A Novel Intramolecular Hydrogen Bond in the Crystal Structure of 5-Hydroxymethyl-2′-deoxyuridine, an Antiviral and Antineoplastic Nucleoside. Conformational Analysis of the Deoxyribose Ring[1]

George I. Birnbaum,*[2a] Roxanne Deslauriers,[2a] Tai-Shun Lin,[2b] George T. Shiau,[2b] and William H. Prusoff[2b]

Contribution from the Division of Biological Sciences, National Research Council of Canada, Ottawa, Ontario, Canada K1A 0R6, and the Department of Pharmacology, Yale University School of Medicine, New Haven, Connecticut 06510. *Received November 19, 1979*

Abstract: The three-dimensional structure of 5-hydroxymethyl-2′-deoxyuridine, an inhibitor of herpes simplex and vaccinia viruses, was determined by X-ray crystallography. The crystals belong to the monoclinic space group $P2_1$, and the cell dimensions are $a = 9.301$ (1), $b = 12.302$ (1), $c = 4.862$ (1) Å, $\beta = 97.18$ (1)°. Intensity data were measured with a diffractometer, and the structure was solved by direct methods; least-squares refinement converged at $R = 0.050$. In the $-CH_2OH$ substituent at C(5), the hydroxy group occupies two different positions. The conformation about the glycosyl bond is anti ($\chi_{CN} = 56.4°$), while the $-CH_2OH$ side chain attached to the sugar ring adopts the gauche-gauche conformation. The deoxyribose ring has the relatively rare C(1′) exo pucker, possibly stabilized by a previously unobserved intramolecular hydrogen bond C(6)–H··· O(4′). ^1H NMR spectroscopy was used to determine the conformation in solution. The spectra indicate an anti conformation about the glycosyl bond and a 43% contribution of the gauche-gauche conformation in the $-CH_2OH$ side chain. The sugar ring may exist in a 32:68 equilibrium of ^3E and ^2E conformers. However, the conventional interpretation may be inadequate since better agreement with observed coupling constants is obtained by assuming equal contributions of ^3E, $_1E$, and ^2E conformers. A survey of 5-substituted 2′-deoxyuridines reveals no correlation between the electronegativity of the substituents and either bond lengths or antiviral properties of these substances.

5-Hydroxymethyl-2′-deoxyuridine (hm^5dU)[3] inhibits the replication of *Escherichia coli* $15T^-$,[4] the reproduction of Ehrlich ascites carcinoma cells[5] as well as a number of other mammalian cells,[6] and the propagation of vaccinia and herpes simplex viruses.[7]

Early studies on the synthesis and metabolism of 5-hydroxymethylpyrimidines and their derivatives have been reviewed by Ulbricht.[8] Baker et al.[9] prepared hm^5dU by base-catalyzed hydroxymethylation of 2′-deoxyuridine, and, recently, hm^5dU has been synthesized by Langen and Bärwolff[5] by selective monobromination of the 5-methyl carbon of thymidine under ultraviolet catalysis. 5-Hydroxymethylpyrimidine 2′-deoxyribonucleotides have also been prepared enzymically.[10]

5-Hydroxymethyl-2′-deoxyuridylate competitively inhibits both prokaryotic and eukaryotic thymidylate synthetase[11-14] and has been found in the DNA of several *Bacillus subtilis* phages in place of thymidine.[15]

Santi and Sakai[12] suggested that the inhibitory effect depends not only on the 5 substituent being of a size that avoids steric hindrance, but also on its electron-withdrawing property. More recently, Egert et al.[16] attempted to demonstrate the influence of 5 substitution on the bond lengths in several uridine derivatives. An X-ray analysis of hm^5dU, yielding both steric and electronic information about the $-CH_2OH$ substituent at C(5), was expected to contribute to our understanding of the effects of such substituents. Furthermore, bearing in mind that its conformation must be related to its biological activity, it is of value to determine not only the crystal structure but also the conformation in solution. As we showed recently,[17] such a parallel study may lead to a more meaningful interpretation of ^1H NMR spectra.

Experimental Section

hm^5dU was prepared by the base-catalyzed hydroxymethylation of 2′-deoxyuridine with paraformaldehyde at 60 °C according to the method reported by Baker et al.[9] with modification.

Colorless prisms were obtained from chloroform–ethanol (4:1); the crystals belong to the monoclinic space group $P2_1$. There are two molecules in a cell with dimensions $a = 9.301$ (1), $b = 12.302$ (1), $c = 4.862$ (1) Å, $\beta = 97.18$ (1)°. The observed and calculated crystal densities are 1.549 and 1.544 g cm^{-3}, respectively. Intensity data were collected on a Picker four-circle diffractometer with Cu Kα radiation. There were 987 unique reflections accessible to the diffractometer ($2\theta \leq 130°$) and all but four were considered observed. The intensities were corrected for Lorentz and polarization factors; absorption corrections were considered unnecessary in view of the low value of μ (10.7 cm^{-1}).

The structure was determined by the multisolution approach to direct methods, and the first E map revealed the positions of the 18 nonhydrogen atoms. Anisotropic refinement of these atoms by block-diagonal least squares converged at $R = 0.11$. A difference Fourier map showed that O(7) was disordered, occupying two distinct positions. Subsequent refinement included occupancy factors for these two sites which converged at 0.70 (3) and 0.28 (3) for the major and minor site, respectively, and were then fixed at 0.71 and 0.29. Most hydrogen atoms were located on difference Fourier maps and two were placed in calculated positions; however, it was not possible to find the hydrogen atoms attached to O(7) and O(5′). The contributions of the hydrogen atoms were included in the calculation of the structure factors, but their parameters were not refined. All scattering factors were taken from the International Tables for X-ray Crystallography,[18] and the oxygen curve was corrected for anomalous dispersion. A weighting scheme was chosen which made the average values of $w(\Delta F^2)$ independent of $|F_o|$ and $\sin^2\theta$. After the final cycle the average parameter shift equalled 0.14σ and the final value of R was 0.050. A stereoscopic representation of the molecule is shown in Figure 1. All coordinates are listed in Table I; anisotropic thermal parameters, as well as a list of observed and calculated structure factors, are provided as supplementary material (see paragraph at the end of the paper).

