JOURNAL OF VIROLOGY, Oct. 1982, p. 175–188
0022-538X/82/100175-14$02.00/0

Vol. 44, No. 1

Polyoma Virus Early and Late mRNAs in Productively Infected Mouse 3T6 Cells

WILLIAM C. HEISER* AND WALTER ECKHART
The Salk Institute, San Diego, California 92138

Received 5 April 1982/Accepted 22 June 1982

We mapped polyoma virus-specific mRNAs isolated from productively infected mouse 3T6 cells on the viral genome by analyzing nuclease S1-resistant RNA-DNA hybrids. The polyoma early mRNAs, which code for the three T antigens, have several 5′ ends near 73 map units (m.u.). During the late phase of infection an additional 5′ end is found near 71 m.u. All of the major early mRNAs have common 3′ ends at 26.01 m.u. There is a minor species of early mRNA with a 3′ end at 99.05 m.u. There are two proximal and two distal splice junctions in the early region which are used to generate three different spliced early mRNAs. There are three late mRNAs encoding the three virion proteins, VP1, VP2, and VP3. The late mRNAs have common 3′ ends at 25.34 m.u. The late mRNAs have heterogeneous 5′ leader sequences derived from the region between 65.53 and 68.42 m.u. The leader sequences are joined to the bodies of the messages coding for VP2, VP3, and VP1 at 66.59, 59.62, and 48.57 m.u., respectively. These results confirm and extend previous analyses of the fine structure of polyoma mRNAs.

Productive infection of mouse cells by polyoma virus provides a favorable system for studying RNA synthesis and processing. The nucleotide sequence of polyoma DNA has been determined (10, 11, 17, 45). The regions of the polyoma genome that code for virus-specific proteins have been analyzed extensively by genetic mapping of the DNA and peptide mapping of the proteins (15, 47). Polyoma mRNA can be translated in vitro (27, 28, 44) to produce the virus-encoded proteins.

The polyoma genome is divided into early and late regions. Early proteins, which are synthesized before viral DNA replication, are translated from mRNAs that map in the region between 72 and 25 map units (m.u.) clockwise on the physical map of polyoma DNA (22, 32). The late virion proteins (VP1, VP2, and VP3), which are synthesized after viral DNA replication, are translated from three separate mRNAs that map between 67 and 25 m.u. counterclockwise on the map (27, 32, 44).

Late polyoma mRNAs have been mapped extensively by determining S1 nuclease resistance of RNA-DNA hybrids (30) and by using electron microscopy (26, 37, 52). A short, untranslated leader sequence has been shown to be spliced to the body of the messages coding for VP3 and VP1 by sequencing cloned, reverse transcripts (48), and RNase fingerprinting has shown that these leaders have heterogeneous 5′ ends (15, 35). In this paper we report results in agreement with this. The early messages have been analyzed less extensively because the amounts of these messages are much lower. The early splice junctions have been deduced from the polyoma sequence (17, 44) by using the "GT/AG" rule (43). The 3′ end of the early messages has been mapped to the region near 25.8 m.u. A small fraction of the early messages has been reported to terminate near 99 m.u. (31).

In this report we describe an analysis of the early and late polyoma virus mRNAs isolated from productively infected mouse 3T6 cells in which we used the technique of heteroduplex formation followed by S1 nuclease digestion developed by Berk and Sharp (3). We describe the locations of the 5′ and 3′ ends and the positions of the splice junctions for the viral messages for the three early virus-specific proteins (T antigens) and the three late virion proteins (VP1, VP2, and VP3). Our results confirm and extend previous analyses of the fine structure of polyoma mRNAs (9, 14, 15, 31, 35, 48).

MATERIALS AND METHODS

Virus stocks and cell cultures. A clone of Swiss mouse 3T6 cells (clone 7-3) grown in Dulbecco modified Eagle medium supplemented with 5% calf serum was used for all experiments. Wild-type polyoma virus and TS25E, a *ts*A mutant, both of which were derived from large plaque virus, have been described previously (12).

Isolation of cytoplasmic RNA. Cytoplasmic RNA was isolated from infected cell cultures essentially as

described by Gibson et al. (20). Early RNA was harvested 20 to 24 h after infection from cells infected by TS25E and incubated at the nonpermissive temperature (39°C) or from cells infected by wild-type polyoma virus and grown at 37°C in the presence of hydroxyurea to block DNA synthesis.

Larger amounts of early cytoplasmic virus-specific RNA were obtained by incubating TS25E-infected 3T6 cells for 36 h at 32°C, followed by incubation for 6 h at 39°C before harvest (7).

Late RNA was harvested 40 to 45 h after infection from cells infected with wild-type polyoma virus, incubated at 37°C for the first 24 h, and incubated subsequently at 33°C.

Polyadenylic acid [poly(A)]-containing RNA was selected by oligodeoxythymidylic acid cellulose chromatography (20), precipitated with ethanol, suspended in water at a concentration of 1 mg/ml, and stored at −20°C.

Preparation of ^{32}P-labeled polyoma DNA restriction enzyme fragments. Polyoma DNA was extracted from infected cells by the method of Hirt (25). The Hirt supernatant was extracted once with phenol and twice with chloroform-isoamyl alcohol (24:1). Superhelical form I DNA was purified by CsCl-ethidium bromide centrifugation. After extraction of the ethidium bromide with CsCl-saturated isopropanol, the DNA was precipitated twice with ethanol, suspended in 10 mM Tris (pH 8.0)–1 mM EDTA, and stored at −20°C.

Removal of terminal phosphates with bacterial alkaline phosphatase and labeling of the DNA 5′ ends with T_4 polynucleotide kinase (Bethesda Research Laboratories, Inc., Rockville, Md.) and [^{32}P]ATP (3,000 Ci/mmol; ICN, Irvine, Calif.) were performed essentially as described by Maxam and Gilbert (38).

End-labeled DNAs were prepared for strand separation by boiling for 3 min in 75% dimethyl sulfoxide and then immediately loaded onto a pre-electrophoresed 10% acrylamide–0.17% bisacrylamide gel in a 0.55× Tris-borate buffer. After electrophoresis for 1 min at 1,900 V, electrophoresis was continued for 15 to 25 h at 300 to 400 V.

Polyoma virus DNA was labeled in vivo with $^{32}P_i$ by the method of Lai et al. (34). Briefly, 3T6 cells were seeded into 100-mm petri dishes, infected with wild-type polyoma virus at a multiplicity of infection of about 20 when the cultures were about 75% confluent, and incubated for 16 to 18 h at 37°C. The cells were washed with phosphate-free Dulbecco modified Eagle medium and incubated in 4 ml of Dulbecco modified Eagle medium supplemented with 3% dialyzed calf serum, 10 μM P_i, and 0.25 mCi of $^{32}P_i$ (ICN) per ml. After a labeling period of 24 to 26 h, viral DNA was isolated as described above. Form II polyoma DNA (specific activity, 2×10^6 to 5×10^6 Cerenkov counts per min per μg) was isolated by CsCl-ethidium bromide equilibrium sedimentation after phenol extraction.

Restriction endonucleases were obtained either from Bethesda Research Laboratories or from New England BioLabs, Inc., Beverly, Mass. The conditions used for digesting the polyoma DNA were those suggested by the suppliers. DNA fragments were separated by electrophoresis through 0.7% agarose or 5% acrylamide (0.25% bisacrylamide, 25% glycerol); in both cases, Tris-borate buffer (89 mM Tris base, 89 mM boric acid, 2.5 mM EDTA, pH 8.3) was used. DNA fragments were electroeluted as described by Friedmann et al. (18), precipitated with ethanol, suspended in buffer containing 10 mM Tris (pH 8.0) and 1 mM EDTA, and stored at −20°C.

