JOURNAL OF VIROLOGY, Jan. 1989, p. 85–93
0022-538X/89/010085-09$02.00/0
Copyright © 1989, American Society for Microbiology
Vol. 63, No. 1

Leader-to-Leader Splicing Is Required for Efficient Production and Accumulation of Polyomavirus Late mRNAs

GUY R. ADAMI,† CHRISTOPHER W. MARLOR,‡ NANCY L. BARRETT, AND GORDON G. CARMICHAEL*

Department of Microbiology, University of Connecticut Health Center, Farmington, Connecticut 06032

Received 6 July 1988/Accepted 22 September 1988

Polyomavirus late mRNA molecules contain multiple, tandem copies of a noncoding 57-base "late leader" exon at their 5′ ends. This exon is encoded only once in the genome. Leader multiplicity arises from leader-leader splicing in giant primary transcripts, which are the result of multiple circuits of the viral genome by RNA polymerase II. We have been interested in learning more about the role of the leader exon in late viral gene expression. We recently showed that an abbreviated-leader mutant virus (ALM) with a 9-base leader exon is nonviable (G. R. Adami and G. G. Carmichael, Nucleic Acids Res. 15:2593–2610, 1987) and has a severe defect in both late pre-mRNA splicing and stability. However, a mutant virus with a different, substituted leader sequence of 51 nucleotides (SLM/MP8) is viable and has no apparent defects. Here we examined further the role of the late leader exon in late pre-mRNA processing. When the leader exon length was gradually reduced from 51 nucleotides to 9 nucleotides in a series of mutants, RNA splicing and stability defects were coupled. In this system there was a minimum exon size of between 33 and 27 nucleotides. Next, a number of mutations were introduced into the 3′ splice site which precedes the late leader. Such mutations blocked leader-leader splicing. Surprisingly, they also interfered with leader-mVP1 body splicing and resulted in unstable primary transcripts. Thus, polyomavirus leader-leader splicing appears to be important for the efficient accumulation of late viral mRNA molecules.

Late polyomavirus pre-mRNA molecules are giant and of variable length, the result of multiple circuits of the viral genome by RNA polymerase II (1, 22). Mature late messages contain multiple, tandem copies of an untranslated 57-base "late leader" sequence near their 5′ ends (21). Leader multiplicity is thought to arise by splicing of leader exons one to another during the processing of giant primary transcripts (13, 21). Figure 1A illustrates how leader-leader splicing could account for one late message having two leader units near its 5′ end. This scheme for late viral pre-mRNA processing is an interesting example of a natural system in which multiple, identical exons can occur in primary transcripts but in which message biogenesis involves a precise pattern of splice site selection involving only some of the available sites. The mechanism of splice site selection in polyomavirus late pre-mRNA splicing is not known.

Our laboratory has been examining the role of the late leader unit in viral gene expression. We have shown that the nucleotide sequence of the late leader appears to be largely irrelevant for virus viability in cultured mammalian cells but that its length is important for both pre-mRNA splicing and stability (2). Stability is measured by comparing the ratios of steady-state levels of early and late transcripts following viral DNA transfections, after controlling for promoter strength with nuclear run-on assays. A mutant (ALM) with a shortened leader exon of only 9 bases is severely defective in both pre-mRNA splicing and stability (3). Since another mutant with a substituted leader sequence of 51 nucleotides (SLM/MP8) grows like the wild type and is not defective in late pre-mRNA processing, we have suggested that the ALM defects result from splice site interference (Fig. 1B) (3). Here we show by studying a series of mutants with leaders of different lengths that the splicing and stability defects are proportional to the proximity of the splice sites which border the leader. In other words, when strong splice sites are too close, neither functions, resulting in RNA degradation. Since the splicing inhibition data showed that decreasing RNA stability and mVP1 splicing are coincident, we sought to determine whether these phenomena are a direct effect of splice site interference. We therefore constructed and studied a number of viruses with mutations in the 3′ splice site which precedes the leader exon. Our prediction was that these mutations would relieve the interference, thus abrogating the splicing and stability defects. Unexpectedly, however, mutations that blocked leader-leader splicing also interfered with leader-mVP1 body splicing. Primary transcripts that could not be spliced from leader unit to leader unit did not efficiently use the mVP1 splice site. Thus, late leader splicing appears to be required for the efficient usage of other late viral splice sites and for the formation of stable mRNA. Blocking leader-leader splicing by splice site juxtaposition or by splice site mutation results in the same phenotype.

MATERIALS AND METHODS

Materials. The plasmids used, pPy (containing the entire wild-type polyomavirus strain 59RA genome), pALM (containing the genome of the abbreviated-leader mutant), and pSLM/MP8 (containing the genome of the substituted-leader mutant), were constructed and propagated as described previously (2). Although strain 59RA contains a 31-base-pair (bp) duplication around the region of the major late start sites (17) and several other base changes from the A2 sequence (18), we have adopted the numbering system of Soeda et al. (18) in our description of probes and RNA splice sites. *Escherichia coli* HB101 and GM48 (F^+ *lac gal ara thr glu phi* T_1^r T_6^r *dam-3 dcm-6* Sup^+; obtained from Henryk Cudny) were used to propagate plasmids by standard procedures

* Corresponding author.
† Present address: Rockefeller University, New York, NY 10021.
‡ Present address: Molecular Therapeutics, Inc., West Haven, CT 06516.

