Nucleotide sequences of the 5' termini of φX174 mRNAs synthesized in vitro

Lloyd H. Smith, Karel Grohmann, and R. L. Sinsheimer

Division of Biology, California Institute of Technology, Pasadena, California 91109, USA

Received 20 August 1974

ABSTRACT

When using φX174 RFI DNA as a template, in vitro, E. coli RNA polymerase synthesizes four major purine triphosphate-containing 5' end sequences. RNase A digests of $\alpha^{32}P$ labeled RNA were further digested with spleen exonuclease to remove the bulk of the oligonucleotides with 5' hydroxyls and then chromatographed on DEAE cellulose to resolve the remaining 5' terminal oligonucleotides. By application of standard separation and sequence techniques, the major 5' end sequences were shown to be: pppApUp(Cp), pppApApApUp(Cp), pppApApApApUp(Cp), and pppGpApUp(Gp).

INTRODUCTION

Bacteriophage φX174 stimulates the synthesis of a number of mRNAs of discrete sizes during infection of E. coli[1,2]. Hayashi et al.[3] have presented evidence that mRNAs with similar molecular weights to some of the in vivo φX mRNAs are synthesized in vitro in the presence of rho factor[4]. Though no evidence has appeared to show what part E. coli RNase III[5] plays in the production of any of the φX mRNAs found in vivo, RNase III does appear to be responsible for the generation of most, if not all, of the T7 early mRNA species observed in vivo from longer precursors arising from three closely spaced initiations[7]. φX mRNA as well seems to be initiated from more than one promotor since (1) Hayashi and Hayashi[8] have reported that both $\gamma^{32}P$ ATP and $\gamma^{32}P$ GTP are incorporated into φX mRNA produced in vitro and (2) Chen et al.[9] have shown that three of the discrete fragments of φX RF DNA produced by $Hind_{II+III}$ endonuclease cleavage contain E. coli RNA polymerase binding sites. The present report confirms the existence of more than one 5' end nucleotide sequence in in vitro φX mRNA, and describes a procedure for the isolation and characterization of the 5' oligonucleotides derived from $\alpha^{32}P$ ribonucleoside triphosphate-labeled in vitro φX mRNA.

MATERIALS AND METHODS

(1) Preparation of ϕX174 RFI: RFI is isolated by the procedure of Komano and Sinsheimer[10] through the phenol extraction and first isopropanol precipitation steps. Subsequently, the precipitate is redissolved, then treated with RNase A (2 μg/OD_{260} unit), 37°C, 2 hr, extracted with phenol five times, and precipitated with isopropanol two times. The resulting precipitate is redissolved in 0.05 M Tris-Cl pH 7.9, 0.005 M EDTA and further purified by gel filtration on Biogel A 1.5 m (Biorad) as described by Komano and Sinsheimer[9] or on a porous glass bead column (Corning CPG 2000 Biorad) in Tris-EDTA buffer. The DNA excluded from these columns is already 95-98% ϕX RFI + II as judged by analytical sedimentation through 3 M CsCl + 0.1 N KOH in a Beckman Model E ultracentrifuge. RFI is purified to homogeneity by CsCl equilibrium banding in the presence of 500 μg/ml propidium diiodide[11]. RFI is separated from propidium diiodide by passage through Dowex AG50W-X2 (Biorad) in 0.5 M Tris-Cl pH 7.9, 0.005 M EDTA, followed by dialysis versus 0.05 M Tris-Cl pH 7.9, 0.005 M EDTA. The resulting RFI preparations sometimes contain traces (0-1%) of RFII.

(2) E. coli K12 RNA polymerase: RNA polymerase is purified according to the method of Burgess and Travers[12]. The Agarose column fraction is used in most of the experiments described here, although identical results are obtained with homogeneous RNA polymerase purified by glycerol gradient sedimentation.

(3) ^{32}P ribonucleoside triphosphates: $\gamma^{32}P$ ATP and $\gamma^{32}P$ GTP are obtained from ICN at specific activities of $\geq$ 75c/mM. $\alpha^{32}P$ ribonucleoside triphosphates are obtained from New England Nuclear Corporation at specific activities of $\geq$ 100c/mM.

