The Crystal and Molecular Structure of α-D-Ribofuro[1',2':4,5]-2-oxazolidone, a Decomposition Product of 5-Azacytidine

BY PHIRTU SINGH AND DEREK J. HODGSON

Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina 27514, U.S.A.

(Received 26 September 1975; accepted 13 February 1976)

 α -D-Ribofuro[1',2':4,5]-2-oxazolidone, C₆H₉NO₅, has been identified by X-ray diffraction methods as a decomposition product of 5-azacytidine obtained from a chloroform-methanol solution. The product is identical with that obtained by boiling 5-azacytidine in distilled water. The compound crystallizes in the monoclinic space group, P2₁, with the unit-cell constants a=7.369 (3), b=8.529 (4), c=6.377 (3) Å, $\beta=113.00$ (3)°. The structure was refined using 1041 counter data to an R of 0.033. 5-Azacytidine, which is the s-triazine riboside, 1- β -D-ribofuranosyl-5-azacytosine, cyclizes by forming a bond between C(2) of the s-triazine base and O(2') of the ribose, yielding the observed unusual N-glycosyl cyclic urethane with the α -anomeric configuration. The five-membered oxazolidone ring is approximately planar with the carbonyl carbon atom and the three atoms attached to it coplanar. The two carbon-nitrogen bonds and the two carbon-oxygen bonds in the oxazolidone ring have different lengths, the bonds involving the carbonyl carbon atom being shorter than those involving ribose carbon atoms. The glycosyl bond length of 1.437 Å is smaller than that normally observed in nucleosides. The dihedral angle O(4')-C(1')-N(1)-C(2) is 72.5°. The ribose has the C(4')-endo puckering which is different from that observed for some α -nucleosides. The conformation around the C(4')-C(5') bond is the usual gauche-gauche.

Introduction

5-Azacytidine (I), an analog of cytidine in which C(5)-H has been replaced by N, is an important anticancer agent (Suhadolnik, 1970) which is effective against various types of leukemia and is presently showing considerable clinical promise (Hrodek & Vesely, 1971). Consequently, the chemical properties of (I) have received considerable attention, and several diverse decomposition products have been detected under a variety of conditions (Pithova, Piskala, Pitha & Sorm, 1965; Piskala & Sorm, 1964; Suhadolnik, 1970).

In an effort to crystallize 5-azacytidine itself from chloroform-methanol solution as a part of our continuing structural investigation of azanucleosides (Singh & Hodgson, 1974*a*, *b*, *c*) we obtained a small quantity of colorless crystals of unknown composition; our subsequent crystallographic analysis (see below) demonstrated that this material was the title compound (II), which had previously been suggested by Pithova *et al.* (1965) as the hydrolysis product (along with guanidine) of (I) in neutral aqueous medium.



The present structural analysis, therefore, confirms that the decomposition of 5-azacytidine in these organic solvents (which may, of course, contain small quantities of water) is the same as that in neutral aqueous solution and confirms the stereochemistry of the product as assigned by Pithova *et al.* (1965).

Experimental

5-Azacytidine was purchased from Calbiochem, Los Angeles, California. Colorless, plate-like single crystals of the title compound were obtained from it as described in the previous section. Preliminary unit-cell constants and the space group were determined from Weissenberg and precession photographs of a crystal of dimensions $0.5 \times 0.3 \times 0.2$ mm, mounted along its long direction, which was chosen as the a axis. Refined cell constants were obtained by a least-squares refinement using the angular settings of 12 accurately centered reflections on a Picker four-circle automatic diffractometer equipped with a pulse-height analyzer and a graphite monochromator. The radiation used was monochromatized Mo $K\alpha$. The intensity data were collected by the $\theta/2\theta$ scan technique. The relevant crystallographic data are presented in Table 1. The data were processed by the method of Corfield, Doedens & Ibers (1967) with the usual corrections in the intensities for background and for Lorentz and polarization factors, and with estimated standard deviations assigned to them (Meyer, Singh, Hatfield & Hodgson, 1972). Absorption corrections were minimal and were not made.

