Production of Polyomavirus Late mRNAs Requires Sequences near the 5' End of the Leader but Does Not Require Leader-to-Leader Splicing

JACQUELINE LANOIX,[†] RICHARD W. TSENG,[‡] AND NICHOLAS H. ACHESON*

Department of Microbiology and Immunology, McGill University, 3775 University Street,
Montreal H3A 2B4, Quebec, Canada

Received 4 April 1991/Accepted 3 June 1991

Polyomavirus late mRNAs contain multiple copies of a 57-nucleotide leader sequence derived by splicing from multigenome-length late transcripts. Inefficient termination of transcription and inefficient polyadenylation allow accumulation of these giant transcripts. In this report, we show that a viable mutant virus, *ins5*, which contains an efficient rabbit β -globin polyadenylation signal, produced late mRNAs whose vast majority contains only one leader. *ins5* virus nevertheless produced as much late mRNA as did wild-type virus and grew as well as did wild-type virus in mouse cells. These results demonstrate that leader-to-leader splicing per se is not required for efficient production of late mRNAs or for efficient virus replication. However, we also found that RNAs lacking critical sequences near the 5' end of the leader did not accumulate as mRNAs and that most late transcripts made during the early part of the late phase, when few late mRNAs are produced, initiated downstream of the 5' end of the leader. These results indicate that a sequence element near the 5' end of the leader is required for proper processing, transport, or stability of late mRNAs and that the control of late mRNA production depends in part on the choice of transcription initiation sites at the late promoter.

Many polyomavirus late mRNAs contain multiple, tandem copies of a 57-nucleotide untranslated leader sequence near their 5' ends (24, 29). However, only one copy of this leader sequence is present on polyomavirus DNA, just downstream of the start sites for late transcription (12, 28). Multiple leaders are generated by splicing of leader minixons spaced one genome length (5.3 kb) apart on giant transcripts of the late strand of polyomavirus DNA (Fig. 1). These multigenome-length transcripts arise by repeated passage of RNA polymerase II around the circular viral genome (1, 2), a result of inefficient termination of transcription on the late DNA strand. Only one-half to two-thirds of RNA polymerases terminate per genome traverse during the late phase of polyomavirus infection (17, 27, 31).

Furthermore, the late polyadenylation signal is inefficiently recognized by the cleavage-polyadenylation machinery in virus-infected cells (2, 23). This leads to the accumulation in the nucleus of multigenome-length RNAs containing uncleaved polyadenylation sites. These giant transcripts can subsequently undergo splicing, which removes the 5.3 kb of RNA between adjacent leaders (Fig. 1). Although up to 12 tandemly arranged leaders have been detected on individual late mRNA molecules, the average number of leaders per mRNA appears to be closer to 3 (17, 24, 29).

We showed that insertion of the efficient rabbit β -globin polyadenylation signal (14) just upstream of the polyomavirus polyadenylation signal led to increased efficiency of cleavage and polyadenylation of polyomavirus L-strand transcripts (22). The resulting viable polyomavirus insertion mutant, which we named *ins5*, produced lower levels of giant read-through RNAs than did wild-type virus (22). We

interpreted this to mean that termination of transcription occurs efficiently on the *ins5* genome downstream of the inserted β -globin polyadenylation signal and that termination is therefore linked to efficient cleavage and polyadenylation of RNAs (see also references 10, 25, and 32).

Adami et al. (4-6) carried out a mutational analysis of the polyomavirus late leader region to determine nucleotide sequence requirements for late mRNA production and virus growth. These authors showed that the minimum length of the late leader required to retain virus viability is approximately 30 nucleotides (nt) but that the nucleotide sequence within this minixen is not critical. Viral genomes with leaders shorter than 30 nt generated late transcripts that were rarely spliced and were rapidly degraded (5). The authors concluded that splicing cannot take place if the ends of the leader minixen are closer than 30 nt apart, perhaps because factors binding to the 3' splice site upstream of the leader and to the 5' splice site downstream of the leader can sterically interfere with each other's function (5). Furthermore, mutations that removed the upstream 3' splice site were unable to rescue mRNA accumulation or virus viability in mutants containing less than the minimum leader length. This result was interpreted to mean that leader-to-leader splicing, which uses the 3' splice site upstream of the leader, is necessary for efficient late mRNA production (6) and, by extension, for efficient virus production.

