Study of Conformational Changes Induced on Substituting NH for O at C(5′) of Thymidine Nucleosides and Nucleotides

Eileen M. Nottoli, Joseph B. Lambert,* and Robert L. Letsinger*

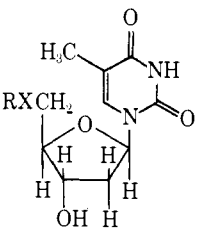
Contribution from the Department of Chemistry, Northwestern University, Evanston, Illinois 60201. *Received October 12, 1976*

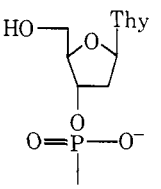
Abstract: The 300-MHz proton magnetic resonance spectra of a series of compounds derived from 5′-amino-5′-deoxythymidine are compared directly with spectra of the corresponding parent thymidine derivatives in order to assess the conformational changes induced on substituting NH for O at C-5′ of thymidine nucleosides and nucleotides. Analysis of chemical shifts and coupling constants indicates that (a) the thymine ring in the 5′-NH analogues has an anti orientation similar to that of the parent compounds, although the average torsional angles (Φ_{CN}) may differ somewhat, (b) the puckering of the furanose ring in the two families is essentially the same (weighted toward 2′-endo), and (c) the proportion of the gauche-gauche rotamer at the C(4′)-C(5′) bond is less in the 5′-NH analogues than in the 5′-O compounds. From $J_{5'P}$ and $J_{3'P}$ for d-TpT and d-Tp(NH)T one may conclude that the population of the gauche′-gauche′ rotamer at the P-X-C(5′)-C(4′) segment is significantly reduced in the 5′-NH compound although the torsional arrangement at the O(3′)-C(3′)segment in the internucleotide link is essentially unchanged.

The conformation of nucleosides can be described in large part by the torsional angles at the C(1′)-N(1) bond, the C(4′)-C(5′) bond, and the furanose ring C-C bonds. X-Ray data have shown that these angles generally fall in narrow ranges for nucleosides and nucleotides in the solid state, corresponding to either an anti or a syn orientation of the base at the glycosidic carbon, either a C(2′)-endo or a C(3′)-endo conformation for the furanose ring, and a gauche-gauche (gg), a gauche-trans (gt), or a trans-gauche (tg) conformation for the protons at C(4′)-C(5′).[1] Nuclear magnetic resonance[2-5] and optical data[5] for nucleosides and nucleotides in solution have been rationalized in terms of equilibrium mixtures involving these favored conformations. The results indicate a strong preference for the anti arrangement of the base ring (H-6 lying over the furanose ring) and for the gg arrangement at C(4′)-C(5′) (O-5′ lying over the furanose ring) for pyrimidine nucleosides and nucleotides in aqueous media. The equilibrium between the C(2′)-endo and C(3′)-endo conformations appears to lie near 1:1 for the ribonucleotides[3] and to favor the C(2′)-endo conformation in deoxyribonucleosides,[3] deoxyribonucleotides [70:30 C(2′)-endo-C(3′)-endo],[4] and ribonucleotides [60:40 C(2′)-endo-C(3′)-endo].[4] Thus the most heavily weighted conformation for thymidine is that represented schematically in Figure 1.

The proton magnetic resonance spectra of d-TpT,[6] d-TpTp,[6] ApA,[7] and d-ApA[7] have recently been completely analyzed. These studies reveal conformations similar to those for nucleosides and mononucleotides. The principal change noted was an increase in the contribution of the gg conformation at the C(4′)-C(5′) bond of the nucleotide unit at the 3′ terminus. At the internucleotide O(5′)-C(5′) bonds the gauche′-gauche′ conformation (both protons at C-5′ gauche to phosphorus) predominates.

Studies in our laboratories have shown that phosphoramidate analogues of polynucleotides, in which one or more of the oxygen atoms at the 5′ positions in the internucleotide links have been replaced by NH, possess interesting chemical and enzymatic properties.[8] The potential utility of such analogues in biological systems is related to the degree to which the geometry of the analogues resembles the geometry of the natural polynucleotides. An analysis of the carbon-13 spectrum of ammonium β,β,β-trichloroethyl-5′-amino-5′-deoxythymidine 5′-phosphate [dTCEp(NH)T] indicated that this substance exists in aqueous solution primarily in the anti conformation characteristic of natural nucleotides.[9] We report in the present paper an analysis of the proton magnetic resonance spectra of d-Tp(NH)T, which serves as a simple model for a section of an analogue chain, and of the related monomer units d(NH)T, d(AcNH)T, and dTCEp(NH)T. The objective was to gain information on the effect of the NH substitution on conformation by directly comparing the ^{1}H NMR spectra of these compounds with the spectra of the corresponding parent compounds [d-TpT, dT, d(AcO)T, and dTCEpT]. This study was augmented by an investigation of the circular dichroism spectra of d-Tp(NH)T and dTCEp(NH)T.