Hybridization. The hybridization procedure used was similar to that described by Berk and Sharp (3). Hybridizations were performed as follows. A 0.25- to 1-μg equivalent of a ^{32}P-labeled polyoma DNA restriction fragment and 1 to 25 μg of poly(A)-containing RNA were mixed in 1.5-ml Eppendorf tubes, ethanol precipitated, and dried under a vacuum. Samples containing double-stranded DNA were suspended in 30 to 50 μl of hybridization buffer containing 0.4 M NaCl, 1 mM EDTA, 40 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)] (pH 6.4), and 80% formamide (deionized with Amberlite MB-3 immediately before use). (The melting temperature for polyoma DNA in this buffer was 47°C, as determined by susceptibility to digestion by endonuclease S1.) The solution was then incubated at 68 to 70°C to denature the DNA, quickly transferred to a circulating water bath at 51°C, and hybridized for 20 to 60 min. Strand-separated DNAs were hybridized in 30 μl of hybridization buffer containing 0.4 M NaCl, 1 mM EDTA, 40 mM PIPES (pH 6.4), and 50% deionized formamide. The solution was heated to 72 to 74°C for 5 min and then hybridized in a circulating water bath at 42°C for 1 to 3 h. Hybridization was stopped by adding 10 volumes of S1 nuclease buffer (0.25 M potassium acetate, pH 4.5, 1 mM $ZnSO_4$, 5% glycerol) or mung bean nuclease buffer (30 mM sodium acetate, pH 4.6, 50 mM NaCl, 1 mM $ZnCl_2$, 5% glycerol) at 0°C. Then 5 to 15 μl of endonuclease S1 per 1,000 μl of volume was added and the samples were digested at 37 to 45°C for 20 min, or 1 U of mung bean nuclease (P-L Biochemicals) per μg of nucleic acid was added and the samples were digested for 60 min at 37°C. We found that mung bean nuclease gave more precise mapping of the 3′ ends of splice junctions than S1 nuclease did. For this reason, mung bean nuclease was also used to digest the unhybridized nucleic acids when the 5′ ends of the early and late messages were mapped. The reactions were stopped by adding 2 volumes of ethanol. After 1 h at −70°C or overnight at −20°C, the samples were pelleted by centrifugation in a Brinkmann microfuge, and the pellets were suspended in 5 to 10 μl of water. The samples were analyzed in the following ways. (i) For untreated samples 10 μl of electrophoresis buffer (10 mM phosphate, pH 7.0, 2 mM EDTA, 75 μg of bromophenol blue per ml, 50% glycerol) was added, and the samples were loaded onto either a 0.7% agarose gel or a 5% acrylamide gel. (ii) For alkali-denatured samples an equal volume of 120 mM NaOH–4 mM EDTA was added, and the samples were incubated for 15 min at room temperature. Then 10 μl of sample buffer (2 mM EDTA, 75 μg of bromophenol blue per ml, 50% glycerol) was added, and the samples were immediately loaded onto a 1% alkaline agarose gel (39). (iii) For glyoxylated samples the samples were glyoxylated by the method of McMaster and Carmichael (40), precipitated with ethanol, suspended in 15 μl of electrophoresis buffer, and loaded onto a 5% acrylamide gel. (iv) End-labeled samples were precipitated with ethanol, washed twice with 70% ethanol, suspended in 3 μl of sequencing electrophoresis solution (97% formamide, 10 mM NaOH, 1 mM EDTA, 0.05% xylene cyanol, 0.025% bromophenol blue),

boiled for 1 to 2 min, and loaded onto an acrylamide–8 M urea sequencing gel.

The lengths of fragments protected by hybridization were determined by their relative mobilities, using restriction digests of polyoma DNA or nucleotide sequencing reactions of polyoma DNA fragments as standards.

RESULTS

Heteroduplex analysis. Cytoplasmic poly(A)-containing RNA was isolated from polyoma-infected 3T6 cells. The polyoma early and late mRNAs were hybridized with various radiolabeled polyoma DNA restriction enzyme fragments. The DNA-RNA hybrids were formed under conditions favoring DNA-RNA hybridization rather than DNA-DNA reannealing (5). The hybrids were treated with single-strand-specific nuclease to remove unhybridized sequences and were analyzed by neutral and alkaline agarose gel electrophoresis. A more precise analysis of the sizes of protected DNA fragments was made on neutral or 8 M urea–polyacrylamide gels.

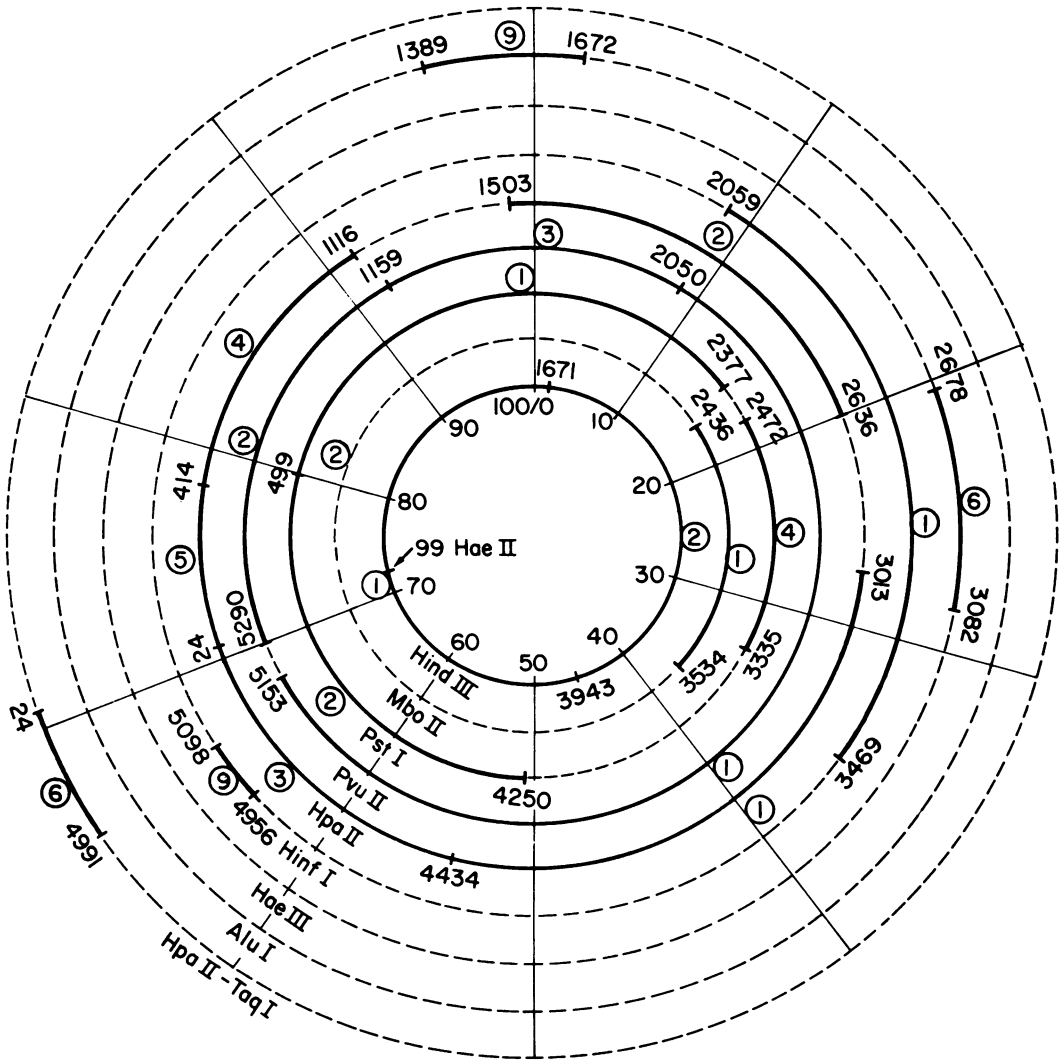

FIG. 1. Recognition sites of restriction enzyme fragments used in mapping the early and late RNAs. The polyoma genome is represented in its circular form. The numbers inside the center circle divide the genome into map units; 0 m.u. is defined as the first nucleotide in the *Eco*RI recognition sequence at nucleotide 1,575. The numbers on each of the concentric circles designate the first nucleotides in the restriction enzyme recognition sequences (not necessarily the nucleotides at which the enzymes cleave the DNA); the sequence numbers are those determined by Deininger et al. (11). The numbers in circles are the numbers of the restriction enzyme fragments, in order of decreasing size.

The size of an RNA transcript can be determined by measuring the size of the viral DNA protected from S1 digestion by hybridization to the RNA. Electrophoresis under neutral conditions shows the total size of the RNA-DNA hybrid. Electrophoresis under denaturing conditions shows the size of each portion of a spliced RNA transcript from which intervening sequences have been removed.

Early transcripts. Figure 1 summarizes the locations of most of the restriction fragments used to map the RNA transcripts on the polyoma genome. (In Fig. 1 the restriction fragment cutting sites are designated by the first nucleotide in the restriction enzyme recognition sequence. However, in the text the cutting sites are designated by the cleavage site on the DNA strand which is complementary to the RNA. We used the sequencing and numbering systems of the polyoma genome described by Friedmann et al. [11, 17] for the large-plaque polyoma strain used in these experiments.)

Polyoma DNA cleaved with restriction enzyme *Hae*II to generate a single full-length linear molecule yielded two protected fragments when it was hybridized to poly(A)-containing RNA isolated early during lytic infection. The lengths of the protected fragments were approximately 2,700 and 2,375 nucleotides (Fig. 2, track 3). Previous in vitro translation experiments have shown that the mRNA that codes for the large T antigen is smaller than the mRNAs that code for the small and medium T antigens (28). Therefore, the smaller fragment was most likely produced by hybridization to the message which codes for the large T antigen, whereas the larger fragment was produced by hybridization to the messages which code for the small and medium T antigens.

To determine whether the early RNAs are transcribed from a contiguous region of the viral genome, we analyzed the hybrids by alkaline agarose gel electrophoresis. As shown in Fig. 2, track 6, there were no fragments corresponding in size to those protected under neutral conditions. Instead, the hybrids contained DNA fragments approximately 2,150 and 575 nucleotides long. This suggests that the 2,700-nucleotide message is composed of noncontiguous transcripts of about 2,150 nucleotides and 575 nucleotides. Because fragments smaller than 400 nucleotides could not be resolved on our gels, we could not detect a smaller fragment (approximately 250 nucleotides long [see below]) that is also protected by early mRNA.

