

Solid-state transformation of 4,2'-anhydro-5-(β -D-arabinofuranosyl)uracil dihydrate to 4,2'-anhydro-5-(β -D-arabinofuranosyl)uracil monohydrate

Jan W. Bats,* Uwe Parsch and Joachim W. Engels

Institut für Organische Chemie, Universität Frankfurt, Marie-Curie-Strasse 11, D-60439 Frankfurt am Main, Germany

Correspondence e-mail: bats@indy2.org.chemie.uni-frankfurt.de

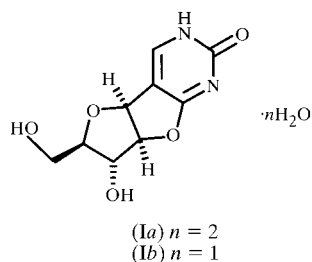
Received 25 February 2000

Accepted 26 May 2000

Crystals of 4,2'-anhydro-5-(β -D-arabinofuranosyl)uracil, (I), obtained from an aqueous solution, were characterized as the dihydrate, $C_9H_{10}N_2O_5 \cdot 2H_2O$, (Ia). In air, these crystals slowly transform to the monohydrate, $C_9H_{10}N_2O_5 \cdot H_2O$, (Ib), but remain crystalline. The solid-state transformation proceeds with the loss of one water molecule and a rearrangement of hydrogen-bonded layers of molecules. The furanose ring in (I) has an approximate $C4'$ -*exo*, $O4'$ -*endo* twist conformation. The central five-membered ring is slightly puckered. The uracil group is planar within experimental uncertainty.

Comment

A number of nucleosides in which the ribofuranosyl moiety is linked to a heterocyclic base through a carbon-carbon bond (hence they are called C-nucleosides) have been found in nature (Suhadolnik, 1970, 1979). The first example was pseudouridine, which was isolated in 1957 (Davis & Allen, 1957;



Cohn, 1957). Its crystal structure has been reported by Hempel *et al.* (1997). Pseudouridine is present ubiquitously in tRNA. Many C-nucleosides are antibiotics and exhibit anti-cancer and/or antiviral activity (Watanabe, 1994). We present here the crystal structures of 4,2'-anhydro-5-(β -D-arabinofuranosyl)uracil as its di-, (Ia), and monohydrate, (Ib). This anhydro C-nucleoside has been used successfully as an intermediate for the specific inversion or replacement of the ribose $O2'$ -hydroxyl group (Reichman *et al.*, 1976; Chu *et al.*, 1978). The structure analysis of (I) was undertaken to establish the

influence of the $O4-C2'$ cyclization on the molecular conformation. Crystals of (I) taken from an aqueous solution corresponded to the dihydrate (Ia). In air, these crystals slowly lost water and transformed to the monohydrate, (Ib), without losing their crystallinity. At room temperature, this transformation proceeds within a few days (see *Experimental*). Thus, the structure determinations of (Ia) and (Ib) were undertaken to clarify the nature of the solid-state transformation.

The molecular conformations in (Ia) and (Ib) are very similar. Torsion angles among corresponding bonds in (Ia) and (Ib) agree within 4° . The five-membered furanose ring has an approximate $C4'$ -*exo*, $O4'$ -*endo* twist conformation. The ring-puckering parameters defined by Cremer & Pople (1975) are $q = 0.351 \text{ \AA}$ and $\varphi = 346.1^\circ$ in (Ia), and $q = 0.373 \text{ \AA}$ and $\varphi = 341.7^\circ$ in (Ib).

The side chain attached to $C4'$ is in a pseudo-equatorial position with respect to the five-membered ring. The central five-membered ring has a $C2',C1'$ twist conformation, with ring puckering parameters $q = 0.114 \text{ \AA}$ and $\varphi = 51.9^\circ$ in (Ia), and $q = 0.094 \text{ \AA}$ and $\varphi = 57.6^\circ$ in (Ib). The uracil ring is planar within experimental uncertainty. A $C4'$ -*exo*, $O4'$ -*endo* sugar conformation has also been observed for one of the two independent molecules of 2,2'-anhydro-1- β -D-arabinofuranosyl-6-methyluracil (Głowska *et al.*, 1986).

