MOLECULAR AND CELLULAR BIOLOGY, Oct. 1991, p. 5291–5300
0270-7306/91/105291-10$02.00/0

Vol. 11, No. 10

Splice Site Choice in a Complex Transcription Unit Containing Multiple Inefficient Polyadenylation Signals

YING LUO AND GORDON G. CARMICHAEL*

Department of Microbiology, University of Connecticut Health Center, Farmington, Connecticut 06030

Received 13 December 1990/Accepted 24 July 1991

The relationship between polyadenylation and splicing was investigated in a model system consisting of two tandem but nonidentical polyomavirus late transcription units. This model system exploits the polyomavirus late transcription termination and polyadenylation signals, which are sufficiently weak to allow the production of many multigenome-length primary transcripts with repeating introns, exons, and poly(A) sites. This double-genome construct contains exons of two types, those bordered by 3′ and 5′ splice sites (L1 and L2) and those bordered by a 3′ splice site and a poly(A) site (V1 and V2). The L1 and L2 exons are distinguishable from one another but retain identical flanking RNA processing signals, as is the case for the V1 and V2 exons. Analysis of cytoplasmic RNAs obtained from mouse cells transfected with this construct and its derivatives revealed the following. (i) V1 and V2 exons are often skipped during pre-mRNA processing, while L1 and L2 exons are not skipped. (ii) No messages contain internal, unused polyadenylation signals. (iii) Poly(A) site choice is not required for the selection of an upstream 3′ splice site. (iv) When two tandem poly(A) sites are placed downstream of a 3′ splice site, the first poly(A) site is chosen almost exclusively, even though transcription can proceed past both sites. (v) Placing a 3′ splice site between these two tandem poly(A) sites allows the more distal site to be chosen. These and other available data are most consistent with a model in which terminal exons are produced by the coordinate selection and use of a 3′ splice site with the nearest available downstream poly(A) site.

In a number of recent studies involving different systems, evidence has been presented that alternate poly(A) site selection can be associated with different 3′ splice site choices (5, 21, 22, 29, 38; see references 39, 42, and 53 for reviews). These studies have most often been done with complex transcription units which have multiple exons, introns, and poly(A) signals. In such systems, alternative pathways of pre-mRNA processing involve either cell-specific (9, 10, 21) or temporal (23, 42, 52) regulation of poly(A) site or splice site selection. One example of cell-specific regulation is the switch between secreted and membrane forms of the immunoglobulin M μ heavy chain. Studies have suggested that this switch is regulated at the level of poly(A) site choice (21, 29, 34, 41, 49). Thus, in the immunoglobulin M system, choice of alternative poly(A) signals may lead to altered splicing pathways. Some recent results, however, have been inconsistent with such a model of poly(A) site dominance and have instead suggested that there may be competition between splicing and polyadenylation (44, 45). In contrast, the expression of calcitonin and the calcitonin-gene related peptide, also controlled in a cell-specific manner from a complex transcription unit containing two poly(A) sites, clearly appears to be regulated at the step of splice site recognition, which in turn leads to the use of the nearest downstream poly(A) signal (10, 22, 38).

An example of temporal regulation of pre-mRNA processing occurs in the human adenovirus major late transcription unit (8, 42, 52). At different times after infection, different mRNA species arising from the same transcriptional promoter are seen. Late in infection there are five distinct late mRNA families (8, 23, 25, 52). Each family contains three to seven mRNAs, due to the use of different 3′ splice sites (12, 19, 42), and each family is associated with the use of a different late poly(A) site (L1 to L5). This differs from early in infection, when only the L1 poly(A) site is used efficiently (8, 52).

We have studied the relationship between splice site and poly(A) site choice in a model system based on the polyomavirus late transcription unit. Like the systems described above, polyomavirus late primary transcripts are complex, with pre-mRNA processing involving the choice of alternative splice sites and poly(A) sites. The major difference between this system and others is that the polyomavirus late splice sites and poly(A) sites chosen during processing are identical to those that are skipped. Therefore, this is a simpler system, as differences in results cannot be ascribed to processing signals of different sequence.

Inefficient polyomavirus late-transcription termination and polyadenylation leads to the production of a heterogeneous collection of late primary transcripts in infected cells, some of which are many times the size of the viral genome (1–3, 56, 57). These transcripts contain two types of exons, noncoding (leader) exons and message body exons. A leader exon occurs at the 5′ end of every late transcript. In multigenome-length transcripts, this exon appears multiple times. During RNA processing, leaders splice to each other, skipping message body splice sites and removing genome-length introns. In addition, a message body 3′ splice site is chosen to attach the final late leader exon to a terminal coding exon. Thus, the 5′ ends of mature late-cytoplasmic messages contain between 1 and 12 tandem copies of the late leader exon, followed by a single coding exon (32, 33, 40, 55). We have studied a series of double-genome constructs designed to produce late viral pre-mRNAs with multiple, identical poly(A) sites but having alternating, distinguishable leader and body exons. Following transfection of these into mouse cells, we have analyzed cytoplasmic RNA and determined which exons splice to which others and which poly(A) sites are chosen. Results from these experiments reveal that,

* Corresponding author.

in this system, splice site choice dominates over poly(A) site choice, with poly(A) site selection in vivo appearing to be coupled to the selection and use of an upstream 3′ splice site.

MATERIALS AND METHODS

Materials. Restriction enzymes, T4 DNA ligase, DNA polymerase I, large fragment (Klenow enzyme), T4 DNA polymerase, and T4 polynucleotide kinase were from New England BioLabs and were used as suggested by the manufacturer. RNase T2 was from Bethesda Research Laboratories. [α-^{32}P]dATP, [α-^{32}P]UTP, [γ-^{32}P]ATP, and avian myeloblastosis virus reverse transcriptase were from New England Nuclear. Polyomavirus strain 59RA has been described (24, 28). The nucleotide sequence of this strain is very similar to that published for strain A3 (20). Plasmid p43.25.67 is wild-type polyomavirus strain A2 with an *Xho*I linker at the *Pvu*II site at nucleotide (nt) 5128, inserted at the *Bam*HI site into pAT153 (58), and was generously provided by A. Cowie and R. Kamen. All recombinant plasmids used here were propagated in *Escherichia coli* JM83.

Cell culture and transfections. Mouse NIH 3T3 cells were obtained from the American Type Culture Collection. The techniques for their propagation and transfection have been described elsewhere (14). Briefly, all transfections were performed by using a modification of the calcium-phosphate transfection procedure of Chen and Okayama (16). Cells were seeded at a concentration of 10^6 cells/150-mm plate in Dulbecco's modified Eagle medium supplemented with 10% bovine calf serum, L-glutamine, penicillin, and streptomycin at 37°C in 5% CO_2 approximately 48 h prior to transfection. Sixteen hours prior to transfection, the cells were serum stimulated by the addition of fresh medium. The total amount of DNA used per transfection was 40 μg. Before transfection, recombinant plasmids were cut with *Eco*RI and dilute ligated with T4 DNA ligase as described previously (6).