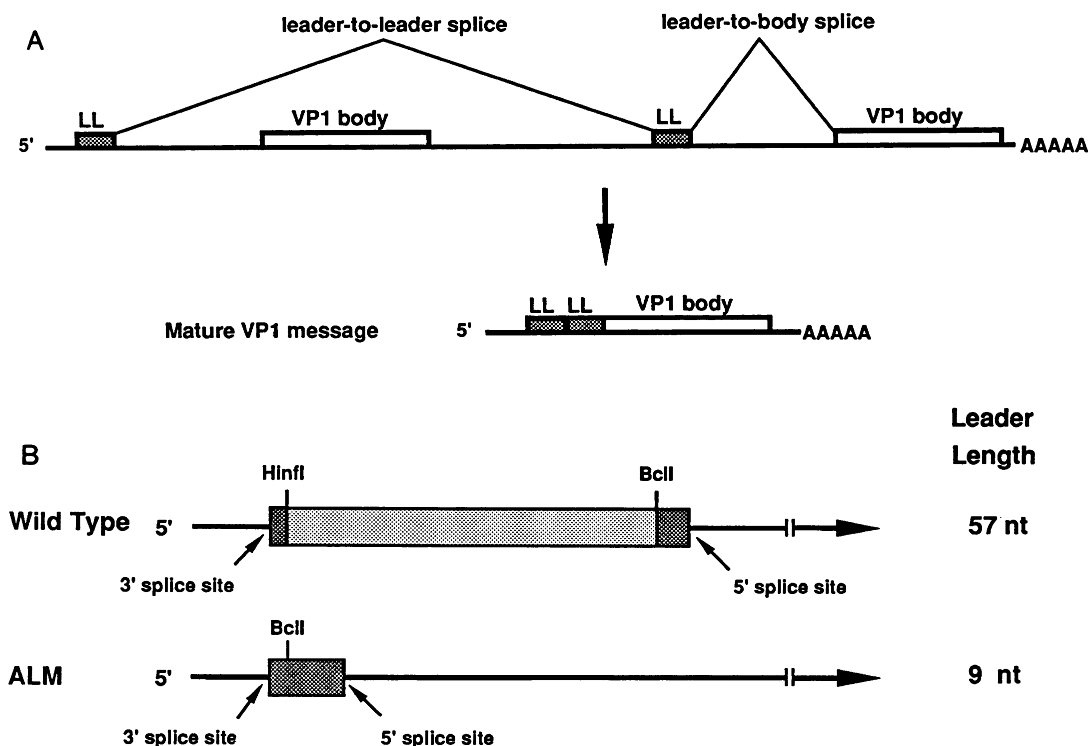

FIG. 1. (A) Model for the processing of polyomavirus late giant primary transcripts and the production of multiple, tandem nontranslated late leader units on mature mRNA molecules. Owing to inefficient termination by RNA polymerase II, late viral transcripts are heterogeneous in length, with many representing multiple circuits of the viral genome by the polymerase. The example shown is for a transcript created by two passes through the viral late region. During processing a message body (mVP1) is spliced to the 3'-proximal late leader unit (LL). In addition, leader-leader splicing removes a genome-length intron, yielding a final mVP1 message with two tandem leader units at its 5' end. (B) Schematic diagram of the leader regions of wild-type polyomavirus and the nonviable, splicing-defective mutant ALM. The heavy lines represent late RNA; the boxes represent the leader units. The ALM late leader unit retains nine nucleotides (nt) from the wild-type leader, as shown. Note that both leaders shown here are flanked by splice sites used in late pre-mRNA processing. In the case of ALM, these splice sites are separated by only nine nucleotides.

(14). Nuclease S1 was from Sigma Chemical Co., and Bal31 nuclease was from Bethesda Research Laboratories, Inc. All other enzymes were from New England BioLabs, Inc. α-^{32}P-labeled deoxynucleoside and ribonucleoside triphosphates and [γ-^{32}P]ATP were from New England Nuclear Corp. or Amersham Corp. Eastman Kodak Co. XAR-5 film was used for autoradiography and densitometry.

Cell culture techniques. Mouse NIH 3T3 cells were grown and transfected with viral DNAs or infected with virus as described previously (2). Intact mutant viral genomes for transfection studies were made by dilute ligation of intact genomes liberated from recombinant plasmids after digestion with *Bam*HI (2). At 30, 40, or 42 h after transfection, nuclei for run-on transcription analysis (7, 8, 15), nuclear RNA (14), or total RNA (isolated by the guanidine thiocyanate procedure [4, 6]) was isolated. RNAs were stored as ethanol precipitates at −20°C until used for nuclease S1 or RNase T_2 analysis.

Preparation of probes and nuclease S1 analysis. Nuclease S1 protection experiments were done with short, end-labeled, single-stranded probes isolated by polyacrylamide-urea gel electrophoresis followed by electroelution and ethanol precipitation (3). The 5' ends were dephosphorylated with calf intestine alkaline phosphatase (Boehringer Mannheim Biochemicals), purified further with an Elutip-d column (Schleicher & Schuell, Inc.), and end labeled with polynucleotide kinase and [γ-^{32}P]ATP (14).

RNA for analysis (generally 1/2 to 1/50 of the RNA isolated from one 100-mm plate) was mixed with labeled probe (0.1 to 1 μg). Hybridization and nuclease S1 digestion were performed as described previously (3), and annealing was done at 36°C. Optimal temperatures for the 1 h of incubation with nuclease S1 varied depending on the probe used and are given in the figure legends. The products were resolved on 6% polyacrylamide–urea sequencing gels. Quantitation was done by densitometry of the autoradiograms. To ensure that nuclease S1 analysis was always done with an excess of probe, we used various amounts of probe in control experiments.

Preparation of probes and RNase T_2 analysis. RNase T_2 protection experiments were done with the same fragments used as probes in the nuclease S1 analysis (3); these fragments were inserted into the multisite cloning region in the vector BlueScribe M13+ (Stratagene). This region is flanked by phage T7 and T3 RNA polymerase promoters. [α-^{32}P]UTP-labeled T3 or T7 transcripts were prepared in vitro as described by Hart et al. (9) and annealed to total RNA prepared by the guanidine thiocyanate method (4) or to poly(A)$^+$ RNA at 60°C for 15 to 24 h. The resulting hybrids were digested with RNase T_2 (Boehringer Mannheim) at a final concentration of 60 U/ml at 34 to 37°C for 2 h and resolved by electrophoresis on 6% polyacrylamide–urea sequencing gels.