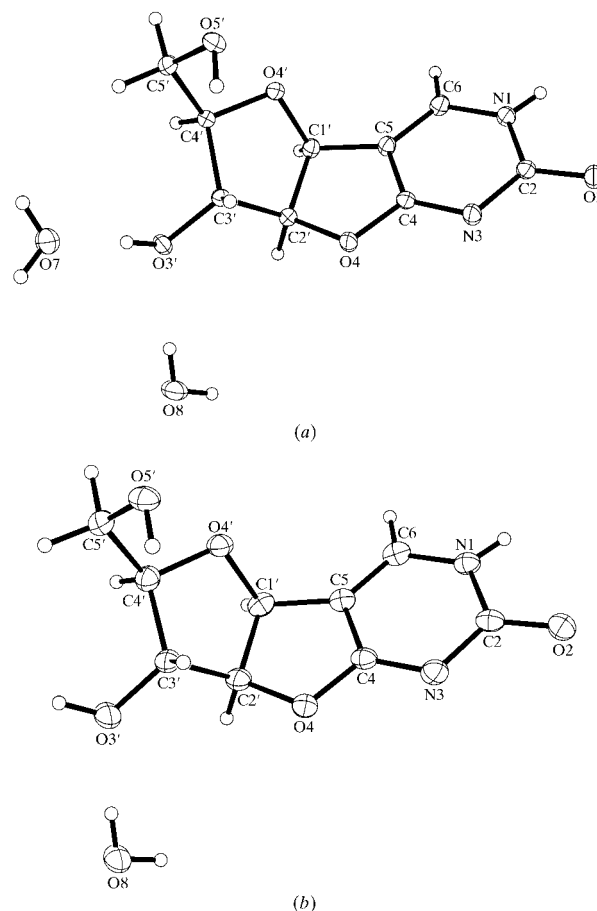


Figure 1

The molecules of (a) (Ia) and (b) (Ib) shown with 50% probability displacement ellipsoids.

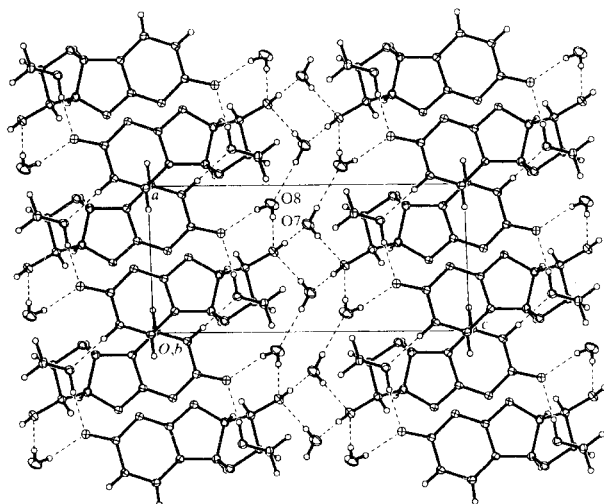


Figure 2
The projection of the crystal packing of (Ia) shown down the *b* axis.

A rather different conformation between a *C4'-endo* envelope and a *C4'-endo,C3'-exo* twist has been observed for the furanose ring in a number of related cyclic nucleosides (Suck & Saenger, 1973; Neidle *et al.*, 1979). An *O4'-endo* envelope conformation was observed by Parkanyi & Kálmán (1995), while a *C3'-exo,C4'-endo* conformation was observed by Purkayastha *et al.* (1990). Thus, the furanose ring can have more than one conformation despite the constraints imposed by the *O4*–*C2'* cyclization.

The C–O bond lengths at *O4* are very asymmetrical. The *C2'*–*O4* lengths of 1.470 (1) Å in (Ia) and 1.471 (3) Å in (Ib) are rather long for C–O single bonds; the *O4*–*C4* lengths of 1.343 (1) Å in (Ia) and 1.343 (3) Å in (Ib), on the other hand, are rather short. Similar bond-length differences have been observed in cyclic nucleosides. The shortening of the *O4*–*C4* bond length has been attributed to overlap of the π orbital at *O4* with the π system of the uracil group (Główka *et al.*, 1986).

The *C4'*–*O4'* and *C1'*–*O4'* bond lengths are almost equal. A shortening of the *C1'*–*O4'* bond due to anomeric effects is generally observed in N-nucleosides. In C-nucleosides, however, such a bond shortening does not occur (Hempel *et al.*, 1997; Birnbaum *et al.*, 1980).

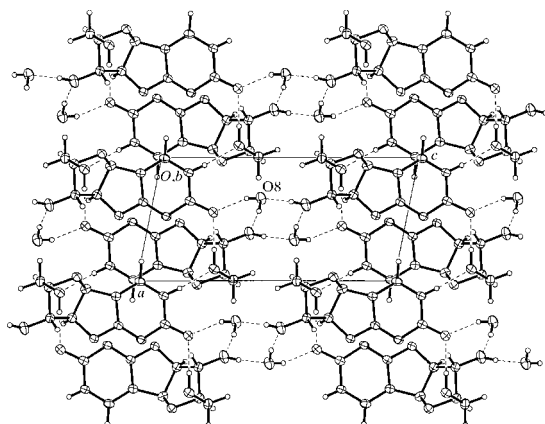


Figure 3
The projection of the crystal packing of (Ib) shown down the *b* axis.

