Acta Cryst C STRUCTURAL CHEMISTRY
ISSN 2053-2296

The α-D-anomer of 2′-deoxycytidine: crystal structure, nucleoside conformation and Hirshfeld surface analysis

Simone Budow-Busse,[a] Yingying Chai,[a] Sebastian Lars Müller,[a] Constantin Daniliuc[b] and Frank Seela[a,c]*

[a]Laboratory of Bioorganic Chemistry and Chemical Biology, Center for Nanotechnology, Heisenbergstrasse 11, 48149 Münster, Germany, [b]Organisch-Chemisches Institut, Westfälische Wilhelms-Universität Münster, Corrensstrasse 40, 48149 Münster, Germany, and [c]Laboratorium für Organische und Bioorganische Chemie, Institut für Chemie neuer Materialien, Universität Osnabrück, Barbarastrasse 7, Osnabrück 49069, Germany. *Correspondence e-mail: frank.seela@uni-osnabrueck.de

Received 24 February 2021
Accepted 30 March 2021

Edited by I. Oswald, University of Strathclyde, United Kingdom

Keywords: α-2′-deoxycytidine; crystal structure; crystal packing; Hirshfeld surface analysis; anomer; nucleic acid chemistry.

CCDC reference: 1938631

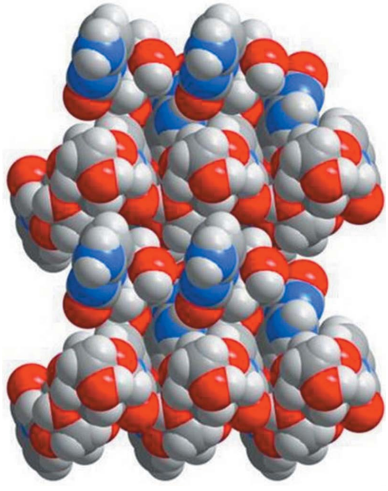
Supporting information: this article has supporting information at journals.iucr.org/c

β-2′-Deoxyribonucleosides are the constituents of nucleic acids, whereas their anomeric α-analogues are rarely found in nature. Moreover, not much information is available on the structural and conformational parameters of α-2′-deoxyribonucleosides. This study reports on the single-crystal X-ray structure of α-2′-deoxycytidine, $C_9H_{13}N_3O_4$ (**1**), and the conformational parameters characterizing **1** were determined. The conformation at the glycosylic bond is *anti*, with χ = 173.4 (2)°, and the sugar residue adopts an almost symmetrical C2′-*endo*-C3′-*exo* twist (^2_3T; *S*-type), with P = 179.7°. Both values lie outside the conformational range usually preferred by α-nucleosides. In addition, the amino group at the nucleobase shows partial double-bond character. This is supported by two separated signals for the amino protons in the ^{1}H NMR spectrum, indicating a hindered rotation around the C4—N4 bond and a relatively short C—N bond in the solid state. Crystal packing is controlled by N—H···O and O—H···O contacts between the nucleobase and sugar moieties. Moreover, two weak C—H···N contacts (C1′—H1′ and C3′—H3′*A*) are observed. A Hirshfeld surface analysis was carried out and the results support the intermolecular interactions observed by the X-ray analysis.

1. Introduction

Nucleosides with an α-configuration at the anomeric carbon are seldom found in nature (Ni *et al.*, 2019). α-Nucleosides are not building blocks of naturally occurring DNA or RNA. However, α-nucleosides have been isolated as constituents of small molecules in living cells, such as vitamin B_{12} (Bonnett, 1963) or a nicotinamide adenine dinucleotide (NAD) derivative isolated from *Azobacter vinelandii* (Suzuki *et al.*, 1965). Also, chemical nucleoside synthesis yields α-nucleosides together with the β-anomers in ratios depending on the structures of the starting materials and the experimental conditions. Protocols were developed for the stereoselective synthesis of α-D nucleosides or by anomerization of β-D anomers. This topic has been reviewed recently by Ni *et al.* (2019).

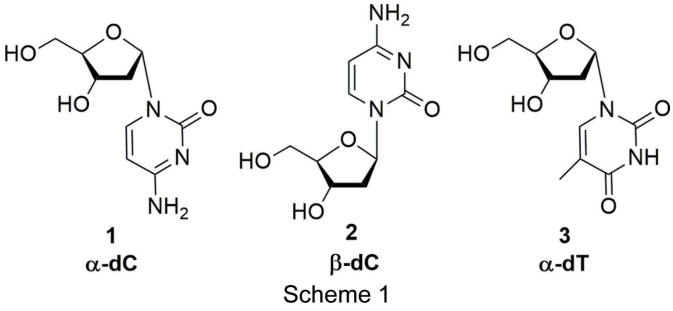
α-Nucleosides were also incorporated into oligonucleotides, replacing single β-nucleosides (Guo & Seela, 2017), or α-oligonucleotides were constructed which are entirely composed of α-nucleosides (Morvan *et al.*, 1990). α-Oligonucleotides form duplexes with an antiparallel orientation, with complementary strands also having an α-configuration (Morvan *et al.*, 1987*a*), while duplexes with a parallel alignment are formed when the



complementary strand is a β-oligonucleotide (Morvan *et al.*, 1987*b*).

Recently, we reported on the stability and recognition of silver-mediated heterochiral DNA with complementary α/β-strands (Chai *et al.*, 2020). Also, silver-mediated homochiral duplexes were constructed in which single residues were replaced by α-dC (**1**) (Scheme 1) (Guo & Seela, 2017). The silver-mediated base pair formed by anomeric α-dC (**1**) with β-dC (**2**) shows significantly higher stability than that formed by the silver-mediated β-dC–β-dC pair. Not only the stability of the metal-mediated base pair is higher, but also the metal-free α-dC–β-dC mismatch is more stable.

Conformational studies on α-nucleosides revealed distinct differences compared to their β-anomeric counterparts (Sundaralingam, 1971; Latha & Yathindra, 1992). In general, the flexibility around the glycosylic linkage, as well as the sugar pucker of α-nucleosides, seems to be more restricted than for β-nucleosides. The different conformational properties of the α/β-anomers were attributed to the differences in the steric interactions between the nucleobase and the sugar moiety (Sundaralingam, 1971; Latha & Yathindra, 1992). However, compared to the number of X-ray analyses of β-nucleosides, studies on α-nucleosides are extremely limited (Sundaralingam, 1971; Latha & Yathindra, 1992). Surprisingly, among the canonical α-nucleosides, only the solid-state conformations of α-cytidine (Post *et al.*, 1977) and α-2′-deoxythymidine (**3**) (Görbitz *et al.*, 2005) have been reported. Moreover, some X-ray studies on modified α-nucleoside analogues have been reported, *e.g.* α-5-acetyl-2′-deoxyuridine (Hamor *et al.*, 1977), α-5-aza-7-deaza-2′-deoxyguanosine (Seela *et al.*, 2002), α-5-iodo-2′-deoxycytidine (Müller *et al.*, 2019) and α-5-octadiynyl-2′-deoxycytidine (Zhou *et al.*, 2019). Among the α-2′-deoxyribonucleosides, the solid-state conformations of the α-anomers of 2′-deoxycytidine (**1**), 2′-deoxyadenosine and 2′-deoxyguanosine are still unknown.



