Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 7142–7146, December 1982
Biochemistry

Changes in the topography of early region transcription during polyoma virus lytic infection

(transcription initiation/open reading frame/tumor antigen regulation)

ROBERT G. FENTON AND CLAUDIO BASILICO

Department of Pathology, New York University School of Medicine, 550 First Avenue, New York, New York 10016

Communicated by Saul Krugman, August 23, 1982

ABSTRACT **We have studied the pattern of transcription of the early region of polyoma virus DNA after the onset of the late phase of lytic infection of mouse cells. Following initiation of viral DNA synthesis, the early/late switch is accompanied not only by efficient production of late mRNAs but also by the appearance of previously unidentified early-strand RNAs which have certain structural features in common with the classical early mRNAs. Stable poly(A)$^+$RNAs have been identified by blot analysis and S1 nuclease mapping that are not detected early during infection or in polyoma virus-transformed cells. One group consists of transcripts whose 5′ ends map 150–200 nucleotides upstream from the major early 5′ ends (at positions 148 and 153 on the polyoma virus genome) but whose splicing pattern and poly(A) addition sites are indistinguishable from those of mRNAs produced early in infection. The 5′ exons of these early region transcripts contain an open translational reading frame that extends from nucleotide positions 5,255 to 124 and is capable of encoding a basic protein of 53 amino acids. Transcription of these RNAs does not appear to be negatively regulated by large tumor antigen. A transcript of 1,800 nucleotides appears to map predominantly between 93 and 26 map units and does not contain sequences present in the early mRNA 5′ exons. These data suggest that, after the onset of polyoma DNA replication, the activation of new early-strand promoters leads to the expression of previously untranscribed viral DNA sequences.**

The polyoma virus (Py) genome, a double-stranded closed circular DNA molecule of 5.3 kilobase pairs, is organized into two distinct transcriptional units, the "early" and the "late" region. The early region comprises nucleotide sequences mapping from approximately 73 to 26 map units (m.u.); transcription of the late region proceeds counterclockwise on the opposite DNA strand from approximately 66 to 26 m.u. (1). Nucleotide sequences between 66 and 73 m.u. are thought to contain mainly noncoding regulatory sequences, including the origin of viral DNA replication, early and late transcriptional promoters, and the so-called enhancer sequences which are required for *in vivo* expression of the early genes (2, 3). Transcription of the early region generates a single nuclear RNA precursor which is spliced differentially to produce three mRNAs encoding the viral early proteins [large, middle, and small tumor antigens (T)] (4) (Fig. 1). Efficient transcription of the late region and production of the three virion proteins requires the initiation of viral DNA replication and thus generally is not detected in transformed cells (1).

It generally has been assumed that transcription of the early region occurs in qualitatively identical manners at early and late times of infection, although few studies have focused on this question (1). We have examined the topography of early region transcription at late times of infection of mouse fibroblasts with Py. Species of cytoplasmic RNAs appeared that are not observed at early times of infection or in Py-transformed cells, and these RNA species constituted approximately 50% of early region RNAs produced at late times. Our results suggest that the early-to-late switch in Py lytic infection is accompanied by the activation of previously quiescent E-strand promoters (in addition to the late promoter) and perhaps the production of novel viral gene products.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Cells and Viruses. To study the viral mRNAs produced during lytic infection, 3T3D mouse fibroblasts were infected as described (6). The R5-1 rat cell line was transformed by ts-a Py DNA and contains a single nontandem insertion of viral DNA, whose structure and transcripts have been described (7, 8). The Py wild-type small-plaque strain and the ts-a Py mutant have been described (9).

Preparation of Cytoplasmic Poly(A)$^+$RNA and Blot Analysis. The procedure for preparing cytoplasmic poly(A)$^+$RNA has been described (6). RNA gel analysis on denaturing agarose gels was performed as described by Alwine *et al.* (10) with certain modifications (6, 8). Viral sequences were identified by hybridization with nick-translated, ^{32}P-labeled Py DNA fragments, including *Pst* I fragment 1 (cloned in pBR322), and the *Hpa* II fragments indicated in the text, which were purified by two sequential polyacrylamide gel electrophoreses and electrolution.

