Structure and Stereochemistry of Nucleic Acid Components and Their Reaction Products. II.[1a] Crystal and Molecular Structure of a Product Obtained in the Reaction of Isocyanates with Uracil [N^1-(N-Methylcarbamoyl)-N^3-methyl-5,6-dihydrouracil]

R. Parthasarathy,*[1b] J. Ohrt,[1b] S. P. Dutta,[1c] and Girish B. Chheda[1c]

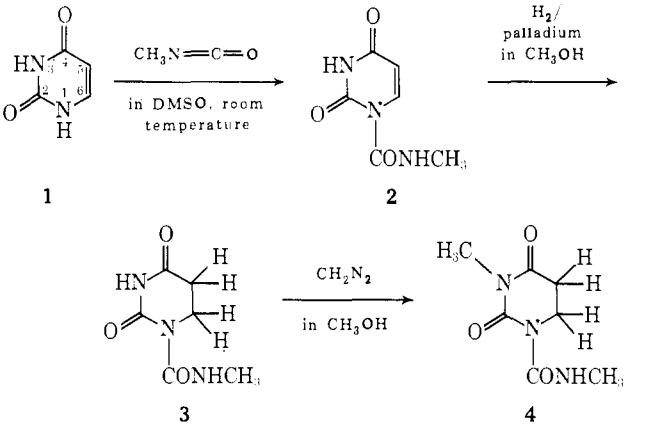
Contribution from the Center for Crystallographic Research and General Clinical Research Center, Department of General Surgery, Roswell Park Memorial Institute, Buffalo, New York 14203. *Received June 6, 1973*

Abstract: Uracil (**1**) reacts with methyl isocyanate to yield N^1-(N-methylcarbamoyl)uracil (**2**) which on hydrogenation and methylation yields a dihydrouracil derivative (**4**). The N-methylcarbamoyl group of **4** is resistant to acid hydrolysis. The structure of **4**, tentatively assigned on chemical grounds as N^1-(N-methylcarbamoyl)-N^3-methyl-5,6-dihydrouracil, was established by X-ray diffraction techniques, confirming this assignment. Crystals of **4** are orthorhombic, space group $Pna2_1$ with cell constants a = 9.160 (3), b = 19.358 (9), and c = 4.811 (3) Å and Z = 4. Using a GE XRD-6 diffractometer, 1066 reflections to the limit 2θ = 165° were measured for the Cu sphere. The complete structure was determined directly by the multisolution technique and refined by using the least-squares method to an R of 0.07. Significant structural features are: (i) C(4) sp^2–C(5) sp^3 bond of 1.538 (11) Å longer than the C(5) sp^3–C(6) sp^3 bond of 1.491 (11) Å, as found in dihydrouracil and dihydrothymine; (ii) the C(sp^2)–N(sp^2) bond of 1.433 (9) Å adjacent to the base in the ureido group is considerably longer than the other C(sp^2)–N(sp^2) bond of 1.323 (9) Å in the ureido group; (iii) the hydrogen H(8) is internally hydrogen bonded to O(2) and is 1.74 Å away from O(2); the angle N(8)–H(8)···O(2) is 136.1°; (iv) the nucleic acid base is puckered such that the carbons at the 5 and 6 positions are displaced toward opposite sides of the least-squares plane through the other four atoms of the base by ±0.22 and ∓0.48 Å respectively; (v) no stacking of the bases was observed in this crystal. The resistance to acid hydrolysis of the N^1 substituent is hypothesized to arise from the internal hydrogen bond from N(8) to the keto oxygen at the 2 position of dihydrouracil.

Although the function of the modified nucleosides is still not fully understood, enough evidence has accumulated to suggest their role in acceptor as well as transfer activity of tRNA.[2] The modification of isopentenyladenosine (6iPeAdo) by iodine[3] and the excision of base Y in tRNA[4] resulted in a poor fidelity of codon reading, with a net loss in the translational accuracy and efficiency of protein synthesis. Thus it is apparent that valuable information on the relationship of the structures of different regions of tRNA to their functions can be obtained by carrying out some selective alterations in the tRNA molecule and then reexamining its acceptor and transfer activity. With this purpose in mind, we have been exploring the reactions of a variety of isocyanates with minor and major nucleic acid bases and nucleosides. In the course of this study we obtained a product formed in the reaction of methyl isocyanate with uracil which required structural clarification. This paper describes the crystal and molecular structure and the stereochemistry of this substance, N^1-(N-methylcarbamoyl)-N^3-methyl-5,6-dihydrouracil.

When uracil **1** was allowed to react with methyl isocyanate at room temperature, a white crystalline product was obtained whose structure was tentatively assigned as **2** by analogy with the synthesis of 1-acetyluracil[5] (see Scheme I). In order to verify the structure of **2**, the compound was subjected to N-methylation with diazomethane. This reaction, however, did not give the desired N^3-methyl derivative of **2**, but gave 3-methyluracil. Therefore, the compound **2** was first hydrogenated to give **3** which was then methylated to give the 3-methyl-(N-methylcarbamoyl)-N^3-methyl-5,6-dihydrouracil (**4**). In order to show that the methylcarbamoyl side chain was attached to N^1 of the uracil moiety in **2**, compound **4** was treated with dilute acid to obtain 3-methyl-5,6-dihydrouracil. The latter reaction, however, failed and as a result the structure of **4** could not be confirmed. In order to establish the structure of **4**, X-ray diffraction studies were carried out and the results are reported here.

Scheme I



Experimental Section

General Methods. Melting points were recorded in a laboratory model MEL-TEMP melting point apparatus and are uncorrected.

