *T* denote envelope and twist forms, respectively. The preferred furanose conformations in crystalline (I) and (II) are highlighted.

 DOI: 10.1107/S0108270106055053

The change in the furanose conformation caused by removal of the $-CH_2OH$ substituent is accompanied by a significant change in the N-glycoside conformation. In (I), the nitrogen base is oriented *syn*/+*sc*, with the C2—O2 bond oriented above the furanose ring, corresponding to an O4′—C1′—N1—C2 torsion angle of 60.8 (2)° (Table 1). In contrast, the N-glycoside conformation is *anti*/*ap* in (II) [O4′—C1′—N1—C2 torsion angle = −162.6 (3)°], orienting the C6—H6 bond above the furanose ring; the C1′—N1 bond is rotated almost 180° in (II) relative to (I). A correlation between the N-glycoside and furanose conformations is thus evident (Saenger, 1984). ^3E conformations leave the top face of the furanose ring unhindered by the $-CH_2OH$ group, but the resulting quasi-axial orientation of the C3′—H3′ bond interferes with a *syn* N-glycoside conformation which would place atom O2 in close proximity with atom H3′. An *anti* N-glycoside conformation eliminates this apparently unfavourable arrangement. The E_3 conformation is achievable in (I) because it lacks the directing influence of an exocyclic $-CH_2OH$. In this ring geometry, the top face of the furanose ring is largely unhindered, allowing the more sterically demanding *syn* geometry about the N-glycoside.

Recent statistical analyses of nucleoside and nucleotide crystal structures (Gelbin *et al.*, 1996) have shown that *N* furanose conformations (C3′-*endo*) almost exclusively prefer *anti* N-glycosyl torsions, while *S* forms (C2′-*endo*) adopt both *anti* and *syn* geometries. Within the C2′-*endo*/*syn* group, virtually all structures were purine nucleosides. In this respect, (I) deviates from expectation. These results suggest that (I) could be useful in the development of NMR probes of N-glycoside conformation [*e.g.* DESERT experiments (Akasaka *et al.*, 1975), with either H1′ or H6 deuterated] or *trans*-glycoside $^3J_{C2',C2/6}$ studies (Kline & Serianni, 1990), potentially serving as a pyrimidine nucleoside reference structure in which a *syn* N-glycoside may be adopted to some extent in aqueous solution.

The crystal structure of β-erythrouridine, (III), has been reported recently (Czechtizky & Vasella, 2001), and some structural parameters are shown in Table 1. Like (I), the furanose conformation in (III) is E_3. By comparison, the crystal structure of β-uridine, (IV) (Green *et al.*, 1975), shows ^3E conformations for the β-ribofuranose ring in two independent forms. Thus, for pyrimidine nucleosides, the erythro- and ribofuranose rings show a consistent difference in their preferred conformation. However, the N-glycoside conformation in (III) is *anti*/−*ac* (O4′—C1′—N1—C2 = −131.4°) (Table 1), unlike that found for (I), and an *anti*/*ap* N-glycoside conformation is observed in the related compound (IV) (O4′—C1′—N1—C2 = −164.41 and −152.96° in the two forms). Thus, the *syn* conformation observed in (I) may be partly caused by crystal packing forces.

Since the furanose conformations are similar in (I) and (III), the differences in the structural parameters in these structures can be attributed mainly to different N-glycosyl conformations, especially for bond lengths and angles in the vicinity of the anomeric C atom. For example, $r_{C1',C2'}$ and the O4′—C1′—N1 bond angle are considerably larger in (I) (Table 1). The C1′—N1 and C1′—O4′ bond lengths, however, are similar in both structures. It is noteworthy that $r_{C2',O2'} < r_{C3',O3'}$ in both (I) and (III), whereas these bond lengths are identical in (II). Presumably, the quasi-axial and quasi-equatorial orientations of the C3′—O3′ and C2′—O2′ bonds, respectively, in (I) and (III) are responsible for this difference, with the former orientation expected to generate a longer bond (Podlasek *et al.*, 1996). The C—O bond lengths, however, will be modulated by differences in hydrogen bonding in the crystal structures, and these secondary effects may complicate interpretations based solely on bond orientation.

