Williams, D. E., & Reisfeld, R. A. (1964) *Ann. N.Y. Acad. Sci. 121*, 373.
Wisdom, C., & Welker, N. E. (1974) *J. Bacteriol. 120*, 748.
Wrigley, C. W. (1969) *Sci. Tools 15*, 17.
Wyatt, G. R. (1951) *Biochem. J. 48*, 584.
Yamamoto, K., Alberts, B., Benzinger, R., Hawthorne, L., & Treiber, G. (1970) *Virology 40*, 734.
Yoo, O. J., & Agarwal, K. I. (1980) *J. Biol. Chem. 255*, 6445.

Ambiguity and Transcriptional Errors as a Result of Modification of Exocyclic Amino Groups of Cytidine, Guanosine, and Adenosine†

B. Singer* and S. Spengler

ABSTRACT: Ribopolynucleotides containing nucleosides with modified exocyclic amino groups were transcribed by using DNA-dependent RNA polymerase in the presence of Mn^{2+} and all four ribonucleoside triphosphates. Nearest-neighbor analysis of the products revealed a variety of effects. N^4-Modified cytidines can act preferentially like uridine (U) as a result of the known tautomerism, but ambiguity is also observed in the case of N^4-hydroxycytidine and N^4-methylcytidine. No ambiguity results from N^4-methoxycytidine which acts only like U. N^4-Acetylcytidine, known to be in the anti conformation, base pairs only with guanosine (G), as expected. N^2-Methylguanosine acts ambiguously in directing all four nucleosides into a transcript. However, it shows preference for incorporation of adenosine (A). Three different size substituents on the N^6 of A did not affect A·U pairing, indicating a conformation in which these modifications must lie anti to the ring N-1. Our results on fidelity may be explained in terms of the tautomerism of cytidine and orientation of the substituent. N^4-Methoxycytidine, predicted to prefer the imino form, appears to have the substituent anti to the ring N-3, resulting in strict U-like behavior. For the amino forms, modifications lying anti to the Watson–Crick side permit normal base pairing. However, substituents in the syn position may block normal pairing; this leads to the ambiguity observed for N^4-hydroxycytidine, N^4-methylcytidine, and N^2-methylguanosine. If the group is both large and syn, e.g., N^6-isopentenyladenosine, the attempt to transcribe the base may result in an inactivating event, such as a frame shift or termination. For this large substituent, the isopentenyl group must be anti when unambiguous base pairing occurs. As a general hypothesis, ambiguity, which may lead to point mutations, will result when hydrogen bonds of the appropriate number or length cannot be formed. This may arise from either steric hindrance and electronic shielding of the sites or loss of the appropriate donor or receptor. The continuation of transcription when noncomplementary nucleotide incorporation occurs may not require that any hydrogen bonds be formed, but that only stacking and other energy considerations are involved.

The misincorporation of noncomplementary bases during in vitro transcription of polydeoxynucleotides has been found to occur (Paetkau et al., 1972; Sirover et al., 1979; Seal et al., 1979), and it is postulated that the polymerases used have an error rate which is influenced by the transcription conditions. Such errors are usually corrected during in vivo transcription. However, when they are not repaired, mutation may result. When the template contains modified bases, the error rate or infidelity of transcription can be greatly increased (Sirover & Loeb, 1974).

Polyribonucleotides containing specifically modified bases can also be transcribed by using DNA-dependent RNA polymerase in the presence of Mn^{2+}, and the transcript can be analyzed for noncomplementary bases. Such a technique has proven useful in studying the effect of many different modified bases in transcription [reviewed by Singer & Kröger (1979)].

In the previous paper from this laboratory, Kröger & Singer (1979b) used as templates polynucleotides containing about 10% of 3-methylcytidine, 3-methyluridine, or 1-methyladenine. These derivatives were chosen because the methyl group blocks one of the essential hydrogen-bonding sites, and it had been considered that all were lethal modifications. However, it was found that all three methylated derivatives act with complete ambiguity and can direct any nucleotide into the transcript. None appeared to stop transcription.

The present series of experiments was designed to investigate whether modification of any exocyclic amino group causes ambiguity even though a single substitution of the N^2 of G, N^6 of A, or N^4 of C should not affect normal base pairing unless the substituent is oriented syn to the base-pairing side. For the purpose of comparison, we also studied several derivatives modified on sites which do not block Watson–Crick base pairing. These included 5-substituted pyrimidines, isoadenosine, 4-thiouridine, and 2′-*O*-methylcytidine. Nearest-neighbor analysis of transcription products was used as a sensitive test for misincorporation.