2330

Table 1. Crystallographic data

$V = 368.9 \text{ Å}^3$
$D_x = 1.577 \text{ g cm}^{-3}$
$D_m = 1.56$ (2)
$Z = 2(C_6 NO_5 H_9)$
Scan range = $1 \cdot 8^\circ + (2\theta_{\alpha 2} - 2\theta_{\alpha 1})^\circ$
Maximum $2\theta = 59^{\circ}$
0 s on each side of peak
$R_{w} = 0.047$
NV = 145

Solution and refinement

The solution of the structure and its subsequent refinement did not prove to be routine since the chemical composition of the crystal was unknown. It was known, however, that 5-azacytosine, 5-azauracil and ribose are among some of the decomposition products of 5-azacytidine. To start the phasing process by direct methods, therefore, it was assumed that one of the species present in the cell was a hydrate of 5-azacytosine, with 3.5 mol of H₂O required to account for the crystallographically calculated molecular weight. Normalized structure amplitudes (E's) were calculated based on the above assumptions, and 126 with magnitudes greater than 1.50 were used as input to the direct-methods phasing program MULTAN (Main, Germain & Woolfson, 1971). Of the 16 multiple solutions, the set with the second lowest R_{Karle} of 15.97% and the sixth highest absolute figure of merit |FOM| of 1.1547 gave the correct solution. The set with the lowest R_{Karle} (15.73%) which had the fourth highest |FOM| of 1.1577 yielded Fourier peaks in the E map which could not be connected to form a reasonable molecule. The *E* map of the set with the correct solution did not have any spurious peaks among its top 12 peaks.

Least-squares refinement of the above 12 peak positions, all with carbon scattering factors and with their positional and isotropic thermal parameters plus a scale factor being refined, gave a conventional residual, $[R=\sum |F_o|-|F_c|/\sum|F_o|]$ of 0.17, and a weighted residual, $R_w \quad \{R_w=[\sum w(|F_o|-|F_c|)^2/\sum w(F_o)^2]^{1/2}\}$ of 0.20 which seemed to indicate that no more nonhydrogen atoms were present in the structure. The fused-ring geometry of the compound was, of course, obvious by now. This fact, coupled with the presence of the exocyclic C=O, OH and CH₂OH groups with differing bond lengths, helped to identify all the non-hydrogen atoms, except O(2'), as to their chemical type (see Fig. 1 for the numbering system).

Since O(2') could, on the basis of the structure of the precursor (I), reasonably be either NH or O, both models were tested. With this site occupied by N, and all other atoms, including the hydrogen atoms, included, least-squares refinement converged to values of R and R_w of 0.046 and 0.066; a subsequent difference Fourier map showed no peak which could be attributed to a hydrogen atom attached to this nitrogen atom, but did have its largest peak (*ca* 0.5 e Å⁻³) at the nitrogen

position. Refinement with an oxygen atom in this site converged to the final values of R=0.033 and $R_w=0.047$.

In the final cycle of least squares a correction for secondary extinction of the type suggested by Zachariasen (1968) was applied in the manner described elsewhere (Meyer *et al.*, 1972); the value of the extinction coefficient, 9 (5) $\times 10^{-9}$, is probably not significantly different from zero.

Although no attempt was made to establish the absolute configuration of the molecule owing to the extremely small anomalous dispersion term, $\Delta f''$, (0.008 e) of oxygen (Cromer & Liberman, 1970) for Mo K α radiation, the atomic coordinates (see below) in Table 2 are given for the D-enantiomer of ribose on the reasonable assumption that the D-configuration of the sugar in the starting compound, 5-azacytidine, remains unchanged during the cyclization process.

The least-squares refinements were carried out on F, the function minimized being $\sum w(|F_o| - |F_c|/g)^2$; weights, w, were taken as $w = 4F_o^2/\sigma^2(F_o)^2$ where $\sigma^2(F_o)^2$ is the variance of $(F_o)^2$.* The form factors for O, N and C were taken from *International Tables for X-ray Crystallography* (1962) and those for H from Stewart, Davidson & Simpson (1965).