The 270-MHz ^1H NMR spectra of D_2O solutions (16 mg/cm^3) were obtained using the Bruker HX-270 system of the Southern New England High Field Facility at New Haven, Conn., which is equipped with a BNC data system and is capable of performing 16K transforms. The spectra were recorded at room temperature with $(CH_3)_3COH$ as the internal reference, which was locked at 1.280 ppm so that the chemical shifts are reported downfield relative to tetramethylsilane with an accuracy of ±0.001 ppm. A local version of the LAOCOON program was used to analyze the spectra.

0002-7863/80/1502-4236$01.00/0 © 1980 American Chemical Society

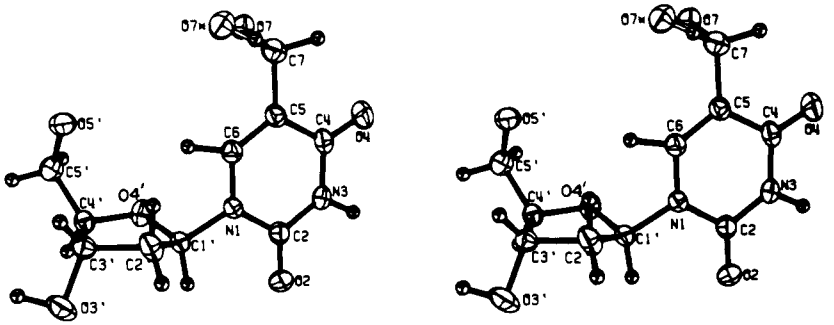


Figure 1. Stereoscopic view of hm^{5}dU; the thermal ellipsoids correspond to 50% probability.

Table I. Fractional Atomic Coordinates[a]

atom	*x*	*y*	*z*
N(1)	6390 (3)	4630 (0)	4141 (6)
C(2)	4979 (4)	4715 (3)	4683 (7)
O(2)	4567 (3)	5351 (3)	6325 (6)
N(3)	4053 (3)	3989 (3)	3215 (7)
C(4)	4404 (4)	3200 (3)	1398 (7)
O(4)	3481 (3)	2565 (3)	362 (7)
C(5)	5900 (4)	3195 (3)	860 (7)
C(6)	6810 (4)	3924 (3)	2260 (7)
C(7)	6383 (4)	2383 (3)	−1114 (8)
O(7)	7097 (4)	1483 (3)	221 (7)
O(7*)	7855 (11)	2361 (9)	−1416 (20)
C(1')	7523 (4)	5310 (3)	5788 (7)
C(2')	8384 (4)	6041 (3)	4117 (8)
C(3')	9888 (4)	6050 (3)	5702 (7)
O(3')	10083 (4)	7005 (2)	7347 (7)
C(4')	9943 (4)	5029 (3)	7515 (7)
O(4')	8533 (3)	4548 (2)	7105 (5)
C(5')	11051 (4)	4198 (3)	6998 (7)
O(5')	11052 (3)	3988 (2)	4094 (5)
H(3)	311	409	322
H(6)	795	404	261
H1(7)	708	276	−224
H2(7)	545	225	−253
H(1')	711	595	664
H1(2')	805	682	369
H2(2')	863	565	239
H(3')	1061	603	431
HO(3')	1115	684	808
H(4')	1017	529	951
H1(5')	1083	352	823
H2(5')	1195	453	768

[a] The coordinates of the nonhydrogen atoms were multiplied by 10^4 and those of the hydrogen atoms by 10^3.

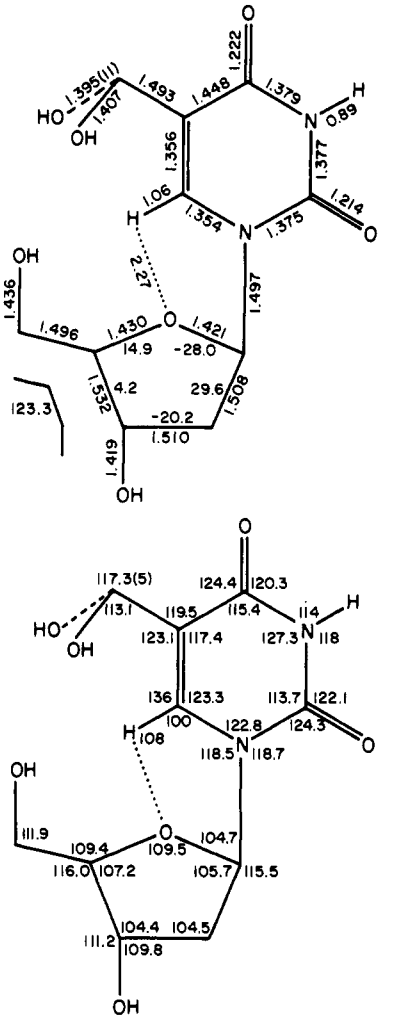


Figure 2. (Top) Bond distances (in ångstroms) and torsion angles (in degrees). Unless otherwise indicated, their estimated standard deviations are 0.004–0.005 Å and 0.4°, respectively. (Bottom) Bond angles (in degrees); unless otherwise indicated, the esd's for nonhydrogen atoms are 0.3°.

Results

Geometry of the Base. The pyrimidine ring is not completely planar, deviations from planarity ranging up to 0.027 Å (Table II). Of the ring substituents, C(1') shows the largest displacement from the mean plane, a phenomenon which is not unusual. As mentioned above, the hydroxy group in the 5 substituent adopts two distinct positions. Somewhat surprisingly, it is the minor position, O(7*), which is close to the mean plane of the ring. The conformation of this side chain can also be expressed by the torsion angles C(6)-C(5)-C(7)-O(7), −78.4°, and C(6)-C(5)-C(7)-O(7*), −2.5°. From the refined occupancy factors, 0.71 and 0.29, one can calculate that the difference in energy between these two conformations (ΔE) is rather low, viz. 0.5 kcal/mol.