S1 Nuclease Analysis. The alkaline S1 nuclease blot procedure utilizing unlabeled viral DNA restriction fragments has been described in detail (6, 11). More precise mapping of the 5′ splice acceptor sites and 5′ termini utilized the modification of the technique described by Weaver and Weissman (12). Purified restriction fragments of Py DNA were labeled with ^{32}P at their 5′ ends as described by Maxam and Gilbert (13). The E strand of end-labeled *Hin*fI fragment 4 was isolated by electrophoresis on denaturing polyacrylamide gels. Each hybridization reaction contained 15,000–25,000 cpm of end-labeled DNA fragment and the amounts of poly(A)$^+$RNA indicated in the text. Hybridization conditions and S1 nuclease digestion were performed as in the alkaline S1 nuclease analysis. Protected DNA fragments were size fractionated on thin 8% polyacrylamide gels containing 50% urea. Gels were dried and exposed for autoradiography as described (6).

RESULTS

Size Heterogeneity of the Early RNAs Produced Late During Lytic Infection. To study the pattern of early region transcription at late times during lytic infection, 3T3D fibroblasts were infected (10 plaque-forming units per cell) with wild-type Py and cytoplasmic poly(A)$^+$RNA was isolated after 36 hr of

Abbreviations: Py, polyoma virus; m.u., map units; T, tumor antigen.

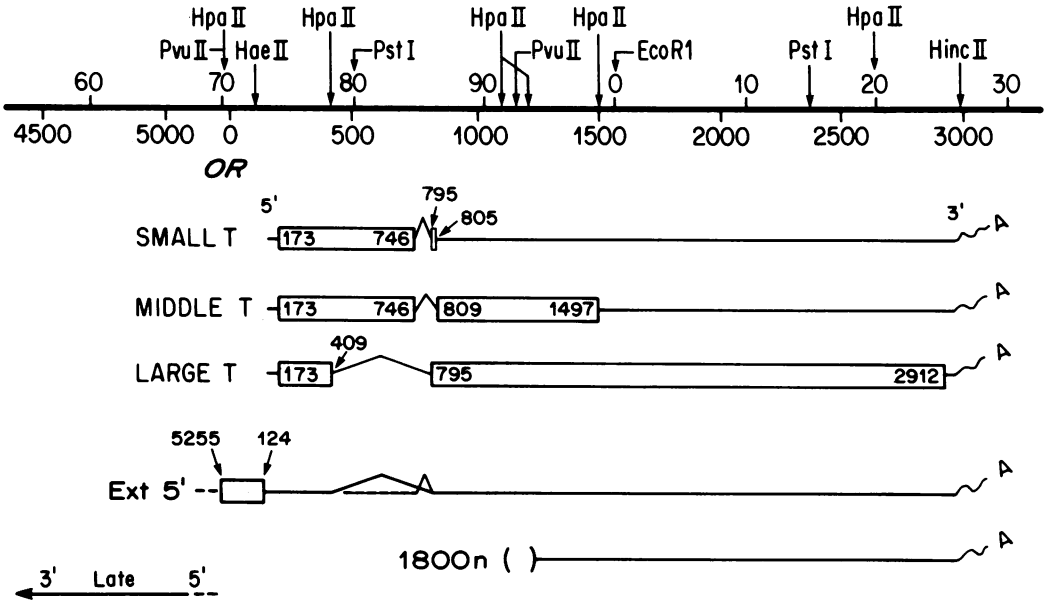

FIG. 1. Linearized map of the Py early region including the segment of viral DNA thought to contain the origin of viral DNA replication and early and late transcriptional control signals. The physical map is divided into map units (0 m.u. at the *Eco*RI site) and nucleotide number. The approximate position of the origin of replication (*OR*) and the sites of cleavage of restriction enzymes used in this study are also shown. The structures of the mRNAs for the three Py early proteins (small, middle, and large T) are indicated with their coding regions (▭), 3′ and 5′ noncoding regions (——), and intervening sequences (^). The numbers within the coding regions represent initiation and termination codons and splice junctions. In addition, the presumed structures of two groups of early region transcripts (Ext 5′ and 1800n, described in detail in the text) are shown. The RNAs with 5′ exons extending into the origin region are indicated as a single RNA, although each of the classical early splices can be utilized (as indicated). These RNAs contain heterogeneous 5′ exons (indicated by ---). We have not determined the position of the 5′ end of the 1800n RNA species or whether this RNA is spliced. The positions of the heterogeneous late mRNA leader sequences (---) and the polarity of late transcription are also shown.

