293-314.
Sturtevant, J. M. (1964), in Rapid Mixing and Sampling Techniques in Biochemistry, Chance, B, et al., Ed., New York, N.Y., Academic Press.
Thomas, C. A. (1954), *J. Am. Chem. Soc. .76,* 6032-6034.
Von Hippel, P. H., and Printz, M. P. (1965), *Fed. Proc. Fed. Am. Soc. Exp. Biol. 24,* 1458-1465.
Williams, M. (1968), Thesis, Yale University.
Yamane, T., and Davidson, N. (1961), *J. Am. Chem. Soc. 83,* 2599-2607.
Yamane, T., and Davidson, N. (1962a), *Biochim. Biophys. Acta 55,* 609-621.
Yamane, T., and Davidson, N. (1962b), *Biochim. Biophys. Acta 55,* 780-782.

New Method for Isolation and Sequence Determination of 5′-Terminal Regions of Bacteriophage φX174 in Vitro mRNAs†

Karel Grohmann, Lloyd H. Smith, and Robert L. Sinsheimer*

ABSTRACT: We have determined the nucleotide sequences of the 5′-terminal oligonucleotides, produced by RNase T_1 digestion of bacteriophage φX174 mRNAs synthesized in vitro. The major sequences are: pppCpGp(Ap), pppApUpCpGp(Cp), pppAp(Ap)$_2$UpCp(Up)$_2$Gp(Gp), and pppAp(Ap)$_3$UpCp(Up)$_2$Gp(Gp). The sequences of several minor 5′-terminal oligonucleotides were also determined. For this research we have devised a simple isolation procedure, for the 5′-terminal oligonucleotides, based upon hydroxylapatite chromatography and two-dimensional thin-layer separation. This method allows for the rapid and quantitative recovery of all oligonucleotides containing 5′-triphosphate end groups and should be generally useful for sequence on 5′ termini of mRNAs.

Methods allowing the rapid and quantitative recovery of the 5′ terminal oligonucleotides from messenger RNAs are of vital importance to studies on the control of transcription. Present techniques for the isolation of these oligonucleotides involve separation based upon the change in net electric charge after dephosphorylation (Konrad, 1973) or selective digestion of oligonucleotides which are not phosphorylated on the 5′ end (Takeya and Fujisawa, 1974; Sugiura et al., 1969; Smith et al., 1974). These methods tend to be tedious or difficult to reproduce. Based on the observation by Soave et al. (1973), that 5S RNAs can be separated on hydroxylapatite columns according to the degree of phosphorylation at their 5′ ends, we have derived a simple isolation procedure for oligonucleotides from the 5′ ends of in vitro mRNAs transcribed from bacteriophage φX174 RF DNA.

Materials and Methods

Hydroxylapatite (grade HT) was obtained from Bio-Rad Laboratories, DEAE-cellulose (DE-52) was obtained from Whatman. Other materials were the same as described in a previous publication (Smith et al., 1974).

In vitro φX174 mRNAs selectively labeled with α- or γ-^{32}P ribonucleoside triphosphates were prepared as described previously (Smith et al., 1974) from φX174 RFI DNA and *Escherichia coli* K12 RNA polymerase holoenzyme. These α-^{32}P-labeled transcripts range in size from 500 to 1000 nucleotides as judged by sedimentation through dimethyl sulfoxide gradients (Smith et al., 1974). They also hybridize almost exclusively (>97%) to φX174 complementary strand, in agreement with results obtained by Hayashi et al. (1963, 1964) for both in vivo and in vitro φX174 mRNAs.

Digestion with RNase T_1. Aliquots of the ^{32}P-labeled RNA were digested in 0.01 *M* Tris-HCl (pH 7.9)-0.01 *M* EDTA using RNase T_1 to substrate ratio of 1:10 (w/w). Incubation at 37° for 40 min gave complete digestion.

Other enzymatic digestions were performed as described previously (Smith et al., 1974).