All bond lengths and bond angles are shown in Figure 2. An examination of the exocyclic bond angles reveals some abnormal values: first, at N(1) where the two angles C(2)-N(1)-C(1') and C(6)-N(1)-C(1') are equal, while in uridine,[19] thymidine,[20] and in other pyrimidine nucleosides[21] the former angle is 4.5–7.5° smaller than the latter. Second, the angle C(5)-C(6)-H(6) is 136°, while N(1)-C(6)-H(6) is 100°. This means that C(6)-H(6) and the glycosyl bond N(1)-C(1') are bent toward each other. Consequently, the distances between O(4') on one hand and C(6) and H(6) on the other are 2.786 and 2.27 Å, respectively, indicating a C-H···O hydrogen bond (see below).

The Deoxyribose. The conformation about the glycosyl bond is, as usual, anti, and the torsion angle χ_{CN} = 56.4° is within the normal range. The $-CH_2OH$ side chain adopts the gauche–gauche conformation (Figure 3), which is the one most commonly observed. The N(1)-C(1')-O(4') bond angle is

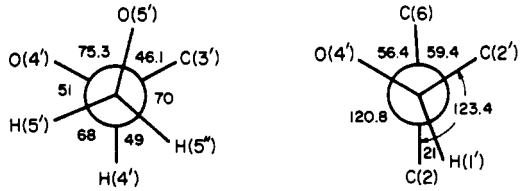


Figure 3. Newman projections along (left) the C(4')-C(5') bond and (right) along the N(1)-C(1') bond.

Table II. Least-Squares Planes and Deviations of Atoms from Them[a]

	plane 1[b]			plane 2[b]	
atom	Δ, Å	atom	Δ, Å	atom	Δ, Å
N(1)	−0.013	C(1')[c]	−0.145		
C(2)	0.008	O(2)[c]	−0.002	C(2')	0.020
N(3)	0.019	O(4)[c]	−0.094	C(3')	−0.024
C(4)	−0.027	C(7)[c]	−0.018	C(4')	0.024
C(5)	0.006	O(7*)[c]	−1.294	O(4')	−0.009
C(6)	0.023	O(7*)[c]	−0.078	C(1')[c]	−0.416

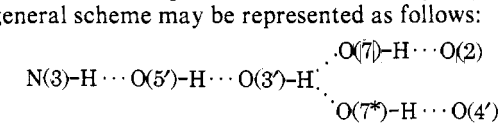
[a] Estimated standard deviations are 0.002-0.004 Å. [b] Plane 1: $0.1078X - 0.6664Y + 0.7378Z + 1.7212 = 0$. Plane 2: $0.3966X - 0.5626Y - 0.7254Z + 2.6487 = 0$. [c] Atoms not included in the calculation of the plane.

about 4° smaller than normal. This also tends to bring H(6) and O(4') closer to each other and thus provides additional evidence for the proposed C-H···O hydrogen bond. Another notable feature of the ring geometry is the nearly equal lengths of the C(1')-O(4') and C(4')-O(4') bonds. The latter bond is normally 0.02-0.05 Å longer than the former.

The most interesting aspect of the sugar ring is its conformation. The ring is an almost ideal envelope with a C(1') exo pucker. The pseudorotation parameters[22] are $P = 129.0°$ and $\tau_m = 32.1°$, and the displacement of C(1') from the mean plane through the other four ring atoms is 0.416 Å. This conformation has been encountered in relatively few cases and will be further discussed below.

Hydrogen Bonds. The C(6)-H···O(4') hydrogen bond described above is unusual in its nature as well as in its geometry (Table III). Nevertheless, our belief that it does exist is supported by several arguments. (a) H(6) has been found to participate in hydrogen bonds in previously determined nucleoside and nucleotide structures.[23] (b) The only obvious reason for the unusual bond angles at C(6), N(1), and C(1') (see above) is to bring H(6) and O(4') closer to each other. A similar angular deformation at C(6) was noted before.[23b] (c) The H···O distance of 2.27 Å is well within the normal range for C-H···O hydrogen bonds. (d) While the C-H···O angle of 108° is very small, it is not unique, particularly for an *intra*molecular hydrogen bond. In a recent crystal-structure analysis,[24] we found an N-H···O angle of 100° in an intramolecular hydrogen bond. More significantly, O-H···O angles of 106.4 (6) and 106.8 (4)° were found in neutron diffraction studies which yield about ten times more precise positions of hydrogen atoms.[25] Such acute angles have been correlated with weak hydrogen bonds.[26]

All -NH and -OH protons are involved in hydrogen bonds. The general scheme may be represented as follows:



Where it was not possible to locate a hydrogen atom, the evidence for a hydrogen bond is supplied by the O···O distance. It should be pointed out that the proton attached to O(3') does not participate in a bifurcated hydrogen bond. It is being accepted either by O(7) or by O(7*), depending on which site is occupied in a given unit cell.