Materials and Methods

Chemicals and Polyribonucleotides. [α-^{32}P]GTP (20–30 Ci/mmol) was purchased from New England Nuclear. N^4-Hydroxy-CDP was prepared according to Janion & Shugar (1968). N^4-Methoxy-CDP was prepared in an analogous way by using methoxyamine (*O*-methylhydroxylamine). N^2-Methyl-GDP, 5-hydroxy-UDP, N^6-(hydroxyethyl)-ADP, and iso-ADP were generous gifts from Dr. A. M. Michelson, Institut de Biologie, Paris. 4-Thio-UDP, N^6-isopentenyl-ADP, and unmodified nucleoside diphosphates were purchased from P-L Biochemicals. We are also indebted to Dr. A. M. Michelson for poly(C,N^4-acC) and poly(A,N^4-acC). Poly(2′-*O*-

† From the Department of Molecular Biology and Virus Laboratory, University of California, Berkeley, California 94720. *Received August 7, 1980.* This investigation was supported by Grant CA 12316 from the National Cancer Institute.

0006-2960/81/0420-1127$01.25/0

Table I: Effect of Modified Nucleosides on Transcription of Poly(C). Fidelity of Transcription When the Exocyclic Amino Group of Cytidine Is Modified[a]

template	modification (%)[b]	^{32}P radioactivity (%) in nearest-neighbor sequence[c]				total transcription[d] % of poly(C)
		CpG	ApG	UpG	total av	
poly(C,mo^{4}C)	3	–	3.6, 4.0	–	3.8	59
	6	–	6.6, 6.9	–	6.8	57
	9	–	10.6, 10.4	–	10.5	33
	16	0.3, 0.3	13.4, 13.8	–	13.9	20
poly(C,ho^{4}C)	8	–	2.7, 2.7	–	2.7	53
	9.5	0.3, 0.2	3.6, 3.7	1.5, 0.8	4.9	35
	16	1.9, 1.1	8.1, 7.4	4.3, 4.2	13.6	14
poly(C,ac^{4}C)	70	–	(1.0, 0.4)[e]	–	0.7	120
poly(C,m^{4}C)	~20	0.6, 0.3	3.0, 1.1	5.7, 1.2	6.0	31

[a] All four nucleoside triphosphates were present in equal amounts. GTP was α-^{32}P labeled. See Materials and Methods. [b] The composition of all polymers was analyzed by LC. There was no detectable uridine in any polymer except poly(C,70% ac^{4}C), which contained 2% U. [c] The nonspecific incorporation directed by the poly(C) carrier is subtracted. Radioactivity in CpG sequences indicates that the modified nucleoside simulated the presence of G. Similarly, ApG and UpG radioactivities indicate simulation of U and A, respectively. [d] Total transcription includes GpG sequences which are not shown in the table. Separate experiments with poly(A) and poly(U) as carriers indicate that ac^{4}C and m^{4}C direct incorporation of GMP, thus behaving as the unmodified base [poly(A) data in Table IV, poly(U) data not shown]. In similar experiments with poly(A,22% ho^{4}C), GMP incorporation is not detected (Table IV). However, with a lower percent of ho^{4}C, GMP is incorporated as well as AMP (Singer & Fraenkel-Conrat, 1970; Fraenkel-Conrat & Singer, 1971). [e] LC analysis of enzyme-digested poly(C,ac^{4}C) showed contamination with about 2% U, probably due to deamination.

methylcytidylate) was purchased from P-L Biochemicals. All other polymers were prepared by using the method described by Singer & Kröger (1978).

Enzymes. Polynucleotide phosphorylase (P-L Biochemicals), RNA polymerase *E. coli* K12 (Miles Laboratories), snake venom phosphodiesterase, bacterial alkaline phosphatase, and acid phosphatase (Worthington) were all commercial preparations.

Analysis of Polynucleotide Composition. About 0.3 absorbancy unit (at λ_{max}) of each polyribonucleotide in 0.1 M Tris,[1] pH 7.3, was incubated at 37 °C for 20 h with 5 μg each of snake venom phosphodiesterase, alkaline phosphatase, and acid phosphatase. All polymers were digested to nucleosides under these conditions. Liquid chromatography (LC) analysis was performed as described by Kröger & Singer (1979b). Elution was with 0.4 M ammonium formate, pH 7, which separated any contaminating uridine from other nucleosides and modified nucleosides. Polymer composition was by peak integration. The composition of polymers of C with very low percentages of N^4-methylcytidine could not be determined with great accuracy due to the closeness of their elutions. All other nucleosides in copolymers were well separated. The composition of polymers is shown in the tables.

Transcription of Polyribonucleotides and Nearest-Neighbor Analysis. The 625-μL standard incubation mixture contained 0.15 absorbance unit of polyribonucleotide, was 0.4 mM each in ATP, CTP, UTP, and [α-^{32}P]GTP (15 μCi), and contained 4 mM $MnSO_4$, 0.8 mM K_2HPO_4, 40 mM β-mercaptoethanol, 15 μg of RNA polymerase. In "noncompetitive" experiments, only one other nucleoside triphosphate was used in addition to [α-^{32}P]GTP. The concentration was then 0.8 mM for each triphosphate in order to maintain a constant 1.6 mM triphosphate concentration. After a 2-h incubation at 37 °C, 75 μL was spotted on DEAE paper disks and washed 7 times with 7% Na_2HPO_4, briefly twice with H_2O, and twice with ethanol. After the paper was dried, the radioactivity on the disk was counted by using toluene scintillation fluid. The count in the aliquot gives a measure of total [^{32}P]GMP incorporation.

