Alternative Splicing of Pre-mRNAs Encoding the Nonstructural Proteins of Minute Virus of Mice Is Facilitated by Sequences within the Downstream Intron

QIHONG ZHAO, ROBERT V. SCHOBORG, † AND DAVID J. PINTEL*

Department of Molecular Microbiology and Immunology, School of Medicine, University of Missouri—Columbia, Columbia, Missouri 65212

Received 6 December 1993/Accepted 20 January 1994

mRNAs R1 and R2 of the parvovirus minute virus of mice encode the two essential viral regulatory proteins NS1 and NS2. Both RNAs are spliced between map units 44 and 46 (nucleotides 2280 and 2399); R2 RNAs are additionally spliced upstream between map units 10 and 39 (nucleotides 514 and 1989), using a nonconsensus donor and poor 3' splice site. The relative accumulation of R1 and R2 is determined by alternative splicing: there is twice the steady-state accumulation of R2 relative to that of R1 throughout viral infection, though they are generated from the same promoter and have indistinguishable stabilities. Here we demonstrate that efficient excision of the large intron to generate R2 is dependent on at least the initial presence, in P4-generated to the size of the intervening exon. Prior splicing of the small intron is unnecessary. Excision of the large intron is also enhanced by improving the polypyrimidine tract within its 3' splice site; however, in contrast, this change renders excision of the large intron independent of the downstream small intron. We suggest that sequences within the small intron play a primary role in efficient excision of the upstream large intron, perhaps as the initial entry site(s) for an element(s) of the splicesome, which stabilizes the binding of required factors to the polypyrimidine tract within the 3' splice site of the large intron.

Alternative splicing mediated by splice site selection is an important mechanism for increasing the diversity of gene expression in eukarvotic cells (reviewed in references 13, 22, and 33). In some cases, alternative splice site selection is determined by the interaction of the general splicing machinery with nonconsensus cis-acting signals (26, 37), with cisacting signals lying at suboptimal positions (12, 14), or with secondary structures within the pre-mRNA molecule (11, 19). In other cases, alternative splice site selection has been shown to be mediated by proteins that are not constituents of the general splicing machinery; these proteins act in either a positive or a negative manner to determine alternative splice site selection (reviewed in references 20, 21, and 29). Many DNA viruses have compact genomes and extensively utilize alternative splicing to increase their coding capacity from overlapping reading frames. Minute virus of mice (MVM), a member of the autonomously replicating subgroup of DNA parvoviruses, has a relatively simple transcription profile in which alternative splicing helps determine the relative steadystate levels of all viral proteins (reviewed in references 3 and 8).

MVM is organized into two overlapping transcription units which produce three major transcript classes, R1, R2, and R3, all of which terminate near the right-hand end of the linear, 5-kb genome (Fig. 1A) (2, 5, 28, 31). Transcripts R1 (4.8 kb) and R2 (3.3 kb) are generated from a promoter (P4) at map unit (m.u.) 4 (2, 28, 31) and encode the viral nonstructural proteins NS1 (83 kDa) and NS2 (24 kDa), respectively (7), which play essential roles in viral replication and cytotoxicity (3, 8). The R2 transcripts differ from R1 by the absence of a large intron between m.u. 10 and 39 (2, 4, 16, 28, 31). This intron utilizes a nonconsensus splice donor at nucleotide (nt) 514 (AA/GCAAG) and has a poor polypyrimidine tract at its 3' splice site (TATAAATTTACTAG\), which overlaps the TATA sequence of the capsid gene promoter P38. This promoter generates the R3 (3.0-kb) transcripts, which encode the overlapping viral capsid proteins VP1 and VP2, utilizing the open reading frame (ORF) in the right half of the genome (2, 4, 16, 18, 28, 31). Three splicing patterns (major, minor, and rare) are alternatively used to excise a small intron at m.u. 44 to 46 from each transcript class, resulting in nine spliced MVM mRNA species (Fig. 1B and C) (4, 16, 23). The major splicing pattern (found in approximately 70% of the spliced molecules of each transcript class) joins D1 at nt 2280 to A1 at nt 2377; a minor splicing pattern (found in approximately 25% of molecules) joins D2 (nt 2317) to A2 (nt 2399); and a rare pattern (found in approximately 5% of molecules) joins D1 (nt 2280) to A2 (nt 2399) (Fig. 1B). The fourth potential splicing pattern, D2-A1 (nt 2317 to 2377), is not detected in vivo (23), presumably because the distance between these sites (60 nt) is shorter than the minimum suggested to be required for successful excision of introns in mammalian cells (27). Unspliced, polyadenylated R1 and R3 comprise a significant portion of viral RNA detected in both nuclear and total RNA preparations throughout MVM infection (Fig. 1B and C) and following transfection of MVM genomic plasmid clones (25, 32), and polyadenylation of MVM RNAs precedes splicing (6).

The relative steady-state levels of the nonstructural proteins NS1 and NS2, which are of critical importance for efficient viral replication, are controlled by both protein stability and the relative steady-state levels of R1 and R2 (9, 32). There is approximately twice the accumulated steady-state level of R2 relative to that of R1 during MVM infection. Since transcripts

^{*} Corresponding author. Phone: (314) 882-3920. Fax: (314) 882-4287. Electronic mail address: pintel@medsci.mbp.missouri.edu.

[†] Present address: Division of Comparative Medicine, Johns Hopkins School of Medicine, Baltimore, MD 21205.

Α.