(1) (a) For part I of this series: J. Ohrt, R. Parthasarathy, and G. B. Chheda, *J. Amer. Chem. Soc.*, **92**, 6604 (1970); (b) Center for Crystallographic Research; (c) General Clinical Research Center.
(2) D. Soll, *Science*, **173**, 293 (1971).
(3) F. Fittler and R. H. Hall, *Biochem. Biophys. Res. Commun.*, **25**, 441 (1966).
(4) R. Thiebe and H. G. Zachau, *Eur. J. Biochem.*, **5**, 546 (1968).
(5) L. B. Spector and E. B. Keller, *J. Chem. Soc.*, 185 (1958).

Ultraviolet measurements were made on a Cary-14 spectrophotometer. Nmr spectra were determined with a Varian A-60A instrument in DMSO-d_6 solution, peaks being measured in δ values downfield from an internal standard of tetramethylsilane. Mass spectra were recorded using a Du Pont 21-491 mass spectrometer at 70 eV. Elemental analyses were carried out by Heterocyclic Chemical Co., Harrisonville, Mo.

N^1-(N-Methylcarbamoyl)uracil (2). To a suspension of 1.12 g (10 mmol) of uracil in anhydrous DMSO (20 ml) was added 1.25 g of methyl isocyanate (22 mmol) and the mixture stirred at room temperature for 20 hr. The reaction mixture was then evaporated to dryness *in vacuo* at 60° and the crude residue was extracted with boiling chloroform (5 × 50 ml). (The insoluble part was found to be unchanged uracil.) The chloroform extract was concentrated to a small volume (25 ml) and cooled. The granular crystalline product was collected on a filter and recrystallized twice from ethyl acetate: mp 338–340°; yield 0.775 g (45.9%); λ max $CHCl_3$ 253 mμ (ε 8200); nmr δ 2.92 (d, 1, J = 8 Hz, 5 H) 9.03 (broad, 1, N*H*CH_3), 11.70 (broad, 1 ring N*H*); mass spectrum *m/e* 169 (M^+), 142 (M^+ − HCN), 126 (M^+ − HNCO), and 112 (M^+ − CH_3NCO).

Anal. Calcd for $C_6H_7N_3O_2$: C, 42.60; H, 4.14; N, 24.85. Found: C, 42.57; H, 4.23; N, 24.92.

N^1-(N-Methylcarbamoyl)-5,6-dihydrouracil (3). To a solution of 500 mg of **2** in 20 ml of methanol was added glacial acetic acid (10 ml) and 150 mg of 5% Pd/charcoal and the mixture was hydrogenated overnight in a Paar apparatus at 50 psi. The catalyst was then filtered and washed with methanol and then the combined filtrates were evaporated to dryness. The crude product did not have any uv absorption maxima above 220 mμ. The product was dissolved in a small volume of chloroform (5 ml), and petroleum ether (bp 60–80°) was added to turbidity; on cooling, the crystalline product was collected on a filter: mp 228–230°; yield, 450 mg (88.9%); nmr δ 2.60 (t, 2, J = 6 Hz, 5 H), 2.80 (d, 3, J = 5 Hz, NHCH_3), 3.93 (t, 2, J = 7 Hz, 6 H), 8.67 (broad, 1, N*H*CH_3), 10.75 (broad, 1, ring N*H*); mass spectrum *m/e* 171 (M^+), 141 (M^+ − $NHCH_3$), 114 (M^+ − CH_3NCO).

Anal. Calcd for $C_6H_9N_3O_3$: C, 42.10; H, 5.26; N, 24.56. Found: C, 41.90; H, 5.17; N, 24.52.

N^1-(N-Methylcarbamoyl)-N^3-methyl-5,6-dihydrouracil (4). To a solution of 500 mg of **3** in 25 ml of methanol was added an ethereal solution of diazomethane (50 ml) prepared from 5 g of N-nitroso-N-methylurea and the mixture was kept overnight at room temperature. The solution was then evaporated to dryness, and the residue was dissolved in chloroform (20 ml), treated with charcoal, filtered, concentrated to 5 ml, and petroleum ether (bp 60–80°) was added to it to turbidity. On cooling at 4° overnight long needles of the product crystallized out: mp 102–103°; yield, 380 mg (70%); mass spectrum *m/e* 185 (M^+), 155 (M^+ − $NHCH_3$), 128 (M^+ − CH_3NCO).

Anal. Calcd for $C_7H_{11}N_3O_3$: C, 45.40; H, 5.94; N, 22.70. Found: C, 45.47; H, 5.96; N, 22.79.

Attempted Methylation of N^1-(N-Methylcarbamoyl)uracil (2). To a solution of 0.420 g of **2** in 25 ml of tetrahydrofuran was added an ethereal solution of diazomethane (50 ml) prepared from 5.0 g of N-nitroso-N-methylurea and kept at room temperature for 16 hr. The clear solution was evaporated to dryness *in vacuo* at 40°. (Tlc of the crude residue showed only one spot matching with 3-methyluracil.) The residue was dissolved in a small volume of methanol, filtered, concentrated, and cooled at 0°. The crystalline product was filtered, washed with ether, and dried *in vacuo:* mp 231–232°; yield, 0.254 g (80.6%); λ max 258 mμ (pH 6.2), 258 (pH 1.5), 283 (pH 11.6). The melting point of the product was not depressed on admixture with authentic 3-methyluracil.

Attempted Hydrolysis of N-Methylcarbamoyl Group of 4. **4** (100 mg) was dissolved in 5.0 ml of 2 N HCl and heated on a steam bath for 2 hr. The solution was then evaporated to dryness *in vacuo*. The residue was dissolved in a small volume (2 ml)of water and cooled at 4°. Long needles crystallized out and were filtered, washed with cold water, and dried *in vacuo:* mp 102–103°; was found to be unchanged **4**; yield, 82 mg (82%). The melting point was not depressed on admixture with authentic **4**. (The melting point of 3-methyldihydrouracil is 127–129°.)