growth at 37°C. Fig. 2 shows blot analysis of this RNA (lane 3) hybridized to the nick-translated *Pst* I fragment 1 early region probe (79.7–15.0 m.u.). As a comparison, cytoplasmic poly(A)$^+$-RNA isolated at early times [i.e., 16 hr after infection in the presence of 1-β-D-arabinofuranosylcytosine (cytosine arabinoside) at 20 μg/ml] or from the R5-1 transformed rat cell line, which contains an intact early region, is shown. Early lytic (lane 2) and R5-1 (lane 1) viral RNAs were of the sizes expected for the middle and small T mRNAs (2,900 nucleotides) and the large T mRNA (2,550 nucleotides) (4, 6) (Fig. 1). The pattern of early region transcription at late times was more complicated than that observed early during infection or in the R5-1 cell line. Two new bands appeared that migrated at positions corresponding to 2,750 and 3,100 nucleotides, approximately 200 nucleotides longer than the large T and the small and middle T mRNAs, respectively. Based on the relative band intensities, these RNA species were present in abundances about equal to those of the classical early mRNAs which continued to be produced at late times. A minor band (2,400 nucleotides long) that migrated faster than the large T mRNA was seen. In addition, a 1,800-nucleotide early region RNA species appeared late in infection and represented about 10% of the early region cytoplasmic transcripts. As shown in lane 4 of Fig. 2, none of these transcripts hybridized to the *Hpa* II fragment 1 late region probe (27.0–53.8 m.u.) which identifies the Py late mRNAs encoding the virion proteins (VP1, VP2, and VP3) (15). Time-course experiments (data not shown) demonstrated that the appearance of these early region RNAs correlated closely in time with the production of high levels of late mRNAs and was similarly blocked by inhibitors of DNA synthesis.

To map these RNAs more precisely, duplicate diazobenzyloxymethyl-paper filters containing early and late lytic RNA samples were hybridized to nick-translated purified Py DNA restriction fragments that represented specific nonoverlapping regions of the early transcription unit. *Hpa* II fragments 4 and 5 (78.1–91.3 and 70.5–78.1 m.u., respectively) are homologous to significant portions of the 5′ exons of early mRNAs; *Hpa* II

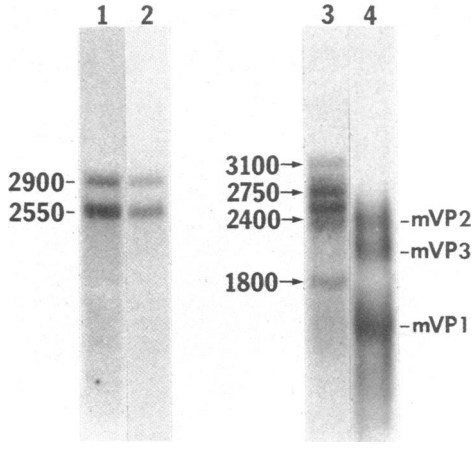

FIG. 2. Identification of early region transcripts produced at late times of Py lytic infection. Cytoplasmic poly(A)$^+$RNA isolated from 3T3D fibroblasts at 16 hr (early) or 36 hr (late) after infection with wild-type Py virus or from a Py-transformed rat cell line (R5-1) was analyzed by the Alwine *et al.* (10) procedure. Lanes: 1, 10 μg of poly(A)$^+$RNA from R5-1 hybridized to the *Pst* I fragment 1 early region probe; 2, 5 μg of early lytic mRNA hybridized to the *Pst* I fragment 1 probe; 3, 5 μg of late lytic RNA hybridized to the *Pst* I fragment 1 probe [the approximate sizes (in nucleotides) of the four early region transcripts are shown; size markers were 28S and 18S ribosomal RNA (4,700 and 1,800 nucleotides, respectively) and the Py early mRNAs]; 4, 1 μg of the same late lytic RNA preparation as in lane 3 but hybridized to the *Hpa* II fragment 1 late region probe (27.0–53.8 m.u.). The positions of the mRNAs encoding the three virion proteins are shown; the heterogeneity of sizes of these mRNAs reflects the complexity of the reiterated 5′ exons of these transcripts (14).

fragment 2 (98.6–19.9 m.u.) hybridizes only to sequences in the 3′ colinear segment (Fig. 1). As shown in Fig. 3, cytoplasmic poly(A)$^+$RNA from early lytic infection (lane 1) hybridized to each of these probes, generating in each case a pattern indistinguishable from that shown in Fig. 2. When each of these probes was hybridized to late lytic RNA (lane 2), the complex late topography of early region transcription was observed. Specifically, four discrete bands between 2,550 and 3,100 nucleotides long were identified by all three probes. Two of these bands (2,550 and 2,900 nucleotides) comigrated with the early mRNAs; the two others migrated at positions corresponding to 2,750 and 3,100 nucleotides, as has been found previously with the *Pst* I fragment 1 probe.