FIG. 1. (A) Genetic map of MVM showing the three major transcript classes, R1, R2, and R3, the locations of the two promoters, P4 and P38, the ORF used by each transcript, and the proteins that each transcript encodes. The line at the bottom shows the nucleotide locations in MVM and the MVM *Hae*III probe (MVM nt 1854 to 2378) used for RNase protection assays (shown as the thick region and enlarged in panel B). (B and C) RNase protection assay of steady-state viral RNA. (B) Unspliced forms and alternative splicing patterns of the R1, R2, and R3 transcripts, and the corresponding RNase-protected regions, designated by thick lines, after hybridization to Sp6-generated antisense MVM *Hae*III probe (MVM nt 1854 to 2378). The acceptor site for the R2 large intron lies at nt 1989, and the initiation site for the R3 RNAs lies at approximately nt 2005. Three splicing patterns are alternatively used to excise the small intron at m.u. 44 to 46 common to all three transcript classes. The major pattern (found in ~70% of spliced molecules of all transcript classes) joins nt 2280 to nt 2377. The minor pattern (found in ~25% of spliced molecules of all transcript classes) joins nt 2317 to nt 2399. A rare pattern, present in ~5% of MVM spliced RNA, joins nt 2280 to nt 2399. M represents protection by those molecules that use the predominant splice donor site at nt 2280; m represents protection by those molecules in which the large intron but not the small intron have been excised (unspliced R2 [R2un]) are very rare. RNase protection assay of total RNA isolated after transfection of A9 cells with wild-type MVM genomic plasmid clone gave essentially similar results (25, 32).

R1 and R2 are both generated from P4 and have similar stabilities (32), the ratio of accumulated steady-state levels of R1 relative to R2 is dependent on the percentage of P4-generated RNA molecules from which the large intron is excised. Therefore, excision of this intron, which utilizes non-consensus 5' and 3' splice sites, is critical in determining the relative steady-state levels of NS1 and NS2 (9, 32) and thus the optimal balance between the essential roles that these proteins play in viral replication and cytotoxicity.

Detailed characterization of the accumulation of MVM RNA throughout infection, and following transfection, of murine cells has previously demonstrated the presence of abundant amounts of polyadenylated, unspliced P4- and P38generated RNAs (unspliced R1 and R3, respectively; Fig. 1B and C); P4- and P38-generated RNAs in which, individually, one of the alternatively spliced small introns between m.u. 44 and 46 have been excised (i.e., mature R1 and R3; Fig. 1B and C); and P4-generated RNAs in which both a small intron and the large intron between nt 514 and 1989 have been excised (doubly spliced, mature R2; Fig. 1B and C) (6, 25, 32). However, P4-generated RNAs in which only the large intron, but not the small intron, have been excised (designated R2un [unspliced R2] in Fig. 1C) are very rare (6, 25, 32); essentially all of the detectable P4-generated molecules that have lost the large intron have also lost the small intron (i.e., doubly spliced, mature R2; Fig. 1B and C) (6, 25, 32). These observations suggested that there may be an ordered splicing pathway for transcripts generated from the P4 promoter in which R2 RNAs are generated from spliced R1. Alternatively, if P4-generated pre-mRNA molecules are first spliced between nt 514 and 1989, these spliced molecules must have extremely short half-lives and the small intron in these molecules must be very quickly excised to generate doubly spliced, mature R2.

In this report, we describe experiments which characterize alternative splicing of the MVM P4-generated RNAs. We demonstrate that efficient excision of the large intron is dependent on at least the initial presence, in P4-generated pre-mRNAs, of sequences within the downstream small intron. These required small intron sequences contain a splice donor and acceptor site, yet efficient excision of the large intron does not require splicing of the small intron. When the poor polypyrimidine tract of the 3' splice site was improved, excision of the large intron was enhanced, and its efficient excision became independent of the downstream small intron sequences. These results suggest that there is a primary requirement of small intron sequences for the efficient excision of the upstream large intron, perhaps as the initial entry site(s) for an element(s) of the splicing apparatus, which may stabilize the binding of essential splicing factors to the polypyrimidine tract within the 3' splice site of the large intron. Because the processing pathway for the parvovirus MVM pre-mRNAs is relatively simple and well defined, it offers an excellent model system particularly valuable for defining a mechanism by which alternative splicing of pre-mRNAs is determined by an interaction between introns.

MATERIALS AND METHODS

Plasmid construction. cDNA joining either D1 (nt 2280) to A1 (nt 2377) or D2 (nt 2317) to A2 (nt 2399) was generated previously (4). To make pMVM(D1-A1) and pMVM(D2-A2), a restriction fragment from nt 2072 to 2652 (XhoI-HindIII) containing either of these cDNAs was inserted into the genomic clone of MVM. The 2377–3636 deletion in $p\Delta S/P$ was constructed by digesting the genomic clone of MVM with StuI (nt 2377) and PmlI (nt 3636), filling in the 3' ends with the Klenow fragment of DNA polymerase I, and religating. Deletions between the NarI (nt 2290) and the PmlI site ($p\Delta N/P$), between the NarI and HpaI (nt 3757) site ($p\Delta N/H$), and between the AseI (nt 2350) and PmlI site ($p\Delta A/P$) were constructed similarly. $p\Delta S$ deletes the StuI fragment (nt 2377 to 2500) from the genomic clone of MVM. The nucleotide sequence of D1 (G/GTACG) (nt 2278 to 2383) was changed to C/CAACT by site-directed mutagenesis using M13mp18 into which the MVM EcoRI fragment (nt 1086 to 3521) had been cloned, exactly as previously described (25). The mutation was confirmed by sequencing and reinserted into the genomic clone of MVM to create pssD1(-). Double mutant pssD1(-)- Δ S/P was created by deleting MVM nucleotide sequences from StuI to PmlI from pssD1(-). $p\Delta N/H$ -mss(+) and $p\Delta N/H-mss(-)$ were constructed by insertion of the MVM Sau3A restriction fragment, spanning nt 2203 to 2508, into the BgIII site of $p\Delta N/H$ at nt 4212 in the sense and antisense orientations, respectively. Mutagenesis of the splice donor site at nt 514 was accomplished by oligonucleotide mutagenesis of a portion of MVM (nt 0 to 1086) cloned in M13mp18, exactly as previously described (25). Nucleotides 513 to 517 were changed from AA/GCA in the wild type to the consensus AG/GTA. The mutation was confirmed by sequence analysis and reinserted into the genomic clone of MVM to create pMVM(514,16). $p\Delta N/P(514,16)$ and pMVM(D1-A1)-(514,16) were created by the inserting EcoRV-XhoI restriction fragment (nt 385 to 2072) from pMVM(514,16) into either $p\Delta N/P$ or pMVM(D1-A1). p652(T4) was created by introducing an ochre termination codon at the XcmI site (nt 652) which terminates the NS1 ORF (ORF3) within the large intron, as previously described (25). The nucleotide sequence of the large intron polypyrimidine tract was changed by site-directed mutagenesis using M13mp18 into which the MVM EcoRI fragment (nt 1086 to 3521) had been cloned, exactly as previously described (25). The mutation was confirmed by sequencing and reinserted into the genomic clone of MVM to create p4Tppt, which was also confirmed by sequencing. $p\Delta N/$ P(4Tppt) was created by inserting the BstEII-XhoI (nt 1886 to 2072) restriction fragment from p4Tppt into $\Delta N/P$ and also confirmed by sequencing.