Crystallographic. After repeated attempts to produce crystals suitable for X-ray diffraction, the title compound (hereafter referred to as compound I for simplicity) was obtained as fine needles from water–propanol. These crystals are orthorhombic, and the systematically absent reflections are: $h0l$ with h odd, $0kl$ with $k + l$ odd; no absences in hkl. These absences are consistent with the space groups *Pnam* and $Pna2_1$. *Pnam* has eight equivalent positions and $Pna2_1$ has four. Since the unit cell volume and the density show that there are only four molecules in the cell, the selection of *Pnam* will demand placing the molecules on special positions, either on the mirrors or on the centers of inversion. Since the intensity statistics did not clearly indicate whether the structure is centric or not and since the molecule did not possess either a center of inversion or a mirror, the acentric space group $Pna2_1$ was tried at first. The structure was readily obtained as described below and refined well. Consequently, the space group for this crystal is $Pna2_1$. The unit cell constants and other crystallographic data of compound I ($C_7H_{11}N_3O_3$) are: a = 9.160 (3) Å, b = 19.358 (9) Å, c = 4.811 (3) Å, V = 853.13 Å^3, ρ_{obsd} (by flotation) = 1.40 g cm^{-3}, ρ_{calcd} = 1.44 g cm^{-3}, Z = 4, molecular weight = 185.18 daltons, μ = 9.80 cm^{-1}, Cu Kα_1 = 1.54051 Å. The unit cell constants were refined from diffractometer data (at 22 ∓ 3°) by a least-squares procedure.

Complete intensity data were obtained to the limit 2θ = 165°, employing Cu Kα radiation. The stationary crystal–stationary counter technique[6] was employed for obtaining the intensities, using a 5° take-off angle; 1066 reflections were measured, of which 302 whose intensities were less than twice the background in that (sin θ/λ) range were considered "unobservable." The crystal used for the data collection had the dimensions 0.2 × 0.1 × 0.4 mm and was mounted with the c^* axis along the ϕ axis of the goniostat. The difference in absorption as a function of ϕ was measured for the axial reflections and was used for correcting approximately for the anisotropy of absorption. This correction was about 15% for most reflections, and was up to 50% for reflections making a small angle with b^*. The data were processed in the usual way.

Phase Determination. The basis of the phase-determining procedure was the multisolution technique as developed by Germain, Main, and Woolfson.[7] Three reflections, 3,16,0, 0 7 1, and 4 5 0 (their $|E|$'s being 3.48, 2.15, and 2.14, respectively), were specified as having a phase of zero in order to define the origin. Three other reflections, namely 1,11,0, 6 6 3, and 5 6 3, with $|E|$'s of 3.57, 2.78, and 2.63, respectively, were chosen as the starting set for the generation of multiple solutions; 32 solutions were obtained. The set with the highest figure of merit[7] of 0.968 yielded the structure when the corresponding E map was calculated. The 13 nonhydrogen atoms were readily located and the R value ($\Sigma||F_o| - |F_c||/\Sigma|F_o|$) for this structure was 0.31. The phases of 0 7 1, 1,11,0, 6 6 3, and 5 6 3 in the final cycle of refinement changed from their starting values of 0, 180, 225, and 45° to 328, 0, 227, and −11°, respectively.

Refinement of the Structure. The atomic coordinates and thermal parameters were refined by several cycles of least squares, employing a block-diagonal approximation. Blocks of 9 × 9 and 4 × 4 were employed for atoms with anisotropic and isotropic thermal parameters, respectively. The hydrogen atoms were located from electron density difference maps after R had reached 0.10. The positional and individual isotropic thermal parameters of the hydrogen atoms were then also allowed to vary in the refinement and the R value was reduced to 0.07. None of the shifts in the final cycle were greater than one-tenth the standard deviations for the nonhydrogen atoms and one-fourth the standard deviations for the hydrogen atoms. The refinement was considered to be complete and the final atomic and thermal parameters and their esd's as obtained from the inverse of the block-diagonal matrix are listed in Tables I and II. A list of observed and calculated structure factors will appear in the microfilm edition (see Supplementary Material Available at end of paper).

The observations were weighted according to the scheme of Evans,[8] and the refinement was carried out by minimizing $[w(|F_o| - (1/k)|F_c|)^2]$. Reflections considered "unobserved" were given zero weight during the refinement and for the R-index calculations. Atomic scattering factors for C, N, and O atoms were those listed in ref 9a. For the hydrogen atoms, the values given by Stewart, Davidson, and Simpson[9b] were used.

Discussion of the Structure

The bond distances and angles in the molecule are illustrated in Figures 1 and 2, respectively. The esd's of the bonds range from 0.008 to 0.011 Å and angles

(6) T. F. Furnas and D. Harker, *Rev. Sci. Instrum.*, **26**, 449 (1955).
(7) G. Germain, P. M. Main, and M. M. Woolfson, *Acta Crystallogr., Sect. A*, **27**, 368 (1971).
(8) H. T. Evans, *Acta Crystallogr.*, **14**, 489 (1961).
(9) (a) "International Tables for X-ray Crystallography," Vol. III, Kynoch Press, Birmingham, England, 1962, pp 202–203; (b) R. F. Stewart, E. R. Davidson, and W. T. Simpson, *J. Chem. Phys.*, **42**, 3175 (1965).