Hydroxylapatite Chromatography. Experiments were routinely performed at room temperature, using disposable columns made from Pasteur pipets (0.5 × 5 cm) previously equilibrated with 0.05 *M* potassium phosphate buffer[1] (pH 6.8). Samples (usually 5-10 μg of RNA digest) were applied in digestion buffer (10 m*M* Tris-HCl (pH 7.9)-10 m*M* EDTA) which often contained traces of phenol from the previous extraction step. The presence of phenol or 10 m*M* EDTA did not interfere with absorption and separation on the column. Samples were loaded by gravity and the columns were washed with 1 ml of 50 m*M* KPO_4 buffer. Elution was achieved with 40 ml of a linear gradient of KPO_4 buffer (pH 6.8), between 0.05 and 0.1 *M*, followed by 10 ml of 0.1 *M* KPO_4 buffer (pH 6.8), and a final wash with 10 ml of 0.5 *M* KPO_4 buffer (pH 6.8). A constant effluent flow rate of 20 ml/hr was provided by means of a peristaltic pump (Buchler Instruments). Fractions of approximately 1 ml were collected. The radioactivity content of effluent fractions was measured by applying aliquots of

† From the Division of Biology, California Institute of Technology, Pasadena, California 91125. *Received November 25, 1974.* This is a report of work supported in part by U.S. Public Health Service Grant No. GM 13554 and in part by U.S. Public Health Service Postdoctoral Research Fellowship No. GM 55249 to K.G.

[1] Abbreviations used are: KPO_4 buffer, potassium phosphate buffer; $TEAHCO_3$, triethylammonium bicarbonate.

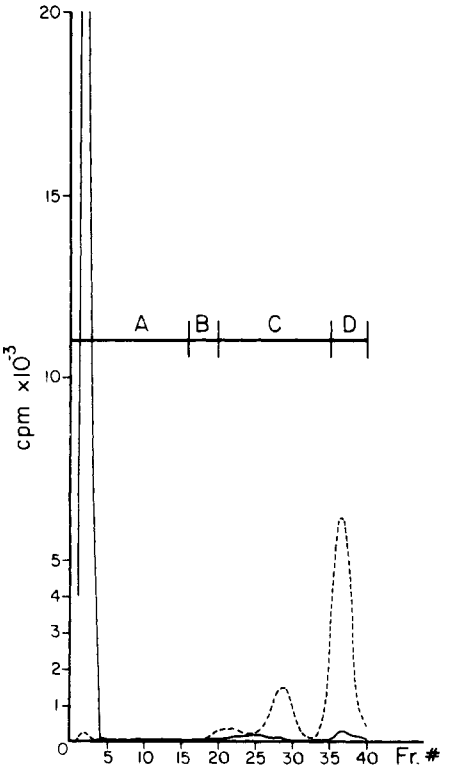

FIGURE 1: Chromatography of oligonucleotides from a combined RNase A and RNase T_1 digest of mixed [γ-^{32}P]ATP- (ca. 3.1 × 10^4 cpm) and [^{3}H]UTP- (ca. 2.6 × 10^5 cpm) labeled, in vitro ϕX174 mRNAs on hydroxylapatite; 1.2-ml fractions were collected and counted in 10 ml of Aquasol (New England Nuclear, Inc.). (—) ^{3}H, (- - -) ^{32}P. Elution steps (KPO_4 buffers, pH 6.8): (A) 10 + 10 ml, 0.05 → 0.1 *M* linear gradient; (B) 5 ml, 0.1 *M*; (C) 10 + 10 ml, 0.1 → 0.25 *M* linear gradient; (D) 5 ml, 0.5 *M* wash.