The model described above proposes that production of multiple-leader mRNAs is essential for expression of late viral proteins and for virus growth. However, our results suggested that *ins5* virus should produce few mRNAs with multiple leaders; yet this virus replicated well and produced high levels of late mRNAs (22). We therefore compared the levels of single- and multiple-leader late mRNAs produced by wild-type polyomavirus and the *ins5* mutant and determined the relative growth capacities of these two virus strains. We found that indeed most of the late mRNAs made by *ins5* virus contained only one leader, although the virus

* Corresponding author.

[†] Present address: Institut du Cancer de Montréal, Montreal H2L 4M1, Quebec, Canada.

[‡] Present address: Pharmingen Inc., San Diego, CA 92121.

FIG. 1. Production of late polyomavirus mRNAs with multiple leaders. Late transcription starts at multiple sites centered about nt 5128 and proceeds clockwise around the circular viral DNA. Transcripts that are not cleaved and polyadenylated at the late poly(A) site can be extended by one or more genome lengths as a result of inefficient termination. A transcript that includes two late poly(A) sites and has been cleaved at the second site is shown. The leader region (nt 5020 to 5076) is shown as an empty rectangle, and the poly(A) site is represented by a filled square. Splicing of adjacent leaders one genome length apart gives rise to mRNAs with multiple tandem leaders near their 5' ends. Shown is VP2 mRNA, which does not undergo further splicing; VP3 and VP1 mRNAs are generated by alternative splicing between the distal leader and a downstream 3' splice site.

grew as well as did wild-type virus. We also found that RNAs that were initiated downstream of the 5' end of the leader and that did not contain multiple copies of the leader did not get processed into stable late mRNAs. Taken together, these findings suggest that sequences near the 5' end of the leader are required for processing, transport, or stability of late mRNAs, but leader-to-leader splicing is not necessary for late mRNA production or virus viability.

MATERIALS AND METHODS

Virus and cells; RNA preparation and DNA-RNA hybridization. Stocks of polyomavirus strain AT3 (27) and insertion mutants *ins5* and *ins6* (22) were grown and titered in primary baby mouse kidney cell cultures. The numbering system that we use for the AT3 strain (27) is adapted from that of Soeda et al. (28) for the A2 strain; nucleotide numbers are virtually identical for these two strains between nucleotides 56 and 5175. Swiss mouse 3T6 cells were infected at approximately 10 PFU per cell, and RNA was labeled with [5-³H]uridine as described previously (31). Nuclear and cytoplasmic poly(A)⁺ RNAs were prepared and hybridized to single-stranded recombinant bacteriophage DNA; the hybrids were treated with RNase T₁, filter bound, and eluted by treatment with S1 nuclease, and the DNA-RNA hybrids were analyzed by gel electrophoresis as described by Tseng and Acheson (31). Bands on fluorograms were quantitated by densitometry, using a Bio-Rad model 620 densitometer. Alternatively, gels were sliced and radioactivity in slices was measured by scintillation counting as described previously (31). Relative molar amounts of different RNA species were calculated by normalizing for the number of uridine residues present in the RNA moiety of each RNA-DNA hybrid.

RESULTS

Experimental approach. Polyomavirus produces three late mRNA species from a common precursor by differential splicing (15, 18). VP2 mRNA is unspliced beyond the most distal leader; splicing removes a 312-nt intron downstream of this leader to produce VP3 mRNA or a 897-nt intron to produce VP1 mRNA. All three mRNA species contain variable numbers of tandem copies of the leader near their 5' ends (17, 24). We analyzed the proportion of VP2 mRNA molecules which contain single or multiple leaders by the approach shown in Fig. 2A. ³H-labeled viral RNAs were hybridized with an excess of a single-stranded DNA fragment which extends upstream of the initiation sites at the late promoter. RNA-DNA hybrids were treated with RNase T₁, bound to nitrocellulose filters via unpaired single-stranded DNA, and released and trimmed with S1 nuclease (31). Hybrid molecules were analyzed by gel electrophoresis. VP2 mRNAs with single leaders will form continuous RNA-DNA hybrids up to the 5' end of the mRNA. However, mRNAs with multiple leaders will form continuous RNA-DNA hybrids only up to the 5' end of the most distal leader, because only a single copy of the leader is present in the DNA; upstream of this leader, the RNA will not be hybridized to DNA. Such hybrids will be cleaved at the 5' end of the most distal leader, giving rise to a protected RNA-DNA hybrid which has been truncated at this position.

