Stereochemistry of Nucleic Acids and Their Constituents. XXIII. Crystal and Molecular Structure of Dihydrouridine "Hemihydrate," a Rare Nucleoside with a Saturated Base Occurring in the Dihydrouridine Loop of Transfer Ribonucleic Acids[1a]

M. Sundaralingam,* S. T. Rao, and J. Abola

Contribution from the Crystallography Laboratory, Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison, Wisconsin 53706. *Received March 15, 1971*

Abstract: The crystal structure of dihydrouridine, a unique rare nucleoside occurring in the dihydrouridine loop of transfer RNAs, has been determined by X-ray diffraction. Crystals of dihydrouridine are orthorhombic, space group $P2_12_12_1$, $a = 8.131$, $b = 11.766$, and $c = 23.016$ Å. There are eight molecules of the nucleoside and four molecules of water per unit cell. The structural solution was obtained by a combination of Patterson search methods and application of the tangent formula. Full-matrix, least-squares refinement of the 35 nonhydrogen atoms with anisotropic temperature factors and the hydrogen atoms with isotropic temperature factors gave a final R value of 0.056 for 1917 reflections measured on a diffractometer. The two independent molecules in the asymmetric unit exhibit important conformational similarities and differences. The conformation about the glycosidic bond is anti for both molecules; in molecule A $\chi_{CN} = 65.5°$ and in molecule B $\chi_{CN} = 57.1°$. The base rings exhibit twist half-chair conformations, but the two rings are puckered in opposite directions. The furanoside rings show a rather novel conformation $^2T_1$ [(C(2′)-endo–C(1′)-exo]. The conformation about the C(4′)–C(5′) bond in molecule A is gauche–trans (gt), while molecule B shows disorder with two alternate staggered arrangements trans–gauche (tg) (88%) and gauche–gauche (gg) (12%). The preponderant conformations (gt and tg) appear to be advantageous for nucleosides in the loop regions of tRNA.

The transfer RNA molecules are characterized by the presence of rare nucleosides usually in the non-double-helical segments of the cloverleaf model.[1b] Although only about 15–20% of the tRNA molecule is made up of rare nucleosides, they constitute about 30–40% of the nucleosides of the loop region. During the past decade considerable detailed information on the stereochemical properties of the common constituents of DNA and RNA have been accumulated through single-crystal X-ray diffraction studies.[2] More recently, the rare nucleic acid constituents have been a subject of great interest, and the crystal structures of a number of these compounds have now been analyzed.[3] Knowledge of the geometries and conformations of the rare nucleosides will aid considerably in molecular model building studies of tRNA. In addition, the hydrogen-bonding and base-stacking properties will be of considerable help in understanding the secondary and tertiary structures of tRNAs. In this paper the crystal and molecular structure of dihydrouridine is presented.

Experimental Section

Dihydrouridine was purchased from Sigma Chemical Co., St. Louis, Mo. Crystals suitable for this investigation were grown from aqueous ethanol solution by Dr. Douglas Rohrer of this laboratory. The crystals were elongated along the a axis and were generally uniform in cross section. Preliminary photographic investigation revealed that the crystals belonged to the orthorhombic system. A crystal measuring 0.2 mm² in cross section and about 0.5 mm in length was mounted on a Picker FACS-1 X-ray diffractometer with the a axis coincident with the ϕ axis of the goniostat. The unit-cell constants were obtained by a least-squares refinement of the angular measurements of ten reflections occurring at medium 2θ levels. These values, together with other pertinent crystal data, are given in Table I.

Table I. Crystal Data for Dihydrouridine

Crystal system	Orthorhombic
Unit cell dimensions, Å	$a = 8.131 \pm 0.001$ $b = 11.766 \pm 0.001$ $c = 23.016 \pm 0.003$
Systematic absences	$h00$, $h = 2n + 1$, $0k0$, $k = 2n + 1$, and $00l$, $l = 2n + 1$
Space group	$P2_12_12_1$
Contents of unit cell	8 ($C_9N_2O_6H_{14}\cdot 0.5H_2O$)
Calculated density, g cm^{-3}	1.538
Measured density, g cm^{-3}	1.538
μ for Cu Kα, cm^{-1}	9.8

For eight formula units of dihydrouridine the calculated density is 1.538 g cm^{-3} which is in agreement with the observed density of 1.538 g cm^{-3}. This indicated that there were two independent molecules of dihydrouridine and one molecule of water in the asymmetric unit.

Intensities for the 2116 reflections with $2\theta \leq 128°$ were measured using nickel-filtered Cu radiation and a θ–2θ scan technique with a 2°/min scan speed. Of the reflections measured 1615 had $I > 1.5\sigma(I)$ and 1917 had $I > 1.3\sigma(I)$, where $\sigma^2(I) = N = B + [0.02(N + B)]^2$, N being the net count and B the net background count for a reflection. The data were corrected for Lorentz and polarization factors but no absorption correction was applied. The linear absorption coefficient (μ) is 9.8 cm^{-1}.

Structure Analysis. Initial attempts to solve the structure using the tangent formula[4] were unsuccessful mainly because of the lack of

(1) (a) Part XXII: R. K. McMullan and M. Sundaralingam, *J. Amer. Chem. Soc.*, **93**, 7050 (1971). A communication on this work has already been published: M. Sundaralingam, S. T. Rao, and J. Abola, *Science*, **172**, 725 (1971). (b) R. W. Holly, J. Apgar, G. A. Everett, J. T. Madison, M. Marguisee, S. H. Merrill, J. R. Penswick, and A. Zamir, *ibid.*, 147, 1462 (1965).

(2) M. Sundaralingam, *Biopolymers*, **9**, 821 (1969).

(3) M. Sundaralingam, Fourth Jerusalem Symposium, "The Purines, Theory and Experiment," B. Pullman and E. Bergmann, Ed., Academic Press, New York, N. Y., 1961.

(4) J. Karle and H. Hauptman, *Acta Crystallogr.*, **9**, 635 (1956).