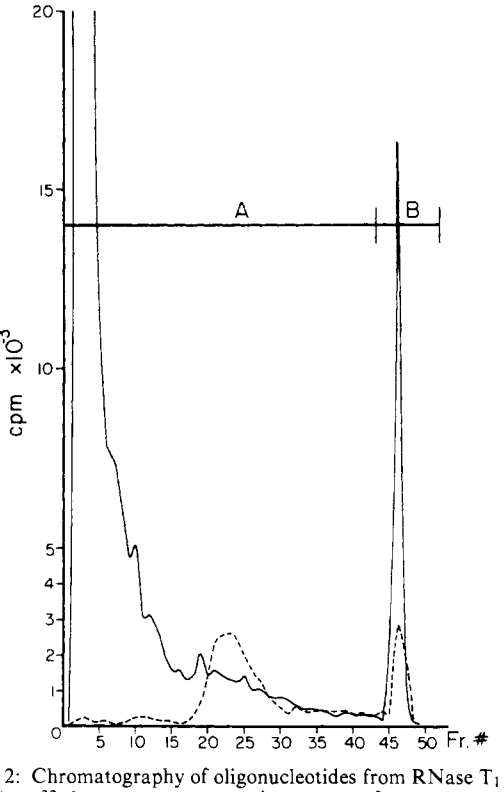

FIGURE 2: Chromatography of oligonucleotides from RNase T_1 digest of mixed [γ-^{32}P]ATP- (ca. 3.6 × 10^4 cpm) and [^{3}H]UTP- (ca. 6.1 × 10^5 cpm) labeled in vitro ϕX174 mRNAs, on hydroxylapatite; 1-ml fractions were collected and counted in 10 ml of Aquasol. (—) ^{3}H, (- - -) ^{32}P. Elution steps (KPO_4 buffers, pH 6.8): (A) 2 ml, 0.05 *M* wash, followed by 20 + 20 ml, 0.05 → 0.25 *M* linear gradient; (B) 10 ml, 0.5 *M* wash.

each fraction onto 3MM paper discs, which were then dried and counted in toluene–Liquifluor (New England Nuclear) counting fluid, using a Beckman LS-200 liquid scintillation counter.

DEAE-Cellulose Chromatography. The three fractions from the 0.5 *M* wash containing the highest level of radioactivity were pooled, diluted with 20 ml of distilled water, and applied by gravity to a 0.5 × 1 cm DEAE-cellulose disposable column poured in a Pasteur pipet. The DEAE-cellulose had previously been equilibrated with 50 m*M* Tris-HCl–1 m*M* EDTA (pH 7.4). The columns were washed with 5 ml of 0.1 *M* triethylammonium bicarbonate (pH 8.0) and then eluted with 4 ml of 1 *M* $TEAHCO_3$ (pH 8.0). The fractions containing the highest level of radioactivity from the 1 *M* $TEAHCO_3$ wash were pooled and evaporated to dryness in 15-ml siliconized centrifuge tubes on a rotary Evapo-mix (Buchler Instruments). Residual $TEAHCO_3$ was removed by repeated evaporation with equal volumes of water. The oligonucleotides were then dissolved in 2–10 μl of 5 m*M* $EDTANa_2$.

Two-dimensional separation of the oligonucleotides was carried out either as described by Barrell (1971) or by electrophoresis on cellulose thin-layer, followed by homochromatography on DEAE-cellulose (Grohmann and Sinsheimer, 1975). Procedures for recovery of spots, redigestion with RNases, and characterization of products have been described previously (Smith et al., 1974).

Results

(a) Evidence for Separation of Oligonucleotides Bearing a 5′-Triphosphate Group on Hydroxylapatite Columns. In vitro ϕX174 mRNAs selectively labeled at the 5′ end with [γ-^{32}P]ATP or [γ-^{32}P]GTP, and uniformly labeled with [5-^{3}H]UTP, were mixed, digested with RNases, and chromatographed on hydroxylapatite columns using potassium phosphate gradients. Figure 1 shows an elution profile of such a sample after digestion with a mixture of RNase A and RNase T_1. Comparison of the ^{3}H and ^{32}P profiles shows that non-5′ terminal oligonucleotides elute at the beginning of the gradient, while oligonucleotides from the 5′ end are selectively retarded and principally begin to elute with 0.25 *M* KPO_4 buffer. (The nature of the two small peaks eluting at lower molarity was not investigated.)

The hydroxylapatite profiles of γ-^{32}P-labeled RNA after digestion with RNase T_1 alone are strikingly different (Figure 2); the γ-^{32}P-labeled oligonucleotides then elute at a lower phosphate concentration beginning at approximately 0.13 *M*. Some ^{3}H-labeled oligonucleotides also elute at phosphate concentrations (0.1–0.2 *M*) where very low radioactivity is observed in the separation of the combined RNase A + T_1 digest.