(4) Preparation of ^{32}P-labeled, in vitro, ϕX mRNA: γ or $\alpha^{32}P$ ribonucleoside triphosphates are evaporated to dryness and redissolved in the standard reaction mixture of 200 μl which contains 0.05 M Tris-Cl pH 7.9, 0.2 M KCl, 0.01 M $MgCl_2$, 0.1 mM EDTA, 0.1 mM dithiothreitol, 100 μg ϕX RFI and 500 μg RNA polymerase (holoenzyme). ^{32}P-labeled triphosphates are present at 0.1 to 0.2 mM, unlabeled triphosphates at 0.5 mM. $\alpha^{32}P$ labeling reactions are allowed to proceed for one minute at 37°C after the addition of RNA polymerase, then terminated by immersion in a boiling water bath for one minute. After equilibration to 25°C, DNase I (RNase free, Worthington) is added to 50 μg/ml and incubated at 37°C for 30 minutes. SDS is then added to 0.2% and the mixture extracted with phenol. $\gamma^{32}P$ labeling reactions are terminated by the addition of DNase I after 30 minutes at 37°C and

allowed to incubate an additional 30 minutes at 37°C. SDS and phenol are added as above. After phenol extraction, the aqueous layers are applied immediately to a 0.7 x 30 cm column of Sephadex G50 (fine) in 0.05 M Tris-Cl pH 7.9, 0.005 M EDTA. Excluded RNA is pooled, its $OD_{260m\mu}$ measured, and precipitated (0.1 volume 3 M sodium acetate pH 5, 2 volumes 100% ethanol are added followed by incubation at -20°C for 4-8 hr before sedimentation at 20,000 rpm in Beckman SW 50.1 rotor). The RNA is redissolved in water. The size of the RNA product from the $\alpha^{32}P$ labeling reactions is heterogeneous (data not shown), ranging from 500-2,000 nucleotides long as judged by sedimentation through dimethylsulfoxide[13].

(5) Digestion of the RNA with enzymes: Aliquots of the ^{32}P labeled RNAs are digested in 0.01 M Tris-Cl pH 7.9, 0.01 M EDTA with pancreatic RNase A (Sigma type III-A) using an enzyme to substrate ratio of 1:15 (μg/μg); incubation at 37°C for 40 minutes gives complete digestion.

For isolation of 5' oligonucleotides from $\alpha^{32}P$ labeled RNA, aliquots of RNA are first digested in a small volume (1-2 μl) with RNase A as above, then further digested by addition of 50-100 μl of bovine spleen exonuclease (purified as per Hilmoe[14], obtained from Worthington as spleen phosphodiesterase) in 0.025 M potassium phosphate pH 6.3, 0.0025 M EDTA and incubation for 1 hr at 37°C. 3 to 5 units spleen exonuclease is added per μg RNA. SDS is then added to 1% followed by extraction with phenol. The aqueous layer is diluted with water to 2 ml and applied to a 0.5 x 1 cm column of Whatman DE52. The column is washed with 0.1 M triethylammonium bicarbonate (TEAB) pH 10 to remove nucleoside monophosphates; the remaining oligonucleotides are eluted with 2 M TEAB pH 10. The TEAB is removed by repeated evaporation with equal volumes of methanol and finally once with 2 ml water. The nucleotides are redissolved in 1-2 μl 0.005 M EDTA and directly applied to cellulose acetate strips for ionophoresis as described by Barrell[15]. Homochromatography (using the 30 minute homomixture) is performed on DEAE cellulose-coated polyethylene sheets (Polygram cel 300 DEAE, Brinkman Inst.). Spots from homochromatography are cut out and counted for Cerenkoff radioactivity in 12 ml 100% ethanol, then scraped off into a suction device[16], washed with additional ethanol, and eluted with 0.25 M NH_4OH.

The eluted oligonucleotides are evaporated to dryness, redissolved in 10 μl 0.02 M EDTA, and then digested to mononucleotides by adding 1 μl of an RNase mixture (100 units/ml T_2 RNase [Sigma], 10 units/ml μ_2 RNase [Calbiochem], 2.5 mg/ml RNase A [Sigma], and 5,000 μ/ml T_1 RNase [Calbiochem] in 0.5 M ammonium acetate pH 4.5). The oligonucleotides are digested for 5 hr

at 37°C. The enzymatic degradation of these oligonucleotides is more rapid than alkaline hydrolysis, and seems to result in less destruction of the purine tetraphosphates.

