Chloroacetaldehyde-Treated Ribo- and Deoxyribopolynucleotides. 2. Errors in Transcription by Different Polymerases Resulting from Ethenocytosine and Its Hydrated Intermediate†

J. T. Kušmierek‡ and B. Singer*

ABSTRACT: Chloroacetaldehyde-modified poly(rC) or poly(dC) was prepared containing either 8–36% 3,N^4-ethenocytidine (ϵC) or 8–36% of a mixture of ϵC and the hydrated ϵC (ϵC·H_2O), with the hydrate greatly predominating (>90%). These ribo- and deoxyribonucleotide templates were transcribed with DNA-dependent RNA polymerases from *Escherichia coli* and calf thymus, in the presence of either Mn^{2+} or Mg^{2+} and all four ribonucleoside triphosphates. All the polymers tested were transcribed with either cation present. In an earlier report from this laboratory [Spengler, S., & Singer, B. (1981) *Nucleic Acids Res. 9*, 365], transcriptional ambiguities resulting from ϵC residues in enzymatically synthesized poly(rC,ϵrC) were studied with *E. coli* DNA-dependent RNA polymerase in the presence of Mn^{2+}. The misincorporations there reported were confirmed when poly(rC,ϵrC) and poly(dC,ϵdC), prepared by reaction of poly(rC) and poly(dC) with CAA, were transcribed in the presence of either Mn^{2+} or Mg^{2+}. We now report that the presence of hydrated ϵC in polymers also leads to misincorporations but with reproducible differences from those found with ϵC alone. Nearest-neighbor analysis of the transcription products showed that the hydrate caused misincorporation of A > U ≫ C while ϵC caused misincorporation of U > A ≫ C. The extent of misincorporation in transcription was less with Mg^{2+} than with Mn^{2+}, but the pattern of ambiguity was the same with both cations and with both ribo- and deoxyribocytidylate polymers. Calf thymus DNA-dependent RNA polymerase IIB was also used to transcribe deoxyribocytidine polymers with Mn^{2+} as the cation. ϵC and ϵC·H_2O both caused a high level of misincorporation of U, A, and C, but the preferred misincorporations differed slightly from those found with *E. coli* DNA-dependent RNA polymerase. For both prokaryotic and eukaryotic enzymes, the type of misincorporation resulting from the loss of hydrogen bonding by modification of the N-3 of C not only differed between ϵC and the hydrated intermediate but also both differed from the transcriptional errors resulting from the presence of 3-methylcytidine in poly(dC) or poly(rC). We conclude that the errors made by these polymerases during transcription do not result primarily from the conditions used (cation, ribo- or deoxyribotemplate) but must be at least in part attributed to the enzyme recognizing some facet of the modified base other than the lack of normal hydrogen bonding.

One of the approaches to an understanding of the mechanism of chemically induced point mutations is to study the transcription of templates containing various modified bases in vitro. The transcribing enzymes available for this purpose differ in their fidelity, when tested for incorporation of noncomplementary nucleotides with unmodified synthetic polynucleotides as templates. The error rate in these in vitro transcriptions ranges from 10^{-4} to 10^{-5} for DNA polymerases through 10^{-3} to 10^{-4} for RNA polymerases to 10^{-2} to 10^{-3} for reverse transcriptases (Mildvan & Loeb, 1979). These error rates are very different from both the 10^{-8}–10^{-10} mutation frequency in vivo (Drake, 1969) and the 10^{-1}–10^{-2} error frequency expected as a result of a difference in free energy between correct and incorrect base pairing. The amplification of the fidelity of transcription by polymerases is due either to the active role of the enzyme in base selection or to the existence of kinetic proofreading, or both (Loeb & Kunkel, 1982). The differences in the accuracy of the functionally different polymerases may be accounted for by the differences in the modes of action of these error-preventing mechanisms. The postreplication repair of mismatches leads to further increase in accuracy of DNA replication in vivo.

On these grounds it can be expected that transcription by different polymerases of templates containing modified bases could lead to different levels of misincorporation. These misincorporations can also be affected by divalent metal ions, since divalent ions replacing Mg^{2+} in transcription of unmodified templates by *Escherichia coli* DNA polymerase I (Sirover et al., 1979) and by *E. coli* DNA-dependent RNA polymerase (Niyogi & Feldman, 1981) enhance the error rate.

Recent interest in the carcinogen vinyl chloride has led to the study of transcription of C- and A-containing deoxypolynucleotides modified by a mutagenic metabolite, chloroacetaldehyde (CAA)[1] (Barbin et al., 1981; Hall et al., 1981). These experiments were performed with *E. coli* DNA polymerase I, which transcribes with high fidelity (Loeb & Kunkel, 1982). Nevertheless, errors were reported.

We have shown in the preceding paper (Kušmierek & Singer, 1982) that reaction of CAA with C and A residues in polynucleotides leads to the formation of intermediates. These hydrated etheno derivatives, ϵC·H_2O and ϵA·H_2O, are under physiological conditions converted relatively slowly to the stable etheno derivatives, ϵC and ϵA. Therefore, in contrast to the results where ϵC- and ϵA-containing templates were prepared by enzymatic copolymerization of appropriate nu-

† From the Department of Molecular Biology and Virus Laboratory, University of California, Berkeley, California 94720. *Received March 29, 1982.* This work was supported by Grant CA 12316 from the National Cancer Institute, National Institutes of Health (B.S.), and by the U.S.–Poland Cancer Program of the National Cancer Institute (J.T.K.).

