# Effects of Methylation on the Stability of Nucleic Acid Conformations: Studies at the Monomer Level†

James D. Engel and Peter H. von Hippel*

ABSTRACT: High-resolution proton magnetic resonance measurements are reported on mono- and dimethylated analogs of adenine, cytosine, and guanine. These measurements demonstrate (for the adenine derivatives) and confirm (for the analogs of cytosine) that rotation of the aminomethyl group is restricted about the nitrogen–ring carbon bond, and that this rotation is characterized by activation enthalpies of +11 to +18 kcal/mol. The two possible rotamers of $m^6A$ and of $m^4C$ lie in the plane of the ring, and show an ~20:1 preference for syn positioning of the methyl group (relative to $N_1$ of $m^6A$ and $N_3$ of $m^4C$), resulting in both cases in preferential interference of the methyl group with Watson–Crick interbase hydrogen bonding. Rotation about the $C_2$–$N_{10}$ bond in $m^2G$ and $m_2^2G$ does not appear to be restricted. Measurements of equilibrium constants ($K$) for cyclic hydrogen-bonded dimer formation between complementary nucleic acid base analogs in organic solvents support these conclusions. Thus, while the equilibrium constants for A·U dimer formation involving the Watson–Crick ($N_1$) and Hoogsteen ($N_7$) sites of adenine are approximately equal, for $m^6A$·U the ratio of $K$ values for $N_7$- and $N_1$-type dimer formation is ~11:1. In G·C dimers, $K$ is reduced *ca.* eightfold for G·$m^4C$ relative to G·C (no "Hoogsteen" pairing is possible for G·C pairs), while the equilibrium constants for G·C and $m^2G$·C dimer formation are approximately equal. These results indicate that the replacement of A or C by $m^6A$ or $m^4C$ in a base-paired polynucleotide structure should destabilize the Watson–Crick double-helical structure by ~1.0–1.8 kcal/mol of methyl substituent. This prediction is consistent with polymer data in the literature, based on melting point depression measurements of ordered structures. The hydrogen bonding studies and the restricted rotation results suggest that little or no helix destabilization should result from the replacement of G by $m^2G$. However, this finding is not supported by the available polynucleotide results, suggesting that despite the favorable situation at the methylated monomer level, the methyl group does not fit satisfactorily into the small groove of the double-helical structure. In the Discussion our results are also utilized to develop simple (and testable) models to explain how methylation of nucleic acid bases might play a role in controlling and stabilizing the active structures of tRNA and ribosomes, as well as to suggest ways that such methylation might help to account for the recognition by the *Escherichia coli* modification–restriction enzymes of the specific DNA base sequences involved in the function of these systems.

In recent years the importance of biological methylation has become increasingly obvious to workers in several related fields. In nucleic acid research, it is now well known that specifically placed methylated bases are found in all tRNA sequences (*e.g.*, see Nishimura, 1972) and these methyl groups have been invoked as contributors to the establishment of the biologically active tertiary conformation of the tRNA molecules. For example, Saneyoshi *et al.* (1969) have suggested that specific alkylation (including methylation) may contribute significantly to the stability of the anti-codon loop of tRNA, since a modified base is always found adjacent to the anti-codon triplet. rRNA has also been found to be specifically methylated (Klagsbrun, 1973), suggesting that methyl groups may also play a role in the structure and interactions of the various protein and nucleic acid components involved in establishing the functional ribosome. Yet in all these studies no concrete molecular mechanism has been put forward to explain how such methylation might actually affect nucleic acid conformation.

Studies of modification and restriction systems (for a recent review see Meselson *et al.*, 1972) have shown that specific nucleotide sequences in bacterial DNA are methylated by their respective host-specificity systems; this methylation protects the genome from endonucleolytic cleavage by the restriction enzymes present in the cell. However if, for example, DNA methylated in an *Escherichia coli* K cell is introduced into a B cell, the K-specific methylation pattern does not protect the foreign DNA from endonucleolytic cleavage (followed by rapid exonucleolytic hydrolysis) by the B-specific restriction enzyme system. It appears that these enzymes recognize specific double-stranded nucleotide sequences in DNA, and the nature of the reaction with these sequences is determined by the presence or absence of methyl groups occurring as monomethyl substituents of the exocyclic amino groups of specific adenine residues (a maximum of two methyls per modification site; one on each strand). Thus these methyl groups must play a crucial role in controlling the specific DNA–protein recognitions and interactions involved in the modification–restriction system, and it is this supposition which prompted the present investigation.

Obviously methyl groups located on particular DNA bases can alter the specificity of DNA–protein interactions in two general (nonexclusive) ways: (i) either the protein, *via* appropriately placed functional groups, recognizes directly the substitution of a methyl for a hydrogen on the amino group of a particular adenine residue (in the large groove of the double helical DNA structure), independent of any secondary modifi-

† From the Institute of Molecular Biology and Department of Chemistry, University of Oregon, Eugene, Oregon 97403. *Received January 28, 1974.* A portion of these results was presented at the Biochemistry/Biophysics Meetings in Minneapolis, Minn., June 1974. This study will be presented (by J. D. E.) to the Graduate School of the University of Oregon in partial fulfillment of the requirement for the Ph.D. degree. It was supported in part by USPHS Research Grants GM-15792 and GM-15423, as well as by a traineeship (to J. D. E.) from USPHS Training Grant GM-00715. The nmr and computational equipment was provided in part by a USPHS Health Sciences Advancement Award to the University of Oregon, and by USPHS Research Grant GM-19286

cations of the double helix which this methylation may introduce: or, (ii) this substitution may have additional effects on the local structure and stability of the DNA, and it is these secondary changes (or their consequences) which are recognized by the protein. (For a general review of the structural basis of DNA–protein interaction specificity, see von Hippel and McGhee, 1972.) In this paper we will show, by nuclear magnetic resonance (nmr) studies of monomer structure and of cyclic hydrogen-bonded dimer formation in organic solvents, that monomethylation of the exocyclic amino groups of adenine and cytosine does indeed perturb the structure of the monomers, and consequently changes the specificity and stability of interbase hydrogen-bonding interactions.

## Experimental Section

*Materials.* 6-Methylamino-9-methyladenine ($m^6A$),[1] 6-dimethylamino-9-ethyladenine ($m_2^6A$), 9-cyclohexyladenine (A), 1-cyclohexyluracil (U), 2-methylaminoguanosine ($m^2G$), 2-dimethylaminoguanosine ($m_2^2G$), 1-ethylcytosine (C), and 2,4-dimethoxypyrimidine were purchased from Cyclo Chemicals. All moved as single chromatographic components on cellulose thin-layer chromatography (tlc) plates (Eastman Kodak Co.) in solvent systems A (95% ethanol–1.0 M ammonium acetate (pH 7.5), 75:30, v/v) and B (1-propanol–concentrated $NH_4OH$–$H_2O$, 60:30:10, v/v).

Methanol-$d_4$ was purchased from Prochem, Ltd., and benzene-$d_6$, $CDCl_3$, dimethyl-$d_6$ sulfoxide ($Me_2SO$), DCl, and methyl-$d_3$ alcohol from Merck, Sharpe and Dohme of Canada, Ltd. Hexamethyldisiloxane (HMD) was obtained from K and K Chemicals. Trifluoroacetic acid-*d* was purchased from NMR Specialties, Inc. Precision bore nmr tubes (Types 524-PP and 507-PP) used in these studies were purchased from Wilmad Glass Co., Inc.

*Synthesis.* 4-Methylamino-1-methylcytosine ($m^4C$) and 4-dimethylamino-1-methylcytosine ($m_2^4C$) were synthesized by the method of Kenner *et al.* (1955) from 2,4-dimethoxypyrimidine. Conversion of the common intermediate (4-methoxy-1-methyl-2-pyrimidinone) to the final products was carried out by heating the intermediate plus a saturated aqueous solution of methylamine or dimethylamine in a sealed Carius tube for 8 hr at 150°. The reaction mixtures were evaporated to dryness, and the products were recrystallized from EtOH–EtOAc.

*$m^4C$. Anal.* Calcd for $C_6H_9N_3O$: C, 51.79; H, 6.52; N, 30.20. Found: C, 51.51; H, 6.61; N, 30.32. Low resolution mass spectral results: parent ion at *m/e* 139. The corrected melting point of the crystalline material was 178–179°. Tlc shows a single component (uv) in both solvent systems A ($R_F$ 0.84) and B ($R_F$ 0.79).

*$m_2^4C$. Anal.* Calcd for $C_7H_{11}N_3O$: C, 54.89; H, 7.24; N, 27.43. Found: C, 54.38; H, 7.35; N, 27.89. Low resolution mass spectral results: parent ion at *m/e* 153. Tlc shows one component (uv) with both solvent systems A ($R_F$ 0.76) and B ($R_F$ 0.78). The corrected melting point of the crystalline product was 178–179°.

*Nmr Measurements.* The spectrometer magnetic field was locked on an internal (solvent) deuterium reference frequency for all experiments except during temperature calibration, when a homonuclear resonance was used as the locking signal. The temperature was calibrated by the variation in the chemical shift of methanol in accord with the revised standards of Van Geet (1970), which were adapted to the 100-MHz spectrometer used in this study. Temperatures were found to be accurate to within ±1° when compared to the reading of a thermistor inserted directly into the nmr tube. The assignments of position and multiplicity for each peak are given in the figure legends. The positions are given in parts per million (ppm) downfield from HMD (≡ 0.0), and the peak multiplicities are denoted as follows: S = singlet, D = doublet, T = triplet, Qd = quadruplet, and Qn = quintuplet.

*Nmr Spectral Assignment.* The observed proton resonance peaks of the nucleoside bases and analogs used in these experiments were assigned primarily from published results (Ts'o *et al.*, 1969). The $C_2$ and $C_8$ protons of the adenine analogs were resolved by deuterium replacement of the slightly acidic $C_8$ proton, using the method of Chan *et al.* (1964). Spectral assignment of the methylamino resonances of $m^6A$ was facilitated by deuterium exchange of the amino hydrogens, which collapses the spin–spin coupling of the amino proton to the methyl group and thus establishes which resonances are due to the exocyclic methylamino group.

*Nmr Instrumentation.* Nmr spectra were recorded on a Varian XL-100 Fourier transform (FT) spectrometer. All spectra used for illustration as figures were recorded in the continuous wave mode. Rates of exchange in the systems under study (*i.e.*, the exocyclic methyl- or dimethylamino groups of adenine, cytosine, and guanine) were calculated from line-width measurements by the method of slow exchange broadening (Shoup *et al.*, 1972b) at negligible saturation, and activation enthalpies for the restricted rotation determined from Arrhenius plots of these data. Spin–lattice relaxation times ($T_1$) were recorded in the FT mode using a modification (all peaks plotted as positive and decaying exponentially with $\tau$, Freeman and Hill, 1972) of the original (. . .180°-$\tau$-90°. . .) pulse technique of Vold *et al.* (1968). The interferograms of the applied pulses were collected and Fourier transformed by a Varian 620/I computer equipped with either a 16K or 8K memory.

## Results

*Restricted Rotation and Rotamer Population Data.* A number of techniques have been used to demonstrate the specificity of interbase hydrogen-bond formation between nucleic acid monomers in organic solvents. Nuclear magnetic resonance has been one of the most powerful of these techniques, and, in parallel with infrared measurements, has been utilized to show, by direct observation of the participating protons, that hydrogen-bond formation in mixed monomer systems occurs predominantly between the "correct" Watson–Crick partners (*e.g.*, Katz and Penman, 1966; Shoup *et al.*, 1966; Hamlin *et al.*, 1966). At about the time these observations were made on hydrogen-bonded dimers, nmr studies of the conformation of nucleoside monomers were also under way. In 1967, Martin and Reese demonstrated restricted rotation of the amino group of dimethylaminoadenine and dimethylaminocytosine, and attributed this behavior to the partial double-bond character of the ring carbon–exocyclic nitrogen bond. Shoup *et al.* (1966, 1971) likewise demonstrated restricted rotation in cytosine by utilization of a $^{15}N$-containing amino group which permitted the visualization of the two amino protons as separate sets of peaks, and later (1972a) these same authors also provided evidence demonstrating restricted rotation in mono- and dimethylaminocytosine. Shoup *et al.* (1972a) and we (this paper) have observed restricted rotation in methylaminocytosine. In addition, in this paper we report results of experiments demonstrating restricted rotation of the amino group in $m^6A$.

[1] The abbreviations used are those recommended by the IUPAC–IUB Commission on Biochemical Nomenclature. A summary of these recommendations may be found in *J. Biol. Chem. 245*, 5171 (1971).