R	X	compd
H-	NH	d(NH$_2$)T
	O	dT
$CH_3C(O)$-	NH	d(AcNH)T
	O	d(AcO)T
$Cl_3CCH_2OP(O^-)$-	NH	dTCEp(NH)T
	O	dTCEpT
	NH	d-Tp(NH)T
	O	d-TpT

Experimental Section

NMR spectra were recorded at 300 MHz on a Varian HR-300 spectrometer at the NMR Center of the University of Akron, Akron, Ohio, which was partially supported by NSF Grant 37725X. Samples run in D_2O were first lyophilized several times from D_2O. Spectral analyses were performed on a Control Data Corp. Model 6400 computer with CalComp Plotter 1136 using the computer programs NMRPLOT,[10] LAOCN3[11] [for d(AcO)T], and ENIT[12] (for all other

Table I. Proton Chemical Shifts (ppm)

Proton	dT	d(NH_2)T	d(AcO)T[a]	d(AcNH)T[a]	dTCEpT	dTCEp(NH)T	d-Tp(NH)T d-Tp-	d-Tp(NH)T -dp(NH)T
H-1′	6.22	6.20	6.28	6.13	6.35	6.26	6.24	6.28
H-2′	2.36	2.34	2.22	2.23	~2.42	2.42	2.39	2.41
H-2″	2.36	2.28	2.27	2.23	~2.42	2.36	2.54	2.37
H-3′	4.44	4.38	4.35	4.24	4.63	4.49	*b*	4.47
H-4′	4.00	3.95	4.10	3.88	~4.18	3.97	4.18	3.98
H-5′	3.82	3.03	4.33	3.44	~4.18	3.22	3.84	3.17
H-5″	3.74	2.94	4.26	3.44	~4.18	3.14	3.79	3.09
CH_3 (base)	1.88	1.89	1.90	1.89	1.95	1.91	1.90	1.90
H-6	7.56	7.35	7.65	7.42	7.73	7.51	7.60	7.67
CH_2 (Cl_3CCH_2)					4.45	4.39		
CH_3 (acetyl)			2.08	1.96				

[a] Measured at 27 °C in D_2O relative to sodium 2,2,3,3-tetradeuterio-4,4-dimethyl-4-silapentanoate, except for d(AcO)T and d(AcNH)T, for which the solvent was CD_3OD and the reference was tetramethylsilane. [b] No signal was observed for this proton; it probably lay under the HOD signal; it could be observed at 75 °C.

compounds). The programs NMRPLOT and ENIT were modified for the specific problems in this work by Mr. Jack Stein of the Vogelback Computer Center, Northwestern University. The average deviation in the error analysis from the ENIT program was 0.10, with a maximum deviation of 0.22. The root mean square error for the calculations for d(AcO)T by LAOCN3 was 0.16.

The circular dichroism spectra were recorded on a Cary 60 spectropolarimeter at 22 °C with the samples in a strain-free 1-cm cell and with a slit program that afforded a 1.5-nm band-pass. The instrument was calibrated with *d*-10-camphorsulfonic acid. Spectra were obtained on ammonium salts in aqueous solution approximately 10^{-4} M in nucleotide units and are reported as molar ellipticity per residue, θ. Concentrations were calculated from extinction coefficients and absorbance values. The extinction coefficients for dTCEpT, dTCEp(NH)T, and d-TpT were taken to be 9.6×10^3, 9.6×10^3 (i.e., the same as for thymidine[13]), and 18.5×10^3,[14] respectively. The extinction coefficient for d-Tp(NH)T (18.7×10^3) was obtained from the hyperchromicity (2%) observed on hydrolysis in aqueous acetic acid (19:1 v/v) and the extinction coefficients of the resulting monomeric components. The elemental analysis was carried out by Micro-Tech, Inc., Skokie, Ill.

Commercial **thymidine** (Sigma Biochemicals) was used without further purification. Reported procedures were followed for the preparations of **5′-amino-5′-deoxythymidine** [d(NH)T] and of **thymidylyl-(3′–5′)-5′-amino-5′-deoxythymidine** [d-Tp(NH)T].[8] The synthesis of **β,β,β-trichloroethyl-5′-amino-5′-deoxythymidine 5′-phosphate** [dTCEp(NH)T] and of the corresponding derivative of **thymidine 5′-phosphate** (dTCEpT) was described previously.[9] **5′-Acetylthymidine** [d(AcO)T] (mp 152–153 °C, lit.[15] mp 150 °C) was prepared by acetylation of 3′-*O*-monomethoxytritylthymidine with Ac_2O, followed by treatment with hot 80% HOAc, by a procedure analogous to that used for synthesis of 5′-*O*-isobutyloxycarbonylthymidine.[16] **5′-Acetylamino-5′-deoxythymidine** [d(AcNH)T] was obtained by acetylating 5′-amino-5′-deoxythymidine (30 min) in pyridine at room temperature: mp 231–232 °C; R_f (MeOH) 0.77; R_f (THF) 0.37. Anal. Calcd for $C_{12}H_{17}N_3O_5$: C, 50.88; H, 6.01. Found: C, 50.33; H, 6.05.

Results

The chemical shift data are summarized in Table I. The line assignments were made by analogy to the assignments for thymidine[17,18] and d-TpT[6] and by examination of the coupling patterns. In each case the lower field proton at C-5′ is designated as the 5′ proton and the higher field proton is designated as the 5″ proton, in accordance with the convention of Davies and Rabczenko.[18] For cases in which the protons at C-2′ can be distinguished, the proton with the larger coupling to H-3′ is designated H-2′ (i.e., cis to H-3′) and the other H-2″, in agreement with data for deoxycytidines specifically deuterated at C-2′.[19] This assignment is also consistent with the observation that for three bond couplings in five-membered rings,

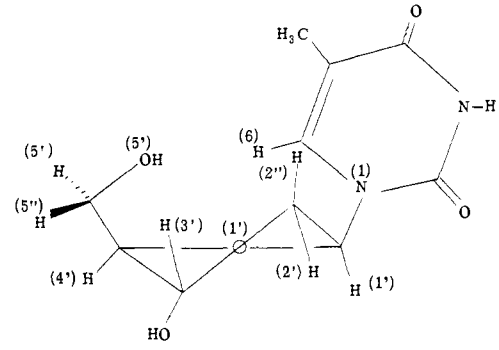


Figure 1. The principal conformation of thymidine.