The positions of the 3′ ends of the early mRNAs were determined by using polyoma DNA restriction fragments that mapped in the region around 25 m.u., where early transcription terminates (32). When early poly(A)-containing

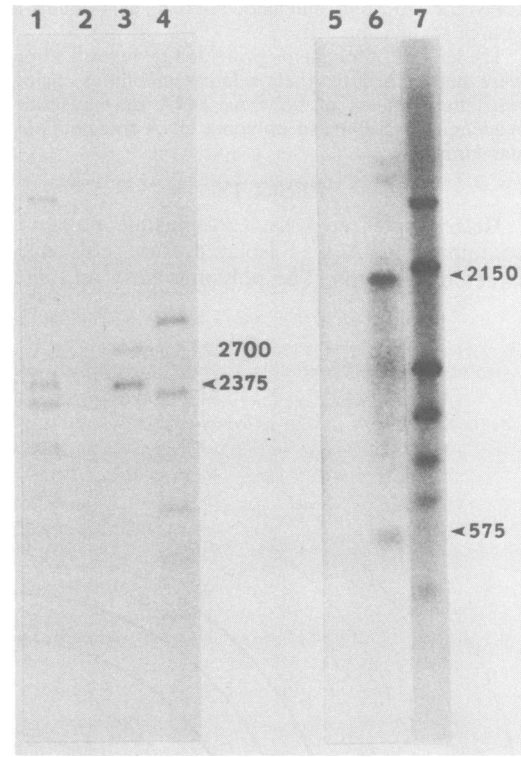

FIG. 2. Nuclease S1 mapping of the early polyoma RNAs. ^{32}P-labeled *Hae*II-digested polyoma DNA (0.5 μg) was hybridized without RNA or with 25 μg of early cytoplasmic poly(A)-containing RNA as described in the text. One-half of each sample was analyzed by electrophoresis on a 0.7% neutral agarose gel (tracks 2 and 3), and the other half was analyzed by electrophoresis on a 1% alkaline agarose gel (tracks 5 and 6). The sizes of the S1-resistant bands were determined by comparison with the migration distances of restriction fragments of the following ^{32}P-labeled polyoma DNA markers: track 1, *Hae*II (5,284 nucleotides), *Hha*I (2,402 and 2,220 nucleotides), and *Pst*I (1,878, 1,544, 915, and 863 nucleotides); tracks 4 and 7, *Hin*dIII (3,023 and 2,272 nucleotides) and *Hpa*II (1,421, 1,133, 885, 702, 390, and 377 nucleotides).

RNA was hybridized with the *Hin*dIII-2 polyoma DNA fragment (nucleotides 1,676 through 3,947) and analyzed by neutral agarose gel electrophoresis, a fragment 1,270 nucleotides long was protected (Fig. 3, track 2). Because the early RNAs include the sequences between 1 and 25 m.u., the end of the *Hin*dIII fragment at nucleotide 1,676 was protected. Therefore, the 3′ end of the early RNA is near nucleotide 2,950. Four other fragments which overlap in this region (*Pvu*II-1, *Hin*fI-1, *Pst*I-4, and *Hae*III-6) were used to map the 3′ end. The fragments protected by the RNAs are shown in Fig. 3, and the results are summarized in Table 1. These

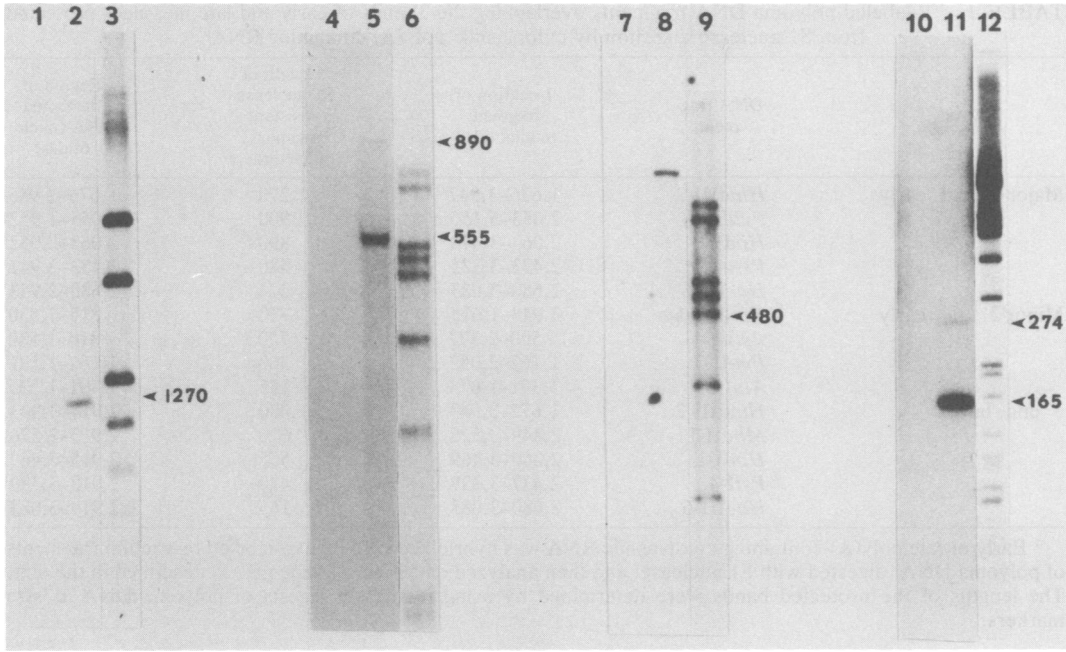

FIG. 3. Nuclease S1 mapping of the 3′ ends of early and late polyoma RNAs. The restriction fragments of ^{32}P-labeled polyoma DNA which overlap in the region of 25 m.u. on the polyoma genome were hybridized without RNA or with 5 to 20 μg of early or late cytoplasmic poly(A)-containing RNA or yeast tRNA, as described in the text. The hybrids were analyzed by electrophoresis on a 0.7% agarose gel (tracks 1 through 3) or on 5% acrylamide gels (tracks 4 through 12). The samples in tracks 4 through 6 were glyoxylated. Tracks 1 and 2, S1-resistant hybrids of *Hin*dIII-2 polyoma DNA hybridized without RNA (track 1) or with early RNA (track 2); tracks 4 and 5, S1-resistant products of *Hin*fI-1 polyoma DNA hybridized with tRNA (track 4) or with late RNA (track 5); tracks 7 and 8, S1-resistant hybrids of *Pst*I-4 polyoma DNA hybridized with tRNA (track 7) or with early RNA (track 8); tracks 10 and 11, S1-resistant hybrids of *Hae*III-6 polyoma DNA hybridized with tRNA (track 10) or with late RNA (track 11); track 3, marker *Hin*dIII and *Hpa*II restriction fragments of polyoma DNA; tracks 6, 9, and 12, marker *Hae*III restriction fragments of polyoma DNA. The lengths of the fragments in tracks 6, 9, and 12 (in nucleotides) are as follows: fragment 1, 718; fragment 2, 673; fragment 3, 540; fragment 4, 514; fragment 5, 488; fragment 6, 404; fragment 7, 314; fragment 8, 209; fragment 9, 199; fragment 10, 170; fragment 11, 134; fragment 12, 117.

results localized the 3′ terminus for the early RNAs near nucleotide 2,952, 10 nucleotides beyond the putative polyadenylation signal (41), AATAAA, which ends at nucleotide 2,942.

The positions of the 5′ ends of the early mRNAs were identified by hybridization with DNA fragments which overlapped the origin of DNA replication near 70 m.u. Early poly(A)-containing RNA hybridized with the *Hpa*II-5 fragment of polyoma DNA (nucleotides 27 through 416) protected two S1-resistant fragments with lengths of 250 and 260 nucleotides (Fig. 4, track 2). In addition to protecting these two fragments, late poly(A)-containing RNA protected a third fragment, which was about 370 nucleotides long (Fig. 4, track 3). Two other fragments were used to map these ends more precisely, the 176-nucleotide *Hpa*II-*Dde*I-9 polyoma DNA fragment (nucleotides 27 through 203) and the 348-nucleotide *Bcl*I-*Nar*I-2 polyoma DNA fragment (nucleotides 5,051 through 102). After strand separation the strand complementary to the early message was identified by nucleotide sequence analysis and was hybridized to poly(A)-containing RNA isolated from polyoma-infected 3T6 cells late in infection. The results of hybridization with *Hpa*II-*Dde*I-9 polyoma DNA are shown in Fig. 4, track 5. These results place two major 5′ initiation sites at nucleotides 167 through 170 (fragments 34 to 37 nucleotides long) and nucleotides 156 and 157 (fragments 47 and 48 nucleotides long). There is also a minor initiation site at nucleotides 138 and 139 (fragments 65 and 66 nucleotides long). The results of hybridization with the *Bcl*I-*Nar*I-2 polyoma DNA fragment are shown in Fig. 4, track 7. These results identify another RNA initiation site, which is used for transcription of early messages late in infection. The RNAs have heterogenous ends located at nucleotides 46 through 49 (fragments 54 to 57 nucleotides long).