The digests are examined for nucleotide composition by streaking half on Whatman DE81 paper and half on Whatman 540 paper. Ionophoresis is performed in the pH 3.5 (no urea) buffer of Barrell[15]. The DEAE ionophoresis allows identification of pppAp and pppGp; the 540 system resolves the four nucleoside 3' monophosphates.

RESULTS

(1) Characterization of γ^{32}P ATP and γ^{32}P GTP 5' oligonucleotides. *In vitro* ϕX mRNAs labeled at the 5' end with either γ^{32}P ATP or γ^{32}P GTP are digested and separated as described above. RNase A digests of γ^{32}P ATP labeled RNA (Figure 1a) show three major 5' end oligonucleotides comprising 95% of all radioactivity on the chromatogram. Figure 1b is the pattern obtained from RNase A digests of γ^{32}P GTP labeled RNA; the major spot accounts for ∿75% of the radioactivity, while the remainder is distributed among several minor oligonucleotides. During the course of this work, we have noticed that these minor γ^{32}P GTP-labeled oligonucleotides, as well as a few minor γ^{32}P ATP-labeled oligonucleotides are present in much higher yields (sometimes amounting to 30-50% of all radioactivity) when the RNA is synthesized by RNA polymerase which has been stored (-20°C 50% glycerol storage buffer) for longer than three or four months. In the present study, freshly prepared RNA polymerase (< two months old) has been used exclusively. Consequently, the minor spots have appeared in very low yields and no attempt was made to determine their sequence.

The positions of the RNase A γ^{32}P ATP and γ^{32}P GTP-labeled oligonucleotides, relative to each other, can be seen from Figure 1c. The major γ^{32}P GTP-labeled spot migrates in the homochromatography dimension as if it were one nucleotide larger than the uppermost γ^{32}P ATP-labeled spot. If the γ^{32}P ATP-labeled RNA is digested with a mixture of RNase A and T_1, three spots of identical mobilities to the spots in Figure 1a are seen (data not shown), indicating the absence of guanosine residues between the 5' pppAp and the pyrimidine at the 3' end.

(2) Digestion with spleen exonuclease. The RNase A digests of α^{32}P-labeled RNAs contain many nucleotides which obscure the positions of the 5' oligonucleotides on the homochromatograms (data not shown). Bovine spleen exonuclease[13] is employed to digest to mononucleotides any oligonucleotides

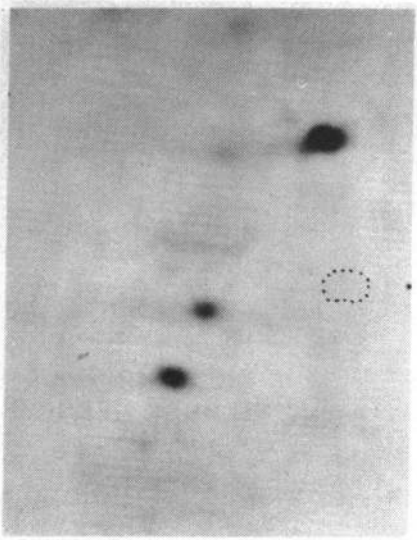
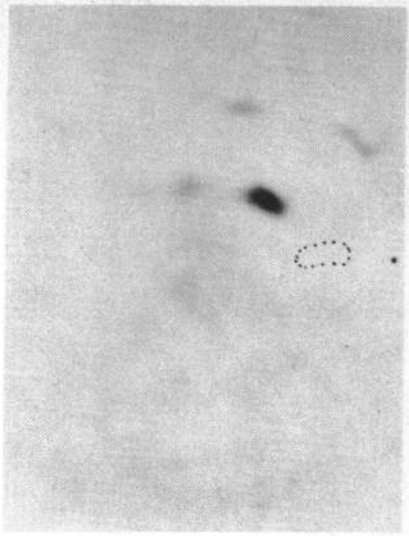
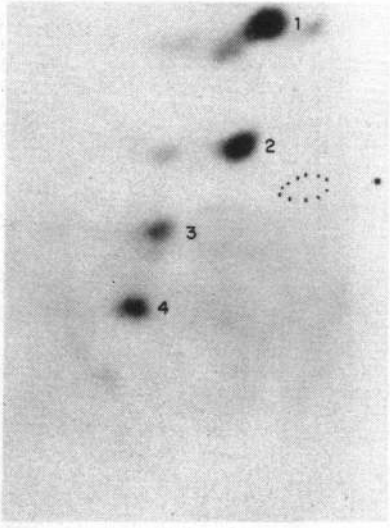