Transfections, RNA isolation, RNase protection assays, and RNA stability assays. Murine A92L cells, the normal tissue culture host for MVM, were grown as previously described (28) and transfected with wild-type and mutant MVM plasmids by using DEAE-dextran exactly as previously described (25). RNA was typically isolated 48 h posttransfection, following lysis in guanidinium hydrochloride, by centrifugation through CsCl exactly as previously described (32). RNase protection assays were performed as previously described (32). The probe used for RNase protection assays was an $[\alpha^{-32}P]UTP$ -labeled, Sp6-generated antisense MVM RNA from nt 1854 to 2378 (corresponding to the MVM HaeIII restriction fragment, diagrammed as A in Fig. 1A). This probe (referred to as the MVM HaeIII probe) extends from before the acceptor site of the large intron to within the small intron common to all MVM RNAs at m.u. 44 to 46 and distinguishes RNA species using either of the alternative small intron donors, designated M for the major splice donor (D1) at nt 2280 and m for the minor splice donor (D2) at nt 2317 (Fig. 1B) (6). For analysis of RNA produced after transfection with pssD1(-) or pssD1(-)- Δ S/P, MVM HaeIII probe-ssD1(-) was used, which is identical to the MVM HaeIII probe except that nucleotide sequence GGTACG (from nt 2278 to 2383) had been changed to CCAACT to be homologous with the mutants used. For analysis of RNA produced after transfection with P4Tppt and $p\Delta N/P(4Tppt)$, an MVM HaeIII probe which contained homogous changes in the polypyrimidine tract was used. RNase protection assay data were analyzed on a Betagen β-scanning phosphor image analyzer, and molar ratios of MVM RNA were determined by standardization to the number of uridines in each protected fragment. Unless otherwise stated in the figure legends, total R1 and R2 (unspliced and spliced R1 and R2) were included in the quantitation of each analysis. RNA stability assays using actinomycin D were performed as previously described (32); briefly, following transfection, cells were washed three times with phosphate-buffered saline and refed with Dulbecco modified Eagle medium plus 5% fetal calf serum, either with or without 40 µg of actinomycin D (Sigma Chemical Co., St. Louis, Mo.) per ml. At 3 and 6 h after addition of the transcription inhibitor, duplicate plates were used for isolating total cellular RNA. The quantities of specific MVM transcripts remaining at each time point were determined by RNase protection assay. These conditions have been previously determined to efficiently inhibit MVM transcription (32).

RESULTS

Efficient excision of the large intron from MVM P4-generated RNA requires sequences within the downstream small



FIG. 2. (A) Map of the small intron cDNA constructs and deletion mutants. The three splicing patterns (M, m, and r) used to excise the small intron are shown on the top as described in the legend to Fig. 1. The two small intron cDNA constructs are shown immediately below; pMVM(D1-A1) removes the small intron joining the donor site at nt 2280 and the acceptor site at nt 2377, and pMVM(D2-A2) removes the small intron joining the donor at nt 2317 and the acceptor at nt 2399. Locations of the small intron donor and acceptor sites and the relevant restriction sites for making deletion mutants in the small intron region are shown on the middle line. All of the deletion mutants are shown below relative to the restriction sites used to make the deletions as described in Materials and Methods. Mutant psD1(-), which destroys D1 by site-directed mutagenesis, is also indicated on the bottom. (B) RNase protection assay showing that the large intron of R2 is inefficiently excised from RNAs generated by the small intron cDNA constructs pMVM(D1-A1) and pMVM(D2-A2). Shown is an RNase protection analysis, using the MVM *Hae*III probe (see Fig. 1B), of 20 μ g of total RNA isolated from A9 cells 48 h after transfection with 20 μ g of wild-type MVM (pMVM) or the small intron cDNA per dish or mock transfection, as indicated. The identities of the protected bands for RNA generated by wild-type pMVM are shown. R1, R2, and R3 generated by utilizing only the minor (m) splicing pattern for pMVM(D1-A1) or the major splicing pattern (M) for pMVM(D2-A2) can be seen. The quantitative results of the ratios of R1 to R2 from three separate experiments are shown in Table 1. (C)

intron. To examine the alternative splicing of P4-generated RNA in more detail, we initially measured the accumulation of the doubly spliced, mature R2 RNA following transfection of MVM genomic plasmid clones in which the small intron region was either partially or completely absent (Fig. 2; Table 1). cDNA clones pMVM(D1-A1) and pMVM(D2-A2) (4), from which the small intron, utilizing either the major (M; 2280 to 2377) or minor (m; 2317 to 2399) splicing pattern, respectively, had been lost (Fig. 2A) but were otherwise wild type, were significantly impaired in generating R2 RNA following transfection. Whereas wild-type MVM generated approximately twice the accumulated steady-state level of R2 relative to that of R1 following transfection, pMVM(D1-A1) and pMVM(D2-A2), which lack small intron sequences, accumulated two- to fourfold more R1 relative to R2, although the ratios of total P4 (R1 + R2) to P38 products (R3) were approximately the same as that of the wild-type MVM (Fig. 2B; Table 1). These results suggested that either the region between and including the small intron internal donor D2 (nt 2317) and acceptor A1 (nt 2377) or a prior splicing event in this region was critical for the efficient accumulation of the doubly spliced, mature R2 (Fig. 2A).