[1] Abbreviations used: mo^{4}C, N^4-methoxycytidine; ho^{4}C, N^4-hydroxycytidine; ac^{4}C, N^4-acetylcytidine; m^{4}C, N^4-methylcytidine; m^{2}G, N^2-methylguanosine; i^{6}A, N^6-isopentenyladenosine; he^{6}A, N^6-(hydroxyethyl)adenosine; m^{6}A, N^6-methyladenosine; Cm, 2′-*O*-methylcytidine; br^{5}C, 5-bromocytidine; ho^{5}U, 5-hydroxyuridine; s^{4}U, 4-thiouridine; Tris, tris(hydroxymethyl)aminomethane.

Nearest-neighbor analysis was performed as described by Kröger & Singer (1979b). With poly(C) as template, there is always some nonspecific incorporation of ^{32}P in CMP, AMP, and UMP, attributed to streaking from the high GMP radioactivity. The percent of such counts compared to the total incorporation of GMP was found to be similar regardless of the amount of poly(C) transcribed (over 2 orders of magnitude). Therefore, even when polymers containing modified bases are poorly transcribed, the same percent of nonspecific radioactivity as in polymers with more transcription can be subtracted as background. In each nearest-neighbor experiment, a poly(C) transcript is included to give this background. Only radioactivity more than 3 times background is considered to represent significant incorporation resulting from the modified base in the polymer.

Results

Fidelity of Transcription in Copolymers with Cytidine. (a) N^4-Substituted Cytidines. The four modifications used were hydroxy, methoxy, acetyl, and methyl. Of these, only the acetyl derivative did not direct [^{32}P]GTP incorporation to any nucleotide other than Gp, thus indicating that it is base pairing normally, even when there is 70% ac^{4}C in the copolymer (Table I).

In contrast, the methyl and hydroxy derivatives directed incorporation of A, C, and U (Table I). ho^{4}C simulates U preferentially since Ap predominates in nearest-neighbor analysis and is the only nucleotide detectable when transcribing the copolymer of C, 8% ho^{4}C with the α-^{32}P label in GTP.

The simulation of U by ho^{4}C is expected as a result of the well-known tautomerism of this derivative. When mo^{4}C, the analogue of ho^{4}C, is in a copolymer with C, the transfer of ^{32}P is almost entirely to Ap (>98%) (Table I). Furthermore, the proportion of total GTP incorporation transferred to Ap parallels the proportion of mo^{4}C in the polymer. This would suggest that mo^{4}C is entirely in the imino form while ho^{4}C is only partially in this form. Increasing the proportion of mo^{4}C or ho^{4}C in poly(C) decreases the rate of transcription, but this is not considered significant since copolymers are generally transcribed less efficiently than homopolymers, probably as a result of changed secondary structure (Kröger & Singer, 1979b).

When copolymers with a C carrier ("carrier" refers to the unmodified or major component of copolymers) are transcribed, only indirect evidence can be obtained as to whether

Table II: Effect of Modified Nucleosides on Transcription of Poly(C). Fidelity of Transcription When the Exocyclic Amino Group of Purines Is Modified[a]

template	modification (%)[b]	^{32}P radioactivity (%) nearest-neighbor sequence CpG	ApG	UpG
poly(C,m^{2}G)	7	0.7	2.1	0.5
poly(C,i^{6}A)	2.6	–	–	1.2
	7	–	–	2.5
	17	–	–	3.8
poly(C,he^{6}A)	14	–	–	7.3
	45	–	–	22[c]
	65	–	–	45[c]
poly(C,m^{6}A)[d]	8	–	–	5.7

[a] All four nucleoside triphosphates were present in equal amounts. GTP was α-^{32}P labeled. [b] Percent modification was determined by LC analysis after digestion of polymers with snake venom phosphodiesterase and phosphatase. No unmodified nucleosides other than C were detected (<0.5%). [c] No ^{32}P transfer occurs to an unmodified nucleotide when two modified nucleotides are adjacent. Therefore, polynucleotides with clustered modified nucleotides will not give quantitative data with nearest-neighbor analysis. [d] Data from Kröger & Singer (1979a,b).