Scheme 1

The single-crystal X-ray analysis of α-2′-deoxycytidine (**1**) was performed in order to obtain a deeper insight into the conformational properties of **1** in the solid-state. This is the second report of an α-anomer of a canonical pyrimidine 2′-deoxyribonucleoside besides α-2′-deoxythymidine (**3**) (Görbitz *et al.*, 2005). The results are compared to the structure of β-dC (Young & Wilson, 1975). For both **1** and **2**, the sugar conformation in solution was determined using a ^{1}H NMR-based method. Moreover, a Hirshfeld surface analysis of **1** was carried out to visualize the packing interactions.

Table 1
Experimental details.

Crystal data	
Chemical formula	$C_9H_{13}N_3O_4$
M_r	227.22
Crystal system, space group	Orthorhombic, $P2_12_12_1$
Temperature (K)	100
a, b, c (Å)	6.8378 (4), 11.4334 (7), 12.7595 (8)
V (Å^3)	997.53 (11)
Z	4
Radiation type	Cu $K\alpha$
μ (mm^{-1})	1.02
Crystal size (mm)	0.22 × 0.18 × 0.16
Data collection	
Diffractometer	Bruker APEXII Kappa CCD
Absorption correction	Multi-scan (*SADABS*; Bruker, 2014)
T_{min}, T_{max}	0.75, 0.85
No. of measured, independent and observed [$I > 2\sigma(I)$] reflections	11824, 1768, 1719
R_{int}	0.037
(sin $\theta/\lambda)_{max}$ (Å^{-1})	0.596
Refinement	
$R[F^2 > 2\sigma(F^2)]$, $wR(F^2)$, S	0.025, 0.061, 1.12
No. of reflections	1768
No. of parameters	161
H-atom treatment	H atoms treated by a mixture of independent and constrained refinement
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.12, −0.18
Absolute structure	Flack x determined using 684 quotients $[(I^+) - (I^-)]/[(I^+) + (I^-)]$ (Parsons *et al.*, 2013)
Absolute structure parameter	0.04 (9)

Computer programs: *SAINT* (Bruker, 2015), *SHELXS2014* (Sheldrick, 2015*a*), *SHELXL2014* (Sheldrick, 2015*b*), *APEX3* (Bruker, 2016) and *XP* (Bruker, 1998).

2. Experimental

2.1. Synthesis and crystallization of α-dC (1)

α-2′-Deoxycytidine (**1**) was synthesized as reported previously (Chai *et al.*, 2019). For crystallization, compound **1** was dissolved in methanol containing 10% water (10 mg in 1 ml) and was obtained as colourless prisms (m.p. 203–204 °C; Yamaguchi & Saneyoshi, 1984) by slow evaporation of the solvent at room temperature. A colourless prism-like specimen of **1** was used for the X-ray crystallographic analysis.

2.2. Refinement

Crystal data, data collection and structure refinement details are summarized in Table 1. The H atoms on N4, O3′ and O5′ were refined freely.

3. Results and discussion

3.1. Molecular geometry and conformation of α-dC (1)

The three-dimensional (3D) structure of α-dC (**1**) is shown in Fig. 1 and selected geometric parameters are presented in Table 2. The 3D structure of **1** clearly indicates the α-orientation of the nucleobase (Fig. 1), which in addition is supported by the Flack parameter (see Table 1; Parsons *et al.*, 2013). Moreover, according to the synthetic pathway, the

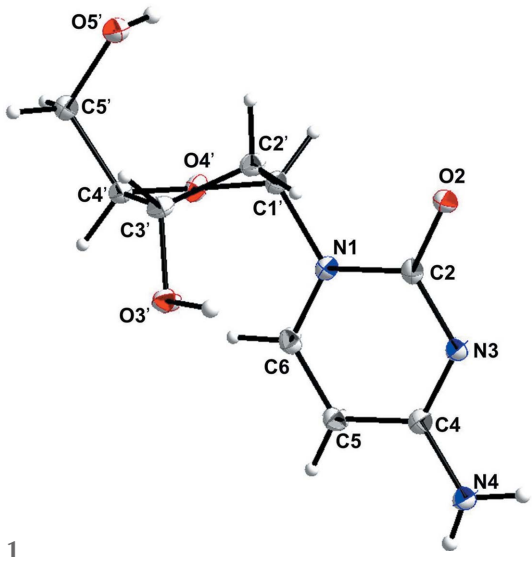


Figure 1
Perspective view of the α-D anomer of 2′-deoxycytidine (**1**), showing the atomic numbering scheme. Displacement ellipsoids are drawn at the 50% probability level and H atoms are shown as small spheres of arbitrary size.

anomeric centre at C1′ shows an *S*-configuration, confirming the α-D anomeric structure of **1**.

The crystal structure of the related canonical β-2′-deoxycytidine (**2**) has been reported previously (Young & Wilson, 1975). The β-anomer **2** shows two conformers (**2a** and **2b**) in the unit cell. As not many single-crystal X-ray analyses of α-2′-deoxyribonucleosides exist, it was of interest to compare the geometric parameters of the α/β-anomers of 2′-deoxycytidine (**1** and **2**).

Table 2
Selected geometric parameters (Å, °).

N1—C1′	1.499 (2)	N4—C4	1.337 (3)
C6—N1—C1′	122.33 (16)	O5′—C5′—C4′	112.26 (16)
C2—N1—C1′	117.08 (15)	C1′—O4′—C4′	110.96 (15)
C4—N3—C2—O2	−178.13 (18)	C1′—C2′—C3′—C4′	−33.03 (18)
N4—C4—C5—C6	175.05 (18)	C3′—C4′—C5′—O5′	55.9 (2)
C2—N1—C1′—O4′	173.39 (16)		

For pyrimidine nucleosides, the orientation of the nucleobase with respect to the sugar moiety (*syn*/*anti*) is defined by the torsion angle χ (O4′—C1′—N1—C2) (IUPAC–IUB Joint Commission on Biochemical Nomenclature, 1983). In the *anti* conformation, atom O2 of the six-membered ring is pointing away from the sugar, while in the *syn* conformation, O2 is pointing towards the sugar ring (Saenger, 1984). The preferred conformation for canonical pyrimidine β-2′-deoxyribonucleosides, including β-dC (**2a** and **2b**), is *anti* (**2a**: χ = 201.2°; **2b**: χ = 222.2°) (Young & Wilson, 1975). In contrast to the broad range of *anti* conformations adopted by β-nucleosides, a rather narrow preferred *anti* range, together with a preference of χ to adopt lower *anti* values, has been reported for α-nucleosides (Latha & Yathindra, 1992). For instance, α-2′-deoxythymidine (**3**) adopts a χ value of 124° (Görbitz *et al.*, 2005). However, in case of the title compound α-2′-deoxycytidine (**1**), an *anti* conformation with χ =