Table III. Distances and Angles for Hydrogen Bonds

	distances, Å		angles, deg,
D *A*	*D*···*A*	H···*A*	*D*-H···*A*
C(6)-H···O(4') (*x*, *y*, *z*)	2.786	2.27	108
N(3)-H···O(5') (−1 + *x*, *y*, *z*)	2.876	2.02	163
O(5')-H···O(3') (2 − *x*, −½ + *y*, 1 − *z*)	2.715		
O(3')-H···O(7) (2 − *x*, ½ + *y*, 1 − *z*)	2.813	1.79	172
O(3')-H···O(7*) (2 − *x*, ½ + *y*, 1 − *z*)	2.615	1.88	126
O(7)-H···O(2) (1 − *x*, −½ + *y*, 1 − *z*)	2.793		
O(7*)-H···O(4') (*x*, *y*, −1 + *z*)	2.876		

Table IV. Proton Chemical Shifts and Coupling Constants[a]

$\delta_{1'}$	6.33	$J_{1'2'}$	6.7
$\delta_{2'}$	2.42	$J_{1'2''}$	6.6
$\delta_{2''}$	2.44	$J_{2'2''}$	−14.0[b]
$\delta_{3'}$	4.51	$J_{2'3'}$	6.5
$\delta_{4'}$	4.08	$J_{2''3'}$	3.8
$\delta_{5'}$	3.89	$J_{3'4'}$	4.1
$\delta_{5''}$	3.81	$J_{4'5'}$	4.3
δ_6	7.93	$J_{4'5''}$	4.4
$\delta_{7'}$	4.40	$J_{5'5''}$	−12.2
$\delta_{7''}$	4.40		

[a] Chemical shifts in parts per million (ppm) from tetramethylsilane; coupling constants in hertz. [b] Arbitrary; when $\delta_i \simeq \delta_j$, the spectrum is independent of J_{ij}.

The bases are not stacked and, apart from the hydrogen bonds, there are no intermolecular contacts shorter than the sum of van der Waals radii.

NMR Analysis. The assignments of the various resonances in the NMR spectra were made by homonuclear decouplings, comparison with previously published data,[27] and computer line-shape simulations. The results are presented in Table IV. The assignments for H(2') and H(2") were based on the observations of the near equalities $J_{1'2''} \simeq J_{2'3'}$ and $J_{2''3'} \simeq J_{3'4'}$.[28] As shown in Figure 4, the computer-simulated spectrum is a good reproduction of its experimental counterpart.

Stable conformations about the glycosyl bond in nucleosides are found in both the syn and the anti ranges.[29] In solution, both conformations may exist in equilibrium.[30] Chemical-shift considerations have been used to determine the syn-anti conformational preference.[31] The anti conformation in 2'-deoxythymidine in D_2O correlates with a chemical shift of 2.37 ppm for H(2'), while 6-methyl-2'-deoxyuridine, which is constrained to exist in a syn conformation, shows a $\delta_{H_{2'}}$ of 2.95 ppm, 0.6 ppm to low field of an anti conformer.[31c] The value observed for hm^{5}dU is 2.42 ppm, thus appearing to favor a predominantly anti conformation. The relative populations of syn and anti conformers could only be estimated by measurement of intramolecular ^{1}H-^{1}H nuclear Overhauser effects[30b,32] or ^{1}H spin-lattice relaxation times on specifically deuterated compounds.[33]

The conformation of the deoxyribose ring may be assessed by assuming a C(2') endo ⇄ C(3') endo equilibrium. The percentage of the 3'-endo conformer in the equilibrium mixture can be estimated[34] with the formula % 3'-endo = $100J_{3'4'}/(J_{1'2'} + J_{3'4'})$, which yields a value of 38%. It should be noted that the coupling constants (and hence the conformation) are very similar to those in deoxyuridine[27a] and in AIU.[17] Another interpretation of these results will be discussed below.

The conformation about the exocyclic C(4')-C(5') bond may be estimated by the expressions:

$$\text{gauche-gauche} = (13 - \Sigma)/10$$

$$\text{gauche-trans} = (J_{4'5''} - 1.5)/10$$

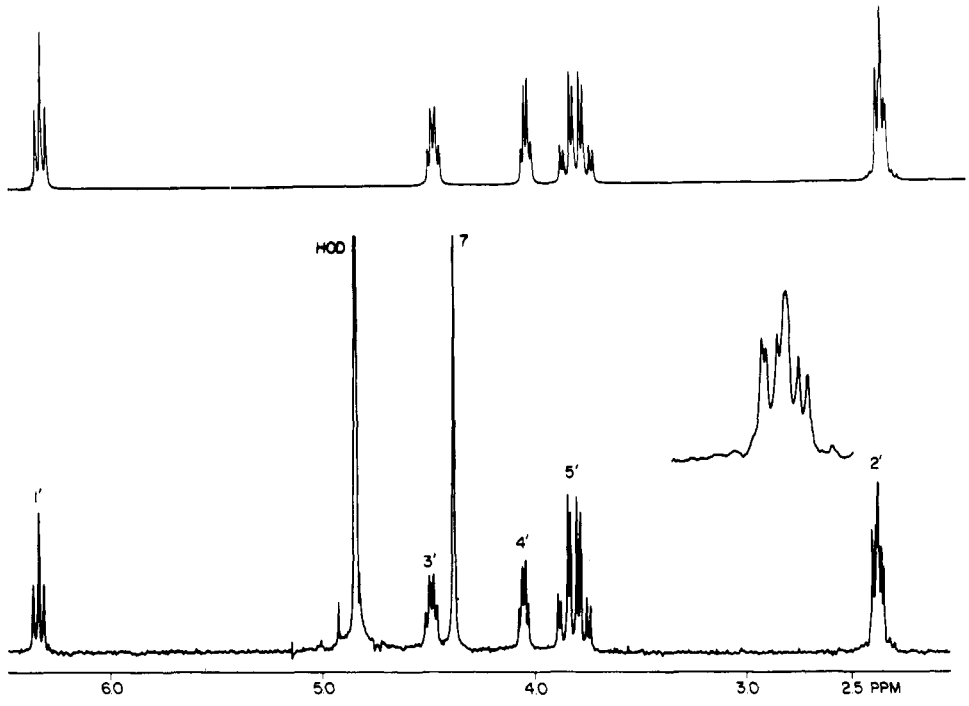


Figure 4. NMR spectrum of hm^{5}dU (bottom) and computer-simulated spectrum (top).