On the other hand, the 1,800-nucleotide species hybridized to the 3′ exon-specific probe *Hpa* II fragment 2 but not to the *Hpa* II fragment 4 and 5 probes. In addition, this RNA species was also found to hybridize to *Hpa* II fragments 7 and 6 (93.4–98.6 m.u. and 19.9–27.0 m.u., respectively; data not shown). These and other experiments including hybridization with an M13 clone containing the E strand of Py *Pst* I fragment 1 indicate that the 1,800-nucleotide RNA has the same polarity as early RNA and maps from approximately 93 to 26 m.u. on the viral genome.

Mapping of the Extended Early RNAs Produced Late During Lytic Infection. We compared the alkaline S1 nuclease patterns of late lytic and early lytic cytoplasmic poly(A)$^+$RNAs. Five micrograms of each RNA sample was hybridized to Py DNA linearized at 58 m.u. by *Bam*HI or to *Pvu* II fragment 2 (70.0–92.1 m.u.), and the resulting hybrids were treated with S1 nuclease. Products were electrophoresed on alkaline agarose gels and transferred to nitrocellulose, and the protected viral DNA segments were identified by hybridization to ^{32}P-labeled Py DNA. The *Bam*HI pattern (not shown) generated by both these RNA samples demonstrated identical 3′ colinear segments (2,140 nucleotides), indicating that the early region transcripts are polyadenylylated at 26 m.u. as reported (6). To map the coordinates of the 5′ exons of the new RNAs more precisely, S1 nuclease analysis was performed with *Pvu* II fragment 2 as the hybridization probe (Fig. 4*A*). Late lytic cytoplasmic RNA preparations contained the same 5′ exons present in early lytic RNA (260 and 600 nucleotides) and, in addition, showed two bands (indicated by the arrows) migrating above the early leader sequences. These segments were 780 and 440 nucleotides long, the size expected for colinear DNA segments extending from the early splice donor sites (at 78.3 and 85.6 m.u.) upstream to the 5′ end of *Pvu* II fragment 2.

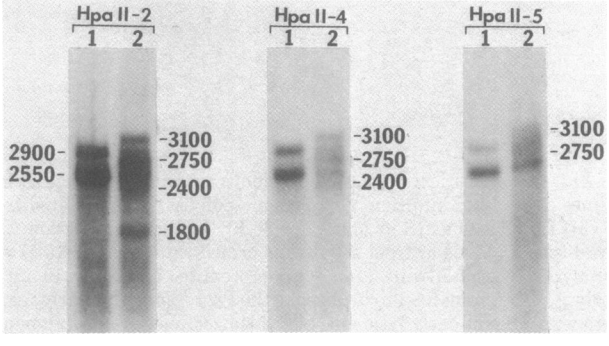

FIG. 3. Hybridization of the early region transcripts with different restriction fragments of Py DNA. Ten micrograms of early lytic (lane 1) or 5 μg of late lytic cytoplasmic poly(A)$^+$RNA (lane 2) was electrophoresed on a denaturing methylmercury gel and transferred to diazobenzyloxymethyl-paper strips. Each strip was hybridized to purified Py DNA *Hpa* II fragments, as specified in the figure, that had been ^{32}P-labeled by nick-translation. The map coordinates of the *Hpa* II probes are given in the text and in Fig. 1.