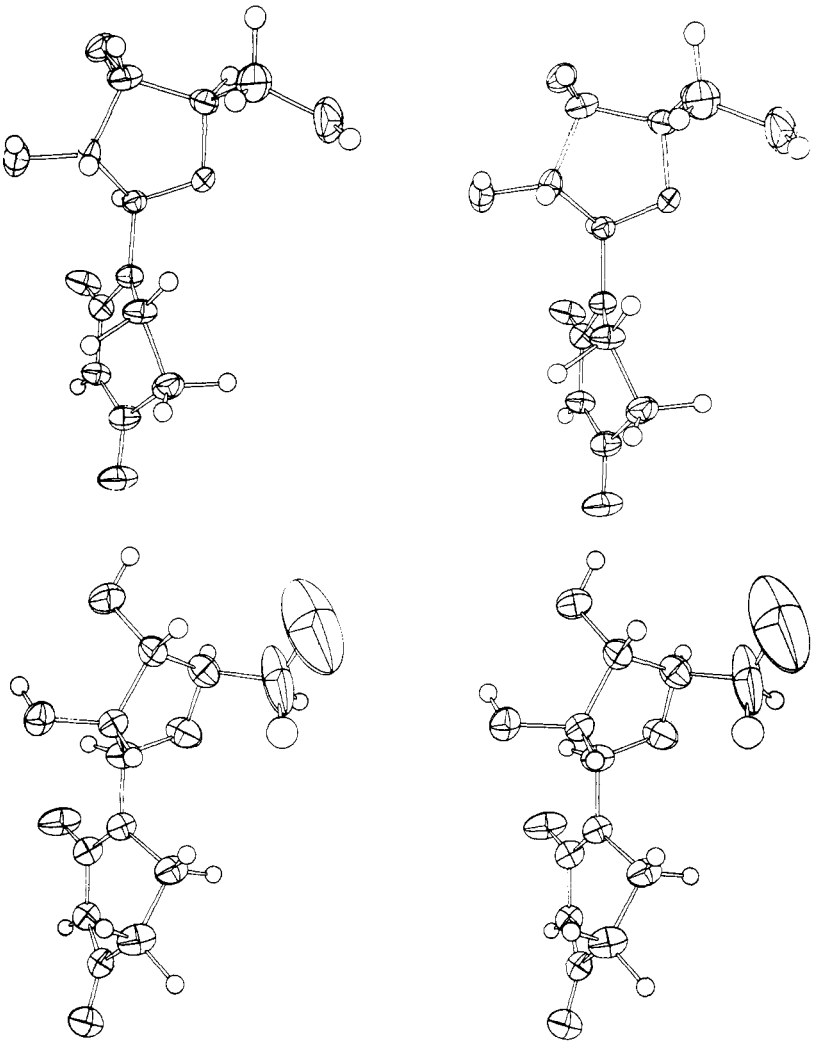

Figure 1. Stereoscopic views of the two molecules A (top) and B (bottom). The thermal ellipsoids of nonhydrogen atoms are at 50% probability level. Hydrogen atoms are shown as open circles.

suitable Σ_2 interactions. The structure was solved by a combination of the Patterson method and phase refinement methods using the tangent formula. An $(E^2 - 1)$ Patterson map, where E is the normalized structure amplitude, was calculated and the peaks around the origin indicated the orientation of one of the pyrimidine rings. This map also provided evidence that the second pyrimidine ring had essentially the same orientation in the unit cell and was stacked with the first ring. The center-to-center vector between the symmetry related rings of one of the molecules was deduced from peaks that were mutually consistent in the three Harker sections. In retrospect, it turned out that this peak represented not the center of the pyrimidine ring but, rather, the ribose. The coordinates of the six atoms of the pyrimidine ring (all treated as carbons) were used to perform one round of structure factor calculation for 234 reflections with $E \geq 1.5$, followed by ten cycles of tangent refinement of the phases. These calculations were performed with programs in the X-ray-67 system of Stewart.[5] The resulting E map showed a number of peaks but did not reveal any identifiable fragment of the molecule. After several unsuccessful attempts to obtain a trial structure from the E map, we finally selected *ten* peaks of which *three* happened to be within 0.5 Å of true atomic sites. The E map produced from this run clearly showed the two pyrimidine rings and their substituents. After additional phase refinement, an E map clearly revealed the positions of all 34 of the atoms in the two nucleosides. The root mean square error in the phase angle in the final run was 13°.

Refinement. The coordinates of the 34 heavy atoms found from the E map were subjected to three cycles of isotropic full-matrix, least-squares refinement[6] using the 1615 reflections with $I > 1.5\sigma(I)$. The R value dropped to 0.22, at which stage a difference electron density map revealed the position of the water molecule. Inclusion of the water molecule in two more cycles of isotropic refinement dropped the R value to 0.11. It was noticed that atom O(5′) of molecule B had a large thermal parameter of 14.8 Å^2. The anisotropic refinement had to be done in stages due to the limited core size of the computer. The parameters of the atoms of molecule A were varied in one cycle and those of molecule B and the water molecule were varied in a successive cycle. One complete round of anisotropic least-squares refinement on all the atoms lowered the R value to 0.08. A difference map was calculated leaving atom O(5′) of molecule B out of the structure factor calculation. The map showed around the O(5′) atom a region of diffuse electron density with two maxima corresponding to two of the three possible staggered arrangements around the C(4′)–C(5′) bond. The two peaks were numerically integrated and the ratio of the two integrals was found to be 7:1. Thus the atom O(5′) of molecule B was assigned to the two sites with occupational values of 0.875 and 0.125, respectively. Refinement was continued, treating the major site with anisotropic thermal parameters and the minor site with isotropic thermal parameters. The difference map gave also the positions of all but four hydrogen atoms. The hydrogen atoms not located were those on C(5′) and O(5′) of molecule B and one of the hydrogen atoms of the water molecule. The hydrogen atoms were included in the refinement with isotropic thermal parameters. Since the parameter to observation ratio was only 4:1, for subsequent refinement 1917 reflections with $I > 1.3\sigma(I)$ were used. Two cycles of refinement

(5) J. M. Stewart, private communication, 1970.

(6) W. R. Busing, K. A. Martin, and H. A. Levy, Oak Ridge National Laboratory Report ORNL-TM-305, Oak Ridge, Tenn., 1962.

Table II. Positional and Thermal Parameters of the Heavy Atoms in Dihydrouridine[a]

