Spatial Configuration of the Bizarre 5′ Terminus of Mammalian mRNA[1]

Chong H. Kim and Ramaswamy H. Sarma*

Contribution from the Institute of Biomolecular Stereodynamics, State University of New York at Albany, Albany, New York 12222. *Received July 23, 1977*

Abstract: In order to obtain information about the spatial configuration of the mRNA 5′ terminus detailed nuclear magnetic resonance studies of $m^7G^{5'}ppp^{5'}Am$, ApA, AmpA, m^7G, $m^7G^{5'}p$, $m^7G^{5'}pp$, $m^7G^{5'}ppp$, m^7I, $m^7I^{5'}p$, and $m^6A^{5'}p$ and the corresponding nonmethylated derivatives were undertaken. NMR studies of 2′-O-methylated and the corresponding nonmethylated derivatives of various purine and pyrimidine nucleosides as well as their mono-, di-, and triphosphate derivatives were also undertaken. The data indicate that the ribose rings of both 7-methylated and nonmethylated guanosine derivatives exist as a ${}^2E \rightleftarrows {}^3E$ equilibrium mixture. However, the 7-methylation causes a shift in ribose pucker toward ${}^3E$ conformation. Irrespective of whether mono-, di-, or triphosphate, or whether 7-methylated or not, all the guanosine nucleotides show an overwhelming preference for g′g′ ($\phi_c = 180°$) orientation about C(5′)–O(5′). The compounds $m^7G^{5'}p$, $m^7G^{5'}pp$, and $m^7G^{5'}ppp$ exist with almost exclusive preference for the gg ($\psi = 60°$) orientation about C(4′)–C(5′), whereas the nonmethylated derivatives exist as a conformational mixture in which gg populations range from 67 to 77%, the remaining being the alternate gt and tg conformers. This is the first observed case of a nucleotide in aqueous solution with a rigid geometry about C(4′)–C(5′). The chemical-shift data for the ribose protons are consistent with two sets of conformation for the phosphodiester linkages. In one set $\omega c_1 \simeq 160°$, $\omega' c_1 \simeq 60°$, $\omega c_2 \simeq 210°$, and $\omega' c_2 \simeq -20°$; in another set $\omega c_1 \simeq 60°$, $\omega c_2 \simeq 20°$, and $\omega' c_2 \simeq 90°$. The 2′-O-methylated monomers exist as a conformational blend in which ${}^2E$ pucker, gg, and g′g′ conformations are preferred indicating that 2′-O-methylation has little effect on conformation of a monomer. However, at the level of a 3′,5′ dimer the 2′-O-methylation causes perturbation of various backbone torsion angles so much so the dimer assumes an extended conformation. The coupling constant data from the 5′,5′ dimer $m^7G^{5'}ppp^{5'}Am$ and constituent monomers indicate that with respect to the local conformations about the sugar ring, C(4′)–C(5′), and C(5′)–O(5′) bonds, the monomers essentially conserve their conformations in the dimer. However, the dimerization chemical-shift data and effect of temperature on chemical shift unmistakably show that the two nucleotidyl units at the 5′ ends of the triphosphate bridge clearly interact with each other as probably in a parallel stacking. This intramolecular stacking interaction must be made feasible by torsional variation around the phosphodiester bond of the triphosphate bridge. Search in the conformation space about the various phosphodiester linkages was undertaken by converting the torsion angles to cylindrical coordinates which were then used to compute the theoretical ring current shifts. The finally derived conformational details for $m^7G^{5'}ppp^{5'}Am$ in the most preferred stacking arrangement are as follows: $\chi_c = 30°$, ribose of $m^7G^{5'}p$- = ${}^3E$, $\psi_c = 60°$, $\phi_c = 180°$, $\omega c_1 = 155°$, $\omega' c_1 = 265°$, $\omega c_2 = 35°$, $\omega' c_2 = 155°$, $\omega c_3 = 295°$, $\omega' c_3 = 112°$, $\phi_1 = 180°$, $\psi_1 = 60°$, $\phi'' = 180°$, ribose of -$p^{5'}Am$ = ${}^2E$, and $\chi_1 = 30°$. Integration of this detailed conformation into the molecular framework of AmpA was attempted to generate the spatial configuration of mRNA 5′-terminus $m^7G^{5'}ppp^{5'}AmpA$. The only stereochemical recourse that was possible to accommodate the intrinsic conformational properties of AmpA and $m^7G^{5'}ppp^{5'}Am$ was to propose a stacked self-intercalating model in which the $m^7G^{5'}p$- part intercalates in between the bases of AmpA. However, contributions from nonintercalating stacked and random oriented conformations for the segment $m^7G^{5'}ppp^{5'}AmpA$ could not be excluded.

The 5′-terminal regions of a wide variety of mRNAs of eucaryotes contain an unusual sequence of nucleosides in which the terminal is 7-methylguanosine.[2] In this structure the 5′-hydroxyl of m^7G is connected through a triphosphate to the 5′-hydroxyl of 2′-O-methylated nucleosides.[2] In Figure 1 is illustrated the observed sequence in various viral and cellular mRNAs including those of cytoplasmic polyhedrosis virus,[3] reovirus,[4,5] vaccinia[6,7] and vesicular somatitis virus (VSV),[8] rabbit reticulocyte globin,[9] HeLa cells,[10] mouse myeloma,[11] and mouse L cells.[9] Muthukrishnan et al.[9,12] and Both et al.[13] have shown that reovirus, VSV, and rabbit reticulocyte mRNAs lose their ability to stimulate protein synthesis if they do not contain the 5′-terminal sequence shown in Figure 1 and that incorporation of the above sequence in these systems restores their ability to carry out in vitro translation in protein synthesis with fidelity.

No rationale is available for the obligatory requirement of m^7G terminal[14] and 2′-O-methylated residues in eucaryotic mRNAs to carry out protein synthesis. Shafritz et al.[15] showed that $m^7G^{5'}p$ specifically inhibited the binding of initiation factor M_3 (IF-M_3) to HeLa cell histone mRNA and vesicular somatitis mRNA. This suggests the possibility that the unusual termini of eucaryotic mRNAs may be involved in an interaction with the initiation factors during the initiation of protein synthesis. Studies in this laboratory and that of Baglioni[16,17] have shown that G, $G^{5'}p$, $G^{5'}pp$, $G^{5'}ppp$, m^7G, and 7-methylated guanosine 2′- or 3′-monophosphates do not inhibit protein synthesis in a wheat-germ cell-free system programmed by exogenous mRNA. However, the mRNA translation was inhibited severely by $m^7G^{5'}p$, $m^7G^{5'}pp$, and $m^7G^{5'}ppp$, the inhibition being higher for the di- and triphosphates. Insofar as the inhibitors resemble the 5′terminus of mRNAs, structurally it appears reasonable to postulate that they and mRNA compete for the initiation factor during the assembly of proteins. In this report we attempt to unravel the possible conformational properties of the 5′ termini of mRNAs by investigating the solution conformations of: (i) 7-methylated guanosine derivatives vis-à-vis the corresponding nonmethylated analogues; (ii) 2′-O-methylated nucleosides and the corresponding 5′-nucleotides vis-à-vis the nonmethylated analogues; (iii) the dimer $m^7G^{5'}ppp^{5'}Am$.

The data so derived in conjunction with that on 2′-O-methylated dinucleoside monophosphates[18,19] are used to propose a possible conformational model for the 5′ terminus of eucaryotic mRNAs.

Experimental Section

Materials. Guanosine, $G^{5'}p$, $G^{5'}pp$, and $G^{5'}ppp$ were purchased from Sigma Chemical, Inc.; m^7G, m^7I, and $m^7I^{5'}p$ were purchased from Terra-Marine Bioresearch (La Jolla, Calif.). The $m^7G^{5'}p$, $m^7G^{5'}pp$, and $m^7G^{5'}ppp$ and the various 2′-O-methylated derivatives and $m^7G^{5'}ppp^{5'}Am$ were obtained from P-L Biochemicals. All the samples were dissolved and lyophilized three times in 99.8% D_2O with a trace of ethylenediaminetetraacetic acid (EDTA). Final sample preparations were performed in 100% D_2O (Bio-Rad) and the pH was adjusted to 7.0 or 5.0 depending upon experimental conditions.

0002-7863/78/1500-1571$01.00/0

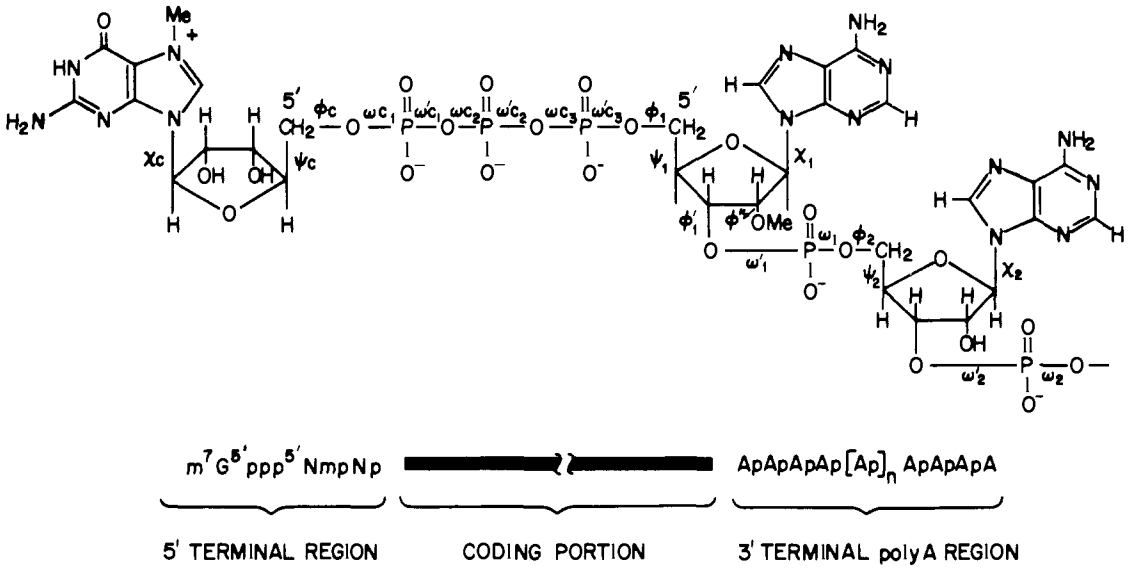

Figure 1. The sequence of mRNA at the 5′ region. The above sequence can be abbreviated as $m^7G^{5'}ppp^{5'}AmpA$. Structures in which AmpA has been replaced with any of the common bases have been observed in mRNA. The conformational nomenclature χ, ψ, ϕ, ω, ϕ', ϕ'', etc., follows definitions according to IUPAC–IUB. The letter c is inserted to distinguish between the cap region and the AmpA region. Thus, χ_c stands for sugar base torsion of $m^7G^{5'}p$- and χ_1 and χ_2 stand for the Amp- and -pA parts, respectively. Nonstandard abbreviations such as $G^{5'}p$, $G^{5'}pp$, etc., are used for 5′-GMP, 5′-GDP, etc., in accordance with their wide use in the original papers.

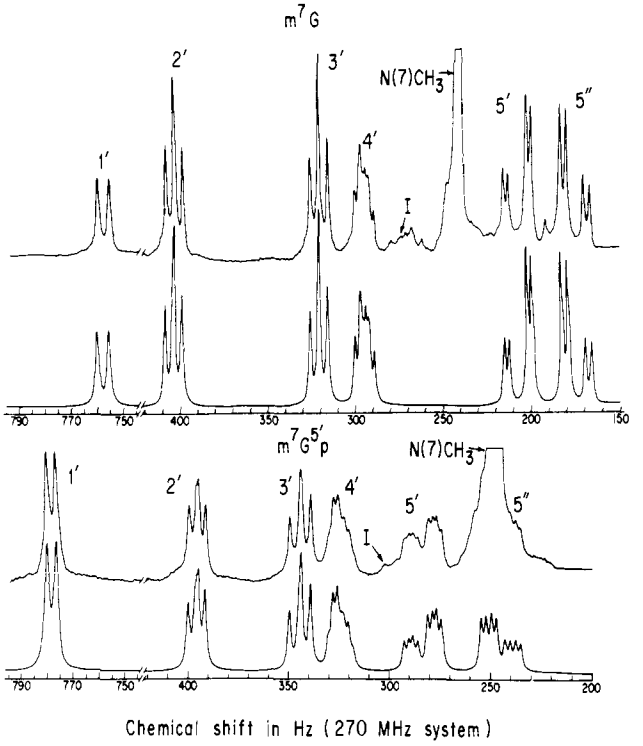

Figure 2. The 270-MHz ^1H NMR spectra, observed and computer simulated, of m^7G and $m^7G^{5'}p$. The shifts are in hertz (270 MHz) from internal TMA; pH 5.0; 19 ± 1 °C.

Measurement of Spectra. ^1H NMR spectra were recorded at 270 MHz using a Brucker-270 HX system equipped with variable temperature assembly and a BNC 12 data system capable of performing 16K transforms. The details of the spectrometer system are given elsewhere.[20,21] The signal positions were measured relative to the internal standard, tetramethylammonium chloride (TMA). ^{31}P NMR measurements were made on $G^{5'}p$, $G^{5'}pp$, $G^{5'}ppp$, and the corresponding 7-methylated analogues using systems described elsewhere.[22,23] ^{31}P NMR spectra of $m^7G^{5'}ppp^{5'}Am$ were not taken because of the lack of availability of the material in sufficient quantities.

Assignments of Spectra. The assignments of the various resonances were intially made by homo- and heteronuclear decoupling as well as from comparison of the spectra from the various derivatives. The initial set of data so obtained was then refined by carrying out simulations using LAOCOON III iterative computer programs.

Results and Discussion

In presenting the results we discuss first the conformational properties of 7-methylguanosine derivatives; second, those of 2′-O-methylated nucleic acid structures; third, those of 2′-O-methylated 3′,5′ dimers; fourth, those of the 5′,5′ dimer $m^7G^{5'}ppp^{5'}Am$; and finally, an integrated picture of the conformational properties of the 5′ terminus of mRNAs.

Conformational Properties of 7-Methylated Derivatives of Guanosine. A complete conformational description of 7-methylated derivatives of guanosine requires the determination of ribose ring conformations, the specification of dihedral angles about C(4′)-C(5′) (ψ_c), C(5′)-O(5′) (ϕ_c) and the various phosphodiester linkages ($\omega c_1/\omega' c_1 \ldots \omega c_3/\omega' c_3$), as well as determination of the relative orientation between the base and sugar (χ_c). In Figure 1 we have illustrated the conformational nomenclature employed.

In Figures 2 and 3 are illustrated the observed and computer-simulated ^1H spectra of m^7G, $m^7G^{5'}p$, $m^7G^{5'}pp$, and $G^{5'}pp$. The assignments are also indicated in the same figures. Table I contains the chemical-shift data for the various 7-methylated derivatives of guanonine and inosine. The coupling constant data are given in Table II.

Lee et al.[20] and Ezra et al.[21] have proposed a method to compute the population distribution of the C(3′)-endo (^3E) and C(2′)-endo (^2E) conformers of the ribose ring in nucleotides from the magnitudes of $J_{1'2'}$, $J_{3'4'}$, and the sum $J_{1'2'} + J_{3'4'}$. The standard sum $J_{1'2'} + J_{3'4'}$ suggested is 9.5 Hz. In the case of 7-methylated nucleotides the observed average sum is 8.4 Hz. The significant reduction in the magnitude of $J_{1'2'} + J_{3'4'}$ from the standard value of 9.5 Hz is a direct result of the presence of a formal (+) charge in the imidazole part of 7-methylated guanosine and the electronegative effect of the (+) charge on $J_{1'2'}$. From a comparison of β-NMN^+ and β-NMNH as well as β-NAD^+ and β-NADH, Sarma and Mynott[22,23] have earlier reported that a formal (+) charge in the neighborhood of C(1′) will lead to significant reduction in the

Table I. Chemical Shifts[a] of Guanosine and Inosine Derivatives in D_2O

Compds	pH	T (°C)	Chemical shifts (δ), Hz							
			1′	2′	3′	4′	5′	5″	H(8)	Me
G	7.2	20	734.4	416.3	328.3	281.9	187.4	169.5	1296.0	
m^7G	7.0	20	758.2	404.3	321.1	295.2	206.1	176.1		241.6
	7.0	80	763.9	410.5	326.0	298.3	205.0	178.1		250.4
$G^{5'}p$	8.0	20	741.4	434.1	353.7	306.2	221.9	214.4	1355.7	
	7.0	80	739.4	425.2	355.5	305.4	229.4	227.5	1335.8	
	5.0	20	741.9	420.6	349.4	311.6	249.5	249.5	1326.8	
$m^7G^{5'}p$	7.0	20	776.9	392.0	353.1	319.2	259.5	222.8		248.9
	7.0	80	781.1	399.1	352.1	323.0	259.6	233.7		255.9
	5.0	20	778.8	395.8	344.2	324.6	282.5	245.8		247.4
$G^{5'}pp$	7.0	20	736.8	414.2	380.3	311.1	284.0	276.0	1332.1	
	7.0	80	740.6	423.2	380.0	320.6	280.0	276.5	1326.6	
	5.0	20	738.9	423.5	364.8	316.3	280.8	275.1	1331.0	
$m^7G^{5'}pp$	7.0	20	771.7	392.8	374.1	321.6	315.0	289.0		246.7
	7.0	80	779.6	402.8	369.4	327.0	302.5	289.0		257.0
	5.0		775.4	401.9	357.7	328.8	314.8	279.1		250.4
$G^{5'}ppp$	7.0	20	737.2	428.0	381.4	317.4	289.5	282.5	1336.1	
	7.0	80	740.4	431.8	380.6	315.8	290.5	286.0	1326.6	
	5.0	20	739.5	427.1	371.5	319.0	289.7	285.5	1336.5	
$m^7G^{5'}ppp$	7.0	20	774.2	403.9	378.0	324.0	314.0	303.0		250.4
	7.0	80	780.1	411.0	374.5	330.5	302.5	296.5		258.1
	5.0	20	776.1	404.9	364.3	331.0	320.7	290.8		251.8
m^7I	7.0	20	802.6	410.9	323.9	305.3	214.3	181.0	1372.2	272.1
$m^7I^{5'}p$	7.0	20	818.2	393.6	353.3	326.0	265.0	225.3	1371.1	278.0
	7.0	80	822.7	402.2	354.8	329.6	276.8	236.0	1375.1	282.7
	5.0	20	834.0	404.1	350.3	335.0	289.4	241.8	1397.5	283.8

[a] Shifts are given relative to internal TMA and are accurate to ±0.1 Hz. Concentration is 0.05 M except for G (0.004 M) and $G^{5'}p$ at pH 8.0 (0.01 M).