Table I. Coordinates and Thermal Parameters[a]

	x	y	z	b_{11}	b_{22}	b_{33}	b_{12}	b_{13}	b_{23}
O(2)	4011 (5)	351 (2)	10613 (12)	158 (7)	17 (1)	353 (22)	14 (6)	111 (27)	34 (11)
O(4)	−242 (6)	847 (3)	14840 (14)	166 (8)	37 (2)	518 (34)	−2 (8)	178 (34)	0 (15)
O(7)	4009 (6)	2247 (2)	6900 (12)	241 (10)	12 (1)	530 (32)	12 (6)	202 (36)	55 (13)
N(1)	3025 (6)	1418 (3)	9747 (13)	123 (8)	14 (2)	354 (26)	10 (6)	54 (31)	17 (13)
N(3)	1983 (6)	673 (3)	12992 (13)	117 (8)	19 (2)	359 (27)	−10 (6)	80 (32)	−4 (13)
N(8)	5048 (6)	1198 (3)	6769 (14)	118 (8)	17 (2)	402 (31)	5 (6)	59 (32)	12 (14)
C(2)	3056 (6)	782 (4)	11055 (15)	83 (7)	25 (2)	273 (27)	13 (7)	7 (29)	−14 (15)
C(3)	2151 (8)	30 (4)	14735 (18)	190 (13)	26 (2)	305 (29)	−26 (10)	74 (44)	22 (20)
C(4)	704 (7)	1010 (4)	13252 (18)	100 (10)	32 (3)	464 (42)	−10 (8)	128 (37)	−50 (20)
C(5)	547 (8)	1637 (4)	11316 (18)	171 (12)	33 (3)	393 (40)	48 (10)	127 (43)	37 (21)
C(6)	2001 (8)	1956 (4)	10778 (18)	158 (11)	24 (2)	406 (39)	44 (8)	169 (43)	20 (18)
C(7)	4081 (7)	1643 (4)	7744 (15)	123 (10)	23 (2)	256 (27)	3 (8)	24 (32)	1 (15)
C(9)	6184 (8)	1404 (4)	4838 (17)	166 (11)	19 (2)	396 (36)	−25 (8)	164 (42)	8 (18)

[a] $T_F = \exp[-(b_{11}h^2 + b_{22}k^2 + b_{33}l^2 + b_{12}hk + b_{13}hl + b_{23}kl)]$. The entries in the table are values × 10^4 for both coordinates and thermal parameters. Standard deviations given in parentheses refer to the last digit.

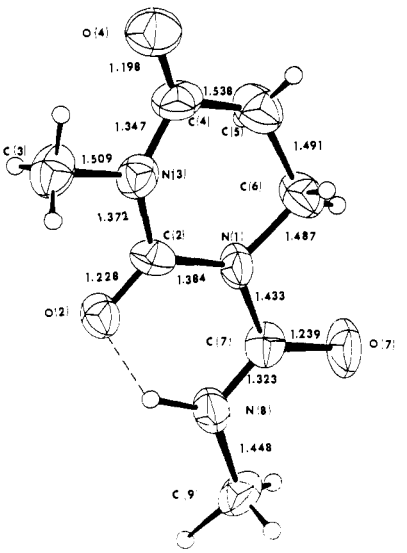


Figure 1. Bond distances in Å.

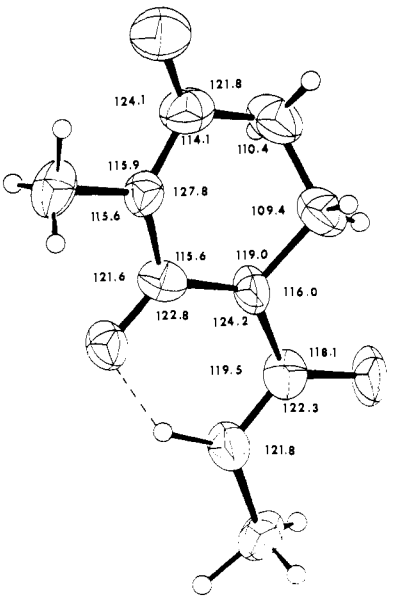


Figure 2. Bond angles in degrees.

Table II. Coordinates (× 10^3) and Thermal Parameters (× 10) for Hydrogen Atoms

	x	y	z	B
H(3a)	146 (6)	10 (4)	1640 (17)	66 (20)
H(3b)	322 (4)	2 (2)	1550 (11)	4 (9)
H(3c)	169 (7)	−29 (4)	1308 (23)	108 (26)
H(5a)	18 (7)	191 (3)	1239 (16)	57 (18)
H(5b)	20 (8)	148 (3)	952 (18)	95 (24)
H(6a)	193 (7)	225 (4)	951 (17)	60 (20)
H(6b)	229 (7)	218 (4)	1275 (18)	82 (24)
H(8)	502 (6)	70 (3)	790 (15)	49 (16)
H(9a)	566 (8)	166 (4)	300 (23)	126 (29)
H(9b)	695 (6)	162 (3)	531 (16)	44 (17)
H(9c)	657 (8)	96 (4)	374 (24)	126 (30)

from 0.6 to 0.7°. When one of the atoms involved in the bonds is a hydrogen, the esd's range from 0.04 to 0.09 Å and 3 to 5°, respectively.

(a) Saturated Base. The bond distances and angles found in the dihydrouracil moiety of the present structure are only in reasonable agreement with those found in dihydrouracil[10] and dihydrouridine hemihydrate.[11-13] The N(1)–C(2) bond distance of 1.384 (9) Å in this structure is considerably longer (by 3–4 std dev) than the corresponding values found in other dihydrouracils (see Table III). It was observed in the present study that the C(4) sp^2–C(5) sp^3 bond of 1.538 (11) Å is actually longer than the C(5) sp^3–C(6) sp^3 bond of 1.491 (11) Å. While this result is difficult to understand, we wish to point out that a similar situation occurs in the structures of dihydrouracil[10] and dihydrothymine[14] and was noticed by the respective authors. In both dihydrouracil and the present structure, the C(5)–C(6) bond distance is significantly shorter than the normal C sp^3–C sp^3 single bond value of 1.533 (3) Å[15] and the C(4) sp^2–C(5) sp^3 bond distance is as long as the usually accepted value for a saturated C–C bond. However, this tendency is not observed in dihydrouridines (see Table III).