FIGURE 1 (a)



(b)



(c)

Figure 1. Autoradiograms. RNase A digests of γ^{32}P-labeled in vitro ϕX mRNA. First dimension ionophoresis is from left to right; homochromatography is from bottom to top. Dotted circle marks position of yellow marker dye. (a) γ^{32}P ATP-labeled RNA, (b) γ^{32}P GTP-labeled RNA, (c) a mixture of γ^{32}P ATP- and γ^{32}P GTP-labeled RNAs.

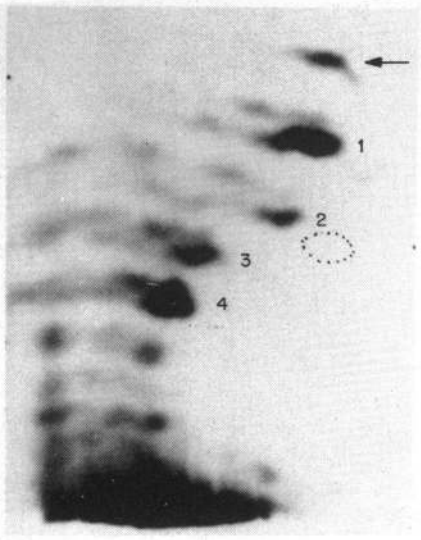
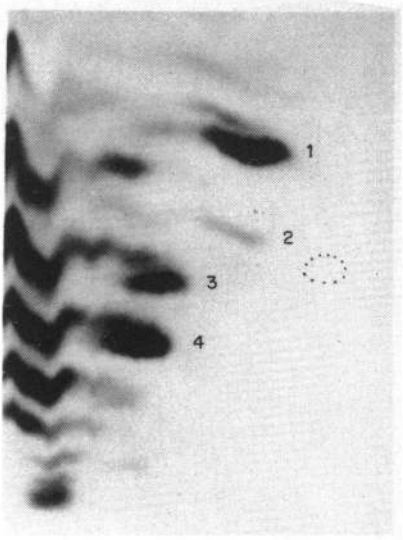
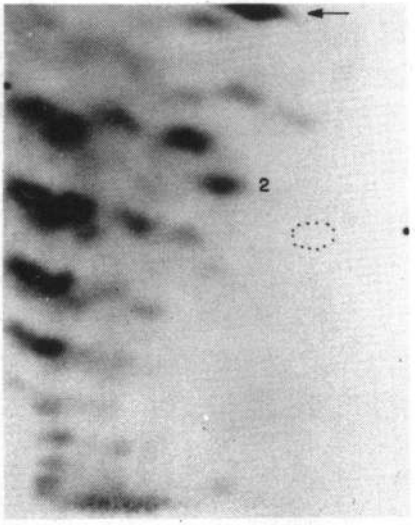


FIGURE 2 (a)

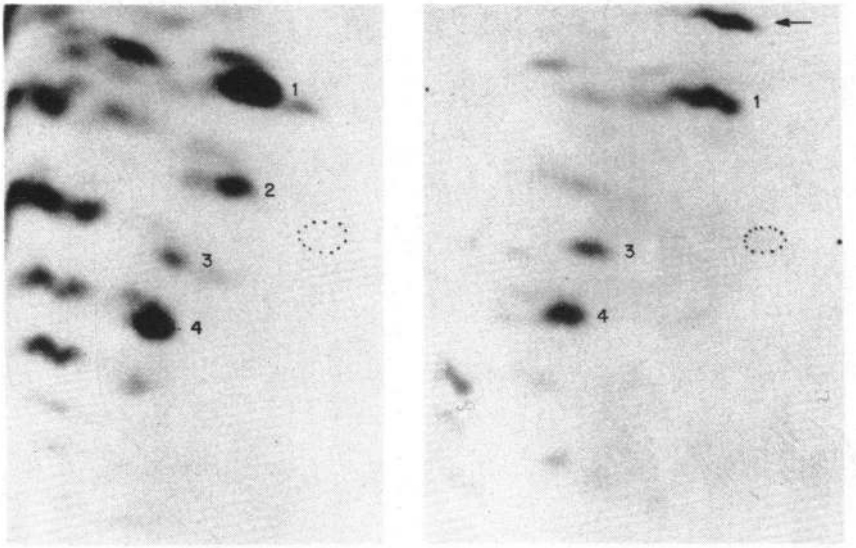


(b)