Table V. Bond Lengths (in Å) in 5-Substituted Deoxyuridines

substituent	σ_m	N(1)-C(2)	C(2)-O(2)	N(1)-C(6)	C(5)-C(6)	N(1)-C(1')	C(1')-O(4')	ref
-H	0	1.375	1.195	1.365	1.325	1.475	1.410	*a*
$-CH_3$	−0.07	1.385	1.206	1.374	1.343	1.480	1.434	19
$-CH(CH_3)_2$	−0.07	1.384	1.227	1.401	1.342	1.438	1.420	*b*
$-CH{=}CH_2$	0.05	1.38	1.22	1.36	1.36	1.51	1.43	*c*
-I	0.35	1.37	1.23	1.37	1.34	1.49	1.42	*d*
-Br	0.39	1.39	1.19	1.42	1.34	1.40	1.42	*e*
$-CH_2OH$	0.00	1.375	1.214	1.354	1.356	1.497	1.421	
$-C{\equiv}CH$	0.21	1.387	1.218	1.369	1.357	1.478	1.422	*f*
-Cl	0.37	1.392	1.204	1.369	1.341	1.476	1.419	*g*
$-CF_3$	0.43	1.39	1.20	1.35	1.35	1.47	1.40	
-F	0.34	1.39	1.20	1.36	1.33	1.48	1.44	*h*

[a] Rahman, A.; Wilson, H. R. *Acta Crystallogr., Sect. B* **1972,** *28,* 2260-2270. [b] Czugler, M.; Kalman, A.; Sagi, J. T.; Szabolcs, A.; Ötvös, L. *Ibid.* **1979,** *35,* 1626-1629. [c] Hamor, T. A.; O'Leary, M. K.; Walker, R. T. *Ibid.* **1978,** *34,* 1627-1630. [d] Camerman, N.; Trotter, J. *Ibid.* **1965,** *18,* 203-211. [e] Iball, J.; Morgan, C. H.; Wilson, H. R. *Proc. R. Soc., London, Ser. A* **1966,** *295,* 320-333. [f] Barr, P. J.; Hamor, T. A.; Walker, R. T. *Acta Crystallogr., Sect. B* **1978,** *34,* 2799-2802. [g] Young, D. W.; Morris, E. M. *Ibid.* **1973,** *29,* 1259-1264. [h] Harris, D. R.; MacIntyre, W. M. *Biophys. J.* **1964,** *4,* 203-225.

where Σ is the observed sum of $J_{4'5'}$ and $J_{4'5''}$.[28] Accordingly, the gauche-gauche conformer contributes about 43% to the equilibrium mixture, while the contribution of each of the other two conformers amounts to approximately 29%.

Discussion

The effect of various 5 substituents of nucleosides on their antiviral activity has been the subject of extensive studies.[35] The property of such substituents which is most often mentioned[12,16,36,37] is their ability to donate or withdraw electrons. In particular, it is assumed that in 5-iodo-dU and in 5-bromo-dU, which are known to be incorporated into DNA,[38] the halogen causes a shift of the keto-enol equilibrium toward the enol tautomer, thus facilitating a mispairing of these bases with guanine.[39] Furthermore, Wataya et al.[40] found that electron withdrawal from the uracil base increased the affinity of 5-substituted 2'-deoxyuridylates for thymidylate synthetase. Such electronic changes would be expected to be reflected in changed bond lengths. In a survey of six 5-substituted uridines, Egert et al.[16a] discerned such changes and concluded that electron-withdrawing groups shorten the N(1)-C(6), C(2)-O(2) and C(1')-O(4') bonds and lengthen the C(5)-C(6), N(1)-C(2), and N(1)-C(1') bonds. Electron-donating substituents should have the opposite effect. In Table V we list these bonds for all 5-substituted 2'-deoxyuridines of whose crystal structures we are aware. The compounds are listed in order of increasing electronegativity of the 5 substituents,[41] for which the Hammett σ_m values[42] are also given. *These data do not reveal any systematic trend.* Moreover, there is no correlation between these data and the antiviral activity.[7] While it is true that in several compounds the bond lengths were determined with rather low precision and that more and better data would be useful, our survey still seems to support the conclusion of De Clercq and Torrence,[35] viz., that there is no uniform mechanism of action for all 5-substituted 2'-deoxyuridine derivatives.

To our knowledge, C(6)-H···O(4') hydrogen bonds have not been observed in the past. The formation of such a bond depends primarily on the O(4')···H(6) distance and much less on the requirement of linearity.[43] Some distortions are necessary in order to achieve a sufficiently short distance between O(4') and H(6): a small N(1)-C(6)-H(6) angle and/or a small χ_{CN}

Table VI. Calculated and Observed Coupling Constants

$\%\ ^2E$	$\%\ ^3E$	$\%\ _1E$	$J_{1'2'}$	$J_{1'2''}$	$J_{2'3'}$	$J_{2''3'}$	$J_{3'4'}$	
62	38	0	6.3	6.2	5.3	4.3	4.0	
33	33	33	6.8	6.1	6.0	3.7	4.5	
62	38	0	6.6	6.5	5.6	4.6	4.0	} corrected for
33	33	33	7.1	6.4	6.3	4.0	4.5	} electronegativity
			6.7	6.6	6.5	3.8	4.1	observed

angle. Nevertheless, it is possible that such hydrogen bonds have been overlooked. A review of our structure analysis of lyxofuranosyluracil,[21c] in which χ_{CN} had the unusually low value of 27.0°, reveals an O(4′)···H(6) distance of 2.31 (4) Å, a C(6)-H···O(4′) angle of 107°, and an N(1)-C(6)-H(6) angle of 111 (2)° (vs. a C(5)-C(6)-H(6) angle of 125 (2)°). These C(6)-H···O(4′) bonds may play a role in stabilizing unusual sugar ring conformations, in analogy to the stabilization of the gauche-gauche conformation of the $-CH_2OH$ side chain by C(6)-H···O(5′) bonds.[23]

NMR spectra of nucleosides are being interpreted on the assumption that in solution there is an equilibrium mixture of the two most stable conformers of the sugar ring, viz., C(2′) endo (type S) and C(3′) endo (type N). In our recent crystal-structure analysis of AIU we found the sugar ring in the unusual O(4′) endo conformation,[17] normally considered to correspond to an energy maximum. Nevertheless, we were able to demonstrate that by assuming a 20% contribution of the O(4′) endo conformer in the solution equilibrium we could improve the agreement between the observed and calculated coupling constants.