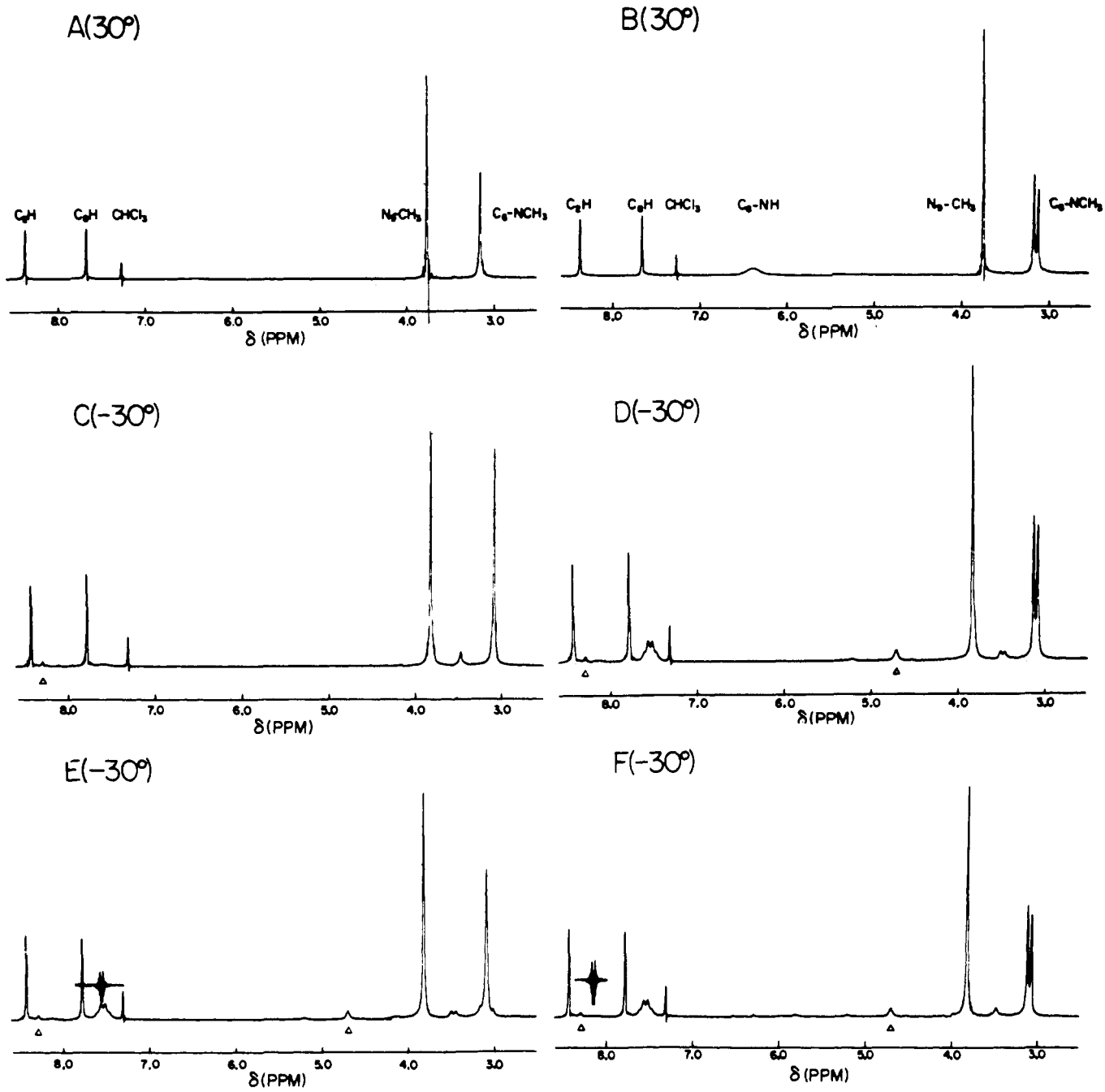
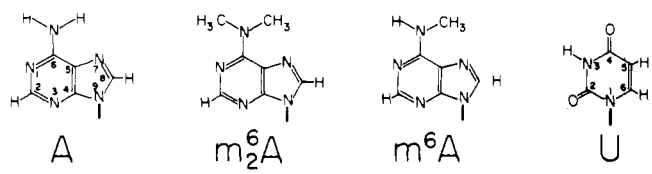



FIGURE 1: Spectra showing restricted rotation of $m^6A$ (0.2 M) at ambient (A, B; 30°) and low (C, D, E, F; −30°) temperatures in $CDCl_3$. In spectra A and C the amino proton had been replaced by a deuteron, as described in Results. Peak assignments for A are: $C_2$-H (8.38 ppm, S), $C_8$-H (7.65 ppm, S), $CHCl_3$ (7.27 ppm, S), $N_9$-methyl (3.76 ppm, S), $C_6$-methylamino (3.15 ppm, S). Irradiation centered over the amino resonances in spectra E and F indicate the frequencies of the saturating ($H_1$) field used to selectively decouple the N-methyl low temperature rotamers (see Results). Impurities are signified by Δ.




*$m^6A$.* The spectra of Figure 1 show the peak positions and shapes corresponding to the various proton resonances of $m^6A$ in $CDCl_3$, as well as the changes observed in lowering the temperature of this system. $m^6A$ was dissolved in $CD_3OD$ for amino group deuteration, and, after drying, the deuterio-$m^6A$ was redissolved in $CDCl_3$. This procedure effectively eliminated competitive D ⇄ H exchange with the solvent. Figure 1A shows a typical nmr spectrum of such a deuterium-exchanged sample of $m^6A$ at ambient temperature. Figure 1B shows the sharply defined doublet obtained for the methylamino group when the undeuterated compound is dissolved directly into $CDCl_3$.

At low temperature, the deuterated $m^6A$ spectrum shows two singlet peaks characteristic of chemical exchange[2] (Figure 1C) at 3.07 and 3.51 ppm, due to the two possible methyl ro-

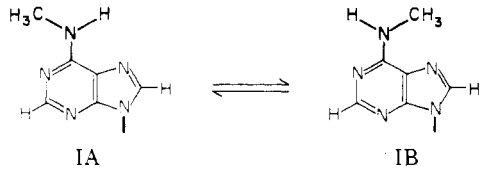
[2] A chemical exchange system is one in which an observable moiety, here a proton or methyl group, is exchanged between two magnetically nonequivalent sites. In this case the exchange is brought about by rotation about the $C_6$-$N_{10}$ bond of adenine. Under slow exchange conditions, two peaks are seen at frequencies characteristic of their respective magnetic environments, and the ratio of the peak heights reflects the relative equilibrium distribution (*i.e.*, populations) of the two sites.

TABLE I: Activation Parameters for Restricted Rotation.

| Compd | Derivative | Solvent | Equilibrium Population Ratio (±0.5) | $\Delta H^{*}$ [a] (kcal/mol) | $\Delta\nu_0$ [b] (ppm) | $T_c$ (°C) | $k_{298}$ ($sec^{-1}$) | Ref |
|---|---|---|---|---|---|---|---|---|
| $m^6A$ | 9-Methyl | $CD_3OD$ | 96:4 | 11.3 | 0.445 | 0 | 53. | This paper |
| | 9-H | $CD_3OD$:TFA-$d$ [d] (50:1) | 88:12 | 12.3 | 0.345 | 20 | 8.7 | This paper |
| $m_2{}^6A$ | 9-Ethyl | $CD_3OD$ | 50:50 | 15.4 | 0.440 | −20 | 20,000 | This paper |
| | 9-H | $CD_3OD$:TFA-$d$ [d] (50:1) | 50:50 | 13.0 | 0.412 | 0 | 265. | This paper |
| | 9-Isopropylidine | $CDCl_3$ | 50:50 | 11.8 | 0.470 | 0 | N.C. [c] | [e] |
| $m^4C$ | 1-Methyl | $CD_3OD$ | 95:5 | | | 0 | | This paper |
| | 1-Methyl | DMF-$d_7$ | 96:4 | | | | | [f] |
| $m_2{}^4C$ | 1-Methyl | $CDCl_3$ | 50:50 | 16.3 | 0.146 | 18 | 208. | This paper |
| | 1-Methyl | $CDCl_3$ | 50:50 | 17.6 | 0.143 | 10 | 90. | [f] |
| | 1-Isopropylidine | $CDCl_3$ | 50:50 | 8.6 | 0.120 | 30 | N.C. [c] | [e] |

[a] Our activation enthalpies and those given by Shoup *et al.* (1972a) were calculated by the method of Shoup *et al.* (1972b). Due to the much greater frequency difference in the two stereoisomer absorption frequencies exhibited in the adenines as compared to the cytosines and to the greater field strength of the spectrometer used in these experiments, we would estimate our precision of activation enthalpy measurements for $m_2{}^6A$ conservatively to ±10%, and for $m_2{}^4C$ and $m^6A$ to ±20%. [b] $\Delta\nu_0$ = the difference in resonance frequency of the stereoisomers of an exocyclic methylamino or dimethylamino nucleic acid derivative under conditions of very slow ($k \ll 1\ sec^{-1}$) exchange. [c] N.C. = not sufficient data for calculation. [d] TFA-$d$, trifluoroacetic acid-$d$. [e] Martin and Reese (1967). [f] Shoup *et al.* (1972a).

tamers depicted as IA and IB. The nondeuterated $m^6A$ spectrum (Figure 1D) shows two doublet peaks at the same chemical shift values, and of the same relative intensities, as observed for the singlets in the deuterated compound. These intensities are directly related to the rotamer population of the two spin states under slow exchange conditions, and from their relative heights we have obtained a population ratio of 96:4 (Table I).



IA IB

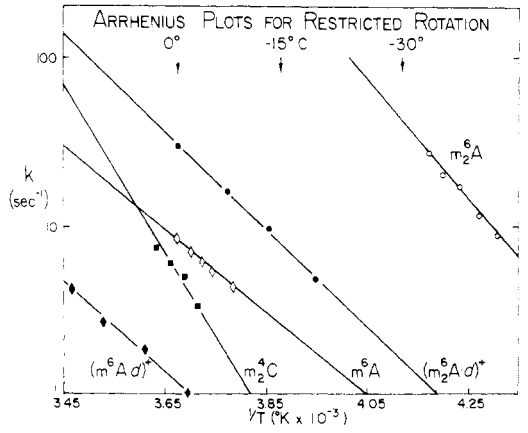



FIGURE 2: Arrhenius plots obtained for restricted rotation of the exocyclic (di)methylamino groups of acid-$m^6A$ [♦; methanol-$d_4$–DCl (50:1)], $m^6A$ [◇; methanol-$d_4$], acid-$m_2{}^6A$ [●; methanol-$d_4$–DCl (50:1)], $m_2{}^6A$ [○; methanol-$d_4$], and $m_2{}^4C$ [■; $CDCl_3$]. $k$ is the rate ($sec^{-1}$) at which a methyl group transfers between magnetic environments as measured by the increase in line widths (see Methods).

The doublets shown in Figure 1D can be collapsed separately (Figure 1E and F) by placing a strong irradiation field at the position of the amino proton resonance frequency, thus spin-

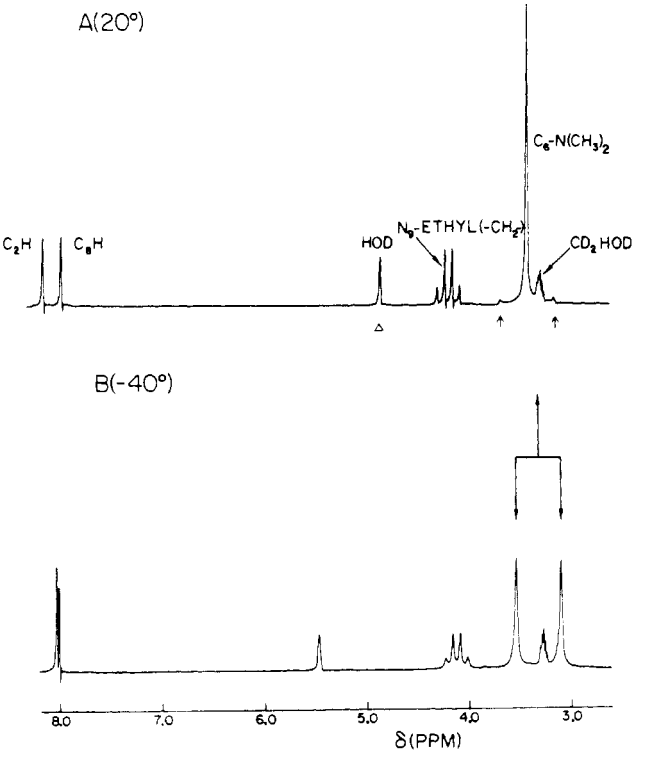



FIGURE 3: Spectra showing restricted rotation of $m_2{}^6A$ (0.2 M) obtained at 20° (A) and −40° (B) in $CD_3OD$. Peak assignments for A are: $C_2$–H (8.12 ppm, S), $C_8$–H (7.94 ppm, S), $C_9$–ethyl (methylene, 4.18 ppm, Qd), $C_6$–dimethyylamino (3.40 ppm, S), and $CD_2HOD$ (3.28 ppm, Qn). Spinning side bands are indicated by ↑, and impurities are noted by Δ.