In order to determine whether the lower molarity required for elution of the γ-^{32}P-labeled T_1 oligonucleotides is a consequence of their larger size or of their interaction with other oligonucleotides present in RNase T_1 digests, we chromatographed the RNA obtained by a combined digestion with RNase T_1 and spleen exonuclease. In this digest all oligonucleotides that are not protected by 5′-triphosphate groups are degraded (Smith et al., 1974). The γ-^{32}P peak (data not shown) now eluted at 0.25 *M* KPO_4. Control experiments indicated that the spleen exonuclease treatment caused no alteration in the integrity of the γ-^{32}P-labeled T_1 oligonucleotides. We therefore conclude that the

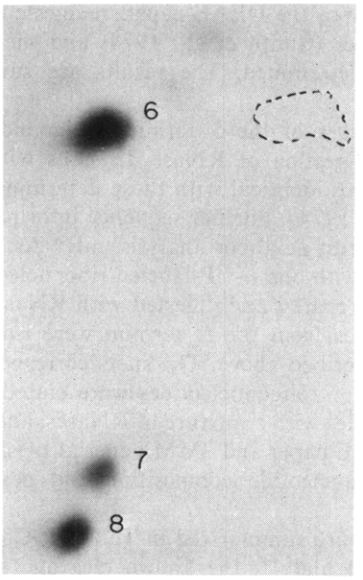

FIGURE 3: Autoradiogram of RNase T_1 digest of [γ-^{32}P]ATP-labeled in vitro φX mRNA. First dimension (electrophoresis) is from left to right; second dimension (homochromatography) is from bottom to top. A mixture of 3% 30′ homo C (Barrell, 1971) and 3% 20′ homo C (4:1 v/v) was used for development in the second dimension. Dashed circle marks position of the Orange G (yellow) dye marker.

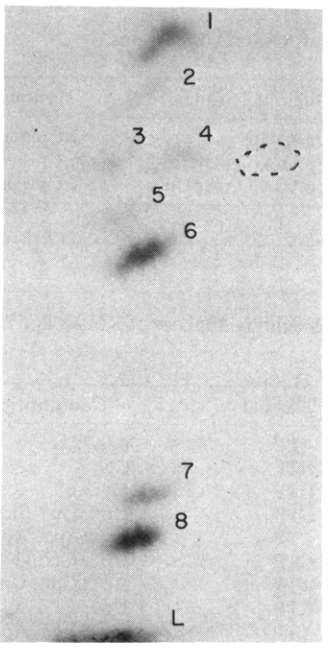

FIGURE 4: Autoradiogram of RNase T_1 digest of φX in vitro φX174 mRNA labeled with all four α-^{32}P-labeled ribonucleoside triphosphates, purified as described in the text. Procedures are the same as used to produce Figure 3.

observed elution with 0.13 *M* KPO_4 of the γ-^{32}P-labeled RNase T_1 oligonucleotides is a consequence of their interactions with nontriphosphorylated oligonucleotides in the digest.

This effect can also explain discrepancies in the elution molarities published for adenosine tetraphosphate (ca. 0.27 *M* KPO_4 buffer (pH 6.8)) (Bernardi, 1964) and those for native partially double-stranded 5S RNA containing a triphosphate group at the 5′ end (0.15 *M* KPO_4 buffer (pH 6.8)) (Soave et al., 1973).

Chromatography under denaturing conditions (60% Me_2SO at room temperature or aqueous buffers at 70°) did not prove helpful. However, the nonterminal oligonucleotides in T_1 digests which elute from hydroxylapatite columns with the terminal fragments are much longer than the terminal fragments. They are 20 nucleotides in length or greater. Thus they can be separated from the terminal fragments by homochromatography. We have therefore adjusted the ionic strength of the gradients used for hydroxylapatite separation (as described under Materials and Methods), so that the terminal fragments are eluted together, with large nonterminal oligonucleotides, after completion of the gradient. They are subsequently separated by homochromatography.