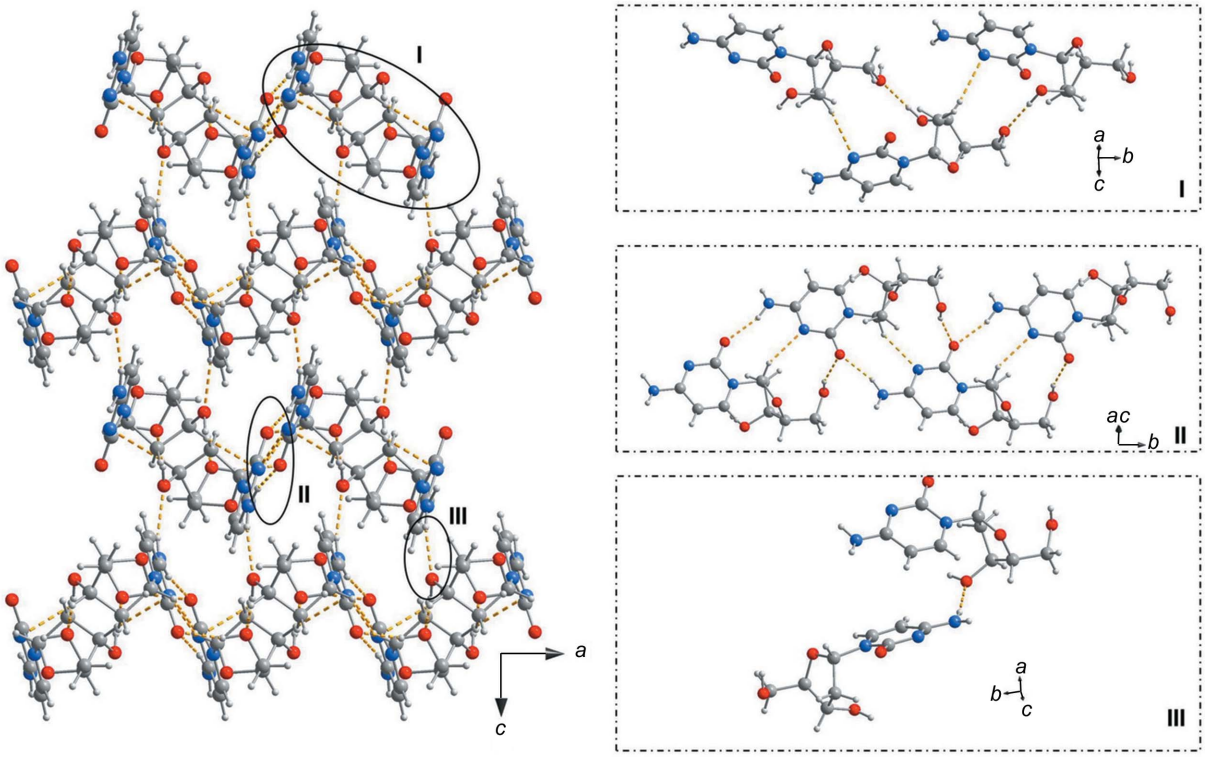


Figure 2
Crystal packing of α-2′-deoxycytidine (**1**), shown along the *ac* plane (ball-and-stick model), and with magnifications of designated areas of the crystal packing, showing the hydrogen-bonding pattern.

Table 3
Hydrogen-bond geometry (Å, °).

D—H···A	D—H	H···A	D···A	D—H···A
N4—H4A···O3′^{i}	0.86 (3)	2.10 (3)	2.949 (2)	168 (2)
N4—H4B···O2^{ii}	0.93 (3)	2.05 (3)	2.937 (2)	159 (2)
C1′—H1···N3^{iii}	1.0	2.46	3.317 (3)	144
C3′—H3A···N3^{iv}	1.0	2.53	3.524 (3)	170
O3′—H3···O5′^{v}	0.94 (3)	1.86 (4)	2.793 (2)	175 (3)
O5′—H5···O2^{iii}	0.91 (4)	1.88 (3)	2.716 (2)	152 (3)

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

173.39 (16)° is observed which is significantly greater. In addition, other solid-state structures of modified pyrimidine α-2′-deoxyribonucleosides with χ values around 168° have been reported recently, which also fall into this range (Zhou *et al.*, 2019; Müller *et al.*, 2019).

The second conformational parameter of interest for nucleosides is the sugar puckering mode. The β-2′-deoxyribofuranosyl moiety shows a preference for two principal sugar conformations, namely C3′-*endo* (*N*) and C2′-*endo* (*S*) (Altona & Sundaralingam, 1972; Sundaralingam, 1971). In contrast, studies on α-2′-deoxy- and ribonucleosides showed that these anomers prefer mainly C3′-*exo*, C2′-*exo* and C4′-*endo* conformations (Latha & Yathindra, 1992; Sundaralingam, 1971). As can be seen in Fig. 1, the sugar moiety of α-dC (**1**) adopts an almost symmetrical C2′-*endo*-C3′-*exo* twist ($^{2}_{3}T$; *S*-type), with a pseudorotational phase angle P = 179.7° and a maximum amplitude τ_m = 33.0° (Altona & Sundaralingam, 1972; Saenger, 1984). Thus, the α-2′-deoxyribofuranosyl moiety of **1** exhibits a C2′-*endo* conformation which is outside the preferred conformational range of α-2′-deoxyribonucleosides. Other examples of α-2′-deoxyribonucleosides with a C2′-*endo* conformation of the sugar residue include α-5-acetyl-2′-deoxyuridine (Hamor *et al.*, 1977) and the α-anomer of 5-aza-7-deaza-2′-deoxyguanosine (Seela *et al.*, 2002). For comparison, the conformers of the canonical β-dC (**2**) exhibit two different conformations, namely C3′-*endo* (*N*-type) for conformer **2a** and C2′-*endo* (*S*-type) for conformer **2b** (Young & Wilson, 1975).

The torsion angle γ (O5′—C5′—C4′—C3′) characterizes the orientation of the exocyclic 5′-hydroxy group relative to the sugar ring (Saenger, 1984). Earlier studies on α-nucleosides indicate that the conformational preference about the C4′—C5′ bond is similar to that of β-nucleosides (Sundaralingam, 1971). For α-dC (**1**), γ is 55.9 (2)°, referring to a +*sc* (*gauche*, *gauche*) conformation which is similar to that found for β-dC (**2**) (56.7 and 62.5°; +*sc*, *gauche*, *gauche*) (Young & Wilson, 1975).

3.2. Hydrogen bonding and molecular packing of α-dC (1)

Fig. 2 displays the crystal packing mode and hydrogen-bonding pattern for the crystal of α-dC (**1**). The corresponding hydrogen-bonding data and symmetry codes are summarized