Table II. Coupling Constants[a] for Guanosine and Inosine Derivatives in D_2O[b]

Compds	pH	T (°C)	Coupling constants, Hz								
			1′2′	2′3′	3′4′	4′5′	4′5″	4′P	5′P	5″P	5′5″
G	7.2	20	5.9	5.2	3.7	3.2	3.8				−12.5
m^7G	7.0	20	4.5	5.2	4.9	2.7	3.6				−13.1
	7.0	80	4.5	5.3	5.0	3.1	4.1				−13.1
$G^{5'}$	8.0	20	6.1	5.2	3.4	4.3	2.9	1.1	4.7	4.7	−11.6
	7.0	80	5.8	5.2	3.8	4.3	4.5	0.9	5.1	5.3	−11.5
	5.0	20	6.0	5.1	3.5	3.4	3.4	2.0	5.1	5.1	−11.8
$m^7G^{5'}p$	7.0	20	3.6	4.9	5.1	2.3	2.0	1.7	4.5	4.5	−12.0
	7.0	80	4.1	5.0	4.9	3.8	3.0	2.0	5.4	5.2	−12.2
	5.0	20	3.5	4.9	5.7	2.4	2.5	1.9	4.4	5.2	−11.8
$G^{5'}pp$	7.0	20	5.4	5.3	4.2	3.5	3.5	1.7	5.5	4.8	−12.0
	7.0	80	5.5	5.2	4.1	3.7	3.7	1.5	5.3	5.3	−12.5
	5.0	20	6.0	5.2	3.5	3.5	3.3	1.6	4.8	4.8	−12.0
$m^7G^{5'}pp$	7.0	20	2.4	4.9	6.2	1.7	1.9	1.7	4.0	5.3	−12.2
	7.0	80	3.1	5.0	5.0	2.7	2.7	2.0	4.4	5.0	−12.0
	5.0	20	2.5	5.0	5.0	1.8	1.8	1.7	3.7	5.0	−11.6
$G^{5'}ppp$	7.0	20	6.1	5.0	3.4	3.1	3.1	1.8	5.8	5.0	−12.5
	7.0	80	5.8	5.3	3.6	3.5	3.5	1.5	5.2	4.9	−12.5
	5.0	20	6.1	5.3	3.5	3.0	3.0	1.8	5.5	5.0	−12.3
$m^7G^{5'}ppp$	7.0	20	2.7	5.1	5.7	1.7	2.0	1.7	4.2	5.3	−12.4
	7.0	80	3.8	5.0	4.8	2.5	2.5	2.0	4.8	5.3	−12.0
	5.0	20	2.8	5.0	4.8	2.0	2.2	1.7	3.8	5.3	−12.0
m^7I	7.0	20	4.2	5.1	5.4	2.6	3.6				−13.0
$m^7I^{5'}p$	7.0	20	3.4	4.8	5.4	2.3	2.2	1.6	3.9	4.7	−11.5
	7.0	80	3.9	4.9	4.9	2.8	2.5	1.9	6.5	5.3	−12.6
	5.0	20	3.1	5.0	5.5	2.1	2.1	1.9	5.2	4.8	−12.0

[a] Coupling constants are accurate to ±0.1-0.2 Hz. [b] Solution conditions are the same as in Table I.

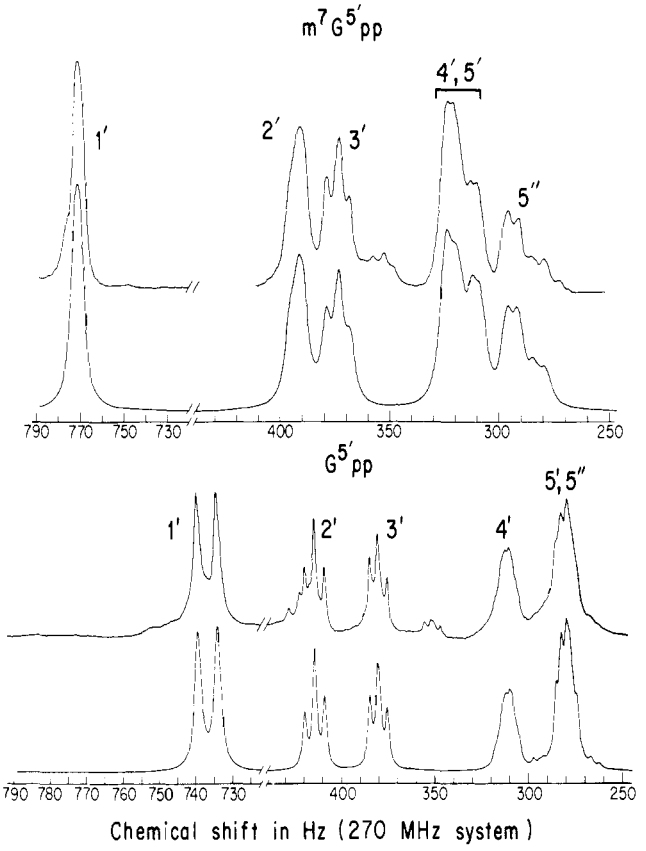

Figure 3. Observed and computer-simulated 270-MHz ^{1}H NMR spectra of m^{7}G$^{5'}$pp and G$^{5'}$pp, pH 7,0. The rest of the details are as in Figure 2.

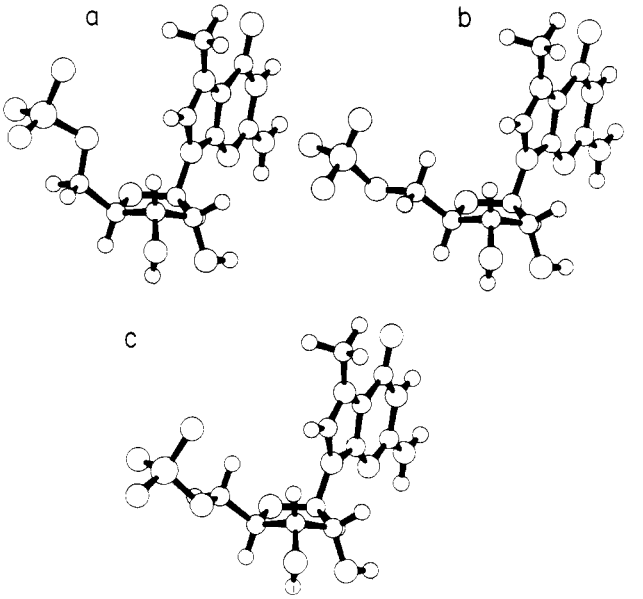

Figure 4. Conformations of m^{7}G$^{5'}$p: $\chi_c = 60°$, and $\phi_c = 180°$. In (a) $\psi_c = 60°$; in (b) $\psi_c = 180°$; and in (c) $\psi_c = 300°$.

value of $J_{1'2'}$. Obviously, one does not expect the value of $J_{3'4'}$ to be affected significantly by changes in the electronegativity characteristics of the base. Hence in the present study $^3E/^2E$ ratios for 7-methylated nucleotides were calculated using the values of $J_{3'4}$ and the standard sum of 9.5 Hz for $J_{1'2'} + J_{3'4'}$. For the nucleosides investigated ribose conformational distribution was computed using a sum of 9.65 Hz (average value in adenosine and guanosine). The results are summarized in Table III.

The data show that, in general, common nucleosides and nucleotides G, G$^{5'}$p, G$^{5'}$pp, and G$^{5'}$ppp at the examined pH and temperature ranges exist as a $^2E \rightleftarrows {}^3E$ equilibrium mixture with a bias of $\simeq$60% for the ^2E conformer. It is important to note that monophosphorylation at the 5′ end of guanosine has no significant effects on the population distribution of ribose conformers. However, diphosphorylation causes a modest increase in ^3E population which disappears on triphosphorylation.

Comparison of the ^3E population for nonmethylated and methylated nucleosides clearly reveals that 7-methylation has an important influence on ribose ring conformational distribution. The data indicate that 7-methylation causes a 13–24% increase in ^3E conformer population and the increase closely parallels the number of phosphate groups introduced, i.e., 13% increase in the nucleoside and 24% in the nucleoside triphosphate. The data indicate that m^{7}G$^{5'}$pp and m^{7}G$^{5'}$ppp display a preference to exist in the ^3E conformation and this is the first observed case in which purine 5′-nucleotides show preference for ^3E pucker in *aqueous solution.* The 5′-phosphorylation on m^{7}I has no effect on ribose conformation contrary to the m^{7}G systems. In the case of 7-methylated di- and triphosphates, elevation of temperature tends to destabilize the ^3E conformers by about 10% with no effect on the monophosphates and the nucleoside. In the case of nonmethylated derivatives temperature has no effect on ribose conformer distribution. Changing the ionization state of the phosphate by lowering the pH to 5.0 causes a slight shift to ^3E pucker for m^{7}G$^{5'}$p. In the case of the di- and triphosphates, this had an opposite effect, i.e., the ^2E population increases. With respect to the common derivatives only G$^{5'}$pp is sensitive to ionization changes. The observed differences in the ribose conformational equilibrium between the guanosine and 7-methylguanosine series as well as the sensitivity of ribose ring conformation in the 7-methyl series to temperature and ionization states of the phosphate are probably related to χ_c and ψ_c variations and this will be discussed later.

The population distribution of conformers about C(4′)–C(5′) (ψ_c) and C(5′)–O(5′)(ϕ_c) is computed from Σ ($J_{4'5'} + J_{4'5''}$) and Σ' ($J_{5'p} + J_{5''p}$) using the expressions in Lee and Sarma[26] and the data are presented in Table III. The data indicate that in all the systems examined, irrespective of constitutional modifications, the C(5′)–O(5′) bond displays an outspoken preference for g′g′ orientation ($\phi_c = 180°$). The orientation about C(5′)–O(5′) is generally insensitive to the state of ionization of the phosphate and temperature except m^{7}I$^{5'}$p which shows a temperature dependence (Table III).

However, the 7-methylated derivatives display a considerably different degree of conformational preference about C(4′)–C(5′) compared to the nonmethylated analogues. The observed 20–25% increase in gg population ($\psi_c = 60°$) upon 5′-phosphorylation of 7-methylguanosine is possibly the result of the electrostatic interaction between the positive charge on N(7) and the negative charges on phosphate groups. In the gg ($\psi_c = 60°$) conformation the distance between the phosphate oxygens and the N(7) is small compared with that in the alternate gt ($\psi_c = 180°$) and tg ($\psi_c = 300°$) conformers. This is illustrated in the perspective in Figures 4a–c. It is important to note that m^{7}G$^{5'}$pp, and m^{7}G$^{5'}$ppp show overwhelming preference for gg orientation about C(4′)–C(5′) and this is the first finding in which a nucleotide shows a rigid geometry about C(4′)–C(5′) in *aqueous solution.* Also, this is a drastic contrast to the case of G$^{5'}$p, G$^{5'}$pp, and G$^{5'}$ppp which show a conformational blend about C(4′)–C(5′), preferring the gg population 67–77% and the remaining being the alternate gt and tg conformers. This points out the impact of 7-methylation on the conformational freedom about the C(4′)–C(5′) bond in the

Table III. Conformational Parameters for Guanosine and Inosine Derivatives[a]

Compds	pH	T (°C)	% conformational preferences about various bonds: C(5′)–O(5′) g′g′ ⇄ g′/t′[b]		C(4′)–C(5′) gg ⇄ g/t[c]		Ribose, ^3E
G	7.2	20			69	31	38
m^7G	7.0	20			76	24	51
	7.0	80			67	33	52
$G^{5'}p$	8.0	20	75	25	67	33	36
	7.0	80	70	30	51	49	40
	5.0	20	71	29	71	29	37
$m^7G^{5'}p$	7.0	20	77	23	97	3	54
	7.0	80	69	31	71	29	52
	5.0	20	74	26	91	9	60
$G^{5'}pp$	7.0	20	71	29	69	31	44
	7.0	80	69	31	65	35	43
	5.0	20	74	26	71	29	37
$m^7G^{5'}pp$	7.0	20	76	24	100	0	65
	7.0	80	75	25	95	5	53
	5.0	20	78	22	100	0	53
$G^{5'}ppp$	7.0	20	68	32	77	23	36
	7.0	80	72	28	69	31	38
	5.0	20	70	30	79	21	37
$m^7G^{5'}ppp$	7.0	20	75	25	100	0	60
	7.0	80	72	28	96	4	51
	5.0	20	76	24	100	0	51
m^7I	7.0	20			77	23	56
$m^7I^{5'}p$	7.0	20	79	21	95	5	57
	7.0	80	64	36	87	13	52
	5.0	20	72	28	98	2	58

[a] Solution conditions are the same as in Table I. [b] Combined population of g′t′ (ϕ = 60°) and t′g′ (ϕ = 300°). [c] Combined population of gt (ψ = 180°) and tg (ψ = 300°).

guanosine 5′-nucleotides. The data on m^7I and $m^7I^{5'}p$ (Table III) indicate that the influence of 7-methylation on the conformational freedom about the C(4′)–C(5′) bond is a general phenomenon for all the 5′-purine nucleotides. The observed small difference in the gg population between 7-methylguanosine and guanosine may originate from a difference in the O(5′)–H···N(3) hydrogen bonding which is possible in syn conformers.[27] Elevation of temperature tends to destabilize the gg population in both $m^7G^{5'}p$ and $G^{5'}p$ with less noticeable trends in the case of di- and triphosphates. Ionization of the phosphate group does not have any pronounced effect on the conformational freedom about C(4′)–C(5′) bond rotation in methylated and nonmethylated systems.

The data show that 5′-phosphorylation of the nucleoside has no effect on the conformational preference about C(4′)–C(5′) in the guanosine series, again suggesting that nucleotides are not any more rigid than nucleosides, contrary to the thesis of Sundralingam.[24,25] The observation of rigid conformation about C(4′)–C(5′) of $m^7G^{5'}p$, $m^7G^{5'}pp$, and $m^7G^{5'}ppp$ is a reflection of the intrinsic and genuine physical properties of these molecules which enable an electrostatic interaction between the base and the phosphate groups.

In order to monitor the electrostatic interaction we have obtained ^{31}P-{^1H} NMR spectra of the pairs $G^{5'}p$ and $m^7G^{5'}p$, $G^{5'}pp$ and $m^7G^{5'}pp$, and $G^{5'}ppp$ and $m^7G^{5'}ppp$. The data are summarized in Table IV. In the case of triphosphates the γ-phosphate shifts to higher field by 23 Hz; the β- and α-phosphates also shift to higher fields by 15 and 10 Hz, respectively. Similar observations are also noticeable for the pair $G^{5'}pp$ and $m^7G^{5'}pp$. The direction of the shift is that expected for electrostatic interaction as has been shown by Sarma and Mynott.[28] It is surprising that the data for $G^{5'}p$ and $m^7G^{5'}p$ (Table IV) do not show a substantial difference in chemical shift of ^{31}P nuclei. This, we believe, is due to the internal compensation of the upfield shift by comparable downfield shift which can originate from torsional changes which accompany 7-methylation of $G^{5'}p$. It should be noted that the conformational changes that accompany the 7-methylation of $G^{5'}p$, $G^{5'}pp$, and $G^{5'}ppp$, though in general similar, are not identical.

Table IV. Chemical-Shift Data (Hz)[a] of Phosphorus in Guanosine Derivatives

Compds	α	β	γ
$G^{5'}p$	−36[b]		
$m^7G^{5'}p$	−34[b]		
$G^{5'}pp$	556	410	
$m^7G^{5'}pp$	566	428	
$G^{5'}ppp$	566	999	400
$m^7G^{5'}ppp$	576	1014	423

[a] Shifts are given relative to internal trimethyl phosphate, $(CH_3O)_3OP$. Concentration is 0.05 M (pH 7.0) at 19 ± 1 °C; 40.5-MHz NMR system. [b] At pH 8.9.

The observation that only C(4′)–C(5′) changes with relatively no effect on C(5′)–O(5′) upon 7-methylation supports the thesis of Lee and Sarma[29] that "*rotation of C(4′)–C(5′) bond is of primary importance in reducing the stress while that of C(5′)–O(5′) bond is secondary during the reorientation of the backbone conformation*". It may be noted that 7-methylation causes an increase in the population of ^3E and gg conformers (Table III) supporting a conformational interdepen-

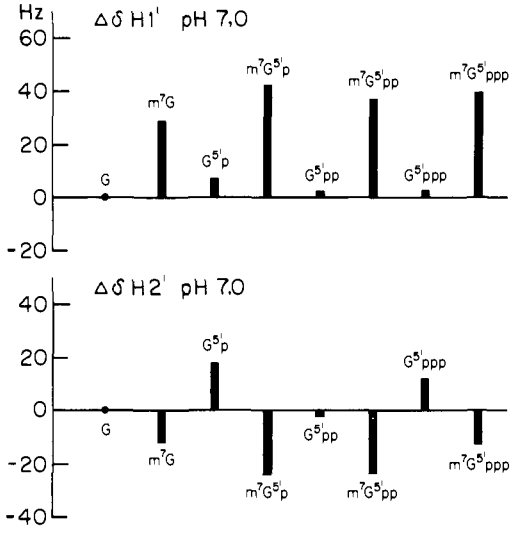

Figure 5. Stick diagrams indicating the difference in chemical shifts for H(1′) and H(2′) among the various derivatives of guanosine. The H(1′) and H(2′) of guanosine were given a value of $\delta = 0$. The $\Delta\delta$s are in hertz (270 MHz).

dence between the two as has been suggested by Hruska[30] for mononucleotides. However, it has been shown elsewhere[18-21] that in dinucleoside monophosphates the ribose conformational distribution is influenced by χ_{CN} changes rather than changes in the orientation about C(4′)–C(5′) bonds. This is discussed in the next section.