It would appear to be of interest to study the effect of hydrogenation at C(5) and C(6) (see Tables III–VI).

(10) D. C. Rohrer and M. Sundaralingam, *Acta Crystallogr., Sect. B*, **26**, 546 (1970).
(11) M. Sundaralingam, S. T. Rao, and J. Abola, *Science*, **172**, 725 (1971).
(12) M. Sundaralingam, S. T. Rao, and J. Abola, *J. Amer. Chem. Soc.*, **93**, 7055 (1971).
(13) D. Suck, W. Saenger, and K. Zechmeister, *Acta Crystallogr., Sect. B*, **28**, 596 (1971).
(14) S. Furberg and L. H. Jensen, *J. Amer. Chem. Soc.*, **90**, 470 (1968).
(15) L. S. Bartell, *J. Amer. Chem. Soc.*, **81**, 3497 (1959). Frequently, it has been observed in amino acid structures that the C–C bonds are considerably less than the value quoted by Bartell. See, for example, D. N. Wright and R. E. Marsh, *Acta Crystallogr.*, **15**, 54 (1962); R. Parthasarathy, *ibid.*, **21**, 422 (1966); L. Golic and W. C. Hamilton, *Acta Crystallogr., Sect. B*, **28**, 1265 (1972).

Table III. A Comparison of Bond Lengths and Standard Deviations in Nucleic Acid Bases Related to Uracil and Dihydrouracil

	Uracil[c]	Thymine[b] monohydrate	Dihydrouracil[10]	Dihydrouridine[13] hemihydrate Mol A	Dihydrouridine[13] hemihydrate Mol B[a]	Dihydrouridine[11,12] hemihydrate Mol A	Dihydrouridine[11,12] hemihydrate Mol B[a]	Dihydrothymine[14] Mol 1[a]	Dihydrothymine[14] Mol 2[a]	I
N(1)–C(2)	1.371 (2)	1.355 (5)	1.335 (5)	1.356 (5)	1.351 (5)	1.351 (6)	1.345 (7)	1.326 (2)		1.384 (9)
C(2)–N(3)	1.376 (2)	1.361 (4)	1.395 (5)	1.389 (4)	1.393 (5)	1.394 (7)	1.393 (7)	1.383 (2)		1.372 (9)
N(3)–C(4)	1.371 (3)	1.391 (3)	1.364 (5)	1.364 (5)	1.371 (4)	1.363 (6)	1.366 (7)	1.358 (3)		1.347 (9)
C(4)–C(5)	1.430 (2)	1.447 (4)	1.515 (6)	1.492 (5)	1.495 (6)	1.484 (7)	1.487 (7)	1.531 (6)	1.555 (8)	1.538 (11)
C(5)–C(6)	1.340 (4)	1.349 (5)	1.507 (6)	1.512 (3)	1.510 (6)	1.515 (7)	1.521 (8)	1.516 (38)	1.521 (33)	1.491 (11)
C(6)–N(1)	1.358 (2)	1.382 (4)	1.464 (5)	1.450 (4)	1.467 (5)	1.449 (6)	1.463 (8)	1.450 (20)	1.372 (56)	1.487 (9)
C(2)–O(2)	1.215 (2)	1.234 (4)	1.222 (4)	1.223 (4)	1.210 (4)	1.219 (7)	1.217 (7)	1.235 (2)		1.228 (8)
C(4)–O(4)	1.245 (3)	1.231 (4)	1.211 (5)	1.229 (5)	1.225 (5)	1.225 (6)	1.223 (7)	1.212 (2)		1.198 (9)

[a] Disordered. [b] R. Gerdil, *Acta Crystallogr.*, **14**, 333 (1961). [c] R. F. Stewart and L. H. Jensen, *ibid.*, **23**, 1102 (1967).

Table IV. A Comparison of Bond Angles and Standard Deviations in Nucleic Acid Bases Related to Uracil and Dihydrouracil

	Uracil[a]	Thymine[b] monohydrate	Dihydrouracil[10]	Dihydrouridine[13] hemihydrate Mol A	Dihydrouridine[13] hemihydrate Mol B[c]	Dihydrouridine[11,12] hemihydrate Mol A	Dihydrouridine[11,12] hemihydrate Mol B[c]	Dihydrothymine[14] Mol 1[c]	Dihydrothymine[14] Mol 2[c]	I
C(6)–N(1)–C(2)	122.7 (1)	122.8 (2)	122.1 (2)	120.3 (2)	118.6 (3)	121.4 (4)	118.9 (4)	122.1 (5)	120.8 (8)	119.0 (6)
N(1)–C(2)–N(3)	114.0 (1)	115.2 (3)	116.1 (2)	116.0 (2)	115.6 (3)	115.2 (5)	115.3 (5)	116.6 (2)		115.6 (6)
C(2)–N(3)–C(4)	126.3 (1)	126.3 (1)	126.7 (2)	126.5 (3)	125.4 (3)	126.6 (4)	125.5 (5)	126.3 (2)		127.8 (6)
N(3)–C(4)–C(5)	115.5 (1)	115.6 (2)	115.1 (3)	115.9 (3)	115.9 (4)	115.3 (4)	116.0 (4)	113.4 (3)	113.9 (4)	114.1 (6)
C(4)–C(5)–C(6)	118.9 (1)	118.2 (3)	112.6 (3)	111.0 (3)	111.0 (3)	111.4 (4)	110.5 (5)	108.1 (15)	104.5 (15)	110.4 (6)
C(5)–C(6)–N(1)	122.3 (1)	121.8 (2)	110.3 (3)	110.8 (3)	109.4 (3)	109.1 (4)	108.7 (5)	108.5	109.5 (14)	109.4 (6)
N(1)–C(2)–O(2)	123.7 (1)	122.7 (2)	124.4 (2)	124.7 (3)	125.4 (4)	125.8 (5)	126.1 (5)	123.9 (2)		122.8 (6)
O(2)–C(2)–N(3)	123.3 (1)	122.1 (1)	119.5 (2)	119.2 (3)	118.9 (3)	119.0 (4)	118.6 (5)	119.5 (2)		121.6 (6)
N(3)–C(4)–O(4)	119.2 (1)	118.3 (2)	120.9 (2)	119.9 (3)	120.4 (3)	120.9 (4)	120.6 (5)	121.1 (2)		124.1 (7)
O(4)–C(4)–C(5)	125.3 (1)	126.1 (4)	123.9 (2)	124.2 (4)	123.7 (3)	123.8 (4)	123.4 (5)	124.1 (3)	121.1 (4)	121.8 (6)