Both crystal structures contain three-dimensional networks of hydrogen bonds (Tables 1 and 2). Layers of 4,2'-anhydro-5-(β -D-arabinofuranosyl)uracil molecules in the *ab* direction occur in both (Ia) and (Ib). Molecules within these layers are connected by intermolecular *N1*–*H01*··*O5'* and *O5'*–*H05'*··*O2* hydrogen bonds. The layers are also stabilized by additional intermolecular weak C–H··*O* and C–H··*N* interactions, which have been included in Tables 1 and 2. The dominant weak interaction is a bifurcated contact involving the *C6*–*H6* bond. The significance of such weak interactions has been outlined by Desiraju & Steiner (1999).

The 4,2'-anhydro-5-(β -D-arabinofuranosyl)uracil layers are connected in the *c* direction by water molecules. The *O8* water molecule occupies very similar positions in structures (Ia) and (Ib). It is connected by two *O*–*H*··*O* hydrogen bonds to the organic layer. Water molecule *O7* in (Ia) connects two adjacent layers and is hydrogen bonded to *O8*. A similar water position is missing in the structure of (Ib). Thus, the transition of the dihydrate structure to the monohydrate structure can be understood as the loss of *O7* from the structure of (Ia). Then adjacent organic layers are displaced by *a*/2. In this way, a new intermolecular *O3'*–*H03'*··*O8* hydrogen bond is formed in (Ib) which replaces the *O3'*–*H03'*··*O7* linkage in (Ia). This displacement of *a*/2 between adjacent layers means that only every second layer should translate. Thus, little overall deformation of the crystal is required. The cell of (Ib) is approximately related to the cell of (Ia) by the relations $a_{1b} = a_{1a}$, $b_{1b} = -b_{1a}$ and $c_{1b} = -c_{1a} - a_{1a}/2$.

Experimental

The title compound was prepared according to the methods of Reichman *et al.* (1976) and Chu *et al.* (1978). Crystallization from water resulted in the dihydrate, (Ia). At room temperature and in air, the crystals slowly lose water and transform to the corresponding monohydrate, (Ib), without losing their crystallinity. The crystal of (Ia) used for the data collection was cut from a large rod taken from an aqueous solution and was immediately cooled to 135 K. The crystal of (Ib) used for the data collection was obtained from water but had been stored in air for several months.

Compound (Ia)

Crystal data

$C_9H_{10}N_2O_5 \cdot 2H_2O$
 $M_r = 262.22$
Monoclinic, $P2_1$
 $a = 6.0315$ (7) Å
 $b = 7.2490$ (8) Å
 $c = 12.8693$ (19) Å
 $\beta = 91.921$ (9)°
 $V = 562.36$ (12) Å³
 $Z = 2$

$D_x = 1.549$ Mg m⁻³
Mo $K\alpha$ radiation
Cell parameters from 207 reflections
 $\theta = 3$ – 23°
 $\mu = 0.135$ mm⁻¹
 $T = 135$ (2) K
Block, colorless
 $0.60 \times 0.54 \times 0.34$ mm

Data collection

Siemens SMART diffractometer
 ω scans
Absorption correction: numerical (SHELXTL; Sheldrick, 1996)
 $T_{min} = 0.916$, $T_{max} = 0.961$
10 536 measured reflections
1970 independent reflections
1942 reflections with $I > 2\sigma(I)$

$R_{int} = 0.039$
 $\theta_{max} = 32.29^\circ$
 $h = -8 \rightarrow 8$
 $k = -9 \rightarrow 10$
 $l = -17 \rightarrow 19$
210 standard reflections
frequency: 600 min
intensity decay: none

Table 1
Hydrogen-bonding geometry (Å, °) for (Ia).

