Factors Influencing the H_6 Chemical Shift in Pyrimidine Nucleosides[1]

Jiří Žemlička* and Jerome P. Horwitz

Michigan Cancer Foundation and Department of Oncology, Wayne State University School of Medicine, Detroit, Michigan 48201. *Received November 5, 1974*

Abstract: The δ-H_6 of a relatively wide spectrum of pyrimidine nucleosides in a variety of solvents has been examined to gain an understanding of the factors responsible for differences in the deshielding of this proton in the anti conformation. It is shown that the δ-H_6 values of all compounds studied can be separated into two groups: those resembling the simple bases, uracil and 1,3-dimethylthymine, and those lying in the range of corresponding unsubstituted nucleosides. With the latter, maximum deshielding occurs wherein H_6 is juxtaposed (in the anti conformation) to both ribofuranose oxygen, $O_{1'}$, and the primary alcohol oxygen, $O_{5'}$, as indicated in structure **11.** The relative importance of $O_{1'}$ and $O_{5'}$ to the overall effect cannot be assessed at this time. Conformational factors involving $C_{4'}$–$C_{5'}$ and $C_{5'}$–$O_{5'}$ bonds are of a greater importance relative to δ-H_6 than, for example, inductive influences. Thus, molecular models show that a gauche-gauche conformation at $C_{4'}$–$C_{5'}$ brings $O_{5'}$ in close proximity to H_6. Indeed, the g,g conformation is a requirement for positioning of H_6 in the electrostatic field of $O_{5'}$ and thereby ensuring the close proximity of both $O_{5'}$ and $O_{1'}$ to H_6 (cf. structure **11**). Additional support for the requisite anti conformation of pyrimidine nucleosides in solution is provided from CD spectral data.

It is now well recognized that nucleosides can exist in syn and anti conformations due to steric hindrance (vide infra) to rotation about the glycosidic bond relative to the sugar.[2] Molecular models of pyrimidine nucleosides in the anti conformation, wherein the 5,6 double bond of the aglycon is oriented toward the 5′ substituent of the sugar, indicate that the furanose ring oxygen, $O_{1'}$, and (pyrimidine) H_6 are in close proximity. By contrast, H_6 is relatively distant from both $O_{1'}$ and the $C_{5'}$ substituent in the syn conformation. Moreover, it has been deduced from crystallographic studies[3] that the major barrier to rotation about the *N*-glycosyl bond is the closeness of approach of O_2 and H_6 of the pyrimidine residue to $O_{1'}$ and $H_{2'}$ in both ribo- and 2′-deoxyribonucleosides. Variations in the pentofuranose conformation (i.e., 2′-endo and 3′-endo) can also influence the number of close contacts, particularly in the syn conformation.[3]

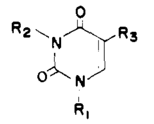
^1H NMR spectroscopy and allied analytical techniques have provided important insights as to the conformation of pyrimidine nucleosides. Thus, from studies of the anisotropic effect of the 2-keto group (O_2) on specific resonance bonds of the sugar,[4,5] from long-range coupling experiments[6,7] along with measurements of nuclear Overhauser effects,[8-12] and from investigation of the influence of conformational chirality on diastereotopic protons,[13] it has been concluded that in solution pyrimidine nucleosides exist almost exclusively in the anti conformation.

Because of the difference in distance between $O_{1'}$ and H_6 the shift of the latter should vary in the two conformations in a manner that reflects a corresponding variance in the through-space electric field effects of the furanose ring oxygen. Indeed, the deshielding of H_6 resulting from the attachment of ribose or 2-deoxyribose to uracil has been ascribed[14] to $O_{1'}$ wherein the nucleosides adopt an anti conformation. A similar explanation has been offered[15] for the observed deshielding of H_6 in pseudouridine. In addition, the possibility of deshielding of the H_6 by 4′-CH_2OH (erroneously referred to as 5′-CH_2OH) where the base is in an anti conformation has been considered.[16] More recently, the influence of the 4′ substituent on the magnitude of the H_6 chemical shift and conformation of some pyrimidine nucleosides has been the subject of a preliminary report.[17] These reports apparently comprise the entire literature on this subject. Accordingly, we undertook a study of a series of pyrimidine nucleosides, along with examples of relatively simple derivatives of uracil and thymine, in various solvents to provide a better understanding of H_6 deshielding.

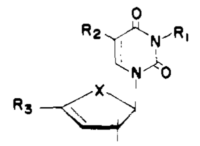
Results and Discussion

It is important to recognize at the outset that the chemical shifts reported in the present paper were determined at levels of concentration ranging from 4 to 12%. It is assumed that values recorded in the literature (Table I) were obtained at comparable concentrations. The line width of the particular signal (H_6 or H_5) was used as a rough indicator of the degree of molecular association (stacking and/or hydrogen bonding). In most cases the observed line widths were between 2 and 3 Hz. For example, it has been reported[10a] that 2′,3′-*O*-isopropylidenecytidine, which is associated in $CDCl_3$, exhibits a line width of its NMR signals of ca. 6 Hz. Thus, the degree of association in thymine and uracil derivatives seems to be substantially smaller.

It is apparent from Table I that the chemical shifts of all compounds examined in this study can roughly be divided into two groups: those resembling the simple bases, uracil and 1,3-dimethylthymine, and those lying in the range of corresponding unsubstituted nucleosides (uridine, thymidine). The H_6 chemical shifts of the 3′,4′-unsaturated nucleosides **2b–d** and the tetrahydrofuryl derivative **10a** approximate those of 1,3-dimethylthymine (**1a**), the cyclopentenyl derivative **2a** (which lacks a ring oxygen), and compounds **1b** and **5a** in which the furanose moiety has a considerable degree of rotational freedom. It is of interest to

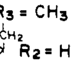


1



2

1a : $R_1 = R_2 = R_3 = CH_3$

1b : R_1 =  $R_2 = H, R_3 = CH_3$

1c : $R_1 = R_2 = R_3 = H$

2a : $R_1 = R_3 = H, R_2 = CH_3, X = CH_2$

2b : $R_1 = R_3 = H, R_2 = CH_3, X = O$

2c : $R_1 = H, R_2 = CH_3, R_3 = COOEt, X = O$

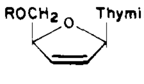
2d : $R_1 = R_2 = CH_3, R_3 = COOCH_3, X = O$

2e : $R_1 = H, R_2 = CH_3, R_3 = CH_2OH, X = O$

2f : $R_1 = R_2 = R_3 = H, X = O$

2g : $R_1 = R_2 = H, R_3 = COOEt, X = O$

2h : $R_1 = R_2 = H, R_3 = CH_2OH, X = O$



3

3a : R = H

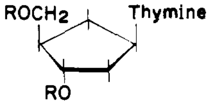
3b : R = Tr

Table I. H_6 Chemical Shifts (δ) of Some Pyrimidine Derivatives and Their Dependence on the Presence of $O_{1'}$, $O_{5'}$, or Another 4′ Substituent