Table III: Effect of Modified Nucleosides on Transcription of Poly(C). Fidelity of Transcription When Modification Does Not Block Watson-Crick Sites[a]

template	^{32}P radioactivity (%) in nearest-neighbor sequence[b] CpG	ApG	UpG	total transcription % of poly(C)
poly(Cm)	–	–	–	27
poly(br^{5}C)	–	(0.9)[d]	–	62
poly(C,38% iso-A)	–	–	27	8
poly(C,11% ho^{5}U)	–	1.6	–	63
poly(C,3% s^{4}U)[c]	–	2.8	–	50
poly(C,6% s^{4}U)[c]	–	3.5	–	30
poly(C,22% s^{4}U)	–	9.4	–	48
poly(C,9.8% U)	–	6.6	–	87
poly(C,A,G,U) 70:5:15:9	6.7	5.8	4.9	17

[a] See footnote *a*, Table I. [b] Data shown are averages of 2–3 separate experiments except where noted by footnote *c*. [c] Single experiment. [d] LC analyses of enzyme-digested poly(br^{5}C) showed contamination with 1% U, probably due to deamination.

the modified cytidine can still act as C. Data obtained by using A as the carrier are given in the next section. However, in the case of ac^{4}C, the fact that there is much G incorporation without misincorporation clearly reflects the unchanged pairing of ac^{4}C with G. Similarly, the stoichiometry of the misincorporation by mo^{4}C indicates that seldom, if ever, can there be normal C·G pairing. m^{4}C directs misincorporation of C, A, and U, but the total is much less than the percent of m^{4}C in the copolymer. Even without further data, it may be deduced that m^{4}C can also base pair with G.

(*b*) *N^2-Methylguanosine.* The polymer used contained 7% m^{2}G, which in transcription directed about 2% A and less than 1% each of U and C into the transcript (Table II). Thus, m^{2}G can simulate U and A, in addition to G. The total transcription was about 35% compared to that of poly(C). Kröger & Singer (1979b) reported that unmodified G or m^{1}G (1-methylguanosine) in copolymers caused a dramatic decrease in the rate of transcription. Under similar conditions, poly(C,12% G) was 8% transcribed while poly(C) was 50% transcribed. The relatively efficient transcription of a polymer containing 7% m^{2}G suggests that m^{2}G does not stop or slow down transcription more than the unmodified G.

(*c*) *N^6-Substituted Adenosines.* Adenosine with two large substituents, hydroxyethyl and isopentenyl, was copolymerized with cytidine. Regardless of the proportion of N^6-modified A, there was no indication of misincorporation (Table II). All [^{32}P]GMP incorporation was found transferred either to Gp or to Up, reflecting the expected behavior of a copolymer of C and A. In the case of he^{6}A, there was a reasonable correlation between the percent he^{6}A and the percent Up radioactivity. However, in polymers containing i^{6}A, the percent Up represented only a fraction of the i^{6}A present. This was particularly noticeable in poly(C,17% i^{6}A), where only 3.8% of the [^{32}P]GMP incorporation was transferred to Up. It is likely that the i^{6}A residue is inactivating as well as behaving like unmodified A. No clear indication of inactivation or interruption of transcription was found with any of the other derivatives studied, but the size of the isopentenyl group, if rotated into the plane (syn to the Watson–Crick site), would sterically interfere and not permit any nucleotide to be added to the transcript.

(*d*) *Modifications Not Blocking Watson–Crick Sites.* Table III presents transcription data derived from a series of polymers containing modified nucleosides which have been reported to base pair normally. These include 5-bromocytidine (Michelson & Monny, 1967; Means & Fraenkel-Conrat, 1971), 2′-*O*-methylcytidine (Rottman et al., 1974; Gerard et al., 1972), isoadenosine (Michelson et al., 1966), 5-hydroxyuridine (Massoulié et al., 1966), and 4-thiouridine (Simuth et al., 1970; Rackwitz & Scheit, 1977).

No misincorporations were detected, and, indeed, the Ap incorporation in transcripts of poly(br^{5}C) was accounted for by the small contamination of uridine. Poly(Cm) appeared to be a better template than that found by Gerard et al. (1972), who used a polymerase from a different bacterium.

Isoadenosine is adenosine with the ribose on the N-3. This does not change the base pairing with U. As was found for polymers containing 2-thiocytidine (Kröger & Singer, 1979a), the substitution of sulfur for oxygen in 4-thiouridine does not cause transcription errors.

Fidelity of Transcription in Copolymers with Adenosine and Uridine. In order to obtain further data on whether modified nucleosides simulated C, polymers were prepared by using U or A as the carrier. In these experiments, [^{32}P]GTP is also used with the other nucleoside triphosphates unlabeled. If there is any GMP incorporation resulting from the presence of the modified nucleoside, the label should then be transferred only to Ap when the carrier is U, or Up if the carrier is A. There can also be transfer to G if two GMP's are adjacent.

Table IV shows the results when A is the carrier. No percentages can be calculated since the normal complementary nucleotide, Up, is not labeled. Of the derivatives in the table, m^{4}C and ac^{4}C both act as C, while neither ho^{4}C nor he^{6}A acts detectably as C. The copolymers, poly(A,7% C) and poly(A,14% C), are included as positive controls, and both direct Gp.