X-ray, NMR, and theoretical calculations on mono- and dinucleoside monophosphates have shown that the preferred glycosidic torsion in these systems lies in the anti domain.[18-21,31-38] NMR data do not allow a quantitative evaluation of the magnitude of χ_{CN} or that of the population distribution of the syn and anti conformers. However, we attempt below to determine the relative differences in χ_{CN} among the common nucleosides, nucleotides, and their N7 methylated analogues.

In Figure 5, we have presented the difference in chemical shifts for H(1′) and H(2′) of the various compounds investigated assuming that the chemical shifts of these protons in guanosine are zero. The chemical shifts of H(1′) and H(2′) are sensitive to χ_{CN} variations.[39,40] Comparison of shift data (Figure 5) for m^7G and $m^7G^{5'}p$ indicates that the introduction of the phosphate group causes downfield shift of H(1′) by 19 Hz and upfield shift of H(2′) by 12 Hz. Comparison of the shift data for $G^{5'}p$ and $m^7G^{5'}p$ (Figure 5) indicates that 7-methylation causes downfield shift of H(1′) by 36 Hz and upfield shift of H(2′) by 42 Hz. These observed shifts are large in magnitude and clearly indicate that 7-methylation and 5′-phosphorylation really affect the magnitude of χ_{CN}. This is entirely expected because such a substitution can lead to electrostatic interaction between N(7)(+) and phosphate negative charges, which in turn will restrict the flexibility about the glycosidic linkage. It should be noted that part of the shift differences can originate from differences in the population distribution of ribose conformers in the various molecules compiled. Inspection of the χ_{CN} vs. chemical-shift difference profiles[39,40] indicates that approach to a χ_{CN} value of about 60° from either side (below 60° or above 60°) will be consistent with the observed shifts.

As can be seen in Figure 4, a χ_{CN} value of about 60° allows the maximum electrostatic interaction between N(7)(+) and negative charges on the phosphate. This conclusion is in agreement with the observation that upon 5′-phosphorylation of m^7G the population of ^3E conformers increases.[39,40] This is probably due to the fact that m^7G exists as an equilibrium blend of conformers in the syn and anti domains and the observed increase of the population of the ^3E conformers is an indication of a conformational purification from a syn and anti mixture to a predominantly anti system.[27] The differences in the chemical shifts of H(1′) and H(2′) among $m^7G^{5'}p$, $m^7G^{5'}pp$, and $m^7G^{5'}ppp$, though small, are very real and primarily due to torsional adjustment of the glycosidic bond so that N(7)(+) could effectively interact with the negative charges on the mono-, di-, and triphosphates. Partial neutralization of the negative charge by lowering the pH to 5.0 as well as elevation of temperature causes shifts of H(1′) and H(2′) to lower field in $m^7G^{5'}p$, $m^7G^{5'}pp$, and $m^7G^{5'}ppp$, the effect being very noticeable in di- and triphosphates. These changes are in line with the previous observation of shifts in ribose population from ^3E to ^2E upon elevation of temperature

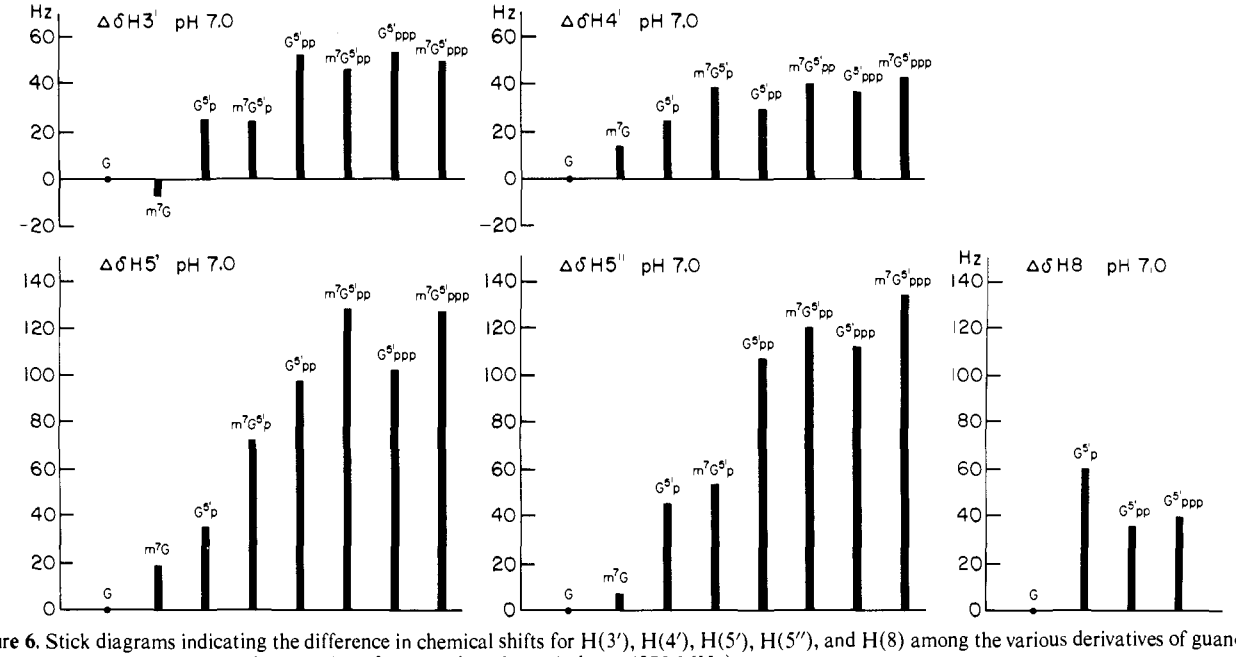

Figure 6. Stick diagrams indicating the difference in chemical shifts for H(3′), H(4′), H(5′), H(5″), and H(8) among the various derivatives of guanosine. The protons of guanosine were given a value of $\delta = 0$. The $\Delta\delta$s are in hertz (270 MHz).

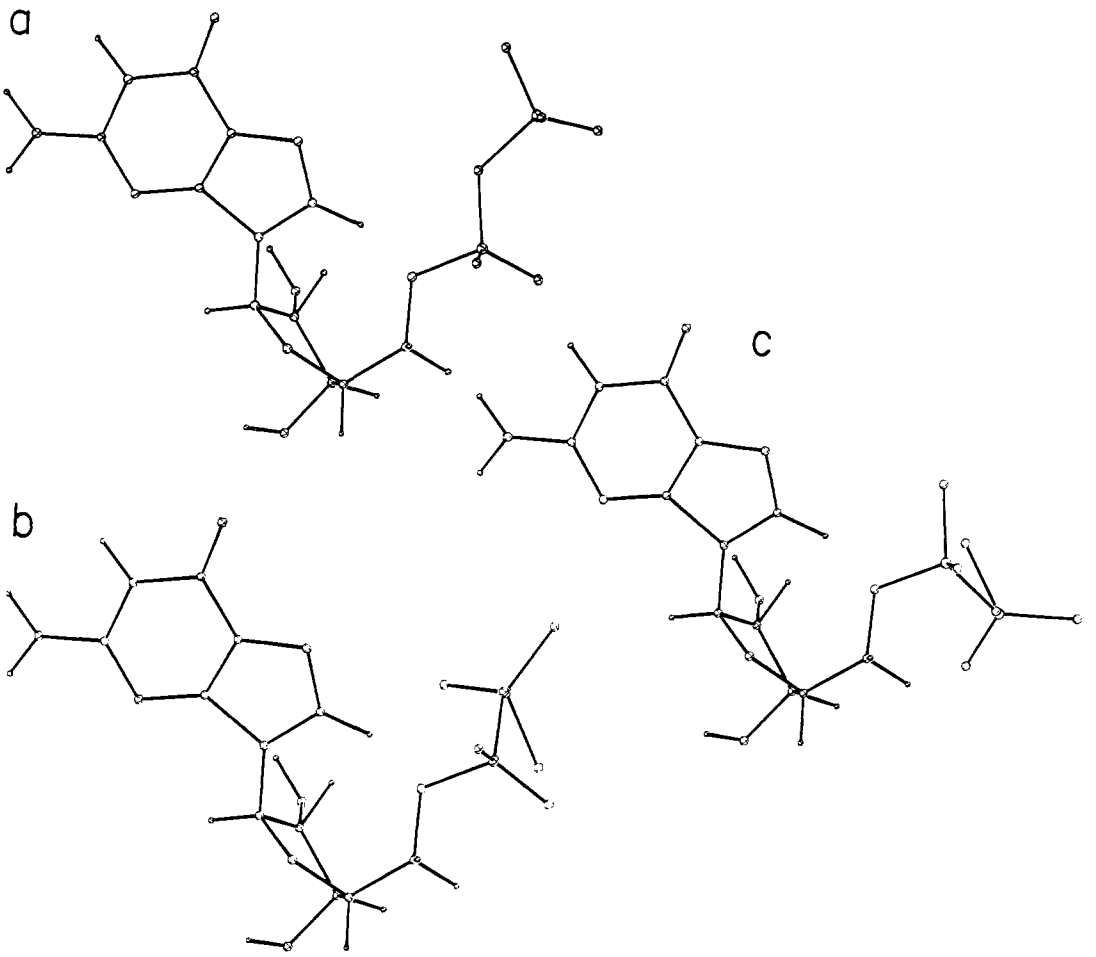

Figure 7. Various perspectives of $G^{5'}pp$. In all perspectives $\chi = 45°$, ribose = ^2E, $\psi = 60°$, and $\phi = 180°$. In (a) the pyrophosphate is all trans, in (b) $\omega c_1 = 160°$ and $\omega' c_1 = -60°$, and in (c) $\omega c_1 = \omega' c_1 = 60°$.

and the protonation of phosphate group. In accordance with the conformational nexus between χ_{CN} and ribose conformation,[20,21] a change of ribose conformation will invariably cause perturbation of χ_{CN} and this in turn reflects on the chemical shifts of H(1′) and H(2′). A comparison of the chemical shifts of H(1′) and H(2′) of G and $G^{5'}p$ indicates that H(1′) and H(2′) are both deshielded upon 5′-phosphorylation. This observation indicates that in going from nucleosides to nucleotides the population of the conformers in the anti domain increases. In going from monophosphate to diphosphate, H(1′) undergoes small shielding and H(2′) undergoes noticeably larger shielding. These are due to a shift in ribose population to ^3E and a small reduction of value of χ_{CN}. In going from diphosphate to triphosphate both H(1′) and H(2′) undergo deshielding, the effect being very noticeable for H(2′). This is entirely expected because in going from $G^{5'}pp$ to $G^{5'}ppp$ the ^3E ribose population decreases.

Conformation of the phosphate backbone is defined by the torsion angles ωc_1, $\omega' c_1$, ωc_2, and $\omega' c_2$. Information about these torsion angles cannot be obtained directly because of the lack of available nuclei which will produce the desired coupling constants. The conclusions have to be reached from the chemical-shift trends of H(5′), H(5″), H(4′), H(3′), and H(8) and hence are qualitative. Furthermore, H(8) and H(3′) will show some small dependence on χ_{CN}.

The examination of data (Figure 6) from $G^{5'}p$ and $G^{5'}pp$ at pH 7, room temperature, indicates that the second phosphorylation causes the downfield shift of H(5′), H(5″), H(4′), and H(3′) and upfield shift of H(8). The effect on H(5′), H(5″), and H(4′) is understandable and is mostly due to the through-the-bond effect of introducing the second phosphate group. The effect on H(3′) and H(8) is an effect through the space. In going from $G^{5'}p$ to $G^{5'}pp$ one neutralizes a negative charge on the α-phosphate and this neutralization causes the upfield shift of H(8). This conclusion is corroborated by the observation that H(8) of $G^{5'}p$ undergoes a shift to a higher field of comparable magnitude upon lowering the pH. The observed downfield shift of H(3′) in going from $G^{5'}p$ to $G^{5'}pp$ can be rationalized only if the second phosphate group become spatially close to H(3′) which is not possible in an all-trans staggered conformation for the pyrophosphate (Figure 7a).

From previous discussion it is clear that the preferred conformation about ψ_c and ϕ_c is about 60 and 180°, respectively. One could generate two different pyrophosphate backbone geometries in which phosphate oxygens will be approximately juxtaposed to H(3′) and cause its deshielding. In one conformation the magnitudes of ωc_1 and $\omega' c_1$ are approximately 160 and −60° (Figure 7b). In another conformation both of these torsion angles are about 60° (Figure 7c). In both conformations the phosphate oxygens are staggered when viewed along the p–p axis. A significant population of the conformations shown in Figure 7b and 7c will also rationalize clearly the higher downfield shift of H(5″) compared to that of H(5′) upon phosphorylation of $G^{5'}p$. In going from $G^{5'}pp$ to $G^{5'}ppp$ the data show that the third phosphate group has *very little effect on the chemical shift of various protons described before.* If the third phosphate group is introduced in the conformation in Figure 7b so that ωc_2 is $\simeq 210°$ and $\omega' c_2$ is $\simeq -20°$ (Figure 8a), the chemical shift of H(3′) will not be affected by the phosphorylation of $G^{5'}pp$ to $G^{5'}ppp$. With respect to the con-

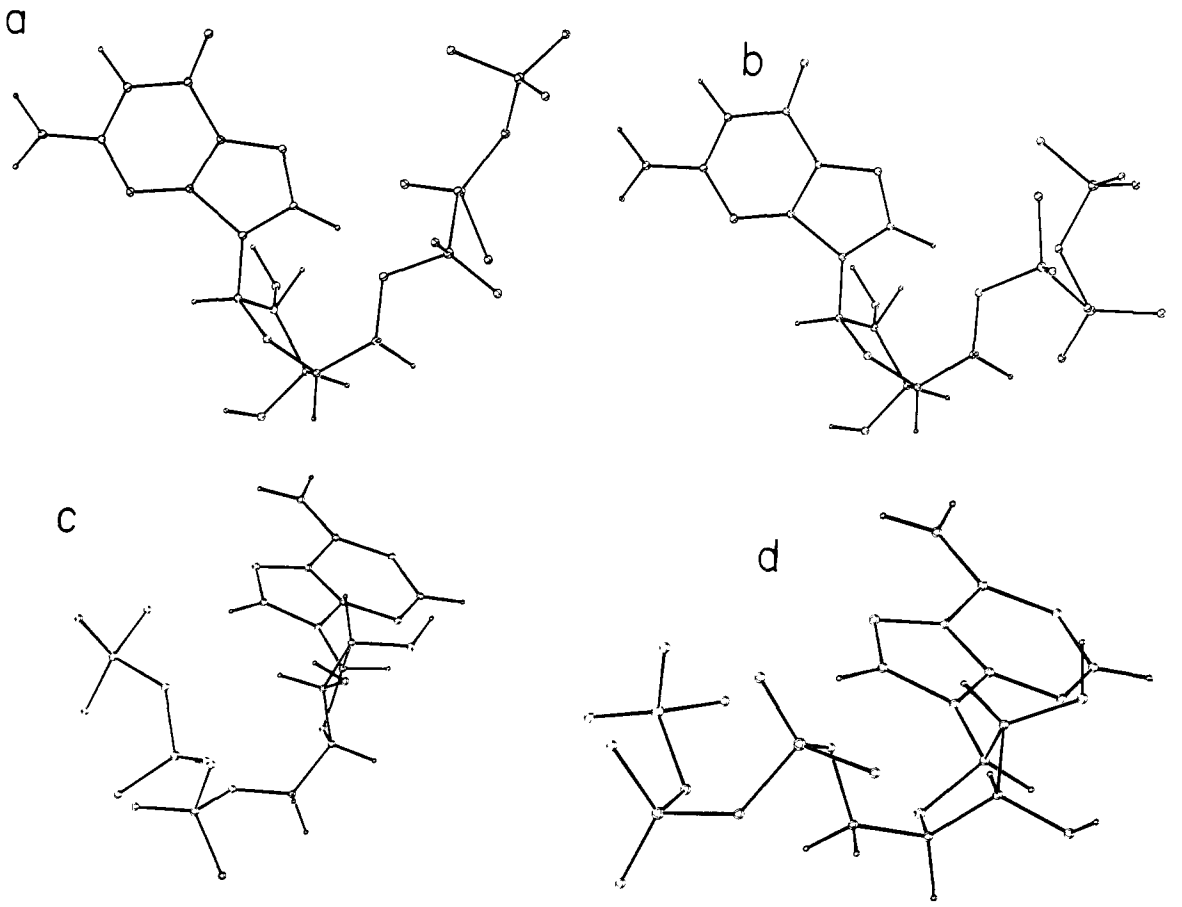

Figure 8. Various perspectives of $G^{5'}ppp$. (a) $\omega c_2 = 210°$, $\omega' c_2 = -20°$; the remaining torsion angles are the same as in Figure 7b. (b) $\omega c_2 = 20°$, $\omega' c_2 = 90°$; the remaining torsion angles are as in Figure 7c. Figures 8c,d are the perspectives of molecules A (^3E) and B (^2E) from x-ray data of Kennard et al. (1971).[41]

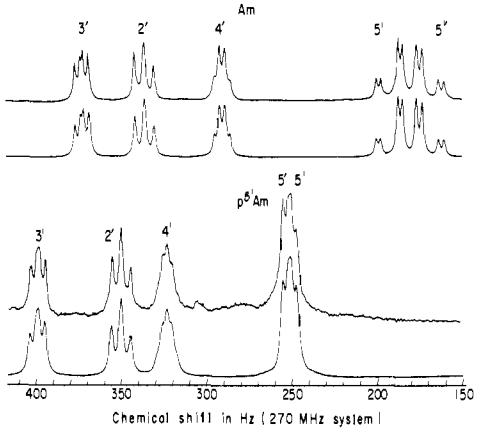

Figure 9. The observed and computer-simulated ^1H NMR (270 MHz) spectra of 2'-*O*-methyladenosine and the corresponding 5'-monophosphate, pH 5.0, 19 ± 1 °C. Chemical shifts are in hertz (270 MHz) from internal TMA.

formation shown in Figure 7c, introduction of the third phosphate so that ωc_2 is $\simeq 20°$ and $\omega' c_2$ is $\simeq 90°$ (Figure 8b), it will not affect the chemical shift of H(3'). Comparison of the chemical-shift data (Figure 6) for H(5'), H(5''), H(4'), and H(3') in $m^7G^{5'}p$, $m^7G^{5'}pp$, and $m^7G^{5'}ppp$ at pH 7, room temperature, shows that the chemical-shift trends are extremely similar to nonmethylated nucleotides. This suggests that the overall molecular geometry for the phosphate backbond in the 7-methylated guanosine series is very much similar to that in the nonmethylated system presented previously.