The locations of the splice junctions in the

TABLE 1. ^{32}P-labeled polyoma DNA fragments overlapping the 3′ ends of early and late messages protected from S1 nuclease digestion by cytoplasmic poly(A)-containing RNA[a]

RNAs	DNA fragment	Location of fragment (nucleotides)	Length of S1 nuclease-resistant fragments (nucleotides)	Region of protected DNA (nucleotides)
Major 3′ end, early	*Hin*dIII-2	1,676–3,947	1,270	1,676–2,945
	*Pvu*II-1	2,053–5,155	900	2,053–2,952
	*Hin*fI-1	2,063–3,472	890	2,063–2,952
	*Pst*I-4	2,473–3,335	480	2,473–2,952
	*Hae*III-6	2,680–3,083	274	2,680–2,953
Minor 3′ end, early	*Hin*dIII-1	3,948–1,675	720	810–1,530
	*Pst*I-1	500–2,377	720	810–1,530
	*Pvu*II-3	1,162–2,052	360	1,162–1,521
	*Alu*I-9	1,391–1,673	135	1,391–1,525
3′ end, late	*Hin*dIII-2	1,672–3,943	1,030	2,914–3,943
	*Mbo*II-1	2,449–3,526	610	2,915–3,526
	*Hin*fI-1	2,060–3,469	555	2,915–3,469
	*Pst*I-4	2,477–3,339	420	2,918–3,339
	*Hae*III-6	2,680–3,083	165	2,919–3,083

[a] Early or late poly(A)-containing cytoplasmic RNA was hybridized with in vivo-labeled restriction fragments of polyoma DNA, digested with S1 nuclease, and then analyzed on 5% acrylamide gels as described in the text. The lengths of the protected bands were determined by using restriction digests of polyoma DNA as size markers.

early messages were determined by analyzing the segments protected after hybridizing early RNA with the following polyoma DNA fragments: *Hin*dIII-1 (nucleotides 3,947 through 1,675), *Pvu*II-2 (nucleotides 5,293 through 1,161), *Pst*I-1 (nucleotides 500 through 2,377), *Pst*I-2 (nucleotides 4,251 through 499), *Hpa*II-4 (nucleotides 417 through 1,118), and *Dde*I-7 (nucleotides 570 through 844). The results (Fig. 5 and Table 2) indicate that there are proximal splice junctions near nucleotides 426 and 761 and distal splice junctions near nucleotides 810 and 824. An examination of the polyoma DNA sequence in these regions shows consensus sequences (43) which place the proximal splice junctions between nucleotides 423 and 424 and between nucleotides 761 and 762 and the distal splice junctions between nucleotides 809 and 810 and between nucleotides 823 and 824.

In addition to the 3′ end at nucleotide 2,952 in the early mRNA, there is also a second, minor 3′ end near nucleotide 1,525. The existence of a 3′ end in this position was initially suggested by the presence of a 720-nucleotide band protected when *Pst*I-1 and *Hin*dIII-1 fragments were hybridized to early RNA (Tables 1 and 2). The location of this 3′ end was mapped more precisely by using polyoma DNA fragments *Pvu*II-3 (nucleotides 1,162 through 2,052) and *Alu*I-9 (nucleotides 1,391 through 1,673), which gave protected fragments 360 and 135 nucleotides long, respectively (Fig. 6, tracks 2 and 5), placing the 3′ terminus at nucleotides 1,521 to 1,525. This 3′ end is 26 to 29 nucleotides beyond a potential polyadenylation signal, AATAAA, which ends at nucleotide 1,496. A 3′ end in this region has also been observed by Kamen et al. (31). Since this message is polyadenylated beyond the termination codons for the small T antigen (at nucleotides 812 through 824) and the medium T antigen (at nucleotides 1,513 through 1,515), it could code for the small T antigen, the medium T antigen, or an as-yet-unidentified protein.

Late transcripts. The same approach employed in mapping the early mRNAs was used to map the late mRNAs. The number of late transcripts was determined by hybridizing in vivo ^{32}P-labeled polyoma DNA linearized with *Hae*II (cut at 73.2 m.u.) or *Eco*RI (cut at 0 m.u.) with poly(A)-containing RNA isolated late in infection. After digestion with S1 nuclease and electrophoresis on neutral or alkaline agarose gels, three fragments, with lengths of approximately 2,200, 1,900, and 1,250 nucleotides, were present (data not shown). Using in vitro translation, Siddell and Smith (44) and Hunter and Gibson (27) have shown that mRNAs of these relative sizes code for VP2, VP3, and VP1, respectively.

The location of the 3′ end of these mRNAs was determined by hybridizing cytoplasmic poly(A)-containing RNA isolated from 3T6 cells late in infection with polyoma DNA fragment *Hin*fI-1 (nucleotides 2,060 through 3,469). In addition to protecting a fragment corresponding in length to the 3′ end of the early message (see above), the late mRNAs protected a 555-nucleotide fragment (Fig. 3, track 5). This placed the 3′

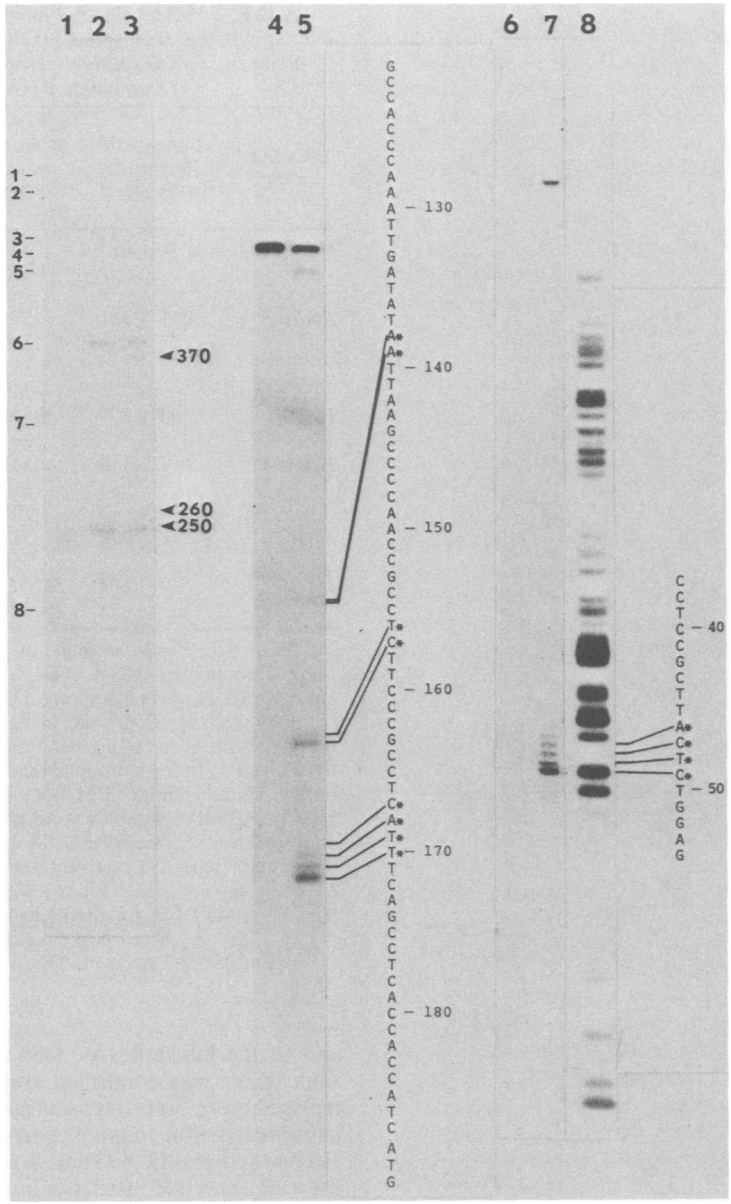

FIG. 4. Nuclease S1 mapping of the 5′ ends of early polyoma RNA. End-labeled polyoma DNA fragments hybridized with poly(A)-containing cytoplasmic RNA isolated from polyoma-infected 3T6 cells and digested with S1 or mung bean nuclease were electrophoresed on 8M urea–acrylamide Tris-borate gels. Tracks 1 through 3, Autoradiogram of S1 digestion products of *Hpa*II-5 polyoma DNA hybridized to tRNA (track 1) or to early (track 2) or late (track 3) RNA and analyzed on a 3.6% acrylamide gel (the numbers on the left indicate the positions of *Hae*III polyoma DNA marker fragments); tracks 4 and 5, autoradiogram of mung bean nuclease digestion products of the strand-separated *Hpa*II-*Dde*I-9 fragment (nucleotides 27 through 203) of polyoma DNA hybridized with tRNA (track 4) or with RNA isolated from *ts*25-infected 3T6 cells shifted to the nonpermissive temperature late in infection (track 5) and electrophoresed on an 8% acrylamide–8M urea gel (the polyoma DNA sequence of the early strand is correlated with the protected fragments); tracks 6 through 8, autoradiogram of mung bean nuclease digestion products of the strand-separated *Bcl*I-*Nar*I-2 DNA fragment (nucleotides 5,051 through 102) of polyoma DNA hybridized with tRNA (track 6) or with RNA isolated from *ts*25-infected 3T6 cells shifted to the nonpermissive temperature late in infection (track 7) and fractionated on an 8% acrylamide sequencing gel. Track 8 shows the G-sequence reaction of the *Bcl*I-*Nar*I-2 fragment complementary to the fragment used in hybridization. The polyoma DNA sequence of the early strand is correlated with the protected fragments.