Interpretation of results is complicated by the presence of multiple transcription start sites at the polyomavirus late promoter (11, 30). As previously shown, polyomavirus strain AT3 uses seven major start sites, spaced approximately 25 nt apart over a 150-nt region (31). RNAs with 5' ends at these positions are diagrammed in Fig. 2B, and RNA-DNA hybrids generated from pulse-labeled nuclear RNA extend to these seven 5' ends (Fig. 3; 22, 31). VP2 mRNAs that contain a single copy of the leader should give rise to RNA-DNA hybrids whose ends correspond to these seven start sites. In contrast, all hybrids generated by multiple-leader VP2 mRNAs will migrate at the position of species 2, which corresponds to the 5' end of the late leader (nt 5076). Since site 2 also corresponds to a transcription start site (11, 30), hybrids which migrate at the position of species 2 could represent either multiple-leader mRNAs or single-leader mRNAs whose 5' end is at nt 5076.

Single-leader mRNAs made by wild-type or *ins5* virus. Figure 3 shows the results obtained when we analyzed two different preparations of cytoplasmic, poly(A)⁺ RNA isolated from cells infected with either wild-type virus or mutant *ins5*. RNAs were labeled with [³H]uridine for 2.25 or 3.5 h approximately 30 h after infection. Wild-type VP2 mRNAs generated a major hybrid band at position 2, as well as less intense bands at positions 3 to 7. Quantitation of these species (Table 1) confirmed that 50 to 70% of the hybrids were in band 2. Since this band could be generated either by multiple-leader mRNAs or by single-leader mRNAs whose 5' ends are at nt 5076, we cannot determine directly the frequency of multiple-leader mRNAs from these data. However, we can conclude that a minimum of 30 to 50% of wild-type VP2 mRNAs contain a single leader (sum of bands 3 to 7). Interestingly, hybrids in band 1 were at or below the detection level (approximately 2% of hybrids); this band represents an RNA species that does not have multiple leaders and whose 5' end lies within the leader region, near nt 5052.

ins5 mRNAs generated a similar pattern of hybrid bands, but band 2 was no longer the major species. In two experi-

FIG. 2. (A) Strategy used to quantitate amounts of single-leader VP2 mRNAs produced; (B) map of 5' ends of late polyomavirus RNAs. Single-stranded (ss) DNA containing the L strand of *HpaII* fragment 3 of polyomavirus (see map at bottom of panel B) is hybridized to labeled RNA, and DNA-RNA hybrids are trimmed with RNase and S1 nuclease. RNAs containing a single copy of the leader form continuous hybrids up to the 5' end of the RNA, giving rise to RNA-DNA duplexes of different sizes denoted by numbers 1 to 7 in panel B. RNAs containing multiple copies of the leader form hybrids which become discontinuous at the 5' end of the most distal copy of the leader, giving rise to RNA-DNA duplexes of uniform size located in band 2 (see panel B). Leaders on RNAs are denoted by rectangles. VP1 and VP3 mRNAs are not scored by this experiment, for the RNA-DNA hybrids that they form are cleaved by nuclease treatment at the 3' end of the distal leader, where single-stranded DNA is looped out as a result of discontinuity between the spliced RNA and the DNA. RT denotes hybrids formed by read-through transcripts which have previously traversed the genome one or more times before this passage of the late promoter region. These read-through transcripts are present only in the nucleus and therefore are not detected in preparations of cytoplasmic mRNAs.

FIG. 3. Relative amounts of single- and multiple-leader VP2 mRNAs made by wild-type (wt) or *ins5* virus. Cytoplasmic polyadenylated RNA was labeled with [3 H]uridine for 2.25 h (mRNA 1) or 3.5 h (mRNA 2); nuclear RNA was labeled for 2 min. RNA was hybridized to the *HpaII* fragment shown in Fig. 2B, nuclease-treated RNA-DNA hybrids were analyzed by electrophoresis, and gels were dried and subjected to fluorography. 3 H-labeled DNA fragments were included as size markers; sizes (in nucleotides) of the *HpaII* fragments are shown on the left. Positions of hybrids corresponding to read-through RNA (RT) (present only in nuclear RNA) and to RNAs with 5' ends at the seven major transcription start sites are shown on the right.

ments, band 2 accounted for one-fourth or one-fifth of the total molar amount of hybrids present. This means that at least 75 to 80% of the VP2 mRNAs made by *ins5* contained only one copy of the leader sequence. Again, since part of the hybrids in band 2 could be generated by single-leader RNAs with 5' ends at nt 5076, this is a minimum estimate of the proportion of single-leader RNAs in the population. If it is assumed that single-leader mRNAs with this 5' end are made in proportion to the relative rate of usage of this start site in nuclear RNA (that is, approximately the same level as RNAs with 5' ends at site 3 [nt 5100]; Fig. 3 and Table 1), we can estimate that about two-thirds of the hybrids in band 2 represent single-leader mRNAs and only one-third represent multiple-leader mRNAs. This would mean that greater than 90% of *ins5* VP2 mRNAs contain a single leader.