Compound	CDCl$_3$[a]	CD$_3$COCD$_3$[a]	Pyridine-d_5[b]	D$_2$O[b]	$O_{1'}$	4′ substituent[c]	Ref[d]
1,3-Dimethylthymine (1a)	6.99 (2.5)	7.35 (3.0)	6.75 (3.0)	7.83 (3.0)	–		28
1-(D-Tetrahydrofuryl-2-methyl)thymine (1b)	7.16				+		29
Uracil (1c)				7.97 (2.0)	–		*e*
1-(Cyclopent-3-en-1-yl)thymine (2a)	6.96 (3.0)				–		(30)
1-(D-2,3-Dihydrofuryl)thymine (2b)	7.00 (2.5)				+		28
Ethyl 3′-deoxy-3′,4′-didehydrothymine-5′-uronate (2c)	7.01[f]		6.84[f]		+		18
Methyl 3′-deoxy-3′,4′-didehydro-3-methylthymidine-5′-uronate (2d)	6.98[f]				+		18
3′-Deoxy-3′,4′-didehydrothymidine (2e)	7.10 (4.5)	7.30 (3.5)			+		18
1-(D-2,3-Dihydrofuryl)uracil (2f)	7.30 (2.5)				+		28
Ethyl 2′,3′-dideoxy-3′,4′-didehydrouridine-5′-uronate (2g)	7.23 (3.0)				+		18
2′,3′-Dideoxy-3′,4′-didehydrouridine (2h)		7.50 (2.0)			+		18
3′-Deoxy-2′,3′-didehydrothymidine (3a)		7.76 (4.0)			+	CH$_2$OH	(31)
3′-Deoxy-2′,3′-didehydro-5′-*O*-tritylthymidine (3b)	7.03–7.48[g]				+	CH$_2$OTr	(31)
1-(*trans*-3-Hydroxy-*cis*-4-hydroxymethylcyclopent-1-yl)thymine (4a)			7.03 (3.0)	7.93 (3.0)	–	CH$_2$OH	(30)
3,4-Di-*O*-acetyl-1-(*trans*-3-hydroxy-*cis*-4-hydroxymethylcyclopent-1-yl)thymine (4b)	6.93 (3.0)				–	CH$_2$OAc	*h*
2-*O*-Acetyl-3,6-anhydro-1-deoxy-4,5-*O*-isopropylidene-1-(thymin-1-yl)-D-mannitol (5a)	7.10[i]				+		32
2-*O*-Acetyl-3,6-anhydro-1-deoxy-4,5-*O*-isopropylidene-1-(uracil-1-yl)-D-mannitol (5b)	7.27[i]				+		32
2′,3′-*O*-Isopropylideneuridine (6a)		7.80 (3.0)	7.75 (3.0)		+	CH$_2$OH	*e*
5′-*O*-Acetyl-2′,3′-*O*-isopropylideneuridine (6b)	7.33 (4.0)				+	CH$_2$OAc	(33)
5′-Deoxy-5′-iodo-2′,3′-*O*-isopropylideneuridine (6c)	7.37 (3.0)		7.42 (3.0)		+	CH$_2$I	(34), 35
1-(β-D-Arabinofuranosyl)uracil (7a)			7.98 (4.5)	8.27 (3.0)	+	CH$_2$OH	*e*
1-(3,5-Di-*O*-acetyl-β-D-arabinofuranosyl)uracil (7b)	7.71[i]				+	CH$_2$OAc	36
1-(3,5-Di-*O*-trityl-β-D-arabinofuranosyl)uracil (7c)	7.49[i]				+	CH$_2$OTr	37
1-(α-D-Ribofuranosyl)uracil (8)			7.80 (2.0)	8.23 (2.5)	+		*e*
1-(3-Deoxy-3-iodo-2,5-di-*O*-trityl-β-D-xylofuranosyl)uracil (9)	7.65[i]				+	CH$_2$OTr	36
1-(D-Tetrahydrofuryl)thymine (10a)	7.12 (2.5)				+		28
Ethyl 3′-*O*-methylsulfonylthymidine-5′-uronate (10b)	7.82 (3.5)				+	COOEt	18
Ethyl 3′-deoxythymidine-5′-uronate (10c)	8.07 (3.5)				+	COOEt	18
Methyl 3-methylthymidine-5′-uronate (10d)	8.07 (2.5)				+	COOMe	28
Thymidine (10e)			7.78 (3.0)	8.03 (2.5)	+	CH$_2$OH	*e*
3′,5′-Di-*O*-methylsulfonylthymidine (10f)		7.53 (4.0)			+	CH$_2$OMs	(38)
3′,5′-Di-*O*-acetylthymidine (10g)	7.30 (3.5)				+	CH$_2$OAc	(39)
3′-Deoxy-3′-iodo-5′-*O*-(*p*-nitrobenzoyl)-thymidine (10h)	7.13[i]				+	CH$_2$OBz-*p*-NO$_2$	36
3′-Deoxythymidine (10i)		7.85			+	CH$_2$OH	(31)
5′-Deoxy-5′-iodo-3′-*O*-acetylthymidine (10j)	7.57[i]				+	CH$_2$I	35
5′-Deoxy-5′-iodothymidine (10k)			7.37[i,j]		+	CH$_2$I	35
3′-Deoxy-3′-iodo-5′-*O*-acetylthymidine (10l)	7.34[i]				+	CH$_2$OAc	36
3′,5′-Dideoxy-3′,5′-diiodothymidine (10m)	7.47[i]				+	CH$_2$I	36
Uridine (10n)			8.09 (3.0)	8.26 (2.0)	+	CH$_2$OH	*e*
2′-Deoxyuridine (10o)			7.92[k]	8.22 (5.0)	+	CH$_2$OH	*e*
5′-*O*-Trityluridine (10p)	7.93 (5.0)				+	CH$_2$OTr	(40)
Ethyl 2′-deoxy-3′-*O*-methylsulfonyluridine-5′-uronate (10q)	8.03 (3.0)				+	COOEt	18
5′-Deoxy-5′-iodouridine (10r)			7.65[i,l]		+	CH$_2$I	36
2′,3′-Di-*O*-acetyl-5′-deoxy-5′-iodouridine (10s)	7.58[i]				+	CH$_2$I	35
3′,5′-Diiodo-2′,3′,5′-trideoxyuridine (10t)	7.67[i]				+	CH$_2$I	36
2′,5′-Di-*O*-trityluridine (10u)	7.66[i]				+	CH$_2$OTr	37
1-(3-Deoxy-3-iodo-2,5-di-*O*-trityl-β-D-ribofuranosyl)uracil (10v)	7.64[i]				+	CH$_2$OTr	36
3′-*O*-Acetyl-2′,5′-dideoxy-2′,5′-diiodouridine (10w)	7.61[i]				+	CH$_2$I	36
3′-*O*-Acetyl-2′,5′-dideoxyuridine (10x)	7.46[i]				+	CH$_3$	36
3′,5′-Di-*O*-trityluridine (10y)	7.56[i]				+	CH$_2$OTr	37
3′-Deoxy-3′-iodo-5′-*O*-tritylthymidine (10z)	7.62[i]				+	CH$_2$OTr	36