* Address correspondence to this author at the Laboratory of Chemical Biodynamics, University of California, Berkeley, CA 94720.

‡ Permanent address: Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw 02532, Poland.

[1] Abbreviations: CAA, chloroacetaldehyde; ϵC, 3,N^4-ethenocytidine or the base residue in a polynucleotide; ϵC·H_2O, hydrated 3,N^4-ethenocytidine or the base residue in a polynucleotide; Tris, tris(hydroxymethyl)aminomethane; HPLC, high-pressure liquid chromatography; DEAE, diethylaminoethyl; EDTA, ethylenediaminetetraacetic acid.

0006-2960/82/0421-5723$01.25/0

cleotides (Spengler & Singer, 1981), the transcription results with CAA-modified polynucleotides are likely to represent data for both types of modification, etheno and hydrated etheno compounds.

The stability of $\epsilon C \cdot H_2O$ allowed us to prepare CAA-modified poly(rC) and poly(dC) templates in which hydrates represented more than 90% of the total modification. In this paper we present results on the effect on transcription of three specific cytosine derivatives: ϵC, $\epsilon C \cdot H_2O$ and m^3C. All three bases lead to misincorporations when present in both ribo- and deoxyribonucleotide templates transcribed by either a prokaryotic or eukaryotic DNA-dependent RNA polymerase, in the presence of either Mn^{2+} or Mg^{2+}.

Materials and Methods

Materials. [α-^{32}P]GTP (20–30 Ci/mmol) was purchased from ICN. Tritium-labeled ATP, GTP, CTP, and UTP (12–26 Ci/mmol) were Schwarz/Mann products. *E. coli* DNA-dependent RNA polymerase (EC 2.7.7.6) (Chamberlin & Berg, 1962; Burgess & Jendrisak, 1975) (380 units/mg of protein) was purchased from Miles, whereas DNA-dependent RNA polymerase IIB from calf thymus (Hodo & Blatti, 1977) was a gift from Bethesda Research Laboratories.

Commercial polynucleotides were from Miles and P-L Biochemicals. Poly(rC) and poly(dC) containing ϵC and $\epsilon C \cdot H_2O$ residues were obtained via reaction with chloroacetaldehyde (Kuśmierek & Singer, 1982). The hydrate represents 90% or more of total modification at the time of transcription. This estimation is based on measured half-lives of $\epsilon C \cdot H_2O$ in polymers. The complete dehydration of $\epsilon C \cdot H_2O$ residues was obtained by heating CAA-modified poly(rC) and poly(dC) in 0.01 M pH 7.25 Tris-HCl buffer at 85–90 °C for 2 h. This treatment approximately doubles the amount of deaminated C residues compared to the level present in commercial poly(rC) and poly(dC). The amount of U residues in heated polymers is on the order of 1%. The total depyrimidination of heated polymers is below 1% and, by HPLC analysis, is equally divided between ϵC and C.

Other polynucleotides used in experiments were prepared by copolymerization of appropriate ribonucleoside 5′-diphosphates or by modification of commercial poly(rC) or poly(dC). Deamination by nitrous acid yielded copolymers of C and U (Singer & Fraenkel-Conrat, 1970), methylation with dimethyl sulfate yielded copolymers of C and m^3C (Ludlum, 1971), and acetylation with acetic anhydride yielded copolymers of C and ac^4C (Michelson & Grunberg-Manago, 1964). Each of these chemically modified polymers contained only a single type of modified base as determined by HPLC analysis.

Transcription of Polynucleotides. Preliminary experiments with each of the polymerases studied, including experiments on kinetics and divalent cation optima, were performed with [^{3}H]ribonucleoside 5′-triphosphates. In these experiments, the 125-μL standard incubation mixture contained 1–6 μCi of varying single [^{3}H]NTPs (the other NTPs were unlabeled), and aliquots were taken at various times of incubation and spotted on DEAE paper disks that were put immediately into 7% Na_2HPO_4 solution. Disks including blank disks (in order to determine background radioactivity) were washed in a bath of 7% Na_2HPO_4 (about 50 disks/L) and shaken gently for 10 min at room temperature, and the washing solution was removed by decantation. After five to seven such washings, the disks were rinsed briefly twice with water and twice with absolute ethanol. After being dried, the disks were placed in scintillation vials, 2 mL of scintillation fluid [4 g of Omnifluor (New England Nuclear) in 1 L of toluene] was added, and the radioactivity was determined in a Beckman liquid scintillation counter. The radioactivity bound to the disks represents the amount of a specific nucleotide incorporated in the newly synthesized polymer.

Transcription experiments, which were used for nearest-neighbor analysis (see below), were performed on a 5-fold larger scale, and the labeled nucleoside triphosphate was [α-^{32}P]GTP. In the experiments with each polymerase studied, the 625-μL standard incubation mixture contained 0.15 absorbance unit of polymer and 0.4 mM each of ATP, CTP, UTP, and [α-^{32}P]GTP (20–50 μCi). Incubation mixtures with the *E. coli* polymerase contained, in addition to the template and four nucleoside 5′-triphosphates, 2 mM $MnSO_4$ (in some cases 4 mM $MnSO_4$) or 10 mM $MgSO_4$, 80 mM Tris-HCl, pH 7.8, 40 mM β-mercaptoethanol, 0.4 mM K_2HPO_4, and 5 units (15 μg of protein) of enzyme. The calf thymus polymerase reaction mixture contained 3 mM $MnSO_4$, 50 mM Tris-HCl, pH 7.9, 100 mM $(NH_4)_2SO_4$, 0.1 mM dithiothreitol, 0.1 mM EDTA, and 0.8 unit[2] of enzyme. All assays were incubated for 60 min at 37 °C.