Harvesting of cytoplasmic RNA. Cells were harvested 44 to 48 h after initiation of transfection. Each plate was rinsed with 1× PBS^- (phosphate-buffered saline [PBS] without Mg^{2+} and Ca^{2+}), and RNA was harvested by the protocol described elsewhere (32). Briefly, cells were scraped into PBS^- and pelleted by centrifugation at 1,000 × *g* for 3 min. The pellet was resuspended in PBS and centrifuged again at 1,000 × *g* for 3 min. The cell pellet was resuspended in 3 ml of Nonidet P-40 lysis buffer and left on ice for 10 min. Nuclei were removed by centrifugation at 1,000 × *g* for 3 min. To the supernatant (the cytoplasmic RNA fraction) was added 1.42 g of guanidine thiocyanate. RNA was separated from DNA by centrifugation through a 5.7 M cesium chloride cushion for 20 h at 110,000 × *g* (18). The pelleted RNAs were drained well and resuspended in 300 μl of 10 mM Tris-Cl, pH 7.0–1 mM EDTA, pH 7.0–0.2% sodium dodecyl sulfate and ethanol precipitated. RNAs were resuspended in 100 μl of double-distilled H_2O.

Construction of plasmids. The structures of the major plasmids used in this study and the construction of probes are shown schematically in Fig. 1, along with the restriction enzyme sites used for their construction.

Plasmid pI-1 was constructed by digesting the BlueScribe vector pBS^+ (Stratagene) with *Bam*HI and *Hin*dIII, blunting the ends with Klenow enzyme, and religating. Plasmid pI-2 was constructed by digesting viral DNA from polyomavirus mutant SLM/MP8R (4) with *Eco*RI and inserting this viral genome into the unique *Eco*RI site of pI-1. Polyomavirus SLM/MP8R is the same as our laboratory wild-type strain, 59RA, except the wild-type late leader exon has been replaced with sequences from a phage M13 cloning vector (4). Plasmid pI-3 was constructed by partial digestion of pI-2 with *Hin*cII, followed by complete digestion with *Eco*RV and religation. This plasmid contains the SLM/MP8R genome with a deletion within the exon coding for VP1. A *Hin*cII site remains in the SLM/MP8R leader. Plasmid pI-6 was constructed by digesting p43.25.67 with *Xho*I and *Xba*I, blunting the ends with Klenow enzyme, and religating. Plasmid pI-13 was constructed by digesting pI-6 with *Bam*HI, electroeluting the polyomavirus fragment from an agarose gel, and inserting this fragment into the unique *Bam*HI site of pI-3. This created a plasmid with two tandem viral genomes but with different leader exons and VP1 exons that can be distinguished. Plasmid pI-5 was constructed by inserting the polyomavirus *Bam*HI p43.25.67 fragment into the *Bam*HI site of the BlueScribe vector pBS^+. Plasmid pI-10 was constructed by digesting pI-5 with *Apa*I and *Xho*I, blunting the ends with T4 DNA polymerase, and religating. This step removed the polyomavirus late polyadenylation site, early coding sequences, and the origin of replication. Plasmid pI-11 was constructed by digesting pI-10 with *Bam*HI, electroeluting the polyomavirus fragment from an agarose gel, and inserting this fragment into the unique *Bam*HI site of pI-3. Plasmid pI-11SmΔ was constructed in the same way as pI-11, except the polyomavirus fragment inserted into pI-3 was from pXX10 (4a), which contains a 26-bp deletion spanning the late poly(A) site (nt 2910 to 2935). In addition, pXX10 carries two base changes which create a *Hin*dIII site at nt 2941 to 2946. Plasmid pI-15 was constructed by digesting plasmid p43.25.67 with *Kpn*I and religating the large fragment. This removed the early coding region, the replication origin, and the late region from the promoter through the VP3 3′ splice site. The VP1 exon remained intact. Plasmid pI-16 was constructed by digesting pI-15 with *Bam*HI, electroeluting the polyomavirus fragment from an agarose gel, and inserting this fragment into the unique *Bam*HI site of pI-3. Plasmid pI-17 was constructed by inserting the 1.6-kb *Eco*RV-*Xba*I fragment (nt 4137 to 2549), after blunting, into the unique *Eco*RV site of pPyBS, which is polyomavirus strain 59RA inserted into the *Bam*HI site of pBS^+.

Riboprobe preparation and T2 ribonuclease protection assays. All riboprobes were made from cones using in vitro transcription with phage T3 or T7 RNA polymerase and [^{32}P]UTP (11). Plasmid pPP1 was made by cloning the *Pst*I-*Xba*I fragment from pI-3 into pBS^+. The riboprobe PP1 measures both the splicing and the 3′-end structures surrounding the V2 exon. Plasmid pPP2 was made by cloning the wild-type polyomavirus *Pst*I-*Hin*dIII fragment from nt 3949 to 4256 into pBS^+. This is the same as a VP1 splicing probe previously described (7). Plasmid pPP3 was made by taking plasmid pI-10, digesting with *Acc*I, and religating. Riboprobe PP3 is colinear with sequences from pI-11 and examines the 3′-end structures of pI-11 messages containing the VP1 splice. Plasmid pPP4 was made by cloning the wild-type polyomavirus *Pst*I-*Xba*I fragment from nt 2549 to 3341 into pBS^+ (plasmid pBSXP1). Riboprobe PP4 measures poly(A) site use downstream of either the wild-type VP1 or truncated V2 exons.

Labeled riboprobes were annealed to cytoplasmic or nuclear RNA at 57°C for 15 to 20 h. The resulting hybrids were digested with T2 RNase at a final concentration of 60 U/ml at 37°C for 2 h. After phenol extraction and ethanol precipitation, the samples were dissolved in denaturing dye and

FIG. 1. (A) Restriction sites used for plasmid or probe construction. A hypothetical duplication of the polyomavirus genome is depicted at the top, along with the positions of the late promoters (P_L) and the borders of the coding exons for the virion structural proteins VP1 and VP3. Restriction enzyme cutting sites used during cloning are shown, but distances are not to scale. The left-hand genome represents strain SLM/MP8R (4), and the right-hand genome represents mutant 43.25.67 (58). The numbering system used in this paper can be converted to that of the A3 strain (20) by subtracting six nucleotides. The SLM leader exon (51 bp) is shown as a solid box, and the wild-type (WT) leader exon (57 bp) is the white box. Thick solid bars denote sequences deleted in the construction of plasmids pI-13, pI-11, pI-11SmΔ, pI-16, and pI-17. In the final clones, the SLM/MP8 and wild-type leader exons are renamed L1 and L2, respectively, and the full-length and truncated VP1 exons are renamed V1 and V2, respectively. (B) Probes used for ribonuclease protection assays. Dotted lines denote sequences deleted in the construction of the truncated V2 exon and probes PP1 and PP3. PP1 is colinear with the V2 exon and its flanking sequences. PP3 is colinear with sequences in pI-11. (C) Schematic diagram of the major constructs used in this study. Before transfection, each was digested with *Eco*RI, which liberated the polyomavirus sequences from the plasmid backbone, and recircularized by using T4 DNA ligase. pI-13 contains two copies of the polyomavirus late region, but the first copy has a substituted late leader exon sequence (L1) (6), while the second has a truncated VP1 coding exon (V2). pI-11 is the same as pI-13, except for a 771-bp deletion spanning the first poly(A) site, with 441 nt downstream of the AAUAAA signal removed. pI-16 is the same as pI-13, except for deletion of sequences including one intact early coding region, replication origin, late promoter, and late leader exon. pI-17 is wild-type polyomavirus with a 1.6-kb duplication of sequences surrounding the late poly(A) site but containing no additional 3′ splice site.

protected bands were resolved on 6% polyacrylamide–urea sequencing gels.