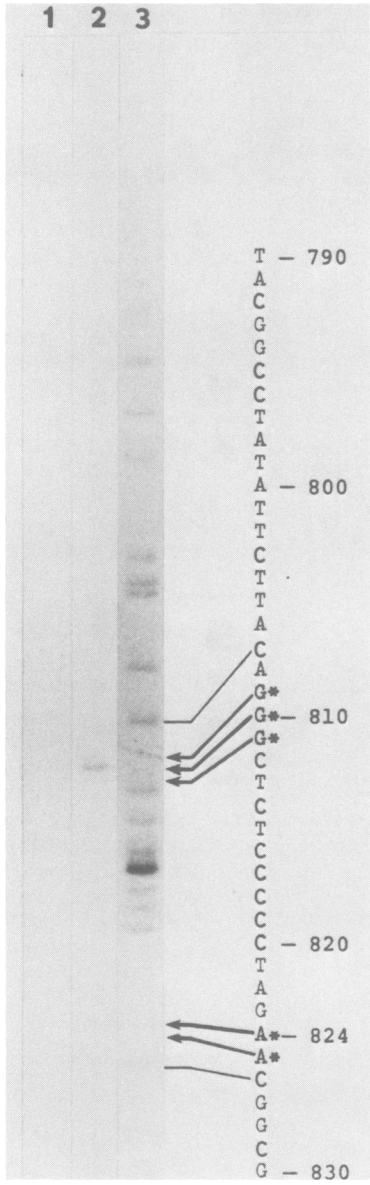

FIG. 5. Mapping the distal splice junctions of early polyoma RNAs by using the *Dde*I-7 fragment of polyoma DNA. The hybridization mixtures contained 10 μg of tRNA (track 1) or early poly(A)-containing cytoplasmic RNA (track 2) and an amount of 5′-end-labeled and strand-separated restriction fragment equivalent to 0.5 μg of full-length polyoma DNA. Hybrids were digested with mung bean nuclease and then separated on an 8% acrylamide–8 M urea sequencing gel. Track 3 shows the G-sequence reaction of the strand of *Dde*I-7 used for hybridization. The sequence at the distal splice junction is aligned with the sequence ladder and with the messenger-protected bands.

TABLE 2. ^{32}P-labeled polyoma DNA fragments overlapping the early splice junctions protected from S1 nuclease digestion by cytoplasmic poly(A)-containing RNA[a]

DNA fragment	Location of fragment (nucleotides)	Length(s) of S1 nuclease-resistant fragments (nucleotides)	Region of protected DNA (nucleotides)
*Hin*dIII-1[b]	3,948–1,675	865	810–1,675
		595	165–1,760
		260	165– 425
*Pvu*II-2[b]	5,293–1,161	595	165– 760
		350	810–1,161
		335	825–1,161
		260	165– 425
*Pst*I-1[b]	500–2,377	1,560	817–2,377
		260	500– 760
*Hpa*II-4[b]	417–1,118	343	417– 760
		310	808–1,118
		295	823–1,118
*Pst*I-2[b]	4,251– 499	335	165– 499
		260	165– 425
*Dde*I-7[c]	570– 844	34, 35, 36	809– 844
		20, 21	824– 844

[a] A 5- to 25-μg amount of early cytoplasmic poly(A)-containing RNA was hybridized with an amount of ^{32}P-labeled polyoma DNA restriction fragment equivalent to 0.5 μg of full-length DNA and digested with S1 or mung bean nuclease. The *Dde*I-7 hybridized samples were analyzed by electrophoresis on an 8% acrylamide–8 M urea sequencing gel (see Fig. 5). All other samples were glyoxylated and then analyzed by electrophoresis on 5% acrylamide gels. The lengths of the S1-resistant bands were determined by using glyoxylated ^{32}P-labeled *Hpa*II, *Hae*III, and *Alu*I polyoma DNA fragments as size markers.

[b] In vivo ^{32}P-labeled polyoma DNA.

[c] 5′-End-labeled polyoma DNA.

end of the late mRNAs near nucleotide 2,914. This result was confirmed and the 3′ end was mapped more precisely near nucleotide 2,917 by using restriction fragments *Hin*dIII-2, *Mbo*II-1, *Pst*I-4, and *Hae*III-6 (Table 1 and Fig. 3). Thus, the early and late messages overlap by about 36 bases. Furthermore, the sequence between nucleotides 2,935 and 2,930 (AATAAA), 13 bases upstream from the 3′ end, is the same as the sequence found near the termination of the early messages.

The 5′ ends of the bodies of the messages coding for VP1, VP2, and VP3 were mapped by hybridizing late RNA with 5′-end-labeled polyoma DNA restriction fragments *Hin*dIII-*Bam*HI-3 (nucleotides 3,944 through 4,657) and *Bam*HI-*Hae*II-2 (nucleotides 4,658 through 103). In each case the messages protected two fragments from digestion with S1 or mung bean nuclease (Fig. 7, tracks 2 and 5). The full-length protected *Hin*dIII-*Bam*HI-3 fragment corre-

sponds to the messages coding for VP2 and VP3. The shorter 204-nucleotide protected fragment corresponds to the 5′ end of the body of the message coding for VP1 and localizes this 5′ end near nucleotide 4,147. The DNA sequence in this region contains a potential distal splicing junction between nucleotides 4,147 and 4,148. The 450- and 75-nucleotide fragments of *Bam*HI-*Hae*II-2 polyoma DNA protected from mung bean nuclease digestion correspond to the 5′ ends of the bodies of the messages coding for VP2 and VP3, respectively. This localizes the 5′ end of the body of VP3 mRNA near nucleotide 4,732 and the 5′ end of the body of VP2 mRNA near nucleotide 5,108. An examination of the DNA sequences in these regions shows consensus sequences (43) which place distal splice junctions between nucleotides 4,732 and 4,733 for VP3 mRNA and between nucleotides 5,101 and 5,102 for VP2 mRNA. These three splice junctions were confirmed by using several other restriction fragments (Table 3).

The number and sizes of the late leaders were initially studied by hybridizing in vivo ^{32}P-labeled polyoma DNA fragments *Hin*dIII-1 and *Hpa*II-3 with late poly(A)-containing cytoplasmic RNA. The results suggested that there are leaders with at least seven different lengths and that one of these leaders, which is about 60 nucleotides long, appears to be present in greater amounts than the other leaders. This is consistent with the size of an amplified portion of the polyoma late leader analyzed by T1 fingerprinting by Flavell et al. (14). The results of hybridizing *Taq*I-*Hpa*II-6 (nucleotides 4,991 through 24) with late RNA confirmed these results (Table 3) and localized the leader sequences between the *Taq*I cutting site at nucleotide 4,991 and the *Hpa*II cutting site at nucleotide 24.

To position the 3′ end of the leader accurately, we hybridized late lytic RNA with polyoma DNA fragment *Hin*f I-9 (nucleotides 4,957 through 5,098), which overlaps the end of the *Taq*I-*Hpa*II-6 DNA fragment between nucleotides 4,991 and 5,098. In addition to protecting the entire *Hin*fI-9 fragment, the RNA protected a 54-nucleotide fragment (Table 3). This positioned the 3′ terminus of the late leader near nucleotide 5,045. An examination of the polyoma DNA sequence in this region shows a prototype splicing junction between nucleotides 5,045 and 5,046.

The positions of the 5′ ends of the late leaders were mapped in the experiment shown in Fig. 8. The strand-separated 349-nucleotide polyoma DNA fragment *Bcl*I-*Nar*I-2 (nucleotides 5,047 through 100) labeled at the 5′ end with ^{32}P (nucleotide 5,047) was hybridized with late poly(A)-containing RNA and analyzed by digestion with mung bean nuclease. The results localized a major 5′ end at nucleotide 5,101; this site is 57 nucleotides upstream from the 3′ terminus of the late leader at nucleotide 5,046. An examination of the DNA sequence in this region shows a consensus sequence for a proximal splice junction between nucleotides 5,101 and 5,102. Together with other results (14, 33, 37, 48), this finding suggests that the late mRNAs contain sequences (between nucleotides 5,101 and 5,046 on the polyoma genome) which are present in greater-than-molar amounts and may be produced by splicing events (1, 4, 30, 35). The results shown in Fig. 8 also localize the 5′ ends of the leaders at nucleotides 5,086 to 5,088, 5,097 or 5,098, 5,103 to 5,105, 5,119, 5,123 to 5,125, 5,128, 5,142 or 5,143, 5,146 or 5,147, 5,151 to 5,153, 5,271 or 5,272, and 5,186 to 5,188. These sites are aligned with the polyoma DNA sequence. There appear to be several preferred initiation sites; these are at nucleotides 5,119, 5,123 to 5,125, 5,146 and 5,147, and 5,151 to 5,153.