After incubation, 75-μL aliquots were spotted on DEAE paper disks and washed as in the small-scale experiments (see above) as a measure of the total radioactivity in the reaction. To the remaining 550 μL, 1.6 mg of yeast RNA was added, and immediately after a mixing, a 30% solution of trichloroacetic acid was added to a final concentration of 5% and placed in ice for a few minutes. The resulting precipitate was centrifuged at 10000 rpm for 10 min, and the pellet was washed 4 times with 0.5 mL of 5% trichloroacetic acid and finally 2 times with ethanol. All precipitations and washings were at 0 °C.

Nearest-Neighbor Analysis of Transcription Products. The pellet containing the newly synthesized ^{32}P-labeled ribopolymer was hydrolyzed with 100 μL of 0.3 M KOH, 18 h at 37 °C. The entire sample, together with 100 μg of each 2′ (3′) nucleotide as additional internal markers, was applied in a 1-cm streak to a Whatman 3 MM 46 × 57 cm paper sheet for paper electrophoresis in a Savant Model LT48A chamber (Savant Instruments, Inc., Hicksville, NY). The electrophoresis was performed in 0.5% (v/v) pyrimidine, 5% (v/v) glacial acetic acid, and 1 mM EDTA, pH 3.5, buffer. The separation was complete for all four 2′ (3′) nucleotides, usually after 5 h at 1300 V, when the visible marker, blue dye xylene cyanol FF (Eastman Organic Chem.), moved 25–26 cm from the origin toward the anode. The dried electrophoresis was washed with acetone and then with 2-propanol in order to remove UV-absorbing pyridine–acetate. The four UV-absorbing areas corresponding to four 2′ (3′) nucleotides were cut out, and ^{32}P radioactivity was determined, by either Cerenkov (without scintillation fluid) or liquid scintillation counting.

The areas on electropherograms between 2′ (3′) nucleotide UV markers did not contain radioactivity, with the exception of the area located slightly ahead of the fastest moving nucleotide, 2′(3′)-UMP. The radioactivity in this area, particularly when the total transcription was low, represented a considerable proportion (up to 5%) of the total radioactivity. The streaking from this area to 2′(3′)-UMP can cause errors in the determination of radioactivity in U^{32}pG sequences. In order to avoid these errors, the 2′(3′)-UMP areas that were cut from the electropherogram were sewed on another sheet of Whatman 3 MM paper and were chromatographed for 18 h (descending) in 2-propanol–1% aqueous $(NH_4)_2SO_4$, 2:1

[2] The definition of the unit by the supplier (Bethesda Research Laboratories) is given in picomoles. Here, we express the enzyme units in nanomoles, so 0.8 unit is equal to 800 supplier's units.

Table I: Effect of Modification of Cytosine Residues on Transcription of Poly(rC) and Poly(dC) by DNA-Dependent RNA Polymerase from *E. coli* with [α-^{32}P]GTP in the Presence of 2 mM Mn^{2+}: Nearest-Neighbor Sequence Analysis[a]

template	source[b]	nearest-neighbor sequence (cpm × 10^{-3})[c] $C^{32}pG$	$A^{32}pG$	$U^{32}pG$	$G^{32}pG$	transcription, % poly(rC) or poly(dC)[d]
ribopolynucleotides						
poly(rC)	A	0.9 (0.1)	2.6 (0.3)	2.3 (0.3)	819 (99.3)	100
poly(rC, 11% m^3rC)	B	*14.6 (4.6)*	*9.4 (3.0)*	*7.6 (2.4)*	285 (90.0)	38
poly(rC, 17% m^3rC)	C	*8.0 (6.0)*	*5.2 (3.9)*	*3.3 (2.5)*	117 (87.6)	16
deoxyribopolynucleotides						
poly(dC)	A	0.8 (0.1)	4.6 (0.5)	1.7 (0.2)	934 (99.2)	100
poly(dC,dT) (1:1)	A	2.7 (0.6)	*258 (56.1)*	0.9 (0.2)	198 (43.1)	49[e]
poly(dC)·poly(dG)	A	0.9 (0.5)	0.9 (0.5)	0.7 (0.4)	195 (98.6)	21[f]
poly(dC, 2.3% dU)	D	1.2 (0.1)	*31.0 (3.3)*	2.3 (0.2)	915 (96.4)	101
poly(dC, 16% m^3dC)	C	*15.9 (7.2)*	*9.5 (4.3)*	*8.9 (4.1)*	185 (84.4)	23
poly(dC, 8% εdC·H_2O+εdC)	E	5.1 (0.9)	*22.9 (4.2)*	*11.6 (2.1)*	505 (92.8)	58
poly(dC, 8% εdC)	E	2.5 (0.5)	*18.9 (3.8)*	*15.8 (3.2)*	464 (92.5)	53
poly(dC, 34% εdC·H_2O+εdC)	E	*1.4 (1.9)*	*6.4 (8.5)*	*4.0 (5.3)*	63.6 (84.3)	8
poly(dC, 34% εdC)	E	1.2 (1.1)	*7.9 (7.5)*	*7.1 (6.7)*	89.6 (84.7)	11