PCR assays. Oligonucleotide (oligo) primers were made by using a Biosearch Cyclone DNA Synthesizer and the phosphoramidite chemistry. Oligo 275 (5′-TATCACCGTACAG CCTTG) anneals across the VP1-late leader splice junction, with the terminal three bases annealing to the late leader and the rest annealing to the VP1 coding exon. Oligo 278 (5′-CCTGACATTTTCTATTTTAAG) anneals to late cDNA molecules immediately upstream of the late leader exon. Oligo 279 (5′-TCCTAGATGAAAATGGAG) anneals to the truncated V2 exon but not to the wild-type V1 exon.

The polymerase chain reaction (PCR) (51) was performed following reverse transcription of the RNA. Oligo 275 or 279 was added to the RNA in a buffer recommended by Perkin Elmer-Cetus. Avian myeloblastosis virus reverse transcriptase (22 U) was added for 1 h at 45°C. *Taq*I DNA polymerase (Amplitaq; Perkin Elmer-Cetus) and oligo 278 were then added, and PCR was continued for 30 cycles. Each cycle consisted of 2 min at 94°C, 1 min at 45°C, and 1.5

min at 72°C. A 10-min incubation at 72°C followed the final cycle. Reaction products were resolved on 6% polyacrylamide–7 M urea gels, and bands were detected by autoradiography. When necessary, quantitation was performed with a Betagen Betascope Blot Analyzer.

RESULTS

Constructs and experimental design. Polyomavirus late transcription termination and polyadenylation are both inefficient. In a productive viral infection, this leads to a heterogeneous collection of late pre-mRNA molecules in the nucleus. Pre-mRNA splicing joins small, noncoding late leader exons to one another, skipping internal message body exons, each of which is flanked by a 3′ splice site and a poly(A) site. We have used these features of polyomavirus late RNA processing to learn more about the relationship between splice site selection and poly(A) site choice. We first constructed a series of plasmids which contain two late transcription units and at least one intact early coding region and DNA replication origin, allowing early region gene expression and DNA replication in transfected mouse cells. Figure 1 illustrates schematically the late region intron-exon structures of these constructs. The basic construct, I-13, has two distinguishable leader exons, L1 (51 nt) and L2 (57 nt). The L2 leader exon is that found in wild-type polyomavirus. The L1 leader is a variant which has been shown previously to function like wild type in a virus and which is not defective in late pre-mRNA processing (4, 6). The two distinguishable message body exons are V1 (the wild-type VP1 exon, about 1,200 nt) and V2 (the VP1 exon with an internal deletion, 81 nt). The splice sites bordering the L1 and L2 exons are identical, as are the sequences flanking the V1 and V2 exons. Construct I-11 differs from I-13 only in having a 771-bp deletion spanning the poly(A) site just downstream of the V1 exon. Construct I-11SmΔ contains a smaller, 26-bp deletion, spanning this same poly(A) site (the last 5 bases of the AAUAAA signal, plus 21 downstream bases). Primary transcripts produced from constructs I-13, I-11, and I-11SmΔ have alternating L1 and L2 leader exons and alternating V1 and V2 exons. As the polyomavirus late promoter requires only a very small region just upstream of the late leader exon for transcription initiation (14, 49a), transcription can initiate with almost equal efficiencies from upstream of either the L1 or L2 leader exons. Two additional constructs were made. Construct I-16 is related to I-13 but carries a deletion of the L2 exon. Construct I-17 contains a single viral genome with a 1.6-kb duplication of the late poly(A) site and downstream flanking sequences.

In all experiments, viral genomes were excised from the vector, recircularized by dilute ligation, and transfected into mouse NIH 3T3 cells. Cytoplasmic RNAs were harvested after 48 h and were analyzed by using T2 ribonuclease protection assays or PCR.

V1 and V2 exons can be skipped during pre-mRNA processing. Cytoplasmic RNAs from wild-type polyomavirus and constructs I-13, I-11, and I-16 were examined for exon skipping by a PCR assay (32). Figure 2 shows the results of such an experiment using oligo 275, which binds to the junction of any L1 or L2 exon spliced to any V1 or V2 exon. PCR of the wild-type RNA described above provided size markers for 1 and 2 tandem L2 leader units. As transcription could initiate from either of the two late promoters in I-13 and I-11 (Fig. 1), we observed both the L1-to-V1 and L2-to-V2 splices in those RNAs. The upper band in each lane represents the splicing of leader exons to one another.

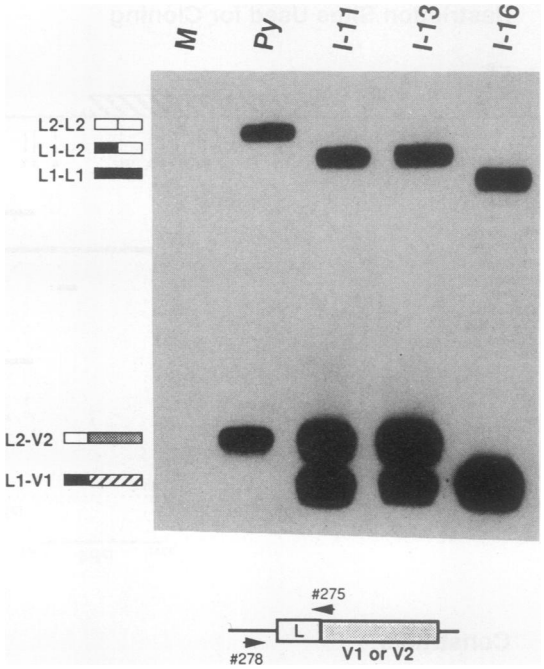

FIG. 2. PCR assay to determine exon structures of cytoplasmic RNA molecules. The assay was performed as described in Materials and Methods using RNA isolated from mock-transfected cells and cells transfected with religated DNAs from constructs from pPyBS (wild-type polyomavirus), I-11, I-13, and I-16. The oligonucleotide probes detect L1 or L2 spliced to either V1 or V2. The expected positions of leader structures are shown on the left. At the bottom is a schematic diagram of the PCR procedure employed, which has been described previously by our laboratory (32).

In the I-13 and I-11 lanes the upper band is intermediate in size from that representing two tandem L1 exons (lane I-16) or two tandem L2 exons (lane Py), indicating that the shorter L1 had spliced to wild-type-length L2. Thus, L1 and L2 exons can splice to each other in long primary transcripts, with a terminal splice to a body exon. Interestingly, the results seen with I-11 and I-13 demonstrate that L1 and L2 exons are not skipped during pre-mRNA splicing (only the intermediate size band observed), while V1 and V2 exons are. Note that the band marked L1-L2 in lanes I-13 and I-11 includes molecules with the structure L2-L1-V1 and L1-L2-V2. These result from transcripts initiating at either promoter. In other PCR experiments using oligonucleotide primers specific for either the V1 or V2 exon, both L1-L2-V2 and L2-L1-V1 structures have been confirmed (data not shown). From these results, the constructs studied appear to provide a simple and convenient system in which to study how exon skipping and splice site selection occur in vivo.