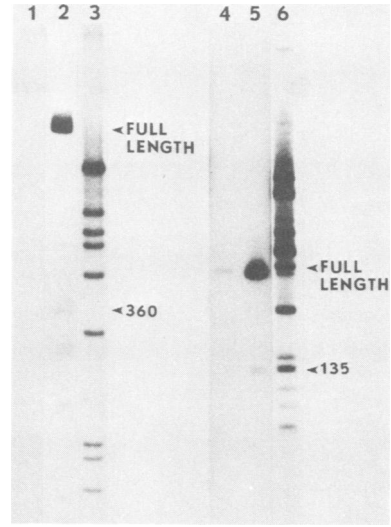

FIG. 6. Nuclease S1 mapping of the 3′ end of the minor polyoma early RNA. Each hybridization mixture containing 25 μg of tRNA (tracks 1 and 4) or early cytoplasmic poly(A)-containing RNA from productively infected 3T6 cells (tracks 2 and 5) was hybridized with ^{32}P-labeled *Pvu*II-3 polyoma DNA (tracks 1 and 2) or with ^{32}P-labeled *Alu*I-9 polyoma DNA (tracks 4 and 5), digested with S1 nuclease, glyoxylated, and then analyzed on 5% acrylamide gels. Track 3 contained glyoxylated *Alu*I restriction fragments of polyoma DNA, and the lengths of these fragments (in nucleotides) are as follows: fragment 1, 534; fragment 2, 518; fragment 3, 423; fragment 4, 390; fragment 5, 338; fragment 6, 333; fragment 7, 331; fragment 8, 303; fragment 9, 283. Track 6 contained glyoxylated *Hae*III restriction fragments of polyoma DNA.

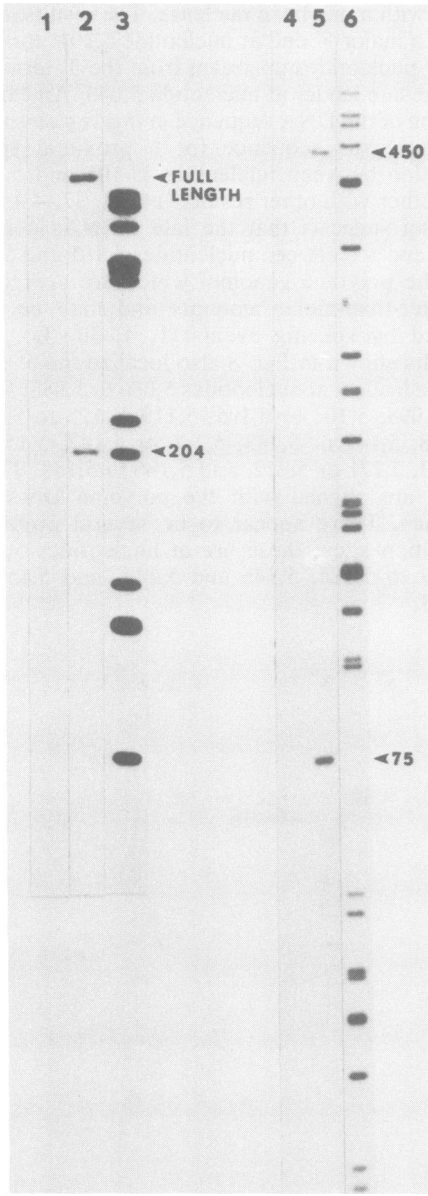

FIG. 7. Mapping the 5′ ends of the bodies of the VP1, VP2, and VP3 messages. Each hybridization mixture contained 1 μg of tRNA or late cytoplasmic poly(A)-containing RNA and an amount of 5′-end-labeled restriction fragment equivalent to 0.5 μg of full-length polyoma DNA. Tracks 1 and 2, *Hin*dIII-*Bam*HI-3 fragment (nucleotides 3,944 through 4,657) hybridized with tRNA and late RNA, respectively, digested with S1 nuclease, and then analyzed on an 8% acrylamide–8 M urea gel; tracks 4 and 5, *Bam*HI-*Hae*II-2 fragment (nucleotides 4,658 through 103) hybridized with tRNA and late RNA, respectively, digested with mung bean nuclease, and then analyzed on a 5% acrylamide–8 M urea gel; track 3, *Hpa*II-*Hin*fI restriction fragments of polyoma DNA (the lengths of these fragments [in nucleotides] are as follows: fragment 1, 577; fragment 2, 556; fragment 3, 524; fragment 4, 522; fragment 5, 456; fragment 6, 377; fragment 7, 376; fragment 8, 362; fragment 9, 348; fragment 10, 221; fragment 11, 203; fragment 12, 153; fragment 13, 142; fragment 14, 141); track 6, *Hpa*I-*Dde*I restriction fragments of polyoma DNA (the lengths of these fragments [in nucleotides] are as follows: fragment 1, 885; fragment 2, 624; fragment 3, 559; fragment 4, 470; fragment 5, 377; fragment 6, 369; fragment 7, 275; fragment 8, 192; fragment 9, 176 [doublet]; fragment 10, 152; fragment 11, 130 [doublet]; fragment 12, 128; fragment 13, 123; fragment 14, 112; fragments 15 and 16, 111; fragment 17, 102; fragment 18, 92 [doublet]; fragment 20, 46 [doublet]; fragment 21, 41).

DISCUSSION

We mapped polyoma-specific mRNAs produced early and late during productive infection, and our results can be summarized as follows. (i) The polyoma early mRNAs are initiated at three regions in the polyoma genome early in infection; there are two major initiation sites, at nucleotides 156 and 157 (73.2 m.u.) and nucleotides 167 to 169 (73.4 m.u.), and one minor initiation site, at nucleotides 138 and 139 (72.9 m.u.). In addition, late in infection there is a fourth initiation site, at nucleotides 46 to 49 (71.1 m.u.). (ii) There are four RNA splice sites in the early region of the genome; these are the proximal splice sites between nucleotides 424 and 425 (78.28 m.u.) and nucleotides 761 and 762 (84.63 m.u.) and the distal splice sites between nucleotides 809 and 810 (85.55 m.u.) and nucleotides 823 and 824 (85.82 m.u.). (iii) The early mRNAs have a major termination site near nucleotide 2,952 (26.0 m.u.) and a minor termination site near nucleotide 1,524 (99.0 m.u.). (iv) The polyoma late mRNAs have leader sequences with at least 10 different 5′ ends, which map between nucleotides 5,086 and 5,202 (66.3 to 68.5 m.u.); the leader sequences between nucleotides 5,044 and 5,101 (65.51 to 66.59 m.u.) are reiterated in the mature mRNAs. (v) The bodies of the mRNAs coding for VP2, VP3, and VP1 have unique 5′ ends at nucleotides 5,101 (66.59 m.u.), 4,732 (59.62 m.u.), and 4,147 (48.75 m.u.), respectively. (vi) The late messages have a single 3′ terminus near nucleotide 2,917 (25.3 m.u.).

The splice sites in the early region are used to generate three different early mRNAs, which code for the three polyoma T antigens. The message produced by splicing out the sequences between nucleotides 425 and 809 has a total length of about 2,400 nucleotides (257 plus 2,143 nucleotides) and, from a comparison with the polyoma DNA nucleotide sequence, codes for a protein which has 782 amino acids and a molecular weight of 87,605, corresponding to the large T antigen. The message produced by splicing out

TABLE 3. Length of ^{32}P-labeled polyoma DNA fragments protected from S1 digestion by late cytoplasmic poly(A)-containing RNA[a]

DNA fragment	Location of fragment (nucleotides)	Length of S1 nuclease-resistant fragments (nucleotides)	Region of protected DNA (nucleotides)	Message(s) protecting the DNA
*Hpa*II-1[b,c]	3,014–4,434	Full length	3,014–4,434	VP2, VP3
		1,130	3,014–4,144	VP1
*Hpa*II-3[b,d]	4,435– 24	670	4,435–5,104	VP2
		300	4,435–4,734	VP3
		153	5,045–5,197	Leader
		145	5,045–5,189	Leader
		124	5,045–5,168	Leader
		106	5,045–5,150	Leader
		84	5,045–5,128	Leader
		60	5,045–5,104	Leader
*Hin*dIII-1[b,d]	3,944–1,671	1,150	3,944–5,093	VP2
		790	3,944–4,733	VP3
		155	5,045–5,199	Leader
		145	5,045–5,189	Leader
		133	5,045–5,177	Leader
		105	5,045–5,149	Leader
		85	5,045–5,129	Leader
		60	5,045–5,104	Leader
*Taq*I-*Hpa*II-6	4,991– 24	155	5,045–5,199	Leader
		145	5,045–5,189	Leader
		130	5,045–5,174	Leader
		108	5,045–5,152	Leader
			4,991–5,098	VP2
		84	5,045–5,128	Leader
		60	5,045–5,104	Leader
*Hin*fI-9[b,d]	4,956–5,098	Full length	4,956–5,098	VP2
		54	5,045–5,098	Leader
*Hin*dIII-*Bam*HI-3[e,f]	3,994–4,657	Full length	3,944–4,657	VP2, VP3
		204	3,944–4,147	VP1
*Bam*HI-*Hae*II-2[e,f]	4,658– 103	450	4,658–5,108	VP2
		75	4,658–4,732	VP3

[a] Late poly(A)-containing cytoplasmic RNA was hybridized with restriction fragments of polyoma DNA, digested with S1 or mung bean nuclease, and then analyzed by electrophoresis as described in the text. The lengths of the protected bands were determined by using restriction digests of polyoma DNA as size markers.
[b] In vivo ^{32}P-labeled polyoma DNA.
[c] Samples were analyzed on neutral or alkaline agarose gels.
[d] Samples were glyoxylated and then analyzed on 5% acrylamide gels.
[e] 5′ End ^{32}P-labeled polyoma DNA fragment.
[f] Samples were analyzed on acrylamide–8 M urea gels.

the sequences between nucleotides 762 and 809 has a total length of about 2,738 nucleotides (595 plus 2,143 nucleotides); this message codes for a protein which has 195 amino acids, terminating at nucleotides 821 to 824, and a molecular weight of 22,785, corresponding to the small T antigen. The message produced by splicing out the sequences between nucleotides 762 and 823 has a total length of about 2,724 nucleotides (595 plus 2,129 nucleotides); this message codes for a protein which has 421 amino acids, terminating at nucleotides 1,513 to 1,515, and a molecular weight of 48,660, corresponding to the medium T antigen.