J_{cis} may be several hertz greater, but not several hertz less, than J_{trans}.[20]

The main difference between the 5′-NH and the 5′-O series of compounds is the large upfield shift (0.8–1.0 ppm) in the resonances for the protons at C-5′ in the amino analogues, a result attributable to the electronegativity difference between nitrogen and oxygen. Another distinctive feature is the upfield shift (~0.2 ppm) in the resonance of the H-6 proton of the thymine ring in the amino analogues.

The similarity in the chemical shifts for the protons in d-Tp(NH)T and d-TpT[21] is striking. For the d-Tp portion the average difference for corresponding protons in this pair amounts to only 0.02 ppm. In the dp(NH)T portion a marked change is found for the 5′ protons (1.0 ppm upfield shift in the nitrogen analogue) as expected; otherwise the resonance positions are very similar (average difference 0.07 ppm).

The proton spectrum of d-Tp(NH)T was run in D_2O at 75 °C as well as at 27 °C. Essentially no change in the chemical shifts was observed for the protons in the deoxyribose rings, each value at the higher temperature being within 0.03 Hz of that at the lower temperature (±0.01 Hz on the average). A small upfield shift (0.06 ppm) was found for H-6 at 75 °C.

The coupling constants for these compounds, determined by standard iterative procedures,[11,12] are listed in Table II. The error is less than ±0.2 Hz. In the case of d-Tp(NH)T each nucleotide unit was analyzed separately. The observed and calculated ^1H spectra for all the compounds have been recorded elsewhere[22] and are included (see paragraph at end of paper regarding supplementary material). Because of overlapping resonances the spectrum of dTCEpT was not analyzed for coupling constants. The *J* values for thymidine agree satisfactorily with those previously reported.[17,18] Both C-2′ and

Table II. Coupling Constants (Hz) and Conformational Populations

						d-TpT[a]		d-Tp(NH)T	
	dT	d(NH_2)T	d(AcO)T	d(AcNH)T	dTCEp(NH)T	d-Tp-	-dpT	d-Tp-	-dp(NH)T
$J_{1'2'}$	*b*	6.8	6.2	*b*	6.9	7.2	*b*	6.3	6.8
$J_{1'2''}$	*b*	6.6	6.8	*b*	6.9	6.1	*b*	7.3	7.0
$J_{1'2'} + J_{1'2''}$	12.9	13.4	13.0	14.1	13.8	13.3	13.8	13.6	13.8
$J_{2'2''}$	*c*	−14.2	−13.5	*c*	−14.6	−14.0	*c*	−14.1	−13.9
$J_{2'3'}$	*b*	7.2	7.5	*b*	7.4	6.0	*b*	7.3	7.0
$J_{2''3'}$	*b*	4.1	3.0	*b*	3.9	3.8	*b*	3.4	3.8
$J_{2'3'} + J_{2''3'}$	11.3	11.3	10.4	10.8	11.3	9.8	11.6	10.7	10.8
$J_{3'4'}$	4.0	4.2	3.2	3.4	4.1	3.2	4.0	3.7	4.6
$J_{4'5'}$	3.5	4.0	5.1	*b*	4.4	3.4	2.5[d]	3.3	4.4
$J_{4'5''}$	5.0	7.6	3.6	*b*	6.6	4.2	4.0[d]	4.4	6.9
$J_{4'5'} + J_{4'5''}$	8.5	11.6	8:7	13.2	11.0	7.6	6.5	7.7	11.3
$J_{5'5''}$	−12.5	−12.5	−12.1	*c*	−14.1	−12.6	*c*	−12.1	−13.9
$J_{3'P}$						6.8		6.7[e]	
$J_{5'P}$					9.9		4.0[d]		9.0
$J_{5''P}$					8.7		4.0[d]		9.5
$J_{5'P} + J_{5''P}$					18.6		8.0[d]		18.5
X_S (2'-endo)		0.62	0.66		0.63			0.63	0.60
% gg	54	28	52	11	34	63	74	62	31
% g'g'					16		82		17

[a] From D. J. Wood, F. E. Hruska, and K. K. Ogilvie, *Can. J. Chem.*, **52**, 3353 (1974). [b] Because $\nu_i - \nu_j$ was very small, only the coupling constant sum could be measured. [c] Not observed. [d] Approximate. [e] Obtained at 75 °C; all others at ambient temperature.

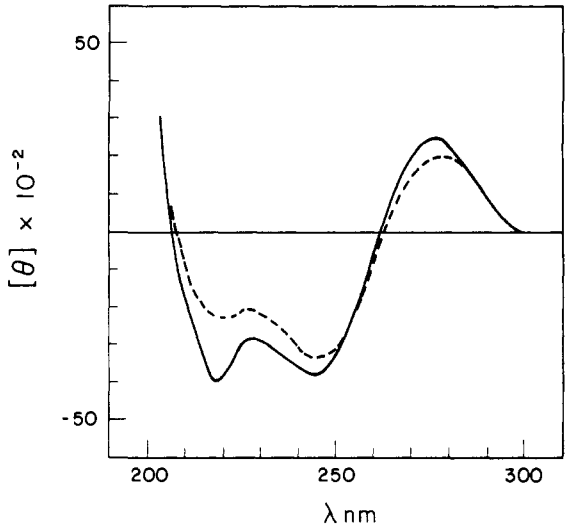


Figure 2. The CD spectra of dTCEp(NH)T (- - -) and of dTCEpT (—) in H_2O at 22 °C.