(c)

Figure 2. Autoradiograms. Spleen exonuclease treated RNase A digests of α^{32}P-labeled in vitro ϕX mRNA. (a) A mixture of in vitro ϕX mRNAs labeled independently with each α^{32}P ribonucleoside triphosphate, (b) α^{32}P ATP-labeled RNA, (c) α^{32}P GTP-labeled RNA.



(FIGURE 2) (d) (e)

Figure 2. (continued) (d) α^{32}P UTP-labeled RNA, (e) α^{32}P CTP labeled RNA.

containing a 5' hydroxyl; 5' phosphorylated oligonucleotides are resistant to such digestion[17].

The commercially available spleen phosphodiesterase (spleen exonuclease) is, however, contaminated with variable amounts of spleen acid ribonuclease and phosphatase[18]. Batches of spleen exonuclease are assayed for phosphatase contamination by digesting aliquots of RNase A digested α^{32}P-labeled RNAs with varying amounts of exonuclease. The digests are separated by one dimensional ionophoresis on DEAE paper and autoradiographed. The presence of phosphatase is indicated by the appearance of inorganic $^{32}PO_4$. Ribonuclease activity is assayed by digesting aliquots of γ^{32}P-labeled RNA with spleen exonuclease and separating the digestion products as above. The presence of RNase is indicated by the appearance of pppAp. If the contaminating ribonuclease and/or phosphatase activities are unacceptably high, the exonuclease activity can be purified by Sephadex G75 filtration (removes RNase) and Sephadex C50 chromatography (removes phosphatase) by the methods of Bernardi and Bernardi[18]. The purified preparations are free of detectable RNase and phosphatase activities, but some loss in exonuclease activity is observed. Results similar to those reported here are obtained with highly purified preparations, although the present data were derived using an uncontaminated commercial batch.

After digestion with spleen exonuclease, followed by SDS-phenol extraction, the DEAE step removes >95% of the total radioactivity and in addition removes salt which would cause extensive streaking of the nucleotides during ionophoresis.

(3) Isolation and sequence of RNase A α^{32}P-labeled 5' purine triphosphate-containing oligonucleotides. Figures 2a-e are autoradiograms of the spleen exonuclease treated RNase A α^{32}P-labeled oligonucleotides after homochromatography. From 2a to e these correspond to RNA labeled with (a) all four α^{32}P ribonucleoside triphosphates, combined after synthesis, (b) α^{32}P ATP, (c) α^{32}P GTP, (d) α^{32}P UTP, (e) α^{32}P CTP. Comparison of these figures to Figures 1a-c enables identification of the three major 5' pppA oligonucleotides and the major 5' pppG oligonucleotide. These four spots have been termed 1-4 from top to bottom. α^{32} P CTP does not label spot 2 and α^{32}P GTP does not label spots 1, 3, or 4. α^{32}P CTP and α^{32}P GTP label a spot (marked by an arrow) which seems to contain a single nucleotide (it is not cleaved by T_2 or alkaline digestion). This spot is neither pppAp or pppGp, nor does it seem to be di- or triphosphorylated adenosine or quanosine.. Its identity is being examined and will be discussed

elsewhere.

