JOURNAL OF VIROLOGY, Feb. 1980, p. 902-908
0022-538X/80/02-0902/07$02.00/0

Vol. 33, No. 2

# NOTES

## Localization of Three Major Capped 5′ Ends of Polyoma Virus Late mRNA's Within a Single Tetranucleotide Sequence in the Viral Genome†

ANDREW J. FLAVELL,‡ ALISON COWIE, JOHN R. ARRAND, AND ROBERT KAMEN*
*Imperial Cancer Research Fund, London WC2A 3PX, England*

The nucleotide sequences at the 5′ ends of polyoma virus late mRNA's were determined by direct RNA sequencing of decapped and 5′-$^{32}$P-labeled RNase T1 oligoribonucleotides. Virus-specific polyadenylated cytoplasmic RNA, which was isolated from mouse cells late during productive infection, was enzymatically or chemically treated to specifically remove the cap structure ($m^7$Gppp). The unblocked 5′ ends of the viral mRNA's were then labeled enzymatically with $^{32}$P, and the RNAs were digested with RNase T1 and fingerprinted. Three oligonucleotides derived from capped termini were isolated, and their sequences were determined to be pGmACAUUUUCUAUUUUAAGp, p(m)AmCAUUUUCUAUUUUAAGp, and p(m)AmUUUUCUAUUUUAAGp. These oligonucleotides comprise a staggered set with members 15, 17, and 18 nucleotides long, which share a common 3′ sequence for 15 nucleotides. The sequences correspond exactly to the polyoma virus DNA sequence (Arrand et al., J. Virol. **33**:606-618, 1980) from 66.79 to 66.46 map units (between 75 and 92 nucleotides preceding the ATG initiation codon for the capsid protein VP2). Previous results showed that the sequence between 13 and 64 nucleotides preceding the VP2 initiation codon corresponds to oligonucleotides reiterated in the leader sequence which is spliced onto the bodies of the three functionally distinct viral late mRNA's (Flavell et al., Cell **16**:357-372, 1979; Legon et al., Cell **16**:373-388, 1979). The three capped oligonucleotides we sequenced are derived from the first large predicted T1 oligonucleotide 5′ to those detected in the leader sequence. The occurrence of a cap at each purine of a single tetranucleotide sequence reflects micro-heterogeneity either in transcriptional initiation or in processing cleavage involved in cap syntheses.

Three functionally distinct viral mRNA's are synthesized specifically at late times during the productive infection of mouse cells by polyoma virus (15, 17, 23). These mRNA's encode the virion proteins VP1, VP2, and VP3 (16, 32). Each late mRNA comprises a colinear transcript of DNA containing the entire coding region spliced to a related family of 5′-terminal leader sequences. We recently found that the leader sequences have two unusual features: (i) the late leaders can have any of at least seven different capped 5′ termini, which all map within the same restricted portion of the viral genome (6); and (ii) the leaders contain repeats of a nucleotide sequence represented uniquely in polyoma virus DNA (22). The leader sequences of the late mRNA's of the closely related simian virus 40 (SV40) have a similarly complex structure, although only a minority of these leaders contain sequence repeats (10, 20, 27, 37). In this paper we report the 5′-terminal nucleotide sequences of three of the several polyoma virus late mRNA leaders. A comparison of these results with the DNA sequence of the late region (1) allows us to specify the exact locations of the 5′ ends of these molecules on the viral genome.

The approach taken was to isolate and then sequence capped RNase T1 oligonucleotides. The heterogeneity of the 5′-terminal capped structures in polyoma virus mRNA's causes individual capped T1 oligonucleotides to be present in low molar yields with respect to oligonucleotides from internal regions. It is thus not

† This paper and the accompanying papers are dedicated to M. G. P. Stoker on the occasion of his retirement as Director of Research of the Imperial Cancer Research Fund.

‡ Present address: Sidney Farber Cancer Institute, Boston, MA 02115.