Construction of leader mutants. Truncated leader mutants with unidirectional deletions were generated by controlled Bal31 digestion. The parent plasmid, pSLM/MP8, consists of

the entire genome of the viable mutant SLM/MP8 (2) inserted into the *Bam*HI site of pBR322. This plasmid DNA was digested with *Sma*I (which cuts only once, inside the late leader unit) and *Eco*RI. The resulting 1,822-bp fragment spanning from the leader unit through the polyomavirus replication origin to the viral *Eco*RI site was gel purified and treated with Bal31 nuclease for various times. *E. coli* DNA polymerase I Klenow enzyme was used to make the ends blunt, and the DNA was digested with *Bgl*I. This digestion yielded a collection of fragments with one fixed *Bgl*I end and the other end variously truncated. In a separate reaction, wild-type plasmid pPK (which contains the entire polyomavirus strain 59RA genome cloned into pMK16* at the unique *Bam*HI site) was grown in *dam* mutant *E. coli* GM48 and digested with *Bcl*I (which cuts at a single site, within the wild-type leader unit). The DNA ends were made blunt with the Klenow enzyme, treated with calf intestine alkaline phosphatase, and digested with *Bgl*I (which cuts only once, in the polyomavirus sequence of pPK). The resulting DNA was mixed with the Bal31-treated material and ligated overnight at 16°C. Competent HB101 cells were transformed, kanamycin-resistant colonies were isolated, and individual mutants were propagated after DNA sequence verification by the dideoxy method.

A series of deletion mutants lacking sequences upstream of the ALM leader exon were constructed in a similar way, starting with pALM DNA (2). This plasmid, grown in *dam* mutant *E. coli*, was digested with *Bcl*I (the ALM leader retains the unique *Bcl*I site) and *Eco*RI. The resulting 1,808-bp fragment was gel isolated and treated with Bal31 nuclease, the ends were made blunt with the Klenow enzyme, and *Bcl*I linkers were attached with T4 DNA ligase. After excess linkers were trimmed with *Bcl*I, the DNA was digested with *Bgl*I and ligated overnight at 12°C to the pPK backbone described above. Following transformation of competent HB101 cells, individual clones were picked, propagated, and verified by dideoxy sequencing. Mutants ALM-4, ALM-11, and ALM-15 were chosen for further study. These contained deletions of 4, 11, or 15 nucleotides upstream of the *Bcl*I site of ALM, respectively (see Fig. 4).

Mutant ALM-3′SS was constructed with a 30-base synthetic oligonucleotide (5′-CATTTTCTATTTTAAAAATG ATCAAGTAAG-3′) to change the AGAG sequence at the ALM leader 3′ splice site junction to AAAA. A 1,170-bp *Bam*HI-*Pst*I fragment (spanning the late leader region) from ALM was cloned into phage M13mp8 DNA. Recombinant plaque were isolated, and single-stranded DNA was prepared and mutated by standard methods (25). Mutants were revealed by hybridization of ^{32}P-labeled oligonucleotides to plaque lifts onto nitrocellulose (14). The mutant chosen for further analysis was verified by dideoxy sequencing. The *Bam*HI-*Pst*I fragment was excised, recloned into pBR322, and propagated in *dam* mutant GM48 cells, and the *Bcl*I-*Bgl*I fragment was isolated and inserted into pPK DNA that had been digested with the same two enzymes.

Transcription rate measurement in isolated nuclei. At 30 h posttransfection, nuclei from mock-transfected, ALM-transfected, or wild-type (pPK)-transfected cells were prepared and incubated with [α-^{32}P]UTP as described previously (7). The labeled RNA was isolated and hybridized to single-stranded probes bound to nitrocellulose filters. For each experiment, approximately 10^8 nuclei were used from NIH 3T3 cells transfected with *Bam*HI-cut, religated DNAs. Following incubation, the reactions were terminated with DNase I, and RNA was isolated essentially as described previously (8, 15). Hybridization was to single-stranded DNA probes specific for early or late transcripts. The early-transcript-specific probe, PR-3 (R. Hyde-DeRuyscher and G. G. Carmichael, Proc. Natl. Acad. Sci., in press), was the viral *Pst*I-*Eco*RI fragment (nucleotides 488 to 1560) cloned into phage M13mp8. The late-transcript-specific probe, RPHD11 (Hyde-DeRuyscher and Carmichael, in press), contains a *Hin*dIII fragment from nucleotides 2921 to 3919 cloned into the *Hin*dIII site of phage M13mp18. Single-stranded DNA (5 μg) was bound to GeneScreen Plus (New England Nuclear) as described by the manufacturer with a Bethesda Research Laboratories slot-blot manifold apparatus. Hybridization and processing of the filters were done as described by Hyde-DeRuyscher and Carmichael (in press).

RESULTS

Minimum leader exon size for both late pre-mRNA splicing and stability in polyomavirus. Mutant ALM, with a 9-base leader exon, is nonviable and produces unstable late primary transcripts which are inefficiently spliced (3). Since a mutant (SLM/MP8) constructed earlier has a 51-base leader with a sequence unrelated to that of the wild type and is indistinguishable from our laboratory wild-type strain 59RA (2), we set out to determine whether the defects in ALM are related to leader length. Therefore, starting with pSLM/MP8, we used Bal31 nuclease to construct mutants with progressively shorter leaders. Figure 2 shows the mutants that were made and sequence verified.

Each of the truncated leader mutants was propagated as a plasmid, liberated from the vector backbone with *Bam*HI, and recircularized with T4 DNA ligase. These DNAs were then transfected into mouse NIH 3T3 cells, and the immunofluorescence burst assay (2) was used to monitor viability. By this assay, mutants with leader exon lengths of 33 nucleotides or longer were found to be viable, while those with leaders of 27 nucleotides or shorter were not (data not shown).