D—H...A	D—H	H...A	D...A	D—H...A
O3'—H03'...O7	0.87 (2)	1.73 (2)	2.600 (1)	178 (2)
O5'—H05'...O2 ⁱ	0.92 (2)	1.80 (2)	2.701 (1)	167 (2)
N1—H01...O5 ⁱⁱⁱ	0.85 (2)	1.88 (2)	2.732 (1)	178 (2)
O7—H07A...O8 ⁱⁱⁱ	0.76 (3)	1.97 (3)	2.736 (1)	177 (3)
O7—H07B...O3 ^{iv}	0.78 (3)	2.08 (3)	2.855 (1)	178 (2)
O8—H08A...O3 ⁱ	0.83 (3)	1.98 (3)	2.807 (1)	174 (3)
O8—H08B...O2 ^v	0.83 (2)	2.05 (2)	2.855 (1)	164 (2)
C2'—H2'...O2 ^v	1.00 (2)	2.62 (2)	3.511 (1)	149 (2)
C3'—H3'...O2 ⁱ	0.97 (2)	2.63 (2)	3.551 (1)	159 (2)
C6—H6...O4 ^{vi}	0.89 (2)	2.57 (2)	3.355 (1)	148 (2)
C6—H6...N3 ^{vi}	0.89 (2)	2.53 (2)	3.362 (1)	155 (2)

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

Refinement

Refinement on F^2
 $R[F^2 > 2\sigma(F^2)] = 0.025$
 $wR(F^2) = 0.076$
 $S = 1.377$
 1970 reflections
 220 parameters
 All H-atom parameters refined

$w = 1/[\sigma^2(F_o^2) + (0.05P)^2]$
 where $P = (F_o^2 + 2F_c^2)/3$
 $(\Delta/\sigma)_{\max} < 0.001$
 $\Delta\rho_{\max} = 0.31 \text{ e } \text{Å}^{-3}$
 $\Delta\rho_{\min} = -0.27 \text{ e } \text{Å}^{-3}$
 Extinction correction: *SHELXL97*
 Extinction coefficient: 0.31 (2)

Compound (Ib)*Crystal data*

$\text{C}_9\text{H}_{10}\text{N}_2\text{O}_5 \cdot 2\text{H}_2\text{O}$
 $M_r = 244.21$
 Monoclinic, $P2_1$
 $a = 6.024 (2) \text{ Å}$
 $b = 7.218 (3) \text{ Å}$
 $c = 12.357 (4) \text{ Å}$
 $\beta = 101.40 (4)^\circ$
 $V = 526.7 (3) \text{ Å}^3$
 $Z = 2$

$D_x = 1.540 \text{ Mg m}^{-3}$
 Mo $K\alpha$ radiation
 Cell parameters from 54 reflections
 $\theta = 3\text{--}23^\circ$
 $\mu = 0.131 \text{ mm}^{-1}$
 $T = 137 (2) \text{ K}$
 Rod, colorless
 $0.56 \times 0.35 \times 0.07 \text{ mm}$

Data collection

Siemens SMART diffractometer
 ω scans
 Absorption correction: numerical
 (*SHELXTL*; Sheldrick, 1996)
 $T_{\min} = 0.946, T_{\max} = 0.991$
 10 608 measured reflections
 1808 independent reflections
 1443 reflections with $I > 2\sigma(I)$

$R_{\text{int}} = 0.1089$
 $\theta_{\max} = 32.01^\circ$
 $h = -8 \rightarrow 8$
 $k = -10 \rightarrow 10$
 $l = -18 \rightarrow 17$
 138 standard reflections
 frequency: 600 min
 intensity decay: none

Refinement

Refinement on F^2
 $R[F^2 > 2\sigma(F^2)] = 0.054$
 $wR(F^2) = 0.110$
 $S = 1.376$
 1808 reflections
 202 parameters

All H-atom parameters refined
 $w = 1/[\sigma^2(F_o^2) + (0.04P)^2]$
 where $P = (F_o^2 + 2F_c^2)/3$
 $(\Delta/\sigma)_{\max} = 0.001$
 $\Delta\rho_{\max} = 0.43 \text{ e } \text{Å}^{-3}$
 $\Delta\rho_{\min} = -0.34 \text{ e } \text{Å}^{-3}$

The transformation from (Ia) to (Ib) was examined for a crystal with dimensions $0.06 \times 0.25 \times 0.80 \text{ mm}$ at room temperature. The crystal was taken from an aqueous solution and was mounted on an Enraf-Nonius CAD-4 diffractometer. Reflections were searched at regular intervals over a period of three days. Initially, all reflections corresponded to the dihydrate. Within hours, however, reflection profiles started to broaden and additional weak very broad reflections appeared which could be indexed as belonging to the monohydrate. After 24 h, reflection intensities of the dihydrate fraction had decreased to about 25% of their starting values and the crystal was already dominated by the monohydrate fraction. After 72 h, reflections of the dihydrate fraction had decreased to about 1.5% of their initial values and the transformation from the dihydrate to the