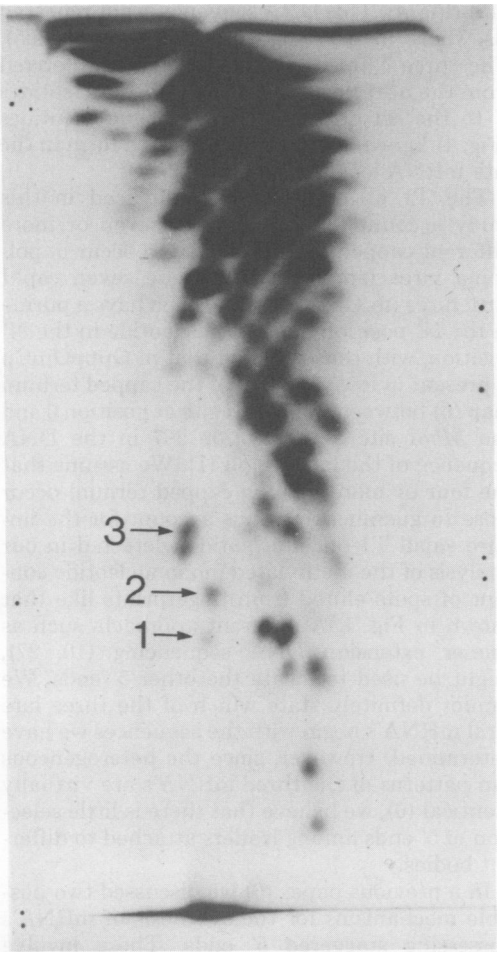


FIG. 1. *RNase T1 fingerprint of enzymatically decapped (4) and 5′ terminally labeled polyoma virus mRNA. Polyadenylated mRNA was isolated (6) from about 5 × 10⁸ 3T6 cells infected with plaque-purified polyoma virus (A2 strain) at 50 PFU/cell and labeled with $^{32}P$ (100 μCi/90-mm culture dish) between 12 and 30 h postinfection. Viral mRNA (20 μg, 300 cpm/μg) was purified by hybridization to polyoma virus DNA cellulose (200 μg of bound DNA) as described previously (6). The desalted RNA precipitate obtained by ethanol precipitation and rinsing with 95% ethanol was dissolved in 20 μl of a buffer containing 50 mM sodium acetate, pH 4.5, and 10 mM 2-mercaptoethanol; 4 U of tobacco acid pyrophosphatase (Bethesda Research Laboratories, Inc.) was added, and the RNA was incubated for 30 min at 37°C. After the addition of 20 μl of 0.2 M Tris-hydrochloride, pH 7.5, and 0.1 U of bacterial alkaline phosphatase (P.L. Biochemicals) treated with diethyl pyrocarbonate to inactivate RNase, as described by Pederson and Haseltine (Methods Enzymol., in press), incubation at 37°C was continued for an additional 30 min. The reaction was terminated by adding 100 μl of a solution containing 0.3 M sodium acetate, pH 5.3, and 0.2% sodium dodecyl sulfate, extracted with phenol, and then extracted with diethyl ether, and the RNA was recovered by precipitation with 2 volumes of ethanol. The decapped RNA, dissolved in a mixture of 225 μl of water and 25 μl of 10× kinase buffer (0.5 M Tris-hydrochloride, pH 8.0, 0.1 M $MgCl_2$, 50 mM dithiothreitol), was transferred to a siliconized glass tube containing 500 μCi of [γ-$^{32}P$]ATP (Radiochemical Centre; 3,000 Ci/mmol). T4 polynucleotide kinase (P.L. Biochemicals; 4 μl; 1 U/μl) was added, and the solution was incubated at 37°C for 30 min. The reaction was terminated by adding 50 μl of 5 M ammonium acetate, 2 μl of 20% (wt/vol) sodium dodecyl sulfate, 20 μg of yeast RNA, and 2 volumes of ethanol. The RNA was recovered by centrifugation, redissolved in 0.3 M sodium acetate (pH 5.3)–0.2% sodium dodecyl sulfate (100 μl), and reprecipitated from ethanol. It was then redissolved in 5 μl of 10 mM Tris-hydrochloride (pH 7.5)–1 mM EDTA containing 1 μg of RNase T1 (Calbiochem) and incubated for 30 min at 37°C. A total of 90% of the sample (2 × 10⁷ cpm [Cerenkov]) was fingerprinted by electrophoresis on cellulose acetate at pH 3.5, followed by homochromatography on a DEAE-cellulose thin layer, as described by Barrell (3). The oligonucleotides marked 1, 2, and 3 contained 5′-terminal 2′-O-methylated nucleotides (see Fig. 2).*

surprising that we could not detect the capped oligonucleotides in fingerprints of uniformly labeled viral mRNA. Therefore, we isolated viral mRNA labeled with $^{32}P$ in vivo in substantially the same manner as reported previously (6) but at low specific radioactivity and after removal of the blocking cap structure, we attempted to label specifically the 5′ ends of the intact mRNA by using T4 polynucleotide kinase and [γ-$^{32}P$]-ATP. To remove the cap structure, we employed both enzymatic and chemical methods. In the former approach we used the technique of Efstratiadis et al. (4), which involves digestion of the pyrophosphate bonds in the capped termini with tobacco acid pyrophosphatase to yield phosphatase-sensitive termini ($m^7GpppN^1mN^2 \ldots \rightarrow pN^1mN^2 \ldots.$). In the chemical approach we employed periodate oxidation of the 3′-terminal ribonucleotide *cis*-diol groups of the caps, followed by base-catalyzed β elimination of the 7-methyl guanoside group from the RNA (8, 36). The decapping procedures were followed by phosphate digestion, giving 5′ hydroxyl termini. Treatment of the RNA with $T_4$ polynucleotide kinase and [γ-$^{32}P$]ATP yielded mRNA's labeled with $^{32}P$ at their 5′ ends ($^{32}pN^1mpN^2p \ldots$). Both chemical and enzymatic methods produced labeled RNA that was somewhat degraded, resulting in the incorporation of $^{32}P$ into internal as well as 5′-terminal oligonucleotides. Figure 1 shows a RNase T1 fingerprint of enzymatically decapped and kinase-labeled polyoma virus late mRNA. Essentially the same fingerprint was obtained with chemically decapped messenger.