Poly(U,m^{4}C) transcripts gave additional evidence that m^{4}C could act as C since [^{32}P]GMP radioactivity was found in Ap. The polymer contained 43% m^{4}C which resulted in high radioactivity in Gp as well.

Fidelity of Transcription under Noncompetitive Conditions. These experiments are designed to amplify mispairing by "forcing", that is, by using only two nucleoside triphosphates: GTP which is complementary to the C carrier and one other. If ambiguity is observed under competitive conditions, then under noncompetitive conditions, each mispairing should be increased. This is clearly the case for ho^{4}C and m^{2}G, but not for mo^{4}C, ac^{4}C, or m^{4}C (Table V). Poly(C,8% ho^{4}C) does

Table IV: Nearest-Neighbor Analysis of [α-^{32}P]GTP Incorporation Directed by Modified Nucleosides in Copolymers with Poly(A)[a]

template	nearest-neighbor sequence[b] (^{32}P cpm × 10^{-3}) C^{32}pG	A^{32}pG	U^{32}pG
poly(A,22% ho^{4}C)	0.9	1.5	2.6
poly(A,20% m^{4}C)	0.8	1.3	*3.4*
poly(A,30% ac^{4}C)	0.6	2.2	*12*
poly(A,12% he^{6}A)	0.4	2.5	1.2
poly(A,7% C)	0.7	2.4	*5.0*
poly(A,14% C)	0.6	2.2	*8.1*
poly(A)	0.4, 0.8	1.5, 2.4	1.2, 2.6

[a] All four nucleoside triphosphates were present in equal amounts. GTP was α-^{32}P labeled. Transfer of label to Gp (GpG) is not included since there is always self-polymerization when the template does not contain C. However, with poly(A,30% ac^{4}C), there were 54 000 cpm in Gp, indicating that clustering of ac^{4}C residues occurred. [b] Significant numbers are in italics. When a poly(A) carrier is transcribed with [α-^{32}P]GTP, radioactivity should be transferred only to Up when any GTP is incorporated as a result of the template containing a nucleoside which stimulates C.

not lead to measurable incorporation of Cp or Up (Table I). In forcing experiments with this polymer and those with higher amounts of ho^{4}C, it is evident that ho^{4}C can direct all nucleotides into transcripts. It is just as evident that mo^{4}C behaves in a very different way, and no nucleotide other than Ap is detected. The presence of m^{4}C does cause misincorporation, but forcing does not increase the amount.

No misincorporation results from noncompetitive conditions that does not occur under competitive conditions. Only the magnitude of misincorporation is changed. This is illustrated by the data in Table V. Poly(C,9.8% U) directs about 6% Ap, and poly(C,70% ac^{4}C) directs only GMP incorporation under both sets of transcription conditions.

Discussion

Experimental results determined by using transcription as an indirect tool to determine whether, and how, a modified base can participate in transcription have yielded data which can be interpreted in terms of orientation of substituents. Although this work utilizes a system which is not part of the normal replication scheme, it is capable of giving information which may be regarded as indicative of mutagenesis. In the two instances where misincorporation resulting from modified bases was studied by using deoxypolynucleotides and Mg^{2+}, no conflicts were reported between such results and those obtained with ribopolynucleotides and Mn^{2+} (Ludlum, 1970, 1971; Gerchman & Ludlum, 1973; Mehta & Ludlum, 1978).

The earlier work from this laboratory on ambiguity dealt with modification of positions which, regardless of orientation, must block normal base pairing. This lack of ability to form two or three hydrogen bonds with any base did not block transcription, but instead the modified base directed any nucleotide into the transcript (Kröger & Singer, 1979b). Results of the present experiments on nearest-neighbor analysis of transcription products when the template contains nucleoside with modified exocyclic amino groups are summarized in Table VI.

When an exocyclic amino group is modified and is in a fixed position, two situations may result. The substituent may be anti to the Watson–Crick sites and not affect base pairing, or it may be syn and therefore act as a block or shield. In addition, rotation can occur from syn to anti and vice versa.