[a] Tetramethylsilane (Me$_4$Si) as an internal standard. Unless stated otherwise the data were derived from 60-MHz spectra. The numbers in parentheses represent the widths of the signals at half-height in hertz. In the case of thymine derivatives, the latter values were not corrected for a long-range H_6–CH_3 coupling which amounts to ca. 1 Hz. [b] Me$_4$Si as an external standard. [c] Only substituents which can interfere with the rotation of the base are listed. [d] References without parentheses indicate the literature from which the H_6 chemical shifts were taken. References in parentheses refer to the procedure used for the preparation of compounds whose NMR spectrum was subsequently determined in our laboratory. [e] Commercial source. The NMR spectrum was taken in our laboratory. [f] Overlapped with $H_{1'}$. [g] Signal overlapped by a trityl envelope. [h] See Experimental Section. [i] A 100-MHz spectrum. [j] Recalculated from the value 7.67 reported[35] for an internal Me$_4$Si. [k] Poor resolution. [l] Recalculated from the value 7.95 reported[35] for an internal Me$_4$Si.

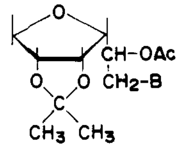
note that the removal of either a 4′-hydroxymethyl or a 4′-carbalkoxy function or replacing $O_{1'}$ by a methylene group all produce a similar effect—an upfield chemical shift of H_6. This has been observed in the case of 4′-carbalkoxy derivatives **10c**, **10d**,[18] and 3′-deoxythymidine (**10i**) relative to the tetrahydrofuryl derivative **10a** in $CDCl_3$ or 1,3-dimethylthymine (**1a**) in CD_3COCD_3. Similarly, H_6 in the carbocyclic analog of thymidine (**4a**) which contains the 4′-hydroxymethyl group but lacks $O_{1'}$ is less deshielded than the same proton in thymidine (**10e**) both in pyridine-d_5 and D_2O. The assumption that the tetrahydrofuryl derivative **10a** exists predominantly in the syn conformation wherein a lesser deshielding of H_6 may be expected is contraindicated by the fact that the CD spectra of **10a** and **10i** both in water and $CHCl_3$ (Figure 1) show a positive Cotton effect at ca. 270 nm (B_{2u} band) which corresponds to that of thymidine (**10e**)[19] and for which an anti conformation is presumed. A more plausible explanation, in which incidentally the more usual anti conformation is the central consideration in both cases, would simply place as in thymidine (**10e**) H_6 in close juxtaposition to both $O_{1'}$ and $O_{5'}$ as illustrated by structure **11**. A similar structure can also be envisioned for carbalkoxy derivatives **10c** and **10d**. Thus, the presence of both $O_{1'}$ and $O_{5'}$ is required for a maximum effect (deshielding) as observed in **10c**, **10d**, and **10i**. When either $O_{1'}$ or $O_{5'}$ is absent, as in, for example, compounds **4a** or **10a**, the overall deshielding effect is considerably weakened and the H_6 signal is shifted upfield toward values corresponding to those of simple bases (**1a**) or compounds **2a** and **1b**. In the latter case (**1b**), a structure similar to **11** may be possible but the nucleobase has a considerable degree of rotational freedom and therefore an important condition for the formation of an intermediate similar to **11**, i.e., an anti conformation of the base, would be more difficult to fulfill. It is not possible at the present time to assess the relative importance (contribution) of $O_{1'}$ and $O_{5'}$ to the overall effect. Models indicate that in the gauche-gauche (g,g) conformation (vide infra) $O_{5'}$ is closer to H_6 than $O_{1'}$. In the case of 3′,4′-unsaturated derivatives **2b**, **2c**, and **2e**, which all have $O_{1'}$ and $O_{5'}$ atoms, the 4′ substituents (carrying $O_{5'}$) lie in the plane of the 3′,4′ double bond which precludes a structure analogous to **11**.