Unused poly(A) sites are not found in mature messages. No wild-type spliced polyomavirus late messages have ever been observed that contain internal unused poly(A) sites. In order to determine whether this was also the case for construct I-13, we analyzed messages from wild-type (polyomavirus) and I-13 transfections by using an RNase T2 protection assay and probe PP4, which spans the poly(A) site (Fig. 3). As the two poly(A) sites are identical, use of either could be scored. Use of the poly(A) site immediately downstream of the V1 exon yielded a band of 420 nt, which was prominent in both wild-type and I-13 RNAs. Use of the poly(A) site downstream of the V2 exon yielded a protected

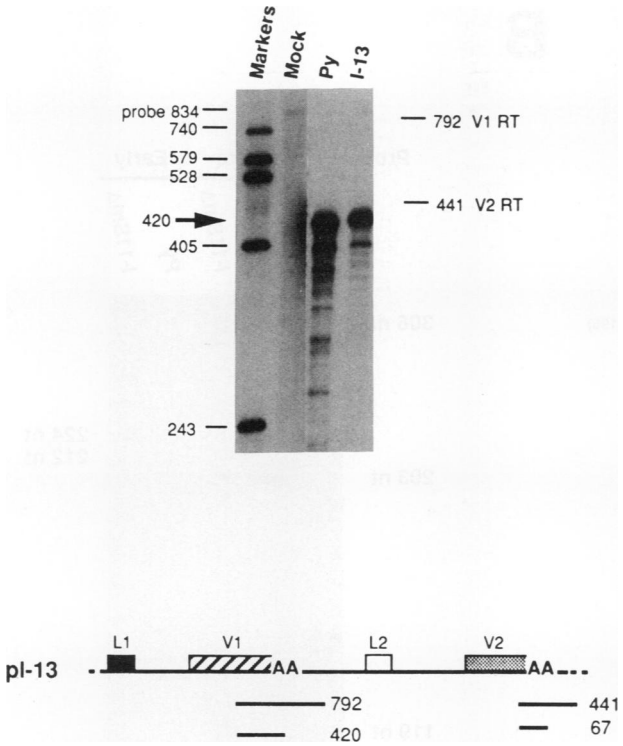

FIG. 3. RNase protection analysis of the 3'-end structures of cytoplasmic RNAs from cells transfected with wild-type polyomavirus DNA and cells transfected with I-13. The Mock lane contains a small amount of undigested probe (834 nt). The position of the band representing 3'-end processing at the late poly(A) site downstream of the V1 exon (420 nt) is shown. Bands representing RNAs containing readthrough sequences downstream of the first (V1 RT, 792 nt) and second (V2 RT, 441 nt) poly(A) sites are not seen. Since the V2 exon is short, the band representing polyadenylation after this exon has run off the gel. At the bottom is an illustration of processing events that would generate the bands observed; I-13 is shown, along with the approximate positions of hybridization of the riboprobe used, PP4, and the sizes of expected protected species.

band of 67 nt, which ran off the gel shown in Fig. 3 but was clearly visible in other experiments. Any messages that used the second poly(A) site of I-13 but which contained an unused first site would give a protected band of 792 nt due to transcriptional readthrough of the V1 exon (Fig. 3, V1 RT). Likewise, messages which contained an unused poly(A) site immediately downstream of the V2 exon (Fig. 3, V2 RT) would give a band of 441 nt. These messages were not detected. Taken together with the PCR results shown in Fig. 2, these results indicate that the use of the V1 or V2 3' splice site use in I-13 is associated with use of the nearest downstream poly(A) site.

V1 3' splice site selection can occur in the absence of a downstream poly(A) site. There are several possible models which could explain why no unused poly(A) sites were seen in mature V1- or V2-spliced messages in I-13. One possibility is that choice of a poly(A) site occurs first, with 3'-end cleavage occurring before splicing. L1 and L2 leader exons would then splice to each other, skipping V1 and V2 3' splice sites and removing all internal, unused poly(A) sites. This would be followed by splicing to the remaining V1 or V2 exon. In this scenario, V1 or V2 exon selection would be determined by poly(A) site choice. In order to test this possibility, we analyzed cytoplasmic RNAs produced from construct I-11, which lacks the poly(A) site downstream of the V1 exon. If the poly(A)-site-first model were correct, we would not expect to see L1-to-V1 splicing in I-11. RNase protection analysis of I-11 cytoplasmic RNA is shown in Fig. 4A. Results using probe PP2, which scores for L1-to-V1 splicing, demonstrate that splicing to the first V1 exon does, in fact, occur in about 70% of the molecules (as determined by using a Betagen Betascope Blot Analyzer). The remaining 30% largely represents use of the minor, alternative VP3 3' splice site, just as we see with wild-type viral RNA (Fig. 1 and data not shown).

Further analysis of these I-11 RNAs indicated that they contain sequences extending past the poly(A) site deletion, all the way to the next leader exon (L2). This was demonstrated by an RNase protection assay using probe PP3, which binds to RNA sequences downstream of the V1 3' splice site (Fig. 4A). About half of these RNAs contain an additional splice, mostly L2 to V2 (data not shown). Further analysis of both the splicing and the 3'-end structures surrounding the V2 exon in I-11 was done by using probe PP1 (Fig. 4A). The 81-nt band represented splicing to the V2 3' splice site, while the 196-nt band mostly represented molecules that contain L2 splices to the alternative VP3 3' splice site. In addition, no molecules were detected that included sequences downstream of the poly(A) site (445-, 560-nt bands), illustrating that construct I-11, like I-13, had no unused poly(A) sites in mature messages. These results with I-11 indicated clearly that V1 3' splice site choice can occur even when the nearby downstream poly(A) site has been removed.

Sequences downstream of the poly(A) site do not signal RNA degradation. The I-11 construct contains a large deletion of the poly(A) site (355 bp upstream of the AAUAAA signal and essentially all downstream sequences before the next genome). Therefore, the unusual message structures found in I-11 might result from the removal of sequences downstream of the poly(A) site that otherwise might serve as a degradation signal of RNAs containing them. In order to test this, a construct with a much smaller poly(A) site deletion, I-11SmΔ, was made. In this construct the deletion removed only the AAUAAA signal and 20 downstream bases, most or all of which are retained in mature, polyadenylated messages. This deletion destroys late polyadenylation (11a). Cytoplasmic RNA made from this construct was compared with wild-type RNA in an RNase protection assay which detected early spliced species (as an internal control for transfection efficiencies and RNA amounts) and late V1-spliced species. Such an assay has been reported in other work from our laboratory (4, 7). The results are shown in Fig. 4B. RNAs containing the V1 splice (protected band of 203 nt) are as abundant as wild type when normalized to the band intensities for the early species. This again demonstrates that poly(A) site choice is not needed for V1 3' splice site selection. The 119-nt band observed for I-11SmΔ with the V1 probe reflects cross hybridization of the probe to messages containing the V2 splice. The late V1 spliced molecules from I-11SmΔ must contain sequences downstream of the V1 3' splice site and upstream of the next leader, because the poly(A) signals immediately downstream of this exon have been removed. This means that L1-V2 splicing occurs in I-11SmΔ just as it does in I-11. Experiments using a variety of other probes spanning the 26-bp poly(A) signal deletion (data not shown) confirmed the presence of abundant I-11SmΔ late messages containing sequences not normally found in mature mRNAs. These