C-5′ have a pair of diastereotopic protons. When the chemical shift difference between the diastereotopic protons is small relative to the coupling constant, as in thymidine, only average values for the coupling constants to other nuclei can be determined. In such cases only the sums of the coupling constants are included in Table II.

An increase in temperature from 27 to 75 °C had little effect on the coupling constants for d-Tp(NH)T. The greatest effect was in the sum $J_{2'3'} + J_{2''3'}$ for the dp(NH)T portion, which was 2.2 Hz greater at 75 °C. The average difference in coupling constants for a given pair of protons at the two different temperatures was 0.26 Hz.

In addition to nuclear magnetic resonance techniques, circular dichroism (CD) spectra proved helpful in studying conformations of nucleosides and polynucleotides.[5] The CD spectra are especially sensitive to changes in Φ_{CN} in monomers and to base stacking in polymers. Thus pyrimidine nucleosides in the anti conformation exhibit a positive long-wavelength (B_{2u} band) Cotton effect, whereas pyrimidine nucleosides assigned a syn conformation on other grounds exhibit a negative long-wavelength Cotton effect.[23] The spectra of dinucleoside monophosphates, dinucleotides, and higher polymers in general differ markedly (e.g., a more intense B_{2u} band) from the sums of the spectra of their constituents. This fact, indicative of a nonrandom arrangement of the bases in the dimers and polymers, has been interpreted as evidence for base-stacked conformations in these substances.[24]

The CD spectra for the phosphoric derivatives dTCEp(NH)T, dTCEpT, d-Tp(NH)T, and d-TpT are pictured in Figures 2 and 3. That for d-TpT agrees satisfactorily with the spectrum previously reported.[24] Two points are noteworthy. First, the spectra of the amino analogues are very similar to those of the parent oxygen compounds. Second, the maximum molar ellipticity for the B_{2u} band per residue is about twice as great in the dimeric compounds as in the monomeric derivatives. These results indicate that the average torsional angles Φ_{CN} must be similar, although probably not the same, in dTCEpT and dTCEp(NH)T and that the relative disposition of the two thymine rings in d-TpT and in d-Tp(NH)T must likewise be similar.

Discussion

Conformation at the Glycosidic Bond. Schweizer et al. reported that the chemical shifts for the 2′ and 3′ protons in β-ribosylpyrimidines in which a carbonyl group is constrained to lie over the sugar ring (e.g., 6-methyluridine, 6-methylcytidine, 6-oxouridine, and 6-oxocytidine) are displaced downfield (~0.47 and 0.16 ppm, respectively) relative to the 2′ and 3′ protons in the natural (no 6-methyl group) nucleosides.[25] These displacements, stemming at least in part from the magnetic anisotropy and bond polarization effects of the keto group,[25] provide a good probe of the conformation at the glycosidic bond of nucleosides that have a carbonyl group in the 2 position of the pyrimidine. As applied to the 5′-amino analogues, the chemical shift at the 2′ proton is the more diagnostic since this atom is well removed from the 5′-NH, and the change in the chemical shift is greater at this position.

The data in Table I show that for a given pair of nucleosides in which the O → NH substitution has been made, the chemical shift at H-2′ has not changed significantly (difference ≤ 0.03 ppm). These data therefore clearly support the view that the thymine ring in the amino analogues has the anti ar-

rangement characteristic of the parent compounds. The same conclusion about the conformation of dTCEp(NH)T was reached on the basis of the carbon-13 spectrum.[9]

Several factors could account for the upfield shifts (0.04–0.23 ppm) observed for H-6 in the amino analogues. Therefore a definite conclusion concerning the relative population of the syn and anti species cannot be drawn from this change alone. It may be noted, however, that the result is consistent with the view that the anti conformation predominates in both the oxygen and nitrogen series of compounds. The upfield shift of H-6 could result either from the difference in the field effects of the 5′-nitrogen and oxygen or from a decrease in the gg conformation in the amino analogues.[26]

Small changes (0.02–0.09 ppm) were found in the resonance positions of H-1′ for the pairs of related compounds in Table 1. The chemical shift at this position is influenced by the magnetic anisotropy of the neighboring keto group.[25,27] Prestegard and Chan[27] pointed out that a rotation of the base toward larger negative angles within the range Φ_{CN} 0 to −90° would be expected to lead to a downfield shift at H-1′. Accordingly, the shifts observed for H-1′ suggest that the torsion angles for the preferred conformers in the 5′-O and 5′-NH series may differ somewhat, although all the compounds still fall in the anti conformational category (Φ_{CN} +10 to −90°). This conclusion is supported by the similarity of the CD spectra of dTCEp(NH)T and d-Tp(NH)T to the spectra of the parent oxygen compounds (Figures 2 and 3).