4

4a : R = H

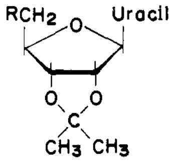
4b : R = Ac



5

5a : B = thymine

5b : B = uracil

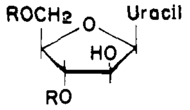


6

6a : R = OH

6b : R = AcO

6c : R = I

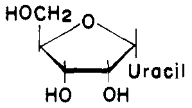


7

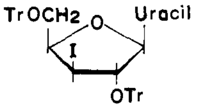
7a : R = H

7b : R = Ac

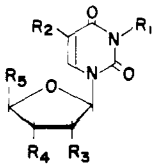
7c : R = Tr



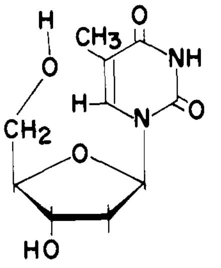
8



9



10a : $R_1 = R_3 = R_4 = R_5 = H$, $R_2 = CH_3$
10b : $R_1 = R_3 = H$, $R_2 = CH_3$, $R_4 = MsO$, $R_5 = COOEt$
10c : $R_1 = R_3 = R_4 = H$, $R_2 = CH_3$, $R_5 = COOEt$
10d : $R_1 = R_3 = H$, $R_2 = CH_3$, $R_4 = OH$, $R_5 = COOCH_3$
10e : $R_1 = R_3 = H$, $R_2 = CH_3$, $R_4 = OH$, $R_5 = CH_2OH$
10f : $R_1 = R_3 = H$, $R_2 = CH_3$, $R_4 = MsO$, $R_5 = MsOCH_2$
10g : $R_1 = R_3 = H$, $R_2 = CH_3$, $R_4 = AcO$, $R_5 = AcOCH_2$
10h : $R_1 = R_3 = H$, $R_2 = CH_3$, $R_4 = I$, $R_5 = AcOCH_2$
10i : $R_1 = R_3 = R_4 = H$, $R_2 = CH_3$, $R_5 = CH_2OH$
10j : $R_1 = R_3 = H$, $R_2 = CH_3$, $R_4 = AcO$, $R_5 = CH_2I$
10k : $R_1 = R_3 = H$, $R_2 = CH_3$, $R_4 = OH$, $R_5 = CH_2I$
10l : $R_1 = R_3 = H$, $R_2 = CH_3$, $R_4 = I$, $R_5 = AcOCH_2$
10m : $R_1 = R_3 = H$, $R_2 = CH_3$, $R_4 = I$, $R_5 = CH_2I$
10n : $R_1 = R_2 = H$, $R_3 = R_4 = OH$, $R_5 = CH_2OH$
10o : $R_1 = R_2 = R_3 = H$, $R_4 = OH$, $R_5 = CH_2OH$
10p : $R_1 = R_2 = H$, $R_3 = R_4 = OH$, $R_5 = TrOCH_2$
10q : $R_1 = R_2 = R_3 = H$, $R_4 = MsO$, $R_5 = COOEt$
10r : $R_1 = R_2 = H$, $R_3 = R_4 = OH$, $R_5 = CH_2I$
10s : $R_1 = R_2 = H$, $R_3 = R_4 = AcO$, $R_5 = CH_2I$
10t : $R_1 = R_2 = R_3 = H$, $R_4 = I$, $R_5 = CH_2I$
10u : $R_1 = R_2 = H$, $R_3 = TrO$, $R_4 = OH$, $R_5 = TrOCH_2$
10v : $R_1 = R_2 = H$, $R_3 = TrO$, $R_4 = I$, $R_5 = TrOCH_2$
10w : $R_1 = R_2 = H$, $R_3 = I$, $R_4 = AcO$, $R_5 = CH_2I$
10x : $R_1 = R_2 = R_3 = H$, $R_4 = AcO$, $R_5 = CH_3$
10y : $R_1 = R_2 = H$, $R_3 = OH$, $R_4 = TrO$, $R_5 = TrOCH_2$
10z : $R_1 = R_3 = H$, $R_2 = CH_3$, $R_4 = I$, $R_5 = TrOCH_2$
$Et = C_2H_5$, $Ms = CH_3SO_2$, $Ac = CH_3CO$,
Tr = triphenylmethyl(trityl), Bz = benzoyl



11

The CD spectra of **10a** and **10i** show a Cotton effect of greater magnitude in chloroform than water. Moreover, the effect is more pronounced with **10a** than **10i** (cf. Figure 1). A similar effect is observed in the case of the unsaturated derivative **2b** (Figure 2); however, the sign of the Cotton effect, as in the case of 2′,3′-dideoxy-2′,3′-didehydrouridine,[19] is reversed. In addition, the influence of the polar solvent on the CD spectra of **10a** and **10i** is opposite to that reported[19] for the 2′,3′-olefinic nucleoside. In the latter case the CD spectra in less polar solvents have been explained in terms of hydrogen bonding of the CH_2OH to the uracil 2-carbonyl group (in syn conformation).[19] Yet, the ir spectrum of the analogous thymidine derivative **3a** in carbon tetrachloride failed to show any intramolecular hydrogen bonding.[20] In this connection it should be noted that the rotation of the

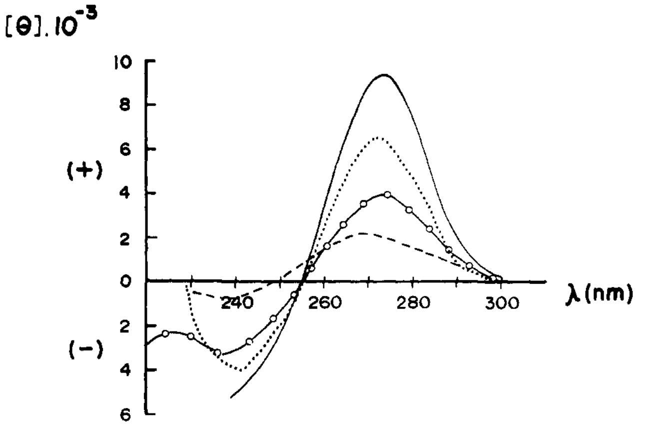


Figure 1. The CD spectra of some pyrimidine nucleosides and related compounds: (---) 1-tetrahydrofurylthymine (**10a**) in water; (——) **10a** in $CHCl_3$; (-O-O-O) 3′-deoxythymidine (**10i**) in water; (···) **10i** in $CHCl_3$.

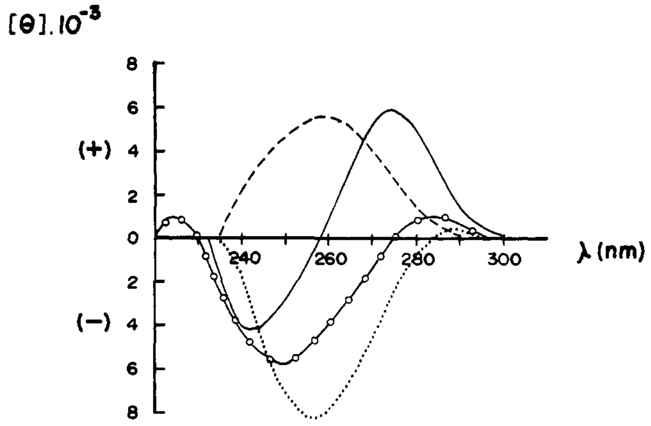


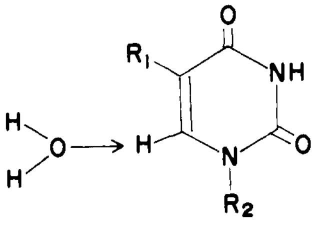
Figure 2. The CD spectra of some pyrimidine nucleoside derivatives and unsaturated analogs: (---) 5′-deoxy-5′-iodo-2′,3′-*O*-isopropylideneuridine (**6c**) in $CHCl_3$; (——) 3′,5′-di-*O*-methylsulfonylthymidine (**10f**) in $CHCl_3$; (-O-O-O) 1-(D-2,3-dihydrofuryl)thymine (**2b**) in water; (···) **2b** in $CHCl_3$.

base about the glycosidic bond in pyrimidine 2′,3′-unsaturated nucleosides is virtually unimpeded by the 4′-hydroxymethyl group.[21] Moreover, the fact that the 2′-hydrogen in these structures lies in the plane of the (2′,3′) olefinic bond precludes any restriction in rotation at $C_{2'}$. By contrast, these considerations do not obtain in either the 3′,4′-unsaturated nucleosides or the tetrahydrofuryl derivative **10a.** In the latter case, as well as in **2b** and **10i,** the observed changes in the intensity of the Cotton effect may be related to a higher proportion of the anti conformer in the less polar ($CHCl_3$) solvent. However, further study is required to clarify this point.