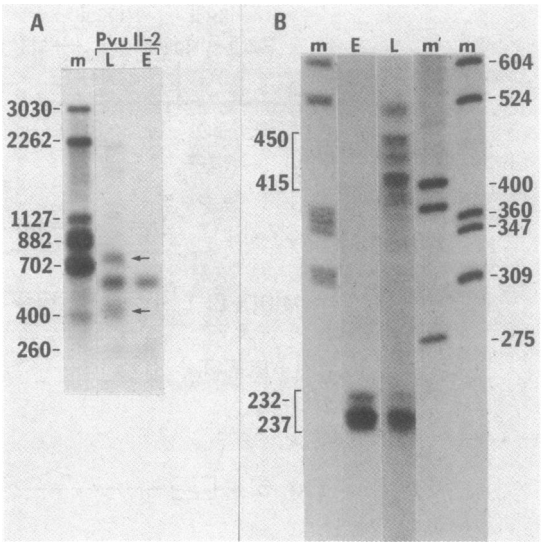

FIG. 4. Mapping of the 5′ ends of the 3,100- and 2,750-nucleotide early region transcripts. (*A*) Alkaline S1 nuclease analysis of the early region transcripts produced late during infection. Ten micrograms of early lytic mRNA (lane E) or 5 μg of late lytic cytoplasmic poly(A)$^+$RNA (lane L) was hybridized to an excess of unlabeled *Pvu* II fragment 2 of Py DNA, and S1 nuclease-resistant fragments were identified by blot hybridization with ^{32}P-labeled Py DNA. The arrows indicate DNA segments rendered S1 nuclease-resistant by hybridization to the extended leader segments of the 3,100- and 2,750-nucleotide RNA species. Markers (lane m) were Py DNA cleaved with *Hin*dIII or *Hpa* II. (*B*) The 5′ ends of the early region transcripts were mapped by using the high-resolution S1 nuclease technique. *Hin*fI fragment 4 (66.4–77.8 m.u.) was labeled with ^{32}P at its 5′ end, and the E-strand DNA was hybridized to 10 μg of early (lane E) or 5 μg of late (lane L) cytoplasmic poly(A)$^+$RNA. S1 nuclease-resistant DNA segments were electrophoresed on denaturing polyacrylamide gels and autoradiography was performed. Markers (m, m′) were Py or pBR322 DNA fragment end-labeled after cleavage with *Hin*fI or *Hpa* II. The DNA segments mapping the early mRNA 5′ ends near 73.3 m.u. (232–237 nucleotides) and the 5′ ends of the early region transcripts (415–450 nucleotides) are shown.

To determine the position of the 5′ ends of these exons, cytoplasmic poly(A)$^+$RNAs were hybridized to 5′ end-labeled E-strand DNA of *Hin*fI fragment 4 of Py DNA (66.4–77.8 m.u.). Hybrids were digested to completion with S1 nuclease, and the products were separated on a thin 8% polyacrylamide gel containing 50% urea (Fig. 4*B*). Early lytic mRNA protected two bands migrating near positions 235 ± 5 nucleotides, corresponding in size to DNA segments extending from the ^{32}P-labeled end of the *Hin*fI fragment 4 probe at 77.8 m.u. to the major 5′ ends of the Py early mRNAs [at positions 148 and 153, as described by Treisman *et al.* (4)]. Late lytic RNA also protected these 5′ colinear segments but, in addition, it protected a heterogeneous group of DNA fragments whose lengths ranged from 415 to 450 nucleotides. In agreement with the alkaline S1 nuclease analysis, these positions mapped 180–215 nucleotides upstream from the major cap sites of the early lytic mRNAs.

It was important to rule out the possibility that previously undetected early region splice sites were utilized at late times of infection and that this mechanism was involved in the appearance of the early mRNAs we describe. Therefore, the early splice acceptor sites were mapped by using 5′ end-labeled *Hin*fI fragment 5 (78.8–88.7 m.u.) as the DNA probe in S1 nuclease experiments. As shown in Fig. 5, early and late RNA samples protected two bands—150 ± 5 and 165 ± 5 nucleotides—that extended upstream from 88.7 m.u. to the two early region splice acceptor sites (at positions 809 and 795, respectively; see Fig.