[a] Footnote *c*, Table III. [b] Footnote *b*, Table III. [c] Disordered.

Table V. Deviation of Atoms from Least-Squares Planes in I[a]

	Å × 10²		
N(1)	**−3**	−29	**4**
C(2)	**7**	**2**	**−4**
N(3)	**−8**	**−3**	**2**
C(3)	−18	8	−28
C(4)	**4**	**3**	49
C(5)	22	**−1**	86
C(6)	−48	−81	**−2**
O(2)	24	26	−20
O(4)	1	10	62
r.m.s. deviation	6	2	4

[a] Boldface numbers represent the corresponding atoms included in defining the least-squares plane.

As is well known, the primary effect of this saturation of the C(5)–C(6) bond is to pucker the base to an approximate half-chair conformation. The C(5) and C(6) atoms are displaced from the least-squares plane through the other four atoms of the base by ±0.22 and ∓0.48 Å, respectively, *i.e.*, toward opposite sides of this plane (see Tables V and VI). This puckering is very similar to that observed in dihydrouracil[10] (Table VI), dihydrouridine,[11–13] and dihydrothymine.[14]

Hydrogenation not only radically changes the stereochemistry at C(5)–C(6), but also redistributes the bonding electrons in the planar part of the molecule, as shown by bond distances and angles. The changes in bond lengths due to hydrogenation on both uracil and thymine are remarkably similar. The most surprising of these changes due to electronic rearrangements are the shortening of the N(1)–C(2) bonds and the approximate equality in length of the C(sp^2)–C(sp^3) and C(sp^3)–C(sp^3) bonds as discussed earlier.

An attempt was made to discover any conformational disorder that might be present in the crystal structure of compound I, similar to the type of disorder found in dihydrothymine[4] (see the discussion on analysis of thermal vibration). It was found that there is no positive evidence for such conformational disorders.

The four-atom plane in the ring through atoms C(2), N(3), C(4), and C(5) seems to be the "best" plane with the least r.m.s. deviations. The second best four-atom plane is the one through C(6), N(1), C(2), and N(3) (see Table V). The torsion angles in the base that exhibit the displacement of the substituent atoms from the appropriate conjugated system are given in Table VII. It is seen from this table that C(7) is least displaced

Table VI. A Comparison of Torsion Angles (deg) in the Ring Portion in Some Nucleic Acid Bases Related to Uracil and Dihydrouracil

	Uracil[a]	Dihydrouracil[10]	Dihydrouridine[11,12] hemihydrate Mol A	Dihydrouridine[11,12] hemihydrate Mol B[b]	Dihydrothymine[14] Mol 1[b]	Dihydrothymine[14] Mol 2[b]	I
C(6)–N(1)–C(2)–N(3)	±0.0	±13.0	−15.6	17.4	−14.0	15.3	±8.4
N(1)–C(2)–N(3)–C(4)	∓0.5	±11.1	−10.6	18.0	−1.8		±19.7
C(2)–N(3)–C(4)–C(5)	±0.8	∓3.0	1.5	−13.8	−14.9	20.2	∓7.2
N(3)–C(4)–C(5)–C(6)	∓0.5	∓26.2	30.4	−22.5	43.2	−46.9	∓30.8
C(4)–C(5)–C(6)–N(1)	±0.0	±45.4	−51.1	51.7	55.4	57.8	±53.5
C(5)–C(6)–N(1)–C(2)	±0.2	∓41.0	45.9	−51.7	43.8	−46.2	∓44.3

[a] Footnote *c*, Table III. [b] Disordered.

Table VII. Torsion Angles (deg) of the Substituents on the Base in I

C(7)–C(1)–C(2)–O(2)	8.4
C(7)–C(1)–C(2)–N(3)	0.4
C(6)–C(1)–C(2)–O(2)	−168.7
C(6)–C(1)–C(2)–N(3)	177.5
C(1)–C(2)–N(3)–C(3)	−170.7
O(2)–C(2)–N(3)–C(3)	6.4
O(2)–C(2)–N(3)–C(4)	−163.2
C(2)–N(3)–C(4)–O(4)	173.0
C(3)–N(3)–C(4)–O(4)	3.5
C(3)–N(3)–C(4)–C(5)	176.7
O(4)–C(4)–C(5)–C(6)	149.0
C(5)–C(6)–N(1)–C(7)	145.7
C(2)–N(1)–C(7)–O(7)	−173.8
C(2)–N(1)–C(7)–N(8)	9.3
C(6)–N(7)–C(7)–O(7)	−4.4
C(6)–N(7)–C(7)–N(8)	178.7
C(1)–C(7)–N(8)–C(9)	−177.2
O(7)–C(7)–N(8)–C(9)	6.0

from the plane of the base, a result analogous to the location of the C(1′) in dihydrouridine.