Results

The positional and thermal parameters derived from the last cycle of refinement are given in Tables 2 and 3 respectively. Deviations of atoms from various leastsquares planes are given in Table 4. The distances and angles associated with the possible hydrogen bonds are shown in Table 5. Bond distances and bond angles are depicted in Fig. 1.

Description and discussion

The unusual structure of the decomposition product of the nucleoside analog 5-azacytidine is depicted in Fig. 2 as a stereo pair, which shows the five-membered

* A list of structure factors has been deposited with the British Library Lending Division as Supplementary Publication No. SUP 31697 (7 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 13 White Friars, Chester CH1 1 NZ, England.



Fig. 1. Bond lengths and angles in the α -D-ribofuro[1',2':4,5]-2-oxazolidone molecule. The e.s.d.'s of angles are 0.2°.

ribose ring fused to another five-membered heterocyclic ring with C(1')-C(2') as the common bond. A rather surprising feature of the structure is the change in the

Table 2. Positional parameters ($\times 10^4$, for H $\times 10^3$) for C₆H₉NO₅

	x	У	Z
N(1)	- 1921 (2)	-2011*	3399 (3)
C(2)	-2817(3)	1905 (4)	1126 (3)
O(2)	- 4457 (2)	-2351(4)	-121(3)
O(2')	- 1627 (2)	-1195 (3)	258 (2)
C(1')	59 (3)	-1412 (4)	4267 (3)
C(2')	173 (3)	-706 (3)	2099 (3)
C(3')	1924 (2)	- 1539 (4)	1892 (3)
O(3′)	1644 (2)	- 1722 (4)	-411 (2)
C(4')	1985 (2)	- 3076 (3)	3145 (3)
O(4')	1516 (2)	- 2607 (3)	5036 (2)
C(5')	3938 (3)	- 3923 (4)	4052 (3)
O(5′)	5524 (2)	- 2982 (3)	5482 (2)
H(N1)	- 302 (5)	-242 (4)	359 (6)
H(C1')	29 (4)	-69(3)	534 (5)
H(C2')	39 (4)	44 (4)	245 (5)
H(C3')	313 (5)	-100 (4)	265 (6)
H(O3')	219 (8)	-190 (10)	-137 (10)
H(C4')	114 (4)	- 373 (4)	230 (5)
H1(C5')	402 (6)	-419 (5)	221 (7)
H2(C5')	414 (4)	-488 (4)	514 (5)
H(O5')	574 (9)	-284 (9)	719 (9)

* The y coordinate of one atom is arbitrarily fixed in polar space group $P2_1$.

Table 4. Deviations in Å of atoms from the leastsquares planes through atoms of the ribose and oxazolidone moieties

	Plane I	Plane II	Plane III
C(1')	-0.026*	0.046*	0.059
C(2')	0.024*	-0.047*	-0.078
C(3')	<i>−</i> 0·016*		
C(4')	0.571		
O(4')	0.017*		
C(5')	0.218		
O(2')		0.035*	0.000*
C(2)		-0.003*	-0.001*
O(2)		0.006	0.000*
N(1)		-0.030*	0.000*

* Atom included in the least-squares calculation of the plane.

anomeric character at C(1') from the β of 5-azacytidine (I) to the α of the decomposition product (II), Fig. 2 [note the *trans* configuration of the hydrogen atoms or of the substituents attached to C(1') and C(4')].

The ribose

The ribose sugar of the present molecule, which is an α cyclic *N*-glycosyl urethane, has features which have been either rarely or never observed in β -nucleosides. For example, the puckering observed for the ribose ring is C(4')-endo envelope, ⁴E, (Table 4, plane I) which is one of the sugar conformations observed

Table 3. Thermal parameters $(U_{ij} \text{ in } \text{Å}^2)$ for C₆H₉NO₅

Anisotropic thermal parameters are $\times 10^4$. Isotropic thermal parameters are $\times 10^2$. The form of the anisotropic thermal ellipsoid is exp $[-2\pi^2(U_{11}h^2a^{*2}+U_{22}k^2b^{*2}+U_{33}l^2c^{*2}+2U_{12}hka^*b^*+2U_{13}hla^*c^*+2U_{23}klb^*c^*)]$.