The hydrogen bonding in the crystal structures of (I) and (II) also differs. In (I), atoms O2′ and O3′ each serve as a donor and a mono-acceptor, and atom O4′ is not involved in hydrogen bonding. In (II), atoms O2′, O3′ and O4′ behave similarly, and atom O5′ serves only as a donor. In (III), however, atoms O2′ and O3′ serve only as acceptors, but like (I) and (II), atom O4′ is not involved in hydrogen bonding. In the nitrogen base, the C4′ NH_2 group serves as a donor in two hydrogen bonds, and atom N3 serves as a hydrogen-bond acceptor in both (I) and (II). The C2 carbonyl serves as a single acceptor in (I) and a double acceptor in (II). In (III), the C2 carbonyl serves as a single acceptor, the C4 carbonyl as a dual acceptor, and atom N3 as a donor. An interesting intermolecular hydrogen-bonding pattern is observed in (I), wherein atoms O2′ and O3′ of one molecule serve as donors to

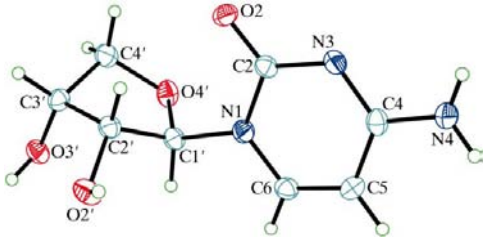


Figure 2
A plot of the molecule of (I). Displacement ellipsoids are drawn at the 50% probability level and H atoms are shown as small spheres of arbitrary radii.

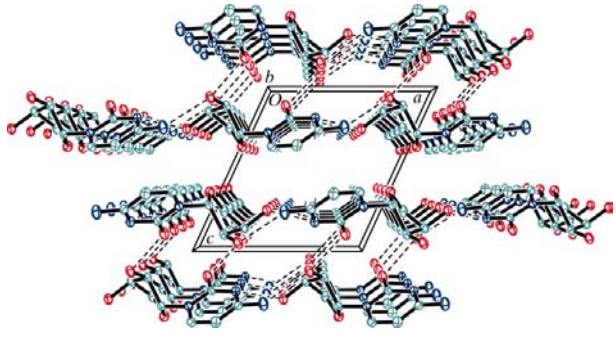


Figure 3
A packing diagram for (I), viewed down the *b* axis, illustrating the hydrogen-bonded bilayers. Dashed lines denote hydrogen bonds.

the C4′-NH_2 and N3 atoms of a second molecule in the crystal lattice. This hydrogen-bonding pattern suggests modes of recognition between nucleosides involving direct base–sugar interactions. This pattern is not observed in (II) and (III).

The structure of (I) forms a hydrogen-bonded bilayer about the *ab* plane (Fig. 3). Each sheet is formed *via* three types of hydrogen bonds, namely N4—H4*A*···O3′($x+1, y+1, z$), N4—H4*B*···O2′($x+1, y, z$) and O3′—H3′···N3($x-1, y, z$). The layers are joined by hydrogen bonding between O2′—H2′O and O2($-x, y+\frac{1}{2}, -z$) of opposite sheets. Details of the hydrogen bonding are summarized in Table 2.

Experimental

Compound (I) was prepared as described previously by Kline & Serianni (1992) and was crystallized from hot ethanol.

Crystal data

$C_8H_{11}N_3O_4$
M_r = 213.20
Monoclinic, $P2_1$
a = 8.5473 (10) Å
b = 6.2423 (7) Å
c = 9.1325 (11) Å
β = 114.701 (4)°
V = 442.68 (9) Å^3
Z = 2
D_x = 1.599 Mg m^{-3}
Cu $K\alpha$ radiation
μ = 1.11 mm^{-1}
T = 200 (2) K
Plate, colourless
0.30 × 0.20 × 0.03 mm

Data collection

Bruker SMART APEX CCD area-detector diffractometer
φ and ω scans
Absorption correction: multi-scan (Sheldrick, 2004)
T_{min} = 0.731, T_{max} = 0.967
5338 measured reflections
1269 independent reflections
1242 reflections with $I > 2\sigma(I)$
R_{int} = 0.023
θ_{max} = 69.6°