[a] All four triphosphates were present in equal molarity. Significant numbers (more than 1% of other than $G^{32}pG$ transfer) are italicized. [b] The composition of all modified polymers was determined by HPLC analysis after enzyme digestion to nucleosides (see Materials and Methods): (A) commercial polymers; (B) polymer prepared by copolymerization of 5′-CDP and 5′-m^3CDP (Kröger & Singer, 1979); (C) modified polymers obtained by reaction of poly(rC) and poly(dC) with dimethyl sulfate, according to Ludlum (1971); (D) polymer partially deaminated by reaction with nitrous acid according to Singer & Fraenkel-Conrat (1970); (E) polymers reacted with CAA (those containing εC·H_2O and εC are not additionally treated after reaction). Of the total modification the hydrate represents 90% or more. Polymers containing εC only have been heated (see Materials and Methods). [c] The numbers in parentheses are the percent of total [^{32}P]GMP incorporation. [d] Total radioactivity incorporated is expressed as the percent of incorporation directed by poly(rC) or poly(dC), respectively (see footnotes e and f). [e] This number represents only part of the transcription of the random copolymer, poly(dC,dT). The transcription of C residues in alternating C-T tracts as well as the transcription of T in T clusters cannot be determined with [α-^{32}P]GTP. [f] This number represents only the transcription of the poly(dC) strand in double-stranded poly(dC)·poly(dG).

Table II: Effect of Modification of Cytosine Residues on Transcription of Poly(rC) and Poly(dC) by DNA-Dependent RNA Polymerase from *E. coli* with [α-^{32}P]GTP in the Presence of 10 mM Mg^{2+}: Nearest-Neighbor Sequence Analysis[a]

template	source[b]	nearest-neighbor sequence (cpm × 10^{-3})[c] $C^{32}pG$	$A^{32}pG$	$U^{32}pG$	$G^{32}pG$	transcription, % poly(rC) or poly(dC)[d]
ribopolynucleotides						
poly(rC)	A	0.8 (0.1)	1.5 (0.2)	1.1 (0.2)	651 (99.5)	100
poly(rC, 11% m^3rC)	B	*6.0 (2.2)*	*4.6 (1.7)*	*7.4 (2.7)*	254 (93.4)	42
poly(rC, 17% m^3rC)	C	*3.6 (3.7)*	*2.9 (3.0)*	*6.4 (6.6)*	83.6 (86.7)	15
deoxyribopolynucleotides						
poly(dC)	A	0.9 (0.1)	4.8 (0.4)	1.5 (0.1)	1229 (99.4)	100
poly(dC,dT) (1:1)	A	2.8 (0.7)	*221 (52.9)*	0.3 (0.1)	193 (46.3)	34[e]
poly(dC)·poly(dG)	A	0.6 (0.3)	0.7 (0.3)	0.5 (0.2)	206 (99.2)	17[f]
poly(dC, 2.3% dU)	D	1.2 (0.1)	*30.6 (3.4)*	3.6 (0.4)	876 (96.1)	74
poly(dC, 16% m^3dC)	C	*11.2 (4.9)*	*4.2 (1.8)*	*11.7 (5.1)*	201 (88.2)	18
poly(dC, 8% εdC·H_2O+εdC)	E	2.5 (0.5)	*16.9 (3.6)*	*7.7 (1.7)*	437 (94.2)	38
poly(dC, 8% εdC)	E	1.4 (0.3)	*11.1 (2.1)*	*16.5 (3.1)*	499 (94.5)	43
poly(dC, 34% εdC·H_2O+εdC)	E	1.0 (0.7)	*7.3 (5.4)*	*3.5 (2.6)*	123 (91.3)	11
poly(dC, 34% εdC)	E	0.7 (0.7)	*3.8 (3.9)*	[g]	88.9 (90.3)	9

[a] For footnotes *a–f* see Table I. [g] Up was lost in analysis of sample. The remaining figures are calculated on the assumption that the proportion of radioactivity into UpG sequences in this case follows the trends observed in other experiments. Therefore, 5% of total incorporation was assumed to be UpG.

(v/v). The nonidentified radioactivity was well separated from the radioactivity comigrating with 2′(3′)-UMP marker and was located between the origin and the 2′(3′)-UMP marker, i.e., in the area where nucleotide diphosphates and triphosphates move in this system. Only the radioactivity that comigrated with the 2′(3′)-UMP marker during rechromotography was taken to represent the radioactivity in $U^{32}pG$ sequences in each case.

Results

Template Activity of Poly(rC) and Poly(dC) Polymers. The kinetic experiments with the polymerases tested showed that for each polymerase the transcription of poly(rC) or poly(dC) reaches a plateau after 30–90 min of incubation. The incubation time of 60 min was chosen for all further experiments (Tables I–IV).