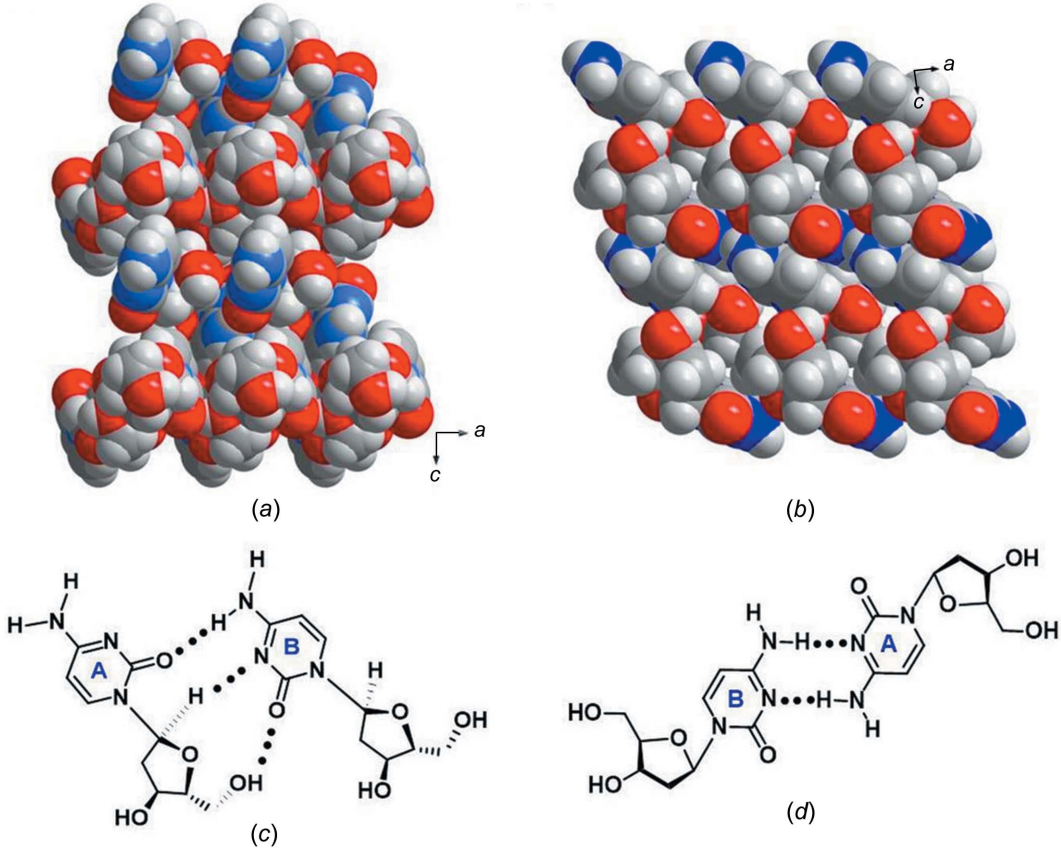


Figure 3
Space-filling models of (*a*) α-dC (**1**) and (*b*) β-dC (**2**). Schematic view of the intermolecular hydrogen-bonding interactions of two nucleosides for (*c*) α-dC (**1**) and (*d*) β-dC (**2**).

Table 4
^{1}H NMR chemical shifts, proton–proton vicinal and geminal coupling constants, and the conformation of nucleoside sugar residues in solution.

	Chemical shift/ppm											
	H-1′	H-2′	H-2″	H-3′	H-4′	H-5′	H-5″	3′-OH	5′-OH	NH	H5	H6
1	6.04 (*dd*)	2.50 (*m*)	1.81 (*dt*)	4.18 (*td*)	4.12 (*td*)	3.38 (*m*)	3.38 (*m*)	5.18 (*d*)	4.82 (*t*)	6.99 (*s*) 7.07 (*s*)	5.69 (*d*)	7.74 (*d*)
2	6.15 (*dd*)	2.10 (*ddd*)	1.92 (*ddd*)	4.19 (*ddd*)	3.76 (*td*)	3.54 (*qdd*)	3.54 (*qdd*)	5.18 (*d*)	4.95 (*t*)	7.10 (*s*) 7.15 (*s*)	5.71 (*d*)	7.78 (*d*)

	Coupling constant/Hz [*J*(H,H)]									Conformation	
	1′2′	1′2″	2′2″	2′3′	2″3′	3′4′	4′5′	4′5″	5′5″	%*N*	%*S*
1	7.5	2.8	−14.1	5.4	2.3	2.1	4.8	4.8	–	21	79
2	7.6	6.0	−13.3	6.0	3.2	3.1	4.0	4.0	−11.8	28	72

Measured in DMSO-d_6 at 298 K; r.m.s. < 0.4 Hz. H-2′ = H-2′$_\beta$; H-2″ = H-2′$_\alpha$. For *PSEUROT* (Van Wijk *et al.*, 1999) calculations, the coupling constants ^3J(H1′–H2′), ^3J(H1′–H2″), ^3J(H2′–H3′), ^3J(H2″–H3′) and ^3J(H3′–H4′) were used

in Table 3. The particular nucleoside units of **1** are connected by hydrogen bonds between (i) the nucleobases, (ii) nucleobases and sugars, as well as (iii) two sugar moieties. The crystal structure is formed by a repetition of nucleoside units which are arranged in chains in a zigzag-like manner within the *ac* plane (Fig. 2). This arrangement is different to that of the canonical *β*-dC (**2**). Space-filling models of *α*-dC (**1**) and *β*-dC (**2**) shown in Figs. 3(*a*) and 3(*b*), respectively, visualize the different crystal packing modes.

In more detail, two cytosine residues of *β*-dC (**2**) (Young & Wilson, 1975) form a mismatch connected by two hydrogen bonds, each between atom N3 and the 4-amino group of the respective second molecule (Fig. 3*d*). Fig. 3(*c*) shows that the situation is completely different for the *α*-anomer of 2′-deoxycytidine (**1**). Only one hydrogen bond is formed between the nucleobases (N4—H4*B*···O2^{ii}) (Fig. 2, motif II, and Table 3). The chains are further stabilized by a nucleobase-to-sugar contact (O5′—H5···O2^{iii}; Fig. 2, motif II, and Table 3) and a sugar-to-sugar contact (O3′—H3′···O5′^{v}; Fig. 2, motif I, and Table 3). In addition, two weak C—H···N contacts with C1′—H1′ (motif II) and C3′—H3′*A* (motif I) as the hydrogen-bond donors and N3 as the acceptor (N3^{iii} and N3^{iv}, respectively) are observed. In most crystal structures of 2′-deoxyribonucleosides, the C—H groups of the sugar moiety do not participate as hydrogen-bond donors in hydrogen bonding. However, a few examples have been reported, *e.g.* *α*-5-iodo-2′-deoxycytidine (Müller *et al.*, 2019) and 7-iodo-5-aza-7-deazaguanosine (Kondhare *et al.*, 2020). Moreover, two neighbouring chains are connected by a N4—H4*A*···O3′ hydrogen bond, as illustrated in Fig. 2 (motif III), thereby generating a network.

3.3. Hirshfeld surface analysis of *α*-dC (1)

To visualize the intermolecular interactions of *α*-dC (**1**) in the solid-state, a Hirshfeld surface analysis was conducted and two-dimensional (2D) fingerprint plots were analysed (Spackman & Jayatilaka, 2009). The *CrystalExplorer* program