(b) Carboxymethylamino Group. The C(7) sp^2–N(1) sp^2 and C(7) sp^2–N(8) sp^2 bonds differ in bond lengths by as much as 0.11 Å which is ten times the standard deviation in bond lengths. These bonds are respectively the longest and shortest of the C(sp^2)–N(sp^2) bonds in this structure. The bond angles around C(7) are very close to the ideal value of 120°.

Another interesting feature of this group in this structure is the strong internal hydrogen bond of N(8) to O(2). The N–C bonds in the carbamoyl linkage may be expected to have double-bond character which will restrict the carbamoyl linkage to two conformations: O(7) or N(8) cis to O(2). In compound I, however, the N(1)–C(7) bond contains little double-bond character and may not be restrictive to the carbamoyl conformation. The conformation where O(2) and O(7) are trans to each other brings N(8) to a position favorable for forming the internal hydrogen bond referred to earlier. The H(8)···O(2) distance is 1.74 Å, considerably shorter than the sum of their van der Waal's radii, and the N(8)–H(8)···O(2) angle is 136.1°. Such internal hydrogen bonds leading to the formation of a six-membered ring have been observed in many structures.[16,17]

(c) Analysis of Thermal Vibration. An attempt was made to determine whether the shortening of the C(5)–C(6) bond compared to a normal C(sp^3)–C(sp^3) bond may be due to a disorder in the structure. It was observed in the structure determination of dihydrothymine[14] that an unreasonably short C(5)–C(6) bond was obtained initially, along with large thermal vibrations for the atoms C(5) and C(6). These results indicated to the authors that refinement was not proceeding correctly and suggested the occurrence of a disordered structure. This disorder consisted of the two modes of puckering of C(5) and C(6) (see Table V and VI) mentioned earlier, and refinement including the disordered structure gave significant improvement of the structural parameters, especially the C(5)–C(6) bonds. In view of the above situation, the thermal parameters for the present structure were examined, but they failed to show any special anomaly in the thermal parameters of C(5) and C(6) atoms. The electron-density difference maps did not indicate any residual electron density greater than 0.15 eA^{-3}, implying no disorder in compound I to the limits of experimental error. Though there was no positive evidence for a disordered structure in the present study, the fact remains that the shortening of the C(5)–C(6) bond persists if no disordered model is used for the refinement (as in the present study and in dihydrouracil[10]), but this anomaly disappears if a disordered model is explicitly used in the refinement (as in dihydrouridine).[12,13]

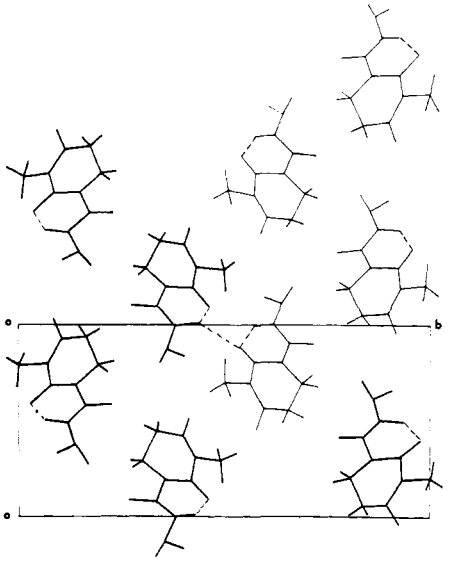


Figure 3. Packing viewed along *c*. There is no stacking of the bases.

(d) Hydrogen Bonding and Packing. There is only one hydrogen (H(8)) of N(8) which can take part in hydrogen bond formation. This hydrogen is involved in a strong internal hydrogen bond to O(2), as discussed earlier. In addition to this interaction H(8) is involved in a weak interaction with O(2′) such that H(8)···O(2′) is 2.48 Å, N(8)–H(8)···O(2′) is 119.3°. The prime on O(2) denotes that this atom is related to the unprimed one by $1 - x, \bar{y}, -1/2 + z$. Though the bifurcated interaction is very weak, such interactions occur quite frequently (in biological molecules) and the specific configuration in which they occur is such that the hydrogen, its donor, and the two acceptors all lie in a plane.[18] In the present case, also, it was found that the hydrogen is nearly in the plane of N(8), O(2), and O(2′), the sum of the angles around H(8) being 357.0°. As for O(2) it has two hydrogen contacts, namely H(8) and H(8″) (the double prime denotes that H(8″) is related to H(8) by $1 - x, \bar{y}, 1/2 + z$) both nearly in the plane through C(2), O(2), N(1), and N(3). The angles C(2)–O(2)···H(8), C(2)–O(2)···H(8″), and H(8)···O(2)···H(8″) are respectively 104.1, 137.4, and 117.7°. From these values, it is interesting to note that O(2) is so oriented that its electrons in the sp^2 orbitals point toward the hydrogens H(8) and H(8′), respectively (see Figure 3).

An important question is whether the saturated bases are stacked on top of each other, thus giving rise to a significant contribution to the stability of the

(16) W. C. Hamilton and J. A. Ibers, "Hydrogen Bonding in Solids," W. A. Benjamin, New York, N. Y., 1968, pp 178–181.
(17) M. Tichy, *Advan. Org. Chem.*, **5**, 115 (1965).
(18) R. Parthasarathy, *Acta Crystallogr., Sect. B*, **25**, 509 (1969).

(crystal) structure. In this case, it is clear from Figure 3 that there is no base stacking,[19] similar to the situation observed in dihydrouridine[13] (see ref 19). The structure seems to be stabilized by a combination of hydrogen interactions, van der Waal's, and dipolar forces.