The transcription of poly(dC) and poly(rC) by *E. coli* DNA-dependent RNA polymerase, studied in the presence of various Mn^{2+} concentrations, has a sharp maximum at 2 mM Mn^{2+}, and both templates are transcribed with equal efficiency at this optimum. The Mg^{2+} optima are broad, 4–25 mM for poly(dC) and 5–15 mM for poly(rC). Poly(rC) is transcribed with about 50% efficiency of transcription of poly(dC) in the optimal range of Mg^{2+} concentration. It is worth noting that both poly(rA) and poly(dA) are efficiently transcribed in the presence of Mn^{2+} (sharp optimum at 2 mM Mn^{2+}); however, only poly(dA) can be transcribed in the presence of Mg^{2+} (optimum 10–20 mM Mg^{2+}). Poly(rA) is not active as a template with the *E. coli* DNA-dependent RNA polymerase at any Mg^{2+} concentrations studied (1–25 mM). With the exception of some experiments that were done in the presence of 4 mM Mn^{2+} (Table IV), all transcriptions of

Table III: Effect of Modification of Cytosine Residues on Transcription of Poly(dC) by DNA-Dependent Polymerase II from Calf Thymus with [α-^{32}P]GTP in the Presence of 3 mM Mn^{2+}: Nearest-Neighbor Sequence Analysis[a]

template	source[b]	nearest-neighbor sequence (cpm × 10^{-3})[c] C^{32}pG	A^{32}pG	U^{32}pG	G^{32}pG	transcription, % poly(dC)[d]
poly(dC)	A	2.2 (0.3)	5.2 (0.6)[e]	1.6 (0.2)	863 (98.9)	100
poly(dC), heated	A	2.6 (0.5)	6.8 (1.2)[e]	1.2 (0.2)	555 (98.1)	65
poly(dC, 2.3% dU)	D	2.1 (0.2)	*22.5 (2.7)*	1.3 (0.2)	816 (96.9)	97
poly(dC,dT) (1:1)	A	2.1 (0.6)	*163 (48.8)*	0.4 (0.1)	169 (50.5)	38[f]
poly(dC,dA) (1:1)	A	0.7 (1.4)[e]	1.0 (2.0)[e]	*24.9 (49.6)*	23.6 (47.0)	6[f]
poly(dC)·poly(dG)	A	1.1 (0.1)	1.8 (0.2)	0.8 (0.1)	777 (99.6)	90[g]
poly(dC, 8% ϵdC·H_2O+ϵdC)	E	1.0 (0.6)	*3.4 (2.0)*	0.7 (0.4)	163 (97.0)	19
poly(dC, 8% ϵdC)	E	1.0 (0.8)	*4.0 (3.0)*	*1.3 (1.0)*	124 (95.2)	15
poly(dC, 34% ϵdC·H_2O+ϵdC)	E	*0.8 (1.3)*	*2.2 (3.7)*	0.5 (0.8)	55.9 (94.2)	7
poly(dC, 34% ϵdC)	E	*0.6 (2.7)*	*1.7 (7.6)*	*0.7 (3.1)*	19.3 (86.6)	3
poly(dC, 16% m^{3}dC)	C	*3.1 (3.6)*	*1.7 (1.9)*	*1.2 (1.4)*	81.8 (93.1)	10
poly(dC, 8% ac^{4}dC)	F	1.3 (0.2)	3.5 (0.6)	0.8 (0.1)	624 (99.1)	72

[a] Assay conditions, including the use of 3 mM Mn^{2+}, were according to Hodo & Blatti (1977), except that all four NTPs were present at a concentration of 0.4 mM each. Significant numbers (more than 1% of other than G^{32}pG transfer) are italicized. [b] (A–E) see footnote b in Table I: (F) polymer partially acetylated according to Michelson & Grunberg-Manago (1964). [c] The numbers in parentheses are the percent of total [^{32}P]GMP incorporation. [d] Total radioactivity incorporated is expressed as the percent of incorporation directed by poly(dC) (see footnotes *f* and *g*). [e] These relatively high numbers most likely represent specific incorporations resulting from deamination of C and A residues present in commercial poly(dC) and poly(dC,dA). Note that the incorporation of A is doubled as the result of heating poly(dC) for 2 h at 85 °C. [f] These numbers represent only part of the transcription of the random single-stranded copolymers poly(dC,dT) and poly(dC,dA) (see footnote *e* in Table I). [g] See footnote *f* in Table I.

Table IV: Ambiguities Resulting from Different Modification of C in Poly(rC) and Poly(dC) during Transcription by DNA-Dependent RNA Polymerase from *E. coli* in the Presence of Mn^{2+} and Mg^{2+} [a]

template	source[b]	^{32}P radioactivity (%), nearest-neighbor sequence[c] 2 mM Mn^{2+} CpG	ApG	UpG	total	10 mM Mg^{2+} CpG	ApG	UpG	total
ribopolynucleotides									
poly(rC, 7% ϵC)[e]	B	0.3	2.2	3.2	5.7	0.2	1.1	2.8	4.1
poly(rC, 9% ϵrC·H_2O+ϵrC)	E	0.5	3.3	1.2	5.0	0.1	1.5	0.6	2.2
poly(rC, 9% ϵrC)	E	0.2	1.6	2.5	4.3	0.1	0.8	2.3	3.2
poly(rC, 36% ϵrC·H_2O+ϵrC)[d]	E	1.2	6.5	2.9	10.6				
poly(rC, 36% ϵrC)[d]	E	0.5	4.4	9.0	13.9				
poly(rC, 17% m^{3}rC)	C	5.9	3.6	2.2	11.7	3.6	2.8	6.4	12.8
poly(rC,rA,rG,rU) (70:5:15:9)[e]	B	7.3	5.6	4.3		7.6	5.0	3.5	
deoxyribopolynucleotides									
poly(dC, 8% ϵdC·H_2O+ϵdC)	E	0.8	3.7	1.9	6.4	0.4	3.2	1.6	5.2
poly(dC, 8% ϵdC)	E	0.3	2.6	3.0	5.9	0.2	0.9	3.0	4.1
poly(dC, 34% ϵdC·H_2O+ϵdC)	E	1.8	8.0	5.1	14.9	0.6	5.0	2.5	8.1
poly(dC, 34% ϵdC)	E	0.9	6.3	6.4	13.6	0.6	2.7	*f*	
poly(dC, 16% m^{3}dC)	C	7.1	3.8	3.9	14.8	4.8	1.4	5.0	11.2