We next used protection from S1 nuclease to examine the levels of splicing of the late leader exon to the message body for mVP1, which encodes the major viral structural protein, VP1. Since the early promoter appears not to be affected by deletions or substitutions in the late leader region (3), for each mutant we measured levels of early viral spliced RNAs as an internal control for transfection efficiencies. Figure 3 shows the results of this experiment. It is clear that mutants with leaders of 33 nucleotides or longer showed the same levels of late RNA splicing (ratio of spliced band to full-length protected probe, about 0.9). Comparison to the early RNA internal controls indicated that steady-state levels of late nuclear RNA (determined by calculating the ratios of total early to total late protected signals) were also about the same for each of these mutants. However, mutant SLM-27, with a 27-base leader exon, had a deficiency of about fivefold in leader-mVP1 body splicing. SLM-22, with a 22-base leader exon, was further impaired, with the mVP1 splice being virtually undetectable in this experiment. This mutant was indistinguishable from ALM, with a 9-base leader exon. Note that for the late splice levels to be visible, more nuclear RNA had to be used in the lanes for SLM-27 and SLM-22 (compare early RNA band intensities). Table 1 presents the densitometric quantitation of these results. In agreement with previous results for ALM with the same probes (3), we found that RNA stability was reduced more than 20- to 40-fold in short-leader mutants (e.g., SLM-22), while the band representing the mVP1 splice was reduced at least an additional 10- to 20-fold. These results showing stability and

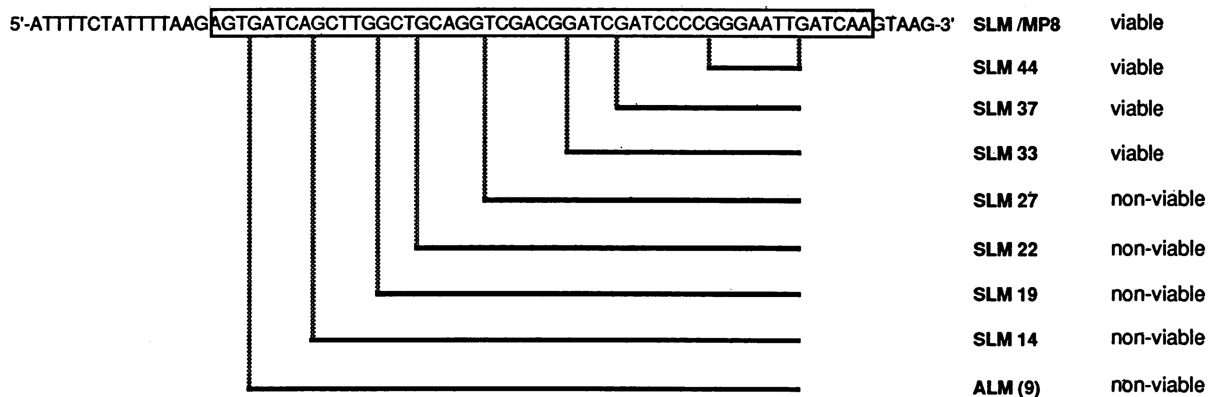

FIG. 2. Leader deletion mutants used to determine the minimum exon length for RNA splicing and stability. The starting virus was SLM/MP8, which has a 51-base leader (boxed) derived from cloning vector M13mp8. This virus is indistinguishable from our laboratory wild-type strain, 59RA, which encodes 57-base late leaders (2, 3). Bal31 nuclease was used to generate unidirectional deletion mutants as described in Materials and Methods. The dark lines represent bases deleted in the indicated mutants. Each virus was then tested for viability, using the immunofluorescence burst assay (2), and the results are shown.

splicing decreasing together with decreasing leader length help to prove that the defects are linked and dependent on leader exon length.

Mutations in the 3′ splice site upstream of the late leader unit do not correct the defects in ALM. We previously suggested that the defects in ALM are due to the juxtaposition of two efficient splice sites (3). In this mutant, splice sites flanking the leader exon are too close together to allow efficient usage of either site. A prediction of this model is that a defective 3′ splice site upstream of the mutant leader unit should relieve both defects. To test this model, we constructed a number of mutants (Fig. 4). Starting with ALM, a series of unidirectional deletion mutants were made with Bal31 nuclease. In these mutants we removed 4, 11, or 15 nucleotides from the 3′ splice site region upstream of the ALM leader exon. In mutant ALM-3′SS, an oligonucleotide was used to change the AGAG sequence at the ALM leader 3′ splice site junction to AAAA, destroying any possible AG dinucleotide motif in this region. Like ALM, all of the upstream splice site mutants were nonviable, as determined by the immunofluorescence burst assay (data not shown). It is important to note that in these mutants, leader-leader splicing could not take place in vivo, owing to the destruction of an involved splicing site. Splice sites involved in leader-body splicing, however, were not altered. Thus, giant nuclear transcripts in mutant-infected cells might be expected to accumulate, while those that only contain a single leader unit might be spliced normally to late message body exons. Alternatively, leader-mVP1 or leader-mVP3 body splicing might occur in giant transcripts, skipping over the mutated leader exons. Figure 5 shows an analysis of late splicing for these mutants. Nuclear RNAs from transfected mouse NIH 3T3 cells were analyzed for mVP1 splicing and RNA stability, and results were normalized to early RNA levels to correct for variable transfection efficiencies. For comparison, nuclear RNA from ALM-transfected cells was analyzed. Surprisingly, late splicing for the mutants remained as severely compromised as for ALM. The ratio of the "unspliced" band to the band representing the mVP1 splice was very high for all of these mutants, whereas the mVP1 splice band was the predominant one for wild-type virus. Mutant transcripts were also less stable than wild-type ones; none of the mutations examined here restored steady-state nuclear RNA levels to near that in the wild type. As it is possible for one RNA processing step to have effects on others (e.g., there is a reported relationship between splicing and polyadenylation with the polyomavirus early region [16]), it is important to determine what the primary flaw is in the splice site mutants. One can imagine a scenario in which RNA that is inefficiently polyadenylated is not a good substrate for the splicing reaction. Figure 6 shows that the splicing defects of the leader upstream splice site mutants were also apparent in poly(A)$^+$ RNA. The mutant primary transcripts that managed to be processed and transported to the cytoplasm still reflected the inefficient usage of the mVP1 splice site. Thus, the primary defect is in late polyomavirus splicing and is not secondary to any defect in polyadenylation: otherwise, one would expect the RNA that is polyadenylated to show no splicing defect.

As a final control, nuclear run-on analyses were carried out to correct for possible late promoter effects in the mutants examined. Although the ALM late promoter is indistinguishable from that of the wild type (3; Fig. 7), this is an important control, since the major late start sites for our strain are 55 and 85 nucleotides upstream of the leader (3), and some of the deletions used might be expected to interfere with late promoter function. Figure 7 shows that for these upstream mutants the late promoter was never significantly compromised, by comparison with the early promoter. Densitometric scanning of these and other exposures revealed that the late promoters of each of the viruses examined here were equivalent, within a factor of 2.