Furanose Ring Conformation. Altona and Sundaralingam[3] have developed a method for interpreting the proton coupling constants in ribonucleosides and ribonucleotides based on the assumptions (a) that the spectra reflect a time average of two equilibrating ring conformers (2′-endo and 3′-endo) and (b) that the C–C–H bond angles are approximately equal. Application of this method to a large number of purine and pyrimidine nucleosides yielded internally consistent results. With an electronegativity correction (+0.3 Hz) to compensate for replacement of OH by H at the 2′-carbon, the method was extended with some success to the analysis of deoxyribose monomers.[3] The calculations predict a marked decrease in $J_{1'2'}$ (from 10.4 to 0.3 Hz), a decrease in the sum $J_{1'2'} + J_{1'2''}$ (from 16.1 to 7.1 Hz), and an increase in $J_{3'4'}$ (from 0.2 to 10.1 Hz) when the equilibrium is shifted from the pure 2′-endo to the pure 3′-endo conformation in deoxyribonucleosides.

Davies and Danyluk derived eq 1 by a related scheme (X_S and X_N are the mole fractions of the 2′-endo and 3′-endo conformers, respectively).[4] The equilibrium values calculated for a series of deoxyribonucleotides by this simple equation agreed well with those obtained by a complete pseudorotational analysis.[4]

$$K_{eq} = X_S/X_N = J_{1'2'}/J_{3'4'} \qquad (1)$$

An inspection of Table II shows that $J_{1'2'}$, $J_{1'2'} + J_{1'2''}$, and $J_{3'4'}$ (the coupling constants sensitive to the 2′-endo/3′-endo equilibrium) differ very little between the 5′-O and 5′-NH series of compounds. The changes that are observed are small compared to the ranges predicted by Altona and Sundaralingam for pure 2′-endo and 3′-endo conformations. One may therefore conclude that the conformation of the furanose ring in the amino analogues is much the same as that in the parent oxygen compounds. Furthermore, calculations of K_{eq} from eq 1 using values of $J_{1'2'}$ and $J_{3'4'}$ from Table II yield equilibrium constants corresponding to about 60% 2′-endo conformer in the amino analogues. This value agrees well with the percent of 2′-endo conformer calculated for deoxyribonucleosides (~60%) and is somewhat less than that estimated for monodeoxyribonucleotides (~70%).[4]

Rotamer Preferences at the Exocyclic Bonds. The exocyclic C(4′)–C(5′) bond is near the point of substitution of NH for O, so it is to be expected that this portion of the molecule might be particularly sensitive to the O → NH structural modification. It has been known for some time that rotameric populations depend on the polar nature of substituents in ethane derivatives.[28] The magnitude of the coupling constants between the protons on C-4′ and C-5′ provides a direct handle on the rotameric population (gg, tg, gt) because of the well-known dependence of vicinal coupling constants on dihedral angle, as expressed by the Karplus equation.

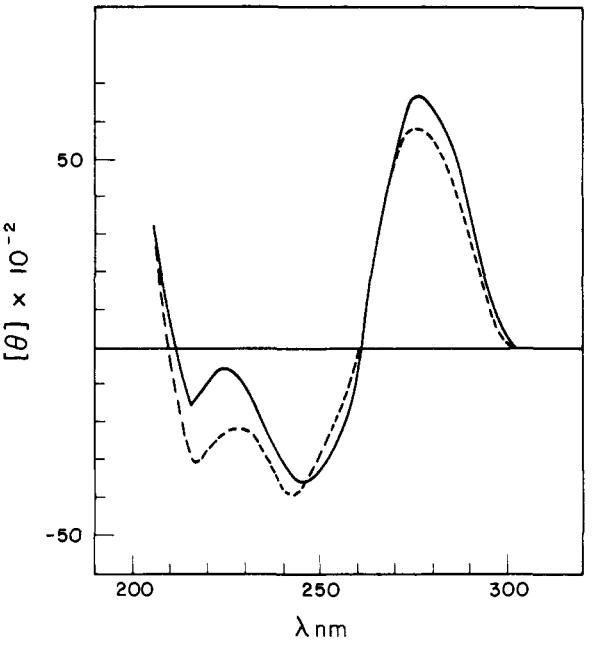
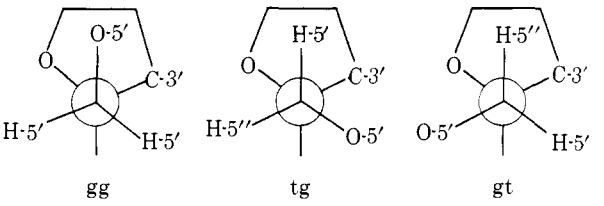


Figure 3. The CD spectra of d-Tp(NH)T (- - -) and d-TpT (—) in H_2O at 22 °C.