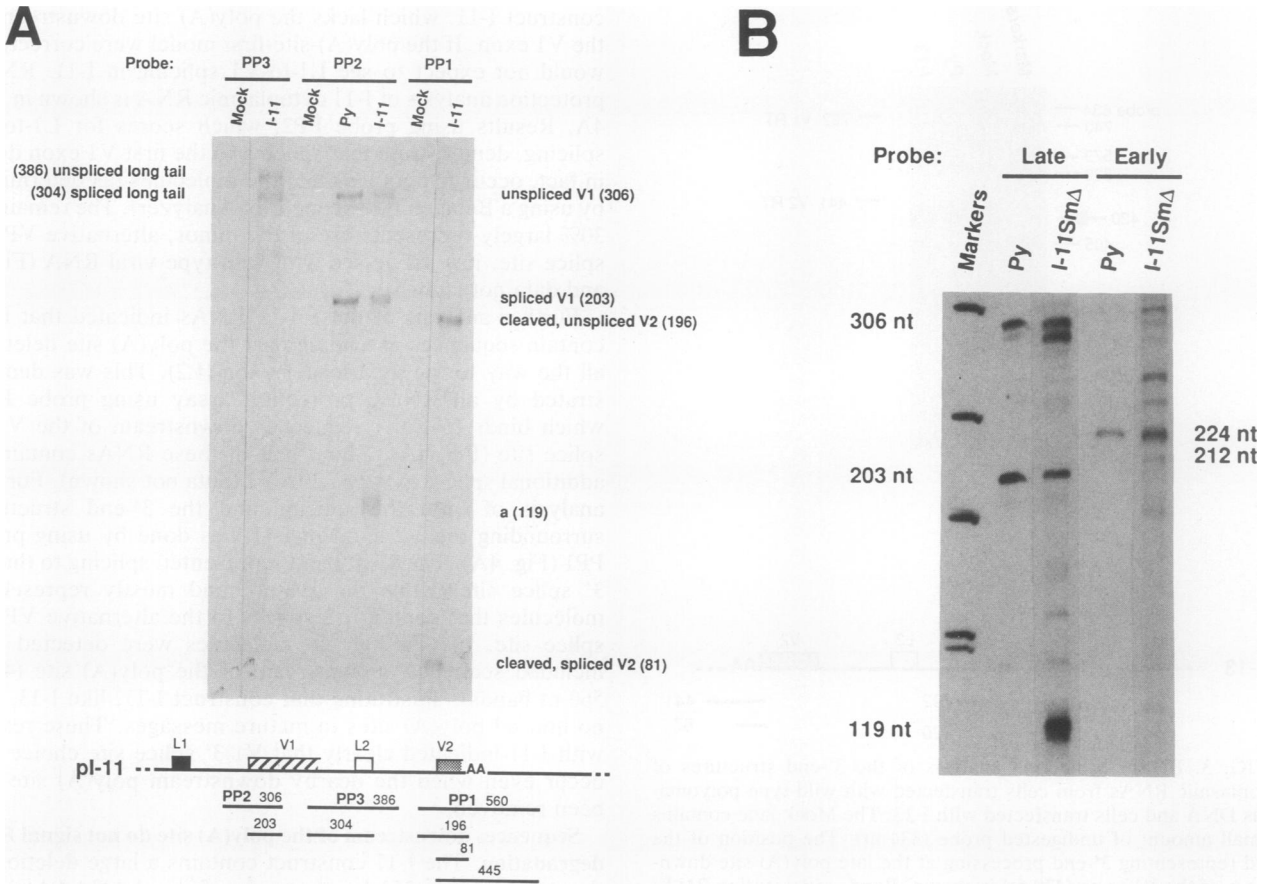

FIG. 4. (A) Riboprobe analysis of the structures of cytoplasmic RNAs from mock-infected cells, wild type-infected cells, and cells transfected with I-11. Gel analysis was as described in the legend to Fig. 3. Interpretations of bands are given beside the figure, and lengths of the fragments are shown in parentheses. The band labeled a represents cross hybridization of probe PP2 with sequences upstream of the V2 exon, and extending into this exon, to the position of the deletion junction. Following autoradiography the relative intensities of bands were quantitated by using a Betagen Betascope 603 Blot Analyzer. At the bottom is shown the structure of I-11 along with the approximate positions of binding of the probes used for this experiment and the sizes predicted for protected species. Probe PP3 is contiguous with I-11 sequences. For probe PP1, protected species of 560 and 445 nt were not detected, and that region of the gel is not shown here. (B) Riboprobe analysis of the structures of cytoplasmic RNAs from wild-type infected cells and cells transfected with I-11SmΔ. Cytoplasmic RNA was analyzed by using probe PP2 (to measure late RNA amounts and V1 and V2 splices) and an early probe which measures early polyomavirus mRNA amounts and splices (4). Bands at 306 and 203 nt represent messages without (306 nt) or with the V1 or V2 splice. The band at 119 nt using the late probe represents cross hybridization of probe PP2 with sequences upstream of the V2 exon of I-11SmΔ and extending into this exon to the position of the deletion junction. Bands at 224 and 212 nt, made with the early probe, represent early viral messages encoding the three T antigens.

results argue strongly that sequences downstream of the poly(A) site do not contain signals for message degradation.

A relationship exists between 3′ splice site selection and poly(A) site choice. If poly(A) site choice does not influence splice site choice, and if the sequences downstream of the poly(A) site do not serve as degradation signals, why do we fail to observe these sequences in spliced V1 exons in I-13 messages? One remaining possibility is that 3′ splice site selection influences poly(A) site choice. In order to investigate this possibility, we examined message structures in I-17, which contains a large (1.6-kb) duplication of the late 3′ processing and polyadenylation region (Fig. 1). 3′-end analysis by an RNase protection assay demonstrated that cytoplasmic messages from this construct were all processed at the upstream poly(A) site (420-nt band, Fig. 5A). A 792-nt band representing molecules containing an unused poly(A) site [indicating use of the downstream poly(A) site] was not seen. PCR analysis (Fig. 5B), however, showed that this construct generated messages with multiple L2 exons at their 5′ ends, indicating that transcription could continue past both poly(A) sites. Choice of the first poly(A) site in constructs containing duplicated polyadenylation signals is influenced by the distance between the two signals. Lanoix et al. (37) showed that in polyomavirus mutant ins4, which had a 144-nt duplication of the late poly(A) signal (including 51 nt upstream and 87 nt downstream of the AAUAAA sequence), both poly(A) sites were used almost equally. This is likely the result of interference between the two sites. We have carried out an extensive analysis of the spacing requirements for poly(A) site choice in this system, with the results to be presented elsewhere (40a). In that work we found that the first poly(A) site was chosen almost exclusively whenever the separation between the two sites was greater than about 150 nt.

Finally, in order to investigate further the relationship between splice site selection and downstream poly(A) site

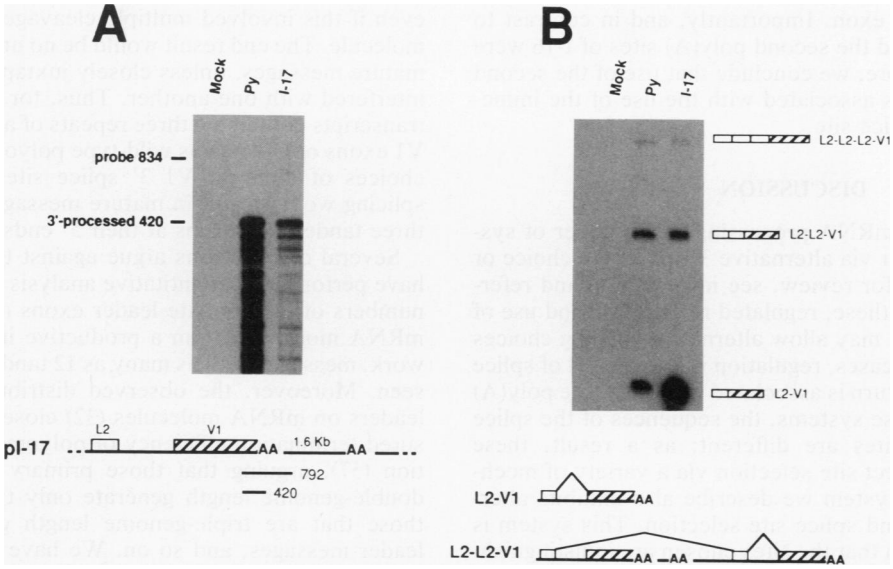

FIG. 5. Analysis of I-17. (A) Analysis of 3'-end structures using riboprobe PP4. The assay was performed as described in the legend to Fig. 3. At the bottom is shown the structure of I-17 along with the approximate positions of binding of probe PP4 and the sizes of expected protected species. (B) PCR analysis of the late leader structures on V1-spliced cytoplasmic messages. Py, wild-type polyomavirus-infected cells. I-17, cells transfected with I-17. The assay was performed as described in Materials and Methods, using oligo 275 and oligo 278. At the bottom is an illustration of splicing events that would generate the bands observed.