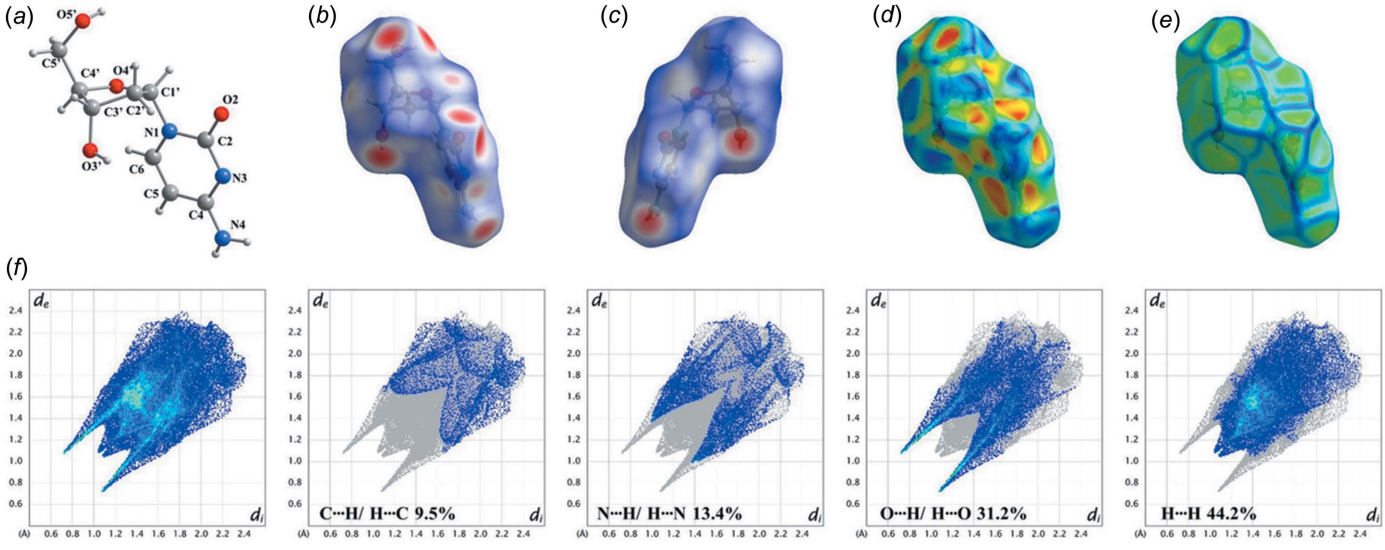


Figure 4
(*a*) Perspective view of *α*-dC (**1**), showing the atomic numbering scheme. The Hirshfeld surface of **1** mapped with (*b*) d_{norm} (−0.5 to 1.5, front view), (*c*) d_{norm} (−0.5 to 1.5, back view), (*d*) shape index and (*e*) curvedness, and (*f*) the corresponding fingerprint plots. Full interactions (left) and the resolved contacts (left, C···H/H···C; middle left, N···H/H···N; middle right, O···H/H···O; right, H···H) are shown, together with the percentages of their contribution to the total Hirshfeld surface area of *α*-anomer **1**.

(Version 17; Spackman & Jayatilaka, 2009; Turner *et al.*, 2017) was used to carry out the Hirshfeld surface analysis mapped over a d_{norm} range from −0.5 to 1.5 Å, shape index (−1.0 to 1.0 Å) and curvedness (−4.0 to 0.4 Å), as well as their associated 2D fingerprint plots (Fig. 4). On the d_{norm} surface of α-dC (**1**), several red areas (intense red spots) are observed (Figs. 4*b* and 4*c*), corresponding to the close contacts of the nucleobase and sugar residue (N—H···O and O—H···O). These interactions are shorter than the sum of the van der Waals radii and show negative d_{norm}. Small and light-red coloured spots are also found (Fig. 4*b*) and can be assigned to the weak contacts with C1′—H1′ and C3′—H3′*A* as hydrogen-bond donors and N3 as acceptor. The results of the Hirshfeld analyses are consistent with the hydrogen-bonding data (Table 3). The shape index (Fig. 4*d*) indicates π–π stacking interactions by the presence of red and blue triangles, and flat surface patches within the curvedness surfaces (Fig. 4*e*) are characteristic for planar stacking. However, as also indicated by Fig. 2, these interactions are less pronounced in the crystal structure of α-dC (**1**).

The 2D fingerprint plots provide a visual summary of the intermolecular contacts in the crystal structure of **1** and can be resolved to particular atom-pair interactions and their relative contributions to the Hirshfeld surface, as illustrated in Fig. 4(*f*). Strong interactions are found for O···H/H···O (31.2%) and N···H/H···N (13.4%), which agrees with the fact that the crystal packing of α-dC (**1**) is largely controlled by N—H···O and O—H···O hydrogen bonds (Table 3).

3.4. Conformation of the α- and β-anomers of 2′-deoxycytidine in solution

For canonical nucleosides with a β-D configuration, numerous studies exist describing their crystal structures and conformation in the solid-state and in solution. Compared to the information available for β-anomeric nucleosides, reports on α-anomers are limited (Poznański *et al.*, 2001). Moreover, the conformational change of anomeric nucleosides from β to α has an effect on the stability of the DNA double helix (Thibaudeau & Chattopadhyaya, 1997).

To ascertain the sugar conformation of α-dC (**1**) in solution, a conformational analysis of the furanose puckering of α-dC (**1**) and, for comparison, of β-dC (**2**) was performed. To this end, high resolution (600 MHz) ^{1}H NMR spectra were measured in dimethyl sulfoxide (DMSO) and coupling constants were determined (Table 4). The conformational analysis of the puckering of the 2′-deoxyribofuranosyl moiety was performed using the *PSEUROT* program (Version 6.3; Van Wijk *et al.*, 1999). This program calculates the population of *N*- and *S*-type conformers on the basis of five ^3J(H,H) coupling constants, namely, ^3J(H1′,H2′), ^3J(H1′,H2″), ^3J(H2′,H3′), ^3J(H2″,H3′) and ^3J(H3′,H4′). The coupling constants are summarized in Table 4 and the spectra are available in the supporting information.

The *PSEUROT* analysis of α-dC (**1**) and β-dC (**2**) revealed that both nucleosides prefer an *S*-type sugar conformation (72 and 79% *S*-type, respectively) in solution. Accordingly, α-dC (**1**) adopts the same sugar conformation (*S*-type) in solution and the solid-state. The *S*-conformation is also the preferred conformation of the canonical sugar residues as constituents of DNA. In this regard, the sugar residue of α-nucleoside **1** fits into the DNA backbone (Fig. 5).

Moreover, the ^{1}H NMR spectra of α-dC (**1**) and β-dC (**2**) show, in both cases, two signals for the amino protons (Table 4 and Figs. S1 and S2 in the supporting information). The appearance of two separated resonances for the amino protons indicates a hindered rotation about the C4—N4 bond due to partial double-bond character. Moreover, in the solid-state structure of α-dC (**1**), the C4—N4 bond is relatively short [1.337 (3) Å; Table 2]. This suggests that the lone electron pair of the amino group is at least partially delocalized into the pyrimidine ring. These observations are consistent with earlier reports of 1-methylcytosine also reporting on the partial double-bond character of the amino group (Rossi & Kistenmacher, 1977; Fonseca Guerra *et al.*, 2014).

4. Conclusion

In this work, the crystal structure of the α-anomeric analogue of 2′-deoxycytidine (**1**) has been studied. α-2′-Deoxyribonucleosides are not widespread in nature as they are not part of canonical DNA. Literature reports on the conformational properties of α-2′-deoxyribonucleosides are also limited. The single-crystal X-ray analysis of α-2′-deoxycytidine revealed conformational properties which are outside the preferred range of α-nucleosides. This is rather unexpected as α-dC (**1**) is a rather 'simple' α-nucleoside without any further modifications at the nucleobase or sugar moiety. The *anti* conformation [χ = 173.39 (16)°] at the glycosylic bond is shifted to a higher χ value and the sugar moiety shows an almost symmetrical C2′-

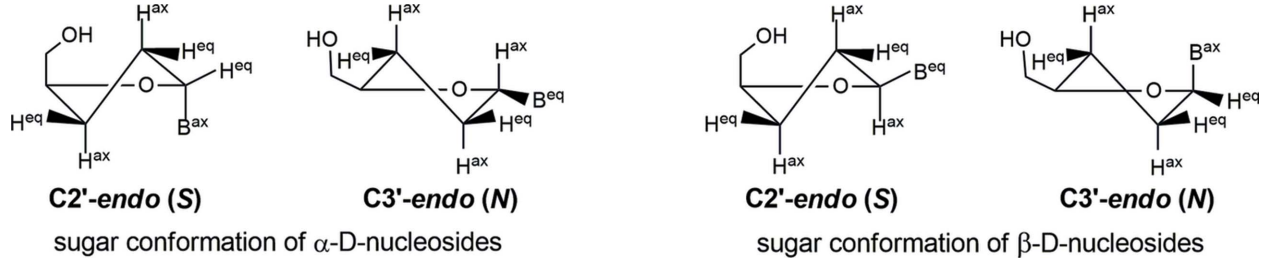


Figure 5
N and *S* conformations of α-D and β-D nucleosides in solution. 'B' corresponds to a nucleobase, with ax indicating axial and eq equatorial.