Similarly, H_6 in the 3′,4′-unsaturated nucleosides of the uracil series (**2f–h**) is distinctly more shielded than the same proton in uridine (**10n**), 2′-deoxyuridine (**10o**), α-uridine (**8**), 1-β-D-arabinofuranosyluracil (**7a**), and the corresponding derivatives **10p, 10q,** and **6a.** In point of fact, the chemical shift values of H_6 in **2f–h** lie in a range close to that of **5b** in which the heterocyclic base is free to rotate about the ribofuranose moiety. On the other hand, structures **6a, 7a, 8, 10n, 10p,** and **10q** all exhibit significant H_6 deshielding in chloroform, acetone, and pyridine. By contrast, the magnitude of the δ-H_5 in a series of uracil derivatives is fairly constant as shown in Table II. Thus derivatives **2f** (or **2g**) and **10q,** which display a substantial difference in corresponding H_6 chemical shifts (Table I), have virtually identical values of the δ-H_5. This is not surprising in view of a much greater distance of H_5 from both $O_{1'}$ and $O_{5'}$. With the exception of **9** the trityl derivatives show consistently lower values of δ-H_5 which is probably a consequence of diamagnetic shielding by phenyl groups.

The possible influence (shielding) of the 3′,4′ double bond on H_6 in, for example, **2b,** which incidentally was considered in an earlier paper,[22] is contraindicated by the close similarity (cf. Table I) of the δ-H_6 in the corresponding saturated derivative **10a.** Finally, it is of interest to note that H_6 of the pyrimidine 2′,3′-unsaturated nucleoside **3a** is appreciably deshielded despite unrestricted rotation of the base (vide supra).[21] Thus, the CD results[19] would seem to lead to a conclusion regarding the conformation of **3a** different from that derived on the basis of the H_6 chemical shift. However, it is important to recognize that conditions such as solvent and concentration, in addition to other considerations,[21] are not rigorously comparable. Moreover, the formation of a structure analogous to **11** would be feasible in an anti conformation assuming that $C_{5'}$ and N_1 occupy a pseudoaxial position (cf. ref 21).

It is apparent from Table I that differences in the H_6 chemical shift, which are pronounced where measurements were recorded in nonhydroxylic solvents, are minimal in water. The normalization by water is ascribed to an intermolecular deshielding phenomenon which modulates the intramolecular effect of $O_{1'}$ and $O_{5'}$ (formula **12**). A similar explanation has been invoked in the case of certain adenosine derivatives wherein the intramolecular effect of $O_{1'}$ on the H_8 chemical shift is less pronounced in water than in dimethyl sulfoxide.[23]



12

Replacement of the 4′-hydroxymethyl moiety by a methyl group, as in 3′-*O*-acetyl-2′,5′-dideoxyuridine (**10x**) or the corresponding iodo derivative **10t,** leads to a profound change. Thus, δ-H_6 in the latter is close to that observed for the same (pyrimidine) proton in 1-(2,3-dideoxy-3,4-didehydro-β-D-erythrofuranosyl)uracil (**2f**). It is not likely that the H_6 of **10x** is significantly influenced by the presence of the 3′-*O*-acetyl group.[24] The observed δ-H_6 shift accompanying such structural modification can readily be accounted for in terms of the absence of $O_{5'}$ which is necessary for a maximum deshielding effect (structure **11**).

The H_6 chemical shift of ethyl 3′-*O*-methylsulfonylthymidineuronate (**10b**) lies appreciably upfield from that expected of this proton in a pyrimidine nucleoside existing in anti conformation.[25] The differences are even more pronounced where this comparison is extended to certain *O*-acyl- (**10g** and **10x**), *O*-acyldeoxyiodo- (**10h, 10j, 10s,** and **10l**), trideoxydiiodo- (**10m, 10t**), and di-*O*-methylsulfonyl derivatives. The assignment of a syn conformation to this group of compounds would, therefore, seem attractive. However, as in the case of the tetrahydrofuryl derivative **10a,** the CD spectra of dimesyl derivative **10f** show a positive Cotton effect comparable to that of thymidine (**10e**)[19] in both chloroform (Figure 2) and water. Our observations do not lend persuasion to an explanation based on differences in inductive effects of the $C_{4'}$ and $O_{5'}$ substituents

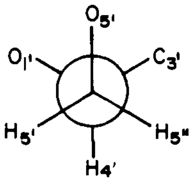
Table II. H_5 Chemical Shifts of Some Uracil Derivatives[a]

Compound	Solvent	H_5 (δ)	Line width[b]
Uracil (1c)	D_2O	6.24	2.5
1-(D-2,3-Dihydrofuryl)uracil (2f)	$CDCl_3$	5.80	3.0
Ethyl 2',3'-dideoxy-3',4'-didehydrouridine-5'-uronate (2g)	$CDCl_3$	5.77	3.5
2',3'-Dideoxy-3',4'-didehydrouridine (2h)	CD_3COCD_3	5.68	2.5
2-*O*-Acetyl-3,6-anhydro-1-deoxy-4,5-*O*-isopropylidene(uracil-1-yl)-D-mannitol (5b)	$CDCl_3$	5.67	
2',3'-*O*-Isopropylideneuridine (6a)	CD_3COCD_3	5.60	2.0
	Pyridine-d_5	5.70	3.0
5'-*O*-Acetyl-2',3'-*O*-isopropylideneuridine (6b)	$CDCl_3$	5.70	*c*
5'-Deoxy-5'-iodo-2',3'-*O*-isopropylideneuridine (6c)	$CDCl_3$	5.77	3.0
	Pyridine-d_5	5.53	3.0
1-(β-D-Arabinofuranosyl)uracil (7a)	D_2O	6.31	3.0
	Pyridine-d_5	5.47	5.5
1-(3,5-Di-*O*-acetyl-β-D-arabinofuranosyl)uracil (7b)	$CDCl_3$	5.55	
1-(3,5-Di-*O*-trityl-β-D-arabinofuranosyl)uracil (7c)	$CDCl_3$	5.44	
1-(α-D-Ribofuranosyl)uracil (8)	D_2O	6.31	3.0
	Pyridine-d_5	5.52	2.5
1-(3-Deoxy-3-iodo-2,5-di-*O*-trityl-β-D-xylofuranosyl)uracil (9)	$CDCl_3$	5.66	
Uridine (10n)	D_2O	6.24	*c*
2'-Deoxyuridine (10o)	D_2O	6.29[d]	5.0
5'-*O*-Trityluridine (10p)	$CDCl_3$	5.37	5.0
Ethyl 2'-deoxy-3'-*O*-methylsulfonyluridine-5'-uronate (10q)	$CDCl_3$	5.78	4.0
5'-Deoxy-5'-iodouridine (10r)	Pyridine-d_5	5.55[e]	
2',3'-Di-*O*-acetyl-5'-deoxy-5'-iodouridine (10s)	$CDCl_3$	5.85	
3',5'-Diiodo-2',3',5'-trideoxyuridine (10t)	$CDCl_3$	5.80	
2',5'-Di-*O*-trityluridine (10u)	$CDCl_3$	5.10	
1-(3-Deoxy-3-iodo-2,5-di-*O*-trityl-β-D-ribofuranosyl)uracil (10v)	$CDCl_3$	5.07	
3'-*O*-Acetyl-2',5'-dideoxy-2',5'-diiodouridine (10w)	$CDCl_3$	5.88	
3'-*O*-Acetyl-2',5'-dideoxyuridine (10x)	$CDCl_3$	5.81	
3',5'-Di-*O*-trityluridine (10y)	$CDCl_3$	5.25	

[a] For references, see Table I. [b] See Table I, footnote *a*. [c] Overlapped with $H_{1'}$. [d] Poor resolution. [e] Recalculated from the value 5.85 reported[36] for an internal Me_4Si.