Analyses of the relative accumulation of R1 and R2 generated by a series of mutants in which various portions of the small intron region were deleted (Fig. 2A) gave similar results. $p\Delta S$ and $p\Delta S/P$, in which A1 at nt 2377 remained intact but sequences downstream of A1 were deleted, both generated RNA from which the small intron was efficiently excised and produced wild-type levels of the doubly spliced, mature R2 RNA (Fig. 2C; Table 1). In contrast, $p\Delta A/P$, in which both small intron acceptors were deleted, and $p\Delta N/P$, in which D2 as well as both small intron acceptors had been deleted, generated RNA from which the small intron was inefficiently excised and the large intron was also inefficiently spliced (Fig. 2C; Table 1). The excision of the large intron was more deficient from RNA generated by $p\Delta N/P$ than by $p\Delta A/P$. In addition, the existing small intron donor at nt 2280 (AG/ GTACG) in the wild-type clone was changed, by site-directed mutagenesis, to AG/CAACT to create pssD1(-). Although this change prevented excision of the small intron using this splice donor (i.e., using the major splicing pattern), the small intron using the remaining donor (D2) at nt 2317 (i.e., using the minor splicing pattern) was efficiently excised, and the large intron was also efficiently excised from RNA generated by this mutant to produce wild-type levels of doubly spliced, mature R2 (Fig. 3; Table 1). Since the large intron can be efficiently excised from RNA generated by deletion mutants $p\Delta S/P$ and $p\Delta S$, but not by $p\Delta A/P$, sequences between nt 2350

 TABLE 1. Ratios of accumulated R1 and R2 following transfection of MVM small intron mutants

Construct	Avg R1/R2 \pm 95% confidence limit (no. of expts) ^a						
pMVM	0.50 ± 0.13 (14)						
pΔS/P	0.51 ± 0.15 (9)						
pΔS	0.26 ± 0.09 (5)						
pΔA/P	1.38 ± 0.21 (5)						
pΔN/P	2.46 ± 0.08 (15)						
pssD1(-)	0.54 ± 0.14 (6)						
pMVM(D1-A1)	3.34 ± 0.31 (3)						
pMVM(D2-A2)	2.0 ± 0.14 (3)						

^{*a*} RNase protection experiments were done as described in Materials and Methods, using the probe described in the legend to Fig. 1.

and 2377, which include the small intron internal acceptor A1, must be required for efficient excision of the large intron; presence of either donor alone (or both together) does not suffice. In addition, since the excision of the large intron from RNA generated by $p\Delta N/P$ is more deficient than that by $p\Delta A/P$, sequences between 2290 and 2350, which include the small intron internal donor D2, must also participate.

The decreased ratio of accumulated $\hat{R}2$ relative to R1 generated by these mutants was not due to alterations in RNA stability. All of the RNAs generated by wild-type MVM (Fig. 2D) (32), $p\Delta N/P$ (which has the largest deletion in the small intron region; Fig. 2D), and the cDNA mutants (data not shown) were similarly stable, at least for 6 h in the presence of actinomycin D under conditions previously established to efficiently inhibit MVM transcription following transfection (25, 32). Therefore, the ratio of accumulated level of R2 relative to that of R1 represents the percentage of molecules that have successfully excised the large intron.

Prior splicing of the small intron is not necessary for efficient excision of the upstream large intron. Results presented so far have demonstrated that small intron sequences, which include the small intron internal donor and acceptor sites (D2-A1; nt 2317 to 2377), are required for the efficient excision of the upstream large intron. Since the internal small intron donor and acceptor pair (D2-A1) is not utilized during viral infection (23) or following transfection of genomic plasmid clones, presumably because the distance between these sites (60 nt) is shorter than the minimum suggested to be required for successful excision of introns in mammalian cells (27), we were able to address whether sequences within the small intron were merely required in *cis*, or whether prior

RNase protection assays of RNAs generated by the small intron deletion mutants confirm results with the cDNA constructs. Shown is an RNase protection analysis, using the MVM HaeIII probe, of 20 µg of total RNA isolated from A9 cells 48 h after transfection with 20 µg of the wild-type MVM (pMVM) or the deletion mutant plasmid DNA (see panel A) per dish or mock transfection, as indicated. The identities of the protected bands for RNA generated by wild-type pMVM (shown on the right) are as described for Fig. 1C. pAS/P and pAS delete A2; therefore, the small intron can be excised only by joining D1 at nt 2280 to A1 at nt 2377 to produce the major (M) forms of transcripts R1, R2, and R3. For RNA generated by pAA/P which deletes both acceptors and therefore from which the small intron cannot be excised, the sizes of the protected bands (shown on the left) for R1, R2, and R3 are 496, 360, and 345 nt, respectively. For RNA generated by pΔN/P, which deletes D2 in addition to both acceptors of the small intron, the sizes of the protected bands (shown on the left) for R1, R2, and R3 are 436, 300, and 285 nt, respectively. The quantitative results of the ratios of R1 to R2 for these mutants are shown in Table 1. (D) Effects of the small intron deletion mutants on MVM RNA splicing are independent of transcript stability. Conditions for assaying stability of MVM RNAs following transfection have been previously established (32) (see Materials and Methods). At 48 h after transfection, A9 cells were treated with 40 µg of actinomycin D per ml, and total cellular RNA was collected 3 and 6 h later. Shown is an RNase protection assay, using the MVM HaeIII probe of 20 µg of total RNA isolated, following actinomycin D treatment, from A9 cells transfected with 20 μ g of wild-type MVM (pMVM) or p Δ N/P (mutant with the largest deletion in the small intron) plasmid DNA per dish or mock transfected, as indicated. The identities of the protected bands for pMVM (wild type) are as described for Fig. 1C, and those for p $\Delta N/P$ are as described for Fig. 2C. Unspliced R1 and unspliced R3 are processed to mature forms within 0.5 h after drug addition (32).