[a] Data from Tables I and II and from additional experiments. In all experiments all four triphosphates were present in equal amounts. [b] Poly(rC, 7% ϵC) (Spengler & Singer, 1981) and poly(rC,rA,rG,rU) (Kröger & Singer, 1979) were prepared with polynucleotide phosphorylase. For other polymers see footnote *b* in Table I. [c] The incorporation directed by the poly(rC) or poly(dC) carriers is subtracted. The relatively high incorporation of radioactivity into ApG sequences directed by poly(rC) (0.2–0.3% for both cations) and poly(dC) (0.5–0.6% for both cations) can be attributed to the presence of U residues in commercial polymers. Heating of poly(rC) and poly(dC), polymers used as the controls for dehydrated CAA-treated polymers, increases incorporation into ApG sequences to 0.4–0.5% and 1.2%, respectively. Incorporation into CpG (0.1–0.2% for both cations) and into UpG (0.1–0.3% for both cations) is approximately the same for unheated and heated poly(rC) and poly(dC). [d] Polymers transcribed in the presence of 4 mM Mn^{2+} (see text). [e] The same polymers were also transcribed in the presence of 4 mM Mn^{2+}. The polymer containing ϵC exhibits significantly lower misincorporations at the higher Mn^{2+} concentration, whereas the ribopolymer [poly(C,A,G,U)] directs approximately the same incorporation into GpG, ApG, and UpG sequences at both levels of Mn^{2+} (see footnote *b* for references). [f] Sample lost (see footnote *g* in Table II).

modified poly(rC) and poly(dC) templates by the *E. coli* polymerase were performed in the presence of 2 mM Mn^{2+} or 10 mM Mg^{2+} (Tables I, II, and IV).

Because of a limited amount of enzyme, no comparative studies of divalent metal ion requirements by calf thymus DNA-dependent RNA polymerase IIB were undertaken. All experiments were performed in the presence of 3 mM Mn^{2+}, which is the concentration recommended by the supplier for assay with denatured DNA as the template. Poly(rC) was almost inactive as a template under these assay conditions. Thus, only modified poly(dC) templates were studied.

When poly(rC) or poly(dC) carriers contain other unmodified bases or modified C residues, the efficiency of transcription by *E. coli* and calf thymus DNA-dependent RNA polymerases is decreased to varying extents (Tables I–III).

Misincorporations Resulting from Modified C Residues in Poly(rC) and Poly(dC). Misincorporations resulting from ϵC, ϵC·H_2O, and m^{3}C were studied. The most extensive study, including transcription of ribo and deoxyribo templates in the presence of Mn^{2+} and Mg^{2+}, was performed with DNA-dependent RNA polymerase from *E. coli*. The results are presented in Tables I and II and in abbreviated form in Table IV. With DNA-dependent RNA polymerase II from calf thymus only deoxypolynucleotides in the presence of Mn^{2+} were transcribed, and the results are presented in Table III.

Transcription of poly(rC,ϵrC) templates by *E. coli* DNA-

dependent RNA polymerase in the presence of Mn^{2+} leads to misincorporation of U > A ≫ C (Table IV). Templates prepared by copolymerization of CDP and ϵCDP give the same results as templates prepared by reaction of CAA with poly(rC) and subsequent dehydration of ϵC·H_2O to ϵC residues. There is also no qualitative difference between misincorporations obtained in the presence of 2 mM Mn^{2+} or 4 mM Mn^{2+} (Spengler & Singer, 1981). ϵC present in poly(dC) templates causes the same misincorporations (U > A ≫ C) as ϵC in poly(rC).

As the proportion of ϵC in poly(rC) or poly(dC) was increased, the misincorporation of U, A, and C also increased. With low levels of ϵC in templates (7–9%), misincorporation of C was borderline. However, transcription of templates containing 34–36% ϵC clearly indicated that ϵC acts like G to a small extent (Table IV). This ability of ϵC to act like G was independently confirmed in "forcing" experiments, where the transcription of poly(rC, ϵrC) was performed in the presence of only two NTPs, [α-^{32}P]GTP and CTP (Spengler & Singer, 1981). The comparison of transcription of ϵC-containing templates by *E. coli* DNA-dependent RNA polymerase in the presence of Mn^{2+} and Mg^{2+} shows that in the presence of Mn^{2+} more misincorporation occurs, but the pattern of misincorporation is the same (Table IV).