Refinement

Refinement on F^2
$R[F^2 > 2\sigma(F^2)]$ = 0.028
$wR(F^2)$ = 0.076
S = 1.09
1269 reflections
146 parameters
H atoms treated by a mixture of independent and constrained refinement
$w = 1/[\sigma^2(F_o^2) + (0.0428P)^2 + 0.1356P]$ where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max} < 0.001$
$\Delta\rho_{max}$ = 0.24 e Å^{-3}
$\Delta\rho_{min}$ = −0.21 e Å^{-3}

H atoms attached to carbon were placed at calculated geometries and allowed to ride on the position of the parent atom, with C—H = 0.95–1.00 Å. Hydroxy H atoms were located at the position that maximizes the electron density and this was coupled with a rotating-group refinement; O—H = 0.84 Å. The two H atoms of N4 were located in a difference map and refined freely, including a parameter for isotropic thermal motion, in subsequent cycles of least-squares refinement; N—H = 0.84 (3) and 0.87 (4) Å. For other H atoms, U_{iso}(H) = 1.2U_{eq}(C) or 1.5U_{eq}(O).

Data collection: *APEX2* (Bruker, 2006); cell refinement: *APEX2* and *SAINT* (Bruker, 2006); data reduction: *SAINT* and *XPREP* (Sheldrick, 2003); program(s) used to solve structure: *XS* (Sheldrick, 2001); program(s) used to refine structure: *XL* (Sheldrick, 2001); molecular graphics: *XP* (Sheldrick, 1998); software used to prepare material for publication: *enCIFer* (Allen *et al.*, 2004).

The diffractometer was purchased using funds provided by the National Science Foundation (award No. CHE-0443233).

Table 1
Comparison of structural parameters (Å, °) in compounds (I)–(III).

	(I)†	(II)‡	(III)§
C1′—C2′	1.548 (3)	1.534 (5)	1.527
C2′—C3′	1.528 (3)	1.518 (5)	1.525
C3′—C4′	1.512 (3)	1.515 (5)	1.499
C1′—N1	1.475 (2)	1.490 (4)	1.474
C1′—O4′	1.414 (3)	1.409 (4)	1.409
C4′—O4′	1.452 (3)	1.461 (4)	1.445
C2′—O2′	1.414 (3)	1.412 (4)	1.398
C3′—O3′	1.422 (2)	1.412 (4)	1.417
C2—O2/C4—O4	1.234 (3)	1.241 (4)	1.223/1.232
C4—N4	1.335 (3)	1.319 (5)	
C4′—O4′—C1′	108.24 (14)	110.2 (3)	108.7
O4′—C1′—N1	110.56 (16)	108.8 (3)	108.1
O4′—C1′—C2′	107.04 (16)	107.0 (3)	107.6
C1′—C2′—C3′	100.67 (15)	101.1 (3)	102.4
C2′—C3′—C4′	99.78 (15)	102.8 (3)	101.0
C3′—C4′—O4′	104.69 (18)	104.1 (3)	105.9
C1′—N1—C2	121.52 (18)	117.2 (3)	119.1
C1′—N1—C6	117.92 (17)	122.7 (3)	119.1
O4′—C1′—N1—C2	60.8 (2) (*syn*; *+sc*)	−162.6 (3) (*anti*; *ap*)	−131.4 (*anti*; *−ac*)
O4′—C1′—N1—C6	−120.29 (18)	18.1 (4)	46.8
C2′—C1′—N1—C2	−60.6 (3)	79.5 (4)	109.0
C2′—C1′—N1—C6	118.28 (19)	−99.7 (4)	−72.8
C1′—C2′—C3′—C4′	−40.83 (18)	37.6 (3)	−35.1
C2′—C3′—C4′—O4′	42.1	−34.3 (3)	36.6
C3′—C4′—O4′—C1′	26.19 (19)	17.0 (4)	−23.5
C4′—O4′—C1′—C2′	−0.97 (18)	7.3 (4)	0.04
O4′—C1′—C2′—C3′	27.00 (18)	−28.3 (3)	22.6
N1—C2—N3—C4	−4.6 (3)	5.2 (5)	3.9
N1—C6—C5—C4	−2.9 (3)	0.9 (6)	−1.5
Furanose			
P (°)	197.8 (E_3; C3′-*exo*; *S*)	7.7 ($^3T_2$; *N*)	199.2 (E_3; C3′-*exo*; *S*)
τ_m (°)	44	39	38