Atom	X	Y	Z	B_{11}	B_{22}	B_{33}	B_{12}	B_{13}	B_{23}
W	3119 (9)	3583 (6)	4815 (3)	284 (14)	169 (8)	28 (1)	−13 (9)	−19 (4)	16 (3)
				Molecule A					
N(1)	−4055 (5)	2128 (3)	1320 (2)	76 (6)	36 (3)	9 (1)	9 (4)	5 (2)	2 (1)
C(2)	−3415 (6)	1068 (4)	1323 (3)	84 (8)	35 (3)	12 (1)	11 (5)	0 (2)	0 (2)
O(2)	−2520 (5)	667 (3)	951 (2)	154 (7)	43 (3)	16 (1)	25 (4)	25 (2)	3 (1)
N(3)	−3823 (5)	408 (3)	1804 (2)	104 (7)	40 (3)	11 (1)	6 (4)	10 (2)	4 (1)
C(4)	−4573 (7)	772 (4)	2300 (2)	102 (9)	49 (4)	11 (1)	11 (5)	0 (3)	6 (2)
O(4)	−4826 (6)	117 (3)	2704 (2)	215 (9)	65 (3)	13 (1)	16 (5)	14 (2)	11 (1)
C(5)	−5021 (7)	1995 (4)	2310 (2)	127 (9)	54 (4)	9 (1)	2 (5)	8 (3)	0 (2)
C(6)	−5395 (6)	2432 (4)	1705 (2)	101 (8)	41 (4)	9 (1)	20 (5)	15 (2)	2 (1)
C(1′)	−3716 (6)	2835 (4)	813 (2)	76 (7)	37 (3)	7 (1)	0 (5)	1 (2)	1 (1)
O(1′)	−3187 (5)	3919 (3)	1009 (1)	130 (6)	40 (3)	9 (1)	−12 (3)	−12 (2)	2 (1)
C(2′)	−5178 (6)	3104 (4)	435 (2)	70 (7)	44 (4)	8 (1)	0 (5)	0 (2)	−3 (1)
O(2′)	−5736 (5)	2164 (3)	105 (1)	111 (6)	65 (3)	12 (1)	−22 (4)	−2 (2)	−5 (1)
C(3′)	−4551 (7)	4137 (4)	100 (2)	103 (8)	44 (4)	8 (1)	23 (5)	2 (2)	3 (1)
O(3′)	−3603 (5)	3817 (3)	394 (1)	128 (6)	59 (3)	7 (1)	7 (4)	1 (2)	0 (1)
C(4′)	−3571 (7)	4750 (4)	564 (2)	129 (9)	30 (3)	9 (1)	3 (5)	2 (2)	2 (1)
C(5′)	−4555 (8)	5712 (5)	837 (2)	189 (9)	48 (4)	13 (1)	17 (6)	3 (3)	−4 (2)
O(5′)	−3768 (6)	6149 (3)	1341 (2)	232 (11)	55 (3)	16 (1)	−28 (5)	17 (2)	−12 (1)
				Molecule B					
N(1)	−136 (6)	2981 (4)	2339 (2)	130 (8)	38 (3)	11 (1)	6 (4)	−2 (2)	−2 (1)
C(2)	218 (8)	2210 (5)	1926 (2)	130 (9)	37 (4)	13 (1)	−4 (6)	−5 (3)	−2 (2)
O(2)	547 (8)	1213 (3)	2007 (2)	329 (13)	29 (3)	17 (1)	8 (5)	−4 (3)	2 (1)
N(3)	283 (6)	2632 (4)	1361 (2)	117 (8)	44 (3)	10 (1)	0 (4)	−4 (2)	−6 (1)
C(4)	448 (7)	3750 (5)	1212 (2)	95 (9)	41 (4)	13 (1)	1 (5)	−4 (3)	0 (2)
O(4)	845 (6)	4020 (3)	719 (2)	165 (8)	60 (3)	15 (1)	−5 (5)	7 (2)	1 (1)
C(5)	173 (9)	4586 (5)	1687 (2)	181 (12)	35 (4)	13 (1)	−9 (6)	−5 (3)	−1 (2)
C(6)	−871 (9)	4065 (5)	2166 (2)	178 (12)	45 (4)	9 (1)	20 (6)	−1 (3)	−4 (2)
C(1′)	−169 (8)	2621 (5)	2943 (2)	140 (10)	36 (4)	11 (1)	15 (5)	−4 (3)	0 (2)
O(1′)	−1811 (5)	2583 (5)	3143 (2)	119 (7)	118 (5)	14 (1)	−34 (5)	−1 (2)	−1 (2)
C(2′)	719 (7)	3422 (4)	3356 (2)	95 (9)	35 (4)	10 (1)	−5 (5)	−5 (3)	−1 (2)
O(2′)	2417 (5)	3242 (3)	3362 (2)	110 (6)	30 (3)	17 (1)	10 (4)	0 (2)	5 (1)
C(3′)	−168 (8)	3212 (5)	3933 (2)	109 (10)	65 (5)	11 (1)	16 (6)	0 (3)	−3 (2)
O(3′)	581 (6)	2313 (4)	4253 (2)	172 (9)	124 (5)	16 (1)	35 (6)	5 (2)	25 (2)
C(4′)	−1877 (8)	2855 (5)	3743 (3)	122 (10)	61 (5)	16 (1)	−15 (6)	7 (3)	−2 (2)
C(5′)	−3125 (9)	3785 (9)	3796 (5)	98 (13)	254 (16)	49 (3)	46 (12)	−11 (5)	−74 (6)
O(5′)I[b]	−3445 (9)	4126 (9)	4295 (5)	224 (20)	394 (6)	81 (2)	−55 (9)	17 (4)	−116 (6)
O(5′)II[b]	−2840 (30)	4840 (20)	3660 (10)	3.2 (3)					

[a] The coordinates and thermal parameters are given × 10^{-4}. The standard deviations refer to the least significant digits. The anisotropic temperature factor is of the form $\exp(-(B_{11}hh + \ldots + 2B_{12}hk + \ldots))$. [b] The occupancy parameters for O(5′)I and O(5′)II are 0.875 and 0.125, respectively. Atom O(5′)II was treated with isotropic thermal parameter only.

on all the atoms concluded the refinement. The final R value was 0.056. The shift–σ ratios were in the range 0.01–0.26 with a mean of 0.12. A final difference electron density map showed no significant residual density, the maximum density being 0.10 e/Å^3.

The weighting scheme used in the final refinement was $1/\sqrt{w} = \sigma = 1.70$, $F_0 \leq 21.0$; $1/\sqrt{w} = \sigma = 1.70 + 0.043(F_0 - 21)$, $F_0 > 21.0$. The scattering factors for the heavy atoms were those of Cromer and Waber[7] and for hydrogens those of Stewart, *et al.*[8] Throughout the refinement, the reflections 0, 2, 0; 2, 0, 0; and 1, 2, 5 were given zero weight, as they were suspected to be suffering from secondary extinction. A stereoscopic[9a] view of the two molecules is shown in Figure 1.

Results

The observed and calculated structure factors will appear in the microfilm edition of this journal.[9b] The atomic positional and thermal parameters are listed in Tables II and III. The estimated errors in the bond distances involving nonhydrogen atoms range from 0.006 to 0.009 Å, and the corresponding estimated errors in bond angles range from 0.3 to 0.4°.

Discussion of the Structure. The Saturated Base. The bond distances and angles in the two base moieties are shown in Figure 2. In general there is excellent agreement between the values in the two molecules and they are comparable to the values found in dihydrouracil.[10] The average discrepancy in bond distances between the two molecules is 0.005 Å with a maximum of 0.014 Å occurring for the N(1)–C(6) bond. The average discrepancy in bond angles is 0.9° with a maximum of 2.1° occurring at the angle C(2)–N(1)–C(6). The carbonyl C=O bond distances are very similar to the values found in the planar unsaturated bases. The conformations about the glycosyl bonds are anti, in molecule A $\chi_{CN}^2 = 65.5°$ and in molecule B $\chi_{CN} = 57.1°$.

An interesting feature of this structure is the finding that the half-chair conformations of the base rings are puckered in opposite directions. The least-squares planes through the six atoms show considerable deviation from planarity, Table IV. In both molecules the largest deviation is shown by atom C(6), while the next largest deviation is exhibited by atom C(5) which is on the opposite side of the mean base plane. In molecule A C(6) is displaced on the opposite side as O(1′) while

(7) D. T. Cromer and J. T. Waber, *Acta Crystallogr.*, **18**, 104 (1965).
(8) R. F. Stewart, E. R. Davidson, and W. T. Simpson, *J. Chem. Phys.*, **42**, 3175 (1965).
(9) (a) C. K. Johnson, Oak Ridge National Laboratory Report ORNL-3794, Oak Ridge, Tenn., 1965. (b) The observed and calculated structure factors will appear following these pages in the microfilm edition of this volume of the journal. Single copies may be obtained from the Reprint Department, ACS Publications, 1155 Sixteenth St., N.W., Washington, D. C. 20036, by referring to author, title of article, volume, and page number. Remit check or money order for $3.00 for photocopy or $2.00 for microfiche.
(10) D. C. Rohrer and M. Sundaralingam, *Acta Crystallogr., Sect. B*, **26**, 546 (1970).