endo-C3′-*exo* twist (^2_3T; *S*-type), with P = 179.7°. The 2′-*endo* conformation is energetically less favoured in α-nucleosides compared to β-nucleosides, where this conformation is the preferred conformation of the DNA constituents. In addition, the C4—N4 bond between the amino group and the nucleobase is relatively short. Together with the appearance of two separated signals for the amino protons in the ^{1}H NMR spectrum, this indicates a hindered rotation around the C4—N4 bond due to partial double-bond character. Within the crystal, the individual nucleoside units of **1** are arranged in chains in a zigzag-like manner (*ac* plane). The crystal packing is controlled by N—H···O and O—H···O contacts between the nucleobase and sugar moieties. Moreover, two weak C—H···N contacts (C1′—H1···N3^{iii} and C3′—H3′*A*···N3^{iv}) are observed.

Although the flexibility at the glycosylic bond and the sugar conformation are generally more restricted for α-nucleosides, α-2′-deoxycytidine (**1**) is an example of an α-nucleoside with properties found outside the energetically favoured conformational range. This work constitutes a useful contribution to the field of nucleic acid chemistry and expands the state of knowledge on α-nucleosides.

Acknowledgements

We thank Dr Peter Leonard for critical reading of the manuscript. We would like to thank Professor Dr B. Wünsch, Institut für Pharmazeutische und Medizinische Chemie, Universität Münster, for providing us with 600 MHz NMR spectra. Funding by ChemBiotech, Münster, Germany, is gratefully acknowledged. Open access funding enabled and organized by Projekt DEAL.

References

Altona, C. & Sundaralingam, M. (1972). *J. Am. Chem. Soc.* **94**, 8205–8212.
Bonnett, R. (1963). *Chem. Rev.* **63**, 573–605.
Bruker (1998). *XP – Interactive molecular graphics*. Version 5.1. Bruker AXS Inc., Madison, Wisconsin, USA.
Bruker (2014). *SADABS*. Bruker AXS Inc., Madison, Wisconsin, USA.
Bruker (2015). *SAINT*. Bruker AXS Inc., Madison, Wisconsin, USA.
Bruker (2016). *APEX3*. Bruker AXS Inc., Madison, Wisconsin, USA.
Chai, Y., Guo, X., Leonard, P. & Seela, F. (2020). *Chem. Eur. J.* **26**, 13973–13989.
Chai, Y., Leonard, P., Guo, X. & Seela, F. (2019). *Chem. Eur. J.* **25**, 16639–16651.
Fonseca Guerra, C., Sanz Miguel, P. J., Cebollada, A., Bickelhaupt, F. M. & Lippert, B. (2014). *Chem. Eur. J.* **20**, 9494–9499.
Görbitz, C. H., Nelson, W. H. & Sagstuen, E. (2005). *Acta Cryst.* E**61**, o1207–o1209.
Guo, X. & Seela, F. (2017). *Chem. Eur. J.* **23**, 11776–11779.
Hamor, T. A., O'Leary, M. K. & Walker, R. T. (1977). *Acta Cryst.* B**33**, 1218–1223.
IUPAC–IUB Joint Commission on Biochemical Nomenclature (1983). *Eur. J. Biochem.* **131**, 9–15.
Kondhare, D., Budow-Busse, S., Daniliuc, C. & Seela, F. (2020). *Acta Cryst.* C**76**, 513–523.
Latha, Y. S. & Yathindra, N. (1992). *Biopolymers*, **32**, 249–269.
Morvan, F., Rayner, B. & Imbach, J.-L. (1990). In *Genetic Engineering*, Vol. 12, edited by J. K. Setlow. New York: Plenum Press.
Morvan, F., Rayner, B., Imbach, J.-L., Chang, D.-K. & Lown, J. W. (1987*a*). *Nucleic Acids Res.* **15**, 4241–4255.
Morvan, F., Rayner, B., Imbach, J.-L., Lee, M., Hartley, J. A., Chang, D.-K. & Lown, J. W. (1987*b*). *Nucleic Acids Res.* **15**, 7027–7044.
Müller, S. L., Zhou, X., Leonard, P., Korzhenko, O., Daniliuc, C. & Seela, F. (2019). *Chem. Eur. J.* **25**, 3077–3090.
Ni, G., Du, Y., Tang, F., Liu, J., Zhao, H. & Chen, Q. (2019). *RSC Adv.* **9**, 14302–14320.
Parsons, S., Flack, H. D. & Wagner, T. (2013). *Acta Cryst.* B**69**, 249–259.
Post, M. L., Birnbaum, G. I., Huber, C. P. & Shugar, D. (1977). *Biochim. Biophys. Acta Nucleic Acids Protein Synth.* **479**, 133–142.
Poznański, J., Felczak, K., Bretner, M., Kulikowski, T. & Remin, M. (2001). *Biochem. Biophys. Res. Commun.* **283**, 1142–1149.
Rossi, M. & Kistenmacher, T. J. (1977). *Acta Cryst.* B**33**, 3962–3965.
Saenger, W. (1984). In *Principles of Nucleic Acid Structure*, edited by C. R. Cantor. New York: Springer-Verlag.
Seela, F., Rosemeyer, H., Melenewski, A., Heithoff, E.-M., Eickmeier, H. & Reuter, H. (2002). *Acta Cryst.* C**58**, o142–o144.
Sheldrick, G. M. (2015*a*). *Acta Cryst.* A**71**, 3–8.
Sheldrick, G. M. (2015*b*). *Acta Cryst.* C**71**, 3–8.
Spackman, M. A. & Jayatilaka, D. (2009). *CrystEngComm*, **11**, 19–32.
Sundaralingam, M. (1971). *J. Am. Chem. Soc.* **93**, 6644–6647.
Suzuki, S., Suzuki, K., Imai, T., Suzuki, N. & Okuda, S. (1965). *J. Biol. Chem.* **240**, PC554–PC556.
Thibaudeau, C. & Chattopadhyaya, J. (1997). *Nucleoside Nucleotides Nucleic Acids*, **16**, 523–529.
Turner, M. J., McKinnon, J. J., Wolff, S. K., Grimwood, D. J., Spackman, P. R., Jayatilaka, D. & Spackman, M. A. (2017). *CrystalExplorer17*. University of Western Australia. http://hirshfeldsurface.net.
Van Wijk, L., Haasnoot, C. A. G., de Leeuw, F. A. A. M., Huckriede, B. D., Westra Hoekzema, A. J. A. & Altona, C. (1999). *PSEUROT 6.3*. Leiden Institute of Chemistry, Leiden University, The Netherlands.
Yamaguchi, T. & Saneyoshi, M. (1984). *Chem. Pharm. Bull.* **32**, 1441–1450.
Young, D. W. & Wilson, H. R. (1975). *Acta Cryst.* B**31**, 961–965.
Zhou, X., Müller, S. L., Leonard, P., Daniliuc, C., Chai, Y., Budow-Busse, S. & Seela, F. (2019). *J. Mol. Struct.* **1190**, 37–46.

supporting information

Acta Cryst. (2021). C**77**, 202-208 [https://doi.org/10.1107/S2053229621003430]

The α-D-anomer of 2′-deoxycytidine: crystal structure, nucleoside conformation and Hirshfeld surface analysis

Simone Budow-Busse, Yingying Chai, Sebastian Lars Müller, Constantin Daniliuc and Frank Seela

Computing details

Data collection: *APEX3* (Bruker, 2016); cell refinement: *SAINT* (Bruker, 2015); data reduction: *SAINT* (Bruker, 2015); program(s) used to solve structure: *SHELXS2014* (Sheldrick, 2015a); program(s) used to refine structure: *SHELXL2014* (Sheldrick, 2015b); molecular graphics: *APEX3* (Bruker, 2016); software used to prepare material for publication: *APEX3* (Bruker, 2016) and *XP* (Bruker, 1998).