† Erythrocytidine (this work). ‡ Cytidine (Ward, 1993). § Erythrouridine (Czechtizky & Vasella, 2001).

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

D—H···A	D—H	H···A	D···A	D—H···A
N4—H4*A*···O3′^{i}	0.87 (4)	2.20 (4)	2.982 (3)	150 (3)
N4—H4*B*···O2′^{ii}	0.84 (3)	2.08 (3)	2.912 (3)	177 (2)
O2′—H2′O···O2^{iii}	0.84	1.88	2.699 (2)	166
O3′—H3′O···N3^{iv}	0.84	2.06	2.825 (2)	151

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

Supplementary data for this paper are available from the IUCr electronic archives (Reference: GZ3050). Services for accessing these data are described at the back of the journal.

References

Akasaka, K., Imoto, T., Shibata, S. & Hatano, H. (1975). *J. Magn. Reson.* **18**, 328–343.
Allen, F. H., Johnson, O., Shields, G. P., Smith, B. R. & Towler, M. (2004). *J. Appl. Cryst.* **37**, 335–338.
Bruker (2006). *APEX2* and *SAINT*. Bruker AXS Inc., Madison, Wisconsin, USA.
Czechtizky, W. & Vasella, A. (2001). *Helv. Chim. Acta*, **84**, 1000–1016.

Flack, H. D. (1983). *Acta Cryst.* A**39**, 876–881.
Gelbin, A., Schneider, B., Clowney, L., Hsieh, S.-H., Olson, W. K. & Berman, H. M. (1996). *J. Am. Chem. Soc.* **118**, 519–529.
Green, E. A., Rosenstein, R. D., Shiono, R., Abraham, D. J., Trus, B. L. & Marsh, R. E. (1975). *Acta Cryst.* B**31**, 102–107. (Erratum: *Acta Cryst.* B**31**, 1221.)
Kempeneers, V., Froeyen, M., Vastmans, K. & Herdewijn, P. (2004). *Chem. Biodivers.* **1**, 112–123.
Kline, P. C. & Serianni, A. S. (1990). *J. Am. Chem. Soc.* **112**, 7373–7381.
Kline, P. C. & Serianni, A. S. (1992). *J. Org. Chem.* **57**, 1772–1777.
Podlasek, C. A., Stripe, W. A., Carmichael, I., Shang, M., Basu, B. & Serianni, A. S. (1996). *J. Am. Chem. Soc.* **118**, 1413–1425.
Rao, S. T., Westhof, E. & Sundaralingam, M. (1981). *Acta Cryst.* A**37**, 421–425.
Saenger, W. (1984). *Principles of Nucleic Acid Structure*, pp. 69–78. New York: Springer-Verlag.
Schoning, K. U., Scholz, P., Guntha, S., Wu, X., Krishnamurthy, R. & Eschenmoser, A. (2000). *Science*, **290**, 1347–1351.
Sheldrick, G. M. (1998). *XP*. Bruker AXS Inc., Madison, Wisconsin, USA.
Sheldrick, G. M. (2001). *XL* and *XS*. Bruker AXS Inc., Madison, Wisconsin, USA.
Sheldrick, G. M. (2003). *XPREP*. Bruker AXS Inc., Madison, Wisconsin, USA.
Sheldrick, G. M. (2004). *SADABS*. University of Göttingen, Germany.
Ward, D. L. (1993). *Acta Cryst.* C**49**, 1789–1792.
Wilds, C. J., Wawrzak, Z., Krishnamurthy, R., Eschenmoser, A. & Egli, M. (2002). *J. Am. Chem. Soc.* **124**, 13716–13721.