DISCUSSION

Role of leader exon length in polyomavirus late pre-mRNA processing. In the polyomavirus system, leader length (but not sequence) appears to be important for both late pre-mRNA splicing and stability. Mutants with leader exons of 33 nucleotides or longer are viable and show no defects or only minor defects in growth, mVP1 splicing, and nuclear RNA stability. Exons of 22 nucleotides or shorter, however, lead to a loss of viability and severe defects in both pre-mRNA splicing and stability. A mutant encoding a 27-base leader exon is nonviable but intermediate in the two RNA processing defects.

We feel that the ALM or other short-leader mutant exons are not "poison" sequences which direct RNA degradation.

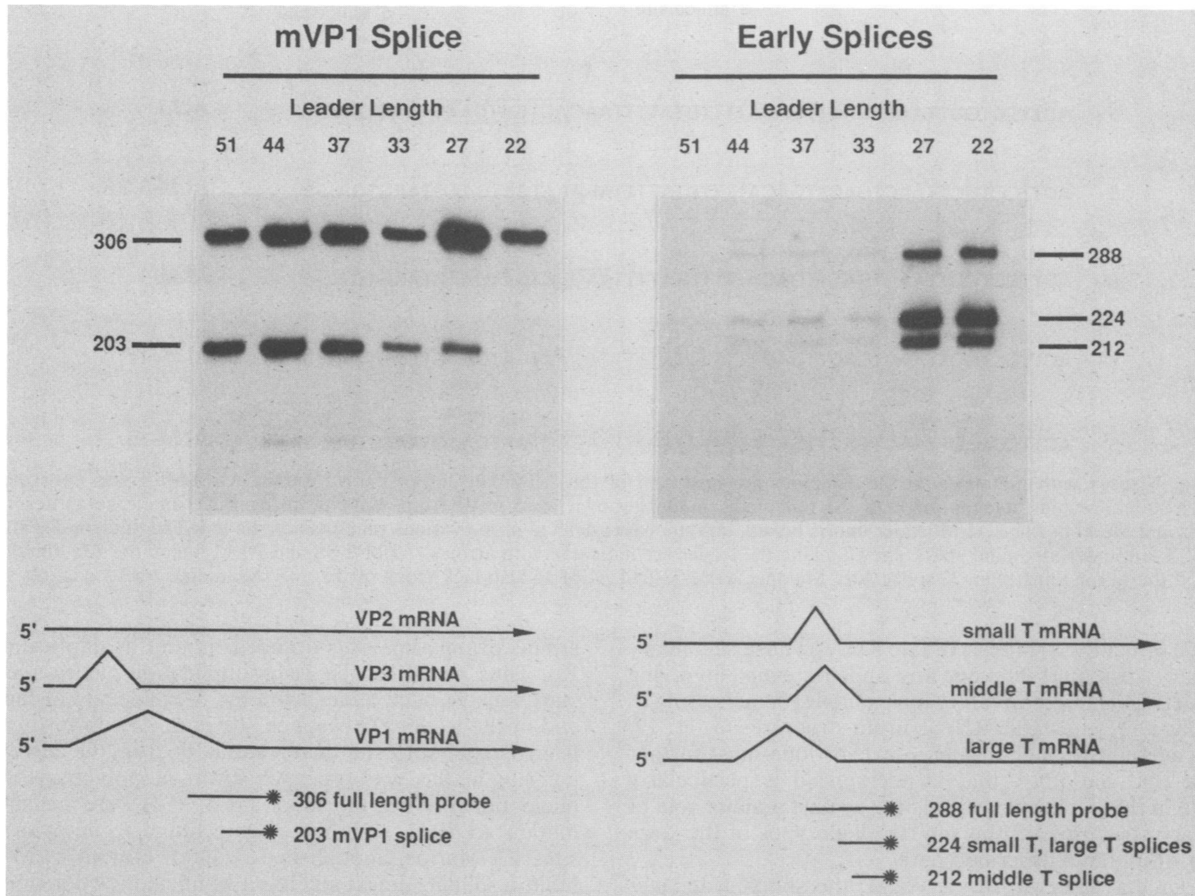

FIG. 3. Analysis of late nuclear RNA accumulation and mVP1 splicing in leader deletion mutants. Nuclear RNA was isolated from cells 40 h after transfection as described in Materials and Methods and subjected to nuclease S1 analysis with 5'-end-labeled, strand-separated, single-stranded probes. The early probe was a 288-nucleotide fragment (*Ava*I-*Hgi*AI fragment, from nucleotides 730 to 1016 [18]), and the late probe was a 306-nucleotide fragment (*Hin*dIII-*Pst*I fragment, from nucleotides 3920 to 4225 [18]). The early probe revealed splices for the messages for small, middle, and large T antigens, while the late probe revealed the mVP1 splice. Although variable amounts of nuclear RNAs (1/2 to 1/50 of the nuclear RNA isolated from one 90-mm petri dish) were used for the different mutants, to be able to visualize levels of polyomavirus transcripts and splices for different mutants, we used the same relative amounts of nuclear RNAs for early versus late RNA analysis in each experiment, to allow comparison and quantitation. Thus, the relative levels of late RNAs produced by different viruses could be determined by comparing the ratios of the early and late signals. Multiple exposures of autoradiograms were made, and relative amounts of bands were determined by densitometric scanning.

Rather, we feel that competence for leader-leader splicing is required for manifestation of the splicing and stability defects. Viruses with leaders of various sequences which are 33 to 96 nucleotides in length are viable, as long as they do not contain frameshifting AUG codons (2), while all mutants tested which have leaders less than 27 nucleotides in length are nonviable and phenotypically identical, as far as we have measured. In experiments to be reported elsewhere (N. Barrett and G. Carmichael, unpublished data), we have analyzed the effects of the wild-type and ALM leader exons on RNA synthesis and stability in a transient expression assay system. In this system, late promoter sequences, including the leaders and 54 nucleotides downstream of the leaders, were fused to the bacterial chloramphenicol acetyltransferase gene; no downstream 3' splice sites were available for splicing to the leaders. These two constructs were indistinguishable in the amounts of chloramphenicol acetyltransferase enzyme or chloramphenicol acetyltransferase mRNA produced following transfection into mouse cells.