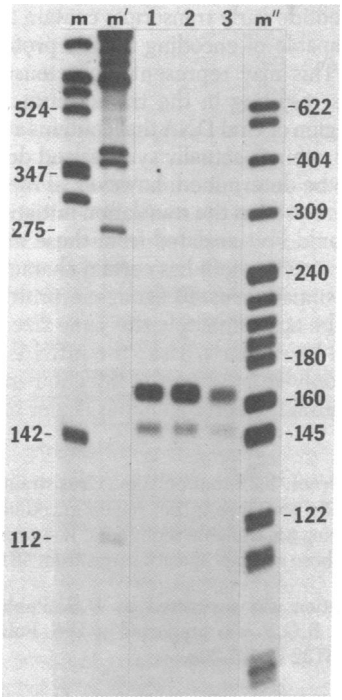

FIG. 5. Analysis of the early mRNA splice acceptor sites by high-resolution S1 nuclease mapping. Purified *Hinf*I fragment 5 (78.8–88.7 m.u.; nucleotides 436–960) was labeled at its 5′ ends by using polynucleotide kinase and hybridized to the specified amounts of cytoplasmic poly(A)$^+$RNA from early or late lytic infection or from the Py-transformed rat cell line R5-1, and the hybrids were treated with S1 nuclease. S1 nuclease-resistant fragments were size fractionated. Size markers were Py DNA end-labeled after cleavage with *Hpa* II (lane m) or *Hinf*I (lane m′) or pBR322 DNA after cleavage with *Hpa* II (lane m″). Lanes: 1, early lytic cytoplasmic poly(A)$^+$RNA (10 μg); 2, late lytic mRNA (10 mg); 3, R5-1 RNA (10 μg).

1) as shown by Treisman *et al.* (4). These data demonstrate that the early region splice acceptor sites utilized at early and late times are identical and that changes in the pattern of early region splicing cannot account for the appearance of the early region mRNAs. Significantly, no bands larger than the 165-nucleotide fragment were observed, indicating that no unspliced RNAs are represented in either RNA sample. These results strongly suggest that leakage of nuclear RNA into the cytoplasm at late times of infection is not responsible for the appearance of the virus-specific RNAs described above.

The nucleotide sequence of the extended 5′ exons contained an open translational reading frame (Fig. 6) that extended from an AUG initiation codon at position 5,255–5,257 to an in-frame termination codon at 122–124 (the UAA sequence within the early region T-A-T-A box) (2). The heterogeneous 5′ termini of the 3,100- and 2,750-nucleotide transcripts mapped just upstream from this initiation codon, suggesting that the open frame could be translated from the 5′ exons of these RNAs (see *Discussion*).

T-Antigen Regulation of Early Transcription Late in the Growth Cycle. The Py *A* gene product (large T) has been shown to regulate viral early transcription *in vivo* in a manner analogous to that described for simian virus 40 (17, 18). To determine whether T also repressed the production of the species of early RNAs found late in the cycle, 3T3 mouse cells infected with ts-a Py (producing a thermolabile large T) at 33°C were shifted to 39.5°C early or late during the infectious cycle. As described (8), this resulted in a 5- to 10-fold increase in the levels of viral mRNAs early in infection, with no substantial change of the qualitative pattern of transcription. Late in infection the complex pattern of early RNA described above was observed at 33°C. Shift to 39.5°C, however, resulted only in an increased level of those RNAs whose 5′ ends mapped at the major early cap site (i.e., near 73.3 m.u.). The amounts of the 5′ extended early RNAs (3,100 and 2,750 nucleotides) did not appear to increase, although S1 nuclease analysis indicated that they continued to be present at a low level (data not shown). The results suggest that large T regulates transcription initiating at the early mRNA cap site near 73.3 m.u. but does not influence the transcription of RNAs whose 5′ ends map heterogeneously upstream from the viral origin of replication.

DISCUSSION

The results presented in this paper show that the early-to-late switch in Py lytic infection is accompanied by distinct changes in the transcription of the early region. After the onset of viral DNA replication, we observed the appearance in the cytoplasm of the infected cells, of new species of viral mRNAs that collectively account for about 50% of the early region transcripts produced at late times (Fig. 1). The 2,400-nucleotide RNA species is relatively minor and migrates 100–150 nucleotides faster than the large T mRNA on denaturing agarose gels. The origin of this RNA species is suggested by recent studies (19) indicating that a minor 5′ end of early region RNA isolated at late times during infection occurs near position 300 on the viral DNA se-

5′ Ends of Novel Early Region Transcripts

```
                                                    Met Thr Thr Ala Val Phe Phe Ser Ile Lys Gln Arg Pro Gly Thr Pro Gly Pro Leu
 UCU CCA GGC CUA GAA UGU UUC CAC CCA AUC AUU ACU AUG ACA ACA GCU GUU UUU UUU AGU AUU AAG CAG AGG CCG GGG ACC CCU GGC CCG CUU
5220        5230            5240        5250        5260        5270        5280              1        10

 Thr Leu Glu Lys Lys Lys Arg Gly Ile Val Glu Ala Ser Arg Gly Asn Leu Ser Lys Gln Asp Trp Arg Leu Gly Gly Ala Val Gly Pro Pro
 ACU CUG GAG AAA AAG AAG AGA GGC AUU GUA GAG GCU UCC AGA GGC AAC UUG UCA AAA CAG GAC UGG CGU CUU GGA GGC GCU GUG GGG CCA CCC
20          30          40          50          60          70          80          90          100         110

 Lys Leu Ile Stp
 AAA UUG AUA UAA UUA AGC CCC AAC CGC CUC UUC CCG CCU CAU UUC AGC
        120         130         140       150 ▲     160 ▲
                                            Major Early 5' Ends
```