X-ray crystallographic data indicate that the acetyl group in N^4-acetylcytidine is fixed in the anti position (Figure 1C)

Table V: "Forced" Misincorporation of Nucleosides Directed by Poly(C) Containing Modified Nucleosides[a]

poly(C) containing	[^{32}P]GTP + NTP	^{32}P radioactivity (%) in nearest-neighbor sequence[b] CpG	ApG	UpG
6% mo^{4}C	CTP	–		
	ATP		6.8 (6.8)	
	UTP			–
9% mo^{4}C	CTP	–		
	ATP		10.1 (10.5)	
	UTP			–
8% ho^{4}C	CTP	*1.0* (nd)		
	ATP		2.4 (2.7)	
	UTP			*3.5* (nd)
9.5% ho^{4}C	CTP	0.9 (0.3)		
	ATP		4.5 (3.9)	
	UTP			*5.2* (1.2)
16% ho^{4}C	CTP	*4.0* (1.5)		
	ATP		6.7 (7.7)	
	UTP			*12* (4.3)
70% ac^{4}C	CTP	–		
	ATP		– (0.7)	
	UTP			–
20% m^{4}C	CTP	0.8 (0.4)		
	ATP		1.4 (2.1)	
	UTP			1.8 (3.5)
7% m^{2}G	CTP	*2.5* (0.7)		
	ATP		3.1 (2.1)	
	UTP			*1.4* (0.5)
9.8% U	CTP	–		
	ATP		6.3 (6.6)	
	UTP			–

[a] [α-^{32}P]GTP and one other nucleoside triphosphate were present in equal amounts, conditions equivalent to those used in "noncompetitive" experiments (Kröger & Singer, 1979b). The total NTP (nucleoside triphosphate) concentration was 1.6 mM, which is the same as for experiments where all four NTPs are used. [b] The nonspecific incorporation directed by the poly(C) carrier is subtracted. The numbers in parentheses are the incorporations under competitive conditions, i.e., all four NTPs. Significant numbers are italicized. nd indicates that no significant incorporation occurred. See Materials and Methods for method of evaluating significance of data.

Table VI: Effect of Amino-Modified Nucleosides on Transcription Using DNA-Dependent RNA Polymerase[a]

modified nucleoside	simulates the presence of A	G	U	C
m^{4}C	±	±	±	++
ac^{4}C	–	–	–	++++
ho^{4}C	+	++	++	±
mo^{4}C	–	–	++++	–
m^{2}G	+	+	++	±
m^{6}A	++++	–	–	–
he^{6}A	++++	–	–	–
i^{6}A	++++	–	–	–

[a] Data on which this summary is based are in Tables I–V. Pluses are subjective. Minuses mean that there was no reproducible incorporation greater than 3 times the background.

(Parthasarathy et al., 1978). Our data confirm this since only GMP is incorporated when poly(C,N^4-acetyl-C) is transcribed. Crystal structures of m^{2}G, m^{4}C, or m^{6}A have not yet been reported. However, NMR data and other physicochemical studies on these modified structures are available although not in complete agreement with our results, possibly a consequence of using free bases or different experimental conditions (Shoup et al., 1972; Engel & von Hippel, 1974). Shoup et al. (1972) conclude, on the basis of NMR at −19 °C in dimethylformamide, that 1-methyl-N^4-methylcytosine is predominately (95%) in the syn configuration. When protonated, the N^4-

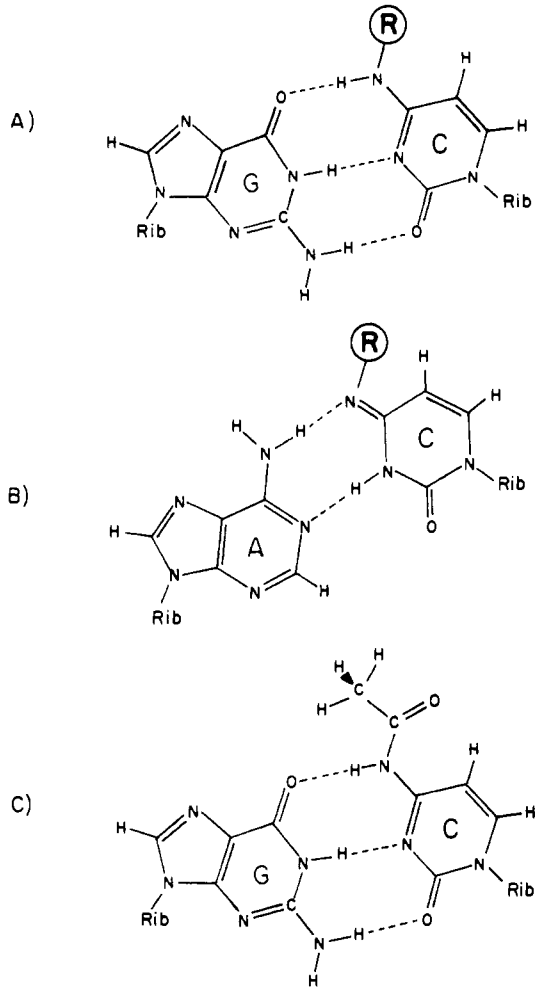

FIGURE 1: Hydrogen bonding of N^4-substituted C with A and G. (A) Normal G·C pairing when cytidine is in the amino form and the substituent is anti; R = OH, CH_3. (B) Pairing with A when modified cytidine is in the imino form and the substituent is anti; R = OH, OCH_3. (C) Pairing of ac^4C with G drawn according to the crystal structure of ac^4C (Parthasarathy et al., 1978).

methyl substituent is anti to the N-3 (Becker et al., 1965). Our data indicate that the methyl group can be anti since the normal complementary base is directed into the transcript. Studies of the secondary structure and base pairing of poly(C,m^4C) support the anti conformation (Brimacombe & Reese, 1966).