Other oligonucleotide spots can be seen in Figures 2a-e which have no relation to the mobilities of the major $\gamma^{32}P$ labeled spots of Figures 1a-c. Nucleotide composition analysis of the spots from Figure 2 (data not shown) reveals that the spots marked 1-4 are the only purine triphosphate containing nucleotides present; these have closely matching mobilities to the $\gamma^{32}P$ labeled spots of Figure 1c and are considered to be identical. The oligonucleotides which do not contain purine-triphosphates, appear in variable amounts from digest to digest. These are, as would be expected with the use of pancreatic ribonuclease, purine-rich; some of these contain 5' di- or monophosphates which may arise from loss of 5' terminal γ or β phosphates or from internal chain damages (perhaps due to radiation damage). Nucleotides arising from chain breakage after disintegration of ^{32}P to ^{32}S may still contain 5' sulfate groups capable of inhibiting attack by spleen exonuclease. However, many of the residual oligonucleotides which do not contain purine-triphosphates yield only 3' monophosphates upon T_2 digestion. It is not known whether these fragments are present because of some feature of their sequence which confers spleen exonuclease resistance, or whether under the conditions described here digestion with spleen exonuclease may be incomplete. These fragments, however, have not interfered with the characterization of the four specific oligonucleotides which contain purine-triphosphate. The radioactivity occasionally present at the transfer origin appears in variable amounts and is of unknown composition; it was not analyzed further.

The nucleotide compositions of the RNase A $\alpha^{32}P$-labeled oligonucleotides from Figures 2a-e are given in Table 1. The nearest neighbor information gained from analysis of each oligonucleotide labeled with each of the $\alpha^{32}P$ ribonucleoside triphosphates separately, together with the known specificity of RNase A for pyrimidines, is sufficient to deduce the sequence of these oligonucleotides. The identity of the nucleotide responsible for the 3' phosphate is also deduced; though not present in the oligonucleotide, the 3' phosphate donor is shown in parenthesis. The sequences derived are consistent with the nucleotide compositions of the 5' oligonucleotides labeled with a combination of all four $\alpha^{32}P$ ribonucleoside triphosphates (combined after synthesis). Spots 3 and 4 contain different numbers of Ap residues. From the ratios of radioactivity in pppAp and Ap after DEAE ionophoresis of RNase $T_2T_1\mu_2A$ digests of spots 3 and 4, the sequences $pppAp(Ap)_2Up(Cp)$ and $pppAp(Ap)_3Up(Cp)$ are deduced.

TABLE 1. Nucleotide Compositions of RNase A 5' Oligonucleotides

Spot	Label	Composition (relative yield)	Sequence
	ATP	pppAp	
	GTP	---	
1	UTP	pppAp	pppApUp(Cp)
	CTP	Up	
	XTP	pppAp,Up	
	ATP	pppGp	
	GTP	pppGp,Up (1.1:1)	pppGpApUp(Gp)
2	UTP	Ap	
	CTP	---	
	XTP	pppGp,Ap,Up	
	ATP	pppAp,Ap (2.0:1)	$pppAp(Ap)_2Up(Cp)$
	GTP	---	
3	UTP	Ap	
	CTP	Up	
	XTP	pppAp,Ap,Up	
	ATP	pppAp,Ap (1.1:1)	
	GTP	---	
4	UTP	Ap	$pppAp(Ap)_3Up(Cp)$
	CTP	Up	
	XTP	pppAp,Ap,Up	

TABLE 2. Relative Yields of 5' Purine Triphosphate-Containing Oligonucleotides

Sequence	Average* Relative Yield
pppApUp(Cp)	11.3 ± 0.2
$pppAp(Ap)_3Up(Cp)$	3.2 ± 0.3
pppGpApUp(Gp)	1.2 ± 0.1
$pppAp(Ap)_2Up(Cp)$	1.0 ± 0.3

* Average of determinations from $\alpha^{32}P$ ATP and $\alpha^{32}P$ UTP labeled spots. Corrections made for multiple phosphates contributed by $\alpha^{32}P$ ATP in spots 3 and 4.

DISCUSSION

We have determined the sequences of the four major 5' purine triphosphate-containing oligonucleotides initiated *in vitro* by *E. coli* RNA polymerase using ϕX174 RFI as a template. The relative yield of each of these 5' oligonucleotides from 1 minute $\alpha^{32}P$ labeling reactions is shown in Table 2. pppApUp(Cp) is by far the most common sequence initiated.

RNase T_1 digestion of $\gamma^{32}P$ ATP-labeled RNA also generates three major oligonucleotides (data not shown). These appear to be related to the three major $\gamma^{32}P$ ATP-labeled RNase A digestion products but are two or more nucleotides longer. Thus, no divergence into multiple sequences is observed for the pppApUp(Cp) initiation at least to the length of four or five nucleotides. This data and hybridization data (Smith and Sinsheimer, in preparation) suggest that the pppApUp(Cp) oligonucleotide arises from initiation at a single site on the ϕX genome.