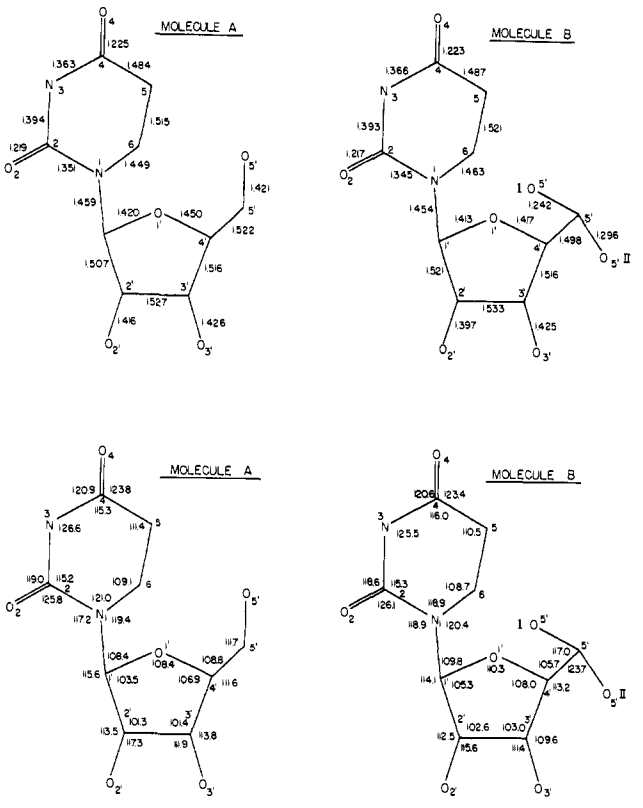

Figure 2. Bond distances and bond angles in the two molecules of dihydrouridine. The estimated standard deviations are given in the text. The disordered positions of O(5′) in molecule B are shown as O(5′) I and O(5′) II.

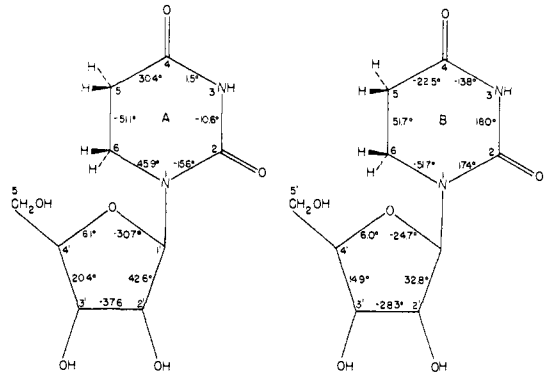

Figure 3. The torsion angles for the base and ribose ring bonds. The furanose ring torsion angles in cytidine 3′-phosphate,[13,14] which may be considered to possess a "standard" $^2T_3$ conformation, are: O(1′) → C(1′) = −19.8°, C(1′) → C(2′) = 36.6°, C(2′) → C(3′) = −38.7°, C(3′) → C(4′) = 28.1°, and C(4′) → O(1′) = −5.6°.[11,12] These angles may be compared with those of the $^2T_1$ conformations found in dihydrouridine.

in molecule B it is displaced on the same side as O(1′). The conformations of the rings with respect to the best four-atom plane and the next best four-atom plane are given in Table IV. The conformations of the base rings are most precisely described by the torsion angles about the ring bonds, Figure 3. It is seen that the signs of the corresponding torsion angles in the two molecules are opposite, again emphasizing the enantiomeric (opposite) relationship of the two bases. It should be noted that the enantiomerism is not exact since the magnitudes of the corresponding torsion angles, in particular those about C(2)–N(3), N(3)–C(4), and C(4)–C(5), are not identical.

Table III. Positional and Thermal Parameters of the Hydrogen Atoms in Dihydrouridine[a]

Atom	X	Y	Z	B
H(W)	387 (9)	409 (6)	502 (3)	5.7 (6)
		Molecule A		
HN(3)	−353 (7)	−21 (4)	178 (2)	3.0 (3)
HC(5)	−601 (7)	207 (4)	248 (2)	3.0 (3)
H′C(5)	−402 (8)	257 (5)	254 (2)	3.9 (3)
HC(6)	−651 (7)	182 (4)	174 (2)	3.0 (3)
H′C(6)	−555 (8)	319 (5)	167 (3)	4.9 (4)
HC(1′)	−289 (9)	238 (6)	61 (3)	5.5 (4)
HC(2′)	−608 (7)	330 (4)	62 (2)	3.0 (3)
HO(2′)	−515 (7)	204 (4)	−8 (2)	3.1 (2)
HC(3′)	−540 (5)	450 (3)	0 (2)	0.9 (2)
HO(3′)	−271 (7)	349 (4)	−24 (2)	3.0 (3)
HC(4′)	−242 (7)	509 (4)	41 (2)	2.0 (3)
HC(5′)	−495 (7)	630 (5)	52 (2)	4.0 (4)
H′C(5′)	−558 (8)	555 (5)	96 (2)	4.2 (4)
HO(5′)	−454 (9)	635 (6)	161 (3)	5.0 (6)
		Molecule B		
HN(3)	58 (7)	217 (5)	111 (2)	3.0 (3)
HC(5)	121 (8)	484 (5)	190 (2)	4.0 (3)
H′C(5)	−47 (8)	524 (5)	144 (2)	4.1 (4)
HC(6)	−216 (7)	397 (5)	207 (2)	3.6 (4)
H′C(6)	−81 (6)	455 (4)	242 (2)	2.4 (5)
HC(1′)	35 (6)	187 (4)	292 (2)	2.0 (5)
HC(2′)	69 (6)	428 (4)	322 (2)	2.0 (4)
HO(2′)	264 (6)	254 (4)	353 (2)	2.0 (3)
HC(3′)	−18 (7)	394 (5)	419 (2)	3.1 (3)
HO(3′)	10 (8)	227 (5)	461 (2)	4.0 (3)
HC(4′)	−241 (6)	222 (4)	384 (2)	2.0 (4)
HC(5′)[b]	−414	353	358	4.0
H′C(5′)[b]	−264	444	354	4.0

[a] The positional coordinates are × 10^{-3}. [b] The positions were geometrically fixed.

The substituent atoms including the ribose C(1′) atom are markedly displaced from the base plane, Table IV. The torsional angles involving the substituent atoms are given in Table V.