Transcriptional ambiguities resulting from ϵC·H_2O are significantly different from those resulting from ϵC. ϵC·H_2O causes more misincorporation of A than of U. Misincorporation of C is lower than misincorporation of U or A but is higher than misincorporation of C directed by ϵC. The pattern of misincorporation resulting from ϵC·H_2O (A > U > C) is not affected by the type of template (ribo or deoxyribo) or by the presence of Mg^{2+} instead of Mn^{2+} in transcription (Table IV).

Transcription of CAA-modified poly(dC) by DNA-dependent RNA polymerase II from calf thymus in the presence of Mn^{2+} results in the misincorporation of A, U, and G. A similar pattern of misincorporation, A > U > C, was obtained for both derivatives, ϵC and ϵC·H_2O. However, ϵC leads to more errors than ϵC·H_2O in templates modified to the same level (Table III).

Templates containing m^3C were included in the study for comparison since much data exist regarding the effect of m^3C or e^3C on fidelity (Ludlum & Wilhelm, 1968; Ludlum, 1970, 1971; Singer & Fraenkel-Conrat, 1970; Fraenkel-Conrat & Singer, 1971; Kröger & Singer, 1979). We find that m^3C causes misincorporations during transcription of poly(rC,m^3rC) and poly(dC,m^3dC) by *E. coli* DNA-dependent RNA polymerase in the presence of either Mn^{2+} and Mg^{2+} ions (Tables I, II, and IV). Although the extent of misincorporation relative to the amount of m^3C is similar, the pattern of misincorporation differs with the divalent ion. In the presence of 2 mM Mn^{2+}, m^3C acts like G > U > A. However, in the presence of 10 mM Mg^{2+}, m^3C acts like A > G > U. When poly(dC,m^3C) is transcribed by calf thymus DNA-dependent RNA polymerase in the presence of 3 mM Mn^{2+}, m^3C acts like G > U > A (Table III). In general, m^3C causes random misincorporation of A, U, and G, with the pattern being biased by the divalent metal. It is worth noting that in transcription of poly(rC,rA,rG,rU) (70:5:15:9) by *E. coli* DNA-dependent RNA polymerase, neither the extent nor the pattern of incorporation is biased by the divalent ion conditions (Table IV; Singer & Spengler, 1981).

Table V: Transcriptional Ambiguities Resulting from Three Different Modifications of the N-3 of C: Comparison of DNA-Dependent RNA Polymerases from *E. coli* and Calf Thymus[a]

modified C residue in template	base simulation when transcription is by enzyme from: *E. coli*[b] G	U	A	calf thymus[c] G	U	A
ϵC·H_2O	+	+++	++	±	++	+
ϵC	±	++	+++	+	+++	++
m^3C	+++	++	++	+++	++	+++

[a] This summary is based on quantitative data in Tables I–IV and in the text. [b] *E. coli* DNA-dependent RNA polymerase data are pooled from experiments where both ribo and deoxyribo templates were transcribed in the presence of either Mn^{2+} or Mg^{2+}. Under all conditions used the type and extent of misincorporation were similar for a derivative, except for m^3C. In this case, the simulation of A was greater with Mg^{2+} than with Mn^{2+}. [c] Calf thymus DNA-dependent RNA polymerase data are taken from Table III where deoxyribo templates were transcribed in the presence of Mn^{2+}.

Discussion

There are four main questions addressed in this investigation: (1) Do ϵC and ϵC·H_2O cause misincorporation, and if so, do they behave alike? (2) Does the divalent cation and specific polymerase used affect the fidelity in transcription? (3) Do the same modified bases have different effects on transcription when present in ribo or deoxyribo polymers? (4) Does the loss of hydrogen-bonding ability of three different derivatives of C differing in configuration lead to the same effects in transcription?

We have found that both ϵC and ϵC·H_2O cause misincorporation in all systems used. However, the pattern for each derivative differs significantly, A > U ≫ C for the hydrate and U > A ≫ C for ϵC (Tables I, II, and IV). Previously there were two reports on transcriptional errors using poly(dA), poly(dC), and the alternating polymers poly(dC-dG) and poly(dA-dT). All were reacted with CAA but not analyzed for the presences of hydrates. These were transcribed by DNA polymerase I from *E. coli*, and misincorporation was measured by the ratio of incorrect to correct nucleotides (Hall et al., 1981; Barbin et al., 1981) and once by nearest-neighbor analysis (Hall et al., 1981). From these two reports it was concluded that ϵC simulates A and T. Despite the fact that an entirely different type of enzyme was used and that it was very likely that the templates contained both derivatives (ϵC and ϵC·H_2O), these results, qualitatively, are not in conflict with our results, since both derivatives miscode for A and T.[3]

The observed pattern of misincorporations for ϵC and ϵC·H_2O is the same in presence of Mg^{2+} and Mn^{2+}, although the extent of misincorporation is less in the presence of Mg^{2+}. The ambiguities resulting from ϵC and ϵC·H_2O transcribed by DNA-dependent RNA polymerase from calf thymus are somewhat different from those obtained in transcription by DNA-dependent RNA polymerase from *E. coli*. The polymerase from calf thymus gives the same pattern of misincorporations for both derivatives (A > U > C), but ϵC causes, with this enzyme, more misincorporation of each nucleotide than does the hydrate (Tables III and V).