Although the coupling constants in the individual rotamers cannot be determined, considerable information can be derived from the observed coupling constant averages, $J_{4'5'}$ and $J_{4'5''}$, and their sum, $J_{4'5'} + J_{4'5''}$. It is immediately evident from Table II that the coupling constants and their sum are invariably larger in the NH species than in the natural O species. This situation contrasts with the general equivalence of coupling constants that define the shape of the five-membered ring, e.g., $J_{1'2'}$ and $J_{3'4'}$. In a comparison of this sort, it is necessary to take into consideration the dependence of the vicinal coupling constant on substituent electronegativity. This dependence has been expressed quantitatively as ${}^3J = 7.9 - 0.7\Sigma\Delta E$, in which $\Delta E = E_X - E_H$ and E is the electronegativity (summed over all substituents X).[29] According to this expression, the coupling constant for the NH species needs to be decreased by about 0.3 Hz in order to be compared directly to the values in the natural O species. Even with this correction (0.6 for the sum $J_{4'5'} + J_{4'5''}$), the difference in $J_{4'5'} + J_{4'5''}$ is 2.5 Hz for dT vs. d(NH_2)T, 3.9 for d(AcO)T vs. d(AcNH)T, and 4.2 Hz for the dpT part of d-TpT compared to the dp(NH)T part of d-Tp(NH)T. It is noted that the coupling constant sum is experimentally identical for the d-Tp parts of d-TpT and d-Tp(NH)T, in which the substitution of NH for O has not been made. The difference between the O and the

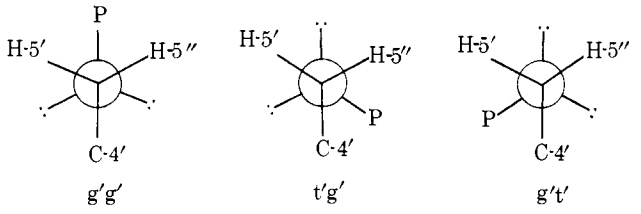
NH coupling constant sum is between 10 and 20 times the possible difference due to experimental inaccuracies in measurement of the coupling constants.

Hruska and co-workers[30] have related the sum of the 4′-5′ couplings to the population of the gauche-gauche conformer by eq 2.

$$\% P_{gg} = 100[13.7 - (J_{4'5'} + J_{4'5''})]/9.7 \quad (2)$$

The rotameric populations for the synthetic NH- and the parent O-type nucleotides and nucleosides calculated by this equation are given in Table II. The coupling constant sum for the NH species has been adjusted by 0.6 Hz to correct for the electronegativity difference between O and NH. It can be seen that the O → NH structural modification decreases the gg population from around 60% to around 25%. Even if the electronegativity correction is doubled to the unreasonable extreme of 1.2 Hz (0.6 Hz per J), the gg population increases by only 6% (28 to 34%). It is therefore apparent that the observed increase in $J_{4'5'} + J_{4'5''}$ on substitution of NH for O results primarily from a decrease in the population of the gg conformer and an increase in the populations of tg and/or gt. These latter conformers possess two protons that have an antiperiplanar relationship, which provides the maximum coupling according to the Karplus expression. The gg conformer, on the other hand, has only gauche relationships between its various proton pairs.

Hruska and Sarma[30] have developed a similar expression for the relationship between the sum of the couplings between



^{31}P and H-5′/H-5″ and the population of the g′g′ rotamer (eq 3) [torsion about the P-X-C(5′)-(C4′)

$$\% P_{g'g'} = 100[25 - (J_{5'P} + J_{5''P})]/20.8 \quad (3)$$

segment]. Again we must assess the influence of electronegativity on the couplings. The values of J_{PH} in $[(CH_3)_2N]_3PO$ and $(CH_3O)_3PO$ are 9.4 and 11.0 Hz, respectively.[31] Consequently, we adopted the correction factor of 1.5 Hz per J. From Table II, it can be seen that the individual $J_{5'P}$ couplings are about twice as large in the NH dinucleotide (around 9 Hz) as in the natural O species (4.0 Hz). This change in the coupling constant reflects a decrease in the g′g′ rotamer population from about 80% to about 20%, according to eq 3. These values are approximate at best, but an alternative approach developed by Davies and Danyluk[4] provides similar results to the Hruska method. It has been noted previously[30] that a reduction in the population of the g′g′ rotamer only occurs when there also is a reduction in the population of the gg rotamer. Our results are in line with these observations.

The coupling between ^{31}P and H-3′ can provide information on the torsional arrangement about the O(3′)-C(3′) bond. In both d-Tp of d-TpT and d-Tp of d-Tp(NH)T, the value of $J_{3'P}$ is 6.8 Hz. Thus the conformation of this segment is not altered appreciably by the O → NH change, as expected for the greater distance from the structural modification. This small value of $J_{3'P}$ indicates that the gauche rotamers are probably more highly populated than the trans rotamer, since the antiperiplanar relationship between P and H-3′ in the latter rotamer would generate a coupling constant in the neighborhood of 20 Hz.

Summary

Substitution of NH for O on C-5′ of nucleotides and nucleosides appears to have little effect on the shape of the five-membered ring or on the syn/anti torsional relationship between the base and the sugar rings. Analysis of the vicinal couplings in the exocyclic portions, however, indicates that the proportion of gg rotamer in the XC(5′)-C(4′) segment decreases on substitution of NH for O, and in dinucleotides the proportion of the g′g′ rotamer in the PX-C(5′) segment also decreases. There is no detectable change, however, in the C(3′)-OP segment of the dinucleotides.

The overall similarity in the conformational features of the 5′-NH and 5′-O series of compounds suggests that the phosphoramidate analogues can serve a useful role as surrogates for nucleotides and polynucleotides in the study of nucleic acid chemistry. Since the preferred conformation at the internucleotide link has been altered, however, one can expect some destabilization in formation of hybrids between natural polynucleotides and complementary phosphoramidate analogues.

Acknowledgments. This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health (GM 10265) and the National Science Foundation (Grant No. CHE-75-05006). We are also grateful to Northwestern University for computer time on the CDC 6400.