excision of the small intron was required, for the efficient excision of the upstream large intron. As discussed above (Fig. 2C; Table 1), both the small intron and the large intron were excised efficiently from RNAs generated by mutant pssD1(-), which destroys only D1, and from RNA generated by mutant $p\Delta S/P$, which deletes only A2 (Fig. 3B). P4 RNAs generated by the double mutant pssD1(-)- $\Delta S/P$, however, which lack both the external donor (D1) and acceptor (A2) of the small intron (Fig. 3A), as predicted, did not utilize the remaining internal donor and acceptor pair D2-A1 (Fig. 3B). Yet even in the absence of excision of the small intron, pssD1(-)- Δ S/P generated a wild-type level of R2 relative to that of R1 (Fig. 3B; the lower amount of R3 generated by this mutant is presumably a consequence of poor transport of unspliced R1 to the cytoplasm, resulting in reduced amounts of NS1 required for P38 transactivation). Therefore, sequences within the small intron can function in *cis* to facilitate the efficient excision of the upstream large intron; i.e., excision of the small intron is not required. These experiments do not imply that the large intron is necessarily excised from the small intron-containing unspliced R1 in vivo but rather demonstrate that the presence of small intron sequences in *cis* is sufficient to promote efficient excision of the large intron from these molecules. It is unlikely that efficient excision of the large intron from RNA generated by $pssD1(-)-\Delta S/P$ is merely an indirect consequence of the inability to splice the small intron (resulting, for example, in

Α.



FIG. 4. Efficient excision of the large intron can be rescued by orientation-dependent insertion of an MVM restriction fragment containing the small intron. (A) Map of $p\Delta N/H$ -mss(+) and $p\Delta N/H$ mss(-). (B) RNase protection assay, using the MVM HaeIII probe, of 20 µg of total RNA isolated 48 h after transfection of A9 cells with 20 µg of wild-type (wt) MVM (pMVM) or mutant plasmid DNA per dish or mock transfection, as indicated. For RNA generated by pMVM, the identities of the protected bands (shown in the middle) are as described for Fig. 1C. For RNA generated by $p\Delta N/H$ and $p\Delta N/H$ -mss(-), in which the remaining original donor at nt 2280 is not used, the protected bands are the same as those for $p\Delta N/P$, as in Fig. 2C. For $p\Delta N/H$ mss(+)-generated RNAs in which the remaining original donor at nt 2280 is used, the protected bands are R1, R2M, and R3M; for those RNAs in which the remaining original donor is not used, the protected bands are the same as those for $p\Delta N/P$, $p\Delta N/H$, and $p\Delta N/H$ -mss(–), i.e., R1, R2, and R3. These protected fragments are 10 nt larger than R1M, R2M, and R3M, respectively, which use the donor at nt 2280. Note that $p\Delta N/H$ -mss(+)-generated R1 molecules in which the remaining original donor site at nt 2280 was used were not detected (see text). Quantitation of the ratio of total R1 to total R2 from four separate experiments, as determined by β-scanning phosphor image analysis, is shown at the bottom. For RNA generated by $p\Delta N/H$ mss(+), the results of the ratio of R1 to R2 (molecules in which the remaining donor at nt 2280 is not used) and those of R1 to R2M (those in which the remaining donor is used) from four separate experiments are also shown at the bottom; 95% confidence limits are included.

donor(s) and acceptor(s) entirely within the insertion (data not shown). The large intron, however, was excised to a much greater extent from RNA in which the remaining original small intron donor at nt 2280 was used (R2M), thus leaving the size of the upstream intervening exon unaltered [Fig. 4; compare R1/R2 (R2 does not use the remaining original donor at nt 2280) with R1/R2M (R2M uses the remaining original donor at nt 2280) for $p\Delta N/H-mss(+)$]. That the large intron is excised much more efficiently from molecules using the remaining original donor at nt 2280 suggests that the size of the intervening exon can affect the interaction between these two introns. In addition, $p\Delta N/H$ -mss(+)-generated R1 molecules in which the remaining original donor site at nt 2280 was used were not detected, suggesting that R1 molecules in which this site had been used were efficiently processed to R2M.

Changing the large intron splice donor to consensus results in an increased ratio of accumulated level of mature R2 relative to that of mature R1, but only in the presence of the downstream small intron. The large intron splice donor at nt 514 (AA/GCAAG) is nonconsensus and is a poor match for the 5' region of U1 small nuclear RNA. When this donor was changed to consensus (AG/GTAAG) so that complementarity to the 5' region of the U1 small nuclear RNA was improved, in an otherwise wild-type construct, to create pMVM(514,16), the ratio of accumulated mature R2 relative to mature R1 was increased, suggesting that excision of the large intron was accelerated [Fig. 5B; compare the ratios of mature R1 to mature R2 for pMVM versus pMVM(514,16)]. [This mutation

В.



increased retention of the small intron-containing unspliced R1 in the nucleus), since, as shown above, the large intron is inefficiently excised from P4 transcripts which are generated by cDNA and deletion mutant constructs and from which the small intron cannot be excised (Fig. 2; Table 1).