(b) Determination of Nucleotide Sequences of the 5′-Terminal Oligonucleotides Produced by RNase T_1 Digestion. The 5′-terminal oligonucleotides isolated by the hydroxylapatite technique from an RNase A digest of ^{32}P in vitro φX174 mRNAs have nucleotide sequences identical with those we have previously determined by the spleen exonuclease procedure (Smith et al., 1974). Therefore we have proceeded to the determination of the 5′-terminal oligonucleotides produced by RNase T_1 digestion. Figure 3 is an autoradiogram of a two-dimensional fingerprint of oligonucleotides produced by digestion of [γ-^{32}P]ATP-labeled φX174 in vitro mRNA with RNase T_1. Three major radioactive spots can be seen (labeled 6–8), together with several minor ones. [The appearance of minor spots is a function of the age of the RNA polymerase preparation as discussed previously (Smith et al., 1974). The RNA polymerase holoenzyme used in the work described here had been stored (−20°, 50% glycerol storage buffer) for 2 months after purification. The minor oligonucleotide spots seen in Figure 3 are greatly reduced in fingerprints of RNA prepared from fresh RNA polymerase and contribute less than 5% of the total radioactivity on the fingerprint.]

Figure 4 is an autoradiogram of a two-dimensional fingerprint of a RNase T_1 digest of φX174 in vitro mRNA, uniformly labeled with ^{32}P by inclusion of all four α-^{32}P-labeled ribonucleoside triphosphates. The oligonucleotides from the 5′ termini were purified by the technique described above. Approximately 7% of the applied radioactivity was eluted from the hydroxylapatite column by 0.5 *M* KPO_4 buffer. Recoveries of label from both hydroxylapatite and DEAE-cellulose columns were over 90%.

As stated above, the large non-terminal oligonucleotides which also elute from the hydroxylapatite column with 0.5 *M* KPO_4 buffer are separated during the fingerprinting. A portion of these remains close to the origin in the first dimension and is therefore cut off during transfer. The remainder (labeled L in Figure 4), which travel in the first dimension in the region of the acid fuchsin (pink) dye marker, do not migrate at all in the homomixture C (Barrell, 1971). The maximum length of oligonucleotides which will migrate in our homomixture (see legend to Figure 3) is 20–25 nucleotides. Nucleotide composition analysis shows that the oligonucleotides which remain at the origin are highly enriched in pyrimidines and that they do not contain significant amounts of triphosphate ends.

A comparison of Figures 3 and 4 shows that we have isolated all three major terminal oligonucleotides starting with ATP (labeled 6–8 on both figures) and several other less frequent terminal fragments (spots labeled 1–5 in Figure 4). Identification of these oligonucleotides is discussed below.

Table I: Products of RNase A Digestions.

Spot	Oligonucleotides	Mononucleotides[a]
6	pppApUp	*n*Cp + 1 Gp
7	pppApApApUp	*n*Cp + *n*Up + 1 Gp
8	pppApApApApUp	*n*Cp + *n*Up + 1 Gp

[a] *n* = 1 or 2.

Table II: Nearest Neighbor Analysis of RNase T_1 5′ Oligonucleotides.

Spot	α-^{32}P Label	Composition (Rel Yield)
6	ATP	ppp*Ap
	GTP	Cp*
	UTP	pppAp*
	CTP	Up*, Gp* (0.8:1)
7	ATP	ppp*Ap*, Ap* (2.4:1)
	GTP	Up*, Gp* (1:0.96)
	UTP	Ap*, Cp*, Up* (1:1.02:0.98)
	CTP	Up*
8	ATP	ppp*Ap*, Ap* (1.2:1)
	GTP	Up*, Gp* (1.02:1)
	UTP	Ap*, Cp*, Up* (1:1.16:1.1)
	CTP	Up*

[a] * denotes position of ^{32}P label.