Supplementary Material Available: proton NMR spectra (Figures 4-20) of d(NH_2)T, dT, d(AcNH)T, d(AcO)T, dTCEp(NH)T, dTCEpT, d-Tp(NH)T, and d-TpT (19 pages). Ordering information is given on any current masthead page.

References and Notes

(1) J. Donohue and K. N. Trueblood, *J. Mol. Biol.*, **2**, 363 (1960); A. E. V. Haschemeyer and A. Rich, *ibid.*, **27**, 369 (1967); M. Sundaralingam, *Biopolymers*, **7**, 821 (1969).
(2) C. D. Jardetzky, *J. Am. Chem. Soc.*, **82**, 229 (1960); **84**, 62 (1962); F. E. Hruska, *Jerusalem Symp. Quantum Chem. Biochem.*, **5**, 345 (1973).
(3) C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, **85**, 2333 (1973).
(4) D. B. Davis and S. S. Danyluk, *Biochemistry*, **13**, 4417 (1974).
(5) P. O. P. Ts'o, "Basic Principles in Nucleic Acid Chemistry", Vol. I, Academic Press, New York, N.Y., 1974.
(6) D. J. Woods, K. K. Ogilvie, and F. E. Hruska, *Can. J. Chem.*, **53**, 2781 (1975).
(7) F. E. Evans, C.-H. Lee, and R. H. Sarma, *Biochem. Biophys. Res. Commun.*, **63**, 106 (1975).
(8) R. L. Letsinger and W. S. Mungall, *J. Org. Chem.*, **35**, 3800 (1970); R. L. Letsinger, J. S. Wilkes, and L. B. Dumas, *Biochemistry*, **15**, 2810 (1976); R. L. Letsinger, B. Hapke, G. R. Petersen, and L. B. Dumas, *Nucleic Acids Res.*, **3**, 1053 (1976).
(9) E. M. Nottoli, J. B. Lambert, and R. L. Letsinger, *Org. Magn. Reson.*, in press.
(10) K. B. Wiberg and B. J. Nist, "The Interpretation of NMR Spectra", W. A. Benjamin, New York, N.Y., 1962.
(11) S. Castellano and A. A. Bothner-By, *J. Chem. Phys.*, **41**, 3863 (1964).
(12) J. D. Swalen and C. A. Reilly, *J. Chem. Phys.*, **37**, 21 (1962).
(13) J. J. Fox and D. Shugar, *Biochim. Biophys. Acta*, **9**, 369 (1952).
(14) T. M. Jacob and H. G. Khorana, *J. Am. Chem. Soc.*, **87**, 368 (1965).
(15) P. T. Gilham and H. G. Khorana, *J. Am. Chem. Soc.*, **80**, 6212 (1958).
(16) K. K. Ogilvie and R. L. Letsinger, *J. Org. Chem.*, **32**, 2365 (1967).
(17) R. U. Lemieux, *Can. J. Chem.*, **36**, 116 (1961).
(18) D. B. Davies and A. Rabczenko, *J. Chem. Soc., Perkin Trans. 2*, 1703 (1975).
(19) B. Fraser-Reid and B. Radatus, *J. Am. Chem. Soc.*, **93**, 6342 (1971).
(20) J. B. Lambert, J. J. Papay, S. A. Khan, K. A. Kappauf, and E. S. Magyar, *J. Am. Chem. Soc.*, **96**, 6112 (1974).
(21) The data for d-TpT are taken from ref 6 for d-TpT in D_2O at 18 °C. Since neither the spectrum of d-TpT (taken at 18 and 65 °C) nor the spectrum of d-Tp(NH)T (taken at 27 and 75 °C) changed much with temperature, a direct comparison of the data for d-Tp(NH)T at 27 °C and for d-TpT at 18 °C is justified.
(22) E. M. Nottoli, Ph.D. Dissertation, Northwestern University, 1976.
(23) D. W. Miles, W. H. Inskeep, M. J. Robins, M. W. Winkley, R. K. Robins, and H. Eyring, *J. Am. Chem. Soc.*, **92**, 3872 (1970).
(24) C. R. Cantor, M. M. Warshaw, and H. Shapiro, *Biopolymers*, **9**, 1059 (1970).
(25) M. P. Schweizer, E. B. Banta, J. T. Witkowski, and R. K. Robins, *J. Am. Chem. Soc.*, **95**, 3770 (1973).
(26) For a discussion of factors influencing the chemical shift of H-6 in pyrimidine nucleosides, see J. Zemlička and J. P. Horowitz, *J. Am. Chem. Soc.*, **97**, 4089 (1975).
(27) J. H. Prestegard and S. I. Chan, *J. Am. Chem. Soc.*, **91**, 2843 (1969).

(28) For example, see M. Hanack, "Conformation Theory", Academic Press, New York, N.Y., 1965, pp 36-41.
(29) C. C. Banwell and N. Sheppard, *Discuss. Faraday Soc.*, **34**, 115 (1962).
(30) D. J. Wood, R. J. Mynott, F. E. Hruska, and R. H. Sarma, *FEBS Lett.*, **31**, 323 (1973); R. H. Sarma, R. J. Mynott, F. E. Hruska, and D. J. Wood, *J. Am. Chem. Soc.*, **95**, 6457 (1973); C.-H. Lee and R. H. Sarma, *J. Am. Chem. Soc.*, **98**, 3541 (1976).
(31) A. Schmidpeter and H. Brecht, *Z. Naturforsch. B*, **24**, 179 (1969).