Our data do not permit us to determine whether there is an mRNA spliced between nucleotides 424 and 824. Such an RNA would code for a protein which has 85 amino acids, terminating at nucleotides 841 to 844 (17), and a molecular weight of about 9,705. Such a protein probably would have escaped detection in previous analyses of T antigens because of its low molecular weight.

The mRNAs that code for the three T antigens are present in different quantities. We estimate that the large-T-antigen message is approximately three to five times more abundant than the sum of the small-T-antigen and medium-T-antigen messages. This is consistent with the relative amounts of the three proteins noted previously (29).

We have not analyzed nuclear transcripts to

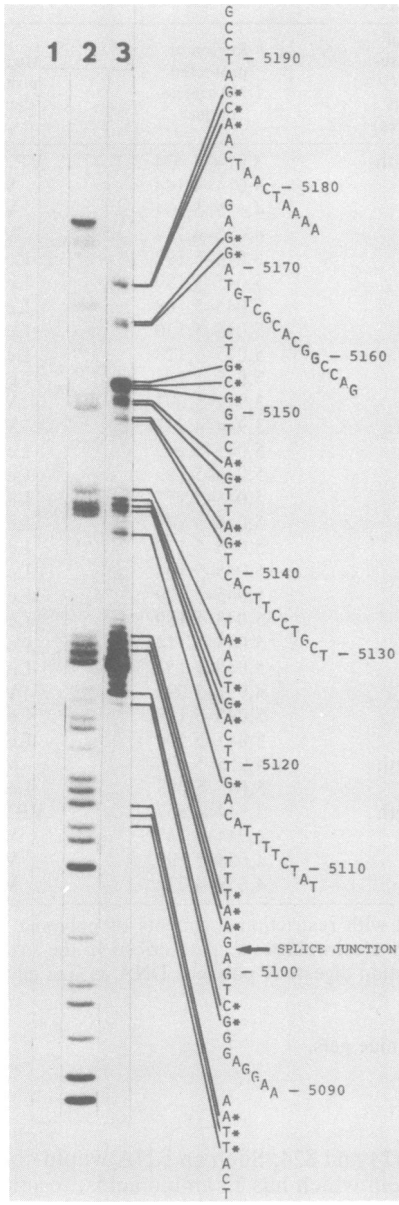

FIG. 8. Mapping the 5′ ends of the polyoma virus late leaders. The strand-separated *Bcl*I-*Nar*I-2 fragment (nucleotides 5,047 through 100) of polyoma DNA was hybridized with 5 μg of tRNA (track 1) or late poly(A)-containing cytoplasmic RNA (track 3) and digested with mung bean nuclease, and the protected fragments were fractionated on an 8% acrylamide–8 M urea gel. Track 2 contained marker fragments (the G-sequencing reaction of the *Hpa*II-*Dde*I-9 strand [nucleotides 27 through 203] complementary to the early message). The polyoma DNA sequence of the late leader is correlated with the protected fragments to the right of track 3.

determine whether they have the same 5′ ends as the mRNAs studied here. However, in other genes the first nucleotide of the mature mRNA has been shown to be the same as the 5′ end of the primary transcript (21, 36, 50, 51). Therefore, it is likely that the 5′ ends of the mRNAs represent sites of initiation of transcription.

We observed multiple protected bands for each initiation site when the 5′ ends of the early mRNAs were mapped. Two possible explanations for this observation are that the RNA polymerase "stutters" at each initiation site and that the exonuclease "nibbles" the ends of the hybrids. We believe the latter explanation to be unlikely for two reasons. First, multiple protected bands generally were not observed when we mapped other features, such as the splice acceptor sites (for example, see Fig. 5, track 2, and Fig. 7, tracks 2 and 5), and second, the use of larger amounts of nuclease for digestion did not change the pattern of bands. Heterogeneous 5′ ends have also been observed in the early mRNAs of simian virus 40 (24, 42).

Kamen et al. (30) have located a termination site for the early mRNAs at 25.8 m.u. and one for the late mRNAs at 25.3 m.u. There is good agreement between our results and the results of these authors. In both cases the early and late mRNAs were found to overlap by about 36 nucleotides at their 3′ ends.

The polyoma DNA sequences near the initiation and termination sites of the mRNAs show a number of distinctive features found in other mRNAs. The DNA sequence around the major 5′ end of the early mRNAs, 5′-^{165} CTCATTT-CAG^{174}-3′, is similar to the consensus sequence, 5′-PyPyCAPyPyPyPyPyPu-3′, (where the adenosine is the capped residue) present in many eucaryotic genes, including mouse β-globin^{maj} and β-globin^{min} (33), adenovirus late genes (51), rabbit β-globin (13), mouse lambda light chain (46), chick ovalbumin (6, 19), and silk fibroin (49). The capped adenosine residue at nucleotide 168 in polyoma virus is 33 nucleotides downstream from the beginning of an adenine-thymine-rich region at nucleotide 135 (5′-^{135}TA-TAATTA^{142}-3′). Adenine-thymine-rich regions are found 27 to 33 nucleotides upstream from unique 5′ ends in other messages (8). The sequence AAUAAA is often found 15 to 30 nucleotides upstream from polyadenylation sites in eucaryotic mRNAs (41). There are AATAAA sequences in polyoma DNA 10 to 30 nucleotides upstream from all three of the 3′ termini which we mapped. There are also sequences similar to the 5′-TTTTCACTGC-3′ sequence described by Benoist et al. (2) in the vicinity of the polyadenylation sites of several eucaryotic genes.

The polyoma late mRNA leaders have been

difficult to study because they are heterogenous and occur in low yields (30). Polyoma late mRNAs with at least 15 different 5′ ends have been inferred previously from an analysis of the 5′ cap structures of the late RNAs (9, 14, 15). By comparing the 5′ cap sequences with the polyoma DNA nucleotide sequence, Cowie et al. (9) have mapped the 5′ ends of the late messages between nucleotides 5,193 and 5,100. Using S1 mapping, Triesman (48) has also localized 5′ ends between nucleotides 5,197 and 5,085. Our observations are in good agreement with these assignments. The sequence of the late leader of the polyoma DNA strain which we have used (Fig. 8) is different by a single nucleotide in the adenine-thymine-rich region from the A2 strain used by others (9, 14, 15, 30, 35) to map the late mRNA leaders. This strain difference does not appear to have a significant effect on the relative abundance of the 5′ ends of the late leaders (compare Fig. 8 with the results of Treisman [48]). This suggests either that the TAATTAAAA sequence found in the A2 strain of polyoma virus does not play a role in positioning the major 5′ ends of the late transcripts or that the thymine-to-cytosine change ($5'\text{-}^{5,183}$TAACTAAAA^{5,175}-3′) has no effect on in vivo transcription of the late messages. Grosschedl et al. (23) have shown that a thymine-to-guanine change in the TATA box reduces the amount of transcript but does not change the initiation site of transcription in vivo.

A more precise description of the structures of the various polyoma mRNAs should emerge from nucleotide sequence analyses of DNA copies of individual RNAs cloned in bacteria. A better understanding of polyoma mRNA biogenesis, including transcriptional control and splicing, may be achieved by studying mutations in the viral DNA which affect mRNA synthesis or processing.

ACKNOWLEDGMENTS

We thank Dan Donoghue for advice in setting up the DNA sequencing gels. We also thank Carolyn Goller for help in preparation of the manuscript.

This investigation was supported by Public Health Service grants CA 13884, CA 14195, and CA 21896 from the National Cancer Institute.