Table III

Spot	Sequence
6	pppApUpCpGp(Cp)
7	pppApApApUpCpUpUpGp(Gp)
8	pppApApApApUpCpUpUpGp(Gp)

(b-1) Nucleotide Sequences of the Major pppA—Oligonucleotides. All spots were eluted and redigested with ribonuclease A. Digests were split into halves and fractionated by electrophoresis on DEAE paper and 3MM paper at pH 3.5 (Smith et al., 1974). The digestion products of each eluted spot were mononucleotides and shorter oligonucleotides containing a 5′-triphosphate end group. These oligonucleotides stay close to the origin on DEAE paper and travel between AMP and P_i on 3MM paper. (Note: these streak badly on 3MM paper.) These oligonucleotides were individually eluted from the DEAE paper, redigested with a mixture of RNases (Smith et al., 1974) and their nucleotide composition determined. The results are summarized in Table I.

The sequences of the 5′-terminal oligonucleotides produced by redigestion of RNase T_1 spots with pancreatic RNase are thus identical with those determined previously (Smith et al., 1974). Further sequence information was deduced by nearest neighbor analysis. mRNAs, each specifically labeled with one α-^{32}P-labeled ribonucleoside triphosphate, were prepared and digested with RNase T_1 and the oligonucleotides from the 5′ termini were isolated by the procedure described above. The spots corresponding to the major pppAp . . . oligonucleotides were eluted, digested to mononucleotides with a mixture of RNases, and electrophoresed on DEAE paper and 3MM paper at pH 3.5 for determination of nucleotide composition and position of the label.

The results are summarized in Table II. Combined data from Tables I and II, the known cleavage specificity of RNases T_1 and A and known nearest 3′ neighbor of the RNase A oligonucleotides (Smith et al., 1974) enable us to determine the sequences of the major pppAp . . . oligonucleotides (Table III). The nucleotide responsible for transfer of ^{32}P to the 3′ terminal GMP is shown in parentheses.

(b-2) Nucleotide Sequences of Remaining Oligonucleotides (Spots 1–5 in Figure 4).

(i) NUCLEOTIDES CONTAINING 5′-PPPCP. Digestion of spots 1 and 3 with either RNase A or mixture of RNases, followed by electrophoresis on DEAE-cellulose paper at pH 3.5, produces mononucleotides and a new spot with a mobility of approximately two times that of pppAp. The R_f of this nucleotide relative to xylene cyanol FF (blue) marker is 0.18 (R_f of pppAp is 0.08 and pppGp is 0.05). Since identical products are formed by cleavage of spots 1 and 3 with RNase A and nonspecific RNases, this nucleotide must be a pyrimidine polyphosphate. The following evidence leads us to the conclusion that this nucleotide is cytosine tetraphosphate (pppCp). (1) [γ-^{32}P]CTP (but not [γ-^{32}P]UTP) is incorporated into in vitro φX174 mRNA in amounts exceeding incorporation of [γ-^{32}P]GTP (Smith and Sinsheimer, in preparation). Fingerprints of RNase T_1 digests of this mRNA show only two spots, which are in the same position as spots 1 and 3. (2) Our presumptive pppCp comigrates in one-dimensional electrophoresis with [γ-^{32}P]pppCp prepared by digestion of [γ-^{32}P]CTP-labeled in vitro φX174

Table IV: Nucleotide Sequences of Minor Oligonucleotides.[a]

Spot	Label	RNase A Products	Nucleotide Composition	Sequence
1	XTP	pppCp, Gp	pppCp, Gp	
	CTP		pppCp	pppCpGp(Ap)
	GTP		pppCp	
	ATP		Gp	
2	XTP	pppApGp, pppCp, Cp, G>	pppAp, pppCp Cp, Gp	Mixture of pppApGp pppCpCpG>*
3	XTP	pppCp, Cp, Gp	pppCp, Cp, Gp	pppCpCpGp*
4	XTP	pppApUp, Gp	pppAp, Up, Gp	pppApUpGp(Gp)
	UTP		pppAp	
	GTP		Up, Gp (1.07:1.0)	
5	XTP	N.C.	pppAp, Ap, Gp	
	ATP		pppAp	pppApApGp
	GTP		Ap	

[a] N. C., not cleaved; G>, 2′,3′-GMP; *, the position of spots 2 and 3 on fingerprints makes it unlikely that more than 1 C is inserted between 5′ and 3′ end nucleotides.

mRNA.