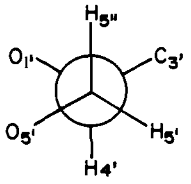
which may influence the electron density on $O_{5'}$ and hence the H_6 chemical shift. Thus, compounds **10c, 10d,** and **10i** do not exhibit the significant differences in H_6 chemical shifts anticipated for nucleosides lacking either $O_{1'}$ or $O_{5'}$ or both (vide supra) though the differences in inductive effects of $C_{4'}$ substituents (CH_2 vs. CO) are indeed substantial. Unlike compounds **10c** and **10d** the H_6 chemical shift of **10i** was measured in CD_3COCD_3. However, the trend of H_6 chemical shift is the same in this solvent as in $CDCl_3$ (cf. derivatives **1a** and **2e**) and thus the comparison remains valid. Furthermore, both compounds **10c** and **10d,** on one hand, and the diacetyl derivative **10g,** on the other, contain a carbonyl group next to $O_{5'}$. The inductive influences of these groups on $O_{5'}$ and thus on H_6 would be expected to be of a similar magnitude. The differences in H_6 chemical shifts are obviously much greater (cf. Table I). Therefore, it would appear that conformational factors involving the $C_{4'}$-$C_{5'}$ and $C_{5'}$-$O_{5'}$ bonds outweigh those resulting from an inductive contribution. An inspection of framework molecular models shows the conformation at the $C_{4'}$-$C_{5'}$ bond which would bring $O_{5'}$ in close proximity to H_6 is a gauche-gauche conformer (g,g, as indicated by the corresponding Newman projection formula **13**). The remaining two conformations, gauche-trans (g,t, formula **14**) and trans-gauche (t,g, formula **15**), are less favorable because the distance between H_6 and $O_{5'}$ in both is increased considerably. Thus, it seems probable that the g,g conformation at the $C_{4'}$-$C_{5'}$ bond is a sine qua non for positioning the H_6 in the electrostatic field of $O_{5'}$ and thereby ensuring the close proximity of both $O_{5'}$ and $O_{1'}$ to H_6 as indicated in formula **11.**

It is possible that the conformation at $C_{5'}$-$O_{5'}$ also plays a role in determining the distance between $O_{5'}$ and H_6. It may, of course, be argued that the latter would be the same in each of the three possible conformations as indicated by projection formulas **16–18.** Nevertheless, it is conceivable that the relative populations of $C_{5'}$-$O_{5'}$ conformers may, in turn, influence those comprising the $C_{4'}$-$C_{5'}$ conformers and thus affect δ-H_6. The decrease or lack of deshielding in compounds **10g, 10f, 10h, 10l, 10m, 10s, 10t,** and **10x** may then be viewed as a consequence of a limited ability to attain a structure comparable to **11** because of a departure of the $C_{4'}$-$C_{5'}$ bond from the usual g,g conformation. Additional studies of coupling constants of $H_{4'}$ and $H_{5'}$ would be necessary to clarify this point. However, it is of interest to note that a g,t conformation of the $C_{4'}$-$C_{5'}$ bond coupled



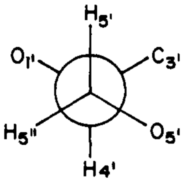
g,g

13



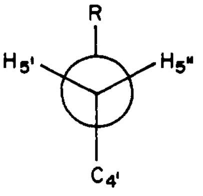
g,t

14



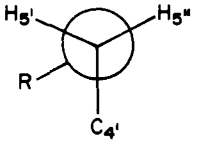
t,g

15



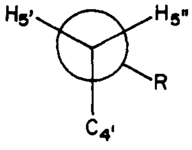
g,g

16



g,t

17



t,g

18

Formula 16–18: R = H, CH_3CO or CH_3SO_2

with a syn orientation of the base has been observed with crystalline 2′-deoxy-3′,5′-di-*O*-acetyl-5-fluorouridine.[26]

The situation is further complicated by the fact that the H_6–$O_{1'}$ and H_6–$O_{5'}$ distances can both be appreciably influenced by changes in the puckering of ribofuranose moiety. Moreover, the latter can also effect the rotameric composition at $C_{4'}$–$C_{5'}$. These factors would account for differences between the H_6 chemical shifts of uridine (**10n**) and 2′,3′-*O*-isopropylideneuridine (**6a**) observed in pyridine or between 5′-deoxy-5′-iodouridine (**10r**) and the corresponding 2′,3′-*O*-isopropylidene derivative **6c.** In view of the considerations cited above, it is not surprising that H_6 in 5′-*O*-acetyl-2′,3′-*O*-isopropylideneuridine (**6b**) is shifted further upfield and lies in the range of that of the 5′-iodo derivative **6c.** The latter shows a positive Cotton effect both in water[19] and $CHCl_3$ and as such is at variance with a view that compound **6c** has a syn conformation.

Comparison of the 3′,5′-di-*O*-acetyl derivative of thymidine (**10g**) and the carbocyclic analog **4b** reveals that H_6 is appreciably more deshielded in the former. This is in agreement with the results obtained with thymidine (**10e**) and its carbocyclic analog **4a** in D_2O and pyridine-d_5. On the other hand, the difference in δ-H_6 between diiodo derivative **10m** and compound **10g** is smaller and indicates again the relative unimportance of inductive effect.

The chemical shifts of trityl derivatives **3b, 7c, 10p, 10u, 10v, 10y,** and **10z** are more difficult to interpret because of the possibility of shielding by diamagnetic phenyl groups. In one report[27] a difference in H_6 chemical shift of some pyrimidine 3′-*O*- or 5′-*O*-tritylnucleosides in CD_3OD was noted but no explanation was offered. Thus, in both thymidine derivatives **3b** and **10z** there is significant departure of the value of H_6 chemical shift from that expected for a compound existing in anti conformation and incorporating both $O_{1'}$ and $O_{5'}$ atoms. On the other hand, the chemical shift of H_6 in 5′-*O*-trityluridine (**10p**) corresponds well to the "expected" value and differs from that found, e.g., in the 5′-*O*-acetyl derivative **10g.** The ditritylated compounds **7c, 10u,** and **10y,** however, exhibit a substantial shielding of H_6. Whether this is caused by changes in ribose puckering or $C_{4'}$–$C_{5'}$ conformational change due to the accumulation of bulky trityl groups in the molecule remains a mute point. As noted above, the diamagnetic shielding by phenyl residues may also influence the value of H_6 chemical shift.