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

D—H...A	D—H	H...A	D...A	D—H...A
O3'—H03'...O8 ⁱ	0.88 (4)	1.79 (4)	2.663 (3)	175 (5)
O5'—H05'...O2 ⁱⁱ	0.93 (5)	1.82 (5)	2.735 (3)	169 (5)
N1—H01...O5 ⁱⁱⁱ	0.84 (4)	1.91 (4)	2.744 (3)	175 (4)
O8—H08B...O2	0.83 (5)	2.06 (5)	2.862 (3)	162 (5)
O8—H08A...O3 ⁱⁱⁱ	0.84 (8)	1.91 (8)	2.728 (3)	165 (7)
C2'—H2'...O2 ^{iv}	0.91 (4)	2.56 (4)	3.406 (4)	155 (3)
C3'—H3'...O2 ⁱⁱ	0.95 (3)	2.59 (3)	3.479 (3)	155 (2)
C6—H6...O4 ^v	0.97 (4)	2.48 (3)	3.337 (4)	146 (2)
C6—H6...N3 ^v	0.97 (4)	2.51 (4)	3.379 (4)	148 (2)

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

monohydrate was approximately complete. The crystal of (Ia) was of good quality with narrow reflection profiles (peak-width at half-height of about 0.45°). Reflection profiles for the crystal of (Ib) were very broad (peak-width at half-height of about 1.2 to 1.4°).

For both compounds, data collection: *SMART* (Siemens, 1995); cell refinement: *SMART*; data reduction: *SAINT* (Siemens, 1995); program(s) used to solve structure: *SHELXTL* (Sheldrick, 1996); program(s) used to refine structure: *SHELXL97* (Sheldrick, 1997); molecular graphics: *XP* in *SHELXTL*; software used to prepare material for publication: CIF in *SHELXL97*.

This work was supported by the Deutsche Forschungsgemeinschaft. UP wishes to thank the Fonds der Chemischen Industrie for a PhD fellowship.

Supplementary data for this paper are available from the IUCr electronic archives (Reference: AV1044). Services for accessing these data are described at the back of the journal.

References

- Birnbaum, G. I., Watanabe, K. A. & Fox, J. J. (1980). *Can. J. Chem.* **58**, 1633–1638.
 Chu, C. K., Reichman, U., Watanabe, K. A. & Fox, J. J. (1978). *J. Med. Chem.* **21**, 96–100.
 Cohn, W. E. (1957). *Fed. Proc.* **16**, 166.
 Cremer, D. & Pople, J. A. (1975). *J. Am. Chem. Soc.* **97**, 1354–1358.
 Davis, F. F. & Allen, F. W. (1957). *J. Biol. Chem.* **227**, 907–915.
 Desiraju, G. R. & Steiner, T. (1999). In *The Weak Hydrogen Bond*. Oxford University Press.
 Glówka, M. L., Iwanicka, I., Gilli, G. & Bertolasi, V. (1986). *Acta Cryst.* **C42**, 1526–1529.
 Hempel, A., Lane, B. G. & Camerman, N. (1997). *Acta Cryst.* **C53**, 1707–1709.
 Neidle, S., Taylor, G. L. & Cowling, P. C. (1979). *Acta Cryst.* **B35**, 708–712.
 Parkanyi, L. & Kálmán, A. (1995). *Z. Kristallogr.* **210**, 711–712.
 Purkayastha, S., Cheer, C. J. & Panzica, R. P. (1990). *J. Heterocycl. Chem.* **27**, 753–757.
 Reichman, U., Chu, C. K., Wempfen, I., Watanabe, K. A. & Fox, J. J. (1976). *J. Heterocycl. Chem.* **13**, 933–935.
 Sheldrick, G. M. (1996). *SHELXTL*. Siemens Analytical X-ray Instruments Inc., Madison, Wisconsin, USA.
 Sheldrick, G. M. (1997). *SHELXL97*. University of Göttingen, Germany.
 Siemens (1995). *SMART* and *SAINT*. Siemens Analytical X-ray Instruments Inc., Madison, Wisconsin, USA.
 Suck, D. & Saenger, W. (1973). *Acta Cryst.* **B29**, 1323–1330.
 Suhadolnik, R. J. (1970). In *Nucleoside Antibiotics*. New York: Wiley Interscience.
 Suhadolnik, R. J. (1979). *Prog. Nucleic Acid Res. Mol. Biol.* **22**, 193–291.
 Watanabe, K. A. (1994). *Chemistry of Nucleosides and Nucleotides*, Vol. 3, edited by L. B. Townsend, pp. 421–535. New York: Plenum Press.