A similar argument may be advanced for the sugar ring conformation in hm^{5}dU. Cheng and Sarma[27c] recently published calculated coupling constants for pentose rings in various envelope conformations. Using their values, we can calculate the expected coupling constants for a 62:38 equilibrium mixture consisting of C(2′) endo and C(3′) endo conformers. The results, listed in the first row of Table VI, show an average deviation from observed values of 0.5 Hz for the five coupling constants and a maximum deviation of 1.2 Hz. Let us assume, however, that the equilibrium mixture contains not only the two most common conformers but also the C(1′) exo envelope which was found in the crystal. Let us further assume that each component contributes equally to the equilibrium mixture. The average deviation for the coupling constants decreases to 0.3 Hz, and the largest deviation is only 0.5 Hz. It should be borne in mind that Cheng and Sarma did not attempt to modify the Karplus equation for the deoxyribose ring system because, in their opinion, the effect of electronegativity contribution is not known with any reliability. However, we can also examine the agreement between coupling constants after first applying an electronegativity correction of +0.3 Hz to the calculated couplings in which H(2′) and H(2″) participate.[44] Again, we see that an assumption of a three-component equilibrium reduces the average deviation of coupling constants from 0.4 to 0.3 Hz, while the maximum deviation decreases from 0.9 to 0.4 Hz.

Thus, there appears to be no doubt that our assumed three-component equilibrium is at least as likely as the two-component system which is derived from the conventional equation. This being our second demonstration of this fact, we strongly suggest that future NMR analyses of nucleosides consider the occurrence of sugar ring conformations other than C(2′) endo and C(3′) endo, particularly if a different conformation was found in a crystal structure. It is of interest to note that, on the basis of experimental NMR results, Shugar and his colleagues independently concluded that in arabinonucleosides the arabinose ring exists either in a two-component ($^2_1T \rightleftarrows {}^3E$) or in a three-component ($^2_3T \rightleftarrows {}^0_1T \rightleftarrows {}^3E$) equilibrium.[45] Clearly, evidence is mounting which shows that deoxyribonucleosides are not the only nucleosides in which the sugar ring is more flexible than hitherto assumed.

Acknowledgments. All crystallographic computations were carried out with programs written by Ahmed et al.[46] Figure 1 was drawn with the ORTEP program of Johnson.[47] We also wish to acknowledge the support of the Southern New England High Field NMR Facility, made possible by a grant from the Biotechnology Resource program of the National Institutes of Health, and United States Public Health Service Grant CA-05262 from the National Cancer Institute. We wish to thank Dr. Ian C. P. Smith for helpful discussion and Dr. A. H. Tench for sending us the results of his X-ray analysis of 5-trifluoromethyl-2′-deoxyuridine prior to publication.

Supplementary Material Available: Anisotropic thermal parameters and a list of observed and calculated structure factors (6 pages). Ordering information is given on any current masthead page.