However, it is expected that 7-methylation will cause some internal torsional adjustment in the phosphate backbone so that the electrostatic interaction between N(7)(+) and negative charges becomes possible. Inspection of Figure 8a,b shows that the γ-phosphate which has the maximum electrostatic interaction (Table IV) resides in the close vicinity of N(7). The triphosphate backbone of crystalline Na_2ATP[41] is considerably different from the ones projected for $G^{5'}ppp$ in aqueous solution at pH 7.0. This becomes evident from a comparison of Figures 8c,d with 8a,b. In the crystal structure both ^2E (Figure 8d) and ^3E puckers (Figure 8c) are encountered, whereas in solution the preferred sugar pucker is ^2E. We have attempted a comparison between solution data on $G^{5'}ppp$ and crystal data on $A^{5'}ppp$ because crystal data on $G^{5'}ppp$ are not available.

Conformational Properties of 2'-O-Methylated Derivatives of Nucleosides and Nucleotides. In Figure 9 the experimentally observed 270-MHz ^1H spectra of Am and $p^{5'}Am$ as well as the corresponding computer simulations are presented. The chemical shifts of various derivatives investigated are tabulated in Table V and the coupling constants are given in Table VI.

Proton magnetic resonance work on 2'-O-methylated uridine and cytidine has demonstrated that 2'-O-methylation of pyrimidine nucleosides has little influence on their three-dimensional structure.[42] Data in Table VII for A and Am and for G and Gm indicate that this is also true of purine nucleosides. They in general show a preference for the ^2E conformation for the ribose and gg conformation about the C(4')–C(5') bond. In the case of 5'-mononucleotides as well, the data (Table VII) indicate that, in general, irrespective of whether

Table V. Chemical Shifts[a] of 2′-O-Methylated Nucleosides and Nucleotides and Corresponding Nonmethylated Nucleosides and Nucleotides

Compds	pH	Chemical shifts (δ), Hz								
		1′	2′	3′	4′	5′	5″	H(2) or H(5)	H(6) or H(8)	Me
N^6-Mep$^{5'}$A	7.0	777.5	425.7	355.0	318.7	234.4	229.0	1332.5	1423.6	43.7
A[b]	8.0	774.3	432.0	334.8	298.3	196.5	175.2	1344.6	1377.0	
Am	7.0	769.3	337.1	373.3	291.3	191.3	171.2	1320.0	1364.8	62.0
p$^{5'}$A[b]	8.0	796.5	437.7	358.3	318.3	225.7	217.6	1371.6	1468.8	
	5.0	793.8	422.3	355.3	324.0	254.6	254.6	1363.5	1428.3	
p$^{5'}$Am	7.0	811.8	364.6	403.3	320.3	232.2	228.0	1364.0	1459.3	73.9
	5.0	810.8	351.0	399.8	323.7	254.0	249.5	1363.3	1433.9	74.3
pp$^{5'}$A[b]	8.0	787.6	421.5	389.1	327.8	296.7	282.9	1337.6	1429.4	
	5.0	786.5	420.9	372.0	331.0	293.7	284.6	1342.2	1424.0	
pp$^{5'}$Am	7.0	817.3	349.0	429.0	322.0	291.5	274.0	1367.0	1445.7	73.9
	5.0	812.2	354.6	415.5	327.5	283.0	276.5	1365.9	1434.7	74.6
G	7.2	734.4	416.3	328.3	281.9	187.4	169.5		1296.0	
Gm	7.0	747.7	340.4	369.6	279.3	185.6	169.4		1297.1	70.1
p$^{5'}$	8.0	741.4	434.1	353.7	306.2	221.9	214.4		1355.7	
p$^{5'}$Gm	7.0	757.1	361.9	397.2	307.2	225.0	219.8		1357.1	74.3
pp$^{5'}$G	7.0	736.8	414.2	380.3	311.1	284.0	276.0		1332.1	
pp$^{5'}$Gm	7.0	759.1	349.8	423.6	308.7	284.5	274.2		1375.0	78.6
ppp$^{5'}$G	7.0	737.2	428.0	381.4	317.4	289.5	282.5		1336.1	
ppp$^{5'}$Gm	7.0	722.4	356.4	422.8	315.3	293.7	285.1		1333.5	77.2
p$^{5'}$U	7.0	749.2	325.9	306.2	282.9	219.0	203.8	746.8	1323.5	
p$^{5'}$Um	7.0	771.5	255.1	350.8	283.2	239.5	218.0	753.0	1334.8	88.6
pp$^{5'}$U[b]	8.0	749.8	324.5	339.1	290.8	284.6	279.2	751.4	1296.8	
pp$^{5'}$Um	7.0	766.6	289.0	365.9	285.0	245.5	242.8	750.5	1304.7	92.6
p$^{5'}$C	7.0	757.3	311.8	306.7	280.5	224.9	211.4	791.1	1328.7	
p$^{5'}$Cm	7.0	775.3	235.7	342.6	276.6	240.7	213.8	797.8	1341.4	93.0
pp$^{5'}$C[b]	8.0	751.9	307.8	334.5	291.3	288.9	286.5	792.7	1295.2	
pp$^{5'}$Cm	7.0	771.4	290.5	355.5	288.2	230.0	221.0	794.9	1307.4	97.0

[a] Shifts are given relative to internal TMA and are accurate to ±0.1 Hz. Concentration is 0.02–0.07 M for methylated derivatives and 0.05–0.1 M for nonmethylated derivatives except for G (0.004 M) and p$^{5'}$G (0.01 M). Temperature is 19 ± 1 °C. [b] Temperature 30 ± 1 °C.

they are purine or pyrimidines, 2′-O-methylation does not seem to have a strong influence on the conformational properties. All these compounds show pronounced preference for g′g′ and gg conformations about the C(5′)–O(5′) and C(4′)–C(5′) bonds. The purine system shows 60–65% preference for the ^2E conformation whereas the pyrimidines usually exist as a 50:50 blend of ^2E: ^3E conformers. Inspection of the ribose population for the pyrimidine mononucleotides indicates that 2′-O-methylation causes a slight shift toward ^3E pucker and this is noticeably absent in the purine derivatives. The data in Table VII also reveal that di- or triphosphorylation does not significantly alter the preferred conformational mode about C(4′)–C(5′), C(5′)–O(5′), and ribose pucker, even though some small changes result from this constitutional modification. The data also reveal that N(6)-methylation has no significant effect on the conformational properties of purine 5′-mononucleotides.

Comparison of the chemical shifts of adenosine vs. p$^{5'}$A, and Am vs. p$^{5'}$Am, and similar comparisons of guanosine series indicate that 5′-phosphorylation causes downfield shift of all the protons. The downfield shift is extremely conspicuous for the H(8) proton (60–90 Hz), clearly indicating that the phosphate group polarizes the C(8)–H(8) bond which is possible only in the anti conformation. This suggests that adenosine and guanosine 5′-monophosphate nucleotides, whether 2′-O-methylated or not, significantly exist in the anti domain. The data in Table V reveal that di- and triphosphorylations do indeed affect the chemical shift of the base protons and that of H(1′) and H(2′), clearly suggesting readjustment of χ_{CN} upon di- and triphosphorylation of purine 5′-mononucleotides. Unfortunately, it is not possible to determine either the direction or magnitude of the changes in χ_{CN} due to this constitutional modification because of lack of precise information about the effect of χ_{CN} variation on chemical shifts. However, the data for H(1′) and H(2′) which are most sensitive to χ_{CN} variation indicate that these torsional adjustments within the anti domain are significantly influenced by the 2′-*O*-methyl group. Thus, in going from p$^{5'}$G to pp$^{5'}$G H(1′) and H(2′) undergo shifts to higher fields whereas in going from p$^{5'}$Gm to pp$^{5'}$Gm H(1′) shifts to lower field and H(2′) shifts to higher field. The trends are very similar in the adenosine series as well. In the uridine and cytidine series one sees a clear trend that upon 2′-O-methylation H(6) is moving to lower field in all systems examined. This observation is consistent with the situation in which the magnitude of χ_{CN} in the nonmethylated series moves from a value of $\simeq 60°$ to lower values in the methylated system as has been demonstrated by Cheng and Sarma.[19] Even though the lower field shift of H(6) is expected to be accompanied by a higher field shift[19] of H(1′), this is not observed in the present case because this change is masked by the very effect of 2′-O-methylation itself, and one cannot compare directly the chemical shifts of H(1′) and H(2′) in, for example, p$^{5'}$U and p$^{5'}$Um to obtain χ_{CN} information.

In conclusion, 2′-O-methylation does not seem to have any significantly noticeable effect on the conformation of the ribose ring or that about C(4′)–C(5′) and C(5′)–O(5′) bonds.

Table VI. Coupling Constants[a] for 2′-O-Methylated Nucleosides and Nucleotides, and Corresponding Nonmethylated Nucleosides and Nucleotides in D_2O[b]

Compds	pH	Coupling constants, Hz								
		1′2′	2′3′	3′4′	4′5′	4′5″	4′P	5′P	5″p	5′5″
N^6-Mep$^{5'}$A	7.0	5.6	5.3	3.7	3.0	3.0	1.7	5.0	5.0	−12.5
A	8.0	6.2	5.2	3.5	2.8	3.6				−12.6
Am	7.0	6.0	5.1	3.2	2.5	3.6				−12.9
p$^{5'}$A	8.0	6.0	5.1	3.5	3.9	2.6	1.7	4.6	4.6	−11.8
	5.0	5.7	5.2	3.7	3.1	3.1	1.7	5.0	5.0	−12.0
p$^{5'}$Am	7.0	6.1	5.1	3.3	3.3	3.1	2.0	5.0	4.8	−12.0
	5.0	6.1	5.2	3.6	3.3	2.9	2.2	5.3	5.2	−12.5
pp$^{5'}$A	8.0	5.0	5.2	4.5	2.6	3.4	1.8	5.5	5.2	−12.8
	5.0	5.5	5.2	3.8	3.5	2.5	2.2	5.0	5.0	−12.8
pp$^{5'}$Am	7.0	5.4	4.8	3.8	2.9	3.4	1.9	5.1	4.8	−12.5
	5.0	6.1	5.1	3.4	3.0	3.0	1.8	5.0	5.0	−12.5
G	7.2	5.9	5.2	3.7	3.2	3.8				−12.5
Gm	7.0	6.0	5.2	3.6	3.1	3.9				−12.8
p$^{5'}$G	8.0	6.1	5.2	3.4	4.3	2.9	1.1	4.7	4.7	−11.6
p$^{5'}$Gm	7.0	6.0	5.1	3.3	3.1	3.5	1.5	4.8	4.9	−12.3
pp$^{5'}$G	7.0	5.4	5.3	4.2	3.5	3.5	1.7	5.5	4.8	−12.0
pp$^{5'}$Gm	7.0	5.6	5.0	4.0	3.0	3.3	2.0	5.7	5.0	−12.0
ppp$^{5'}$G	7.0	6.1	5.0	3.4	3.1	3.1	1.8	5.8	5.0	−12.5
ppp$^{5'}$Gm	7.0	5.7	5.0	3.3	3.6	3.7	1.8	5.7	5.0	−12.3
p$^{5'}$U	7.0	5.1	4.8	4.1	2.3	2.8	1.7	3.8	5.2	−11.8
p$^{5'}$Um	7.0	4.7	5.3	4.7	2.8	2.7	2.0	4.1	4.8	−12.0
pp$^{5'}$U	8.0	4.0	5.2	5.1	2.8	3.0	2.6	5.5	5.5	−11.9
pp$^{5'}$Um	7.0	3.7	5.2	5.3	4.3	3.4	1.5	5.8	5.1	−12.0
p$^{5'}$C	7.0	4.4	4.5	4.6	2.5	2.9	1.0	3.8	4.9	−12.0
p$^{5'}$Cm	7.0	4.0	5.2	5.3	2.3	2.5	1.4	4.1	4.6	−12.3
pp$^{5'}$C	8.0	3.6	5.2	5.6	3.1	3.1	2.6	5.4	5.4	−11.9
pp$^{5'}$Cm	7.0	3.5	5.7	5.3	3.5	3.5	2.0	5.8	4.7	−12.0

[a] Coupling constants are accurate to ±0.1–0.2 Hz. [b] Solution conditions are the same as in Table V.

However, it seems to influence the actual time-average magnitude of χ_{CN} in the anti domain.

An extended discussion of the orientation of the 2′-*O*-methyl group in 2′-O-methylated nucleic acid derivatives has been presented elsewhere.[18,19] Using arguments presented there, one could conclude that C(2′)–O(2′) torsion (ϕ'', Figure 1) in the collection of molecules examined in the present study populates in domains in which ϕ'' is $\simeq$180–150° and $\simeq$60–90° with little or no contribution from conformers in the $\phi'' \simeq 300°$ domain. One presumably has an equilibrium between conformers in these domains. This equilibrium is coupled to the ribose conformation.[18,19]

The geometry of di- and triphosphate backbones in the 2′-O-methylated derivatives can be monitored, using the procedure presented earlier. Chemical-shift data in Table V indicate that diphosphorylation and triphosphorylation cause a downfield shift of H(3′) to approximately the same magnitude in nonmethylated and methylated derivatives. This clearly suggests that 2′-O-methylation has no significant effect on the orientation of the phosphate groups and retains the geometry that was found in the nonmethylated analogues (vide supra).

Conformational Properties of Dinucleoside Monophosphates and Their 2′-O-Methylated Derivatives. It was shown in an earlier section that 2′-O-methylation has no significant effect on the aqueous solution conformation of mononucleotides. However, the situation is very different in the case of 3′,5′ dimers. Details of this work have been published from our laboratory.[18,19] Because the findings are very relevant in describing the spatial configuration of the mRNA 5′ terminus, $m^7G^{5'}ppp^{5'}NmpN$, a summary of these findings is given below.

Comparison of the conformational properties of ApA and AmpA[18] shows that the 2′-*O*-methyl group has the following effects.

(1) It causes a shift toward ^2E pucker for the ribose ring of the 3′-nucleotidyl units. The shift is about 8% for the Amp segment and the ^2E:^3E distribution is 50:50.

(2) It decreases the g′g′ populations about C(5′)–O(5′) (ϕ).

(3) It severely influences the torsion about C(3′)–O(3′) (ϕ').

(4) It has a strong influence on the torsional preferences about O(3′)–P(ω') and O(5′)–P(ω). This conclusion is reached from the conclusions of Sundaralingam[24,25] that the principal site of flexibility lies in the phosphodiester bonds and the reported[20,21] nexus between phosphodiester torsions and stacking interactions.

The overall effect of all these variations is that the 2′-*O*-methyl group causes AmpA to adopt an extended structure compared to ApA. This was further confirmed by measurement of ring current shifts.[18] In aqueous solution it is known[20] that ApA exists as an equilibrium blend of stacked and extended arrays. The usually preferred stacked conformation for ApA is a right-handed stack in which the adenines are anti, both sugars ^3E, $\phi' \simeq 210°$, $\omega' \simeq 290°$, $\omega \simeq 300°$, $\psi \simeq 60°$, and $\phi \simeq 180°$. A perspective of this structure is shown in Figure 10. AmpA can be described as an equilibrium blend of a stacked array and an extended conformation. The effect of the 2′-*O*-methyl group is to increase the population of the ex-

Table VII. Conformational Parameters for 2′-O-Methylated Nucleosides and Nucleotides, and Corresponding Nonmethylated Nucleosides and Nucleotides[a]

Compds	pH	% conformational preferences and various bonds: C(5′)-O(5′) g′g′ ⇄ g′/t′		C(4′)-C(5′) gg ⇄ g/t		Ribose, ^3E
N^6-Mep$^{5'}$A	7.0	72	28	79	21	39
A	8.0			75	25	36
Am	7.0			78	22	33
p$^{5'}$A	7.0	76	21	75	25	38
	5.0	72	28	77	23	39
p^{c}5′Am	7.0	73	27	75	25	35
	5.0	70	30	77	23	38
pp$^{5'}$A	8.0	69	31	79	21	47
	5.0	72	28	79	21	40
pp$^{5'}$Am	7.0	73	27	76	24	40
	5.0	72	28	79	21	36
G	7.2			69	31	38
Gm	7.0			69	31	37
p$^{5'}$G	7.0	75	25	67	33	36
p$^{5'}$Gm	7.0	74	26	73	27	35
pp$^{5'}$G	7.0	71	29	69	31	44
pp$^{5'}$Gm	7.0	69	31	76	24	42
ppp$^{5'}$G	7.0	68	32	77	23	36
ppp$^{5'}$Gm	7.0	69	31	66	34	35
p$^{5'}$U	7.0	77	23	89	11	43
p$^{5'}$Um	7.0	77	23	85	15	49
pp$^{5'}$U	8.0	67	33	81	19	54
pp$^{5'}$Um	7.0	68	32	62	38	56
p$^{5'}$C	7.0	78	22	86	14	48
p$^{5'}$Cm	7.0	78	22	92	8	56
pp$^{5'}$C	8.0	68	32	77	23	59
pp$^{5'}$Cm	7.0	70	30	69	31	56

[a] Solution conditions are the same as in Table V.

tended array at the expense of the stacked conformation, i.e., the population of the extended conformation will be the predominant one compared to the stacked one. Assuming a simple equilibrium between stacked and extended arrays one can compute the population of these species using methodology developed by Lee et al.[20] and Ezra et al.[21] Such calculations show that the total population of the extended arrays in AmpA is ≃75%. In Figure 11 are shown the perspectives of the conformations of AmpA along with the torsion angles. Comparison of the torsion angles for the two conformations in Figure 11 reveals that small changes in a few of them have an important effect on the molecular topology of the structure. It should be noted that Figure 11 is only one of the possibilities for the many extended arrays in aqueous solution. This particular model is selected because it fits in with the intercalated model of Tsai et al. (vide infra ref 48) and explains the fact that 2′-O-methylation causes shifts of sugar pucker to ^2E and increases the magnitude of ϕ'. It is indeed possible to generate an extended structure for AmpA in which the sugar of Amp- is ^3E and ϕ' and χ assume other values.