The complexity of the fingerprint is consistent with the extent of degradation of the RNA after the decapping and labeling procedure.

The oligonucleotides from the fingerprint shown in Fig. 1 (as well as the corresponding oligonucleotides from equivalent fingerprints of other preparations of enzymatically or chemically decapped and kinase-treated viral mRNA) were eluted, and a portion of each was digested to completion with nuclease P1 to obtain the 5′-terminal radioactive mononucleotide. Oligonucleotides derived from capped structures yield a 2′-*O*-methylated nucleotide ($^{32}pN^1mpN^2$ ... Gp $\rightarrow$ $^{32}pN^1m$ plus unlabeled mononucleotides). The 5′-mononucleotides were identified by two-dimensional thin-layer chromatography, as described by Nishimura (25), and by one-dimensional electrophoresis on Whatman no. 1 paper at pH 3.5; the former fractionation resolves all mononucleotides except pC and pGm, which are separated in the latter system. Only three of the large T1 oligonucleotides yielded a methylated mononucleotide (Fig. 2), and these are indicated in Fig. 1. Oligonucleotide 1 contained $^{32}$pGm, and oligonucleotides 2 and 3 contained $^{32}$pmAm. Several of the very small T1 oligonucleotides also contained minor amounts of labeled methylated mononucleotide, suggesting that they were mixtures of oligonucleotides derived from capped termini and from internal positions; these oligonucleotides were too impure for further characterization. Curiously, nearly all of the other T1 oligonucleotides contained $^{32}$pA.

The majority of the radioactivity eluted from oligonucleotide spots 1, 2, and 3 (Fig. 1) was digested partially with nuclease P1, and the products were fractionated in a manner similar to that used to produce the original fingerprint. This generates a wandering spot pattern of radioactive partial digestion products, and the sequence of the original oligonucleotide can be determined from the characteristic mobility shifts obtained (33). The partial nuclease P1 fingerprints are shown in Fig. 3, together with the deduced sequence of each. The three oligonucleotides are closely related. They have a common 3′ sequence (UUUUCUAUUUUAAGp) but differ at their 5′ ends. Oligonucleotide 3 has an additional $^{32}$p(m)AM, oligonucleotide 2 has $^{32}$p(m)AmC, and oligonucleotide 1 has the further trinucleotide $^{32}$pGmACA. When compared with the sequence of the E strand of polyoma virus DNA (1), which has the same polarity as late mRNA's (18), the sequences of the three oligonucleotides match exactly the sequence of the DNA between 75 and 92 nucleotides preceding the ATG initiation codon for capsid protein VP2 (nucleotides 198 through 215) (Fig. 4). This DNA sequence is not repeated elsewhere in polyoma virus DNA (1, 33a–35). The three capped oligonucleotides are derived from the next possible large T1 oligonucleotides 5′ to the set of four large T1 oligonucleotides (Fig. 4) known to occur in reiterated form in the late mRNA leader sequence (22).

The T1 oligonucleotides sequenced in this study account for three of the seven or more different capped termini known to occur in polyoma virus late mRNA's. These seven cap I structures ($m^7GpppN^1mN^2p$) each have a purine in the $N^1$ position and any nucleotide in the $N^2$ position, with the exception that $m^7$GpppGmCp is present in low yield. All of the capped termini map (6) between the *Hpa*II site at position 0 and the *Mbo*I site at nucleotide 267 in the DNA sequence of the late region (1). We assume that the four or more missing capped termini occur close to guanines and thus account for the impure small T1 oligonucleotides detected in our analysis of the methylated mononucleotide content of spots eluted from fingerprints like that shown in Fig. 1. A different approach, such as primer extension DNA sequencing (10, 27), might be used to locate the other 5′ ends. We cannot definitely state which of the three late viral mRNA's begin with the sequences we have determined. However, since the heterogeneous cap patterns of the three mRNA's are virtually identical (6), we believe that there is little selection of 5′ ends among leaders attached to different bodies.

In a previous paper (6) we discussed two possible mechanisms for the synthesis of mRNA's possessing staggered 5′ ends. These involve either the nonspecific initiation of transcription followed by capping of all nascent 5′ ends or the nonspecific processing of nuclear transcripts followed again by capping. We are still unable to determine which of these two possible mechanisms operates here, but favor the former mechanism for two reasons. First, we observe only purine-containing caps in polyoma virus-specific late mRNA's, and transcriptional initiation by RNA polymerase II in eucaryotes is thought to occur at purine nucleotides (2, 29, 30). Second, transcriptional initiation in SV40 is thought to occur in the region of the DNA genome analogous to that containing these capped termini (7, 21). Similarly, the capped ends of adenovirus late mRNA's (38) are thought to correspond to transcriptional start points. If the former mechanism is indeed the correct one, then we have mapped a major promoter region of polyoma virus DNA and have shown that transcriptional initiation may occur at a number of purine nucleotides in this restricted region.