References and Notes

(1) Issued as NRCC No. 18297.
(2) (a) National Research Council; (b) Yale University.
(3) Abbreviations employed are: hm^{5}dU, 5-hydroxymethyl-2′-deoxyuridine; AIU, 5-iodo-5′-amino-2′,5′-dideoxyuridine.
(4) Green, M.; Barner, H. D.; Cohen, S. S. *J. Biol. Chem.* **1957,** *228,* 621–631.
(5) Langen, P.; Bärwolff, D. *Acta Biol. Med. Germ.* **1975,** *34,* K7–K10.
(6) Waschke, S.; Reefschläger, J.; Bärwolff, D.; Langen, P. *Nature (London)* **1975,** *255,* 629–630.
(7) De Clercq, E.; Descamps, J.; Torrence, P. F.; Krajewska, E.; Shugar, D. *Proc. Int. Congr. Chemother., Zurich, 10th* **1978,** 352–354.
(8) Ulbricht, T. L. V. *Prog. Nucleic Acid Res. Mol. Biol.* **1965,** *4,* 189–216.
(9) Baker, B. R.; Schwan, T. J.; Santi, D. V. *J. Med. Chem.* **1966,** *9,* 66–72.
(10) (a) Flaks, J. G.; Cohen, S. S. *J. Biol. Chem.* **1959,** *234,* 1501–1506. (b) Roscoe, D. H.; Tucker, R. G. *Virology* **1966,** *29,* 157–166. (c) Alegria, A. H.; Kahan, F. M.; Marmur, J. *Biochemistry* **1968,** *7,* 3179–3186.
(11) Kampf, A.; Barfknecht, R. L.; Shaffer, P. J.; Osaki, S.; Mertes, M. P. *J. Med. Chem.* **1976,** *19,* 903–908.
(12) Santi, D. V.; Sakai, T. T. *Biochem. Biophys. Res. Commun.* **1971,** *42,* 813–817.
(13) Flaks, J. G.; Cohen, S. S. *J. Biol. Chem.* **1959,** *234,* 2981–2986.
(14) Maley, F. *Arch. Biochem. Biophys.* **1962,** *96,* 550–556.
(15) (a) Kallen, R. G.; Simons, M.; Marmur, J. *J. Mol. Biol.* **1962,** *5,* 248–250. (b) Roscoe, D. H.; Tucker, R. G. *Biochem. Biophys. Res. Commun.* **1964,** *16,* 106–110. (c) Okubo, S.; Strauss, B. S.; Stodolsky, M. *Virology* **1964,** *24,* 552–562.
(16) (a) Egert, E.; Lindner, H. J.; Hillen, W.; Gassen, H. G. *Acta Crystallogr., Sect. B* **1978,** *34,* 2204–2208. (b) *Ibid.* **1979,** *35,* 920–923. (c) Hillen, W.; Egert, E.; Lindner, H. J.; Gassen, H. G.; Vorbrüggen, H. *J. Carbohydr. Nucleosides, Nucleotides* **1978,** *5,* 23–32.
(17) Birnbaum, G. I.; Lin, T.-S.; Shiau, G. T.; Prusoff, W. H. *J. Am. Chem. Soc.* **1979,** *101,* 3353–3358.
(18) Ibers, J. A.; Hamilton, W. C., Eds. "International Tables for X-ray Crystallography"; Kynoch Press: Birmingham, U.K., 1974; Vol. IV.
(19) Green, E. A.; Rosenstein, R. D.; Shiono, R.; Abraham, D. J.; Trus, B. L.; Marsh, R. E. *Acta Crystallogr., Sect. B* **1975,** *31,* 102–107.
(20) Young, D. W.; Tollin, P.; Wilson, H. R. *Acta Crystallogr., Sect. B* **1969,** *25,* 1423–1432.
(21) (a) Furberg, S.; Petersen, C. S.; Rømming, C. *Acta Crystallogr.* **1965,** *18,* 313–320. (b) Birnbaum, G. I.; Darzynkiewicz, E.; Shugar, D. *J. Am. Chem. Soc.* **1975,** *97,* 5904–5908. (c) Ekiel, I.; Darzynkiewicz, E.; Birnbaum, G. I.; Shugar, D. *J. Am. Chem. Soc.* **1979,** *101,* 4724–4729.
(22) Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972,** *94,* 8205–8212.
(23) (a) Rubin, J.; Brennan, T.; Sundaralingam, M. *Biochemistry* **1972,** *11,* 3112–3128. (b) Wang, A. H.-J.; Barrio, J. R.; Paul, I. C. *J. Am. Chem. Soc.* **1976,** *98,* 7401–7408. (c) Srikrishnan, T.; Parthasarathy, R. *Acta Crystallogr., Sect. B* **1978,** *34,* 1363–1366. (d) Hillen, W.; Egert, E.; Lindner, H. J.; Gassen, H. G. *Biochemistry* **1978,** *17,* 5314–5320.
(24) Shugar, D.; Huber, C. P.; Birnbaum, G. I. *Biochim. Biophys. Acta* **1976,** *447,* 274–284.
(25) (a) Sequeira, A.; Bernal, I.; Brown, I. D.; Faggiani, R. *Acta Crystallogr., Sect. B* **1975,** *31,* 1735–1739. (b) Panagiotopoulos, N. C.; Jeffrey, G. A.; La Placa, S. J.; Hamilton, W. C. *Ibid.* **1974,** *30,* 1421–1430.
(26) Brown, I. D. *Acta Crystallogr., Sect. A* **1976,** *32,* 24–31.
(27) (a) Schleich, T.; Blackburn, B. J.; Lapper, R. D.; Smith, I. C. P. *Biochemistry* **1972,** *11,* 137–145. (b) Wood, D. J.; Ogilvie, K. K.; Hruska, F. E. *Can. J. Chem.* **1975,** *53,* 2781–2790. (c) Cheng, D. M.; Sarma, R. H. *J. Am. Chem. Soc.* **1977,** *99,* 7333–7348.
(28) Wood, D. J.; Hruska, F. E.; Ogilvie, K. K. *Can. J. Chem.* **1974,** *52,* 3353–3366.
(29) Donohue, J.; Trueblood, K. N. *J. Mol. Biol.* **1960,** *2,* 363–371.
(30) (a) Smith, I. C. P.; Blackburn, B. J.; Yamane, T. *Can. J. Chem.* **1969,** *47,* 513–515. (b) Nanda, R. K.; Tewari, R.; Govil, G.; Smith, I. C. P. *Can. J. Chem.* **1974,** *52,* 371–375.
(31) (a) Dugas, H.; Blackburn, B. J.; Robins, R. K.; Deslauriers, R.; Smith, I. C. P. *J. Am. Chem. Soc.* **1971,** *93,* 3468–3470. (b) Schweizer, M. P.; Banta, E. B.; Witkowski, J. T.; Robins, R. K. *Ibid.* **1973,** *95,* 3770–3778. (c) George,

A. L.; Hruska, F. E.; Ogilvie, K. K.; Holy, A. *Can. J. Chem.* **1978**, *56*, 1170–1176.
(32) Son, T.-D.; Chachaty, C. *Biochim. Biophys. Acta* **1973**, *335*, 1–13.
(33) (a) Akasaka, K.; Imoto, T.; Shibata, S.; Hatano, H. *J. Magn. Reson.* **1975**, *18*, 328–343. (b) Zens, A. P.; Bryson, T. A.; Dunlap, R. B.; Fisher, R. R.; Ellis, P. D. *J. Am. Chem. Soc.* **1976**, *98*, 7559–7564.
(34) Davies, D. B.; Danyluk, S. S. *Biochemistry* **1975**, *14*, 543–554.
(35) For a recent review, see De Clercq, E.; Torrence, P. F. *J. Carbohydr. Nucleosides Nucleotides* **1978**, *5*, 187–224.
(36) Torrence, P. F.; Huang, G.-F.; Edwards, M. W.; Bhooshan, B.; Descamps, J.; De Clercq, E. *J. Med. Chem.* **1979**, *22*, 316–319.
(37) (a) Hillen, W.; Gassen, H. G. *Biochim. Biophys. Acta* **1978**, *518*, 7–16. (b) Mertes, M. P.; Chang, C. T.-C.; De Clercq, E.; Huang, G.-F.; Torrence, P. F. *Biochem. Biophys. Res. Commun.* **1978**, *84*, 1054–1059.
(38) Prusoff, W. H.; Ward, D. C. *Biochem. Pharmacol.* **1976**, *25*, 1233–1239.
(39) Freese, E. *J. Mol. Biol.* **1959**, *1*, 87–105.
(40) Wataya, Y.; Santi, D. V.; Hansch, C. *J. Med. Chem.* **1977**, *20*, 1469–1473.
(41) (a) Huheey, J. E. *J. Phys. Chem.* **1965**, *69*, 3284–3291. (b) *Ibid.* **1966**, *70*, 2086–2092.
(42) Hansch, C.; Leo, A.; Unger, S. H.; Kim, K. H.; Nikaitani, D.; Lien, E. J. *J. Med. Chem.* **1973**, *16*, 1207–1216.
(43) Baur, W. H. *Acta Crystallogr., Sect. B* **1972**, 1456–1465.
(44) Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1973**, *95*, 2333–2344.
(45) Ekiel, I.; Remin, M.; Darzynkiewicz, E.; Shugar, D. *Biochim. Biophys. Acta* **1979**, *562*, 177–191.
(46) Ahmed, F. R.; Hall, S. R.; Pippy, M. E.; Huber, C. P. *J. Appl. Crystallogr.* **1973**, *6*, 309–346.
(47) Johnson, C. K. Oak Ridge, Tenn., 1970, ORTEP Report ORNL-3794 (2nd revision).