Conformational Properties of the Dimer m^{7}G$^{5'}$ppp$^{5'}$Am. In Figure 12 are shown the experimental 270-MHz ^{1}H NMR spectra of m^{7}G$^{5'}$ppp$^{5'}$Am at 5 and 80 °C. Computer simulation was arrived at by simulating m^{7}G$^{5'}$p- and -p$^{5'}$Am separately and finally by combining the two separate simulations to obtain the complete computer-simulated spectra. These are also illustrated in Figure 12.

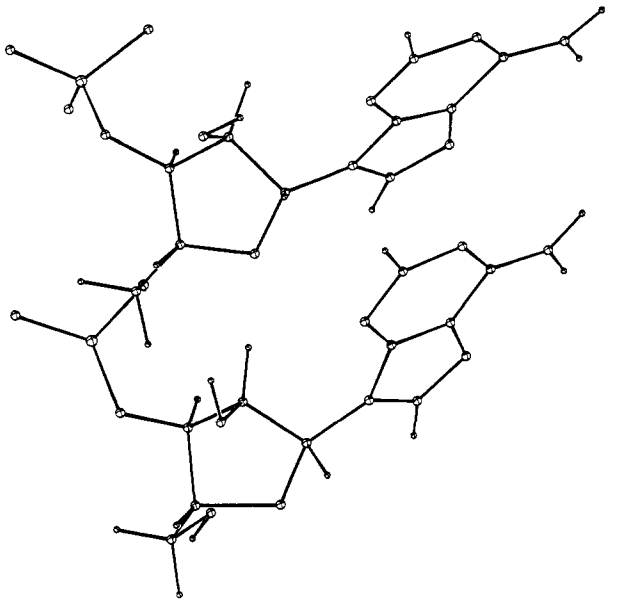

Figure 10. Perspective of a right-handed stacked ApA; both the sugar rings are in ^3E, $\chi_1 = 40°$, $\phi' = 210°$, $\omega_1' = 290°$, $\omega_1 = 300°$, $\phi = 180°$, $\psi = 60°$, and $\chi_2 = 50°$.

Table VIII. Chemical Shifts[a] and Coupling Constants[b] for m^{7}G$^{5'}$-ppp$^{5'}$Am in D_2O

	5 °C		80 °C	
	m^{7}G$^{5'}$p-	-p$^{5'}$Am	m^{7}G$^{5'}$p	-p$^{5'}$Am
$\delta 1'$	702.5	765.3	747.0	794.3
$\delta 2'$	334.0	302.0	377.3	326.5
$\delta 3'$	317.5	385.5	347.1	391.5
$\delta 4'$	329.0	321.5	319.0	315.8
$\delta 5'$	319.0	309.0	321.0	293.5
$\delta 5''$	292.5	288.5	288.0	287.5
$\delta 2$		1333.5		1369.6
$\delta 8$		1401.6		1420.3
δCH_3	220.9	60.6	243.7	69.8
$J_{1'2'}$	3.1	5.5	4.0	5.7
$J_{2'3'}$	5.0	5.2	5.2	5.3
$J_{3'4'}$	5.8	3.8	4.8	3.8
$J_{4'5'}$	1.8	2.6	2.5	3.0
$J_{4'5''}$	2.1	2.6	2.5	3.0
$J_{4'5'} + J_{4'5''}$	3.9	5.2	5.0	6.0
$J_{4'p}$	1.5	1.6	1.7	1.8
$J_{5'p}$	3.8	4.8	4.5	5.0
$J_{5''p}$	4.3	4.8	5.0	5.0
$J_{5'p} + J_{5''p}$	8.1	9.6	9.5	10.0
$J_{5'5''}$	−11.5	−12.1	−12.0	−12.5

[a] Shifts are given relative to internal TMA and are accurate to ±0.1 Hz. [b] Coupling constants are accurate to ±0.1–0.2 Hz. Concentration is ~0.008 M and pH 7.0.

The chemical-shift and coupling constant data at 5 and 80 °C are given in Table VIII. The conformational parameters (Table IX) suggest that both m^{7}G$^{5'}$p- and -p$^{5'}$Am nucleotidyl units prefer gg and g′g′ conformations about C(4′)–C(5′) and C(5′)–O(5′) bonds. However, the guanosine component displays ^3E sugar pucker whereas adenosine displays ^2E. In order to determine whether dimerization from the component monomers causes conformational changes we have compared the conformational properties of monomers at pH 5 (Tables III and VII) with that of dimer (Table IX). Such comparisons lead to a very interesting finding. It was found that with respect to the local conformations about the sugar ring and C(4′)–C(5′) and C(5′)–O(5′) bonds the mononucleotides conserve their conformation in the dimer. However, the m^{7}G$^{5'}$p- part displays

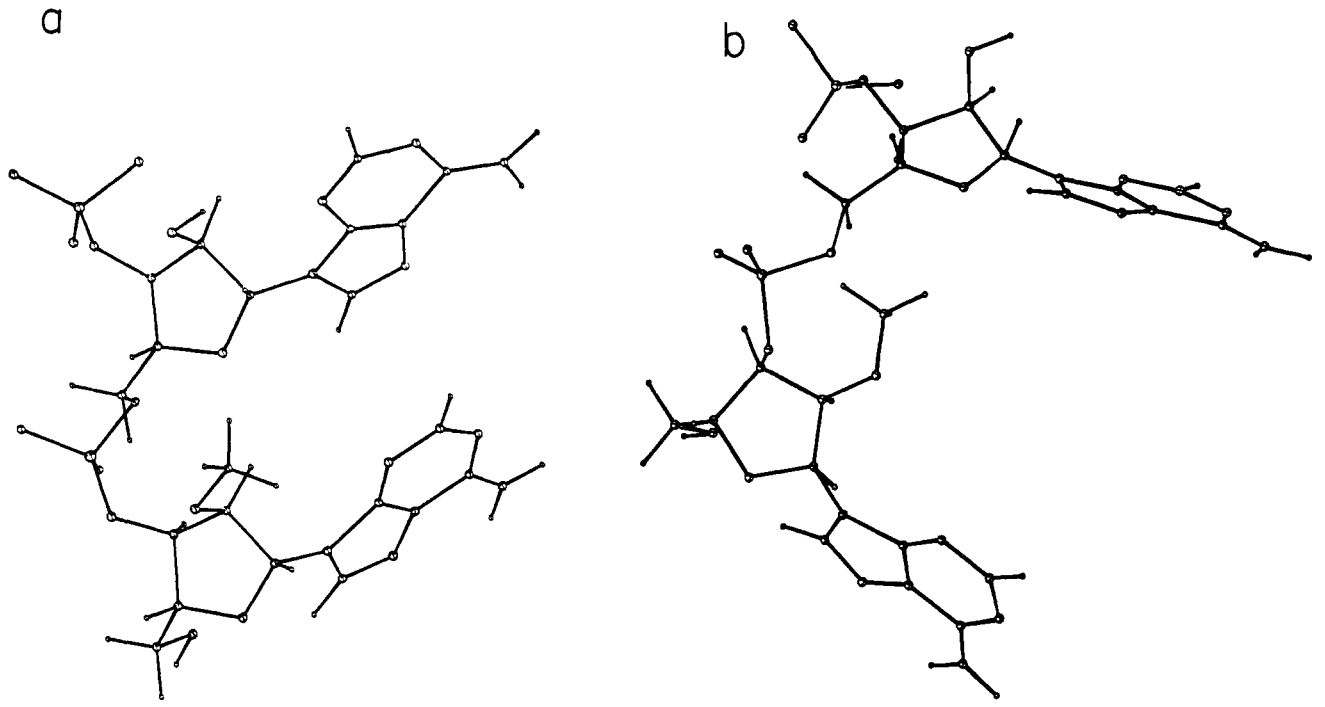

Figure 11. In (a) and (b), respectively, are shown the perspectives of the stacked and extended conformations of AmpA. Conformational details in (a) are the same as in Figure 10 and $\phi'' = 90°$. In (b) the ribose of Amp- is ^2E, that of -pA = ^3E; $\chi_1 = 40°$, $\phi' = 230°$, $\phi'' = 180°$, $\omega_1' = 270°$, $\omega_1 = 220°$, $\phi = 195°$, $\psi = 60°$, and $\chi_2 = 100°$.

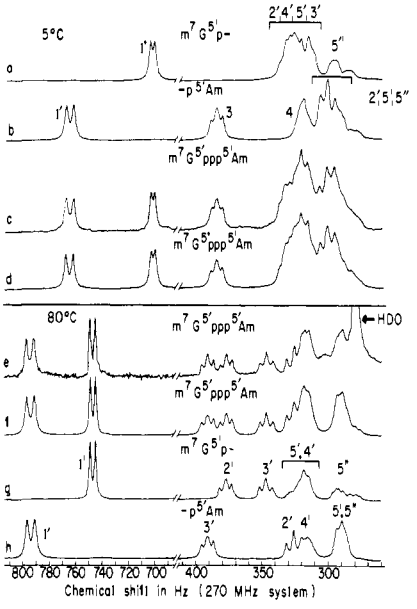

Figure 12. The observed 270-MHz ^{1}H NMR spectra of m^{7}G$^{5'}$ppp$^{5'}$Am, pH 7.0, at 5 °C (c) and 80 °C (e). Both spectra were analyzed by computer simulation of m^{7}G$^{5'}$p- and -p$^{5'}$Am parts separately (a and b at 5 °C and g and h at 80 °C) and combining them to produce complete simulation.

Table IX. Conformational Parameters for m^{7}G$^{5'}$ppp$^{5'}$Am[a]

		% conformational preferences about various bonds				
		C(5')-O(5') g'g' ⇄ g'/t'		C(4')-C(5') gg ⇄ g/t		Ribose, ^3E
5 °C	m^{7}G$^{5'}$p-	81	19	100	0	61
	-p$^{5'}$Am	74	26	88	12	40
80 °C	m^{7}G$^{5'}$p-	75	25	90	10	51
	-p$^{5'}$Am	72	28	79	21	40

[a] Solution conditions are the same as Table VIII.

Table X. Dimerization Effect on Chemical Shift,[a] Temperature Effect on Chemical Shift[b]

	m^{7}G$^{5'}$p-	-p$^{5'}$Am	m^{7}G$^{5'}$p-	-p$^{5'}$Am
AH(8)	c	27.6	c	18.7
AH(2)		22.8		36.1
2'-*O*-Methyl		11.8		9.2
7-Methyl	25.7		22.8	
H(1')	62.6	39.8	44.5	29.0
H(2')	60.6	46.3	43.3	24.5
H(3')	46.3	28.1	29.6	6.0
H(4')	1.5	6.9	−10.0	−5.7
H(5')	1.7	−18.0	2.0	−15.5
H(5'')	−0.7	−11.5	−4.5	−1.0

[a] Value in monomer − that in dimer. [b] Value for dimer at 80 °C − that for dimer at 5 °C. [c] GH(8) exchanges with D_2O and could not be located.

a small but real shift toward ^3E pucker in the dimer. This shift indicates a reduction in χ_{CN} upon dimerization as can be inferred from the interrelationship between ribose pucker and χ_{CN}.[20,21] Unfortunately, in the dimer one cannot obtain the information about χ_c and χ_1 (Figure 1) from the dimerization trends for H(1') and H(2') (vide supra). This is because intramolecular stacking interactions between the m^{7}G$^{5'}$p- and -p$^{5'}$Am nucleotidyl units (vide infra) will affect the magnitudes of δH(1') and δH(2') in the dimer. However, it is reasonable and sound to assume that χ_c and χ_1 are in the anti domain.

Knowledge about the spatial configuration of m^{7}G$^{5'}$-ppp$^{5'}$Am can be obtained by comparing the shift data in the dimer and the components. These dimerization data are presented in Table X. The data indicate that the chemical shifts of the monomers are considerably different from that of dimer. This could originate if the two nucleotidyl units at the 5' ends of the triphosphate bridge interact with each other and change the effective field felt at various protons. The H(4'), H(5'), and

Figure 13. The two perspectives of $m^7G^{5'}p$. The left one shows the B side of purine and the right one the A side. The ribose is ^3E, $\psi = 60°$, $\phi = 180°$, and $\chi_c = 60°$.

H(5″) of the $m^7G^{5'}p$- residue undergo very few changes as a result of this interaction. The H(5′) and H(5″) of the -$p^{5'}$Am residues undergo shifts to lower field. The remaining protons of both nucleotidyl units are shifted to higher fields by as much as 60 Hz. From the observation that H(8) and H(2) of the -$p^{5'}$Am moiety and the 7-methyl group of the $m^7G^{5'}p$- moiety are shifted to higher fields, it is immediately apparent that the two aromatic systems interact with each other probably as in a parallel stacking.

An important observation is that *the interaction between bases takes place while the individual nucleotidyl units maintain conformations which are very similar to those in the respective monomers for ribose, C(4′)-C(5′), and C(5′)-O(5′) bonds. Hence, this intramolecular stacking interaction must be made feasible by torsional variation around the phosphodiester bonds of the triphosphate bridge.* In order to describe the stacking possibilities between the two ring systems, it is necessary to label the sides of purine moieties by the letters A and B as has been done before.[43] These are shown in Figure 13 for $m^7G^{5'}p$.

Now, using the previously determined conformations for the ribose rings, ψ_c, ϕ_c, ψ_1, and ϕ_1, and using the anti domain for χ_c and χ_1 one could search in the conformation space about $\omega c_1/\omega' c_1$, $\omega c_2/\omega' c_2$, and $\omega c_3/\omega' c_3$ to determine the stacking possibilities. This search was undertaken by converting the torsion angles to x,y,z coordinates which were then used along with ORTEP-II to generate perspectives which avoid bad van der Waals contacts. This search produced four possible stacking arrangements for $m^7G^{5'}ppp^{5'}Am$. They are: (1) A-A stacking—the A surface of the $m^7G^{5'}p$- moiety faces the A surface of the -$p^{5'}$Am moiety (Figure 14). (2) A-B stacking—the A surface of the $m^7G^{5'}p$- moiety faces the B surface of the -$p^{5'}$Am moiety (Figure 15); (3) B-A stacking—the B

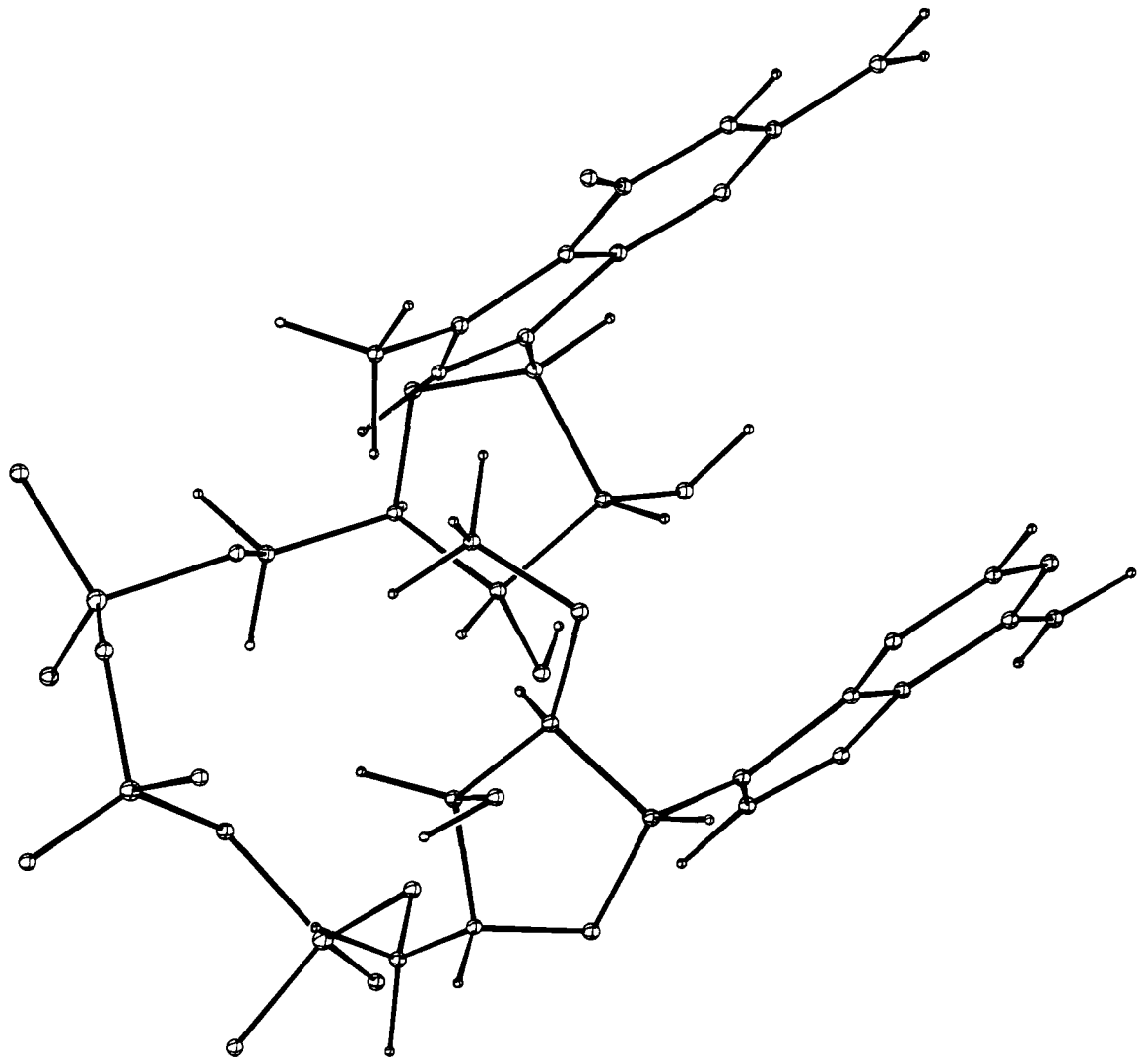

Figure 14. A perspective of $m^7G^{5'}ppp^{5'}Am$ in the A-A stacked array. The conformational details are as follows: $\chi_c = 30°$; ribose of $m^7G^{5'}p$- = ^3E; $\psi_c = 60°$; $\phi_c = 180°$; $\omega_1 = 155°$; $\omega' c_1 = 265°$; $\omega c_2 = 35°$; $\omega' c_2 = 155°$; $\omega c_3 = 295°$; $\omega' c_3 = 112°$; $\phi_1 = 180°$; $\psi_1 = 60°$; $\phi'' = 180°$; ribose of -$p^{5'}$Am = ^2E; $\chi_1 = 30°$.