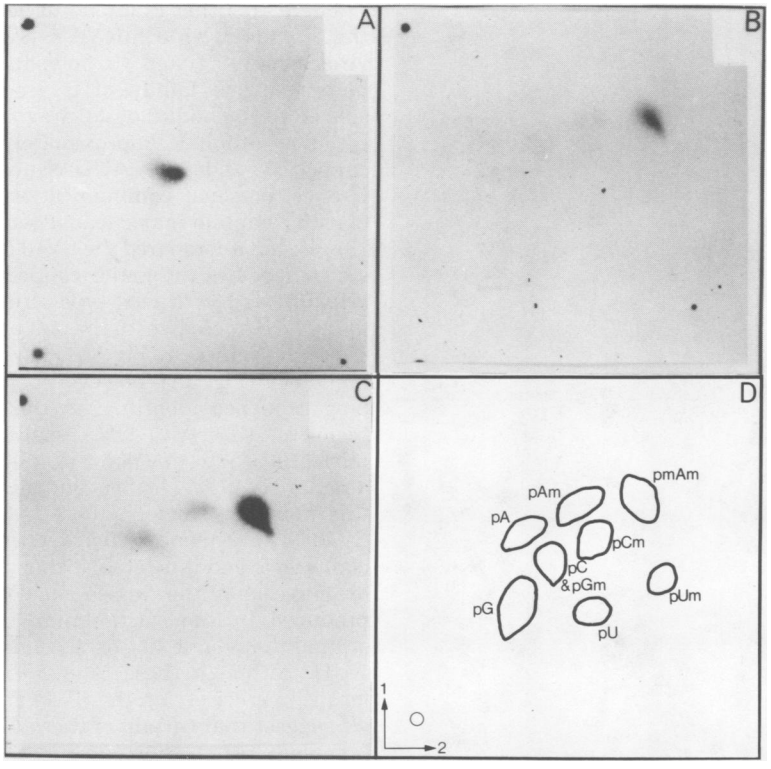


FIG. 2. *Two-dimensional fractionation of nuclease P1 digests of 5'-$^{32}$P-labeled RNase T1 oligonucleotides derived from decapped and kinase-treated polyoma virus mRNA. Cytoplasmic polyadenylated late viral mRNA (see legend to Fig. 1) was chemically decapped by using a modification of the method described by Steinschneider and Fraenkel-Conrat (36). The RNA (10 μg) was dissolved in 40 μl of water to which were added 5 μl of 1 M sodium acetate (pH 5.0), 0.5 μl of 0.1 M EDTA, and 10 μl of freshly prepared 10 mM $NaIO_4$; after incubation for 2 h in the dark room temperature, 1 μl of ethylene glycol, 10 μl of 20% (wt/vol) sodium acetate (pH 5.3), and 150 μl of ethanol were added. The oxidized RNA was recovered by centrifugation washed with 95% ethanol, redissolved in 50 μl of 0.33 M aniline hydrochloride, pH 5.0, and incubated in the dark for 2 h at room temperature. The decapped RNA was precipitated by the addition of 5 μl of 5 M NaCl and 2 volumes of ethanol. The RNA pellet was redissolved in 50 μl of 50 mM Tris, pH 7.5, containing 0.1 U of bacterial alkaline phosphatase, incubated at 37°C for 30 min, extracted with phenol (three times) and diethyl ether, and then reprecipitated. The recovered RNA was then 5'-labeled with [γ-$^{32}$P]ATP (100 μCi) and polynucleotide kinase as described in the legend to Fig. 1. After the labeling reaction, the [$^{32}$P]RNA was further purified by hybridization to polyoma DNA cellulose (6), and the radioactive RNA was digested with RNase T1 and fingerprinted as described in the legend to Fig. 1. The fingerprint obtained was virtually identical to the fingerprint of enzymatically decapped RNA shown in Fig. 1. The 26 clearly resolved spots were eluted, and each oligonucleotide was digested to completion with nuclease P1 (2 μg; Calbiochem) in 50 mM sodium acetate, pH 6. The P1 digestions were fractionated on cellulose thin-layer plates (10 by 10 cm; Merck), using the solvent systems described by Nishimura (25). Nonradioactive 5'-mononucleotides (5 μg each of pA, pG, pU, pC, pAm, pmAm, pGm, pUm, and pCm) were run as internal markers on each plate. After the two-dimensional chromatography, the dried plates were fluorographed (Fuji Rx film, preflashed; Fuji Mach II intensifying screens), and the spots on the resulting films were aligned with tracings of the patterns of marker mononucleotides. (D) Example of the marker pattern. Two of the oligonucleotides (B and C), corresponding to those marked 2 and 3 in Fig. 1, contained a $^{32}$P-labeled nucleotide which comigrated with $N^6$,2'-O-dimethyladenosine 5'-phosphate (pmAm) and 2'-O-methyladenosine 5'-phosphate (pAm). A third oligonucleotide (oligonucleotide 1) (A) contained a $^{32}$P-labeled nucleotide which comigrated with either 2'-O-methyl guanosine 5'-phosphate (pGm) or pC, which are not fully separated in the system used. To resolve this ambiguity, the mononucleotide was eluted and further fractionated by electrophoresis on Whatman no. 1 paper at pH 3.5 (data not shown). The radioactivity comigrated with pG, demonstrating that oligonucleotide 1 contained $^{32}$pGm. Electrophoretic analysis of the mononucleotides from oligonucleotides 2 and 3 confirmed their identification as pmAm. The 23 other T1 oligonucleotides each contained only $^{32}$pA, and these were thus not derived from capped termini.*