2′-Deoxycytidine

Crystal data

$C_9H_{13}N_3O_4$
M_r = 227.22
Orthorhombic, $P2_12_12_1$
a = 6.8378 (4) Å
b = 11.4334 (7) Å
c = 12.7595 (8) Å
V = 997.53 (11) Å^3
Z = 4
$F(000)$ = 480
D_x = 1.513 Mg m^{-3}
Cu $K\alpha$ radiation, λ = 1.54178 Å
Cell parameters from 9144 reflections
θ = 5.2–66.9°
μ = 1.02 mm^{-1}
T = 100 K
Prism, colourless
0.22 × 0.18 × 0.16 mm

Data collection

Bruker APEXII Kappa CCD diffractometer
Radiation source: fine-focus sealed tube, fine-focus sealed tube
Graphite monochromator
Detector resolution: 8.3333 pixels mm^{-1}
φ and ω scans
Absorption correction: multi-scan (SADABS; Bruker, 2014)
T_{min} = 0.75, T_{max} = 0.85
11824 measured reflections
1768 independent reflections
1719 reflections with $I > 2\sigma(I)$
R_{int} = 0.037
θ_{max} = 66.9°, θ_{min} = 5.2°
h = −8→8
k = −13→12
l = −15→15

Refinement

Refinement on F^2
Least-squares matrix: full
$R[F^2 > 2\sigma(F^2)]$ = 0.025
$wR(F^2)$ = 0.061
S = 1.12
1768 reflections
161 parameters
0 restraints
Primary atom site location: structure-invariant direct methods
Hydrogen site location: mixed
H atoms treated by a mixture of independent and constrained refinement
$w = 1/[\sigma^2(F_o^2) + (0.0234P)^2 + 0.3118P]$ where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max} < 0.001$

$\Delta\rho_{max}$ = 0.12 e Å^{-3}
$\Delta\rho_{min}$ = −0.18 e Å^{-3}

Absolute structure: Flack *x* determined using 684 quotients [(I+)-(I-)]/[(I+)+(I-)] (Parsons *et al.*, 2013)
Absolute structure parameter: 0.04 (9)

Special details

Geometry. All esds (except the esd in the dihedral angle between two l.s. planes) are estimated using the full covariance matrix. The cell esds are taken into account individually in the estimation of esds in distances, angles and torsion angles; correlations between esds in cell parameters are only used when they are defined by crystal symmetry. An approximate (isotropic) treatment of cell esds is used for estimating esds involving l.s. planes.

Refinement. Reflections were merged by SHELXL according to the crystal class for the calculation of statistics and refinement.
_reflns_Friedel_fraction is defined as the number of unique Friedel pairs measured divided by the number that would be possible theoretically, ignoring centric projections and systematic absences.
The hydrogens at N4, O3′ and O5′ atoms were refined freely.

Fractional atomic coordinates and isotropic or equivalent isotropic displacement parameters (Å^2)

	x	*y*	*z*	U_{iso}*/U_{eq}
N1	0.8363 (2)	0.56682 (14)	0.36354 (13)	0.0141 (4)
N3	0.9151 (2)	0.75960 (14)	0.30990 (13)	0.0150 (4)
N4	0.8602 (3)	0.91354 (15)	0.42076 (14)	0.0179 (4)
H4A	0.851 (4)	0.941 (2)	0.484 (2)	0.025 (7)*
H4B	0.912 (4)	0.960 (2)	0.368 (2)	0.041 (8)*
O2	0.9632 (2)	0.60359 (12)	0.20186 (11)	0.0168 (3)
C2	0.9078 (3)	0.64392 (17)	0.28828 (15)	0.0139 (4)
C4	0.8596 (3)	0.79804 (17)	0.40431 (15)	0.0143 (4)
C5	0.8004 (3)	0.72036 (16)	0.48517 (15)	0.0159 (4)
H5A	0.7702	0.748	0.5535	0.019*
C6	0.7887 (3)	0.60552 (17)	0.46094 (16)	0.0154 (4)
H6	0.7467	0.5512	0.5126	0.018*
C1′	0.8145 (3)	0.44122 (16)	0.33182 (15)	0.0150 (4)
H1	0.9404	0.4117	0.3019	0.018*
C2′	0.6507 (3)	0.42608 (17)	0.25202 (15)	0.0154 (4)
H2A	0.6779	0.3599	0.2042	0.018*
H2B	0.632	0.4981	0.2101	0.018*
C3′	0.4725 (3)	0.40127 (17)	0.32044 (16)	0.0155 (4)
H3A	0.3738	0.3528	0.2823	0.019*
C4′	0.5609 (3)	0.33435 (17)	0.41243 (15)	0.0143 (4)
H4	0.4897	0.3557	0.4781	0.017*
C5′	0.5584 (3)	0.20248 (17)	0.39946 (15)	0.0170 (4)
H5B	0.6242	0.166	0.4603	0.02*
H5C	0.4211	0.175	0.3985	0.02*
O3′	0.3873 (2)	0.50722 (12)	0.35979 (11)	0.0179 (3)
H3	0.381 (5)	0.559 (3)	0.303 (3)	0.054 (9)*
O4′	0.7623 (2)	0.37426 (11)	0.42000 (11)	0.0155 (3)
O5′	0.6537 (2)	0.16570 (12)	0.30536 (11)	0.0187 (3)
H5	0.786 (5)	0.167 (3)	0.316 (3)	0.050 (9)*

Atomic displacement parameters (Å²)