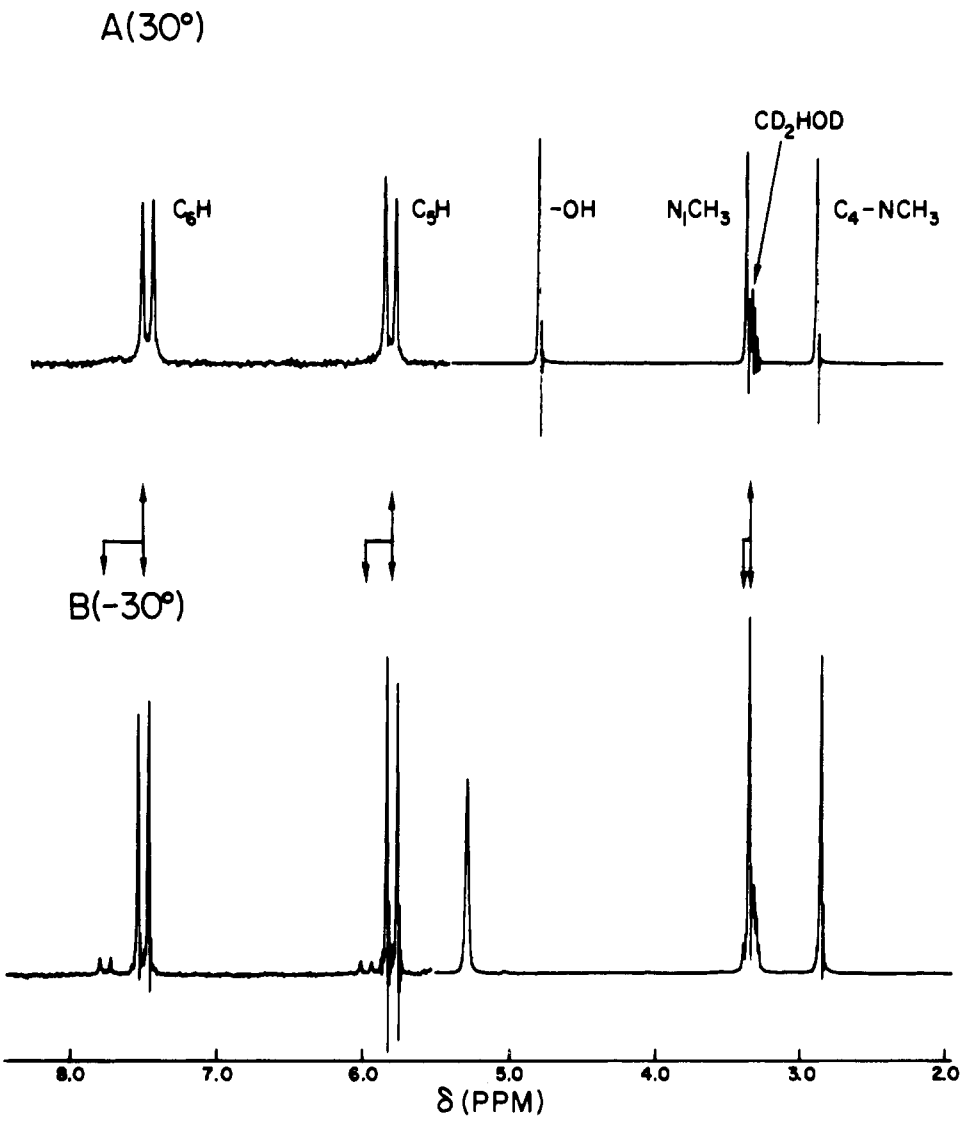



FIGURE 4: Spectra showing restricted rotation of $m^4C$ (0.2 M) at 30° (A) and −30° (B) in $CD_3OD$. Peak assignments for A are: $C_6$-H (7.38 ppm, D), $C_5$-H (5.73 ppm, D), $CD_3OH$ (4.76 ppm, S), $N_1$-methyl (3.29 ppm, S), $CD_2HOD$ (3.26 ppm, Qn), $C_4$-methylamino (2.81 ppm, S). The spectrum amplitude was increased (×10) from 5.5 ppm downfield to permit easier visualization of the lower field peaks.

spin decoupling one or the other of the isomers of the methyl group by saturation of the individual coupling protons. Although we cannot resolve the minor isomer amino coupling proton due to nitrogen quadrupole and consequent line broadening effects (Figure 1F), we can see the effect of saturating the spin levels by narrow-band irradiation (decoupling) at the position of this "virtual" peak, by observing the collapse to a singlet of the small isomer methyl spin–spin doublet. This establishes both the approximate position of the amino proton resonance of the minor isomer and the identity of the minor isomer as a methyl group resonance.

These spectra clearly indicate the presence of two distinct methylamino isomers (rotamers) at low temperature, which we attribute to positioning of the methyl group syn and anti to the $N_1$ of the purine ring (A and B, respectively). Steric considerations (see Discussion) led us to attribute the predominant isomer peak (at 3.07 ppm) to the methyl group in the syn position, and studies of interbase hydrogen-bonding interations between $m^6A$ and U (see below) confirm this assignment.[3] Arrhenius plots of ln $k$ (where $k$ is the rate of exchange of the methyl group from one magnetic environment to another) *vs.* $T^{-1}$ yield an activation enthalpy of +11.3 kcal/mol (Figure 2). The parameters we have measured which characterize the restricted rotation of exocyclic methylamino groups in the nucleic acid derivatives, as well as relevant literature values, are listed in Table I.

*$m_2^6A$.* Typical spectra, depicting the nmr behavior of $m_2^6A$ at ambient and low temperatures, are shown in Figure 3. Our measured activation enthalpy for this compound, together with that obtained by Martin and Reese (1967), is presented in Table I. We find that our value is approximately 4 kcal/mol higher, but since these workers did not record the methods and

[3] For later comparisons of these results with those obtained from base-pair formation and homopolymer melting point ($T_m$) depression data, we should point out that the equilibrium enthalpy change for $m^6A$, obtained by measuring the activation enthalpy for both the exchange rates from syn to anti and from anti to syn (I), is only +1.5 (±0.3) kcal/mol of methylated base. Furthermore, there is no measurable change in the population ratio between −30 and −60° in the restricted rotation results, confirming that the equilibrium $\Delta H°$ for the methyl isomerization process is very small. Thus, as a first approximation, temperature-dependent effects on rotamer equilibrium populations may be neglected, and results obtained at the various temperatures by the different techniques used may be directly compared.

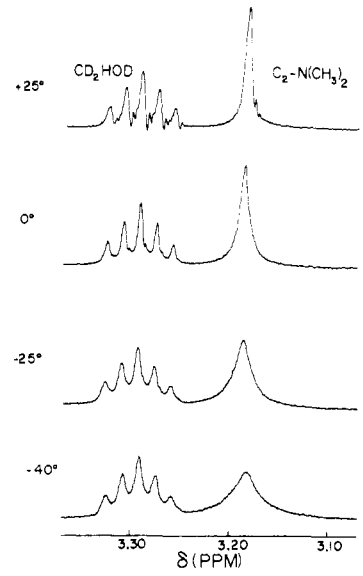
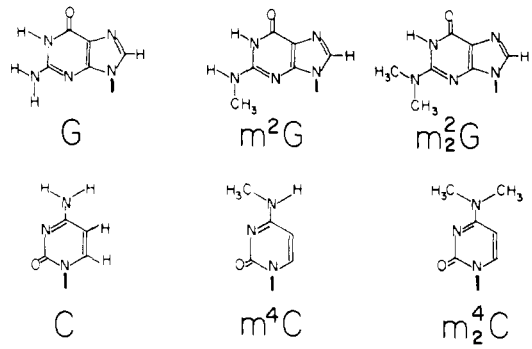



FIGURE 5: Spectra of $m_2^2G$ (0.1 M) at a variety of temperatures in $CD_3OD$:TFA-*d* (50:1). The dimethyl resonance is centered approximately 10.8 Hz upfield from the impurity in the solvent ($CD_2HOD$, Qn).

approximations used to obtain their value, we cannot comment on the possible sources of this discrepancy.

*$m^4C$.* The spectra shown in Figure 4 depict the resonance positions of the protons of $m^4C$. At low temperature (Figure 4B) the $C_5$ doublet resolves into two unequally populated sets




of peaks (as previously observed by Shoup *et al.*, 1972a), which, in agreement with these workers, we attribute to the two methyl isomers (population ratio 95:5; Table I). As expected on the basis of the $C_5$-proton resonances, the spectra of Figure 4 also show that, at low temperature, the $C_6$- and $N_1$-methyl protons are resolved into two sets of peaks with similar population ratios. Based on steric reasoning (see Discussion), and in agreement with the conclusions reached by Shoup *et al.* (1972a) on the basis of similar arguments, we assign the larger peaks to the syn methyl isomer.

Further inspection of Figure 4 reveals an apparent anomaly, since the $C_4$-methyl resonances themselves do not show the low temperature separation into two sets of peaks, even though such separation is a major feature of the $C_6$-methylamino resonance of $m^6A$, and would certainly be expected for $m^4C$ on the basis of the behavior of the $C_5$-, $C_6$- and $N_1$-methyl proton resonances. (The absence of separation of the $C_4$-methylamino resonances of $m^4C$ was confirmed by rerunning the spectrum in $CDCl_3$, which is completely featureless in the region of interest.) This observation led us to carry out a number of control experiments, seeking possible antifactual origins for the differential shifts of the $C_5$-, $C_6$- and $N_1$-methyl proton peaks into two sets at low temperature. These experiments, plus consideration of various studies in the literature, have led us to rule out all other interpretations of the double peaks, and to conclude that the failure to observe a separation of the $C_4$-methylamino resonance of $m^4C$ must be due to a fortuitous compensation and cancellation of magnetic effects at this particular locus.

*$m_2^4C$.* This compound was made to permit direct comparison with the work of Shoup *et al.* (1972a), and, in confirmation of their results (spectra not shown), we find that at low temperature the $C_4$-dimethylamino resonance separates into two equally populated peaks. The activation enthalpy we have determined for the restricted rotation about the $C_4$-$N_7$ bond from these data is shown in Table I and Figure 2, and differs from the value obtained by Shoup *et al.* (1972a) by ~8%. This comprises good agreement for such determinations, and shows not only that our measurements of the activation parameters for the restricted rotation of methylated adenines are consistent with values for related compounds in the literature, but also that our computations of activation energies (summarized in Table I) are consistent with those measured by the slow-exchange coalescence method by others. [A discussion of activation enthalpies measured by CW and FT techniques is presented by Johnson (1966), wherein he points out that such calculated parameters sometimes vary greatly depending on which method is used.]

*$m_2^2G$ and $m^2G$.* Spectra of $m_2^2G$ at several temperatures are shown in Figure 5. The $C_2$-dimethylamino resonance clearly shows an appreciable broadening with decreasing temperature, but two observations indicate that this broadening is artifactual in origin, or at least does not reflect the isomerization of the aminomethyl groups as a consequence of restricted rotation about the $C_2$-$N_{10}$ bond. First, Figure 5 shows that, in parallel to the broadening of the $C_2$-dimethylamino resonance, the $CD_2HOD$ (solvent impurity) peaks broaden similarly, suggesting that a good part of the effect must be due to increasing inhomogeneity with decreasing temperature. In addition, if any part of the broadening of the $C_2$-dimethylamino peak is *not* due to this artifact, it shows no signs of saturation over a 75° temperature range, indicating that, if an exchange process is actually involved, it is characterized by a much lower energy of activation than that observed in the $m^6A$, $m_2^6A$ and $m_2^4C$ spectra, which go from initial broadening to coalescence over a temperature range of less than 25°. The $m^2G$ spectra (not shown) behave identically. These observations suggest that $m^2G$ and $m_2^2G$ show no restricted rotation measurable on the nmr time scale.