Growth of *ins5* virus. Given that *ins5* produces predominantly single-leader VP2 mRNAs (and, presumably, VP1 and VP3 mRNAs), we wanted to determine whether this virus replicates as well as does wild-type virus, which produces many late mRNAs with multiple leaders. We therefore infected mouse 3T6 cells or primary baby mouse kidney cells with a high multiplicity (10 to 20 PFU per cell) of either wild-type AT3 virus, *ins5* virus, or pseudo-wild-type *ins6* virus (which contains the same rabbit β -globin insert as does *ins5* but in the reverse orientation). Infectious virus yield was measured at times varying from 6 to 150 h postinfection. Maximum yields were achieved between 48 and 72 h postinfection. Table 2 summarizes the results of five experiments in which growth of *ins5* and *ins6* was compared. *ins5* replicated as well as or better than *ins6* virus in both cell

TABLE 1. Distribution of late RNAs with different 5' ends^a

RNA-DNA hybrid band	% of RNA species made by:					
	Wild-type virus			<i>ins5</i> virus		
	Nuclear RNA	VP2 mRNA 1 ^b	VP2 mRNA 2	Nuclear RNA	VP2 mRNA 1	VP2 mRNA 2
7	7	8	5	8	11	19
6	3	4	— ^c	4	8	10
5	5	7	—	8	10	9
4	26	17	17	29	24	28
3	18	11	9	16	17	13
2	18	52	69	17	25	19
1	23	2	—	18	4	2

^a Fluorograms shown in Fig. 3 and those from similar experiments were subjected to densitometry. For each RNA sample, optical densities in bands corresponding to the seven major late RNA 5' ends were expressed as a percentage of the sum of optical densities in all seven bands.

^b mRNA 1 and mRNA 2 refer to two independent experiments.

^c —, not detectable.

types. Yields of unsubstituted AT3 strain wild-type virus were comparable with those obtained for both *ins5* and *ins6* (data not shown). We conclude that the ability of polyomavirus to replicate in cultured cells is insensitive to changes in the relative levels of single-leader and multiple-leader late mRNAs produced.

Changes in transcription start sites during the late phase. A striking observation from the experiments shown in Fig. 3 and Table 1 is that no single-leader VP2 mRNAs with 5' ends at nt 5052 (RNA species 1), in the middle of the leader region, could be detected. However, many nuclear mRNA precursors with this 5' end were made in cells infected with either wild-type or *ins5* virus (Fig. 3). Late mRNAs with this 5' end have nevertheless been detected in polyomavirus-infected cells (15, 29, 30); we conclude that those mRNAs must have contained multiple leader sequences. This implies that some sequence element upstream of nt 5052, presumably within the 5' half of the leader or just upstream of the leader, is required for correct processing or transport of mRNA or for RNA stability.

These considerations led us to ask whether the production of late mRNAs might be controlled by the choice of transcription start sites on the late strand. It is known that late-strand transcripts are made in the nucleus during the early phase of infection, or in the presence of inhibitors of DNA replication, but little or no late mRNAs are produced under these conditions (3, 7, 16). It occurred to us that late-strand RNAs made during the early phase might not be efficiently processed into mRNA if their transcription initiated mostly at site 1, within the leader region, and if

predominantly single-leader RNAs were produced. We therefore examined the 5' ends of pulse-labeled nuclear RNA at earlier times in the infectious cycle to determine the distribution of transcription start sites in this RNA. Unfortunately, levels of viral transcription during the early phase (before 12 h postinfection) are so low that insufficient [³H] uridine is incorporated during short pulse-labeling periods to allow analysis of nuclear RNAs by hybridization to unlabeled probes. We therefore analyzed pulse-labeled nuclear RNAs made during the early part of the late phase (16 to 18 h postinfection) and compared them with RNAs made well into the late phase (28 to 30 h postinfection).