When RNA polymerases were first described, it was shown that Mn^{2+} was a better cation than Mg^{2+} with certain ribopolynucleotide templates (Karstadt & Krakow, 1970). As a consequence, Mn^{2+} has been generally used in such transcription studies [see Singer (1982) and Singer & Kuśmierek

[3] We are reporting in a paper in press (Singer et al., 1982) that poly(dC) containing hydrated ϵC when copied by DNA polymerase I from *E. coli* does not cause misincorporations, while the presence of ϵC leads to misincorporation of T only.

(1982) for reviews]. We find that poly(rC) is also transcribed quite efficiently in the presence of Mg^{2+}, although the optimal ionic concentrations differ. However, this is not a general phenomenon since poly(rA) is completely inactive as a template in the presence of Mg^{2+}, as also noted by Karstadt & Krakow (1970) using another bacterial polymerase. Nevertheless, in the case of templates with a high content of C we were able to directly compare transcriptional errors induced in both ribo- and deoxyribopolynucleotides using either Mg^{2+} or Mn^{2+}. There is no significant difference in patterns of ambiguity when ϵC or $\epsilon C{\cdot}H_2O$ in ribo or deoxyribo templates is transcribed in the presence of the same divalent cation (Tables I, II, and IV). There are reports stressing the error proneness of divalent cations other than Mg^{2+} in transcription (Sirover et al., 1979; Niyogi & Feldman, 1981). However, these authors were studying incorporation, directed by unmodified bases, of one incorrect base among thousands of correct bases. Our experiments do indicate that Mn^{2+} slightly increases the frequency of misincorporations when modified nucleosides are present. However, all misincorporations found with Mn^{2+} are also found with Mg^{2+}. The use of Mn^{2+} does not lead to misincorporation when the modified base can form proper hydrogen bonds (Singer & Spengler, 1981). This is now confirmed with a deoxypolynucleotide containing ac^4dC transcribed by the calf thymus polymerase (Table III).

When we included in this study m^3C, which is also blocked on an essential Watson–Crick site, the ambiguity observed differed from that of ϵC or $\epsilon C{\cdot}H_2O$. It was previously shown that in ribopolynucleotides m^3C, m^3U, and m^1A cause random misincorporation of nucleotides but that the "randomness" was different for each derivative (Kröger & Singer, 1979). The same misincorporations resulting from m^3rC are now confirmed with deoxypolynucleotides containing m^3dC. m^3C, like the etheno derivatives, is substituted on a site that is essential for hydrogen-bond formation and is therefore incapable of forming specific Watson–Crick hydrogen-bond base pairs. The lack of information for the polymerase, however, is evidently not complete because misincorporations do result from each of the three modifications and differ (Table V). Misincorporation can be affected to some extent by the type of transcribing enzyme and divalent cation. Although all three derivatives have a blocked N-3 position, their structures are quite different. ϵC is rigid and planar while the substituent $\epsilon C{\cdot}H_2O$ is likely to be flexible. The methyl group on m^3C is smaller than the etheno ring and may also have some flexibility. We conclude that errors made by polymerases during transcription of modified bases lacking the ability to form hydrogen bonds must be at least in part attributed to the enzyme recognizing some specific facet of the modified base.

Acknowledgments

We are most appreciative of Dr. H. Fraenkel-Conrat's interest and criticisms. We are indebted to M. K. Gleason for performing many of the transcription experiments.

References

Barbin, A., Bartsch, H., Leconte, P., & Radman, M. (1981) *Nucleic Acids Res. 9*, 375.

Burgess, R. R., & Jendrisak, J. J. (1975) *Biochemistry 14*, 4634.

Chamberlin, M., & Berg, P. (1962) *Proc. Natl. Acad. Sci. U.S.A. 48*, 81.

Drake, J. W. (1969) *Nature (London) 221*, 1132.

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

Hall, J. A., Saffhill, R., Green, T., & Hathway, D. E. (1981) *Carcinogenesis (N.Y.) 2*, 141.

Hodo, H. G., & Blatti, S. P. (1977) *Biochemistry 16*, 2334.

Karstadt, M., & Krakow, J. S. (1970) *J. Biol. Chem. 245*, 746.

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

Kuśmierek, J. T., & Singer, B. (1982) *Biochemistry* (preceding paper in this issue).

Loeb, L. A., & Kunkel, T. A. (1982) *Annu. Rev. Biochem. 51*, 429.

Ludlum, D. B. (1970) *Biochim. Biophys. Acta 213*, 142.

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

Ludlum, D. B., & Wilhelm, R. C. (1968) *J. Biol. Chem. 243*, 2750.

Michelson, A. M., & Grunberg-Manago, M. (1964) *Biochim. Biophys. Acta 91*, 92.

Mildvan, A. S., & Loeb, L. A. (1979) *CRC Crit. Rev. Biochem. 6*, 219.

Niyogi, S. K., & Feldman, R. P. (1981) *Nucleic Acids Res. 9*, 2615.

Singer, B. (1982) in *Molecular and Cellular Mechanisms of Mutagenesis* (Lemontt, J. F., & Generoso, W. M., Eds.) pp 1–42, Plenum Press, New York.

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

Singer, B., & Spengler, S. (1981) *Biochemistry 20*, 1127.

Singer, B., & Kuśmierek, J. T. (1982) *Annu. Rev. Biochem. 51*, 655.

Singer, B., et al. (1982) *Proc. Natl. Acad. Sci. U.S.A.* (in press).

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

Spengler, S., & Singer, B. (1981) *Nucleic Acids Res. 9*, 365.