It is an obvious possibility that all three pppAp oligonucleotides are due to initiation at a common promotor; RNA polymerase might begin transcription variably within a tract of T residues. Why the polymerase would initiate RNAs with one, three, and four adenosines at the 5' end, and none with two adenosines, from the same promotor, however, cannot be explained at this time. Alternatively, Chen *et al.* observed that three of the ϕX RF cleavage products produced by the *Haemophilus* restriction enzyme $Hin_{dIII+III}$ (R2,4,6.3) contain RNA polymerase binding sites; therefore, all of the pppAp ends may not arise from a single location.

The 5' end sequences present in *in vitro* ϕX mRNA are not the same as those determined for fd[19], another small single stranded DNA containing *E. coli* bacteriophage which start mainly with pppApUpG and pppGpUp. pppGpApUp has, however, been determined as a major starting sequence for ϕ 80[20]. T7 early mRNA synthesized *in vivo*[21] and *in vitro*[6] begins with the similar sequence pppApUpCpGp. pppApUpCpApGp is a starting sequence found in *in vitro* mRNA directed by λ DNA, as are the sequences pppApUpGp, pppApCpGp, and pppGpUpUp[22].

ACKNOWLEDGMENTS

The authors acknowledge the excellent technical assistance of Charles Romeo.

This research was supported by grant number GM 13554 from the National Institutes of Health.

REFERENCES

1 Hayashi, Y. and Hayashi, M. (1970) Cold Spring Harbor Symp. Quant. Biol. 35, 171-178
2 Clements, B. and Sinsheimer, R. L. (1974) J. Virol., submitted
3 Hayashi, M., Hayashi, M. N., and Hayashi, Y. (1970) Cold Spring Harbor Symp. Quant. Biol. 35, 174-177
4 Roberts, J. W. (1969) Nature 224, 1168-1174
5 Robertson, H. D., Webster, R. E., and Zinder, N. D. (1968) J. Biol. Chem. 243, 82-91
6 Rosenberg, M., Kramer, R. A., and Steitz, J. A. (1974) J. Mol. Biol., in press
7 Dunn, J. J. and Studier, F. W. (1973) Proc. Nat. Acad. Sci. 70, 1559-1563
8 Hayashi, M. N. and Hayashi, M. (1972) J. Virol. 9, 207-215
9 Chen, C.-Y., Hutchison, C. A. III, and Edgell, M. H. (1973) Nature New Biol. 243, 233-236
10 Komano, T. and Sinsheimer, R. L. (1968) Biochim. Biophys. Acta 155, 295-298
11 Radloff, R., Bauer, W., and Vinograd, J. (1967) Proc. Nat. Acad. Sci. 57, 1514-1521
12 Burgess, R. R. and Travers, A. A. (1971) Proc. Nuc. Acid Res. 2, 851-863
13 Sedat, J. W. and Sinsheimer, R. L. (1970) Cold Spring Harbor Symp. Quant. Biol. 35, 163-170
14 Hilmoe, R. (1961) Biochemical Preparations 8, pg. 105. J. Wiley, New York
15 Barrell, B. G. (1971) Proc. Nuc. Acid Res. 2, 751-779
16 Brownlee, G. G. and Sanger, F. (1969) Eur. J. Biochem. , 395-399
17 Bernardi, A. and Cantoni, G. L. (1969) J. Biol. Chem. 244, 1468-1476
18 Bernardi, A. and Bernardi, G. (1968) Biochim. Biophys. Acta 155, 360-370
19 Sugiura, M., Okamoto, T., and Takanami, M. (1969) J. Mol. Biol. 43, 299-315
20 Takanami, M., Okamoto, T., Sugiura, M. (1970) Cold Spring Harbor Symp. Quant. Biol. 35, 179-187
21 Kramer, R. A., Rosenberg, M., and Steitz, J. A. (1974) J. Mol. Biol., in press
22 Blattner, F. R. and Dahlberg, J. E. (1972) Nature New Biol. 237, 227-232