Figure 4B shows that the major 5' end in pulse-labeled nuclear RNA at 28 h postinfection was at site 4 (nt 5128; 35% of all 5' ends in this experiment), as already documented (30, 31; Fig. 3). This places the major transcription start site at that location. However, at 16 h postinfection (Fig. 4A), site 1 (nt 5052) was the major transcription start site, accounting for nearly 60% of all 5' ends in this experiment (site 4 accounted for only 12% of 5' ends). At intermediate times, the relative level of 5' ends at site 1 was between the values determined at 16 and 28 h (data not shown). These experiments demonstrate that most RNA polymerases initiate transcription at site 1 during the early part of the late phase and that initiation at site 4 becomes predominant only later during the viral infectious cycle. We conclude that the production of late polyomavirus mRNAs is controlled, at least in part, by the choice of transcriptional start site and that this choice changes as the virus cycle progresses.

DISCUSSION

We have shown that 30 to 50% of late VP2 mRNAs produced by wild-type polyomavirus, and at least 80% of VP2 mRNAs produced by mutant *ins5*, contain only one leader. Nevertheless, *ins5* produces as much late mRNA as does wild-type virus and grows at least as well as does wild-type virus in cultured mouse cells. We conclude that leader-to-leader splicing is not required for efficient late mRNA production or for efficient growth of polyomavirus.

In contrast, Adami et al. (6) concluded, on the basis of the study of deletion mutants of polyomavirus, that leader-to-leader splicing is required for efficient production and accumulation of late viral mRNAs, and thus for virus viability. They suggested that the requirement for a minimum length of the late leader minixon could be explained by steric hin-

TABLE 2. Maximum yields of *ins5* and pseudo-wild-type *ins6* virus grown in cultured mouse cells^a

Expt no.	Cell type	Harvest time (h)	Virus yield (PFU/ml)	
			<i>ins5</i>	<i>ins6</i>
1	3T6	48	3 × 10 ⁷	2 × 10 ⁷
2	3T6	72	8 × 10 ⁶	2 × 10 ⁷
3	BMK	48	1 × 10 ⁹	5 × 10 ⁸
4	BMK	72	2 × 10 ⁹	8 × 10 ⁸
5	BMK	55	7 × 10 ⁸	5 × 10 ⁸

^a Approximately 10⁷ cells per 88-mm petri dish were infected with 10 to 20 PFU/cell and incubated at 37°C in 10 ml of medium plus 2% calf serum. After harvest, virus infectivity was determined by plaque assays in primary baby mouse kidney (BMK) cells.

FIG. 4. Change in major transcription start site as a function of time after infection. Nuclear RNA labeled for 2 min with [^3H]uridine 16 h (A) or 28 h (B) after infection with wild-type virus was hybridized, and nuclease-treated hybrids were analyzed by gel electrophoresis as for Fig. 3. Gel lanes were sliced into 1-mm sections, and radioactivity was determined by scintillation counting. Only that portion of the gel containing hybrids which correspond to the seven transcription start sites is shown. Numbers above the peaks represent the different RNA species shown in Fig. 2 and 3 and Table 1.

drance between factors which bind to the 3' splice site upstream of the leader and to the 5' splice site downstream of the leader (5). This led them to postulate that elimination of the upstream 3' splice site would rescue mutants which contain less than the minimum leader length, because it would eliminate this steric hindrance effect (6). These double mutants were nevertheless nonviable and produced unstable mRNA precursors. This apparent requirement for the 3' splice site suggested to them that leader-to-leader splicing was necessary for proper mRNA processing (6).

A different interpretation of the results of Adami et al. emerges in the light of the data presented here. Even though many primary transcripts are initiated at site 1, within the leader sequence, little or no single-leader mRNAs with this 5' end accumulate in infected cells. We propose that sequences near the proximal end of the late leader are required for correct RNA processing, transport, or stability; these sequences are absent from transcripts that begin at site 1 and do not read through the leader region on subsequent rounds of transcription, but are present in transcripts that read through the leader region after traversing the entire genome. The critical sequences are also absent from all transcripts produced by the double-deletion mutant constructed by Adami et al. (6), therefore explaining the inability of these mutants to produce stable late mRNAs.