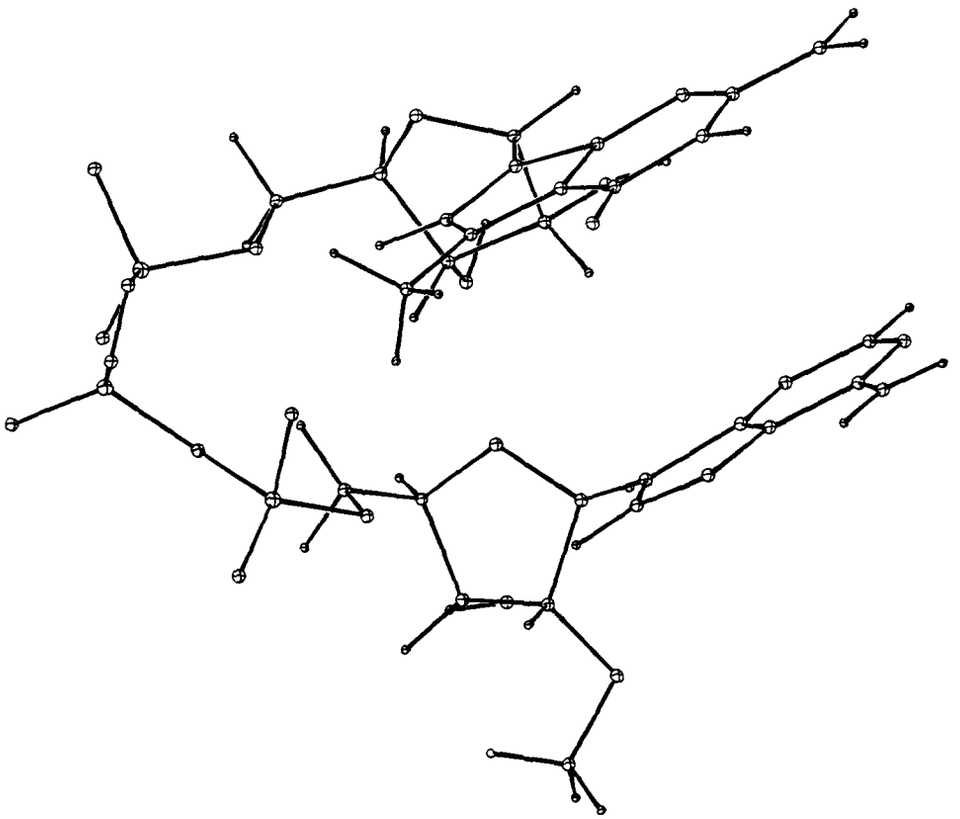

Figure 15. A perspective of $m^7G^{5'}ppp^{5'}Am$ in the A-B stacked array. The conformational details are as follows: $\chi_c = 35°$; ribose of $m^7G^{5'}p$- = ^3E; $\psi_c = 60°$; $\phi_c = 180°$; $\omega c_1 = 200°$; $\omega' c_1 = 250°$; $\omega c_2 = 60°$; $\omega' c_2 = 120°$; $\omega c_3 = 200°$; $\omega' c_3 = 355°$; $\phi_1 = 180°$; $\psi_1 = 60°$; $\phi'' = 180°$; ribose of -$p^{5'}Am$ = ^2E; $\chi_1 = 80°$.

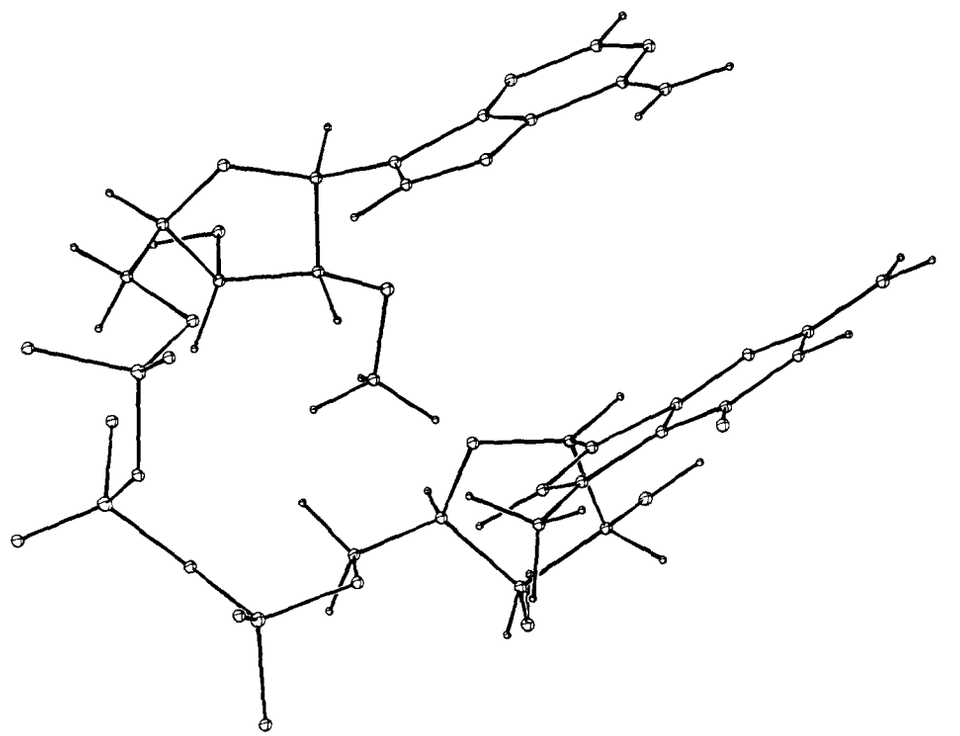

Figure 16. A perspective of $m^7G^{5'}ppp^{5'}Am$ in the B–A stacked array. The conformational details are as follows: $\chi_c = 30°$; ribose of $m^7G^{5'}p$- = ^3E; $\psi_c = 60°$; $\phi_c = 180°$; $\omega c_1 = 335°$; $\omega' c_1 = 235°$; $\omega c_2 = 25°$; $\omega' c_2 = 130°$; $\omega c_3 = 310°$; $\omega' c_3 = 100°$; $\phi_1 = 180°$; $\psi_1 = 60°$; $\phi'' = 180°$; ribose of -$p^{5'}Am$ = ^2E; $\chi_1 = 40°$.

surface of the $m^7G^{5'}p$- moiety faces the A surface of the -$p^{5'}Am$ moiety (Figure 16); (4) B–B stacking—the B surface of the $m^7G^{5'}p$- moiety faces the B surface of the -$p^{5'}Am$ moiety (Figure 17).

The four stacking arrangements depicted in Figures 14–17 are generated while maintaining the same conformation for the ribose rings and C(4′)–C(5′) and C(5′)–O(5′) bonds; the only changes are in χ_c, χ_1, ωc_1, $\omega' c_1$, ωc_2, $\omega' c_2$, ωc_3, and $\omega' c_3$.

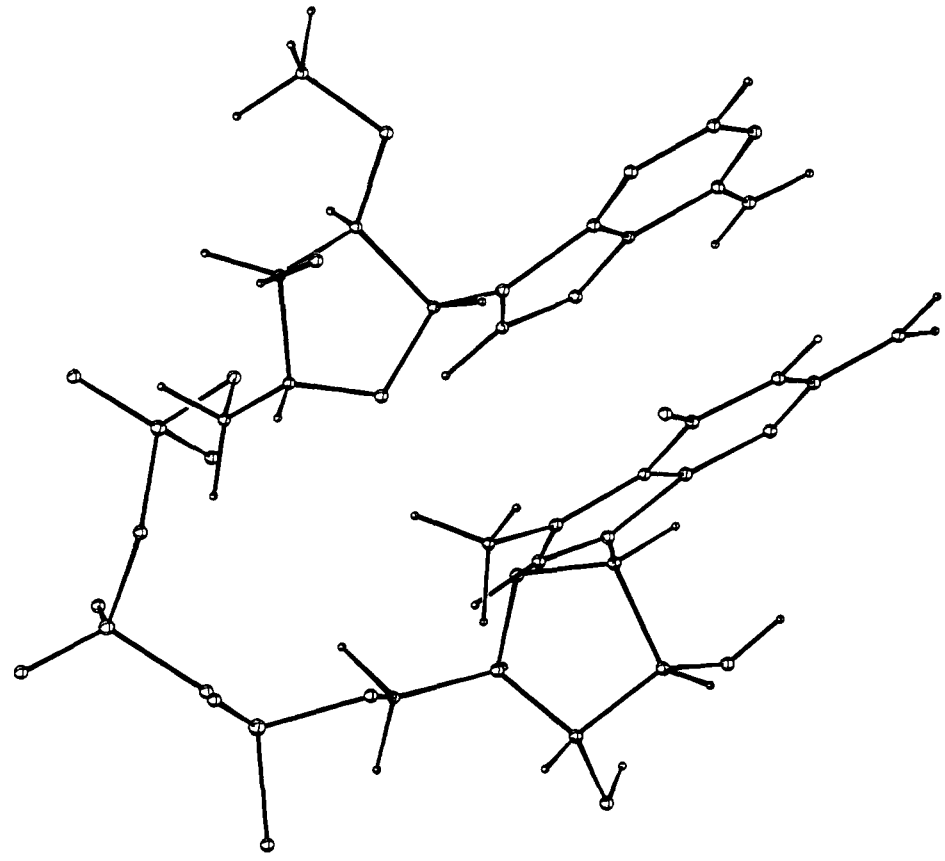

Figure 17. A perspective of $m^7G^{5'}ppp^{5'}Am$ in the B-B stacked array. The conformational details are as follows: $\chi_c = 30°$; ribose of $m^7G^{5'}p- = {}^3E$. $\psi_c = 60°$; $\phi_c = 180°$; $\omega c_1 = 330°$; $\omega' c_1 = 230°$; $\omega c_2 = 60°$; $\omega' c_2 = 110°$; $\omega c_3 = 205°$; $\omega' c_3 = 335°$; $\phi_1 = 180°$; $\psi_1 = 60°$; $\phi'' = 180°$; ribose of $-p^{5'}Am = {}^2E$; $\chi_1 = 30°$.

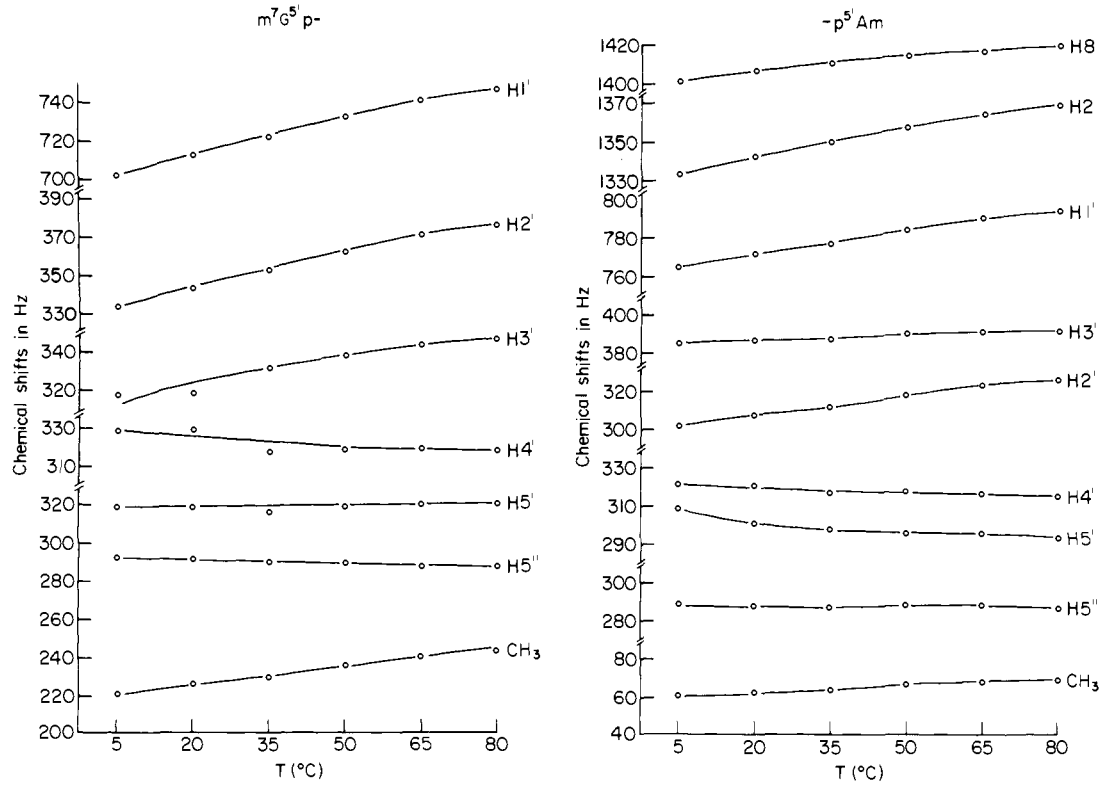

Figure 18. Temperature profiles of the chemical shifts of $m^7G^{5'}p-$ and $-p^{5'}Am$ parts of $m^7G^{5'}ppp^{5'}Am$, pH 7.0.

The employed torsion angles are given in the respective figure legends. In A-A and B-B stacking the O(1′) of both ribose rings are oriented in opposite directions whereas in A-B and B-A stacking they are oriented in the same direction.

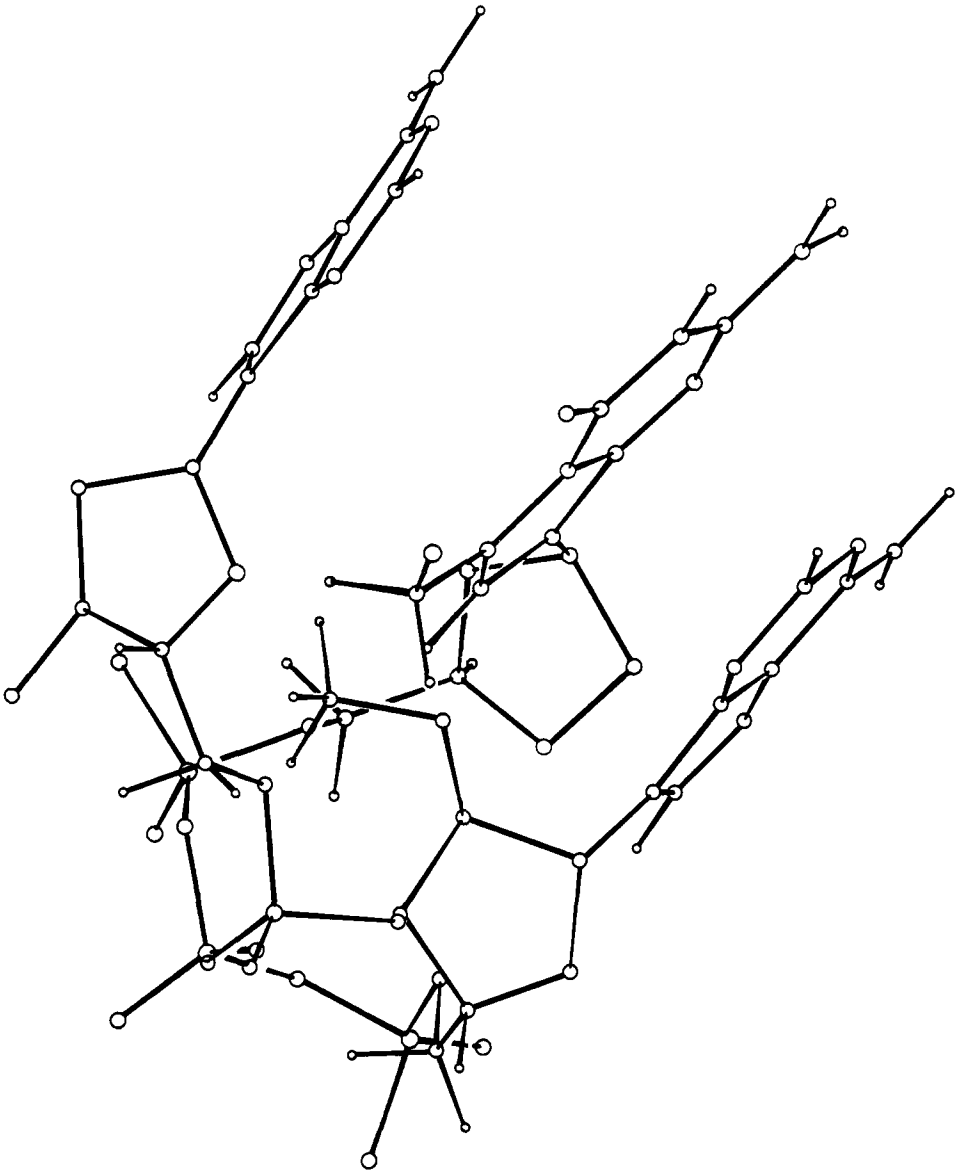

Figure 19. This is the proposed model for the 5′ terminus $m^7G^{5'}ppp^{5'}AmpA$ of mRNA. For clarity certain of the atoms were omitted and the intercalating $m^7G^{5'}ppp$-part is shown in red. The ribose of the $m^7G^{5'}ppp$-part is ^3E, that of the Amp-part is ^2E, and that of -pA is ^3E. The various torsion angles are: $\chi_c = 40°$; $\psi_c = 60°$; $\phi_c = 180°$; $\omega c_1 = 155°$; $\omega' c_1 = 265°$; $\omega c_2 = 35°$; $\omega' c_2 = 130°$; $\omega c_3 = 315°$; $\omega' c_3 = 112°$; $\phi_1 = 180°$; $\psi_1 = 60°$; $\chi_1 = 40°$; $\phi_1' = 230°$; $\phi'' = 180°$; $\omega_1' = 270°$; $\omega_1 = 220°$; $\phi_2 = 195°$; $\psi_2 = 60°$; $\chi_2 = 100°$. The designations of the torsion angles are illustrated in Figure 1.