Acid Hydrolysis of the N^1 Substituent. The resistance to acid hydrolysis of the N^1 substituents is surprising since it is known that the hydrogenation of carbons 5 and 6 of uracil facilitates the removal of the sugar from uridine.[20] We were looking for any structural features that might explain this resistance of compound I to acid hydrolysis. Two features in this structure are interesting to note in this connection, namely the long N(1)–C(7) bond and the internal hydrogen bond from N(8) to O(2). The effect of both inter- and intramolecular hydrogen bonding is known not only in altering reaction kinetics, but also in influencing reaction paths.[21,22] Hence, this internal hydrogen bond might be involved in making the N^1 substituent resistant to hydrolysis.

(19) The term "base stacking" seems to be used by different people with different meanings. For example, the structure analysis of dihydrouridine was carried out by two independent groups,[12,13] one of whom[13] consider there is no base stacking in dihydrouridine, but the other[12] suggests that "the base stacking configuration observed here is in many respects similar to that observed in the known planar pyrimidine systems with the carbonyl oxygen atoms O(2) of the two molecules lying either over or close to the rings of adjacent bases." This ambiguity can be removed somewhat by giving a more precise definition for the base stacking in terms of the degree of overlap of the area of the conjugated system on adjacent bases when projected normal to them. The distances between the various atoms on adjacent bases are important, but there are too many such distances to be specified. Hence a stacking distance which is the *shortest* distance between the overlapping π systems might be an appropriate quantity to quote in discussing base stacking. It is realized that the most appropriate quantity for discussive base stacking is the corresponding decrease in energy on forming aggregates, but this is not readily calculable; the percentage degree of overlap and the stacking distance might serve a useful purpose as a crude index of this energy.

(20) L. Haavaldsen, S. G. Laland, and J. M. McKee, *Biochim. Biophys. Acta*, 33, 201 (1959).

Acknowledgments. R. P. and J. O. thank Dr. D. Harker for valuable and stimulating discussions. S. P. D. and G. B. C. thank Drs. A. Mittelman, E. D. Holyoke, and G. P. Murphy for their keen interest and encouragement in this work. This work was partially supported from Grant CA 14185 from the National Cancer Institute of the U. S. Public Health Service. Our thanks go to Mrs. Winiewicz for excellent technical assistance.

Supplementary Material Available. A listing of observed and calculated structure factors will appear following these pages in the microfilm edition of this volume of the journal. Photocopies of the supplementary material from this paper only or microfiche (105 × 148 mm, 20× reduction, negatives) containing all of the supplementary material for the papers in this issue may be obtained from the Journals Department, American Chemical Society, 1155 16th St., N.W., Washington, D. C. 20036. Remit check or money order for $3.00 for photocopy or $2.00 for microfiche, referring to code number JACS-73-8141.

(21) G. C. Pimental and A. L. McClellan, "The Hydrogen Bond," W. H. Freeman, San Francisco, Calif., 1959, pp 184–188.

(22) S. N. Vinogradov and R. H. Linnel, "Hydrogen Bonding," Van Nostrand-Reinhold, New York, N. Y., 1971, p 267.

Conformational Analysis of Cytidine, 1-β-D-(Arabinofuranosyl)cytosine and Their *O'*-Methyl Derivatives by Proton Magnetic Resonance Spectroscopy

M. Remin and D. Shugar*

Contribution from the Department of Biophysics, Institute of Experimental Physics, University of Warsaw, 02-089 Warszawa, and Institute of Biochemistry and Biophysics, Academy of Sciences, 02-532 Warszawa, Poland. *Received June 7, 1973*

Abstract: The pmr spectra of cytidine, 1-β-D-(arabinofuranosyl)cytosine, and a series of their *O'*-methyl derivatives have been subjected to detailed computer analyses. Conformational analyses profited from the "neighbor anisotropy effect" of the *O'*-CH_3 bonds, since CNDO/2 calculated changes in electron density due to O'-methylation were negligible. It was shown for the vicinal protons that, with a given pentose ring puckering, there is a marked disparity between the values of the dihedral angles calculated from standard stereochemical models and those derived from crystallographic data, and this was taken into account in the present analysis. The correlation between the changes in chemical shifts, due to etherification of one of the pentose hydroxyls, with the values of coupling constants, pointed to the existence of preferred puckered forms, *viz.* C_3'-*endo*–C_2'-endo for the ribose ring, and C_2'-*exo*–C_3'-exo for the arabinose ring. In cytidine derivatives the conformation of the exocyclic 5'-CH_2OH group exhibits a preference for the form gauche–gauche (60–70%), and approximately equal populations for the other two forms. For arabinosylcytosine the gauche–gauche population is in the range 20–35% and the gauche–trans, 40–50%. The influence of a 5'-OCH_3 on the chemical shift of the cytosine H_6 demonstrated the marked preference for the form anti in both nucleosides. It also proved possible to predict preferred conformers for the various *O'*-CH_3 groups. The breadths of the H_5 and H_6 signals were essentially unaffected over the pH range for cytosine ring protonation.

The pioneering studies of Jardetzky[1] and Lemieux[2] on the conformations of nucleosides and nucleotides in solution by means of nmr spectroscopy have since undergone extensive developments, including the use of simulation techniques for analyses of spectra,[3] interpretation of coupling constants between ^1H, ^{13}C,

(1) C. D. Jardetzky, *J. Amer. Chem. Soc.*, **84**, 62 (1962), and references therein.

(2) R. U. Lemieux, *Can. J. Chem.*, **39**, 116 (1961).

(3) F. E. Hruska, A. A. Grey, and I. C. P. Smith, *J. Amer. Chem. Soc.*, **92**, 4088 (1970); T. Schleich, B. J. Blackburn, R. D. Lapper, and I. C. P. Smith, *Biochemistry*, **11**, 137 (1972).