Since it was shown that most of the leader minixon can be substituted with unrelated sequences without any effect on mRNA production or virus viability (5), we conclude that the sequences required for mRNA accumulation lie at or just upstream of the 5' end of the leader. The required sequences may, in fact, be part of the 3' splice site at that position;

however, our results show that RNA splicing per se is not required, for single-leader VP2 mRNAs undergo no splicing but clearly accumulate as the majority species in *ins5*-infected cells. It may be that assembly of splicing factors at this 3' splice site facilitates entry of mRNA precursors into the proper processing pathway or stabilizes the precursors during processing or transport to the cytoplasm (e.g., see reference 26). It may also be that the required sequences have nothing to do with splicing. Further mutational analysis of the sequence requirements upstream of the leader, in the presence of wild-type leader sequences, will be necessary to resolve this question.

We note that if the sequences required for proper RNA processing or stability are in fact upstream of the leader, RNAs that are initiated at site 2, which is very close to the 5' end of the leader (nt 5076), should also not accumulate as single-leader species. Our experiments could not distinguish between multiple-leader species and single-leader species whose 5' ends lie at this site (Fig. 3); therefore, we do not know what proportion of the RNA-DNA hybrids in band 2 detected from wild-type- or *ins5*-infected cells were generated by single-leader mRNAs.

Hyde-deRuyscher and Carmichael (17) recently showed that single-leader VP1, VP2, and VP3 mRNAs are underrepresented in the cytoplasm of polyomavirus-infected cells, based on predictions from a simple model of transcription termination and RNA processing which we had previously outlined in detail (2). This model predicts that the percentage of single-leader late mRNAs in the total late mRNA population should be equivalent to the percent termination efficiency per genome traverse by RNA polymerase II or, alternatively, to the percent efficiency of cleavage and polyadenylation at the late polyadenylation site. In one experiment, their quantitation of single-leader mRNAs showed a fivefold-lower accumulation than predicted, although data from a second experiment reveal less than a twofold deficit when a similar calculation is done. Single-leader mRNAs thus appear to be produced at a lower rate, or are less stable, than multiple-leader RNAs by a factor of 2 to 5. Hyde-deRuyscher and Carmichael proposed that transcripts containing only a single leader are preferentially degraded in the nucleus. Despite this finding, our results clearly establish that *ins5* virus can produce sufficient amounts of late mRNAs to replicate as well as wild-type virus even when 80% or more of those mRNAs contain only one leader. Therefore, the deficit in accumulation of single-leader mRNAs observed in their experiments cannot by itself explain the nonviability of their deletion mutants described above (5, 6).

The absence of single-leader mRNAs with 5' ends at site 1 led us to examine the distribution of transcription start sites on late-strand RNAs during the early part of the late phase of virus replication. We found that the majority of late-strand transcripts made 16 h after infection was initiated at site 1. In contrast, by 28 h after infection, the major initiation site had shifted to site 4, 75 nt upstream. This suggests a possible mechanism for the control of late mRNA production. It is known that the late promoter is active during the early phase of the virus growth cycle, or in the presence of inhibitors of DNA replication, but very few late mRNAs accumulate during this period (3, 7, 16). In addition, late genes are not efficiently expressed in many polyomavirus-transformed cell lines, even though the late promoter region is intact (7, 13). If most L-strand RNAs made under these conditions are initiated within the leader region, and if termination or cleavage-polyadenylation of L-strand transcripts is efficient,

then most of the precursor RNAs made would not be able to generate late mRNAs, as they would not contain multiple copies of the leader and would be lacking the sequences near the proximal end of the leader required for correct processing or stability. As the late phase (signalled by the beginning of viral DNA replication) is entered and progresses, more and more late transcripts are initiated at upstream sites, especially site 4. These transcripts contain the critical leader-proximal sequences and could therefore give rise to stable mRNAs. In addition, reduction in the efficiency of termination or cleavage-polyadenylation would give rise to multigenome-length transcripts which are processed to mRNAs with multiple leaders. Multigenome-length transcripts contain the critical leader-proximal sequences even if initiated downstream of the end of the leader; additionally, their increased stability (17) would allow them to accumulate more readily.

This model differs critically from that proposed by Kern et al. (21) and Hyde-deRuyscher and Carmichael (17), who suggested that late transcripts made in transformed cells, during the early phase of productive infection, or in the presence of inhibitors of DNA replication are unable to be processed into mRNA because they contain only one leader, and that leader-to-leader splicing is necessary for the accumulation of late mRNA species. Results from both this report and that of Hyde-deRuyscher et al. (17) show that single-leader late mRNAs do in fact accumulate and are functional; we therefore propose that the lack of production of late mRNAs during the early phase is a direct result of initiation of transcription downstream of the critical leader-proximal sequences, combined with production of single-leader transcripts by efficient termination or cleavage-polyadenylation (or, in the case of transformed cells, by the absence of tandem copies of integrated viral DNA; see reference 21).