Ribose. With respect to the best four-atom plane (Table VI) through C(1′)–O(1′)–C(3′)–C(4′) the ribose conformation is C(2′)-endo in both molecules with C(2′) deviations of 0.632 and 0.495 Å, respectively, for molecules A and B. A novel conformation C(1′)-exo is found for the ribose with respect to the next best four-atom plane. Whenever no combinations of four atoms of the furanoside ring lie in a plane, it is useful to describe the conformation as a twist of a bond with respect to a three-atom plane. The three atoms are selected so that they are common both to the best four-atom plane and the next best four-atom plane.[11] For both molecules of dihydrouridine the three-atom plane is O(1′)–C(4′)–C(3′) and therefore the conformations of the ribose rings are C(2′)-endo–C(1′)-exo or abbreviated as $^2T_1$. The abbreviated nomenclature describes that the ribose has a twist (T) conformation as opposed to the envelope (E) conformation, and the superscript (2) denotes that the major puckering (with respect to the best four-atom plane) is C(2′)-endo and the subscript (1) denotes that the minor puckering (with respect to the next best four-atom plane) is C(1′)-exo.[12]

(11) It turns out that C(2′) and C(1′) show the largest and next largest deviations, respectively, from the plane through the five ring atoms of the ribose. Thus, the choice of the three-atom plane O(1′)–C(4′)–C(3′) can also be made on this basis.

(12) For further details of the abbreviated nomenclature for sugar puckering see M. Sundaralingam, *J. Amer. Chem. Soc.*, **93**, 6644 (1971).

Table IV. Least-Squares Planes through the Bases in Dihydrouridine[a]

Atom	Plane 1, dev, Å		Plane 2, best four-atom plane dev, Å		Plane 3, next best four-atom plane dev, Å	
	A	B	A	B	A	B
N(1)	*−0.014*	*+0.186*	0.197	0.133	*0.020*	*0.034*
C(2)	*−0.071*	*−0.078*	*−0.003*	*−0.030*	*−0.041*	*−0.071*
N(3)	*0.122*	*0.178*	*−0.007*	*0.060*	*0.004*	*0.075*
C(4)	*0.043*	*−0.051*	*−0.006*	*−0.055*	*−0.022*	*−0.038*
C(5)	*−0.249*	*−0.234*	*0.003*	*0.025*	−0.167	−0.022
C(6)	*0.300*	*0.366*	0.729	0.759	0.491	0.653
C(2)	−0.176	−0.241	−0.202	−0.298	−0.177	−0.334
O(4)	0.193	−0.234	−0.037	−0.250	0.024	−0.180
C(1′)	−0.292	−0.502	0.248	0.005	−0.011	−0.154
O(1′)	−1.338	0.641	−0.588	1.278	−0.961	1.089
Rms Δ	0.180	0.200	0.005	0.045	0.033	0.057
Rms σ	0.006	0.006	0.006	0.006	0.006	0.006
Conformation of base[b]			${}_6H_1$	${}^6H^1$	${}_6H^6$	${}^6H_5$

[a] Atoms used in fitting the plane are italic. A and B refer to the two crystallographically independent molecules. [b] The abbreviated nomenclature means the following: ${}_6H^5$ [C(6)-exo-C(5)-endo], ${}^6H_5$ [C(6)-endo-C(5)-exo], ${}_6H_1$ [C(6)-exo-N(1)-exo], ${}^6H^1$ [C(6)-endo-N(1)-endo]. Endo means displacement is on the same side of O(1′) while exo means displacement is on the opposite side of O(1′).

Table V. Torsion Angles in the Base Involving Substituent Atoms

Angle	A, deg	B, deg
C(6)–N(1)–C(2)–O(2)	165.4	−165.9
O(2)–C(2)–N(3)–C(4)	168.5	−158.9
C(2)–N(3)–C(4)–O(4)	−177.8	164.2
O(4)–C(4)–C(5)–C(6)	−150.7	159.6
C(6)–N(1)–C(1′)–C(2′)	−50.2	−60.9
O(2′)–C(2′)–C(3′)–O(3′)	−40.0	−33.7
C(1′)–N(1)–C(2)–O(2)	3.3	−1.1
C(1′)–N(1)–C(2)–N(3)	−177.6	−177.8
C(1′)–N(1)–C(6)–C(5)	−152.4	143.7

Conformation about C(4′)–C(5′) Bond.[14,15] An important stereochemical parameter of consequence in the determination of polynucleotide chain conformations is the conformation about the exocyclic bond C(4′)–C(5′). Molecule A assumes the gauche–trans conformation. Molecule B exhibits disorder, and the two alternative conformations trans–gauche and gauche–gauche occur in the ratio 7:1. Conformational analysis of the common β nucleosides has shown that the preferred conformation about the C(4′)–C(5′) bond is gauche–gauche, although the gauche–trans and the trans–gauche conformations are occasionally exhibited.[2] But, the gauche–gauche conformation does not appear to be favored for dihydrouridine (which possesses the saturated base), rather the alternative conformations, especially the gauche–trans, appear to be preferred (see also ref 16). This finding is important since it may for the first time provide an insight into the structural role of dihydrouridine in the loops of tRNAs. In the current models for double helical nucleic acids and polynucleotides

Table VI. The Deviations of Atoms from Some Least-Squares Planes through Atoms of the Ribose[a] Rings

	Best four-atom plane		Next best four-atom plane		Three-atom plane O(1′)–C(4′)–C(3′)	
	A	B	A	B	A	B
O(1′)	*−0.034*	*−0.033*	*0.080*	*0.057*	*0.000*	*0.000*
C(1′)	*0.021*	*0.020*	0.555	0.440	0.144	0.139
C(2′)	−0.632	−0.495	*−0.074*	*−0.051*	−0.521	0.383
C(3′)	*−0.020*	*−0.018*	*0.115*	*0.081*	*0.000*	*0.000*
C(4′)	*0.033*	*0.031*	*−0.021*	*−0.087*	*0.000*	*0.000*
C(5′)	−1.143	−1.141	−1.540	−1.450	−1.239	−1.230
N(1)	−0.644	−0.756	0.180	−0.114	−0.456	−0.575
O(2′)	−0.477	−0.063	0.450	0.644	−0.521	−0.383
O(3′)	1.255	1.285	1.405	1.370	1.284	1.303
O(5′)	−1.237	−1.254	−1.841	−1.774	−1.368	−1.406
		−2.361		−2.547		−2.426
Rms Δ	0.028	0.026	0.099	0.071	0.000	0.000
Rms σ	0.005	0.006	0.005	0.006	0.005	0.006
Conformation[b]	${}^2E$	${}^2E$	${}_1E$	${}_1E$	${}^2T_1$	${}^2T_1$

[a] Atoms fitted to the plane are shown in italics. [b] The abbreviations used are: E for envelope form and T for twist form. ${}^2E$ = C(2′)-endo, ${}_1E$ = C(1′)-exo, ${}^2T_1$ = C(2′)-endo–C(1′)-exo, and ${}^2T^3$ = C(2′)-endo–C(3′)-endo.