FIG. 6. Amino acid sequence of the putative Py gene product encoded within the extended 5′ exons of the early region transcripts observed at late times of lytic infection. The RNA sequence of the extended 5′ exons is shown with the nucleotide coordinates and the junction of *Hpa* II fragments 3 and 5 (corresponding to position 0) (2). The bracket on the top line indicates the region containing the heterogeneous 5′ ends of the new RNA species (between positions 5,230 and 5,265). The major 5′ ends of the early lytic mRNAs are designated by the vertical arrows at positions 148 and 153 (corresponding to 73.3 m.u.). The amino acid sequence of the protein encoded by the open reading frame is aligned with the coding region. This putative protein contains 53 amino acids (including the NH_2-terminal methionine) with 10 basic and 3 acidic amino acids.

quence, 150 nucleotides downstream from the major early mRNA cap sites. This position corresponds to an *in vitro* transcription start site (19) although the efficiency of initiation is much less than at the major early promoter. Significantly, a consensus T-A-T-A box is located 25–30 base pairs upstream from this position (2). The precise structure of the 1,800-nucleotide RNA species remains to be determined, but it appears to be colinear with E-strand viral DNA from near 93 m.u. to about 26 m.u. This RNA species does not contain sequences homologous to the early mRNA 5′ exons, suggesting that it does not arise at the classical early promoter. A promoter for this RNA may exist near 93 m.u. or, alternatively, it could be located further upstream on the genome. Splicing of a small 5′ exon to a position near 93 m.u. could generate this RNA species.

The group of RNAs whose 5′ ends extend 180–215 nucleotides upstream from the early mRNA cap site is particularly interesting. These RNAs (approximately 3,100 and 2,750 nucleotides long) have splice junctions and the poly(A) addition site in common with early mRNAs, but their 5′ ends map heterogeneously on the late side of the viral origin of replication. The 5′ ends of these transcripts were mapped to between positions 5,230 and 5,265 on the Py genome. This region of the viral DNA also encodes the heterogeneous late mRNA 5′ exons (14) and contains important regulatory elements of both the early and late promoters, including "enhancer" sequences (3). The fact that the synthesis of both of these groups of RNAs is activated by DNA replication suggests that a common mechanism may be responsible for activating E- and L-strand promoters in this region of the viral DNA. An alternative interpretation is that these early region transcripts arise from large tandem E-strand transcripts that may be present in the nucleus at late times of infection; internal cleavage could occur to generate the 5′ ends observed. This possibility seems unlikely because of recent findings from many laboratories that the 5′ ends of eukaryotic mRNAs are generated by transcription initiation (20–22), not endonucleolytic cleavage of a larger precursor.

It recently was reported that late in simian virus 40 infection the 5′ termini of the early RNA shift upstream by about 40–50 nucleotides (23, 24). It has been suggested that the upstream shift is mediated by large T (23); however, binding of T to site I on the simian virus 40 DNA molecules is not necessary for the upstream shift, and *in vitro* transcription initiating at the upstream sites is also inhibited by large T, albeit at higher concentrations (24). Similar to the simian virus 40 findings, our data indicate that differential regulation of early region transcription by large T occurs at late times of infection. Temperature shift-up experiments with ts-a-mutant Py-infected cells showed that mRNAs with 5′ ends mapping at the major early cap sites (near position 73.3) were selectively overproduced at the nonpermissive temperature. The steady-state levels of the RNA species whose 5′ termini map upstream were not increased. Therefore, large T represses transcription initiating at the major early cap sites (at positions near 150) but not at the heterogeneous upstream sites. However, we cannot rule out the possibility that transcription from the upstream promoter can only take place from replicating viral DNA molecules. Our data are compatible with the hypothesis that large T mediates the upstream shift in transcription only indirectly, by allowing viral DNA replication. The changes in the pattern of early transcription may be due to the presence, after replication, of new DNA templates that have undergone conformational changes allowing initiation of transcription at new sites. This hypothesis is consistent with the failure to observe these new transcripts in transformed cells.