Another possible explanation for the observed length dependence is that short-leader mutants (e.g., ALM and

TABLE 1. Relationship between late leader exon length and nuclear RNA stability and splicing[a]

Mutant	Relative level of late nuclear RNA[b]	Relative level of mVP1 splice[c]
SLM/MP8	100	100
SLM-44	98	98
SLM-37	77	88
SLM-33	32	62
SLM-27	3.6	22
SLM-22	1.3	3.3

[a] Nuclear RNA was isolated from cells 40 h after transfection as described in Materials and Methods and subjected to nuclease S1 analysis with 5'-end-labeled, strand-separated, single-stranded probes (Fig. 3). Band intensities from different exposures of the autoradiogram shown in Fig. 3 were determined by densitometry.

[b] The intensities of autoradiogram bands with the mVP1 probe (Fig. 3) were summed and divided by the sum of intensities of the "early" nuclease S1 bands. Values have been normalized to SLM/MP8. These data have not been corrected for any possible late promoter defects.

[c] Ratios of the relative intensities of spliced bands to unspliced bands with the mVP1 probe. Values have been normalized to SLM/MP8.

3' splice site 5' splice site

5'- ACTTCCTGCTTAACTGACTTGACATTTTCTATTTTAAGAGTGATCAAGTAAGTGA...-3' ALM

5'- ACTTCCTGCTTAACTGACTTGACATTTTCTATTTTAAAAATGATCAAGTAAGTGA...-3' ALM-3'SS

5'- ACTTCCTGCTTAACTGACTTGACATTTTCTATTTTAAGACTGATCAAGTAAGTGA...-3' ALM-4

5'- ACTTCCTGCTTAACTGACTTGACATTTTCTATTTTAAGACTGATCAAGTAAGTGA...-3' ALM-11

5'- ACTTCCTGCTTAACTGACTTGACATTTTCTATTTTAAGACTGATCAAGTAAGTGA...-3' ALM-15

FIG. 4. Viruses with mutations in the 3′ splice site upstream of the ALM late leader exon. Mutant ALM-3′SS was constructed by oligonucleotide-directed mutagenesis, and the remaining mutants shown were made from ALM by using Bal31 nuclease as described in Materials and Methods. The ALM leader unit is boxed, and the bases deleted in the various mutants are shaded. The nucleotides altered in ALM-3′SS are underlined, and splice sites flanking the leader are indicated with arrows. ALM-3′SS and ALM-4 lack the AG dinucleotides present at the normal upstream 3′ splice site. Mutants ALM-11 and ALM-15 also lack much of the polypyrimidine tract.

SLM-22) direct the synthesis of late RNAs whose secondary or tertiary structures prevent any splicing event involving the leader. This might involve, for example, sequestering 5′ or 3′ splice sites in stable RNA hairpin structures (5, 11, 19, 20). To account for our results, however, one must further propose that the RNA that is not spliced is particularly unstable in the nucleus. Also arguing against a major role of secondary structure are the similar phenotypes of the large number of mutants described here.

The defects observed in the ALM, the short-leader mutants, and the upstream deletion mutants are also unlikely to be due to a shortened distance between 5′ ends and leader splice sites. Our wild-type strain of polyomavirus (59RA) contains a 31-bp duplication (17) which places the two major late start sites 55 and 85 nucleotides upstream of the leader unit or 112 and 142 nucleotides upstream of the 5′ splice site at the distal end of the leader (3). The wild-type A2 and A3 strains of polyomavirus do not contain this duplication (17, 18). Thus, in ALM, the 31-bp duplication places many late start sites 94 nucleotides from the 5′ splice site at the distal end of the leader. This splice site is not much further away from a major start site than it would be from the corresponding site in the wild-type A2 strain of polyomavirus (112 nucleotides). When the major late 5′ ends were measured for mutant ALM-11, they were again found to be at the same sites (C. Marlor, unpublished results). Mutant SLM-27 exhibits a splicing defect and has a 30-nucleotide deletion in the late leader. This, coupled with the 31-bp duplication in our strain, places many late start sites for this mutant almost exactly the same distance from the 5′ splice site as in wild-type strain A2 (112 nucleotides). Furthermore, the cap site of a major late adenovirus type 2 transcript is only 41 nucleotides upstream of a 5′ splice site used for the production of the tripartite late leader of this virus (24). Finally, in

mVP1 Splice — M, ALM, 3'SS, ALM-4, ALM-11, ALM-15, SLM/MP8 — 306, 203

Early Splices — M, ALM, 3'SS, ALM-4, ALM-11, ALM-15, SLM/MP8 — 288, 224, 212

FIG. 5. Nuclease S1 analysis of mutations in the 3′ splice site upstream of the ALM leader exon. Cells were transfected with wild-type (SLM/MP8), ALM, or mutant DNA, and nuclear RNA was isolated 40 h later. Nuclease S1 analysis was carried out as described in the legend to Fig. 3. The mutations did not restore nuclear RNA stability or efficient mVP1 splicing.