Further evidence comes from sequence analysis of spot 1. Upon digestion of spot 1 with either RNase A or mixture of RNases, only GMP and pppCp are produced. Upon nearest neighbor analysis of spot 1, pppCp is labeled with ^{32}P by GTP and CTP only. If the mRNA is synthesized with [α-^{32}P]ATP the GMP of spot 1 is labeled; therefore, we conclude that sequence of spot 1 is pppCpGp(Ap). The small amount of radioactivity in spot 3 precluded nearest neighbor analysis. Digestion of spot 3 with either RNase A or mixed RNases produced pppCp, CMP, and GMP. Based upon these results and the position of spot 3 on fingerprint, we suggest that sequence of spot 3 is pppCpCpGp.

(ii) NUCLEOTIDE SEQUENCES OF REMAINING MINOR OLIGONUCLEOTIDES were determined by the same technique as described above. Insufficient amount of label in these spots provided only partial data for nearest neighbor analysis. Results for oligonucleotides 1–5 are summarized in Table IV.

The pppGpApU 5′ terminus previously described (Smith et al., 1974) yields pppGp upon RNase T_1 digestion. This nucleotide is also isolated by the hydroxylapatite column procedure described above, but has higher mobility than the yellow dye marker in both dimensions during fingerprinting. Therefore it is sometimes (as is the case in Figure 4) cut off during transfer.

Discussion

The work presented in this paper shows that chromatography on the hydroxylapatite can be used as a simple procedure for isolation of oligonucleotides bearing triphosphate groups at the 5′ end. In case that contamination with long oligonucleotides would prove to be a problem, they can be removed before hydroxylapatite chromatography by digestion with spleen exonuclease, as we have shown previously (Smith et al., 1974). Together with our previous publication (Smith et al., 1974) we can state that there are five major initiation sequences of ϕX174 in RNAs in vitro (spots 1, 6, 7, and 8 and pppGpApU [Smith et al., 1974]). Work is in progress (L. H. Smith and R. L. Sinsheimer, manuscript in preparation) to localize these sites on ϕX174 genetic map. We are also extending this technique to the determination of the starting nucleotide sequences of in vivo ϕX174 mRNAs. Preliminary results strongly indicate that the major in vitro and in vivo 5′ terminal nucleotide sequences are identical, with the single exception of pppCpGp, which we observe only in vitro.

The occurrence of the 5′ pyrimidine triphosphate sequence in in vitro ϕX174 mRNA may be an artifact; pyrimidine triphosphate termini have rarely been observed in other in vitro systems (Geiduschek and Hasselkorn, 1969). However, the occurrence of pppCp oligonucleotides in amounts comparable to the purine triphosphate oligonucleotides and their presence in RNA made with fresh RNA polymerase indicate that such oligonucleotides derive from major initiation sites in vitro on ϕX174 RFI.

Acknowledgment

The authors acknowledge the excellent technical assistance of Ms. Carole Latham and Ms. Janice M. Wodnicki.

References

Barrell, B. G. (1971), *Proced. Nucleic Acid. Res. 2,* 751.
Bernardi, G. (1964), *Biochim. Biophys. Acta 91,* 686.
Geiduschek, P. E., and Hasselkorn, R. (1969), *Annu. Rev. Biochem. 38,* 647.
Grohmann, K., and Sinsheimer, R. L. (1975), *Anal. Biochem.* (in press).
Hayashi, M., Hayashi, M. N., and Spiegelman, S. (1963), *Proc. Natl. Acad. Sci. U.S.A. 50,* 665.
Hayashi, M., Hayashi, M. M., and Spiegelman, S. (1964), *Proc. Natl. Acad. Sci. U.S.A. 51,* 351.
Konrad, M. W. (1973), *Anal. Biochem. 54,* 213.
Smith, L. H., Grohmann, K., and Sinsheimer, R. L. (1974), *Nucleic Acids Res. 1,* 1521.
Soave, C., Nucca, R., Sala, E., Viotti, A., and Galante, E. (1973), *Eur. J. Biochem. 32,* 392.
Sugiura, M., Okamoto, T., and Takanami, M. (1969), *J. Mol. Biol. 43,* 299.
Takeya, T., and Fujisawa, H. (1974), *Biochim. Biophys. Acta 324,* 110.