LITERATURE CITED

1. **Acheson, N. H., E. Buetti, K. Scherrer, and R. Weil.** 1971. Transcription of the polyoma virus genome: synthesis and cleavage of giant late polyoma-specific RNA. Proc. Natl. Acad. Sci. U.S.A. **68**:2231–2235.
2. **Benoist, C., K. O'Hare, R. Breathnach, and P. Chambon.** 1980. The ovalbumin gene: sequence of putative control regions. Nucleic Acids Res. **8**:127–142.
3. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **11**:721–732.
4. **Birg, F., J. Favaloro, and R. Kamen.** 1977. Analysis of polyoma virus nuclear RNA by miniblot hybridization. Proc. Natl. Acad. Sci. U.S.A. **74**:3138–3142.
5. **Casey, J., and N. Davidson.** 1977. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. Nucleic Acids Res. **4**:1539–1552.
6. **Chambon, P., C. Benoist, R. Breathnach, M. Cochet, F. Gannon, P. Gerlinger, A. Krust, M. LeMeur, J. P. LePennee, J. L. Mandel, K. O'Hare, and F. Perrin.** 1979. Structural organization and expression of ovalbumin and related chicken genes, p. 55–78. *In* T. R. Russell, K. Brew, H. Faber, and J. Schults (ed.), From gene to protein: information transfer in normal and abnormal cells. Academic Press, Inc., New York.
7. **Cogen, B.** 1978. Virus-specific early RNA in 3T6 cells infected by a tsA mutant of polyoma virus. Virology **85**:222–230.
8. **Corden, J., B. Wasylyk, A. Buchwalder, P. Sassone-Corsi, C. Kedinger, and P. Chambon.** 1980. Promotor sequences in eukaryotic protein-coding genes. Science **209**:1406–1414.
9. **Cowie, A., C. Tyndall, and R. Kamen.** 1981. Sequences at the capped 5′ ends of polyoma virus late region mRNAs: an example of extreme terminal heterogeneity. Nucleic Acids Res. **9**:6305–6322.
10. **Deininger, P., A. Esty, P. LaPorte, and T. Friedmann.** 1979. Nucleotide sequence and genetic organization of the polyoma region: features common to the polyoma early region and SV40. Cell **18**:771–779.
11. **Deininger, P. L., A. Esty, P. LaPorte, H. Hsu, and T. Friedmann.** 1980. The nucleotide sequence and restriction enzyme sites of the polyoma genome. Nucleic Acids Res. **8**:855–860.
12. **Eckhart, W.** 1979. Complementation and transformation by temperature-sensitive mutants of polyoma virus. Virology **38**:120–125.
13. **Efstradiadis, A., F. C. Kafatos, and T. Maniatis.** 1977. The primary structure of rabbit β-globin mRNA as determined from cloned DNA. Cell **10**:571–585.
14. **Flavell, A. J., A. Cowie, J. R. Arrand, and R. Kamen.** 1980. Localization of three major capped 5′ ends of polyoma virus late mRNA's within a single tetranucleotide sequence in the viral genome. J. Virol. **33**:902–908.
15. **Flavell, A. J., A. Cowie, S. Legon, and R. Kamen.** 1979. Multiple 5′-terminal cap structures in late polyoma virus RNA. Cell **16**:357–371.
16. **Fried, M., and B. E. Griffin.** 1977. Organization of the genomes of polyoma and SV40. Adv. Cancer Res. **24**:67–113.
17. **Friedmann, T., A. Esty, P. LaPorte, and P. Deininger.** 1978. The nucleotide sequence and genome organization of the polyoma early region: extensive nucleotide and amino acid homology with SV40. Cell **17**:715–724.
18. **Friedmann, T., P. LaPorte, and A. Esty.** 1978. Nucleotide sequence studies of polyoma DNA. J. Biol. Chem. **253**:6561–6567.
19. **Gannon, F., K. O'Hare, F. Perrin, J. P. LaPennee, C. Benoist, M. Cochet, R. Breathnach, A. Royal, A. Garapin, B. Cami, and P. Chambon.** 1979. Organization and sequences at the 5′ end of a cloned complete ovalbumin gene. Nature (London) **278**:428–434.
20. **Gibson, W., T. Hunter, B. Cogen, and W. Eckhart.** 1977. Altered virion proteins of a temperature-sensitive mutant of polyoma virus ts59. Virology **80**:21–41.
21. **Grez, J., H. Land, K. Kiesecke, G. Schultz, A. Jung, and A. E. Sippel.** 1981. Multiple mRNA's are generated from the chicken lysozyme gene. Cell **25**:743–752.
22. **Griffin, B. E., M. Fried, and A. Cowie.** 1974. Polyoma DNA: a physical map. Proc. Natl. Acad. Sci. U.S.A. **71**:2077–2081.
23. **Grosschedl, R., B. Wasylyk, P. Chambon, and M. Birnstiel.** 1981. Point mutation in the TATA box curtails expression of sea urchin H2A histone gene in vivo. Nature (London) **294**:178–180.
24. **Haegeman, G., and W. Fiers.** 1981. Characterization of the

5'-terminal cap structures of early simian virus 40 mRNA. J. Virol. 35:955–961.
25. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. 26:365–369.
26. **Horowitz, M., S. Bratosin, and Y. Aloni.** 1978. Polyoma infected cells contain at least three spliced late RNAs. Nucleic Acids Res. 5:4663–4675.
27. **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.
28. **Hunter, T., M. A. Hutchinson, and W. Eckhart.** 1978. Translation of polyoma virus T antigens in vitro. Proc. Natl. Acad. Sci. U.S.A. 75:5917–5921.
29. **Hutchinson, M. A., T. Hunter, and W. Eckhart.** 1978. Characterization of T-antigens in polyoma-infected and transformed cells. Cell 15:65–77.
30. **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.
31. **Kamen, R., J. Favaloro, J. Parker, R. Treisman, L. Lania, M. Fried, and A. Mellor.** 1980. A comparison of polyoma virus transcription in productively infected mouse cells and transformed rodent cell lines. Cold Spring Harbor Symp. Quant. Biol. 43:63–75.
32. **Kamen, R., D. M. Lindstrom, H. Shure, and R. Old.** 1974. Virus-specific RNA in cells productively infected or transformed by polyoma virus. Cold Spring Harbor Symp. Quant. Biol. 39:187–198.
33. **Konkel, D. A., S. M. Tilghman, and P. Leder.** 1978. The sequence of the chromosomal mouse β-globin major gene: homologies in capping, splicing, and poly(A) sites. Cell 15:1125–1132.
34. **Lai, C. J., R. Dhar, and G. Khoury.** 1978. Mapping the spliced and unspliced late lytic SV40 RNAs. Cell 14:971–982.
35. **Legon, S., A. J. Flavell, A. Cowie, and R. Kamen.** 1979. Amplification of the leader sequence of "late" polyoma virus mRNAs. Cell 16:373–388.
36. **Luse, D. S., and R. G. Roeder.** 1980. Accurate transcription initiation on a purified mouse β-globin DNA fragment in a cell-free system. Cell 20:691–699.
37. **Manor, N., M. Wu, N. Baran, and N. Davidson.** 1979. Electron microscopic mapping of RNA transcribed from the late region of polyoma virus DNA. J. Virol. 32:283–303.
38. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. 65:499–560.
39. **McDonnell, M. W., M. N. Simon, and F. W. Studier.** 1977. Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J. Mol. Biol. 110:119–146.
40. **McMaster, G. K., and G. G. Carmichael.** 1977. Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. Proc. Natl. Acad. Sci. U.S.A. 74:4835–4838.
41. **Proudfoot, N., and G. Brownlee.** 1976. 3' Non-coding region sequences in eukaryotic messenger RNA. Nature (London) 252:359–362.
42. **Reddy, V. B., P. K. Ghosh, P. Lebowitz, M. Piatak, and S. M. Weissman.** 1979. Simian virus 40 early mRNA's. I. Genomic localization of 3' and 5' termini and two major splices in mRNA from transformed and lytically infected cells. J. Virol. 30:279–296.
43. **Rogers, J., and R. Wall.** 1980. A mechanism for RNA splicing. Proc. Natl. Acad. Sci. U.S.A. 77:1877–1879.
44. **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.
45. **Soeda, E., J. A. Arrand, N. Smolar, J. E. Walsh, and B. E. Griffin.** 1980. Coding potential and regulatory signals of the polyoma virus genome. Nature (London) 283:445–453.
46. **Tonegawa, S., A. Maxam, R. Tizard, O. Bernard, and W. Gilbert.** 1978. Sequence of a mouse germ-line gene for a variable region of an immunoglobulin light chain. Proc. Natl. Acad. Sci. U.S.A. 75:1485–1489.
47. **Tooze, J.** 1980. DNA tumor viruses: molecular biology of tumor viruses, part 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
48. **Treisman, R.** 1980. Characterization of polyoma late mRNA leader sequences by molecular cloning and DNA sequence analysis. Nucleic Acids Res. 8:4867–4888.
49. **Tsujimoto, Y., and Y. Suzuki.** 1979. Structural analysis of the fibroin gene at the 5' end and its surrounding regions. Cell 16:425–436.
50. **Wasylyk, B., C. Kedinger, J. Corden, O. Brison, and P. Chambon.** 1980. Specific in vitro initiation of transcription on conalbumin and ovalbumin genes and comparison with adenovirus-2 early and late genes. Nature (London) 285:367–373.
51. **Ziff, E. B., and R. M. Evans.** 1978. Coincidence of the promotor and capped 5' terminus of RNA from the adenovirus 2 major late transcription unit. Cell 15:1463–1475.
52. **Zuckermann, M., H. Manor, J. Parker, and R. Kamen.** 1980. Electron microscopic demonstration of the presence of amplified sequences at the 5' ends of the polyoma virus late mRNAs. Nucleic Acids Res. 8:1505–1513.