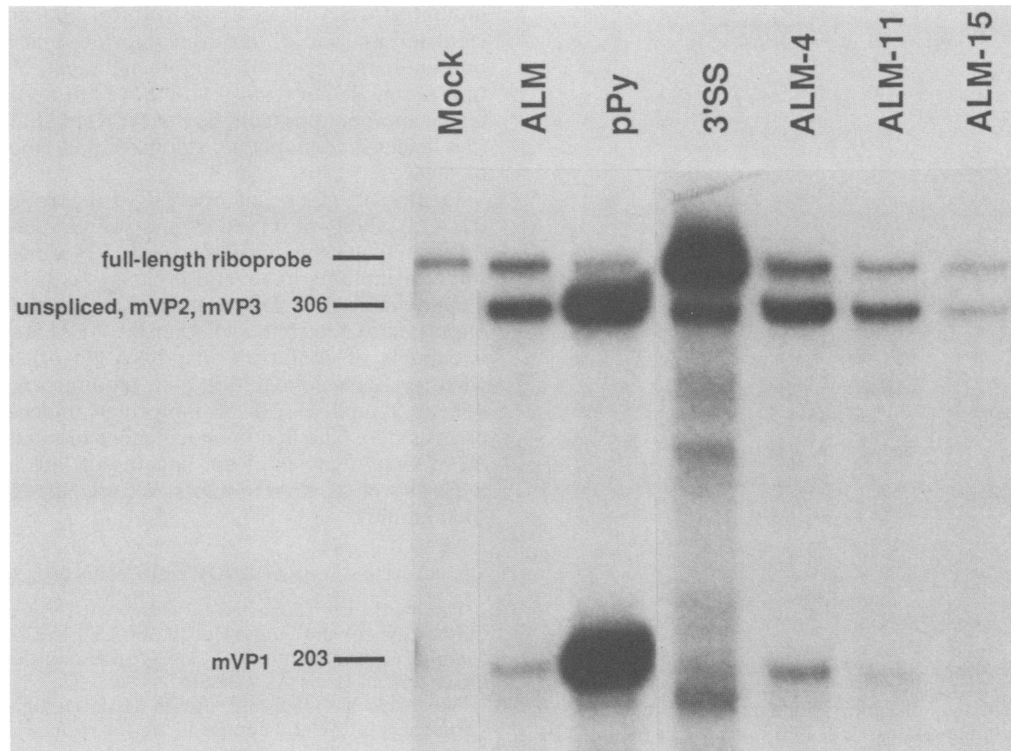

FIG. 6. Splicing defects of the leader upstream splice site mutants observed by using poly(A)$^+$ RNA. Total cell RNA from cells transfected with *Bam*HI-liberated, dilute-ligated wild-type (pPy), ALM, or other mutant DNA was isolated and subjected to oligo(dT)-cellulose chromatography, and the fraction of the mRNA molecules containing the mVP1 splice was determined by riboprobe analysis as described in Materials and Methods. The riboprobe used here was a *Hin*dIII-*Pst*I fragment, from nucleotides 3920 to 4225, cloned into the polylinker cloning site of BlueScribe. The plasmid was digested with *Eco*RI, and ^{32}P-labeled transcripts were made with phage T3 RNA polymerase. These transcripts were 348 nucleotides long, 306 nucleotides of which were complementary to polyomavirus late primary transcripts. The full-length riboprobe band represents incomplete digestion with RNase T_2. The band at 306 nucleotides represents polyomavirus transcripts not containing the mVP1 splice, and the band at 203 nucleotides represents transcripts containing the mVP1 splice. 3'SS = ALM-3'SS.

experiments to be reported elsewhere, we have analyzed the splicing patterns of other mutants with altered spacing in the leader region, and all results obtained are fully consistent with the model presented here.

The simplest explanation for the leader length dependence observed is that the defects are the result of splice site interference. A factor(s) binding at one splice site could prevent by steric hindrance the binding of a factor(s) at the other nearby splice site or could interfere with functional spliceosome assembly. This block in splicing would lead to RNA degradation. However, our system does not allow us to define a general rule for a minimum exon size required for RNA splicing and stability. For example, RNA processing for the troponin T gene has been examined by Breitbart and Nadal-Ginard (5). In this gene there are many very short exons, as small as 6 bases in length. In this system, however, the splice sites bordering the short troponin T exons may be of different efficiencies and used at different times during pre-mRNA splicing. Thus, the troponin T mini-exons may be functionally longer than 22 bases, while the polyomavirus late leader sites are not.

In an attempt to investigate the observation that splice site crowding causes RNA degradation in this system, we introduced mutations into the 3' splice site which precedes the ALM leader exon. If the splicing and RNA stability defects in ALM were coupled and were both due to splice site interference, then we would expect that mutations that eliminate one of the splice sites involved would correct both lesions. This was not the result observed. mVP1 splicing was not significantly more efficient for the upstream 3' splice site mutants than for ALM, and nuclear RNA stability was increased marginally, if at all. The conclusion is that splice site crowding inhibits leader-leader splicing. This in turn inhibits leader-mVP1 splicing and leads to RNA degradation. A block in leader-leader splicing by splice site deletion or mutation produces the same phenotype. Whether splice site crowding alone is sufficient for RNA instability, as originally postulated (3), has not been shown.

Late splice site selection in polyomavirus pre-mRNA molecules. The unexpected results obtained with the upstream leader splice site mutants have forced us to consider an important role for the late leader exon in the selection and usage of late splice sites. In giant primary transcripts, late leader and message body exons are repeated (Fig. 1). To account for the generation of mature mRNAs with multiple leaders but only one message body exon, we propose that only the coding exon closest to the 3' end of the primary transcript is chosen for splicing to the nearest upstream late leader unit. However, this cannot occur efficiently without leader-leader splicing (or splice site commitment). Thus, only after all leader-leader splices are made can efficient mVP1 or mVP3 splicing occur. This splicing prevents the usage of "internal" message body splice sites in giant transcripts and the production of giant mRNAs. When leader-leader splicing is blocked, late transcripts cannot be spliced efficiently and are channeled to a degradation path-

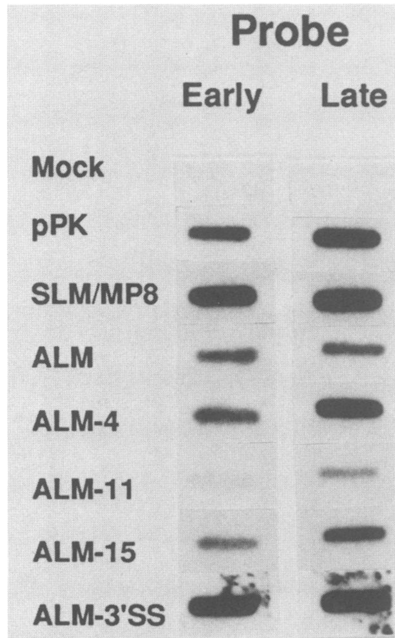

FIG. 7. Nuclear run-on assay to measure late promoter strengths of late leader upstream splice site mutants. Viral genomes propagated as *Bam*HI inserts in plasmids were excised, recircularized by ligation under dilute conditions, and used to transfect mouse NIH 3T3 cells. Transcription rate measurements were made as described in Materials and Methods. The early-transcript-specific probe, PR-3 (Hyde-DeRuyscher and Carmichael, in press), is the viral *Pst*I-*Eco*RI fragment (nucleotides 488 to 1560) cloned into phage M13mp8. The late-transcript-specific probe, RPHD11 (Hyde-De-Ruyscher and Carmichael, in press), contains a *Hin*dIII fragment from nucleotides 2921 to 3919 cloned into the *Hin*dIII site of phage M13mp18. Since ALM early gene expression is indistinguishable from that of the wild type (2), the early probes were used as an internal control. The exposure time was 71 h. The efficiency of late promoter function was determined by comparing the relative levels of signals bound to the early and late probes.

way. We are currently constructing and analyzing a number of other viral mutants that will help us to determine the order of splicing events in polyomavirus late pre-mRNA processing.