Fig. 2. Stereoscopic view of the α -D-ribofuro[1',2':4,5]-2-oxazolidone molecule. Hydrogen atoms are shown as open circles of arbitrary size; thermal ellipsoids of other atoms are drawn at the 40% probability level.

 Table 5. Possible hydrogen bonds and short van der

 Waals contacts

(a) Distances and angles associated with the possible hydrogen bonds

$A-H\cdots B$	A-H	H · · · <i>B</i>	• A ••• B	$A-\mathbf{H}\cdot\cdot\cdot E$
$N(1) - H \cdots O(5')$	0.93 Å	1·96 Å	2·82 Å	152°
$O(3')-H\cdots O(2)$	0.86	2.31	2.86	121
$O(3')-H\cdots O(4')$	0.86	2.23	2.97	143
$O(5')-H\cdots O(2)$	1.04	1.83	2.85	166

(b) Short van der Waals contacts (Å)

Distances shown are those less than the sum of their respective van der Waals radii, H=1.20, O=1.40, C=1.70 Å (Pauling, 1963). See also Hopfinger (1973).

$H(N1) \cdots C(5')$	2.70	$H1(C5') \cdots O(2')$	2.52
$H(C1') \cdots H(O3')$	2.27	$H2(C5') \cdots H(C3')$	2.17
$H(C3') \cdots H2(C5')$	2.17	$H(O5') \cdots C(2)$	2.45
$H(O3') \cdots HC(1')$	2.27	$H(O5') \cdots O(2')$	2.57
$H(O3') \cdots C(1')$	2.64		

for α nucleosides (Sundaralingam, 1971), but in a tabulation (Sundaralingam, 1973) of conformational parameters for the sugar puckering in some eighty β -nucleosides, β -dinucleosides and β -polynucleotides there is not a single listing exhibiting the C(4')-endo puckering. It is noteworthy, however, that this conformation has been observed in the cyclic β -nucleoside cyclo ara-C (Brennan & Sundaralingam, 1973). The conformation of the C(5')-O(5') bond around the C(4')-C(5')-O(5') dihedral angle being 61.4 and the C(3')-C(4')-C(5')-O(5') 54.6°.

Some of the bond lengths in the α -D-ribose moiety are slightly, but significantly, different from those normally observed in the β -isomers. Thus the C(1')– C(2') bond, at 1.540 (3) Å, is 0.021 Å longer than the mean length of this bond in β -nucleosides (Sundaralingam, 1973). Most of this lengthening can be attributed to the anomeric change at C(1') since it has also been observed in other α -sugars, such as α -D-2'-amino-2-deoxyadenosine (Rohrer & Sundaralingam, 1970a) and in α -pseudouridine (Rohrer & Sundaralingam, 1970b). It is also noteworthy that the usual inequality of the C(1')–O(4') and C(4')–O(4') bonds observed in the β -nucleosides, in which the former is shorter than the latter, is retained in this α -'nucleoside', although to a lesser extent.

The ring bond angles in the ribose moiety differ quite significantly (by as much as $\sim 2.5^{\circ}$) from those observed in other nucleoside structures (Singh & Hodgson, 1974*a*; Sundaralingam, 1973). This, to a large extent, is due to the unusual puckering, ⁴E, of the sugar since sugar bond angles are known to be sensitive to the mode of puckering and also perhaps to the fused-ring geometry of the molecule.