Conclusion

The results reported herein identify two ranges of δ-H_6 values in pyrimidine nucleosides: one corresponding to the δ-H_6 in bases such as uracil and 1,3-dimethylthymine and the other corresponding to those of unsubstituted nucleosides. Moreover, δ-H_6 is subject to marked solvent effects. Thus, the observed differences in δ-H_6 are significantly larger in certain nonhydroxylic solvents than in water.

It is apparent that $O_{1'}$ and $O_{5'}$ both contribute to the deshielding of H_6. However, δ-H_6 is also sensitive to modification of the carbohydrate moiety. Among the unsaturated nucleosides, the position of the olefinic linkage is seen to have a profound effect on δ-H_6. Esterification of the OH groups in certain nucleosides, particularly 5′-deoxy-5′-iodo and 3′,5′-di-*O*-methylsulfonyl derivatives, leads to a prominent upfield shift. The possibility that these effects can be ascribed to an anti → syn conformational shift is deemed unlikely on the basis of CD data.

Experimental Section

The NMR spectra were measured on a Varian A-60A apparatus. The results are summarized in Tables I and II. CD spectra were obtained using a JASCO optical rotatory dispersion recorder, Model ORD/UV-5 in a CD modification SS-10 (Sproul Scientific, Boulder Creek, Calif.) between 500 and 200 nm. The CD data were digitized by hand after a smooth curve had been drawn through the data. The results were plotted as molar ellipticities $[\theta]$ against the wavelength (Figures 1 and 2). Starting materials were either commercial products or were prepared by conventional procedures. Diacetyl derivative **4b** was obtained by acetylation of **4a** (15 mg) with acetic anhydride (0.2 ml) in pyridine (0.1 ml) for 2 days at room temperature. The crude product was purified by preparative thin-layer chromatography on a 2-mm thick, loose layer of silica gel (5 × 20 cm) (70–325 mesh ASTM, Merck, Darmstadt, Germany) containing 1% of fluorescent indicator (Leuchtpigment ZS Super, Riedel-DeHaën, Hannover, Germany) in chloroform–methanol (9:1).[28] Evaporation of the eluate of the main uv-absorbing band gave an amorphous **4b:** NMR ($CDCl_3$, Me_4Si) δ 6.93 (d, 1, H_6), 5.48 (m, 1, "anomeric" H), 2.04 (s, 6, CH_3 of acetyl), 1.93 (d, 3, CH_3 of thymine).

Acknowledgments. The authors are indebted to Mr. N. Cvetkov and Mr. D. Marks for measuring the NMR and CD spectra and gratefully acknowledge the generous samples of compounds **2a** and **4a** provided by Dr. K. C. Murdock, American Cyanamide Co., Pearl River, N.Y. Our thanks are also due to Dr. J. Pitha, Gerontology Research Center, National Institute of Child Health and Human Development, Baltimore City Hospitals, Baltimore, Md., for information on ir spectral measurements of compound **3a.** The authors are also indebted to Dr. H. Follmann, Department of Chemistry, University of Marburg/Lahn, Germany, for a preprint of his publication (ref 23) and Dr. J. V. Freisler for the preparation of some derivatives used in the present study.

This investigation was supported in part by U.S. Public Health Service Grant No. FR-5529, GM 19108, and CA 13858 from the National Cancer Institute and in part by an institutional grant to the Michigan Cancer Foundation from the United Foundation of Greater Detroit.

References and Notes

(1) Presented in part at the 161st National Meeting of the American Chemical Society, Los Angeles, Calif., April 1971, CARB 035.
(2) J. Donohue and K. N. Trueblood, *J. Mol. Biol.*, **2**, 363 (1960).
(3) A. E. V. Haschemeyer and A. Rich, *J. Mol. Biol.*, **27**, 369 (1967).
(4) F. E. Hruska, *J. Am. Chem. Soc.*, **93**, 1795 (1971).
(5) M. P. Schweitzer, J. T. Witkowski, and R. K. Robins, *J. Am. Chem. Soc.*, **93**, 277 (1971).
(6) F. E. Hruska, *Can. J. Chem.*, **49**, 2111 (1971).
(7) F. E. Hruska, K. K. Ogilvie, A. A. Smith, and H. Wayborn, *Can. J. Chem.*, **49**, 2449 (1971).
(8) For a review, see G. E. Bachers and T. Schaefer, *Chem. Rev.*, **71**, 617 (1971).
(9) R. E. Schirmer, J. P. Davis, J. H. Noggle, and P. A. Hart, *J. Am. Chem. Soc.*, **94**, 256 (1972).
(10) (a) P. A. Hart and J. P. Davis, *J. Am. Chem. Soc.*, **93**, 751 (1971); (b) *ibid.*, **94**, 2572 (1972).
(11) J. P. Davis, *Tetrahedron*, **28**, 1155 (1972).
(12) J. P. Davis and P. A. Hart, *Tetrahedron*, **28**, 2883 (1972).
(13) H. J. Fritz, R. Machat, and R. R. Schmidt, *Chem. Ber.*, **105**, 642 (1972).
(14) J. H. Prestegard and S. I. Chan, *J. Am. Chem. Soc.*, **91**, 2843 (1969).
(15) A. A. Grey, I. C. P. Smith, and F. E. Hruska, *J. Am. Chem. Soc.*, **93**, 1765 (1971).
(16) C. Formoso, *Biochemistry*, **11**, 4031 (1972).
(17) H. Follmann, 9th International Symposium on Chemistry of Natural Products, Ottawa, Canada, June 1974, Abstracts, p 32B.
(18) The structure of **10c** has not been rigorously proven although all available evidence points to a β-D epimer: J. Žemlička, J. V. Freisler, R. Gasser, and J. P. Horwitz, *J. Org. Chem.*, **38**, 990 (1973).
(19) 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).
(20) (a) J. Pitha, personal communication; (b) *Biochemistry*, **9**, 3678 (1970).
(21) The 4′-hydroxymethyl group cannot impede rotation of the base in **3a** unless $C_{5'}$ and N_1 occupy a pseudoaxial position with both $C_{1'}$ and $C_{4'}$ eclipsed above the plane defined by $O_{1'}$, $C_{2'}$, and $C_{3'}$. In this connection, it is recognized that CD (as uv) spectra reflect an excited state of the molecule and further that conformations in normal and excited state need not necessarily be the same.[22]
(22) J. Žemlička, R. Gasser, and J. P. Horwitz, *J. Am. Chem. Soc.*, **92**, 4744 (1970).
(23) H. Follmann and G. Gremels, *Eur. J. Biochem.*, **47**, 187 (1974).
(24) Note, for example, very small or essentially no difference of H_6 in isomeric 2′- and 3′-*O*-acyluridines found in both D_2O and CD_3COCD_3: (a) H. P. M. Fromageot, B. E. Griffin, C. B. Reese, J. E. Sulston, and D. H. Trentham, *Tetrahedron*, **22**, 705 (1966); (b) S. Chládek and J. Žemlička, *J. Org. Chem.*, **39**, 2187 (1974).