Other groups have examined the activity of the polyomavirus late promoter in cells transiently infected (8, 9) or stably transformed (19, 20) with chimeric plasmids. Deletion analysis showed that a fragment of polyomavirus DNA extending from nt 5130 (just upstream of the major late transcription start site at 5128) to 5022 (just upstream of the 3' end of the leader region) has substantial promoter activity (8, 20). However, deletion of the proximal 34 nt (5130 to 5096) resulted in a 10- to 100-fold drop in promoter activity (8, 9). Interestingly, addition of nucleotides 5021 to 4900 to the basal fragment, which restores the 3' end of the leader region and the 5' splice site just downstream, increased promoter activity by a factor of 10 (8). These results indicate that elements downstream of the major transcription start sites are important in late promoter activity; it is not clear whether these elements directly affect transcription or may be involved in RNA stability or transport. In no case were sequences at the 3' splice site just upstream of the leader specifically deleted to test their effect on mRNA production.

Some of the constructs used in the studies discussed above produced mRNAs with 5' ends within the late leader region (19, 20). Our results suggest that mRNAs that do not contain multiple leaders should not survive processing or transport unless their 5' ends lie upstream of the leader region. Interestingly, constructs that generated mRNAs whose 5' ends were within the leader did not include the 5' splice site immediately downstream of the leader (19, 20). However, constructs that did include sequences downstream of the leader generated mRNAs whose 5' ends were at the major start site near nucleotide 5128, some 50 nt upstream of the leader (8). These differences mirror the

differences that we found in transcription start site choice between the early and the late part of the late phase in productively infected cells. They suggest that usage of start sites upstream of the leader may depend on the presence of sequences downstream of the leader and that the requirement we suggest here for sequences upstream of the leader to allow correct processing or transport of mRNA may not be necessary if sequences immediately downstream of the leader (perhaps the 5' splice site) are not present.

ACKNOWLEDGMENTS

We thank Lynn Kelly for excellent technical assistance and Z. Ali-Khan for use of his densitometer.

This work was supported by the Medical Research Council of Canada (grant MA-7281). J. Lanoix was supported by the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (Quebec), by a McGill University McConnell studentship, and by a McGill University Faculty of Medicine studentship.

REFERENCES

1. Acheson, N. H. 1978. Polyoma virus giant RNAs contain tandem repeats of the nucleotide sequence of the entire viral genome. *Proc. Natl. Acad. Sci. USA* **75**:4745-4748.
2. Acheson, N. H. 1984. Kinetics and efficiency of polyadenylation of late polyomavirus nuclear RNA: generation of oligomeric polyadenylated RNAs and their processing into mRNA. *Mol. Cell. Biol.* **4**:722-729.
3. Acheson, N. H., and F. Miéville. 1978. Extent of transcription of the E strand of polyomavirus DNA during the early phase of productive infection. *J. Virol.* **28**:885-894.
4. Adami, G. R., and G. G. Carmichael. 1986. Polyomavirus late leader region serves an essential spacer function necessary for viability and late gene expression. *J. Virol.* **58**:417-425.
5. Adami, G. R., and G. G. Carmichael. 1987. The length but not the sequence of the polyoma virus late leader exon is important for both late RNA splicing and stability. *Nucleic Acids Res.* **15**:2593-2610.
6. Adami, G. R., C. W. Marlor, N. L. Barrett, and G. G. Carmichael. 1989. Leader-to-leader splicing is required for the efficient production and accumulation of polyomavirus late mRNAs. *J. Virol.* **63**:85-93.
7. Beard, P., N. H. Acheson, and I. H. Maxwell. 1976. Strand-specific transcription of polyomavirus DNA early in productive infection and in transformed cells. *J. Virol.* **17**:20-26.
8. Bourachot, B., M. Yaniv, and P. Herbolme. 1989. Control elements situated downstream of the major transcriptional start site are sufficient for highly efficient polyomavirus late transcription. *J. Virol.* **63**:2567-2577.
9. Cahill, K. B., and G. G. Carmichael. 1989. Deletion analysis of the polyomavirus late promoter: evidence for both positive and negative elements in the absence of early proteins. *J. Virol.* **63**:3634-3642.
10. Connelly, S., and J. L. Manley. 1988. A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev.* **2**:440-452.
11. 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.* **23**:6305-6322.
12. Deininger, P., A. Esty, P. LaPorte, and T. Friedmann. 1979. Nucleotide sequence and genetic organization of the polyoma late region: features common to the polyoma early region and SV40. *Cell* **18**:771-779.
13. Fenton, R. G., and C. Basilico. 1981. Viral gene expression in polyomavirus-transformed cells and their cured revertants. *J. Virol.* **40**:150-163.
14. Gil, A., and N. J. Proudfoot. 1984. A sequence downstream of AAUAAA is required for rabbit β -globin mRNA 3'-end formation. *Nature (London)* **312**:473-474.