Engel & von Hippel (1974) believe that the methylated amino group of G freely rotates, and our data can be interpreted as agreeing with this structure. Figure 2A shows that if the methyl group is rotated into the plane of the base pair then it is probable that a G·C pair is not formed, and ambiguity results. In the orientation of Figure 2B, G·C pairing occurs. On the other hand, they report that in N^6-methyladenine the methyl group is syn to the ring N-1 which would block or shield any pairing. If this conformation existed in polymers, then N^6-isopentenyl-A could not act as A, which it does in transcription. The other N^6 derivatives, m^6A and he^6A, also act only as A. We conclude, on the basis of transcription results, that all three substituents are fixed in the anti configuration in a polynucleotide (Figure 3). If the isopentenyl group is rotated out of the plane, size alone would prevent even non-hydrogen-bonded helix formation. This also may occur (Table II).

Two modified derivatives of special interest are N^4-hydroxycytidine and N^4-methoxycytidine. Both have been found to exhibit a tautomeric shift to the imino form (Brown

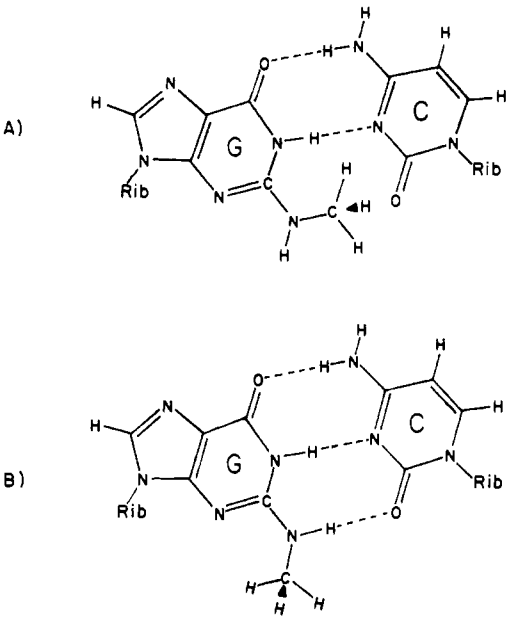

FIGURE 2: Possible hydrogen bonding of m^2G with C. (A) Methyl group rotated into plane of Watson–Crick pairing, leading to possible shielding of N-3 of C. (B) Methyl group rotated out of plane of Watson–Crick pairing, allowing normal base pairing.

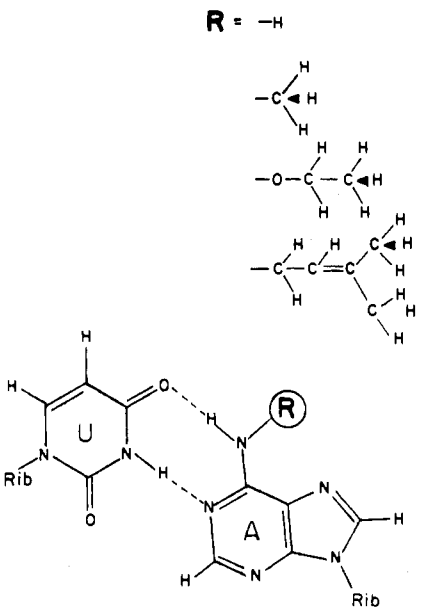

FIGURE 3: Hydrogen bonding of N^6-substituted A with U. R = substituents as shown above the base pair. Rotation of any substituent except the hydrogens would probably prevent the formation of any hydrogen bonds, by shielding (m^6A) or gross steric interference (he^6A, i^6A).

et al., 1968), and, in various types of biochemical experiments, they act like U to a significant extent (Figure 1B) (Budowsky et al., 1971, 1972; Müller et al., 1978; Sabo et al., 1977). The crystal structures of 1,5-dimethyl-N^4-hydroxycytosine and of 1-methyl-N^4-hydroxycytosine hydrochloride have been reported by Shugar et al. (1976) and Birnbaum et al. (1979) to be only in the imino form, but it is unclear from these results whether the amino form (normal C configuration) (Figure 1A) exists to any extent in the crystal and whether rotation of the hydroxyamino group could occur. We find that, in polynucleotides, N^4-hydroxycytidine acts primarily as U (imino) but can direct GMP incorporation. This would indicate that the amino form exists, although there appears to be about a 10:1 preference for the imino tautomer. In addition, there is

nonspecific incorporation of UMP and CMP. This ambiguity suggests that rotation occurs with a reasonable frequency, even for the imino form. Birnbaum et al. (1979) do not discuss whether the rotation from syn to anti occurs.