Efficient excision of the large intron can be rescued by insertion of an MVM restriction fragment containing the small intron. Since the MVM small intron can function in *cis* to facilitate the efficient production of the doubly spliced, mature R2 RNA, we examined whether efficient excision of the large intron, in a small intron deletion mutant, could be rescued when the small intron was inserted at another site downstream in the viral genome. Efficient excision of the large intron was rescued from RNA generated by $p\Delta N/H$ -mss(+), in which an MVM Sau3A restriction fragment (nt 2203 to 2508) containing all small intron donor and acceptor sites was inserted into $p\Delta N/H$, 455 nt downstream of the NarI-HpaI deletion junction (Fig. 4A; the deletion in $p\Delta N/H$ extends to nt 3757 in the capsid gene region [compared with the deletion in $p\Delta N/P$, which extends to nt 3636], and the splicing pattern of RNAs generated by $p\Delta N/H$ is identical to that of $p\Delta N/P$ [see Materials and Methods]). However, efficient excision was recovered only when the MVM Sau3A fragment was inserted in the same orientation as in the viral genome, such that the small intron sequences appear in the RNA (i.e., the sense orientation; Fig. 4B).

RNAs generated by $p\Delta N/H$ -mss(+) are spliced between the remaining original small intron donor at nt 2280 and acceptor(s) in the insertion (Fig. 4B), as well as between small intron



	AG/GTAAG							
pMVM(514,16)				D1 D2 A	1 A2			
	514	Angla	1989		3000		4000	5000
p∆N/P(514,16) –	514		1989	D1	Δ2290-3636	_		
	AG/GTAAG							
pMVM(D1-A1)-(514,16)	514		1989		A2			
	A G /GTAAG			D1/A1				



also introduces a translation termination signal into the NS1 ORF (ORF3) within the large intron. Translation termination codons in the exons flanking the large intron inhibit its excision (25); however, as previously shown (25) and as shown in Fig. 5 for mutant p652(T4), in which an ochre codon was introduced at nt 652, translation termination mutations within the large intron have no such effect.]

For RNAs generated by mutants $p\Delta N/P$ and pMVM(D1-A1), both of which lack small intron sequences, excision of the large intron was not enhanced by changing its donor site to consensus [mutants $p\Delta N/P(514,16)$ and pMVM(D1-A1)-(514,16); Fig. 5; compare the ratios of total R1 to total R2]. This result suggests that there is a primary requirement of the

FIG. 5. Changing the large intron donor to consensus results in an increased ratio of accumulated level of mature R2 relative to that of mature R1, but only in the presence of the small intron. (A) pMVM(514,16) was created by changing the nonconsensus donor at nt 514 to consensus following site-directed mutagenesis as described in Materials and Methods. Double mutant $p\Delta N/P(514,16)$ or pM-VM(D1-A1)(514,16) was created by introducing the large intron consensus donor mutation into $p\Delta N/P$ or pMVM(D1-A1), which lacks the small intron, as described in Materials and Methods. (B) RNase protection assay, using the MVM HaeIII probe, of 20 µg of total RNA isolated 48 h after transfection of A9 cells with 20 µg of wild-type MVM (pMVM) or mutant plasmid DNA per dish or mock transfection, as indicated. p652(T4) was created by introducing an ochre termination codon into the NS1 ORF (ORF3) within the large intron as previously described (25). Note that there is an increased level of unspliced R1 generated by the consensus donor mutant pMVM (514,16). The identities of the protected bands are as described in the legends to Fig. 2B and C. Quantitation of the ratio of total R1 to total R2 from five separate experiments, determined by β-scanning phosphor image analysis, is shown at the bottom and includes 95% confidence limits. For RNA generated by pMVM and pMVM(514,16), quantitation of the ratios of mature R1 to mature R2 and of unspliced R1 to mature R1 from five separate experiments is also shown. n/a, not applicable, since these mutants make only one species of R1.

small intron sequences for the efficient excision of the large intron, regardless of the nature of the large intron donor. Changing the large intron donor at nt 514 to consensus in the absence of the small intron led to an even greater reduction in the generation of the doubly spliced, mature R2 than that seen in the parent mutants $p\Delta N/P$ and pMVM(D1-A1), which bear the nonconsensus donor (Fig. 5), suggesting that this change had an inhibitory effect on excision of the large intron in the absence of the small intron. In addition, pMVM(514,16)generated an increased amount of unspliced R1 [such that the total accumulated R1 relative to R2 for pMVM(514,16) was similar to that of the wild-type; Fig. 5; compare the ratio R1 unspliced to R1 mature for pMVM versus pMVM(514,16)].

Improving the polypyrimidine tract at the 3' splice site of the large intron facilitates its excision and confers its efficient excision independent of the downstream small intron sequences. The large intron also contains a poor polypyrimidine tract (TATAAATTTACTAG\) within its 3' splice site. When the pyrimidine content of this region was improved (TA TTTTTTTTCCAG\), in an otherwise wild-type clone, to create p4Tppt (Fig. 6A), the ratio of accumulated levels of R2 relative



limits.

to R1 was increased by approximately ninefold, suggesting that excision of the large intron was rendered more efficient (Fig. 6B).

In contrast to changes at the large intron donor site, improving the polypyrimidine tract of the large intron rendered its efficient excision independent of downstream small intron sequences. As shown above, excision of the large intron from RNA generated by $p\Delta N/P$, which lacks small intron sequences, was inefficient. In the experiment shown in Fig. 6B, the R1/R2 ratio for RNA generated by $\Delta N/P$ was slightly less than the average value shown in Table 1. However, when the polypyrimidine tract of the large intron in $p\Delta N/P$ was improved $[p\Delta N/P(4Tppt);$ Fig. 6A], efficient excision of the large intron was regained (Fig. 6B). Excision of the large intron from RNA generated by $p\Delta N/P(4Tppt)$ was four times more efficient than that from RNA generated by wild-type pMVM, although only approximately half as efficient as that for RNA generated by p4Tppt. These results suggest that downstream small intron sequences are required to compensate for a poor 3' splice site, perhaps as an entry site for an element(s) of the splicing machinery which may stabilize the binding of essential splicing factors to the polypyrimidine tract of the large intron.