Mechanism-Based Enzyme Inactivation Using an Allyl Sulfoxide–Allyl Sulfenate Ester Rearrangement[1]

Michael Johnston,[†] Ronald Raines,[†] Christopher Walsh,*[†] and Raymond A. Firestone*[‡]

Contribution from the Departments of Chemistry and Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and the Membrane and Arthritis Research Department, Merck, Sharp and Dohme Research Laboratories, Rahway, New Jersey 07065.
Received December 3, 1979

Abstract: 2-Amino-4-chloro-5-(*p*-nitrophenylsulfinyl)pentanoic acid (**1**) has been synthesized and shown to induce mechanism-based inactivation of two pyridoxal phosphate dependent enzymes: (1) cystathionine γ-synthetase, which catalyzes a γ-replacement reaction in bacterial methionine biosynthesis; and (2) methionine γ-lyase, which catalyzes a γ-elimination reaction in bacterial methionine breakdown. The inactivations are irreversible and display saturation kinetics. Each enzyme incorporates roughly 1 mol of tritium per mol of enzyme monomer when inactivated by 2-amino-4-chloro-5-(*p*-nitro[^{3}H]phenylsulfinyl)pentanoic acid (**1a**), confirming that the modification of each protein is covalent and stoichiometric. Substoichiometric labeling (0.12 mol of tritium per mol of enzyme monomer) is given when methionine γ-lyase is fully inactivated by 2-amino-4-chloro-5-[^{3}H]-5-*p*-nitrophenylsulfinyl)pentanoic acid (**1b**). Both enzymes, inactivated by **1**, are susceptible to reactivation by thiols. Inactivated cystathionine γ-synthetase recovers 25% of its catalytic activity upon incubation with excess dithiothreitol, while methinonine γ-lyase is 100% reactivated by dithiothreitol, mercaptoethanol, and mercaptopropionate. Reactivation generates *p*-nitrophenylthiolate anion, which forms, in the case of methionine γ-lyase, stoichiometrically with enzyme reactivated. Both enzymes are "protected" from inactivation by **1** in the presence of thiols, which simultaneously generates *p*-nitrophenylthiol. In the presence of dithiothreitol, the protection reaction gives *p*-nitrophenylthiol production with pseudo-first-order kinetics. 2-Amino-4-chloro-5-(*p*-tolylsulfinyl)pentanoic acid (**2**) and 2-amino-4-(*p*-nitrophenylsulfinyl)-5-chloropentanoic acid (**3**), the reverse regioisomer of **1**, have also been prepared and give no evidence of inactivation of either enzyme. The data are taken to indicate a novel form of suicide inactivation (Scheme II) wherein β-carbanion-assisted γ-halide elimination generates an allyl sulfoxide–enzyme–pyridoxal adduct (**4**) which undergoes spontaneous 2,3-sigmatropic rearrangement to an electrophilic allyl sulfenate ester (**5**). The latter is then captured by an enzymic nucleophile to give an inactive enzyme **6**, which may be a mixed disulfide or, less likely, a sulfenamide.

Introduction

Considerable interest has been generated in the last several years in mechanism-based enzyme inactivators, also called suicide substrates.[2] Much of this interest results from the fact that the targeted enzyme uses some portion of its catalytic mechanism to "unmask", from an otherwise chemically *unreactive* group in the inactivator, a functionality *reactive* for alkylation of the enzyme. The reactive species is generated only in the enzyme's active site, and, thus, suicide substrates promise greater in vivo selectivity than do conventional affinity reagents.

A variety of functional groups have been used for mechanism-based inactivations, including acetylenes, olefins, nitriles, and β-halo substitutions,[2] which become activated usually by rearrangement or elimination to generate electrophiles susceptible to Michael-type addition by an active-site nucelophile. Certain functionalized penicillins, such as the clavulanates and penicillin sulfones,[3] cyclopropylamines,[4] fluoro- and nitrodeoxyuridylates,[5] and such drugs as allopurinol[6] are all known to function as specific suicide substrates, exemplifying the rich chemical diversity of this class of reactions.

It has occurred to us that a novel strategy for the generation of an electrophile in situ might use a sigmatropic rearrangement, wherein the first partner in the rearrangement would be unreactive to nucleophilic addition but the second, rearranged partner would serve to derivatize the enzyme. In this regard, we have chosen the 2,3-sigmatropic rearrangement of allyl sulfoxides[7a,b] (unreactive to nucleophiles) to allyl sulfenate esters (highly reactive to nucleophilic addition)[8a,b] as a likely mode of suicide-substrate inactivation (eq 1). A reagent of the type imagined will have greatest potential selectivity if designed such that the allyl sulfoxide is generated only within the active site of the targeted enzyme. Therefore, we have further con-

[†] Massachusetts Institute of Technology.
[‡] Merck, Sharp and Dohme Research Laboratories.

0002-7863/80/1502-4241$01.00/0 © 1980 American Chemical Society