15. Heiser, W. C., and W. Eckhart. 1982. Polyomavirus early and late mRNAs in productively infected mouse 3T6 cells. *J. Virol.* **44**:175–188.
16. Hyde-deRuyscher, R. P., and G. G. Carmichael. 1988. Polyomavirus early-late switch is not regulated at the level of transcription initiation and is associated with changes in RNA processing. *Proc. Natl. Acad. Sci. USA* **85**:8993–8997.
17. Hyde-deRuyscher, R. P., and G. G. Carmichael. 1990. Polyomavirus late pre-mRNA processing: DNA replication-associated changes in leader exon multiplicity suggest a role for leader-to-leader splicing in the early-late switch. *J. Virol.* **64**:5823–5832.
18. Kamen, R., J. Favaloro, and J. Parker. 1980. Topography of the three late mRNAs of polyomavirus which encode the virion proteins. *J. Virol.* **33**:637–651.
19. Kern, F. G., and C. Basilico. 1985. Transcription from the polyoma late promoter in cells stably transformed by chimeric plasmids. *Mol. Cell. Biol.* **5**:797–807.
20. Kern, F. G., L. Dailey, and C. Basilico. 1985. Common regulatory elements control gene expression from polyoma early and late promoters in cells transformed by chimeric plasmids. *Mol. Cell. Biol.* **5**:2070–2079.
21. Kern, F. G., P. Delli Bovi, and C. Basilico. 1987. A reiterated leader sequence is present in polyomavirus late transcripts produced by a transformed rat cell line. *J. Virol.* **61**:4055–4059.
22. Lanoix, J., and N. H. Acheson. 1988. A rabbit β -globin polyadenylation signal directs efficient termination of transcription of polyomavirus DNA. *EMBO J.* **7**:2515–2522.
23. Lanoix, J., R. W. Tseng, and N. H. Acheson. 1986. Duplication of functional polyadenylation signals in polyomavirus DNA does not alter efficiency of polyadenylation or transcription termination. *J. Virol.* **58**:733–742.
24. Legon, S., A. J. Flavell, A. Cowie, and R. Kamen. 1979. Amplification in the leader sequence of late polyoma virus mRNAs. *Cell* **16**:373–388.
25. Logan, J., E. Falck-Pedersen, J. E. Darnell, and T. Shenk. 1987. A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse β -globin gene. *Proc. Natl. Acad. Sci. USA* **84**:8306–8310.
26. Ryu, W.-S., and J. E. Mertz. 1989. Simian virus 40 late transcripts lacking excisable intervening sequences are defective in both stability in the nucleus and transport to the cytoplasm. *J. Virol.* **63**:4386–4394.
27. Skarnes, W. C., D. C. Tessier, and N. H. Acheson. 1988. RNA polymerases stall and/or prematurely terminate nearby both early and late promoters on polyomavirus DNA. *J. Mol. Biol.* **203**:153–171.
28. Soeda, E., J. R. 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.
29. Treisman, R. 1980. Characterization of polyoma late mRNA leader sequences by molecular cloning and DNA sequence analysis. *Nucleic Acids Res.* **8**:4867–4888.
30. Treisman, R., and R. Kamen. 1981. Structure of polyoma virus late nuclear RNA. *J. Mol. Biol.* **148**:273–301.
31. Tseng, R. W., and N. H. Acheson. 1986. Use of a novel S1 nuclease RNA-mapping technique to measure efficiency of transcription termination on polyomavirus DNA. *Mol. Cell. Biol.* **6**:1624–1632.
32. Whitelaw, E., and N. Proudfoot. 1986. α -Thalassemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3'-end processing in the human α -2 globin gene. *EMBO J.* **5**:2915–2922.