The oxazolidine moiety

The five atoms C(1'), C(2'), O(2'), C(2) and N(1) of the ring are coplanar within ± 0.05 Å, as seen from Table 4 (plane II). The carbonyl oxygen atom O(2)

is only 0.006 Å away from this plane. A more perfect plane (plane III) is calculated through the three ring atoms N(1), C(2) and O(2') and the carbonyl oxygen O(2) which shows them to be off of this plane by less than ± 0.001 Å, thus establishing the coplanarity of the three bonds attached to C(2). This facilitates the considerable delocalization of electrons in the three bonds attached to C(2) which is manifested in their bond lengths. Thus, according to Pauling's (1960) values for the single-bond radii of C, N and O, the N(1)-C(2) bond is shortened by 0.13 and the C(2)-O(2') bond is shortened by 0.08 Å, both showing considerable double bond character. The length of the carbonyl bond C(2)-O(2), at 1.220 (3) Å, is in the range normally found for this type of bond in purines and pyrimidines where it is delocalized with the ring (see, for example, Singh & Hodgson 1974d; and references therein). The glycosyl bond N(1)–C(1'), at 1.437 (3) Å, is one of the shortest ever observed (Singh & Hodgson, 1974c). The bond $\dot{C}(2')$ –O(2'), 1.447 (2) Å, being involved in a cyclic system in which O(2') is attached to two carbon atoms, is slightly longer than C(3')-O(3'), 1.410 (3) Å.

Packing and hydrogen bonding

The packing of the molecules is dictated mainly by the formation of hydrogen bonds in a head to tail fashion involving the ribose hydroxyl groups [O(3')-H]and O(5')-H] and the oxazolidone NH and carbonyl groups. O(3')-H and O(5') of one molecule hydrogen bond to O(2) and N(1)-H respectively of the molecule related to it by the +a translation while O(2) of one molecule hydrogen bonds to O(5')-H of a molecule related to it by a (-a, -c) translation, thus forming a hydrogen-bonded network of molecules lying in the ac plane. There are no hydrogen bonds connecting molecules along the **b** direction along which, therefore, van der Waals interactions are the main stabilizing force. There are no hydrogen bonds connecting two ribose or two oxazolidone moieties. The hydrogen atom attached to O(3') may be involved in a bifurcated hydrogen bond interacting with both O(2) and O(4'). The oxazolidone ring oxygen atom O(2') is not involved in any hydrogen-bond formation, which is not surprising since ether oxygen atoms are not known to form hydrogen bonds readily. The distances and angles associated with the hydrogen bonds [Table 5(a)] are within the range normally observed in crystals. There are no short contacts involving only non-hydrogen atoms [Table 5(b)]. There are, however, a substantial number of short contacts involving hydrogen atoms with each other and with carbon and oxygen atoms. None of these is unusually short if one considers the van der Waals radius of hydrogen to be as short as 1.05 and those of $C(sp^3)$ as 1.65, $C(sp^2)$ as 1.50 and O(sp) as 1.30 Å (Hopfinger, 1973). Interestingly, the shortest $C \cdots H$ contact $[C(2) \cdots H(O5')]$ of 2.45 Å involves an sp^2 hybridized carbon; part of this short contact, however, is probably a result of a large experimentally observed O(5')-H distance (1.04 Å).

We are grateful to our colleagues Professor Ernest Eliel and Professor Robert McKee for helpful discussions. This investigation was supported by Public Health Service Research Grant No. CA-15171-02 from the National Cancer Institute.

References

- BRENNAN, T. & SUNDARALINGAM, M. (1973). Biochem. Biophys. Res. Commun. 52, 1348–1353.
- CORFIELD, P. W. R., DOEDENS, R. J. & IBERS, J. A. (1967). Inorg. Chem. 6, 197-204.
- CROMER, D. T. & LIBERMAN, D. (1970). J. Chem. Phys. 53, 1891–1898.
- HOPFINGER, A. J. (1973). Conformational Properties of Macromolecules, p. 39. New York: Academic Press.

HRODEK, O. & VESELY, J. (1971). Neoplasma, 18, 493-503.

International Tables for X-ray Crystallography (1962). Vol. III, pp. 201–209. Birmingham: Kynoch Press.