Our model helps to explain the relative paucity of late messages with single leaders. Most polyomavirus late messages have three to five tandem leader units at their 5′ ends (13, 21), even though late transcription termination or 3′ processing efficiency has been estimated to be 50% or greater (12, 23). If leader-leader splicing is favored over leader-body splicing, then giant primary transcripts could be efficiently resolved into mRNA molecules, without the usage of internal mVP1 or mVP3 splice sites, creating giant cytoplasmic messages. Thus, this model might predict that primary transcripts containing only a single leader exon should be spliced relatively inefficiently and should be unstable. Splicing of leaders, regardless of their sequence, is the important aspect of our model. Other recent work in our laboratory is consistent with this notion (Hyde-DeRuyscher and Carmichael, in press), while the results of Lanoix and Acheson (12) suggest that, at least in their system, multiple spliced leaders are not necessary for efficient virus replication. These latter authors constructed a mutant virus genome with the rabbit β-globin 3′ processing/polyadenylation signal inserted at the 3′ end of the polyomavirus late region. This mutant grows well in tissue culture, appears to direct efficient termination of late transcription, and should produce predominantly single-leader late messages. We cannot at this time reconcile our results with those of Lanoix and Acheson. Using another approach, however, Kern et al. (10) suggested that leader-leader splicing stabilizes polyomavirus late RNA molecules.

Finally, it will be of interest to determine the cause of RNA instability observed in many of our late leader mutants. A link between splicing defects and RNA instability has been seen for mutants in several other systems but has not been well defined (see Discussion of reference 3). It is also important to determine whether the ALM leader region itself is capable of conferring instability on other transcripts or whether RNA degradation also requires an efficient downstream 3′ splice site. A molecular understanding of the precise steps affected in short-leader mutants awaits a careful in vitro analysis of the splicing of late transcripts from these as well as other mutants. We are currently carrying out such studies.

ACKNOWLEDGMENTS

This work was supported by grant MV-256 from the American Cancer Society and Public Health Service grant CA 45382 from the National Cancer Institute. G.C. was an Established Investigator of the American Heart Association.

We thank Karyn Cahill, Robin Hyde-DeRuyscher, and Ellen Carmichael for helpful comments on the manuscript.

LITERATURE CITED

1. **Acheson, N.** 1978. Polyoma virus giant RNAs contain tandem repeats of the nucleotide sequence of the entire viral genome. Proc. Natl. Acad. Sci. USA **75**:4754–4758.
2. **Adami, G. R., and G. G. Carmichael.** 1986. Polyomavirus late leader region serves an essential spacing function necessary for viability and late gene expression. J. Virol. **58**:417–425.
3. **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.
4. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl (ed.).** 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
5. **Breitbart, R. E., and B. Nadal-Ginard.** 1987. Developmentally induced, muscle-specific trans factors control the differential splicing of alternative and constitutive troponin T exons. Cell **49**:793–803.
6. **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.
7. **Greenberg, M. E., and E. B. Ziff.** 1984. Stimulation of 3T3 cells induces transcription of the c-fos proto-oncogene. Nature (London) **311**:433–438.
8. **Groudine, M., M. Peretz, and H. Weintraub.** 1981. Transcriptional regulation of hemoglobin switching in chicken embryos. Mol. Cell. Biol. **1**:281–288.
9. **Hart, R. P., M. A. McDevitt, H. Ali, and J. R. Nevins.** 1985. Definition of essential sequences and functional equivalence of elements downstream of the adenovirus E2A and the early simian virus 40 polyadenylation sites. Mol. Cell. Biol. **5**:2975–2983.
10. **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.
11. **Khoury, G., P. Gruss, R. Dhar, and C.-J. Lai.** 1979. Processing and expression of early SV40 mRNA: a role for RNA conformation in splicing. Cell **18**:85–92.
12. **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.

13. **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.
14. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
15. **Nevins, J. R.** 1980. Definition and mapping of adenovirus 2 nuclear transcription. Methods Enzymol. **65**:768–785.
16. **Norbury, C. J., and M. Fried.** 1987. Polyomavirus early region alternative poly(A) site: 3′-end heterogeneity and altered splicing pattern. J. Virol. **61**:3754–3758.
17. **Ruley, H. E., and M. Fried.** 1983. Sequence repeats in a polyoma virus DNA region important for gene expression. J. Virol. **47**:233–237.
18. **Soeda, E., J. Arrand, N. Smolar, J. Walsh, and B. Griffin.** 1980. Coding potential and regulatory signals of the polyomavirus genome. Nature (London) **283**:445–453.
19. **Solnick, D.** 1985. Alternative splicing caused by RNA secondary structure. Cell **43**:667–676.
20. **Solnick, D., and S. I. Lee.** 1987. Amount of RNA secondary structure required to induce an alternative splice. Mol. Cell. Biol. **7**:3194–3198.
21. **Treisman, R.** 1980. Characterization of polyoma late mRNA leader sequences by molecular cloning and DNA sequence analysis. Nucleic Acids Res. **8**:4867–4888.
22. **Treisman, R., and R. Kamen.** 1981. Structure of polyoma virus late nuclear RNA. J. Mol. Biol. **148**:273–301.
23. **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.
24. **Zain, S., J. Sambrook, R. J. Roberts, W. Keller, M. Fried, and A. R. Dunn.** 1979. Nucleotide sequence analysis of the leader segments in a cloned copy of adenovirus 2 fiber mRNA. Cell **16**: 851–860.
25. **Zoller, M. J., and M. Smith.** 1983. Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. Methods Enzymol. **100**:468–500.