Ring Stacking Interactions between Thiamin and Planar Molecules as Seen in the Crystal Structure of a Thiamin Picrolonate Dihydrate Complex

Whanchul Shin,[1] James Pletcher,* Gary Blank, and Martin Sax*

Contribution from the Biocrystallography Lab, VA Hospital, Pittsburgh, Pennsylvania 15240, and Crystallography Department, University of Pittsburgh, Pittsburgh, Pennsylvania 15260. *Received September 10, 1976*

Abstract: Thiamin has been crystallized as the salt of picrolonic acid, 3-methyl-4-nitro-1-(*p*-nitrophenyl)-2-pyrazolin-5-one. The structure, which was determined by x-ray diffraction techniques, shows that the picrolonate anion forms two different types of stacking interactions with the neutral pyrimidine ring of thiamin but does not exhibit any planar overlap with the positively charged thiazolium ring. Even though the stacking interactions dominate the crystal packing, thiamin maintains the characteristic F conformation. The analysis of this structure supports the inherent stability of the F conformation. The crystal structure was determined using diffractometer data obtained by the θ-2θ scan technique with Cu radiation from a crystal with space group symmetry $P\bar{1}$ and unit cell parameters a = 10.730 (3), b = 11.161 (2), c = 12.719 (3) Å; α = 108.50 (7), β = 100.62 (10), and γ = 107.88 (7)°. The structure was solved by direct methods and refined by full-matrix least squares to a final R = 0.064 for all 4182 independent reflections and R = 0.045 for the 3142 observed reflections.

Thiamin is a precursor of thiamin pyrophosphate (TPP) which is a coenzyme for enzyme systems catalyzing the transfer of aldehyde or acyl groups such as pyruvate decarboxylase and transketolase. Neither the mode of coenzyme binding nor the nature of the enzyme catalytic site is completely known, although a number of investigators have probed various aspects of this problem. The current understanding of the coenzyme binding has emerged from studies of pyruvate decarboxylase or transketolase using TPP analogues and inhibitors,[2-5] from the UV and CD spectra of the TPP-transketolase complex,[6-8] and from studies of the interaction between thiamin and indoles using NMR and UV techniques.[9,10] Although lacking in detail, the binding of the coenzyme to the enzyme is generally considered to involve an ionic interaction with the pyrosphosphate group, a hydrophobic interaction with the 2′- and 4-methyl substituents, and a charge-transfer interaction between the positively charged thiamin and tryptophan residues in the enzyme.

In the studies which originally suggested the charge-transfer interaction, Biaglow et al.[9] concluded that both the pyrimidine and the thiazolium rings are involved in π-π interactions with indoles which serve to delocalize the positive charge on thiamin. In order to account for the details of the interaction of the indole with both rings of thiamin in solution, thiamin was pictured as assuming a V conformation ($\phi_T \approx \pm 90°$; $\phi_P \approx \mp 90°$)[11] which, on the average, formed a close contact between the six-membered ring of indole and the methylene bridge carbon at the apex of the V. The V conformation, with the 4′-amino group adjacent to the active C(2) on the thiazolium ring, is the form first proposed by Schellenberger[2,12] as the active form of the enzyme bound coenzyme. Results from more than a dozen crystal structures present a different picture for the preferred conformation of thiamin with respect to its C(3,5′) methylene bridge. Basically two conformations have been observed. The S form ($\phi_T \approx \pm 100°$, $\phi_P \approx \pm 150°$) is characteristic of thiamin when C(2) is substituted; the F form ($\phi_T \approx 0°$; $\phi_P \approx \pm 90°$) is characteristic of thiamin when C(2) is free of substituents. These results are also at variance with theoretical calculations which indicate a relatively large apparent freedom of rotation about the two bonds to the methylene bridge carbon, especially when C(2) is unsubstituted.[13]

In an effort to help resolve the conflicting viewpoints about the thiamin conformation and to investigate the parameters controlling its conformation, we have been attempting to crystallize thiamin with molecular fragments that could possibly associate with TPP in the enzyme binding site and to observe how they interact and how such molecules might influence the conformation of thiamin. Tryptophan and other indoles are compounds of obvious interest. Although suitable crystalline complexes have not yet been obtained with any of the indoles, thiamin does crystallize readily with picrolonic acid, 3-methyl-4-nitro-1-(*p*-nitrophenyl)-2-pyrazolin-5-one, in the form of its anion. Picrolonate has been used as a reagent for the identification and assay of thiamin[14] and it has the apparent structural characteristics[15] that enable it to form a molecular complex with thiamin. This structure does provide the opportunity to examine the influence that would be exerted on the thiamin conformation by incorporating a large organic anion into the crystal structure and to examine the mode of association between thiamin and a planar, heterocyclic ring system.

Experimental Section

Pale yellow, bladed crystals grew from an approximately equimolar mixture of thiamin chloride and sodium picrolonate in aqueous solution when it was allowed to stand open to the atmosphere at room temperature. This procedure was adapted from a method for thiamin analysis described by Alexandrova and Alexandrov.[14] The thiamin chloride was prepared by dissolving thiamin chloride hydrochloride (Nutritional Biochemicals) in water, titrating with 1 equiv of sodium hydroxide, and precipitating with acetone at 0 °C. The precipitate was collected by suction filtration, washed with aqueous acetone and acetone, and then air-dried at room temperature. The sodium picrolonate was prepared by dissolving picrolonic acid (Sigma) in water and ti-