	U^{11}	U^{22}	U^{33}	U^{12}	U^{13}	U^{23}
N1	0.0156 (8)	0.0115 (8)	0.0152 (8)	−0.0007 (7)	0.0008 (7)	0.0004 (7)
N3	0.0166 (8)	0.0129 (8)	0.0154 (8)	−0.0008 (6)	−0.0002 (7)	−0.0003 (7)
N4	0.0262 (10)	0.0129 (9)	0.0147 (8)	−0.0017 (7)	0.0015 (8)	−0.0008 (7)
O2	0.0211 (7)	0.0146 (7)	0.0149 (7)	−0.0002 (6)	0.0040 (6)	−0.0016 (6)
C2	0.0124 (9)	0.0144 (10)	0.0150 (10)	−0.0007 (7)	0.0003 (8)	0.0027 (7)
C4	0.0131 (9)	0.0139 (9)	0.0157 (9)	0.0001 (8)	−0.0015 (8)	−0.0001 (8)
C5	0.0187 (11)	0.0162 (10)	0.0129 (9)	−0.0001 (9)	0.0001 (8)	−0.0009 (8)
C6	0.0167 (10)	0.0162 (9)	0.0134 (9)	−0.0003 (8)	0.0008 (8)	0.0019 (8)
C1′	0.0196 (10)	0.0110 (9)	0.0145 (10)	−0.0005 (8)	0.0025 (8)	0.0008 (7)
C2′	0.0206 (11)	0.0127 (9)	0.0130 (9)	0.0001 (8)	0.0003 (8)	−0.0010 (8)
C3′	0.0180 (10)	0.0132 (9)	0.0152 (9)	−0.0003 (8)	−0.0022 (8)	−0.0027 (8)
C4′	0.0149 (9)	0.0127 (9)	0.0154 (9)	−0.0011 (8)	0.0018 (8)	−0.0018 (8)
C5′	0.0191 (10)	0.0138 (10)	0.0181 (10)	−0.0007 (8)	0.0039 (8)	−0.0001 (8)
O3′	0.0222 (7)	0.0142 (7)	0.0173 (7)	0.0044 (6)	0.0008 (6)	0.0004 (6)
O4′	0.0181 (7)	0.0133 (7)	0.0151 (7)	−0.0022 (5)	−0.0021 (6)	0.0032 (5)
O5′	0.0191 (8)	0.0156 (7)	0.0214 (7)	−0.0009 (6)	0.0025 (7)	−0.0037 (6)

Geometric parameters (Å, º)

N1—C6	1.359 (3)	C1′—H1	1.0
N1—C2	1.392 (3)	C2′—C3′	1.525 (3)
N1—C1′	1.499 (2)	C2′—H2A	0.99
N3—C4	1.337 (3)	C2′—H2B	0.99
N3—C2	1.352 (3)	C3′—O3′	1.435 (2)
N4—C4	1.337 (3)	C3′—C4′	1.526 (3)
N4—H4A	0.86 (3)	C3′—H3A	1.0
N4—H4B	0.93 (3)	C4′—O4′	1.454 (2)
O2—C2	1.254 (2)	C4′—C5′	1.517 (3)
C4—C5	1.420 (3)	C4′—H4	1.0
C5—C6	1.351 (3)	C5′—O5′	1.429 (2)
C5—H5A	0.95	C5′—H5B	0.99
C6—H6	0.95	C5′—H5C	0.99
C1′—O4′	1.407 (2)	O3′—H3	0.94 (3)
C1′—C2′	1.524 (3)	O5′—H5	0.91 (4)
C6—N1—C2	120.59 (16)	C1′—C2′—H2A	111.2
C6—N1—C1′	122.33 (16)	C3′—C2′—H2A	111.2
C2—N1—C1′	117.08 (15)	C1′—C2′—H2B	111.2
C4—N3—C2	119.66 (17)	C3′—C2′—H2B	111.2
C4—N4—H4A	120.2 (16)	H2A—C2′—H2B	109.1
C4—N4—H4B	116.9 (17)	O3′—C3′—C2′	111.57 (16)
H4A—N4—H4B	120 (2)	O3′—C3′—C4′	108.37 (16)
O2—C2—N3	121.88 (17)	C2′—C3′—C4′	102.54 (16)
O2—C2—N1	118.67 (17)	O3′—C3′—H3A	111.3
N3—C2—N1	119.45 (17)	C2′—C3′—H3A	111.3

N4—C4—N3	117.71 (18)	C4′—C3′—H3A	111.3
N4—C4—C5	120.30 (18)	O4′—C4′—C5′	109.27 (16)
N3—C4—C5	121.99 (18)	O4′—C4′—C3′	105.60 (15)
C6—C5—C4	117.28 (18)	C5′—C4′—C3′	114.21 (16)
C6—C5—H5A	121.4	O4′—C4′—H4	109.2
C4—C5—H5A	121.4	C5′—C4′—H4	109.2
C5—C6—N1	120.77 (18)	C3′—C4′—H4	109.2
C5—C6—H6	119.6	O5′—C5′—C4′	112.26 (16)
N1—C6—H6	119.6	O5′—C5′—H5B	109.2
O4′—C1′—N1	109.30 (15)	C4′—C5′—H5B	109.2
O4′—C1′—C2′	106.64 (16)	O5′—C5′—H5C	109.2
N1—C1′—C2′	111.26 (16)	C4′—C5′—H5C	109.2
O4′—C1′—H1	109.9	H5B—C5′—H5C	107.9
N1—C1′—H1	109.9	C3′—O3′—H3	106.4 (19)
C2′—C1′—H1	109.9	C1′—O4′—C4′	110.96 (15)
C1′—C2′—C3′	103.04 (15)	C5′—O5′—H5	109 (2)

C4—N3—C2—O2	−178.13 (18)	C2—N1—C1′—C2′	−69.1 (2)
C4—N3—C2—N1	2.1 (3)	O4′—C1′—C2′—C3′	27.75 (19)
C6—N1—C2—O2	175.01 (18)	N1—C1′—C2′—C3′	−91.35 (18)
C1′—N1—C2—O2	−5.2 (3)	C1′—C2′—C3′—O3′	82.77 (18)
C6—N1—C2—N3	−5.2 (3)	C1′—C2′—C3′—C4′	−33.03 (18)
C1′—N1—C2—N3	174.55 (18)	O3′—C3′—C4′—O4′	−90.74 (17)
C2—N3—C4—N4	−176.89 (18)	C2′—C3′—C4′—O4′	27.35 (18)
C2—N3—C4—C5	2.8 (3)	O3′—C3′—C4′—C5′	149.18 (16)
N4—C4—C5—C6	175.05 (18)	C2′—C3′—C4′—C5′	−92.73 (19)
N3—C4—C5—C6	−4.6 (3)	O4′—C4′—C5′—O5′	−62.1 (2)
C4—C5—C6—N1	1.4 (3)	C3′—C4′—C5′—O5′	55.9 (2)
C2—N1—C6—C5	3.3 (3)	N1—C1′—O4′—C4′	109.54 (17)
C1′—N1—C6—C5	−176.44 (19)	C2′—C1′—O4′—C4′	−10.83 (19)
C6—N1—C1′—O4′	−6.8 (3)	C5′—C4′—O4′—C1′	112.58 (17)
C2—N1—C1′—O4′	173.39 (16)	C3′—C4′—O4′—C1′	−10.69 (19)
C6—N1—C1′—C2′	110.68 (19)		

Hydrogen-bond geometry (Å, º)

D—H···*A*	*D*—H	H···*A*	*D*···*A*	*D*—H···*A*
N4—H4*A*···O3′^{i}	0.86 (3)	2.10 (3)	2.949 (2)	168 (2)
N4—H4*B*···O2^{ii}	0.93 (3)	2.05 (3)	2.937 (2)	159 (2)
C1′—H1···N3^{iii}	1.0	2.46	3.317 (3)	144
C3′—H3*A*···N3^{iv}	1.0	2.53	3.524 (3)	170
O3′—H3···O5′^{v}	0.94 (3)	1.86 (4)	2.793 (2)	175 (3)
O5′—H5···O2^{iii}	0.91 (4)	1.88 (3)	2.716 (2)	152 (3)

Symmetry codes: (i) *x*+1/2, −*y*+3/2, −*z*+1; (ii) −*x*+2, *y*+1/2, −*z*+1/2; (iii) −*x*+2, *y*−1/2, −*z*+1/2; (iv) −*x*+1, *y*−1/2, −*z*+1/2; (v) −*x*+1, *y*+1/2, −*z*+1/2.