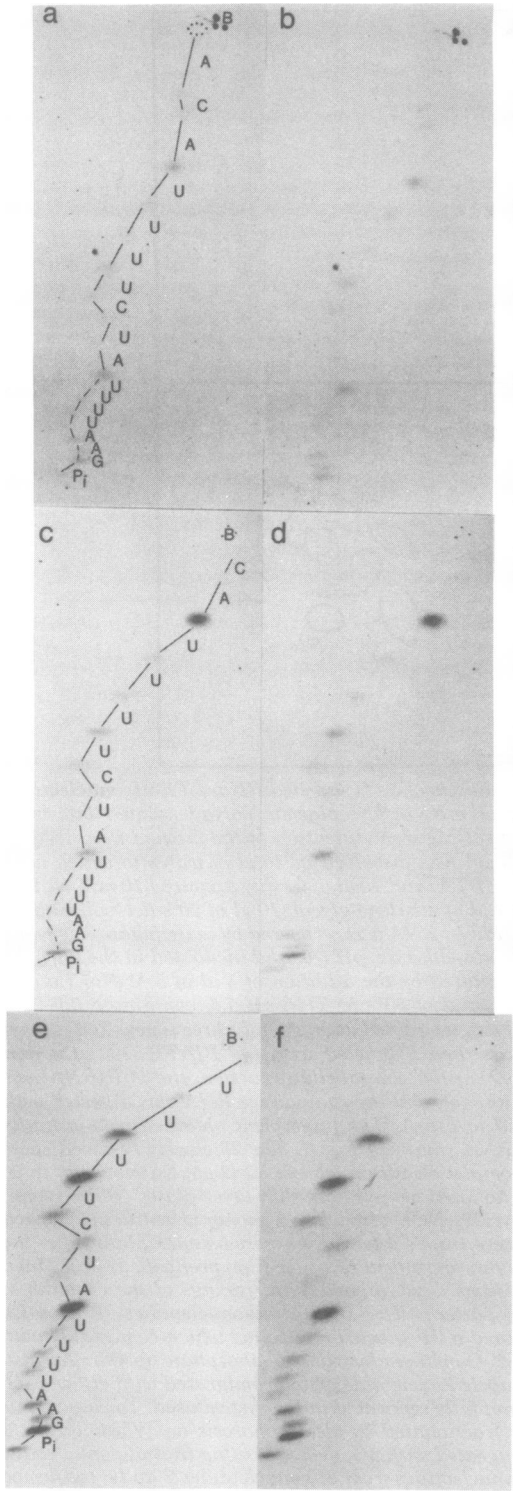


FIG. 3. *Fingerprints of partial digestion products of oligonucleotide spots 1 through 3 (Fig. 1). Oligonucleotide spots 1 through 3 in Fig. 1 were eluted as described by Barrell (3). They were then digested with P1 nuclease (0.2 μg/ml; Sigma) in 20 μl of sodium acetate, pH 6.0, at 20°C for 30 min. The enzyme was destroyed by heating at 100°C for 2 min, and the digests were fingerprinted in a manner similar to that employed in the original fingerprint in Fig. 1. (a) and (b) Fingerprint of partial digestion products of oligonucleotide spot 1 from Fig. 1 together with the deduced sequence. (c) and (d) Fingerprint of partial digestion products of oligonucleotide spot 2 from Fig. 1 together with the deduced sequence. (e) and (f) Fingerprint of partial digestion products of oligonucleotide spot 3 from Fig. 1 together with the deduced sequence. The position of the xylene cyanol FF dye marker (B) is shown.*