choice, we analyzed RNA processing in I-16. Construct I-16 differs from I-17 in that it contains a 3' splice site just upstream of the second poly(A) site. RNase protection assays with probes PP4 (Fig. 6A) and PP1 (data not shown) revealed that mature I-16 messages, like messages from the other constructs, do not contain internal, unused poly(A) sites. Many of these messages, however, contained multiple L1 exons at their 5' ends (Fig. 2 and 6B). Results of the PCR assay shown in Fig. 6B indicate that the V1 exon can be skipped by L1 splicing to a downstream V2 exon instead of

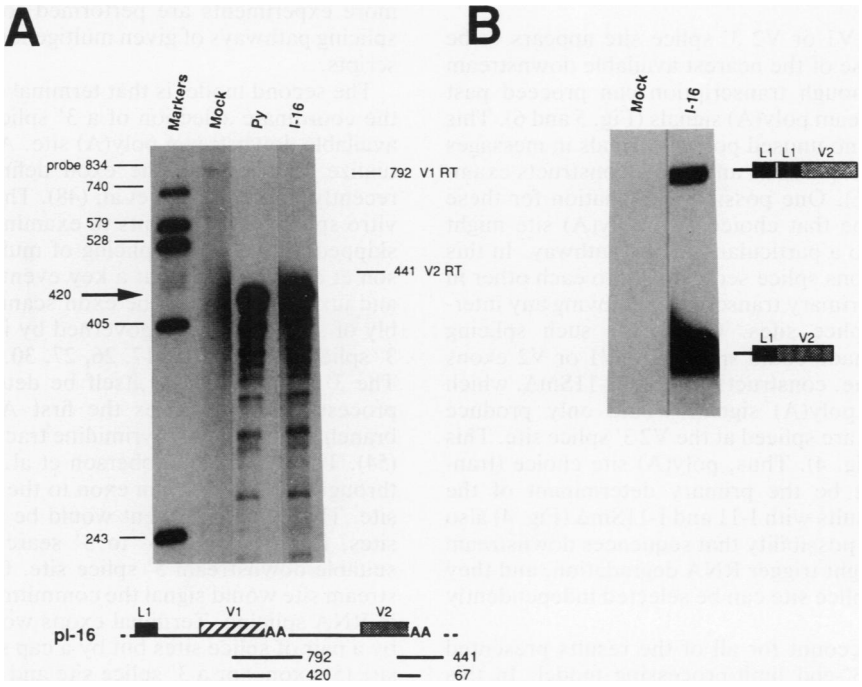

FIG. 6. Analysis of I-16. (A) Analysis of 3'-end structures using riboprobe PP4. The assay was performed as described in the legend to Fig. 3. At the bottom is shown the structure of I-16 along with the approximate positions of binding of probe PP4 and the sizes of expected protected species. (B) PCR analysis of the L1-spliced structures on V2-spliced cytoplasmic messages. WT, wild-type infected cells. I-16, cells transfected with I-16. The assay was performed as described in Materials and Methods, using oligo 275 and oligo 279. Oligo 279 allows detection only of messages with L1 exons spliced to the V2 exon.

to a downstream L1 exon. Importantly, and in contrast to I-17, both the first and the second poly(A) sites of I-16 were used in vivo. Therefore, we conclude that use of the second poly(A) site of I-16 is associated with the use of the immediate upstream 3′ splice site.

DISCUSSION

Regulation of pre-mRNA processing in a number of systems appears to occur via alternative 3′ splice site choice or poly(A) site choice (for review, see introduction and reference 53). In some of these, regulated recognition and use of polyadenylation sites may allow alternative splicing choices to be made. In other cases, regulation is at the level of splice site choice, which in turn is associated with alternate poly(A) site use. In all of these systems, the sequences of the splice sites and poly(A) sites are different; as a result, these differences might affect site selection via a variety of mechanisms. The model system we describe also exhibits alternative poly(A) site and splice site selection. This system is different, however, in that the sites chosen are indistinguishable from the sites that are ignored.

In constructs I-13 and I-11, L1 and L2 exons splice sequentially to one another and are rarely skipped. However, V1 and V2 exons are frequently skipped. Skipping of V1 and V2 exons is almost certainly not the result of inefficient recognition of their 3′ splice sites. In work to be presented elsewhere (40a), we have investigated the effects of different splice site and poly(A) site sequences on VP1 exon skipping in an otherwise wild-type polyomavirus backbone. Replacing the VP1 3′ splice site with the late leader 3′ splice site had no effect on the frequency of exon skipping. Rather, VP1 exon skipping could be eliminated either by replacing the normal poly(A) site with an efficient one or by placing a 5′ splice site between the VP1 3′ splice site and the poly(A) site.

The selection of a V1 or V2 3′ splice site appears to be associated with the use of the nearest available downstream poly(A) site, even though transcription can proceed past identical, far-downstream poly(A) signals (Fig. 5 and 6). This association results in no unused poly(A) signals in messages from wild-type polyomavirus or any of the constructs examined here (Fig. 3 to 6). One possible explanation for these observations would be that choice of a poly(A) site might commit a transcript to a particular splicing pathway. In this model, L1 and L2 exons splice sequentially to each other in multigenome-length primary transcripts, removing any internal V1 or V2 3′ splice sites. Only after such splicing commitments were made could splicing to V1 or V2 exons occur. If this were true, constructs I-11 and I-11SmΔ, which remove an internal poly(A) signal, would only produce mature messages that are spliced at the V2 3′ splice site. This was not observed (Fig. 4). Thus, poly(A) site choice (transcript length) cannot be the primary determinant of the splicing pathway. Results with I-11 and I-11SmΔ (Fig. 4) also rendered unlikely the possibility that sequences downstream of the poly(A) site might trigger RNA degradation, and they showed that the V1 splice site can be selected independently of the poly(A) site.

Two models can account for all of the results presented here. The first is a 3′-end limit-processing model. In this model, all splicing commitments would first be made in multigenome-length primary transcripts. Splicing would presumably occur according to the hierarchical strengths of the splice sites, as well as their positions in the transcript. Then, cleavage would occur at all remaining polyadenylation sites, even if this involved multiple cleavage events on the same molecule. The end result would be no unused poly(A) sites in mature messages, unless closely juxtaposed poly(A) signals interfered with one another. Thus, for example, in primary transcripts containing three repeats of a genome with L1 and V1 exons only (such as wild-type polyomavirus transcripts), choices of different V1 3′ splice sites during pre-mRNA splicing would result in mature messages with one, two, or three tandem L1 exons at their 5′ ends.