DISCUSSION

The P4-generated pre-mRNAs of the autonomous parvovirus MVM undergo alternative splicing which determines the ratio of the accumulated levels of mRNAs R1 relative to R2 (6, 16, 23, 25, 28, 32). Since mature P4 transcripts R1 and R2 generated by wild-type MVM and small intron mutants are similarly very stable (Fig. 2D) (32), the accumulated ratio of R2 relative to R1 represents the percentage of molecules that have successfully excised the large intron. This event is critical in determining the relative steady-state levels of the viral nonstructural proteins NS1 and NS2 and therefore in optimizing the balance between the essential roles that these two proteins play in viral replication and cytotoxicity. No viral protein participates in this excision, which must therefore be accomplished solely by the interactions between cellular factors and viral *cis*-acting signals.

In this report, we have shown that efficient production of the doubly spliced, mature R2 RNA of MVM depends on at least the initial presence of sequences within the downstream small intron. Sequences within the MVM small intron required for efficient excision of the upstream large intron include the small intron internal donor and acceptor sites; however, the large intron is efficiently excised from RNAs which include these sequences but from which the small intron cannot be excised (Fig. 3). This finding demonstrates that a prior splicing event within the small intron is unnecessary and is consistent with the hypothesis that the small intron acts as a primary entry or binding site for the splicing apparatus. Although one possible consequence of the inability of pssD1(-)- Δ S/P-generated RNAs to splice the small intron in these experiments is that these pre-mRNAs could be sequestered in the nucleus, it is unlikely that efficient excision of the large intron is an indirect consequence of the retention of unspliced R1 in the nucleus. The inability to excise the small intron does not necessarily lead to efficient excision of the large intron, since the large intron is inefficiently excised from P4 transcripts generated by pMVM(D1-A1), $p\Delta N/P$, and $p\Delta A/P$, in which the small intron is not excised.

These experiments do not imply that in vivo the large intron is necessarily excised directly from unspliced R1 but rather demonstrate that the presence of small intron sequences, in *cis*, is sufficient to promote the efficient excision of the large intron. In contrast to other well-characterized systems, in which only one splice donor site of the downstream intron is necessary and sufficient for the efficient excision of the upstream intron (26, 30, 35), results with MVM small intron mutants demonstrate that sequences in addition to the splice donor(s) of the downstream small intron are required for efficient excision of the upstream intron. Perhaps this is due to the nature of the small intron, which itself is alternatively spliced in a complex manner. A detailed analysis designed to determine the nature of the sequences within the small intron which are required for excision of the large intron is currently in progress.

Efficient excision of the large intron can be regained from RNA generated by a small intron deletion mutant into which a restriction fragment containing the MVM small intron was inserted downstream of the deletion junction in an orientationdependent manner (so that small intron signals appear in the RNA). RNAs generated by $p\Delta N/H$ -mss(+) are spliced between the remaining original small intron donor at nt 2280 and acceptor(s) in the insertion and between small intron donor(s) and acceptor(s) entirely within the insertion; however, the large intron is excised to a much greater extent from RNA in which the remaining original small intron donor at nt 2280 has been joined to an acceptor within the downstream insertion, leaving the size of the upstream intervening exon unaltered. Since both types of molecules were available, this observation suggests that the size of the intervening exon affects the interaction between these two introns.

Surprisingly, significantly more R2M than R1 was generated by $p\Delta N/H$ -mss(+) than by wild-type MVM. This could be due to the presence of three donor sites downstream of the large intron in this molecule, which might more efficiently recruit element(s) of the spliceosome or stabilize their interaction with the pre-mRNA, or alternatively, splicing between the remaining donor site at nt 2280 and an acceptor in the downstream insertion is more efficient than that between donors and acceptors entirely within the insertion, and this may also contribute to efficient excision of the large intron. The requirement of the remaining donor at nt 2280 was not absolute, however, because excision of the large intron was rescued to a level approaching that of the wild type when the remaining small intron D1 site of $p\Delta N/H$ -mss(+) was altered [to create pssD1(-)/ Δ NH-mss(+)] (data not shown). This result is not in disagreement with the conclusion that the size of the intervening exon affects the interaction between these two introns, because we have shown that sequences in addition to the small intron donor site(s) affect excision of the upstream large intron (Fig. 2 and 3; Table 1); perhaps the altered D1 site in RNA molecules generated by pssD1(-)/ Δ N/H-mss(+) retains the ability to bind element(s) of the spliceosome. Alternatively, a cryptic splice donor upstream of the small intron insertion may be used in RNAs generated by pssD1(-)/ Δ N/H-mss(+).

Why does efficient excision of the MVM large intron require downstream small intron sequences? This intron must be either intrinsically deficient, as has been shown for intervening sequence 4 of the rat preprotachykinin pre-mRNA (15), or actively repressed, as has been shown for the Rous sarcoma virus genomic RNA (1, 17, 34), for RNA encoding the yeast ribosomal protein L32 (10), and for RNA encoded by the *Drosophila* sex lethal gene (reviewed in references 20 and 21).

The requirement of downstream small intron sequences for excision of the upstream large intron is not a consequence of the presence of the nonconsensus donor site of the large intron. Although changing this site to consensus in the presence of small intron sequences [pMVM(514,16)] enhances splicing of the large intron, this change does not enhance excision of the large intron in the absence of small intron sequences [mutants $p\Delta N/P(514,16)$ and pMVM(D1-A1)-(514,16)] but rather leads to an even greater decrease in the production of the doubly spliced, mature R2 than that seen in the parent mutants bearing the nonconsensus donor. While sequences within the small intron play a primary role in the efficient excision of the upstream large intron, the nature of the large intron donor site is still critical, in wild-type MVM, for determining the relative accumulated levels of R1 and R2 and therefore for optimizing the relative steady-state levels of the two essential viral nonstructural proteins NS1 and NS2.