There are other examples of heterogeneity at the 5′ ends of viral mRNA's. SV40, a primate virus closely related to polyoma virus, has a heterogeneous family of 5′ leader sequences spliced to the bodies of its late mRNA's (10, 20, 27, 37), although approximately 65% of the capped 5′ ends are AUU sequences (12; Y. Groner, personal communication) derived primarily from one major leader sequence (10, 13, 19). We have compared the SV40 DNA sequence which specifies the major capped end found in cells infected with wild-type virus (11, 13) and found no homology with the polyoma virus DNA sequence shown in Fig. 4. There is a striking homology (Fig. 5A), however, between the capping sequence identified in this paper and a region of the SV40 DNA sequence about 60 nucleotides 5′ to this point (5, 28). The 5′ end of a minor SV40 late leader sequence maps within this homology region (10, 27). Moreover, the SV40 DNA sequence which determines the most common 5′ end is deletable, (13a, 24, 31), and an examination of the capped ends of the mRNA's produced by one such mutant (13a) demonstrated that about 40% of the ends were ACA or AUU. Although these 5′ ends have not been localized precisely on the SV40 DNA sequence, we suggest that certain of them derive from the homology region shown in Fig. 5. The mRNA's synthesized from the early regions of adenovirus 2 also display 5′-terminal heterogeneity (L. T. Chow, T. R. Broker, and J. B. Lewis, submitted for publication). E. Ziff and his collaborators (personal communication) have determined the 5′ ends of mRNA's transcribed from early region II of adenovirus 2 to find capped A and G termini derived from adjacent nucleotides in the DNA sequence. Thus, the micro-heterogeneity in either transcriptional initiation or mRNA processing demonstrated for polyoma virus is common among viruses, but has not as yet been reported for cellular mRNA's, which in all instances thus far described have apparently unique capped termini. When the DNA se-

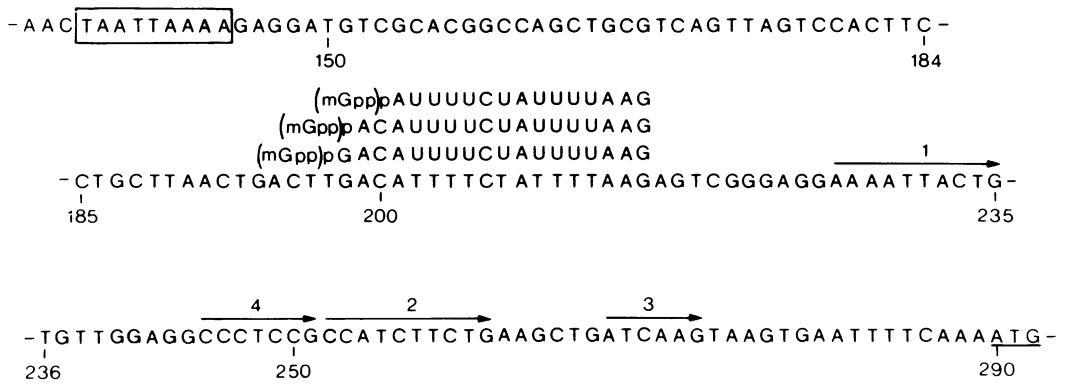


FIG. 4. *Comparison of the RNA sequences of capped T1 oligonucleotides with the DNA sequence from the portion of the late region which determines the leader sequences and the amino terminus of capsid protein VP2. The sequence of the E strand of polyoma virus DNA between 131 and 299 nucleotides from the HpaII-5-HpaII-3 junction as determined by Arrand et al. (1) is shown. The sequences of the three capped T1 oligonucleotides determined in this paper are aligned above the DNA sequence. The DNA sequences which correspond to the four large T1 oligonucleotides which comprise the leader sequence of the late mRNA's are indicated by the numbers 1 through 4 (22). It should be noted that leader oligonucleotide 3 has the sequence ATCAAAG in the mRNA's (22), most probably because it spans the junction between nucleotides 272 and 216, which is formed by the splice producing the leader sequence amplification (17, 22). The ATG initiation codon for capsid protein VP2 at nucleotides 290 to 292 (14) is underlined. The AT-rich sequence (nucleotides 137 to 143) indicated is discussed in the text. The numbers below the DNA sequence refer to the distances (in nucleotides) from the HpaII-3-HpaII-5 junction.*

quences determining the capped ends of various rodent globin messengers (19) are compared with the polyoma virus sequence, however, a high degree of homology is detected within the first 12 transcribed nucleotides (Fig. 5B). The chicken ovalbumin RNA sequence (9) is less homologous, sharing only the first four nucleotides with the polyoma virus sequence. The sequences preceding the cap site are similar among the globin genes (19) but unrelated to the viral sequences.

D. Hogness and M. Goldberg (personal communication) and Gannon et al. (9) have noted that sequences related to the canonical octanucleotide TATAAATA occur some 25 nucleotides preceding the base determining the capped mRNA end of many vertebrate and invertebrate mRNA's and suggested that this sequence may be the eucaryotic equivalent of the Pribnow box (26), which is a common feature of procaryotic promoters. There is a TAATTAAAAG sequence (nucleotides 134 through 143 [Fig. 4]) in the polyoma virus sequence at 54 to 57 nucleotides before the mRNA 5' ends localized in this paper. This sequence is similar to the canonical sequence and very similar to the TATAAAAG found in the putative adenovirus 2 major late promoter region (38). However, the sequence is much farther from the cap sites than in other known examples, and thus its significance is unclear. We are now constructing deletion mutants lacking the "Hogness box" to test the hypothesis that it is involved in the initiation of late transcription.