The observation that the base protons of the $-p^{5'}Am$ moiety and the 7-methyl group of the $m^7G^{5'}p$- moiety are shielded (Table X) can be rationalized on the basis of any of the above stacked structures or on the basis of a blend of all four forms.

Table XI. Cylindrical Coordinates z, $\rho(5)$, and $\rho(6)$ for Protons of $m^7G^{5'}p$- Part of $m^7G^{5'}ppp^{5'}Am$ for the A–A Stack Shown in Figure 14[a]

Protons on $^7G^{5'}p$-	z	$\rho(5)$	$\rho(6)$	Projected shielding	Obsd shielding
H(1′)	5.7	2.7	3.7	0.19	0.232
H(2′)	3.4	2.0	3.5	0.37	0.224
H(3′)	3.7	3.2	5.3	0.08	0.171
H(4′)	5.7	4.8	6.6	0.0	0.005
H(5′)	5.2	5.4	7.5	0.0	0.006
H(5″)	6.8	4.8	6.9	0.0	−0.002
H(8)	5.7	2.0	4.0	0.19	
$-CH_3(7)$	5.8	3.3	4.1	0.09	0.095

[a] Also given are the projected and observed shielding in parts per million.

Is it possible to make a distinction among the four possible structures? We believe that a distinction can be made with a semiquantitative use of ring current theory and following the shift trends of the interior protons H(1′), H(2′), H(3′), H(4′),

Table XII. Cylindrical Coordinates z, $\rho(5)$, and $\rho(6)$ for Protons of $-p^{5'}m$ Part of $m^7G^{5'}ppp^{5'}Am$ for the A-A Stack Shown in Figure 14[a]

Protons $-p^{5'}Am$	z	$\rho 5$	$\rho 6$	Projected shielding	Obsd shielding
H(1′)	6.4	2.5	3.9	0.1	0.147
H(2′)	3.7	2.6	4.5	0.14	0.171
H(3′)	3.8	4.9	6.7	0.0	0.104
H(4′)	6.3	5.4	7.1	0.0	0.025
H(5′)	4.7	5.8	7.8	0.0	−0.067
H(5″)	6.0	5.8	7.9	0.0	−0.043
H(2)	5.8	3.8	2.0	0.1	0.084
H(8)	5.7	2.7	4.7	0.08	0.102
$-CH_3(2')$	2.2	3.7	4.9	0.0	0.044

[a] Also given are the projected and observed shieldings in parts per million.

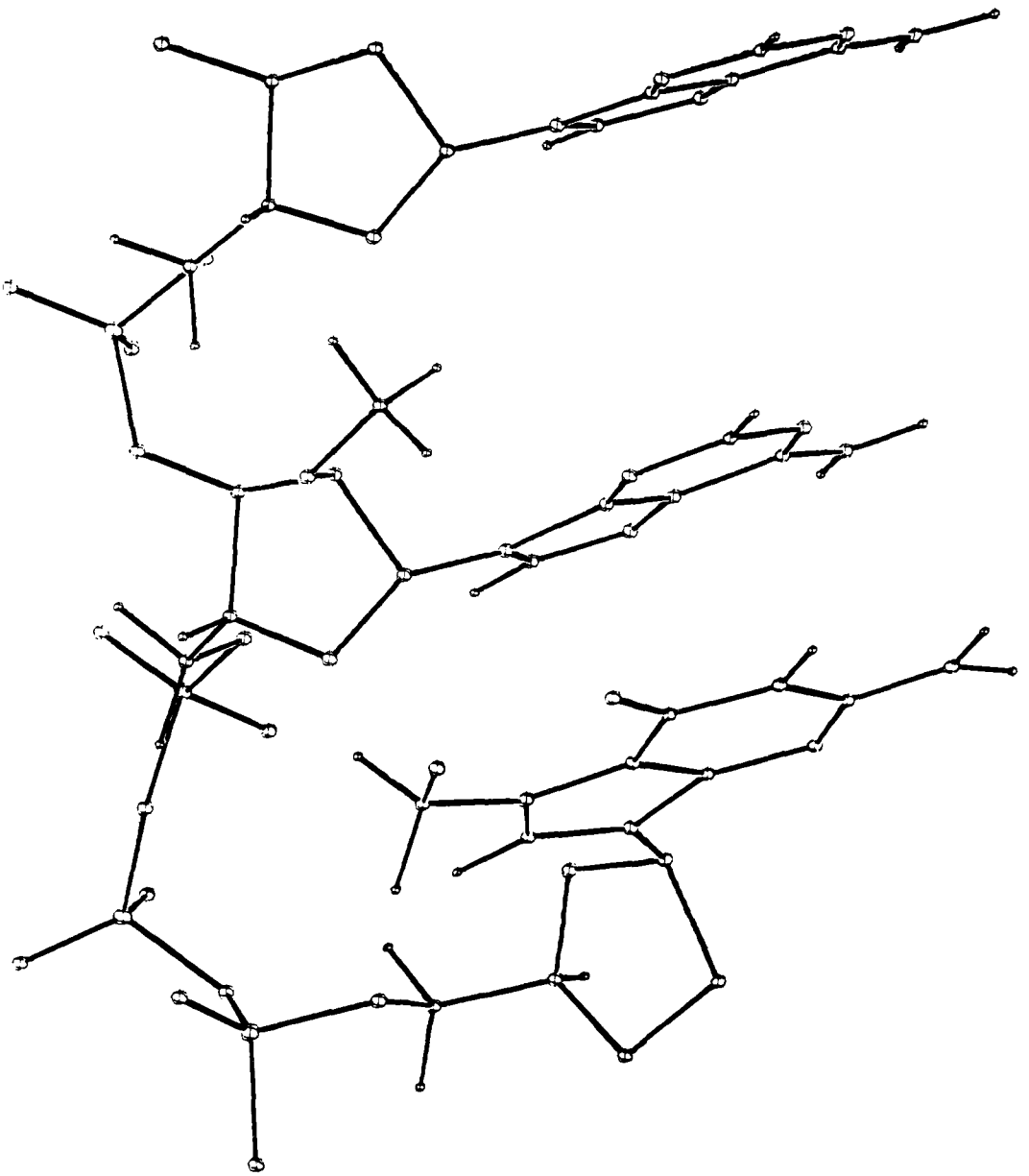

Figure 20. A model for $m^7G^{5'}ppp^{5'}AmpA$ where the $m^7G^{5'}ppp^{5'}Am$ part forms B-B stack. For clarity certain atoms are omitted and the $m^7G^{5'}ppp$- part is shown in red. The three sugar rings are in ^3E conformation and the various torsion angles are: $\chi_c = 20°$; $\psi_c = 60°$; $\phi_c = 180°$; $\omega c_1 = 330°$; $\omega' c_1 = 230°$; $\omega c_2 = 60°$; $\omega' c_2 = 110°$; $\omega c_3 = 205°$; $\omega' c_3 = 320°$; $\phi_1 = 180°$; $\psi_1 = 60°$; $\chi_1 = 30°$; $\phi_1' = 210°$; $\phi'' = 90°$; $\omega_1' = 290°$; $\omega_1 = 300°$; $\phi_2 = 180°$; $\psi_2 = 60°$; $\chi_2 = 40°$. The designations of the torsion angles are illustrated in Figure 1.

H(5′), and H(5″). However, before the ring current theory is applied it is necessary to carefully evaluate the contribution to the data in Table X from torsional processes and molecular events other than stacking interactions. For example, the shifts of H(5′) and H(5″) could be affected by torsional changes in the phosphate backbone as a dimer is formed from the monomers. Hence, they are of little use to determine stacking interactions from ring current considerations.

A very important observation in the dimerization data in Table X is that the H(2′)'s of $m^7G^{5'}p$- and -$p^{5'}Am$ have undergone significant shielding to higher fields. This observation can be rationalized only on the basis of the presence of significant populations of A–A stacking (Figure 14). Inspection of the perspectives (Figures 14–17) clearly reveals that only in an A–A stack do these two protons lie in the inside of the stack. Thus, H(2′) of the $m^7G^{5'}p$- moiety can be shielded by the adenine of the -$p^{5'}Am$ moiety and the H(2′) of -$p^{5'}Am$ can be shielded by the guanine ring of $m^7G^{5'}p$-. *This mutual shielding is possible only in an A–A stack.* The observed difference (Table X) in the magnitude of shielding between the sets of H(2′) protons of two moieties is merely a reflection of the difference in the ring current fields of adenine and guanine moieties. The ring current shielding abilities of guanine are less than those of adenine.[44] In the present case the guanine ring contains a positive charge and that will further dilute the effect of guanine. The data in Table X clearly show that the adenine shields the H(2′) of $m^7G^{5'}p$- considerably more effectively than guanine shields the H(2′) of -$p^{5'}Am$. This is indeed expected.

One may attempt to evaluate quantitatively the above arguments from ring current calculations. We have emphasized elsewhere[43] the extreme caution that should be employed in applying this type of calculation and the information, while valuable, should be used only for arriving at general trends. We outline below how the dimerization shifts in Table X can be theoretically calculated.

The various torsion angles in Figures 14–17 are at first converted to x,y,z coordinates. Then these coordinates were used to calculate the x,y,z coordinates of the centers of the five- and six-membered rings of the purine residues. In the next step,

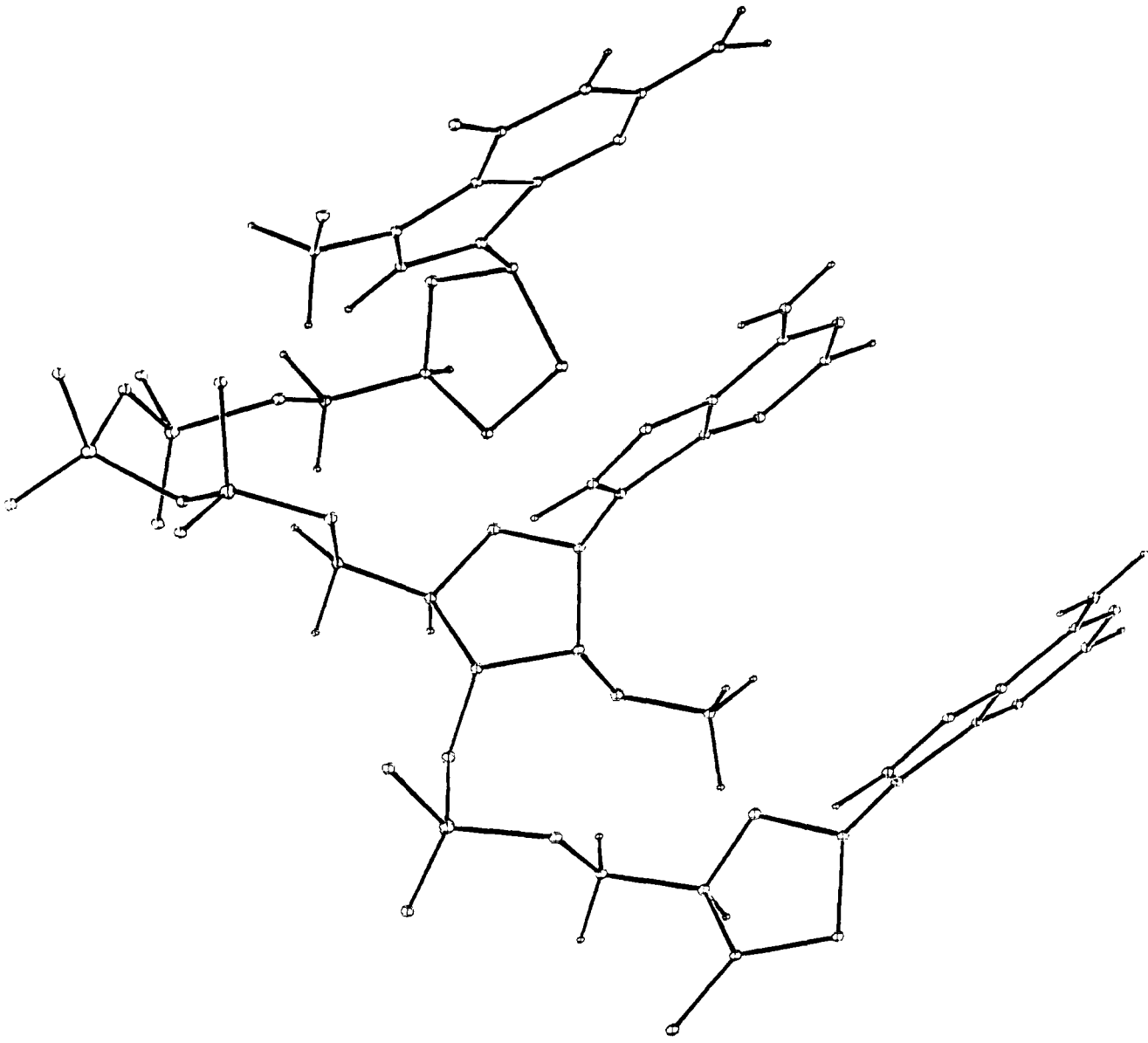

Figure 21. A model for $m^7G^{5'}ppp^{5'}AmpA$ when the $m^7G^{5'}ppp^{5'}Am$ part forms A-B stack. For clarity certain atoms are omitted and the $m^7G^{5'}ppp$- part is shown in red. The three sugar rings are in ^3E conformation and the various torsion angles are: $\chi_c = 30°$; $\psi_c = 60°$; $\phi_c = 180°$; $\omega c_1 = 200°$; $\omega' c_1 = 250°$; $\omega c_2 = 60°$; $\omega' c_2 = 120°$; $\omega c_3 = 200°$; $\omega' c_3 = 345°$; $\phi_1 = 180°$; $\psi_1 = 60°$; $\chi_1 = 70°$; $\phi_1' = 210°$; $\phi'' = 90°$; $\omega_1' = 290°$; $\omega_1 = 300°$; $\phi_2 = 180°$; $\psi_2 = 60°$; $\chi_2 = 80°$. The designations of the torsion angles are illustrated in Figure 1.

a second set of x,y,z coordinates for the various atoms was generated using the centers of the five- and six-membered rings as origins. Once this is accomplished, the second set of the x,y,z coordinates was used to generate the cylindrical coordinates z, ρ_5, and ρ_6 which were used in conjunction with references[44] to arrive at the projected shieldings in the various stacked arrays. A comparison of the projected and observed shieldings in A-A (Figure 14), A-B (Figure 15), B-A (Figure 16), and B-B (Figure 17) stacks clearly revealed that only in the A-A stack is there a general agreement between projected and observed shieldings. The data for the A-A stacked arrays are summarized in Tables XI and XII. While the data indicate general agreement between the projections and observations, the following points should also be considered. Data for the shielding of the $m^7G^{5'}p$- protons by the adenine ring (Table XI) indicate that the projected value for H(2′) is slightly larger than observed and for H(1′) it is smaller than observed. This is entirely expected because upon dimerization the χ_{CN} of $m^7G^{5'}p$- decreases. Such a decrease in χ_{CN} alone will cause the H(1′) to be *shielded* and H(2′) to be *deshielded*,[39,40] i.e., for H(1′) the observed shielding will be the sum of shielding from χ_{CN} reduction and stacking and hence will be larger than the projected one which takes into consideration only the stacking. In the case of H(2′), the reverse is true, i.e., the large shielding originating from stacking (i.e., 0.37 ppm, Table XI) will be internally compensated by the deshielding originating from the χ_{CN} reduction of $m^7G^{5'}p$- upon dimerization.

Examination of the data for the protons of $-p^{5'}Am$ indicates that, in general, the projections are smaller than observed ones. This is also expected because the conformation of the ribose of the $-p^{5'}Am$ part is ^2E and it is established[44] that when a ^2E pucker, as in B-DNA, is involved in the structures, the projections are always smaller than observed.