In the case of N^4-methoxy-C, the transcription results are striking since this derivative acts only as U, and only AMP is incorporated. No misincorporation of GMP incorporation was observed. On this basis, we conclude that N^4-methoxy-C exists only in the imino form and that the substituent is anti to the ring N-3.

A general hypothesis may be derived from these studies on fidelity in transcription. Ambiguity, rather than termination, will result if the appropriate number of hydrogen bonds cannot be formed. This may be due to steric hindrance or shielding, as well as substitution of a hydrogen bonding site. Regardless of the mechanism leading to ambiguity, the result will be point mutation.

References

Becker, E. D., Miles, H. T., & Bradley, R. B. (1965) *J. Am. Chem. Soc. 87*, 5575.

Birnbaum, G. I., Kulikowski, T., & Shugar, D. (1979) *Can. J. Biochem. 57*, 308.

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

Brown, D. M., Hewlins, M. J. E., & Schell, P. (1968) *J. Chem. Soc. C*, 1925.

Budowsky, E. I., Sverdlov, E. D., & Spasokukotskaya, T. N. (1971) *FEBS Lett. 17*, 336.

Budowsky, E. I., Sverdlov, E. D., & Spasokukotskaya, T. N. (1972) *Biochim. Biophys. Acta 287*, 195.

Engel, J. D., & von Hippel, P. H. (1974) *Biochemistry 13*, 4143.

Fraenkel-Conrat, H., & Singer, B. (1971) *Biological Effects of Polynucleotides* (Beers, R. F., & Braun, W., Eds.) p 13, Springer-Verlag, New York.

Gerard, G. F., Rottman, F., & Boezi, J. A. (1972) *Biochem. Biophys. Res. Commun. 46*, 1095.

Gerchman, L. L., & Ludlum, D. B. (1973) *Biochim. Biophys. Acta 308*, 310.

Janion, C., & Shugar, D. (1968) *Acta Biochim. Pol. 15*, 107.

Kröger, M., & Singer, B. (1979a) *Biochemistry 18*, 91.

Kröger, M., & Singer, B. (1979b) *Biochemistry 18*, 3493.

Ludlum, D. B. (1970) *J. Biol. Chem. 245*, 477.

Ludlum, D. B. (1971) *Biochim. Biophys. Acta 247*, 412.

Massoulié, J., Michelson, A. M., & Pochon, F. (1966) *Biochim. Biophys. Acta 114*, 16.

Means, G. E., & Fraenkel-Conrat, H. (1971) *Biochim. Biophys. Acta 247*, 441.

Mehta, J. R., & Ludlum, D. B. (1978) *Biochim. Biophys. Acta 521*, 770.

Michelson, A. M., & Monny, C. (1967) *Biochim. Biophys. Acta 149*, 88.

Michelson, A. M., Monny, C., Laursen, R. A., & Leonard, N. J. (1966) *Biochim. Biophys. Acta 119*, 258.

Müller, W., Weber, H., Meyer, F., & Weissmann, C. (1978) *J. Mol. Biol. 124*, 343.

Paetkau, V., Coulter, M. B., Flintoff, W. F., & Morgan, A. R. (1972) *J. Mol. Biol. 71*, 293.

Parthasarathy, R., Ginell, S. L., De, N. C., & Chheda, G. B. (1978) *Biochem. Biophys. Res. Commun. 83*, 657.

Rackwitz, H. R., & Scheit, K. H. (1977) *Eur. J. Biochem. 72*, 191.

Rottman, F., Friderici, K., Comstock, P., & Khan, M. K. (1974) *Biochemistry 13*, 2762.

Sabo, D. L., Domingo, E., Bandle, E. F., Flavell, R. A., & Weissmann, C. (1977) *J. Mol. Biol. 112*, 235.

Seal, G., Shearman, C. W., & Loeb, L. A. (1979) *J. Biol. Chem. 254*, 5529.

Shoup, R. R., Miles, H. T., & Becker, E. D. (1972) *J. Phys. Chem. 76*, 64.

Shugar, D., Huber, C. P., & Birnbaum, G. I. (1976) *Biochim. Biophys. Acta 447*, 274.

Simuth, J., Scheit, K. H., & Gottschalk, E. M. (1970) *Biochim. Biophys. Acta 204*, 371.

Singer, B., & Fraenkel-Conrat, H. (1970) *Biochemistry 9*, 3694.

Singer, B., & Kröger, M. (1978) *Anal. Biochem. 90*, 590.

Singer, B., & Kröger, M. (1979) *Prog. Nucleic Acid Res. Mol. Biol. 23*, 151.

Sirover, M. A., & Loeb, L. A. (1974) *Nature (London) 252*, 414.

Sirover, M. A., Dube, D. K., & Loeb, L. A. (1979) *J. Biol. Chem. 254*, 107.