Several observations argue against this model. First, we have performed a quantitative analysis of the distribution of numbers of tandem late leader exons on polyomavirus late mRNA molecules from a productive infection (32). In this work, messages with as many as 12 tandem leader units were seen. Moreover, the observed distribution of numbers of leaders on mRNA molecules (32) closely parallels the measured termination efficiency of polyomavirus late transcription (57), arguing that those primary transcripts that are double-genome length generate only two-leader messages, those that are triple-genome length generate only three-leader messages, and so on. We have also failed to detect any nuclear RNA molecules that contain splices to V1 or V2 3′ splice sites but which are not also cleaved at the nearest downstream poly(A) site. Finally, we (40a) and others (56) have used a variety of probes to try to detect internal 3′ end cleavage intermediates in nuclear RNA samples. Such "amputated transcripts," as have been reported in the immunoglobulin heavy-chain system (49), have never been observed. It is possible, however, that in our system the processing intermediates are so unstable as to be undetectable, even in nuclei late after productive viral infection, where late primary transcripts are abundant. Further, we are not aware of any clear demonstration that multiple cleavage or polyadenylation events can occur on the same molecule in vivo. This model must remain a viable one, however, until more experiments are performed to elucidate the detailed splicing pathways of given multigenome-length primary transcripts.

The second model is that terminal exon selection involves the coordinate selection of a 3′ splice site with the nearest available downstream poly(A) site. A useful way to conceptualize this is using the exon definition model described recently by Robberson et al. (48). These authors designed in vitro splicing experiments to examine why exons are rarely skipped during RNA splicing of multiexon RNAs. Robberson et al. suggested that a key event in splice site selection and juxtaposition may be exon scanning. The initial assembly of spliceosomes is governed by interactions at and near 3′ splice sites (13, 15, 17, 26, 27, 30, 31, 35, 36, 46, 47, 50). The 3′ splice site may itself be determined by a scanning process that recognizes the first AG downstream of the branch point and polypyrimidine tract, regardless of distance (54). The model of Robberson et al. extends such a search through the downstream exon to the next available 5′ splice site. Thus, the first event would be recognition of 3′ splice sites, followed by a 5′ to 3′ searching process to find a suitable downstream 5′ splice site. Once found, this downstream site would signal the commitment of both sites for use in RNA splicing. Terminal exons would thus be defined not by a pair of splice sites but by a cap structure and a 5′ splice site (5′ exons) or a 3′ splice site and poly(A) site (3′ exons). In the case of 3′ exons, definition of the exon borders would be associated with the commitment of both polyadenylation and splicing to occur. Supporting this model, Niwa et al. (43) have recently shown that the presence of an intron can stimulate the rate of polyadenylation at a downstream

poly(A) site in vitro, with maximal stimulation requiring a functional 3′ splice site.

In the above model, 3′ splice site recognition is the initial event of terminal exon definition and might also be important for the subsequent use of the poly(A) site. Results from I-11 and I-11SmΔ clearly demonstrate that V1 3′ splice site choice can be independent of poly(A) site choice; this site can be used whether or not a poly(A) site is positioned immediately downstream of it. However, it is nevertheless quite possible that use of this splice site still requires the recognition of a downstream RNA processing signal. In I-11 and I-11SmΔ, the V1 3′ splice site might be functionally paired with a downstream leader 5′ splice site, which allows the recognition of the V1 exon as an internal rather than as a terminal exon.

Poly(A) site use in our system appears to be related to selection of the upstream 3′ splice site. As the results from I-17 show, only the poly(A) site nearest to the 3′ splice site is normally used, even though transcription proceeds past both. This is also consistent with a 5′ to 3′ searching mechanism. If the exon definition model is correct, we predict that insertion of a 3′ splice site between two poly(A) sites in a primary transcript should allow the previously unused downstream poly(A) site to be chosen. This is what we see when we compare the mRNAs produced by constructs I-17 and I-16, and it again illustrates the dominance of 3′ splice choice over poly(A) site choice.

In the exon definition model, both borders of an exon must be recognized in order for either to be available for RNA processing. If the 3′ border of an exon [either a 5′ splice site or a poly(A) site] were recognized inefficiently, then the upstream 3′ splice site would also be ignored, leading to exon skipping. In our system, results from I-13 clearly show that there is no spliced V1 or V2 exon with an unused poly(A) site downstream. This means that the choice of the V1 or V2 3′ splice site is coupled to the use of the nearest downstream poly(A) site. The intrinsic weakness of the polyomavirus late poly(A) site would explain why V1 and V2 are often skipped in the constructs studied here. In the absence of a poly(A) site to define the terminal exon, a further downstream 5′ splice site (such as the L2 leader 5′ splice site in the long tail of some I-11 messages) can be used instead. This explains what we observed in I-11. Thus, in order for proper splicing and polyadenylation to occur, both a terminal 3′ splice site (here V1 or V2) and a polyadenylation site must normally be chosen coordinately.

Finally, we speculate that the simplicity of the polyomavirus-based system described here has allowed us to uncover an underlying relationship between 3′ splice site choice and poly(A) site use that might exist in a number of transcription units. Only when both choices are made can either process normally be completed. In different systems, regulation might occur at either border of an alternative terminal exon.

ACKNOWLEDGMENTS

We thank N. Barrett, E. Carmichael, X. Cheng, R. Hyde-DeRuyscher, A. Roome, and A. Wahl for many helpful comments on this manuscript.

This work was supported by grant CA45382 from the National Cancer Institute.

REFERENCES

1. **Acheson, N.** 1976. Transcription during productive infection with polyoma virus and SV40. Cell **8**:1–12.
2. **Acheson, N.** 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.** 1978. Polyoma giant RNAs contain tandem repeats of the nucleotide sequence of the entire viral genome. Proc. Natl. Acad. Sci. USA **75**:4754–4758.
4. **Adami, G., 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.

4a. **Adami, G., R. Hyde-DeRuyscher, and G. Carmichael.** Unpublished data.

5. **Adami, G., and J. R. Nevins.** 1988. Splice site selection dominates over poly(A) site choice in RNA production from complex adenovirus transcription units. EMBO J. **7**:2107–2116.
6. **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.
7. **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.
8. **Akusjärvi, G., and H. Persson.** 1981. Controls of RNA splicing and termination in the major late adenovirus transcription unit. Nature (London) **292**:420–426.
9. **Alt, F. W., A. L. W. Bothwell, M. Knapp, E. Siden, E. Mather, M. Koshland, and D. Baltimore.** 1980. Synthesis of secreted and membrane-bound immunoglobulin mu heavy chain is directed by mRNAs that differ at their 3′ ends. Cell **20**:293–301.
10. **Amara, S. G., R. M. Evans, and M. G. Rosenfeld.** 1984. Calcitonin/calcitonin gene-related peptide transcription unit: tissue-specific expression involves selective use of alternative polyadenylation sites. Mol. Cell. Biol. **4**:2151–2160.
11. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl.** 1987. Current protocols in molecular biology. Greene Publishing Associates, New York.

11a. **Barrett, N., and G. Carmichael.** Unpublished data.