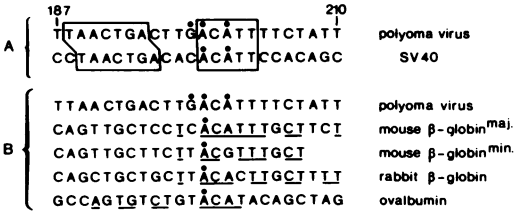


FIG. 5. *Comparison of the DNA sequences in the regions surrounding the cap sites in corresponding mRNA's between polyoma virus DNA and SV40 DNA and between polyoma virus DNA and several cellular genes. (A) The polyoma virus DNA E strand sequence (the same sense as late mRNA) from nucleotides 187 to 210 (1) (Fig. 4) is compared with the sequence of the SV40 DNA E strand between nucleotides 233 and 256 (5) or 168 and 190 (28). The dots indicate the nucleotides corresponding to the three capped ends of polyoma virus late mRNA's and to the predicted capping sites for SV40 DNA. Boxes indicate the regions of identical nucleotide sequences. (B) The polyoma virus DNA sequence (see above) is compared with the sequences of mouse β-globins (19), rabbit β-globin (19), and ovalbumin (9). Dots indicate nucleotides corresponding to capped ends of mRNA's. Underlining indicates nucleotides identical to those found in the polyoma virus DNA sequence.*

We thank John Vournakis for his advice and for providing an initial sample of tobacco acid pyrophosphatase. Jane Walsh provided excellent technical assistance.

## LITERATURE CITED


1. **Arrand, J. R., E. Soeda, J. E. Walsh, N. Smolar, and B. E. Griffin.** 1980. Polyoma virus DNA: sequence from the late region that specifies the leader sequence for late mRNA and codes for VP2, VP3, and the N terminus of VP1. J. Virol. **33**:606–618.
2. **Bajszar, G., O. P. Samarina, and G. P. Georgiev.** 1976. On the nature of 5′-termini in nuclear pre-mRNA of Erlich carcinoma cells. Cell **9**:323–332.
3. **Barrell, B.** 1974. Fractionation and sequence analysis of radioactive nucleotides, p. 751–779. *In* G. L. Cantoni and D. R. Davies (ed.), Procedures in nucleic acid research. Harper & Row, Publishers, New York.
4. **Efstratiadis, A., J. N. Vournakis, H. Doris-Keller, G. Chaconas, D. K. Dougall, and F. C. Kafatos.** 1977. End labeling of enzymatically decapped mRNA. Nucleic Acids Res. **4**:4165–4174.
5. **Fiers, W., R. Contreras, G. Haegeman, R. Rogiers, A. Van de Voorde, H. Van Heuverswyn, J. Van Herreweghe, G. Volckaert, and M. Ysebaert.** 1978. Complete nucleotide sequence of SV40 DNA. Nature (London) **273**:113–120.
6. **Flavell, A. J., S. Legon, A. Cowie, and R. Kamen.** 1979. Multiple 5′-terminal cap structures in late polyoma virus RNA. Cell **16**:357–372.
7. **Ford, J. P., and M.-T. Hsu.** 1978. Transcription pattern of in vivo-labeled late simian virus 40 RNA: equimolar transcription beyond the mRNA 3′ terminus. J. Virol. **28**:795–801.
8. **Furiuchi, Y., S. Muthukrishnan, and A. J. Shatkin.** 1975. 5′-Terminal $m^7G(5')ppp(5')G^mp$ in vivo: identification in reovirus genome RNA. Proc. Natl. Acad. Sci. U.S.A. **72**:742–745.
9. **Gannon, F., K. O'Hare, F. Perrin, J. P. Le Pennec, C. Benoist, M. Cochet, R. Breathnach, A. Royal, A. Garapin, B. Cami, and D. Chambon.** 1979. Organisation and sequences at the 5′ end of a cloned complete ovalbumin gene. Nature (London) **278**:428–434.
10. **Ghosh, P. K., V. B. Reddy, J. Swinscoe, P. Lebowitz, and S. M. Weissman.** 1978. Heterogeneity and 5′-terminal structures of the late RNAs of simian virus 40. J. Mol. Biol. **126**:813–846.
11. **Groner, Y., P. Carmi, and Y. Aloni.** 1977. Capping structures of simian virus 40 19S and 16S mRNAs. Nucleic Acids Res. **4**:3959–5968.
12. **Haegeman, G., and W. Fiers.** 1978. Characterization of the 5′-terminal capped structures of late simian virus 40-specific RNA. J. Virol. **25**:824–830.
13. **Haegeman, G., and W. Fiers.** 1978. Localization of the 5′ terminus of late SV40 mRNA. Nucleic Acids Res. **5**: 2359–2371.

13a. **Haegeman, G., H. van Heuverswyn, D. Gheysen, and W. Fiers.** 1979. Heterogeneity of the 5′ terminus of late mRNA induced by a viable simian virus 40 deletion mutant. J. Virol. **31**:484–493.