(25) A similar anomaly has been noted for methyl 3'-*O*-methylsulfonyl-*N*-methylthymidine-5'-uronate.[18] Contrastingly, such a difference is not seen in the corresponding uridine derivative **10q**.

(26) W. Saenger and D. Suck, *Nature (London)*, **242**, 610 (1973).

(27) W. Regel, E. Stengele, and H. Seliger, *Chem. Ber.*, **107**, 611 (1974.)

(28) J. Žemlička, R. Gasser, J. V. Freisler, and J. P. Horwitz, *J. Am. Chem. Soc.*, **94**, 3213 (1972).

(29) J. Defaye, M. Naumberg, and T. Reyners, *J. Heterocycl. Chem.*, **6**, 229 (1969).

(30) K. C. Murdock and R. B. Angier, *J. Am. Chem. Soc.*, **84**, 3758 (1962).

(31) J. P. Horwitz, J. Chua, M. A. DaRooge, M. Noel, and I. L. Klundt, *J. Org. Chem.*, **31**, 205 (1966).

(32) G. Giovanetti, L. Nobile, M. Amorosa, and J. Defaye, *Carbohydr. Res.*, **21**, 320 (1972).

(33) D. M. Brown, A. R. Todd, and S. Varadarajan, *J. Chem. Soc.*, 2388 (1956).

(34) D. M. Brown, A. R. Todd, and S. Varadarajan, *J. Chem. Soc.*, 868 (1957).

(35) J. P. H. Verheyden and J. G. Moffatt, *J. Org. Chem.*, **35**, 2319 (1970).

(36) J. P. H. Verheyden and J. G. Moffatt, *J. Org. Chem.*, **35**, 2868 (1970).

(37) A. F. Cook and J. G. Moffatt, *J. Am. Chem. Soc.*, **89**, 2697 (1967).

(38) A. M. Michelson and A. R. Todd, *J. Chem. Soc.*, 951 (1953).

(39) R. E. Beiz and D. M. Visser, *J. Am. Chem. Soc.*, **77**, 736 (1955).

(40) P. A. Levene and R. S. Tipson, *J. Biol. Chem.*, **104**, 385 (1934).

Synthetic Spectroscopic Models. Intramolecular Stacking Interactions between Indole and Connected Nucleic Acid Bases. Hypochromism and Fluorescence[1-3]

Kiyoshi Mutai, Bruce A. Gruber, and Nelson J. Leonard*

Contribution from the Roger Adams Laboratory, School of Chemical Sciences, University of Illinois, Urbana, Illinois 61801.

Received January 27, 1975

Abstract: Stacking interactions between indole, as a neutral representative of tryptophan, and the nucleic acid bases have been observed in aqueous solution at 25° by means of hypochromism and fluorescence emission. This was accomplished by synthesizing and utilizing compounds in which indole and the nucleic acid bases adenine, cytosine, guanine, and thymine are connected by a three-atom or four-atom bridge, particularly the trimethylene bridge. The degree of interaction between indole and the purine bases was found to be of the same order as that between two purine bases themselves. For the Ind-$(CH_2)_3$-Base models which allow plane parallelism of the two units, the percentage of internally stacked vs. unfolded conformations was determined from fluorescence quantum yield and lifetime measurements, which gave a decreasing order of complexation with indole of adenine $\approx$ guanine $>$ thymine $\gg$ cytosine. The equilibrium between stacked and unfolded conformations for the indole/adenine, guanine, or thymine cases indicates ΔG near zero. On the basis of our results, total fluorescence quenching of the indole of tryptophan in a polypeptide or protein is to be expected if it comes into close proximity with a base moiety of a nucleic acid or if intercalation occurs.

The binding of proteins to nucleic acids involves electrostatic forces, hydrogen bonding, and π-overlap or stacking interactions,[4-6] all of which depend upon the accommodating sizes, shapes, and spacings of the interacting units. Among the specific stacking interactions[3,4,7-18] which may contribute to the positioning of protein with respect to nucleic acid, that of tryptophan or related indolic compounds with nucleic acid bases has been demonstrated (1) by complexation of the indole derivative with DNA, RNA, or poly A,[19-23] (2) by the quenching of tryptophan fluorescence in the binding of aminoacyl-tRNA synthetases and tRNA's,[24-28] (3) by ^1H NMR studies of aqueous solutions, especially acidic solutions, of tryptophan and other indole derivatives with nucleic acid bases,[29,30] and (4) by reflectance and luminescence studies of complex formation between tryptophan and nucleic acid components in aggregates formed in frozen aqueous solutions.[31-33]

Although the accumulated information is impressive, we sought to avoid certain of the limitations inherent in each set of experiments (and no doubt substituting different limitations of our own) by selecting suitable spectroscopic models for the observation of stacking interactions between indole (as an uncharged tryptophan) and the nucleic acid bases in dilute, neutral aqueous solution. We therefore chose a system that would permit intramolecular stacking, but not hydrogen-bonding, interactions which would be detectable by both ultraviolet and fluorescence spectroscopy. In the past, we have used polymethylene bridges, and in particular the trimethylene bridge, $-(CH_2)_3-$, as synthetic spacers to study intramolecular interactions between nucleic acid bases.[34] These bridges also provide the possibility of further controlling the inter-ring interactions by attachment of the chain to different positions on the heterocyclic termini.[3,16] Accordingly, we have synthesized compounds in which indole and the nucleic acid bases adenine, cytosine, guanine, and thymine are connected by a three-atom or four-atom bridge.[35] Attachments are at the 1 or 3 position of the indole and at the 9 position of adenine and guanine, the 1 position of cytosine and thymine, and also the N^6 position of adenine. The simple bases rather than the nucleosides were chosen so that we could survey the heteroaromatic interactions in the absence of additional factors involving the carbohydrate and phosphate linkages. With these models we could determine the degree of interaction between indole and nucleic acid base with respect to that between two nucleic acid bases, the degree of quenching of indole fluorescence by a nucleic acid base, and the equilibrium between folded or stacked conformations and open conformations.

Synthesis. General procedures for the linking of two different heterocyclic bases by a polymethylene bridge have been described previously,[34] including those for alkylation of adenine at the 9 position and cytosine and thymine at the 1 position. We have adapted these procedures by first preparing 3-(indol-3-yl)propyl and 4-(indol-3-yl)butyl bromides (**3a,b**) from the corresponding acids **1a,b** by reduction to the alcohols **2a,b** and displacement and then using the bromides to prepare the corresponding (indol-3-yl)alkyladenine, -cytosine, and -thymine products (**4–6**, Scheme I). While the alkylation of thymine may lead to mixtures of