As described in *Results*, the extended 5′ exons of the 3,100- and 2,750-nucleotide early transcripts contain an open reading frame that is capable of encoding a basic protein of 53 amino acids (Fig. 6). This may represent the extension of the early mRNA 5′ ends resulting in the transcription of a previously unexpressed region of viral DNA that contains a structural gene. Whether this protein is actually synthesized during lytic infection remains to be determined; however, if the 5′ penultimate AUG codon were used as the translation initiation site (25) then the protein should be translated from these mRNAs. The putative 53-amino acid protein has certain characteristics in common with the simian virus 40 agnogene protein (26). The Py protein would be approximately the same size as the agnoprotein and would be similarly basic; the mRNAs encoding them are expressed only at late times of infection and both proteins are encoded within the heterogeneous leader sequences of viral mRNAs.

Note Added in Proof. Different polyoma virus strains or isolates differ in their nucleotide sequence in the region encoding the putative 53-amino acid protein. All available sequences, however, contain an open reading frame whose coding capacity varies from 48 to 72 amino acids.

This investigation was supported by U.S. Public Health Service Grant CA 16239. R.G.F. was supported by U.S. Public Health Service Training Grant 5T32 GM07-308.

1. Tooze, J. (1980) *DNA Tumor Viruses: Molecular Biology of Tumor Viruses*, Part 2 (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
2. Soeda, E., Arrand, J. R., Smolar, N., Walsh, J. E. & Griffin, B. E. (1980) *Nature (London)* **283**, 445–453.
3. Tyndall, C., La Mantia, G., Thacker, C. M. & Kamen, R. (1981) *Nucleic Acids Res.* **9**, 6231–6250.
4. Treisman, R., Cowie, A., Favaloro, J., Jat, P. & Kamen, R. (1981) *J. Mol. Appl. Genet.* **1**, 83–92.
5. Rio, D., Robbins, A., Myers, R. & Tjian, R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5706–5710.
6. Fenton, R. G. & Basilico, C. (1981) *J. Virol.* **40**, 150–163.
7. Colantuoni, V., Dailey, L., Della Valle, G. & Basilico, C. (1982) *J. Virol.*, in press.
8. Fenton, R. G. & Basilico, C. (1982) *Virology*, in press.
9. Basilico, C., Gattoni, S., Zouzias, D. & Della Valle, G. (1979) *Cell* **17**, 645–659.
10. Alwine, J. C., Kemp, D. J. & Stark, G. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5350–5354.
11. Favaloro, J., Treisman, R. & Kamen, R. (1980) *Methods Enzymol.* **65**, 718–749.
12. Weaver, R. F. & Weissman, C. (1979) *Nucleic Acids Res.* **6**, 1175–1193.
13. Maxam, A. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 460–499.
14. Treisman, R. H. (1980) *Nucleic Acids Res.* **8**, 4867–4888.
15. Kamen, R. I., Favaloro, J. M. & Parker, J. T. (1980) *J. Virol.* **33**, 637–651.
16. Kamen, R., Favaloro, J., Parker, J., Treisman, R., Lania, L., Fried, M. & Mellor, M. (1980) *Cold Spring Harbor Symp. Quant. Biol.* **44**, 63–75.
17. Reed, S. I., Stark, G. R. & Alwine, J. C. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3083–3087.
18. Alwine, J. C., Reed, S. I. & Stark, G. R. (1977) *J. Virol.* **24**, 22–27.
19. Jat, P., Roberts, J. W., Cowie, A. & Kamen, R. (1982) *Nucleic Acids Res.* **10**, 871–887.
20. Hagenbuchle, O. & Schibler, U. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2283–2286.
21. Ziff, E. B. & Evans, R. M. (1978) *Cell* **15**, 1463–1475.
22. Darnell, J. E. (1982) *Nature (London)* **297**, 365–371.
23. Ghosh, P. K. & Lebowitz, P. (1981) *J. Virol.* **40**, 224–240.
24. Hansen, U., Tenen, D. G., Livingston, D. M. & Sharp, P. A. (1981) *Cell* **27**, 603–612.
25. Kozak, M. (1980) *Cell* **22**, 7–8.
26. Jay, G., Nomura, S., Anderson, C. & Khoury, G. (1981) *Nature (London)* **291**, 346–349.