The conformation of the furanoside rings can be precisely described in terms of the torsion angles about the ring bonds, Figure 3. Considering only the signs, the conformations of the two molecules are similar, but the magnitudes of the angles in molecule A, with the exception of that about the C(4′)–O(1′) bond, are significantly larger than those of molecule B. This indicates that molecule A shows a greater puckering than molecule B. For a comparison with the furanoside ring conformations (${}^2T_1$) found in dihydrouridine, the torsion angles in a "standard" C(2′)-endo conformation, ${}^2T_o$ [C(2′)-endo–C(3′)-exo][13] are given in Figure 3.

(13) M. Sundaralingam and L. H. Jensen, *J. Mol. Biol.*, 13, 930 (1965).
(14) M. Sundaralingam, *J. Amer. Chem. Soc.*, 87, 599 (1965).
(15) E. Shefter and K. N. Trueblood, *Acta Crystallogr.*, 18, 1067 (1965).

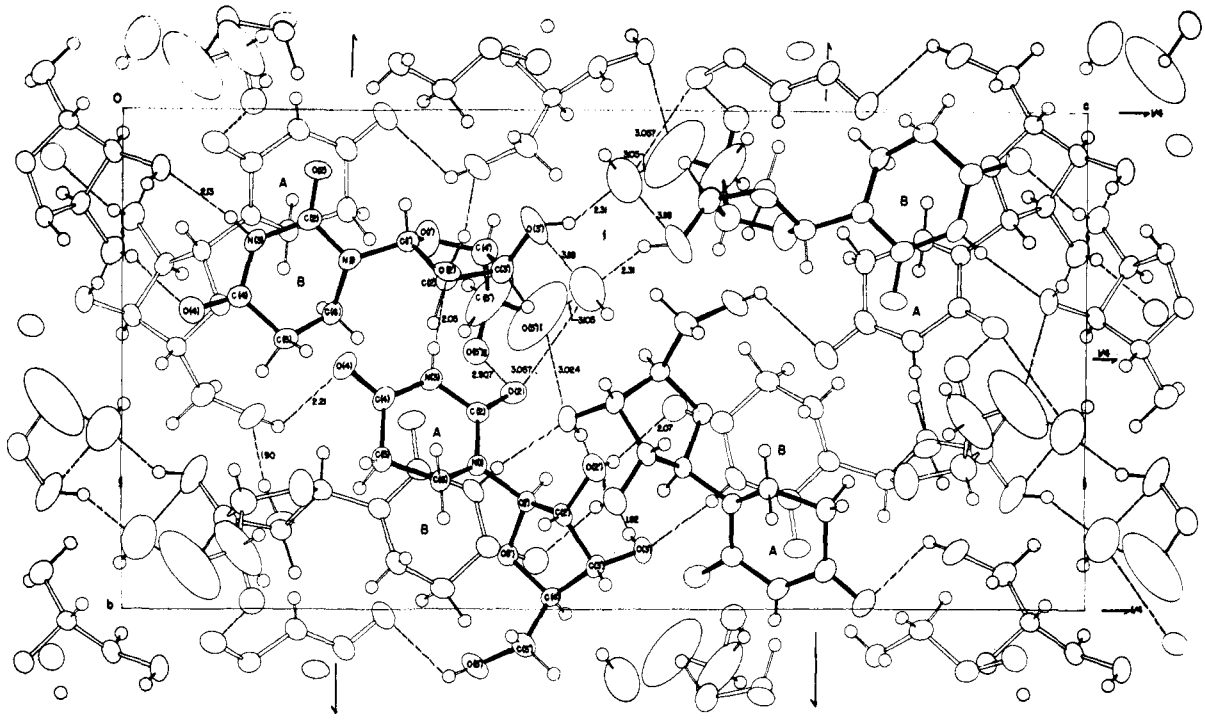

Figure 4. Hydrogen bonding scheme as viewed down the *a* axis. O(5′) I and O(5′) II in molecule B are the two disordered sites with an occupancy ratio of 7:1. The distances between hydrogen atoms and the corresponding acceptor atom are shown except in the case of W and O(5′) B where the distances between heavy atoms are given. Note the environs of the disordered oxygens O(5′) B I and O(5′)B II.

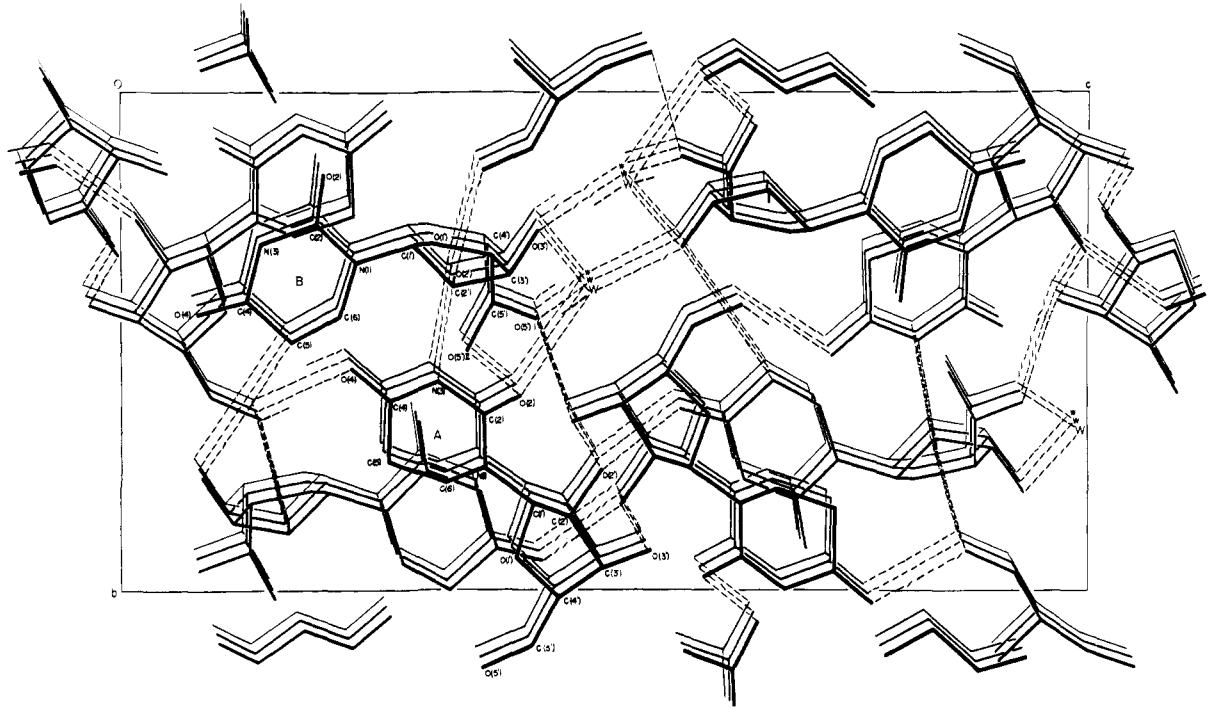

Figure 5. Skeletal drawing of the packing viewed down the *a* axis. The hydrogen bonding networks at three different levels along the *a* axis are shown. The hydrogen atoms are not shown.

the C(4′)–C(5′) conformation is gauche–gauche[2,17] whereas consideration of molecular models suggests that for the formation of a loop in a polynucleotide chain it is advantageous to have the gauche–trans or the trans–gauche conformations. Alternatively, loops in a polynucleotide chain can be generated by rotations about the ester P–O bonds (see also ref 3). Consequently, the preferred conformation of dihydrouridine may disrupt the regular helical sugar–phosphate chain conformation to initiate loop formation. The torsion angles about the C(4′)–C(5′) bond and other angles involving the ribose ring are given in Table VII.