12. **Berget, S. M., C. Moore, and P. A. Sharp.** 1977. Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. Proc. Natl. Acad. Sci. USA **74**:3171–3175.
13. **Bindereif, A., and M. Green.** 1987. An ordered pathway of snRNP binding during mammalian pre-mRNA splicing complex assembly. EMBO J. **6**:2415–2424.
14. **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.
15. **Chabot, B., and J. A. Steitz.** 1987. Multiple interactions between the splicing substrate and small nuclear ribonucleoproteins in spliceosomes. Mol. Cell. Biol. **7**:281–293.
16. **Chen, C., and H. Okayama.** 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. **7**:2745–2752.
17. **Cheng, S.-C., and J. Abelson.** 1987. Spliceosome assembly in yeast. Genes Dev. **1**:1014–1027.
18. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
19. **Chow, L. T., R. E. Gelinas, T. R. Broker, and R. J. Roberts.** 1977. An amazing sequence arrangement at the 5′ ends of adenovirus 2 messenger RNA. Cell **12**:1–8.
20. **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.
21. **Early, P., J. Rogers, M. Davis, K. Calame, M. Bond, R. Wall, and L. Hood.** 1980. Two mRNAs can be produced from a single immunoglobulin gene by alternative RNA processing pathways. Cell **20**:313–319.
22. **Emeson, R. B., F. Hedjran, J. M. Yeakley, J. W. Guise, and M. G. Rosenfeld.** 1989. Alternative production of calcitonin and CGRP mRNA is regulated at the calcitonin-specific splice acceptor. Nature (London) **341**:76–80.
23. **Falck-Pedersen, E., and J. Logan.** 1989. Regulation of poly(A)

site selection in adenovirus. J. Virol. **63**:532–541.
24. **Feunteun, J., L. Sompayrac, M. Fluck, and T. L. Benjamin.** 1976. Localization of gene functions in polyoma virus DNA. Proc. Natl. Acad. Sci. USA **68**:283–288.
25. **Fraser, N. W., and E. Ziff.** 1978. RNA structures near poly(A) of adenovirus-2 late messenger RNAs. J. Mol. Biol. **124**:27–51.
26. **Frendewey, D., and W. Keller.** 1985. Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. Cell **42**:355–367.
27. **Frendewey, D., A. Kramer, and W. Keller.** 1987. Different small nuclear ribonucleoprotein particles are involved in different steps of splicing complex formation. Cold Spring Harbor Symp. Quant. Biol. **52**:287–298.
28. **Freund, R., G. Mandel, G. G. Carmichael, J. P. Barncastle, C. J. Dawe, and T. L. Benjamin.** 1987. Polyomavirus tumor induction in mice: influences of viral coding and noncoding sequences on tumor profiles. J. Virol. **61**:2232–2239.
29. **Galli, G., J. W. Guise, M. A. McDevitt, P. W. Tucker, and J. R. Nevins.** 1987. Relative position and strengths of poly(A) sites as well as transcription termination are critical to membrane versus secreted μ-chain expression during B-cell development. Genes Dev. **1**:471–481.
30. **Grabowski, P., S. Seiler, and P. Sharp.** 1985. A multicomponent complex is involved in the splicing of messenger RNA precursors. Cell **42**:345–353.
31. **Grabowski, P. J., and P. A. Sharp.** 1986. Affinity chromatography of splicing complexes: U2, U5 and U4/U6 small nuclear ribonucleoprotein particles in the spliceosome. Science **233**: 1294–1299.
32. **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.
33. **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.
34. **Kemp, D. J., G. Morahan, A. F. Cowan, and A. W. Harris.** 1983. Production of RNA for secreted immunoglobulin chains does not require transcription termination 5′ to the M exons. Nature (London) **301**:84–86.
35. **Konarska, M. M., and P. A. Sharp.** 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. Cell **46**:845–855.
36. **Konarska, M. M., and P. A. Sharp.** 1987. Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. Cell **49**:763–774.
37. **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.
38. **Leff, S. E., R. M. Evans, and M. G. Rosenfeld.** 1987. Splice commitment dictates neuron-specific alternative RNA processing in calcitonin/CGRP gene expression. Cell **48**:517–524.
39. **Leff, S. E., and M. E. Rosenfeld.** 1986. Complex transcriptional units: diversity in gene expression by alternative RNA processing. Annu. Rev. Biochem. **55**:1091–1117.
40. **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.
40a. **Luo, Y., and G. Carmichael.** Unpublished data.
41. **Mather, E. L., K. J. Nelson, J. Haimovich, and R. P. Perry.** 1984. Mode of regulation of immunoglobulin mu and delta chain expression varies during B-lymphocyte maturation. Cell **36**:329–338.
42. **Nevins, J. R., and M. C. Wilson.** 1981. Regulation of adenovirus-2 gene expression at the level of transcriptional termination and RNA processing. Nature (London) **290**:113–118.
43. **Niwa, M., S. D. Rose, and S. M. Berget.** 1990. In vitro polyadenylation is stimulated by the presence of an upstream intron. Genes Dev. **4**:1552–1559.
44. **Peterson, M. L., and R. P. Perry.** 1986. Regulation of production of μ_m and μ_s mRNA requires linkage of the poly(A) addition sites and is dependent on the length of the μ_s-μ_m intron. Proc. Natl. Acad. Sci. USA **83**:8883–8887.
45. **Peterson, M. L., and R. P. Perry.** 1989. The regulated production of μ_m and μ_s mRNA is dependent on the relative efficiencies of μ_s poly(A) site usage and the Cμ4-to-M1 splice. Mol. Cell. Biol. **9**:726–738.
46. **Pikielny, C. W., and M. Rosbash.** 1986. Electrophoresis of ribonucleoproteins reveals an ordered assembly of yeast splicing complexes. Nature (London) **324**:341–345.
47. **Reed, R., and T. Maniatis.** 1985. Intron sequences involved in lariat formation during Pre-mRNA splicing. Cell **41**:95–105.
48. **Robberson, B. L., G. J. Cote, and S. M. Berget.** 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol. Cell. Biol. **10**:84–94.
49. **Rogers, J., N. Fasel, and R. Wall.** 1986. A novel RNA in which the 5′ end is generated by cleavage at the poly(A) site of immunoglobulin heavy-chain secreted mRNA. Mol. Cell. Biol. **6**:4749–4752.
49a. **Roome, A., and G. Carmichael.** Unpublished data.
50. **Ruskin, B., and M. Green.** 1985. Specific and stable intron-factor interactions are established early during in vitro pre-mRNA splicing. Cell **43**:131–142.
51. **Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, and N. Arnheim.** 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science **230**:1350–1354.
52. **Shaw, A., and E. B. Ziff.** 1980. Transcripts from the adenovirus major late promoter yield a single early family of 3′ coterminal mRNAs and five late families. Cell **22**:905–916.
53. **Smith, C. W. J., J. G. Patton, and B. Nadal-Ginard.** 1989. Alternative splicing in the control of gene expression. Annu. Rev. Genet. **23**:527–577.
54. **Smith, C. W. J., E. B. Porro, J. G. Patton, and B. Nadal-Ginard.** 1989. Scanning from an independently specified branch point defines the 3′ splice site of mammalian introns. Nature (London) **342**:243–247.
55. **Treisman, R.** 1980. Characterization of polyoma late mRNA leader sequences by molecular cloning and DNA sequence analysis. Nucleic Acids Res. **8**:4867–4888.
56. **Treisman, R., and R. Kamen.** 1981. Structure of polyoma virus late nuclear RNA. J. Mol. Biol. **148**:273–301.
57. **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.
58. **Tyndall, C., G. LaMantia, C. Thacker, and R. Kamen.** 1981. A region of the polyoma virus genome between the replication origin and late protein coding sequences is required in cis for both early gene expression and viral DNA replication. Nucleic Acids Res. **9**:6231–6250.