MAIN, P., GERMAIN, G. & WOOLFSON, M. M. (1971). MUL-TAN. A Computer Programme for the Automatic Solution of Crystal Structures, Dept. of Physics, Univ. of York.

- MEYER, M. H., SINGH, P., HATFIELD, W. E. & HODGSON, D. J. (1972). Acta Cryst. B28, 1607–1613.
- PAULING, L. (1960). *The Nature of the Chemical Bond*, 3rd ed. Ithaca: Cornell Univ. Press.

- PISKALA, A. & SORM, F. (1964). Coll. Czech. Chem. Commun. 29, 2060–2076.
- PITHOVA, P., PISKALA, A., PITHA, J. & SORM, F. (1965). Coll. Czech. Chem. Commun. 30, 2801–2811.
- ROHRER, D. C. & SUNDARALINGAM, M. (1970a). J. Amer. Chem. Soc. 92, 4956-4962.
- ROHRER, D. C. & SUNDARALINGAM, M. (1970b). J. Amer. Chem. Soc. 92, 4950–4955.
- SINGH, P. & HODGSON, D. J. (1974a). Biochemistry, 13, 5445–5452.
- SINGH, P. & HODGSON, D. J. (1974b). J. Amer. Chem. Soc. 96, 5276–5278.
- SINGH, P. & HODGSON, D. J. (1974c). J. Amer. Chem. Soc. 96, 1239–1241.
- SINGH, P. & HODGSON, D. J. (1974d). Acta Cryst. B30, 1430-1435.
- STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). J. Chem. Phys. 42, 3175–3187.
- SUHADOLNIK, R. J. (1970). Nucleoside Antibiotics, pp. 271-297. New York: John Wiley.
- SUNDARALINGAM, M. (1971). J. Amer. Chem. Soc. 93, 6644-6647.
- SUNDARALINGAM, M. (1973). Conformation of Biological Molecules and Polymers, edited by E. D. BERGMAN & B. PULLMAN, pp. 417–455. New York: Academic Press.

ZACHARIASEN, W. H. (1968). Acta Cryst. A24, 212-216.

Acta Cryst. (1976). B32, 2333

The Crystal and Molecular Structure of 2-Acetamido-2,3-dideoxy-D-*threo*-hex-2-enono-1,4-lactone, C₈H₁₁NO₅

BY ŽIVA RUŽIĆ-TOROŠ AND BISERKA KOJIĆ-PRODIĆ 'Ruđer Bošković' Institute, P.O. Box 1016, 41001 Zagreb, Yugoslavia

(Received 19 January 1976; accepted 17 February 1976)

2-Acetamido-2,3-dideoxy-D-*threo*-hex-2-enono-1,4-lactone crystallizes in the space group $P_{2_12_12_1}$ with a=6.178, b=6.925, c=21.523 Å, Z=4. The structure was solved by direct methods and refined to an R of 0.037. The five-membered lactone ring is planar within experimental error. The molecules are packed into a three-dimensional network by intermolecular O(3)-H···O(5) (2.777), O(4)-H···O(3) (2.854) and N-H···O(1) (3.033 Å) hydrogen bonds.

Introduction

Treatment of 2-acetamido-2-deoxy-D-aldono-1,4-lactone (aldono = glucono, mannono or galactono) with methanolic KOH gives a mixture of two unsaturated lactones in approximately equal proportions. These two lactones were identified as 2-acetamido-2,3dideoxy-D-threo-hex-2-enono-1,4-lactone and its Derythro isomer.



The preparations, the IR and NMR spectral characteristics and the inhibitory activities were described by Pravdić & Fletcher (1971) and Pokorny, Zissis, Fletcher & Pravdić (1975).

The results of the present structure determination of the D-threo isomer confirm the configuration and conformation given by Pravdić & Fletcher (1971).

Experimental

The space group (Table 1) was determined from Weissenberg photographs recorded with Cu $K\alpha$ radiation. The diffraction symmetry and space-group extinction determined $P2_12_12_1$ uniquely.

The intensities were collected on a Philips PW 1100 computer-controlled four-circle diffractometer with