It should be pointed out explicitly that these guanine derivatives were investigated in a solvent consisting of methanol-$d_4$ and trifluoroacetic acid-*d* (added to increase solubility). In this solvent these nucleosides are ionized, with a deuteron located at $N_7$ of the purine ring (Shapiro, 1968) and an overall positive charge on the aromatic system. Direct measurements could not be obtained on neutral $m_2^2G$ or $m^2G$, because the solubility of these compounds in appropriate organic solvents is too low to permit generation of clear spectra, even by collecting and summing many transients in the FT mode. As a related comparison, we have demonstrated that cationic $m^6A$ and $m_2^6A$, dissolved in the same (acid-methanol) solvent used for the guanine derivatives, show kinetic parameters for restricted rotation about the $C_6$-$N_{10}$ bond which are virtually unchanged from the values for the neutral species [*e.g.*, for acid-$m^6A$ in methanol-$d_4$-trifluoroacetic acid-*d* (50:1) we obtained: $\Delta H^* = 11.3$ kcal/mol, $\Delta\nu_0 = 34.5$ Hz, and $T_c = 20°$; compare the $m^6A$ data, Table I, and Figure 2]. Finally, the measurements of cyclic dimer formation reported in the next section do involve

neutral guanine monomers, and provide independent support for the suggestion that, unlike $m^6A$ and $m^4C$, the methyl group of $m^2G$ does *not* show restricted rotation *accompanied* by preferential syn orientation relative to the Watson-Crick hydrogen-bonding loci.

*Evidence from Hydrogen-Bonded Dimer (Base-Pair) Formation in Organic Solvents.* Information on the effect of exocyclic amino group methylation of adenine, cytosine, and guanine can also be obtained from mixing curves of these monomers with their complements to measure the affinity and specificity with which they form hydrogen-bonded base pairs.[4] The mixing curves we obtained closely resemble those described by Katz and Penman (1966), except that it is apparent that our mixing curves are *nonlinear* in proton chemical shift as a function of base added (see Figures 6–8). We have exploited this nonlinearity to obtain equilibrium constants for hydrogen-bonded dimer base-pair formation, using a modified form of the theoretical approach outlined by Katz (1969) and by Newmark and Cantor (1968), and thus have obtained independent information about preferential exocyclic methylamino group positioning as reflected in the altered base pairing affinities induced by these groups.

In its simplest form, a mixing or "continuous variation" experiment (Job, 1928) is carried out by keeping constant the *total* concentration of the two components (A and B) under consideration, while varying the molarities of the *individual* species over the total concentration range (*i.e.*, $M = [A_{tot}] + [B_{tot}]$; $M$ = constant). For an individual data point on a mixing curve, the total concentration of species A (which we set equal to $X$) is the sum of the concentration of free and complexed A, or $[A_{total}] = [A] + [AB] = X$. Therefore $M = X + [B] + [AB]$, and the fraction of A in complex ($f_{AB}$) is $[AB]/([A] + [AB])$. We assume initially (see Katz, 1969, and Nagel and Hanlon, 1972) that only trivial amounts of AA and BB dimers are formed under the conditions of the experiment. This restriction can be relaxed later, when we compute the final concentrations of all species. The reaction leading to complex formation is simply $A + B \rightleftarrows AB$, and is characterized by an association constant $K$, with $f_{AB} = K[B]/(1 + K[B])$. Following further substitution and expansion, we obtain

$$[AB] = \frac{[KM + 1] \pm \sqrt{(KM + 1)^2 - 4K^2(MX - X^2)}}{2K} \quad (1)$$

In this equation, the concentration of complex [AB] is expressed in terms of $K$, $M$, and $X$, with the latter two parameters given for a particular mixing curve data point.

Values of $K$ and [AB] are derived from the nmr mixing curves as follows: under conditions of fast exchange (which are met for all of these mixing experiments) only one resonance is observed for a given proton (*e.g.*, attached to species A), and the chemical shift of this resonance is thus given by $\delta_{A,obs} = \delta_A f_A + \delta_{AB} f_{AB}$, where $\delta_A$ and $\delta_{AB}$ are the chemical shifts of the A proton in the free and complexed condition, and $f_A$ and $f_{AB}$ are the fractional populations of these two states. For a mixing curve in which the concentration of A increases from left to right (*e.g.*, in Figure 6 in which we use guanosine and cytosine to illustrate the analysis, with guanosine as component A and cytosine as component B), the intercept of the mixing curve with the left-hand ordinate (*i.e.*, $\delta_{AB}$) represents the chemical shift of the A proton at infinite dilution of A in a solution which is $M$ (molar) in species B. The actual value of $\delta_{AB}$ will depend upon $K$, $M$, and $[A_{total}]$, and can be determined together with $K$ by an appropriate computer program (see below). The intercept with the left-hand ordinate for the "dilution curve" of A represents the chemical shift of the A proton at infinite dilution in the free form (defined as $\delta_A$). [As an example, consider the intercepts of the dashed (complexed) and solid (free) curves depicted in Figure 8A.]

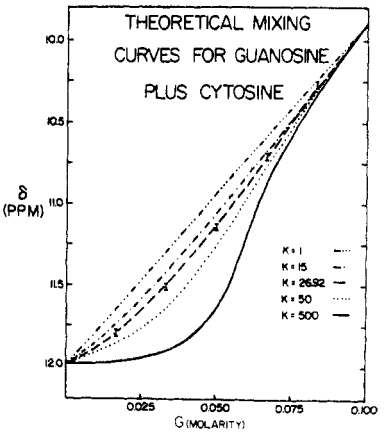



FIGURE 6: Theoretical mixing curves superimposed over the actual data (error bars) for guanosine and cytosine in $Me_2SO$-$d_6$:$CDCl_3$ (1:1) at 32° (±1°). All lines were plotted according to eq 1 (see Results).

For a given mixing curve, we begin interpretation by making a rough (linear) extrapolation of the experimental data to obtain an initial (trial) value of $\delta_{AB}$ (estimation of $\delta_A$ is straightforward, since it involves only linear extrapolation using points which here do fall on a straight line), and then by computer iteration we refine and resolve a best least-squares fit for $K$ to all of the original data (that is, we find the line derived from eq 1 which provides the least deviation from all data points for a given $K$). This trial value of $\delta_{AB}$ is refined by an iteration procedure involving the variation of $\delta_{AB}$ over a reasonable range around the initial estimate, and by obtaining a best-fit association constant for each assumed value of $\delta_{AB}$. The final value of $\delta_{AB}$ [that providing the best fit for all the experimental data to a single association constant chosen as the minimum in the least-squares deviation from eq 1], together with the final association constant, is obtained by appropriate computer routines.

This procedure is illustrated in Figure 6, in which the theoretical mixing curves are plotted for a dimerization (AB) system for values of $K$ ranging from 1 to 500 $M^{-1}$. Note that at these concentrations (0.1 M) of total mixed species, the theoretical curves are essentially linear for $K < 5$ $M^{-1}$. Superimposed on this figure, to illustrate the analysis, we plot the downfield chemical shifts observed for the real case when we study the imino proton of guanosine as the G monomers are progressively replaced by C, resulting in G·C dimer formation (at $[G_{tot}] + [C_{tot}] = M = 0.1$ M in $Me_2SO$-$d_6$-$CDCl_3$; 1:1). The accuracy of the data for these experiments is ±2. Hz, and, as Figure 6 shows, a fairly definitive best fit is obtained to the theoretical curve for $K = 27$ $M^{-1}$.

In a similar way, we have solved the equations governing the (trimer) association of substituted adenines with two uracil residues to obtain best-fit association constants. The experi-

[4] The term "mixing curve," as used here, should not be confused with mixing curves measured by uv or ir spectroscopy. Both observe the same phenomena, but the time resolution of nmr and optical spectroscopy are very different. Thus the nmr experiment displays the monomer-dimer equilibrium *via* a *single* peak for each proton under conditions of fast exchange, and the peak position depends on the relative concentrations of the two species. In contrast, the uv or ir experiment shows characteristic peaks for *both* individual species, and the relative peak sizes monitor the equilibrium between monomer and dimer.

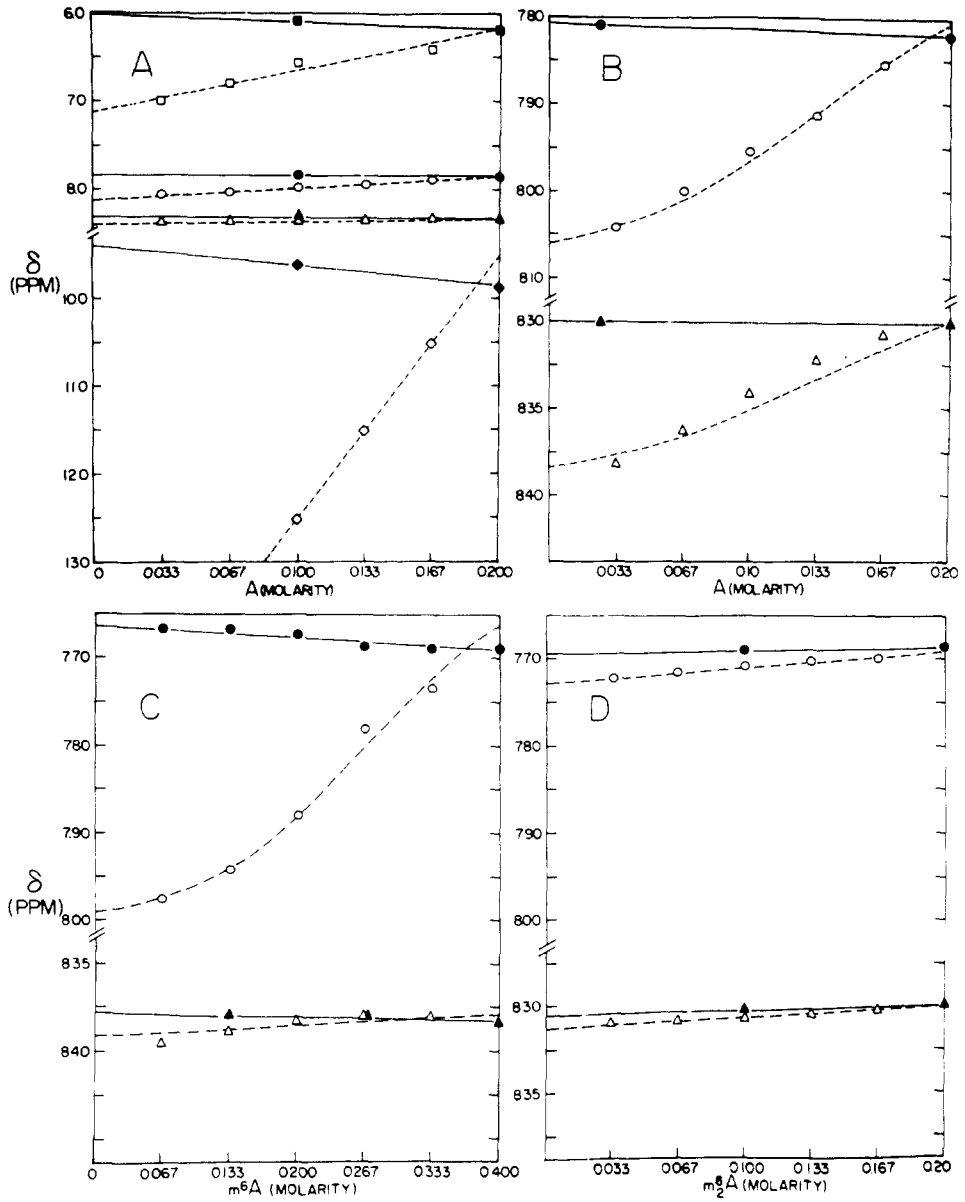
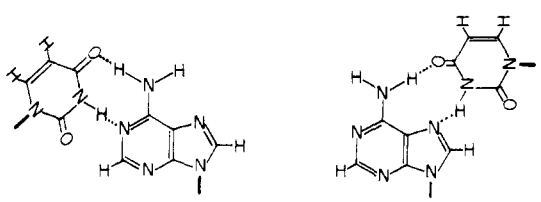



FIGURE 7: Mixing curves of A + U (A, B), $m^6A$ + U (C), and $m_2{}^6A$ + U (D) in $CDCl_3$. All spectra were recorded in the CW mode at 20° (±1°), and are referred to the frequency offset from the internal (HMD) reference. □ = A ($C_6$-$NH_2$); ○ = A ($C_8$-H); Δ = A ($C_2$-H); ◇ = U ($N_1$-H). -▲-▲- (filled symbols, solid lines) = dilution of base with no added complementary base; -Δ-Δ- (open symbols, dashed lines) = dilution of base with complementary base added to make the *total* base concentration to the initial value (*e.g.*, every point on the dashed lines in A is at a *total* base concentration of 0.200 M). Imino proton peak positions were reproducible to 0.05 ppm and amino protons were accurate to 0.02 ppm. $C_2$- and $C_8$-proton chemical shift values were reproducible to 0.002 ppm. The abscissa in Figure 7A (*i.e.*, the concentration of adenosine) is reversed for the U results.

ments are summarized in Figure 7, and the results of the analysis are described below.