All in all our analysis of the dimerization data using ring current theory clearly indicates that in aqueous solution the dimer $m^7G^{5'}ppp^{5'}Am$ populates significantly in A-A stacked arrays. Because of the inherent flexibility of the various phosphodiester linkages, in aqueous solution it may be that the dimer $m^7G^{5'}ppp^{5'}Am$ exists as a conformational blend of the various stacked arrays and extended forms. Our major conclusion is that *among the various conformational possibilities in aqueous solution the A-A stack predominates and the*

molecule has an inherent proclivity to form A–A stacked arrays. The temperature data on $m^7G^{5'}ppp^{5'}Am$ summarized in Table X and Figure 18 closely parallel the dimerization trends and these observations are essentially in agreement with the observations of Ts'o[45] and co-workers that elevation of temperature causes destacking and at higher temperatures the dimer tends to approximate to monomer conformations.

Spatial Configuration of the mRNA Terminus $m^7G^{5'}ppp^{5'}$AmpA. From the discussion above one knows the geometric features of the segments $m^7G^{5'}ppp^{5'}Am$ and AmpA. Now using the stereochemical principles which govern the assembly of larger structures from smaller units in aqueous solution[46,47] one should be able to assemble the structure of the mRNA 5′ terminus $m^7G^{5'}ppp^{5'}$AmpA.

An important feature of the conformation of AmpA (Figure 11a,b) is that the A side of the adenine of the 3′-nucleotidyl unit (to whose free 5′ end the $m^7G^{5'}ppp$- will be attached to form the $m^7G^{5'}ppp^{5'}$AmpA structure) is facing the B side of the adenine of the 5′-nucleotidyl unit, i.e., the dimer AmpA prefers A–B stacking. Hence, the only free side that is available for stacking for the guanosine moiety in $m^7G^{5'}ppp^{5'}$AmpA is the B side of Amp-. If the AmpA segment of $m^7G^{5'}ppp^{5'}$AmpA maintains a conformation similar to free AmpA, the only available stacking possibilities for the $m^7G^{5'}ppp^{5'}Am$ segment in $m^7G^{5'}ppp^{5'}$AmpA are the A–B and B–B stackings (Figures 15 and 17). This is unsatisfactory because the dimer $m^7G^{5'}ppp^{5'}Am$ displays an intrinsic proclivity to form A–A stacked arrays (Figure 14). *The only stereochemical recourse available for $m^7G^{5'}ppp^{5'}$AmpA to form A–A stacked arrays between $m^7G^{5'}p$- and Amp- residues is for the $m^7G^{5'}p$- to intercalate in between the bases of the AmpA segment.* Obviously for $m^7G^{5'}p$- to intercalate, there should be enough space available between the two adenines of AmpA. This is indeed true because, as has been discussed before, the 2′-*O*-methyl group of AmpA causes AmpA to adopt an extended structure. In such an extended structure (Figure 11b) there is sufficient space between the adenine moieties so that the $m^7G^{5'}p$- moiety of $m^7G^{5'}ppp^{5'}$AmpA can intercalate between the adenines. We propose such an intercalated model for the $m^7G^{5'}ppp^{5'}$AmpA segment of the mRNA 5′ terminus and a perspective of this structure is shown in Figure 19. Some important and unique features of this structure are the following.

(1) The three bases form a closely stacked array contributing to its stability. The stacking is such that the $m^7G^{5'}ppp^{5'}Am$ part forms A–A stacks. All the torsion angles in the $m^7G^{5'}ppp^{5'}$AmpA structure are given in the legend of Figure 19. Comparison of these torsion angles with those in the A–A stacked dimer $m^7G^{5'}ppp^{5'}Am$ (see the legend of Figure 14) shows that in the trimer χ_c, χ_1, $\omega' c_2$, and ωc_3 have undergone changes. These changes were introduced because one anticipates that the three bases in Figure 19 will form a closely packed stacked array in which the distance between the base planes will be about 3.4 Å. In the A–A stack for the dimer (Figure 14) this distance was calculated to be 5.4 Å.

(2) The structure enables us to rationalize the fundamental role of the 2′-*O*-methyl group which is invariably present in the structure of the mRNA 5′ terminus. We believe that its role is to induce a conformational change so much so that the backbone of the AmpA segment assumes an extended spatial configuration enabling the $m^7G^{5'}p$- residue to intercalate in between the adenine bases.

(3) In the intercalated trimer (Figure 19) the 2′-*O*-methyl, 7-methyl, H(2′), and H(3′) of the Amp- and $m^7G^{5'}p$- parts are in close vicinity creating hydrophobic interactions and a water-free channel.

(4) The AmpA segment in the intercalated complex (Figure 19) displays conformational features akin to those of dinucleoside monophosphates which carry intercalated ethidium bromide.[48] In accordance with Sobells' discovery[48] we find mixed sugar pucker and perturbation of χ_{CN}, ϕ, ϕ', ω', and ω.

(5) The structure enables us to rationalize the necessity of a long triphosphate tail because such a long structure is necessary to bend around so that the $m^7G^{5'}p$- part can be in between the adenines.

(6) In the proposed structure the 2-amino group of $m^7G^{5'}p$- is freely exposed. This amino group has been implicated to play a crucial part in protein synthesis[19] probably via interaction with the initiation factors, and such interactions become accessible when the amino group is on the periphery of the intercalated complex.

Even though in the legend of Figure 19 we give a complete listing of the probable torsion angles in $m^7G^{5'}ppp^{5'}$AmpA, those around the phosphodiester backbones are obtained from computer search to satisfy the intrinsic proclivity of $m^7G^{5'}ppp^{5'}Am$ to form A–A stacked arrays. Further, we have assumed that in the trimer the distance between the bases is about 3.4 Å. This intricate and elaborate methodology is necessitated by the inability of NMR methods to give directly the torsion angles about the phosphodiester bonds. However, the evidence presented is reasonable, bordering on the compelling, that the overall spatial configuration of the mRNA terminus $m^7G^{5'}ppp^{5'}$AmpA is a self-intercalating one which imparts a characteristic feature to the terminus. Obviously, one cannot rule out contributions from nonintercalating structures.

While the proposed structure of $m^7G^{5'}ppp^{5'}$AmpA (Figure 19) makes an important contribution to the aqueous solution conformational blend, there are important candidates which can make a significant, contribution to the conformational mixture. This is because our analysis of data has indicated that AmpA can also exist in a stacked array (Figure 11a) and that for the dimer $m^7G^{5'}ppp^{5'}Am$, molecular topologies other than the predominant A–A stacks are accessible.

Hence, one could integrate a stacked structure of AmpA (Figure 11a) into $m^7G^{5'}ppp^{5'}$AmpA so that the $m^7G^{5'}ppp^{5'}Am$ segment maintains A–B and/or B–B stacked arrays in the trimer. Perspectives of these nonintercalating contributions to the conformational blend of $m^7G^{5'}ppp^{5'}$AmpA are shown in Figures 20 and 21. A few of the torsion angles employed (see legends of Figures 20 and 21) have undergone small deviations from that encountered in A–B and B–B stacks of the dimer (Figures 15 and 17). This is necessary to achieve some overlapping and approximate parallel stackings among the base planes. Our studies on nucleic acid structures[18-21,46,47] have indicated that unstacked arrays also contribute to the conformational blend in aqueous solution. Hence, it is anticipated that structures such as $m^7G^{5'}ppp^{5'}$AmpA in aqueous solution will also show some population of conformers of random type with little stacking interactions. Finally, it should be noted that the proposed structures for the dimer and trimer can be highly influenced by salts such as Mg^{2+} or polyamines because of the presence of large numbers of phosphate groups.

Acknowledgments. This research was supported by Grant No. CA12462 from the National Cancer Institute of the National Institutes of Health and PCM 75-16406 from the National Science Foundation. This research was also supported in part by National Institutes of Health Grant No. 1-P07-PR00798 from the Division of Research Resources. We thank Dr. G. Govil and Dr. M. M. Dhingra for several of the computer programs which were of use in this study. The authors thank the University Computer Center, particularly Mr. J. Quinn, for assistance in the course of this work. We are indebted to Dr. H. Bushweller, Dr. S. Hoogasian, and Dr. J. Zubieta for their help with the ORTEP programs.

References and Notes

(1) A preliminary communication on this subject has been published: *Nature (London)*, **270**, 223 (1977).
(2) F. Rottman, A. J. Shatkin, and R. P. Perry, *Cell*, **3**, 197 (1974).
(3) Y. Furuichi and K.-I. Miura, *Nature (London)*, **253**, 374 (1975).
(4) Y. Furuichi, M. Morgan, S. Muthukrishnan, and A. J. Shatkin, *Proc. Natl. Acad. Sci. U.S.A.*, **72**, 372 (1975).
(5) Y. Furuichi, S. Muthukrishnan, and A. J. Shatkin, *Proc. Natl. Acad. Sci. U.S.A.*, **72**, 742 (1975).
(6) T. Urushibara, Y. Furuichi, C. Nishmura, and K.-I. Miura, *FEBS Lett.*, **49**, 385 (1975).
(7) C. M. Wei and B. Moss, *Proc. Natl. Acad. Sci. U.S.A.*, **72**, 318 (1975).
(8) G. Abraham, D. P. Rhodes, and A. K. Banerjee, *Nature (London)*, **255**, 37 (1975).
(9) S. Muthukrishnan, G. W. Both, Y. Furuichi, and A. J. Shatkin, *Nature (London)*, **255**, 33 (1975).
(10) Y. Furuichi, M. Morgan, A. J. Shatkin, W. Jenlinek, M. Daldítt-Georgiev, and J. E. Darnell, *Proc. Natl. Acad. Sci. U.S.A.*, **72**, 1904 (1975).
(11) J. Adams and S. Cory, *Nature (London)*, **255**, 28 (1975).
(12) S. Muthukrishnan, M. Morgan, A. K. Banerjee, and A. J. Shatkin, *Biochemistry*, **15**, 5761 (1976).
(13) G. W. Both, A. K. Banerjee, and A. J. Shatkin, *Proc. Natl. Acad. Sci. U.S.A.*, **72**, 1189 (1975).
(14) Because the 5′ terminus of mRNA combines the structures of nucleotide coenzymes and nucleic acids the original investigators[2-13] have employed clever nomenclature that does not conform to IUPAC-IUB recommendations. Thus: m^7G = 7-methylguanosine; $m^7G^{5'}p$ = 7-methylguanosine 5′-monophosphate; $m^7G^{5'}pp$ = 7-methylguanosine 5′-diphosphate; $m^7G^{5'}ppp$ = 7-methylguanosine -triphosphate; $m^7G^{5'}ppp^{5'}AmpA$ = structure in Figure 1; m^6Am = N^6-monomethyl-2′-*O*-methyladenosine; $N^{6'}p$ or $p^{5'}N$, etc., stand for 5′-NMP, etc., where N is A, G, U, or C. Nm is 2′-O-methylated nucleoside.
(15) D. Shafritz, J. A. Weinstein, B. Safer, W. Merrick, L. A. Weber, E. D. Hickey, and C. Baglioni, *Nature (London)*, **261**, 291 (1976).
(16) E. D. Hickey, L. A. Weber, and C. Baglioni, *Proc. Natl. Acad. Sci. U.S.A.*, **73**, 19 (1976).
(17) E. D. Hickey, L. A. Weber, C. Baglioni, C. H. Kim, and R. H. Sarma, *J. Mol. Biol.*, **109**, 173 (1977).
(18) H. Singh, M. H. Herbut, C. H. Lee, and R. H. Sarma, *Biopolymers*, **15**, 2167 (1976).
(19) D. M. Cheng and R. H. Sarma, *Biopolymers*, **16**, 1687 (1977).
(20) C. H. Lee, F. S. Ezra, N. S. Kondo, R. H. Sarma, and S. S. Danyluk, *Biochemistry*, **15**, 3627 (1976).
(21) F. S. Ezra, C. H. Lee, N. S. Kondo, S. S. Danyluk, and R. H. Sarma, *Biochemistry*, **16**, 1977 (1977).
(22) R. H. Sarma and R. J. Mynott, *J. Am. Chem. Soc.*, **95**, 1641 (1973).
(23) R. H. Sarma and R. J. Mynott, *J. Am. Chem. Soc.*, **95**, 7470 (1973).
(24) M. Sundaralingam, *Jerusalem Symp. Quantum Chem. Biochem.*, **5**, 417 (1973).
(25) M. Sundaralingam, *Struct. Conform. Nucleic Acids Protein–Nucleic Acid Interact. Proc. Annu. Harry Steenbock Symp., 4th, 1974*, 487 (1976).
(26) C. H. Lee and R. H. Sarma, *J. Am. Chem. Soc.*, **98**, 3541 (1976).
(27) R. H. Sarma, C. H. Lee, F. E. Evans, N. Yathindra, and M. Sundaralingam, *J. Am. Chem. Soc.*, **96**, 7337 (1974).
(28) R. H. Sarma and R. J. Mynott, *Org. Magn. Reson.*, **4**, 577 (1972).
(29) C. H. Lee and R. H. Sarma, *Biochemistry*, **15**, 697 (1976).
(30) F. E. Hruska, *Jerusalem Symp. Quantum Chem. Biochem.*, **5**, 345 (1973).
(31) S. H. Kim, H. M. Berman, N. C. Seeman, and M. D. Newton, *Acta Crystallogr., Sect. B*, **29**, 703 (1973).
(32) J. M. Rosenberg, N. C. Seeman, S. H. Kim, F. L. Suddath, H. B. Nicholas, and A. Rich, *Nature (London)*, **243**, 150 (1973).
(33) B. Hingerty, E. Subramanian, S. D. Stellman, S. B. Broyde, T. Sato, and R. Langridge, *Biopolymers*, **14**, 227 (1975).
(34) J. Rubin, T. Brennan, and M. Sundaralingam, *Biochemistry*, **11**, 3112 (1972).
(35) W. K. Olson, *Biopolymers*, **12**, 1787 (1973).
(36) N. Yathindra and M. Sundaralingam, *Biopolymers*, **12**, 2261 (1973).
(37) S. B. Broyde, R. H. Wartell, D. Stellman, B. Hingerty, and R. Langridge, *Biopolymers*, **14**, 1597 (1975).
(38) H. Berthod and B. Pullman, *Biochim. Biophys. Acta*, **232**, 595 (1971).
(39) C. Giessner-Prettre and B. Pullman, *J. Theor. Biol.*, **65**, 189 (1977).
(40) C. Giessner-Prettre and B. Pullman, *J. Theor. Biol.*, **65**, 171 (1977).
(41) O. Kennard, N. W. Isaacs, W. D. S. Motherwell, J. C. Coppola, D. L. Wampler, A. C. Larson, and D. G. Watson, *Proc. R. Soc. London, Ser. A*, **325**, 401 (1971).
(42) F. E. Hruska, A. Mak, N. Singh, and D. Shugar, *Can. J. Chem.*, **51**, 1099 (1973).
(43) F. E. Evans and R. H. Sarma, *Biopolymers*, **13**, 2117 (1974).
(44) C. Giessner-Prettre, B. Pullman, P. N. Borer, L. S. Kan, and P. O. P. Ts'o, *Biopolymers*, **15**, 2277 (1976).
(45) P. O. P. Ts'o, N. S. Kondo, M. P. Schweizer, and D. P. Hollis, *Biochemistry*, **8**, 997 (1969).
(46) F. E. Evans and R. H. Sarma, *Nature (London)*, **263**, 567 (1976).
(47) R. H. Sarma and S. S. Danyluk, *J. Quantum Chem., QBS* **4**, 269 (1977).
(48) C. C. Tsai, S. C. Jain, and H. M. Sobell, *Proc. Natl. Acad. Sci. U.S.A.*, **72**, 628 (1975).

Proton Inventories of a Serine Protease Charge-Relay Model in an Aprotic Solvent

John L. Hogg,* Ray Morris III, and Noel A. Durrant

Contribution from the Department of Chemistry, Texas A&M University, College Station, Texas 77843. *Received August 1, 1977*

Abstract: Solvent isotope effects in H_2O-D_2O mixtures have been used to probe the operation of a catalytic system which has been suggested as a model for the charge-relay chain of serine protease enzymes. The absence of any significant solvent isotope effects coupled with the shapes of the proton inventories rules out any significant proton transfer mechanisms for the imidazole promoted or imidazole-benzoate anion promoted reaction of *p*-nitrophenyl acetate in acetonitrile containing 1 M H_2O. In the present study it was shown that the observed reaction is not hydrolysis as previously suggested. Nucleophilic attack by hydrogen-bonded dimers is proposed as the mechanism of catalysis.

Since the discovery of a catalytic triad of amino acids at the active site of the serine protease, α-chymotrypsin,[1] there have been numerous attempts to verify the operation of the so-called charge-relay system.[2] Attempts to model the aspartic acid-histidine-serine charge-relay chain of the enzymes have met with varying degrees of success.[3,4] Even some enzymatic studies themselves have cast doubt on the operation of the charge-relay chain[5,6] although the interpretation of the proton inventory for the deacetylation of acetyl-α-chymotrypsin has been questioned.[7] Bender et al. have recently reported some success in modeling the intramolecular cooperativity of the carboxylate–imidazole–hydroxyl triad.[8]

Perhaps the most interesting intermolecular charge-relay model to be studied is that of Haake and coworkers.[9] They investigated the reaction of *p*-nitrophenyl acetate in 1 M water in acetonitrile as a function of imidazole and benzoate ion concentrations. This dipolar, highly aprotic solvent system may mimic the somewhat hydrophobic environment of many enzymes. Their results suggested a high degree of cooperativity between benzoate anion, imidazole, and water in the catalysis.

The observed rate constant, eq 1, contained terms suggesting modes of catalysis such as those represented in eq 2–4. The

$$k_{obsd} = k_1[\mathrm{Im}] + k_2[\mathrm{Im}]^2 + k_3[\mathrm{Im}][RCO_2^-] \quad (1)$$

0002-7863/78/1500-1590$01.00/0