(16) In dihydrothymidine (J. Konnert, I. L. Karle, and J. Karle, *Acta Crystallogr., Sect. B*, 26, 770 (1970)) the disordering of O(5′) was between the gauche–trans and the gauche–gauche conformations in the ratio 2:1.

(17) S. Arnott, *Progr. Biophys. Mol. Biol.*, 21, 267 (1970).

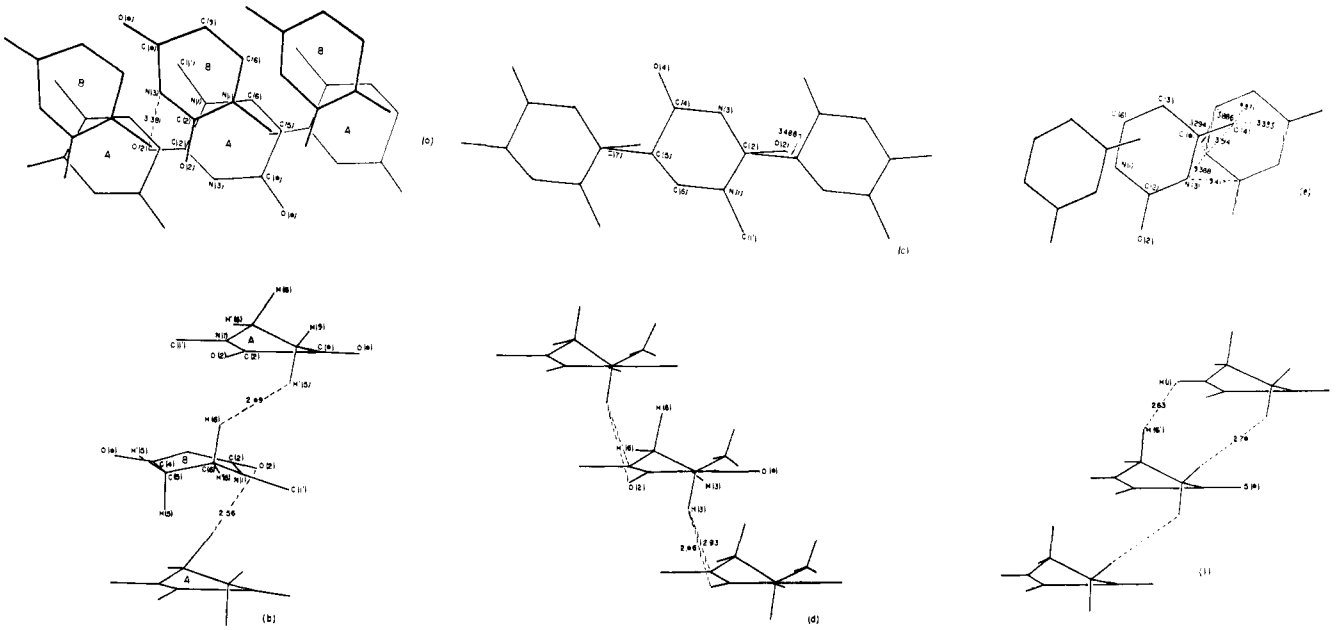

Figure 6. Stacking configurations in a and b, dihydrouridine, c and d, dihydrothymidine,[16] and e and f, dihydrouracil.[10] The first view is at right angles to the plane C(2)–N(3)–C(4)–C(5) and the second view is edge on to this plane. Some of the shortest contacts are also marked.

Hydrogen Bonding. The hydrogen bonding scheme in the structure is shown in Figures 4 and 5. The distances and angles are listed in Table VIII. All hydrogen atoms that are covalently bonded to nitrogen or oxygen atoms are engaged in hydrogen bonding. For convenience of discussion the hydrogen bonding is divided into the following types: A–A, B–B, A–B, and A,B–water, where A and B designate the crystallographically independent nucleosides.

Table VII. Some Torsional Angles Involving the Ribose Moieties

	A, deg	B, deg
N(1)–C(1′)–O(1′)–C(4′)	−153.9	−148.0
N(1)–C(1′)–C(2′)–C(3′)	161.0	153.3
C(1′)–O(1′)–C(4′)–C(5′)	126.9	127.5
C(1′)–C(2′)–C(3′)–O(3′)	84.1	89.2
O(2′)–C(2′)–C(3′)–O(3′)	−40.0	−33.7
C(2′)–C(3′)–C(4′)–C(5′)	−98.6	−101.8
O(3′)–C(3′)–C(4′)–C(5′)	141.1	139.4
O(1′)–C(4′)–C(5′)–O(5′)	51.3	167.7 (I), 77.1 (II)
C(3′)–C(4′)–C(5′)–O(5′)	169.1	−65.2 (I), 41.0 (II)

There are two strong hydrogen bonds between A molecules, one involving the ribose hydroxyls O(3′)A···O(2)A and the other is between the base and the ribose O(5′)A···O(4)A. On the other hand, there are no strong hydrogen bonds between B molecules. The only weak hydrogen bond between B molecules involves the disordered oxygen O(5′)B II and O(2)B of 2.907 Å. In addition to this hydrogen bond, O(5′)B II shows the following contacts to A molecules: O(5′)B II···O(3′)A = 2.936 Å, O(5′)B II···N(3)A = 2.984 Å. O(5′)B I is involved in two contacts, O(5′)B I···O(3′)A = 3.024 Å and O(5′)B I···W = 3.105 Å. Now the disordering of O(5′)B may be explained as due to the fact that in both positions O(5′)B I and O(5′)B II, *i.e.*, for the conformations trans–gauche and gauche–gauche, respectively, there are several possible weak "hydrogen" bonded contacts with neighboring atoms.

The water molecule is involved in a reasonably strong hydrogen bond with O(3′)B. Also, the water shows two other contacts: W···O(3′)B = 3.118 Å and W···O(2)A = 3.057 Å. The only donor sites of the bases, N(3), are involved in hydrogen bonding to the ribose oxygens, N(3)A···O(2′)B = 2.819 Å and N(3)B···O(3′) = 2.944 Å. The remaining A–B hydrogen bonds are O(2′)A···O(4)B = 2.682 Å and O(2′)B···O(5′)A = 2.783 A. Thus the A–B hydrogen bonding dominates the crystal structure.