*A·U Base Pairs.* The original experiments of Katz and Penman (1966) showed that when adenine and uracil are mixed in a suitably nonpolar organic solvent, the predominant association involves coplanar hydrogen-bonded A·U base pairing with little stacking of the aqueous solution type (*e.g.*, see Solie and Schellman, 1968). Our results are consistent with this conclusion, and are depicted in Figure 7, which shows the mixing curve for A and U in the presence and absence of the complementary base. It has been shown that adenine can engage in hydrogen-bonded complex formation with two uracil residues simultaneously in organic solvent (*e.g.*, see Nagel and Hanlon, 1972). In such trimers, one U occupies the "Watson-Crick" site (IIA), and the other the "Hoogsteen" site (IIB). The mixing curves we obtained with A and U are qualitatively similar to those obtained by Katz and Penman (1966), and our curves



IIA IIB

show the same small chemical shift differences observed by Katz (1969) for the $C_2$ and $C_8$ protons upon mixing A with U (Figure 7B). These differences are small compared to the amino- and imino-proton chemical shift effects (Figure 7A), but they are nevertheless distinct, and easily and accurately measured. A downfield chemical shift is characteristic of complex formation (Figure 7, dash lines) and does not occur when dilution is carried out in the absence of the complementary

TABLE II: Comparison of Association Constants of Cyclic Dimer Formation in Organic Solutions.[a]

| Complexing Species | $K$ ($M^{-1}$) | Association Ratios[b] | |
|---|---|---|---|
| | | Equilibrium Constant Compared | Value of Ratio |
| [c] G·C | 33 | $K_{G\cdot C}/K_{G\cdot m^4C}$ | 7 |
| G·m$^4$C | 5 | | |
| m$^2$G·C | 50 | $K_{m^2G\cdot C}/K_{m^2G\cdot m^4C}$ | 6 |
| m$^2$G·m$^4$C | 8 | | |
| m$_2^2$G·C | <1 | $K_{G\cdot C}/K_{m_2^2G\cdot C}$ or $K_{G\cdot C}/K_{G\cdot m_2^4C}$ | 33 |
| G·m$_2^4$C | <1 | | |
| [d] A·U$_H$ | 19 | $K_{A\cdot U,H}/K_{A\cdot U,WC}$ | 1.1 |
| A·U$_{WC}$ | 18 | | |
| m$^6$A·U$_H$ | 17 | $K_{m^6A\cdot U,H}/K_{m^6A\cdot U,WC}$ | 11.0 |
| m$^6$A·U$_{WC}$ | ~1.5 | | |
| m$_2^6$A·U | <1 | $K_{A\cdot U}/K_{m_2^6A\cdot U}$ | >19 |

[a] All association constants were calculated according to eq 1 after equilibration of the sample at the desired temperature. [b] Association ratios are obtained by comparing the association constants of different pairs of hydrogen-bonded dimers in the *same* solvent. [c] Mixing curves done in $Me_2SO$-$d_6$:benzene-$d_6$ (1:1) at 31 ± 1°C. *Ratios* are accurate to ±10%. [d] Mixing curves done in $CDCl_3$ at 20 ● 1°. Absolute values of the association constants are relatively poor, but the *ratios* obtained are accurate to better than ±20% (see footnote 7).

base (Figure 7, solid lines). Thus the $C_2$ and $C_8$ protons of adenine serve as excellent probes of their chemical environments, even though these protons are not directly involved in hydrogen bonding. We conclude (in agreement with Katz, 1969) that a significant downfield chemical shift of the $C_2$ proton implies that a Watson–Crick hydrogen bonded structure has formed, and that an analogous downfield chemical shift of the $C_8$ proton implies hydrogen bonding interactions of the Hoogsteen type.

Inspection of structures IIA and IIB shows that this interpretation is chemically reasonable; *i.e.*, the magnetic environment of the $C_2$ proton of adenine should certainly be perturbed by a uracil residue engaged in Watson–Crick dimer formation, while the $C_8$ proton should be most sensitive to a hydrogen-bonded structure of the Hoogsteen type. In fact, when we examine cyclic dimers formed between guanosine and cytosine (where only Watson–Crick dimers are seen), the $C_8$ proton is seen to shift by less than 1.5 Hz while in the A plus U mixing curve the same proton shifts downfield by over 22 Hz. This latter observation is totally consistent with the above conclusion that in a matched pair of bases (*e.g.*, either G plus C or A plus U), the effect of Watson–Crick hydrogen bonding on the magnetic environment of a distant proton is negligible. We thus feel justified in attributing a chemical shift of the $C_2$ proton primarily to Watson–Crick dimer formation in the A plus U case, and likewise in attributing a chemical shift of the $C_8$ proton primarily to Hoogsteen hydrogen bonding.

Figure 7B shows the $C_2$ and $C_8$ proton chemical shift data superimposed on the best-fit theoretical curves to these data obtained as outlined above, on the assumption that binding of U is exclusive to either one site or the other.[5] The values of $K$ obtained in this way are approximately equal for both adenine binding sites (Table II), confirming the finding of Nagel and Hanlon (1972) that the intrinsic affinity constants for U binding to the two A sites ($K_{WC}$ and $K_H$) are approximately the same. (This conclusion is also implicit in the earlier results of Kyogoku *et al.*, 1967.) The deviation of the experimental data from the theoretical curves in Figure 7B indicates that the assumption of exclusive binding is clearly incorrect, and thus also that the absolute values of $K$ obtained are probably low (see footnote 5). On the other hand, these deviations should have no significant effect on the *relative* values of $K_{WC}$ and $K_H$, nor should there be any effect on the same parameters measured for methylaminoadenine described below.

We have made the following use of the $C_2$ and $C_8$ proton environmental probes in connection with the methylation problem. Since the restricted rotation data on methylaminoadenine have been interpreted in terms of an equilibrium preference for one of the two possible rotamers of the methyl group at low temperature, we surmise that observation of the relative chemical shifts of the $C_2$ and $C_8$ protons in a mixing experiment of m$^6$A and U should produce a drastically reduced downfield shift for the proton ($C_2$ or $C_8$) on the side preferentially populated by the methyl group, while the proton on the other side ($C_8$ or $C_2$) should be relatively unshifted from the position observed in free (unmethylated) adenine. In Figure 7C, the mixing experiment previously carried out with A plus U is repeated with m$^6$A plus U. The results indicate that the m$^6$A·U complex does indeed exist almost exclusively as a Hoogsteen-type hydrogen bonded structure in organic solution, since almost no change in the chemical shift of the $C_2$ proton is observed with added complementary base, while the $C_8$ proton shifts approximately as far as in a normal A·U$_2$ complex. Thus, these experiments show that at ambient temperature the methyl group preferentially occupies a position blocking the hydrogen-bonding interactions characteristic of native double-helical nucleic acids. The quantitative extent of this preference, as measured by ratios of relevant dimerization constants obtained by fitting the data of Figure 7C to theoretical curves, is summarized in Table II. By these same criteria, Figure 7D indicates, as ex-

[5] Since we assume exclusive binding to either site, in the initial formulation of this trimeric association we know that the concentration of free U is overestimated since the equation governing U binding is $\delta_{U,obs} = f_U\delta_U + f_{WC}\delta_{WC} + f_H\delta_H$. But for our purposes, we have assumed it to be either $\delta_{U,obs} = f_U\delta_U + f_{WC}\delta_{WC}$ or $\delta_{U,obs} = f_U\delta_U + f_H\delta_H$ depending on whether we observe the Watson–Crick or Hoogsteen site. Thus we can see from eq 1 that by ignoring one association while measuring the other, we have overestimated the amount of free U, leading to an underestimate of the association constant, $K$.

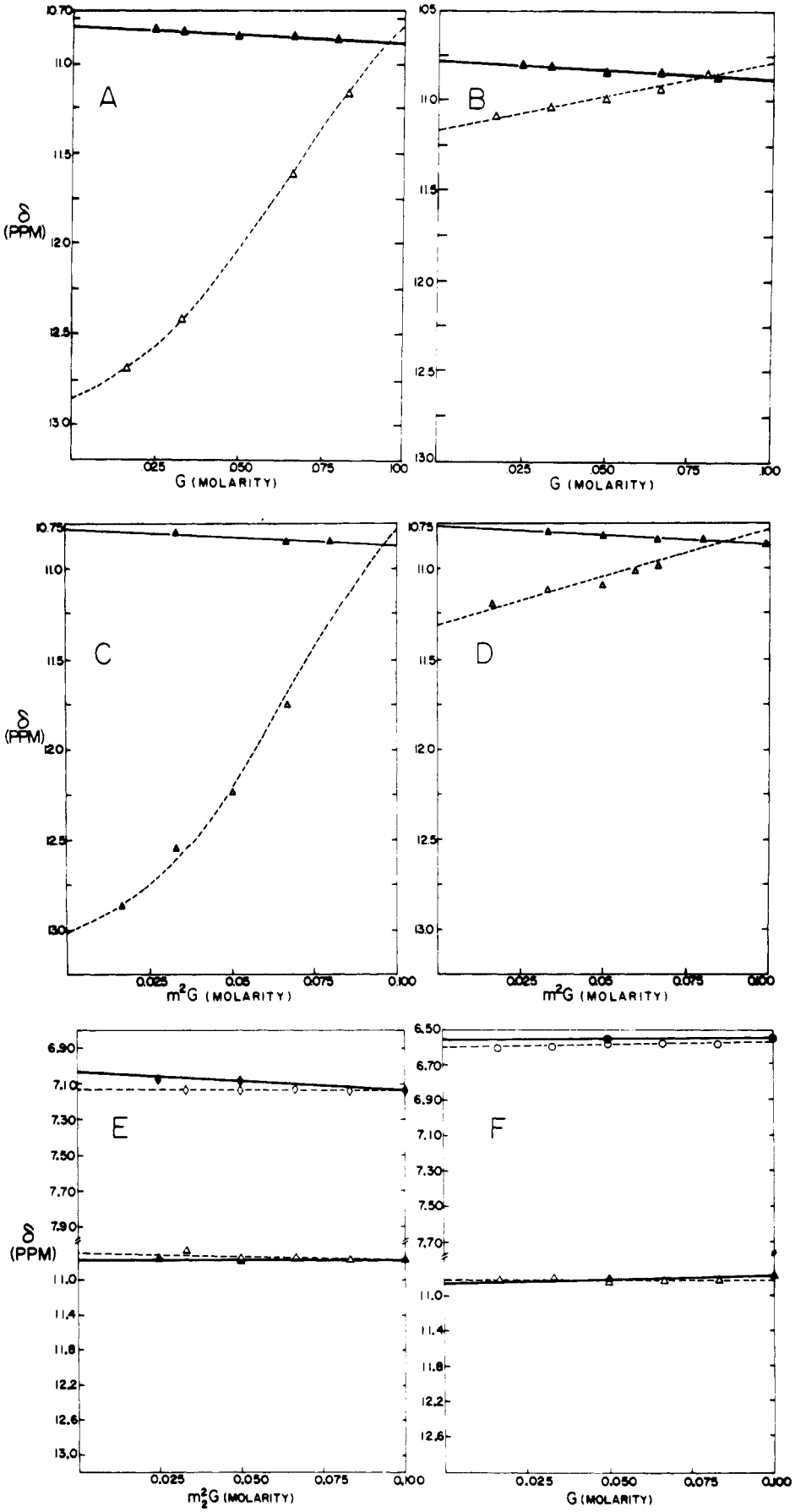



FIGURE 8: Mixing curves of G + C (A), G + m⁴C (B), m²G + C (C), m²G + m⁴C (D), $m_2^2$G + C (E), and G + $m_2^4$C (F) in $Me_2SO$-$d_6$:benzene-$d_6$ (1:1). All spectra were recorded in the CW mode at 32 ± 1°C, and are referred to the frequency offset from HMD. ○ = G ($C_2$-$NH_2$); △ = G ($N_1$-H); ◇ = C ($C_4$-$NH_2$). Open and solid data points and dashed and solid lines have the same meaning as in Figure 7. The imino resonance position was reproducible to 0.05 ppm, and the amino resonances were reproducible to 0.02 ppm.

pected, that essentially no hydrogen-bonded dimer formation occurs between $m_2^6A$ and U. Thus this experiment serves as a useful control (Table II).