14. **Hewick, R. M., A. Mellor, A. E. Smith, and M. D. Waterfield.** 1980. Partial amino-terminal sequences of the polyoma nonhistone proteins VP1, VP2, and VP3 synthesized in vitro. J. Virol. **33**:631–636.
15. **Horowitz, M., S. Bratosin, and Y. Aloni.** 1978. Polyoma infected cells contain at least three spliced late RNAs. Nucleic Acids Res. **5**:4663–4676.
16. **Hunter, T., and W. Gibson.** 1978. Characterization of the mRNA's for the polyoma virus capsid proteins VP1, VP2, and VP3. J. Virol. **28**:240–253.
17. **Kamen, R., J. Favaloro, and J. Parker.** 1980. Topography of the three late mRNA's of polyoma virus which encode the virion proteins. J. Virol. **33**:637–651.
18. **Kamen, R., J. Sedat, and E. Ziff.** 1976. Orientation of the complementary strands of polyoma virus DNA with respect to the DNA physical map. J. Virol. **17**:212–218.
19. **Konkel, D. A., S. M. Tilghman, and P. Leder.** 1978. The sequence of the chromosomal mouse $\beta$-globin major gene: homologies in capping, splicing and poly(A) sites. Cell **15**:1125–1132.
20. **Lai, C.-J., R. Dhar, and G. Khoury.** 1978. Mapping the spliced and unspliced late lytic SV40 RNA's. Cell **14**: 971–982.
21. **Laub, O., S. Brabsin, H. Horowitz, and Y. Aloni.** 1979. The initiation of transcription of SV40 DNA at late times after infection. Virology **92**:310–323.
22. **Legon, S., A. Flavell, A. Cowie, and R. Kamen.** 1979. Amplification of the leader sequence of "late" polyoma virus mRNAs. Cell **16**:373–388.
23. **Manor, H., M. Wu, N. Baran, and N. Davidson.** 1979. Electron microscopic mapping of RNA transcribed from the late region of polyoma virus DNA. J. Virol. **32**:293–303.
24. **Mertz, J., and P. Berg.** 1974. Viable deletion mutants of SV40: selective isolation by means of a restriction endonuclease from *H. parainfluenzae*. Proc. Natl. Acad. Sci. U.S.A. **71**:4879–4883.
25. **Nishimura, S.** 1972. Minor components in transfer RNA. The characterisation, location and function. Prog. Nucleic Acid Res. Mol. Biol. **12**:49–85.
26. **Pribnow, D.** 1975. Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. Proc. Natl. Acad. Sci. U.S.A. **72**:784–788.
27. **Reddy, V. B., P. K. Ghosh, P. Lebowitz, and S. M. Weissman** 1978. Gaps and duplicated sequences in the leaders of SV40 16S RNA. Nucleic Acids Res. **5**:4195–4214.
28. **Reddy, V. B., B. Thimmappaya, R. Dhar, K. N. Subramanian, B. S. Zain, P. K. Ghosh, M. L. Celma, and S. M. Weissman.** 1978. The genome of simian virus 40. Science **200**:494–502.
29. **Schibler, U., and R. P. Perry.** 1976. Characterization of the 5′-termini of mRNA in mouse L cells: implications for processing and cap formation. Cell **9**:121–130.
30. **Schminke, C. D., K. Herrmann, and P. Hausen.** 1976. Size of primary transcripts in Ehrlich ascites cells as measured by tetranucleotide determination. Proc. Natl. Acad. Sci. U.S.A. **73**:1994–1998.
31. **Shenk, T., J. Carbon, and P. Berg.** 1976. Construction and analysis of viable deletion mutants of simian virus 40. J. Virol. **18**:664–671.
32. **Siddell, S. G., and A. E. Smith.** 1978. Polyoma virus has three late mRNA's: one for each virion protein. J. Virol. **27**:427–431.
33. **Silberklang, M., A. M. Gillum, and V. L. RajBhandary.** 1977. The use of nuclease P1 in sequence analysis of end group labeled RNA. Nucleic Acids Res. **4**:4091–4108.

33a. **Soeda, E., J. R. Arrand, and B. E. Griffin.** 1979. Polyoma virus DNA. The early region and its T-antigen. Nucleic Acids Res. **7**:839–857.

34. **Soeda, E., J. R. Arrand, and B. E. Griffin.** 1980. Polyoma virus DNA: complete nucleotide sequence of the gene which codes for polyoma virus capsid protein VP1 and overlaps the VP2/VP3 genes. J. Virol. **33**:619–630.
35. **Soeda, E., J. R. Arrand, N. Smolar, and B. E. Griffin.** 1979. Sequence from early region of polyoma virus DNA containing viral replication origin and encoding small, middle and (part of) large T antigens. Cell **17**:357–370.
36. **Steinschneider, A. and H. Fraenkel-Conrat.** 1966. Studies of nucleotide sequences in tobacco mosaic virus ribonucleic acid. IV. Use of aniline in stepwise degradation.
37. **Villarreal, L. P., R. T. White, and P. Berg.** 1979. Mutational alterations within the simian virus 40 leader segment generate altered 16S and 19S mRNA's. J. Virol. **29**:209–219.
38. **Ziff, E. B., and R. M. Evans.** 1978. Coincidence of the promoter and capped 5′-terminus of RNA from the adenovirus 2 major late transcription unit. Cell **15**: 1463–1475.