Table VIII. Hydrogen Bond Lengths and Angles in Dihydrouridine

Atoms A	Atoms H	Atoms B	Symmetry code[a]	A···B, Å	H···B, Å	A–H···B, deg
Scheme A—H···B						
N(3)A	HN(3)A	O(2′)B	1	2.819	2.05	169.7
O(2′)A	HO(2′)A	O(4)B	2	2.682	2.07	151.3
O(3′)A	HO(3′)A	O(2′)A	2	2.685	1.82	168.9
O(5′)A	HO(5′)A	O(4)A	3	2.757	2.21	147.7
N(3)B	HN(3)B	O(3′)A	2	2.944	2.13	172.2
O(2′)B	HO(2′)B	O(5′)A	4	2.783	1.90	157.8
O(3′)B	HO(3′)B	W	5	2.859	2.31	147.1
Short Contacts and Possible Hydrogen Bonds[b]						
O(5′)B I		O(3′)A	6	3.024		
O(5′)B II		O(2)A	1	2.907		
		O(3′)A	6	2.936		
		N(3)A	3	2.984		
W		O(3′)B	7	3.118		
		O(5′)B I	8	3.105		
		O(2)A	1	3.057		

[a] Symmetry operation code no. and equivalent positions: 1, $-x, -1/2+y, -z$; 2, $1/2+x, 1/2-y, -z$; 3, $1-x, -1/2+y, -z$; 4, $-x, 1/2+y, 1/2-z$; 5, $-1/2+x, 1/2-y, 1-z$; 6, $1/2-x, 1-y, -1/2+z$; 7, $1/2+x, 1/2-y, 1-z$; 8, $1+x, y, z$. [b] Since the corresponding hydrogen atoms were not located, an exact assignment of the hydrogen bonding scheme is not possible.

Base Stacking. An important question regarding dihydrouridine is whether the saturated base can participate in parallel stacking of the bases. In fact, the base stacking configuration observed here is in many respects similar to that observed in the known planar pyrimidine systems[18] with the carbonyl oxygen atoms O(2) of the two molecules lying either over or close to the rings of adjacent bases. Figure 6 shows some intermolecular contacts between atoms of stacked bases in dihydrouridine and related saturated bases in views normal and parallel to the bases.

(18) C. E. Bugg, J. Thomas, M. Sundaralingam, and S. T. Rao, *Biopolymers*, 10, 175 (1971).

Acknowledgments. It is a pleasure to thank Dr. Douglas Rohrer for preparing the crystals. We gratefully acknowledge the United States National Institutes of Health for generous support of this research through Grant No. GM 17378. Further thanks are due to the University Computing Center for use of their facilities and supported by grants from the Wisconsin Alumni Research Foundation with funds administered by the University Research Committee.

Interaction of Metal Ions with Polynucleotides and Related Compounds. XVIII. The Multiplicity of Reactions of Copper(II) with Inosine and Its Derivatives

N. A. Berger and G. L. Eichhorn

Contribution from the Laboratory of Molecular Aging, Gerontology Research Center, National Institutes of Health, National Institute of Child Health and Human Development, Baltimore City Hospitals, Baltimore, Maryland 21224. *Received March 18, 1971*

Abstract: Inosine displays remarkable versatility in its coordination to copper ions under various conditions. Proton magnetic resonance studies in D_2O indicate that copper(II) binds to both the N-7 and N-1, O-6 areas of inosine at intermediate pH, preferring N-7 at pH 3, but N-1 at pH 7 and above. The N-1 (O-6) site of inosine is also favored by heating to 80°. Blocking with methyl groups limits Cu(II) binding to N-7 in 1-methylinosine and N-1 in 7-methylinosine. In dimethyl sulfoxide Cu(II) preferentially binds to the hydroxyl groups of the ribose portion of inosine and its 1-methyl and 7-methyl derivatives, as indicated by pmr broadening and continuous variation studies in the visible, demonstrating 2:1 copper–nucleoside binding. The visible spectra are compatible with the occurrence of copper–copper bonds. These characteristics of the spectra of Cu(II)–ribose complexes are absent in deoxyinosine and triacetylinosine.

Copper has profound effects on the structures of polynucleotides and DNA. Under appropriate conditions Cu(II) is capable of bringing about both the unwinding and rewinding of the double helix.[1-4] Unwinding is aided by coordination of Cu(II) ions to electron-donor sites on the purine and pyrimidine bases resulting in withdrawal of electrons and weakening of interstrand hydrogen bonds.[1] Rewinding and renaturation require that some of the bases on opposite strands be held together, or in register, until conditions favor the return of the double helix.[2] The bases may be held in register by copper ions coordinated between bases of the opposing strands.

In earlier studies in which nmr was used to determine the nature of the Cu(II) binding sites in polynucleotides and their constituents several of these coordination sites were identified.[5-8] When paramagnetic copper ions bind to a ligand, the protons in the vicinity of the paramagnetic ion are rapidly relaxed resulting in the broadening of their characteristic proton magnetic resonance peak.[9,10] Using this technique, Cu(II) ions were demonstrated to bind to multiple sites on the adenine base with preference for a given site influenced by molecular conformation and association, which in the AMP isomers varies with the position of the phosphate on the ribose.[7] Cu(II)-induced broadening of the H-8 and H-2 resonances of 5′-IMP was observed and attributed to the possibility of a chelate involving N-7 and O-6 as proposed earlier[11,12] or also coordination at N-1.[8] Both the H-8 and H-2 of poly(I) were also broadened by Cu(II), although the broadening of the H-8 in advance of H-2 indicates preferential coordination at N-7 on the polymer.[8] The present study was undertaken to delineate more clearly under what conditions the various Cu(II) binding sites on inosine (see formula) are favored. The nucleosides are used to eliminate the complication of metal ion bound to

(1) G. L. Eichhorn, *Nature (London)*, **194**, 474 (1962).
(2) G. L. Eichhorn and P. Clark, *Proc. Nat. Acad. Sci. U. S.*, **53**, 586 (1965).
(3) G. L. Eichhorn and E. Tarien, *Biopolymers*, **5**, 273 (1967).
(4) G. L. Eichhorn and Y. A. Shin, *J. Amer. Chem. Soc.*, **90**, 7323 (1968).
(5) M. Cohn and T. R. Hughes, Jr., *J. Biol. Chem.*, **237**, 176 (1962).
(6) G. L. Eichhorn, P. Clark, and E. D. Becker, *Biochemistry*, **5**, 245 (1966).
(7) N. A. Berger and G. L. Eichhorn, *ibid.*, 10, 1847 (1971).
(8) N. A. Berger and G. L. Eichhorn, *ibid.*, **10**, 1857 (1971).
(9) N. C. Li, R. L. Scruggs, and E. D. Becker, *J. Amer. Chem. Soc.*, **84**, 4650 (1962).
(10) D. R. Eaton and W. D. Phillips, *Advan. Magn. Resonance*, 1, 103 (1965).
(11) E. Frieden and J. Alles, *J. Biol. Chem.*, **230**, 797 (1957).
(12) A. T. Tu and C. G. Friedrich, *Biochemistry*, 7, 4367 (1968).