*G·C Base Pairs.* The effect of exocyclic methylamino groups on hydrogen bond formation between guanosine and cytosine is shown in Figure 8. All the G + C mixing experiments were carried out in a solution of $Me_2SO$-$d_6$:benzene-$d_6$ (1:1), which solubilizes the usually insoluble guanosine derivatives while maintaining interbase hydrogen bonding. These results are analogous to, but different in focus than, the experiments of Newmark and Cantor (1968). Figure 8A shows the results of mixing G and C, displaying the expected downfield shift indicative of hydrogen bonding when the complementary base is added.

Figure 8B depicts the result when G is mixed with $m^4C$. A drastic reduction in the magnitude of the downfield chemical shift (and thus the association constant, Table II) provides evidence that the methylamino group of cytosine has an equilibrium preference for blocking the normal Watson–Crick hydrogen bonding site (the only one available) of a G·C base pair. Thus the amino-bound methyl group of cytosine shows the preferential syn positioning "inside" the nucleic acid double helix which is also characteristic of $m^6A$, and aminomethylation of C likewise disrupts Watson–Crick hydrogen bonding. This preferred rotamer stereochemistry for $m^4C$ was previously proposed on the basis of monomer restricted rotation studies, as pointed out above, by Shoup *et al.* (1972a).

In contrast, Figure 8C shows that when $m^2G$ is added to C, the resultant complex appears to behave nearly identically with G·C (compare with Figure 8A and Table II); *i.e.,* methylation of guanosine on the exocyclic amino group contributes essentially no measurable destabilization to a hydrogen bonded complex in this solvent. In Figure 8D, the mixing of $m^2G$ and $m^4C$ shows behavior analogous to that of $m^4C$ + G (Figure 8B), further supporting the conclusion that $m^2G$ causes very little destabilization of the hydrogen-bonded complex in this organic solution. In Figures 8E and F, the mixing curves for $m_2^4C$ + G and for C + $m_2^2G$ show very little base–base interaction. Thus the mixing of molecules which can form no hydrogen bonds due to total blockage of the potential hydrogen-bonding sites again serves as a useful control to our conclusions (see Table II).

## Discussion

*Preferential Positioning of Exocyclic Methylamino Substituents.* Crystallographic studies have shown that the exocyclic amino groups of adenine, cytosine, and guanine are each approximately coplanar with their respective heterocyclic rings, and that this coplanarity is maintained when these groups are involved in interbase hydrogen bonding (*e.g.,* see Voet and Rich, 1970). Furthermore, the substituents added by amino group alkylation are also coplanar to the rings. This confirms that the exocyclic nitrogen–ring carbon bond is partially double bond in character, as generally observed for aromatic exocyclic amino groups. As expected on this basis, the bond lengths of the $C_6$–$N_{10}$ bonds in unsubstituted (Ringertz, 1972), monomethylated (Sternglanz and Bugg, 1973), and dimethylated (Sundaralingam and Arora, 1969) adenine derivatives are all significantly shorter than a carbon–nitrogen single bond; and the C–N bonds of adenine and cytosine show restricted rotation with activation enthalpies for isomerization comparable to those of the partially double-bonded peptide linkage (*e.g.,* see Schellman and Schellman, 1964).

These observations (restricted rotation, partial double-bond character, and coplanarity of substituents with the heterocyclic aromatic ring) raise the question of preferential positioning of chemically different substituents of the nitrogen about the C–N bond; *i.e.,* for $m^6A$ does the syn (to $N_1$) or the anti rotamer represent the preferred (lower free energy) form, or are these rotamers energetically equivalent? This is a significant question, since in nucleic acid polymers, preferential syn positioning of the methylamino moiety (on the $N_1$ side of A or G, and on the $N_3$ side of C) would interfere with normal interbase Watson–Crick hydrogen bonding, while anti positioning would place these methyl groups into the large (for A and C) or small (for G) groove positions relative to the hydrogen-bonded base pairs. Moreover, even if the free energy change favoring base-pair formation were sufficient to overcome a preferential syn positioning of the methyl groups, thus "forcing" the methyls into the less favored anti position, a component of thermodynamic instability would be introduced which would manifest itself as a net weakening of the base-pairing interaction and thus as a lowered melting temperature of the double helix (see below).

These restricted rotation studies, together with the measurements of the specificity and relative affinity of cyclic hydrogen bonded dimer formation, indicate that the methyl group of $m^6A$ lies preferentially in the syn position relative to $N_1$ of the purine ring, and that of $m^4C$ in the syn position relative to $N_3$ of the pyrimidine ring. These studies also show that the methyl of $m^2G$ tends to lie in the anti position relative to $N_1$ of guanine.[6]

Inspection of CPK molecular models of the bases in question provides some rationale for these preferential methyl positionings: thus the methylamino group of $m^6A$ shows more steric interference with the $N_7$ nonbonding orbital than with the nonbonding orbital of the $N_1$ ring nitrogen. Similarly, more steric interference of the methyl group in $m^4C$ is seen with the $C_5$ hydrogen than with the $N_3$ lone pair. These observations, based on molecular models, are also consistent with the general finding (*e.g.,* see Eliel, 1970) that in heterocyclic rings the unshared electron pair of a ring nitrogen engenders less steric repulsion to a neighboring exocyclic substituent than does a hydrogen atom. In $m^2G$, inspection of molecular models reveals more steric interference between the methylamino group and the $N_1$ hydrogen than between the methyl and the unshared electrons of the $N_3$ nitrogen, in accord with the expectations cited above.

*Thermodynamic Consequences.* The restricted rotation data on the preferential positioning of the methyl groups of $m^6A$ and $m^4C$ can be used to calculate a free energy of destabilization of the Watson–Crick hydrogen-bonded base pairs resulting from the insertion of the methyl group. If we assume that interbase hydrogen-bonded dimer formation of the Watson–Crick type is prevented when the methylamino group is in the syn position (though of course dimer formation at the hoogsteen site between A and U monomers, involving the nonbonding orbital of the $N_7$ of adenine, can proceed) while Watson–Crick hydrogen bonding is unimpeded when the methyl is in the anti position, then we may write

$$\Delta(\Delta G^\circ)_{WC} = -RT \ln ([\text{anti}]/[\text{total}]) \qquad (2)$$

where $\Delta(\Delta G^\circ)_{WC}$ represents the change in the free energy of

[6] This conclusion is consistent with the findings of Chantot and Guschlbauer(1972) who showed that the gel formation in solutions of guanine-containing nucleotides (which is indicative of the formation of a cyclic tetrameric hydrogen-bonded structure) occurs at lower nucleotide concentrations in $m^2G$- than in G-containing solutions. These cyclic tetramers (see Gellert *et al.* (1962) for structure) could only form with $m^2G$ if the methyl group lies in the "outside" position.

TABLE III: Methylation-Induced Instability of Watson–Crick Hydrogen-Bonded Polymers. Comparison of Free-Energy Changes $[\Delta(\Delta G^\circ)_{WC}]$ Calculated from Rotational Isomerization, Cyclic Dimer Formation, and Melting Temperature Depression.[a]

| Bases | $\Delta(\Delta G^\circ)_{WC}$ (kcal/mol) Calculated from Rotational Isomers[b] | Comparative Associating Systems | $\Delta(\Delta G^\circ)_{WC}$ (kcal/mol) Calculated from Cyclic Dimer Formation[c] | $\Delta(\Delta G^\circ)_{WC}$ (kcal/mol) Calculated from Hydrogen-Bonded[d] Polymers |
|---|---|---|---|---|
| $m^6A$ | +1.9 (96:4) | $m^6A$·U-A·U | +1.5 | +0.9 |
| $m^4C$ | +1.8 (95:5) | G·$m^4C$-G·C | +1.2 | |
| | | $m^2G$·$m^4C$-G·C | +0.9 | |
| | | I·$m^4C$-I·C | | +0.8 |
| $m^2G$ | ~0 | $m^2G$·G-G·C | −0.1 | >2. |

[a] $\Delta(\Delta G^\circ)_{WC}$ values are calculated at 300°K, in kcal/mole of methylated base. [b] From eq 2. Population ratios (syn:anti) obtained at −40° are given in parentheses. [c] From eq 3, using association data in Table II and Results. [d] From eq 4. References, assumptions, and polynucleotide systems studied are in Discussion.

interaction of the Watson–Crick base pairs due to the methyl group, and [anti] and [total] correspond to the concentrations of the relevant anti and total nucleoside base isomers, respectively. In this formulation we assume that methylation has no other effects on the base–base interaction; some aspects of this assumption will be tested by the considerations to follow. We should also note that eq 2 dictates that monomethylation of an amino group potentially involved in interbase hydrogen bonding must destabilize the dimer complex even in the *absence* of preferential positioning of the methyl on the syn side. Thus totally random positioning of the methyl will change $(\Delta G^\circ)_{WC}$ by $-RT \ln (1/2)$ (~ +0.4 kcal mol base pairs at 300°K), simply as a consequence of the loss of one of the two equivalent potential hydrogen-bond donor groups. Values of $\Delta(\Delta G^\circ)_{WC}$ calculated by eq 2, utilizing the restricted rotation data on the relative populations of the methylated isomers of $m^6A$ and $m^4C$, are summarized in Table III.

For comparison, and as a system one step closer to the "real" situation in the native DNA double helix, we can also determine $\Delta(\Delta G^\circ)_{WC}$ by utilizing the equilibrium constants calculated for hydrogen-bonded dimer formation in various organic solvents with methylated and unmethylated bases. For the A·U (and presumably by analogy for the A·T) reaction, Nagel and Hanlon (1972; see also Kyogoku *et al.*, 1967) have shown that the intrinsic equilibrium constants are approximately equal for the formation of the 1:1 dimer, with either the Watson–Crick or the Hoogsteen binding positions of A. Therefore, at any particular point on the mixing curve, we expect approximately equal concentrations of A·U Watson–Crick and Hoogsteen dimers. However, if A is replaced by $m^6A$ in these systems, the distribution of Watson–Crick- and Hoogsteen-type dimers should be biased toward the Hoogsteen form (indicated by a significant chemical shift of the $C_8$ proton) by an equilibrium distribution of the methyl groups which favors the syn isomer (*i.e.*, preferentially blocking the Watson–Crick base pairing positions). Such a redistribution of the base-pairing equilibria is demonstrated in Figure 7, and the intrinsic equilibrium constants for dimer formation calculated in Results (Table II) can be used quantitatively to deduce $\Delta(\Delta G^\circ)_{WC}$ on these grounds. Thus

$$\Delta(\Delta G^\circ)_{WC,m^6A} = -RT \ln [K_{(m^6A\cdot U)_{WC}}/K_{(A\cdot U)_{WC}}] \quad (3)$$

where $\Delta(\Delta G^\circ)_{WC,m^6A}$ is the change in the free energy for A·U Watson–Crick base-pair formation accompanying the introduction of a methyl group onto the exocyclic amino of A, and $K_{(m^6A\cdot U)_{WC}}$ and $K_{(A\cdot U)_{WC}}$ are the *intrinsic* equilibrium constants for $m^6A$·U and A·U dimer formation (*in the Watson–Crick mode*), respectively.[7] These values are listed in Table III, and are in good agreement with those determined from the restricted-rotation rotamer-population data for $m^6A$.

Similar calculations can be made more straightforwardly from the $m^2G$·C and G·$m^4C$ data (Table II), since no complicating trimolecular structure can form. Values of $\Delta(\Delta G^\circ)$ calculated on this basis are also compared with values derived from restricted rotation data (for $m^4C$) in Table III. The agreement is quite reasonable, given the errors involved in both measurements, thus buttressing our central hypothesis that the primary determinant of the change in stability of the hydrogen-bonded dimers is the preferential positioning of the methyl group rather than possible intra-ring electronic rearrangements due to the insertion of this group which might be reflected in changed hydrogen bond strengths, and thus in altered equilibrium constants.

The data of Kyogoku *et al.* (1967), obtained by infrared measurement of equilibrium constants in organic solvents for hydrogen-bonded dimer formation between various $N_1$-blocked uracil and $N_9$-blocked adenine derivatives, can also be used to strengthen this last conclusion. Their measurements did not distinguish between Watson–Crick and Hoogsteen dimer formation, and thus for our purposes can be used only to indicate the *ratios* of equilibrium constants obtained for the formation of methylated and unmethylated dimer pairs, respectively. These workers show that the ratios of the association constants for (A·U)-($m^6$A·U), (A·5,6 dihydro-U)-($m^6$A·5,6 dihydro-U), (A·T)-($m^6$A·T), and (A·5 BrU)-($m^6$A·5 BrU) are all approximately constant, as are the ratios of these equilibrium constants when other adenine derivatives are paired with these variously modified uracils. Furthermore, the ratios, in all cases, are very close to 2, which is the expected result if the methyl group simply blocks one or the other hydrogen-bonding position without altering the *intrinsic* affinities of these positions for base-pair formation.

*Polynucleotide Data.* Thus our results show, at least to the level of cyclic hydrogen-bonded dimer formation in organic sol-

[7] We should note that only *ratios* of equilibrium constants, measured for dimer formation involving methylated and unmethylated monomers in the *same* organic solvent, enter into eq 3. Thus the calculated values of $\Delta(\Delta G^\circ)$ should be independent of the nature of the solvent, which means that the solvent can be manipulated (within the limits of other experimental constraints such as solubility) to put the equilibrium constants into the optimal range for measurement.

vents, that the free energy effects on base-pair stability of a methyl group substitution on the exocyclic amino group of A, C, and G are consistent with expectations based on the preferential positioning of these moieties as syn or anti rotamers with respect to the "sides" (hydrogen-bonding positions) of the purine or pyrimidine rings invoved in Watson-Crick double-helix formation. Of course the ultimate question of biological interest is how the insertion of a methyl group into an otherwise complementary nucleotide sequence might alter the local geometry and conformational stability of the nucleic acid double helix. To approach this question, we turn to the data available on exocyclic monomethylamino derivatives of complementary polynucleotide structures.

Based on the monomer and dimer results summarized in Table III, we predict that the substitution of A by $m^6A$ (or C by $m^4C$) should decrease the free energy of the double helix (relative to single-stranded forms) by ~1.5 kcal/mol of substituent, while the replacement of G by $m^2G$ should decrease helix stability by less than 0.1 kcal/mol. If we accept the comparison of the cyclic hydrogen-bonded dimer data with the methyl rotamer population data cited above as strong evidence that methylation does not appreciably alter the strengths of the hydrogen bonds between the partners involved in complementary base-pair formation *within* a double helical polynucleotide, the fact that we are now concerned with a polymeric structure might enter in two additional ways: (i) the overall stereochemistry of the double-helical DNA structure might not accommodate a methyl group in the large groove (for $m^6A$ and $m^4C$) or in the small groove (for $m^2G$) without some strain due to steric interference between the methyl group and the atoms comprising the "walls" of the grooves; and (ii) the methyl groups themselves might alter the stacking interactions of the modified bases with the bases located directly above and below them in the polymeric double-helix.

These aspects of the effects of methylation are unfortunately difficult to isolate from the data available in the literature, since the substituents of interest have generally been inserted onto every residue (*i.e.*, into every base pair). Thus, in addition to possible changes in helix stability due to the insertion of isolated amino substituents, possible *substituent* stacking interactions must also be considered. However, the limited literature data do suggest that changes in melting temperatures (which we will use as a measure of change in helix stability) are generally approximately linear functions of the fraction of substituent incorporated into partially modified strands (*e.g.*, see Inman and Baldwin, 1964; Fink and Crothers, 1972; J. D. Engel and P. H. von Hippel, in preparation), so for present purposes we will neglect this possible complication in going further.[8]

In most available studies, changes in the melting temperature ($T_m$) of double helical polymers which contain either only fully modified bases, or only unmodified bases, have been measured. These changes can be approximately related to the change in the free energy of stabilization of the double helix per mole of modified base pairs (in units of kcal/mole) by an empirical equation

$$\Delta(\Delta G^\circ)_{WC} = (\Delta T_m)(\text{fraction of base pairs modified})(1/50) \quad (4)$$

where $\Delta T_m$ is the melting temperature of the modified double helix subtracted from that of the control (unmodified "parent" complementary polynucleotide), and ($1/50$) is the change (in kcal/mole) in helix stability corresponding to a 1° decrease in melting temperature.[9] This empirical equation can also be deduced directly as one consequence of a general theory of nucleic acid structure destabilizaton, which provides a relation between the observed $T_m$ depression and the extent and type of substituent incorporated into the double helix. Such a theory has been partially provided (for base substitutions resulting in *total* mismatching of interbase hydrogen bonding) by Fink and Crothers (1972), and a further treatment and extension, including supporting experiments, will be presented elsewhere (J. D. Engel and P. H. von Hippel, in preparation). However, for present purposes the use of eq 4 can be justified by the following polynucleotide observations relevant to these studies.

The most directly useful findings on the effects of monomethylation of the amino group of adenine on the stability of the poly(A)·poly(U) double helix are those of Ikeda *et al.* (1970), who compared the melting temperatures of poly(A)·poly(U) and poly($m^6A$)·poly(U), and also of poly($2NH_2A$)·poly(U) and poly($2NH_2m^6A$)·poly(U). They showed that under the conditions used, 1:1 complexes were formed in all the systems studied and that the stability of the poly($m^6A$)·poly(U) helix (Griffin *et al.*, 1964) is considerably less than that of the poly(A)·poly(U) helix (see also Ikehara *et al.*, 1972). They also demonstrated that the effects on melting temperatures of adding an amino group at the $C_2$ position of adenine (thus permitting the formation of three interbase hydrogen bonds rather than two), and of monomethylation of the amino group at the $C_6$ position, are independent and approximately additive. This observation indicates strongly that even though Hoogsteen base-paired structures can form perfectly well at the monomer level, 1:1 *polymer* complexes form only in the Watson-Crick mode, with Hoogsteen bonding serving to add a third strand to an already stable double-helical structure.

[8] Some measurements of monomer "stacking" equilibrium constants in aqueous solution suggest that, at least in some of these systems, methylation does alter stacking interactions appreciably. Thus, for example, Broom *et al.* (1967) have shown that adenosine, $N^6$-methyladenosine, and $N^6$-dimethyladenosine self-associate in water with equilibrium constants of 4.5, 13, and 22 $M^{-1}$, respectively. However, we would expect such measurements to lead to overestimates of changes which might be observed in the native double helix, since the monomers, of course, are much more free than the polymerized bases to rearrange themselves to maximize stacking interactions, while the polymerized bases are restrained from making optimal use of the new stacking possibilities introduced by the added methyl groups by the topological requirements imposed by the sugar-phosphate backbone and the need to maintain appropriate interbase hydrogen-bonding distances and direction. In addition, the *change* in the free energy of helix formation on introducing a nucleotide analog into the structure depends on the free energy *difference* between the double helix and the products of the "melting" reaction. If increased stacking between methylated bases is accompanied by increased stacking in the single-stranded methylated structure formed in the melting process, changes in stacking interactions would be expected to have little effect on the *magnitude* of the $T_m$ change due to methylation.

[9] This conversion factor may be obtained as follows (*e.g.*, see Crothers and Zimm, 1964). At $T_m$, $\Delta G^\circ \equiv 0$; therefore the effect on the free energy of helix formation of a modification in helix structure (such as methylation) is obtained by the relation: $\Delta(\Delta G^\circ) = \Delta H^\circ - T_{m,mod}\Delta S^\circ$, where $T_{m,mod}$ is the melting temperature of the modified polymer, and $\Delta H^\circ$ and $\Delta S^\circ$ are the changes in standard enthalpy and entropy of helix formation (at pH 7.0) for the unmodified double helix. For $\Delta H^\circ$ and $\Delta S^\circ$, we have used −7.0 kcal/mol and −20 eu/mol, respectively. These represent the average values applicable to the temperature range 30–100°, from the most recent data of Shiao and Sturtevant (1973) for native calf thymus DNA. The error resulting from ignoring the small changes in $\Delta H^\circ$ and $\Delta S^\circ$ with temperature is less than ±5% (especially since both parameters change in the same direction with temperature). We use the enthalpy and entropy changes for the formation of a DNA helix involving unmodified bases since the thermodynamic data for the relevant substituted double-helical (and mostly ribopolynucleotide) systems are not available.

(See also the discussion of this point in Voet and Rich (1970).) The 1:1 polymer–monomer double-helical structure formed by $m^6m^9A$ (monomer) and poly(U) also appears to occur in the Watson-Crick mode (Hoffman and Pörschke, 1973), providing further support for this conclusion.

Ikeda *et al.* (1970) found, in comparing the above polymer systems, that the addition of the 2-amino group to the adenine residues *increased* the $T_m$ of the complex by 29° (±2°) above the corresponding $T_m$ for poly(A)·poly(U) [or poly($2NH_2A$)·poly(U)], and that the monomethylation of adenine residues *decreased* $T_m$ by 43° (±2°) below the corresponding value for poly(A)·poly(U) [or poly($2NH_2A$)·poly(U)]. By eq 4, the latter alteration in adenine structure corresponds to a destabilization of the Watson-Crick helix by ~+0.9 kcal/mol of methyl group inserted into the helix, in reasonable agreement with the values derived from the simpler systems summarized in Table III. [We may note that this value would be larger, and thus agreement with monomer data improved, if a larger (and perhaps somewhat more realistic) value of $\Delta H^\circ$ were used for the formation of the stacked $m^6A$·U base pair.[9]]

Quantitative polymer data directly relevant to the effects of the monomethylation of the $C_2$-amino group of G, or the $C_4$-amino group of C, on the stability of an isolated G·C base pair are not available, but the information which does exist is compatible (for $m^4C$) with the model developed here. Thus Brimacoombe and Reese (1966) have shown that the partial replacement (39%) of C bases by $m^4C$ in poly(I)·poly(C) destabilizes the double-helical structure by 15°, with no double helical structure formed at room temperature when all the C residues have been replaced by $m^4C$. Assuming that here also the $T_m$ depression will be approximately linear in mole fraction $m^4C$ incorporated in the polymer, we can estimate a $T_m$ depression for the *totally* methylated poly($m^4C$)·poly(I) double helix of ~40° from these data. [It is important to note that mixing curves of the partially methylated systems show that the $m^4C$ residues, as well as the C residues, are involved in hydrogen bonding with the inosine bases at temperatures below $T_m$, thus demonstrating that destabilization is due to a decreased stability of the $m^4C$:I pair rather than to complete loss of base pairing (loop formation, bulges, etc.) at these loci.] In terms of eq 4, this $T_m$ depression corresponds to a value of $\Delta(\Delta G^\circ)_{WC}$ of ~+0.8 kcal/mol, again in reasonable accord with estimates based on monomer data (Table III).

This approximate compatibility of polymer-based ($T_m$ depression) and monomer-based (rotomer population and dimer formation) data suggest that the methyl groups of both $m^6A$ and $m^4C$ fit into the large groove of the double-helical structure without much steric distortion, and also without making substantial additional contributions to the *relative* stability of the structure *via* modified stacking interactions.

The situation in which G in poly(G) is partially replaced with $m^2G$ is less straightforward. Pochon and Michelson (1969), as well as Ikehara and Hattori (1971), have shown that fully methylated poly($m^2G$) does not form a stable double helix with poly(C). Measurements of $T_m$ for the copolymer poly($m^2G$·G)·poly(C) system (containing various input ratios of $m^2G$ to G) have not been made; but mixing curves indicate that the 1:1 complex (involving hydrogen bonding between $m^2g$ and C as well as between G and C bases) is stable at room temperature when the ratio of $m^2G$:G < 70:30 (Ikehara and Hattori, 1971). Their data imply a melting temperature decrease for the fully methylated complex [relative to the melting of poly(G)·poly(C)], in excess of 100°, again assuming linearity of $T_m$ depression. From eq 4, this corresponds to a change in double-helix stability in excess of +2 kcal/mol of methyl substituent, which is clearly *much* larger than the value of $\Delta(\Delta G^\circ)_{WC}$ expected from monomer data (Table III), and suggests that factors other than preferential rotamer positioning must contribute to double-helix stability here.

That these factors may involve a steric inability of the methyl group to fit into the small groove of the double helix is suggested by the fact that neither poly(2-methyl I) nor poly(2-methyl thio I) can form a stable double-helical structure with poly(C) (Ikehara and Hattori, 1972a,b). This suggests that *any* substituent (other than an amino group) in the $C_2$ position of poly(G) will tend to interfere with Watson-Crick hydrogen bond formation with poly(C) *at the polynucleotide level.*[10] Inspection of CPK molecular models tends to support this conclusion, indicating that there is less space between the bases at the 2 position of G in the G·C pair than at the comparable locus in the A·T pair.

In concluding this section, we believe that the restricted rotation-methyl isomer population data generated on monomers can be applied to polymers as well (at least for $m^6A$ and $m^4C$), and thus may be used to make at least semiquantitative predictions about the effects of methylation on the local stability of the nucleic acid double helix. This finding may have significant consequences for the structure, stability, and function of the heavily methylated tRNA and rRNA molecules, as well as helping to elucidate the molecular basis of the interaction of specific DNA sequences with modification–restriction enzymes.

*Effects of Alkylation on tRNA and rRNA Structure.* The most extensive modification of biologically active nucleic acid polymers occurs in tRNA. Alkylated, thiolated, and otherwise modified nucleotide bases occur in every natural tRNA, and modified bases account for a much higher percentage of the total bases in this RNA than in any other nucleic acid component (Hall, 1971). Nishimura has recognized that very similar base sequences are alkylated in different tRNAs. On this basis, he has proposed that alkylation facilitates formation of precise anticodon conformations (which are presumably necessary for anticodon:codon recognition) *via* "stabilization" (by strengthening "stacking" interactions?) of the tertiary structure of the anticodon loop (Nishimura, 1972).

Certainly alkylation does increase monomer stacking (see footnote 8, and also Broom *et al.*, 1967; Pörschke and Eggers, 1972). However, based on the results of this paper, we propose that another, and perhaps more important, consequence of some types of tRNA alkylation might be the *destabilization* of certain tertiary structures which can occur in unmodified tRNA, and that this selective destabilization might differentially stabilize the biologically active structure in the modified tRNA.

For example, Gefter and Russell (1969) have isolated a tRNA from certain mutant bacteria which differs from the wild-type in the absence of a specific alkyl substituent on the adenosine residue directly adjacent to the anticodon. This tRNA ($tRNA_{E.\ coli}^{Tyr}$[$su^+$ III]) will not transfer amino acids or bind to ribosomes, although it retains the ability to accept an amino acid in the reaction with its cognate synthetase (with identical kinetics of synthetase reaction for both species). Such results might most straightforwardly be interpreted by suggesting that modification of bases adjacent to the anticodon has lit-

[10] We may note that the small groove is somewhat smaller, and the large groove somewhat larger, in the A form than in the B form double-stranded polynucleotide structure. Thus these conclusions, which are based primarily on polyribonucleotide data, may have to be somewhat modified when extended to the B form double helix of DNA.

tle or no effect on the bulk of the tRNA conformation, but serves primarily to define the anticodon triplet by destabilizing *in*appropriate double-helical structures (involving this region) with either additional bases in the messenger (mRNA) sequence, or with other parts of the tRNA molecule.

The rRNA of both prokaryotes and eukaryotes is methylated primarily at the 2′-hydroxyl position of the ribose moiety (Nichols and Lane, 1966; Brown and Attardi, 1965). However, the bases themselves are also appreciably methylated in rRNA, and Iwanami and Brown (1968) have shown that the largest fraction of this methylation occurs as $m^4C$, with a smaller fraction of $m^6A$ and $m_2{}^6A$ and other modified species also being generated. Our results may be coupled with these findings to suggest that such methylation may have an appreciable effect in directing and controlling the delicate interactions (*e.g.*, between loops of rRNA and between rRNA regions and the ribosomal proteins), which must be crucial in forming the active ribosome.

*Inferences for the Recognition of Specific DNA Sequences by Modification–Restriction Enzymes.* The finding of preferential syn positioning of the methyl group of $m^6A$ provides at least the beginnings of a possible, and testable, mechanistic explanation of the specificity of bacterial host modification–restriction systems.

Briefly, the known molecular facts about the best understood of these systems (the *E. coli* B and K systems: see Meselson *et al.*, 1972; Eskin and Linn, 1972; Lautenberger and Linn, 1972) are these. (1) The same enzymes (or at least shared enzyme subunits) are responsible for both function. (2) Modification and restriction occur only with native, double-stranded DNA. (3) In the modification mode, the system methylates (by methyl group transfer) a total of two specific adenine nucleotides per recognition site, these groups being located in close proximity on opposite DNA strands. The *E. coli* B and K systems methylate different sequences, and thus provide no interstrain protection against restriction. (4) Methylation occurs soon after replication, and the progeny strand in replication is preferentially modified (rather than restricted). (5) Completely unmethylated (unmodified) DNA which contains the recognition sequence specific for the enzyme is preferentially restricted; *i.e.*, subjected to specific endonucleolytic cleavage. (6) The nucleotide sequences of all restriction sites so far sequenced have proven to be centrosymmetric, with cleavage occurring about a pseudo-dyad axis on the double helix as either a straight or a staggered cut, the latter leaving complementary single-stranded ends up to five residues in length. (We should note that this statement applies to the *E. coli* B and K systems only by analogy, since the specific sites associated with the recognition function in these systems have not yet been sequenced.)

The preferential syn positioning of the methyl groups of $m^6A$ suggest that the native double-helical structure of DNA is locally destabilized by 1–2 kcal/mol of methyl group (Table III), relative to the same nucleotide sequence in unmodified DNA. This methyl-induced destabilization tends to favor forms in which the methylated adenine moieties are not hydrogen bonded to the complementary thymidine residues of the opposite chain. This conclusion, coupled with the above facts about bacterial modification–restriction systems, suggests that a partially open or less stable sequence is recognized preferentially by the modification methylase, while the more stable (unmethylated) native sequence is preferred by the restriction endonuclease.

The centrosymmetric nature of the site further suggests that this sequence could form transient "doodad" or "clover leaf" structures involving short, partially hydrogen-bonded hairpins protruding at right angles from the native DNA helix (see Gierer, 1966; Sobell, 1973). Such a configuration has been suggested by Meselson *et al.* (1972) as one which might be recognized by restriction–modification enzymes. Our present finding of local destabilization of the double helix introduced by methylation of the exocyclic amino group of adenine residues provides a possible energetic rationale for this proposal, since methylation would stabilize such a structure relative to the "straight" double-helix if the methylated adenines are located at the ends or "corners" of the loops in non-hydrogen-bonded positions. Specifically, this hypothesis predicts that the modification enzymes prefer (and the restriction enzymes fail to recognize) such a structure, and that methylation induced local instability of the double helix is a necessary (but obviously not sufficient) component in the protection of the host DNA against the cell's own restriction functions.

Acknowledgments

We are grateful to Professor Frederick Dahlquist, who provided much helpful advice on nmr interpretation, and particularly to our colleague Dr. James McGhee, whose constant advice, comments, and constructive criticisms were invaluable throughout this work.

References

Brimacoombe, R. L. C., and Reese, C. B. (1966), *J. Mol. Biol. 18,* 529.

Broom, A. D., Schweizer, M. P., and Ts'o, P. O. P. (1967), *J. Amer. Chem. Soc. 89,* 3612.

Brown, G. M., and Attardi, G. (1965), *Biochem. Biophys. Res. Commun. 20,* 298.

Chan, S. I., Schweizer, M. P., Ts'o, P. O. P., and Helmkamp, G. K. (1964), *J. Amer. Chem. Soc. 86,* 4182.

Chantot, J. F., and Guschlbauer, W. (1972), *in* The Purines: Theory and Experiment (Proceedings of the Fourth Jerusalem Symposium), Jerusalem, Israel Academy of Sciences and Humanities, p 205.

Crothers, D. M., and Zimm, B. H. (1964), *J. Mol. Biol. 9,* 1.

Eliel, E. L. (1970), *Accounts Chem. Res. 3,* 1.

Eskin, B., and Linn, S. (1972), *J. Biol. Chem. 247,* 6183.

Fink, T. R., and Crothers, D. M. (1972), *J. Mol. Biol. 66,* 1.

Freeman, R., and Hill, H. W. (1972), *J. Chem. Phys. 54,* 3367.

Gefter, M. L., and Russell, R. L. (1969), *J. Mol. Biol. 39,* 145.

Gellert, M. F., Lipsett, M., and Davies, D. R. (1962), *Proc. Nat. Acad. Sci. U. S. 48,* 2013.

Gierer, A. (1966), *Nature (London) 212,* 1480.

Griffin, B. E., Haslam, W. J., and Reese, C. B. (1964), *J. Mol. Biol. 10,* 353.

Hall, R. (1971), The Modified Nucleosides in Nucleic Acids, New York, N. Y., Columbia Press.

Hamlin, R., Lord, R. C., and Rich, A. (1966), *Science 148,* 1754.

Hoffman, G. W., and Pörschke, D. (1973), *Biopolymers 12,* 1611.

Ikeda, K. Frazier, J., and Miles, H. T. (1970), *J. Mol. Biol. 54,* 59.

Ikehara, M., and Hattori, M. (1971), *Biochemistry 10,* 3585.

Ikehara, M., and Hattori, M. (1972a), *Biochim. Biophys. Acta 269,* 27.

Ikehara, M., and Hattori, M. (1972b), *Biochim. Biophys. Acta 272,* 29.

Ikehara, M., Hattori, M., and Fukai, T. (1972), *Eur. J. Biochem. 31,* 329.

Inman, R., and Baldwin, R. L. (1964), *J. Mol. Biol. 8,* 452.
Iwanami, Y., and Brown, G. M. (1968), *Arch. Biochem. Biophys. 126,* 8.
Job, P. (1928), *Ann. Chim. (Paris) 9,* 113.
Johnson, C. S., Jr. (1966), *Advan. Magn. Resonance 1,* 35.
Katz, L. (1969), *J. Mol. Biol. 44,* 279.
Katz, L., and Penman, S. (1966), *J. Mol. Biol. 15,* 220.
Kenner, G. W., Reese, C. B., and Todd, Sir A. R. (1955), *J. Chem. Soc.,* 855.
Klagsbrun, M. (1973), *J. Biol. Chem. 248,* 2612.
Kyogoku, Y., Lord, R. C., and Rich, A. (1967), *Proc. Nat. Acad. Sci. U. S. 57,* 250.
Lautenberger, J., and Linn, S. (1972), *J. Biol. Chem. 247,* 6176.
Martin, D. M. G., and Reese, C. B. (1967), *Chem. Commun.,* 1275.
Meselson, M., Yuan, R., and Heywood, J. (1972), *Annu. Rev. Biochem. 41,* 447.
Nagel, G. M., and Hanlon, S. (1972), *Biochemistry 11,* 823.
Newmark, R. A., and Cantor, C. R. (1968), *J. Amer. Chem. Soc. 90,* 5014.
Nichols, J. L., and Lane, B. G. (1966), *Biochim. Biophys. Acta 119,* 649.
Nishimura, S. (1972), *Progr. Nucl. Acid Res. Mol. Biol. 12.* 49.
Pochon, F., and Michelson, A. M. (1969), *Biochim. Biophys. Acta 182,* 17.
Pörschke, D., and Eggers, F. (1972), *Eur. J. Biochem. 26,* 490.
Ringertz, H. G. (1972), *in* The Purines: Theory and Experiment, (Proceedings of the Fourth Jerusalem Symposium), Jerusalem, Israel Academy of Sciences and Humanities, p 61.
Saneyoshi, M., Harada, F., and Nishimura, S. (1969), *Biochim. Biophys. Acta 190,* 264.
Schellman, J. A., and Schellman, C. (1964), *in* The Proteins, 2nd ed., Vol. 2, Neurath, H., Ed., New York, N. Y., Academic Press, Chapter 7, p 1.
Shapiro, R. (1968), *Progr. Nucl. Acid Res. Mol. Biol. 8,* 73.
Shiao, D. D. F., and Sturtevant, J. M. (1973), *Biopolymers 12,* 1829.
Shoup, R. R., Becker, E. D., and McNeel, M. L. (1972b), *J. Phys. Chem. 76,* 71.
Shoup, R. R., Becker, E. D., and Miles, H. T. (1971), *Biochem. Biophys. Res. Commun. 43,* 1350.
Shoup, R. R., Miles, H. T., and Becker, E. D. (1966), *Biochem. Biophys. Res. Commun. 23,* 194.
Shoup, R. R., Miles, H. T., and Becker, E. D. (1972a), *J. Phys. Chem. 76,* 64.
Sobell, H. B. (1973), *Advan. Genet. 17,* 411.
Solie, T. N., and Schellman, J. A. (1968), *J. Mol. Biol. 33,* 61.
Sternglanz, H., and Bugg, C. E., (1973), *Science 182,* 833.
Sundaralingam, M., and Arora, S. K. (1969), *Proc. Nat. Acad. Sci. U. S. 64,* 1021.
Ts'o, P. O. P., Kondo, N. S., Schweizer, M. P., and Hollis, D. P. (1969), *Biochemistry 8,* 997.
Van Geet, A. L. (1970), *Anal. Chem. 42,* 679.
Voet, D., and Rich, A. (1970), *Prog. Nucl. Acid Res. Mol. Biol. 10.* 183.
Vold, R. L., Waugh, J. S., Klein, M. P., and Phelps, D. E. (1968), *J. Chem. Phys. 48,* 3831.
von Hippel, P. H., and McGhee, J. D. (1972), *Annu. Rev. Biochem. 41,* 233.