The MVM large intron also contains a very poor polypyrimidine tract. In contrast to changes at the large intron donor, improving the polypyrimidine tract of the large intron rendered its efficient excision independent of downstream small intron sequences $[p\Delta N/P(4Tppt);$ Fig. 6]. Although improvement of either the large intron donor or 3' splice site in the presence of the small intron resulted in enhanced excision of the large intron to levels greater than that seen for the wild type, only improvement of the large intron 3' splice site rendered its efficient excision independent of downstream small intron sequences. This finding suggests that downstream small intron sequences are required to compensate for an otherwise

3' splice site, perhaps as an initial entry site for elements of the splicing machinery which stabilize the binding of essential splicing factors to the polypyrimidine tract of the large intron. In a perhaps similar situation, U1 binding to the downstream 5' donor site in pre-mRNA generated by the rat preprotachykinin gene (15) has been shown to be required for efficient excision of an upstream intron by stabilizing the binding of the 65-kDa subunit of U2AF to a poor polypyrimidine tract.

In MVM, the purine-rich TATA box of the viral P38 promoter, which generates RNAs coding for the viral capsid proteins, is positioned within the polypyrimidine tract of the 3' splice site of the large intron. Perhaps the interaction between this poor polypyrimidine tract and downstream small

intron sequences is an adaption in order to successfully accomplish both required functions in the highly compact MVM genome.

There is growing evidence that efficient excision of introns can be facilitated by downstream introns (26, 30, 35), presumably through protein-protein interactions across the intervening exon (15). Because the processing pathway for parvovirus MVM pre-mRNAs is relatively simple and well defined, the results presented here establish a system which will be particularly valuable for characterizing a mechanism by which alternative splicing of pre-mRNAs is controlled in such a manner and exons are defined in mammalian cells.

ACKNOWLEDGMENTS

We thank Greg Tullis and Mark Hannink for helpful discussion and important suggestions throughout the course of this work and Lisa Burger for excellent technical assistance.

This work was supported by NIH grants AI21302 and AI00934 to D.J.P. Q.Z. and R.V.S. were partially supported by the University of Missouri Molecular Biology Program.

REFERENCES

- Arrigo, S., and K. Beeman. 1988. Regulation of Rous sarcoma virus RNA splicing and stability. Mol. Cell. Biol. 8:4858–4867.
- Astell, C. R., E. M. Gardiner, and P. Tattersall. 1986. DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and comparison with the DNA sequence of the fibrotropic prototype strain. J. Virol. 57:656-669.
- Berns, K. I. 1990. Parvovirus replication. Microbiol. Rev. 54:316– 329.
- Clemens, K. E., D. R. Cerutis, L. R. Burger, G. Q. Yang, and D. Pintel. 1990. Cloning of minute virus of mice cDNAs and preliminary analysis of individual viral proteins expressed in murine cells. J. Virol. 64:3967–3973.
- Clemens, K. E., and D. Pintel. 1987. Minute virus of mice (MVM) mRNAs predominantly polyadenylate at a single site. Virology 160:511-514.
- Clemens, K. E., and D. Pintel. 1988. The two transcription units of the autonomous parvovirus minute virus of mice are transcribed in a temporal order. J. Virol. 62:1448–1451.
- Cotmore, S. F., and P. Tattersall. 1986. Organization of nonstructural genes of the autonomous parvovirus minute virus of mice. J. Virol. 58:724–732.
- Cotmore, S. F., and P. Tattersall. 1987. The autonomously replicating parvoviruses of vertebrates. Adv. Virus Res. 33:91–174.
- Cotmore, S. F., and P. Tattersall. 1990. Alternate splicing in a parvoviral nonstructural gene links a common amino-terminal sequence to downstream domains which confer radically different localization and turnover characteristics. Virology 177:477–487.
- Dabeva, M. D., M. A. Post-Beittenmiller, and J. R. Warner. 1986. Autogenous regulation of splicing of the transcript of a yeast ribosomal protein gene. Proc. Natl. Acad. Sci. USA 83:5854–5857.
- D'Orval, B. C., Y. D. Carata, P. Sirara-Pugnet, M. Gailego, E. Brody, and J. Marie. 1991. RNA secondary structure repression of a muscle-specific exon in Hela cell nuclear extracts. Science 252:1823-1828.
- Goux-Pelletan, M., D. Libri, Y. d'Aubenton-Carafa, M. Fiszman, E. Brody, and J. Marie. 1990. *In vitro* splicing of mutually exclusive exons from the chicken β-tropomyosin gene: role of the branch point location and very long pyrimidine tract. EMBO J. 9:241– 249.
- Green, M. R. 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. Annu. Rev. Cell. Biol. 7:559–599.
- Helfman, D. M., and W. M. Ricci. 1989. Branch point selection in alternative splicing of tropomyosin pre-mRNAs. Nucleic Acids Res. 17:5633-5650.

 Hoffman, B. E., and P. J. Grabowski. 1992. Control of alternative splicing by the differential binding of a small nuclear ribonucleoprotein particle. Genes Dev. 6:2554–2568.

2859

- Jongeneel, C. V., R. Sahli, G. K. McMaster, and B. Hirt. 1986. A precise map of splice junctions in the mRNAs of minute virus of mice, an autonomous parvovirus. J. Virol. 59:564–573.
- Katz, R. A., and A. M. Skalka. 1990. Control of retroviral RNA splicing through maintenance of suboptimal processing signals. Mol. Cell. Biol. 10:696–704.
- Labieniec-Pintel, L., and D. Pintel. 1986. The minute virus of mice P₃₉ transcription unit can encode both capsid proteins. J. Virol. 57:1163–1167.
- 19. Libri, D., A. Piseri, and M. Y. Fiszman. 1991. Tissue-specific splicing *in vivo* of the β -tropomyosin gene—dependence on an RNA secondary structure. Science **252**:1842–1845.
- Maniatis, T. 1991. Mechanisms of alternative pre-mRNA splicing. Science 251:33–34.
- Mattox, W., L. Ryner, and B. S. Baker. 1992. Autoregulation and multifunctionality among trans-acting factors that regulate alternative pre-mRNA processing. J. Biol. Chem. 267:19023–19026.
- McKeown, M. 1992. Alternative mRNA splicing. Annu. Rev. Cell. Biol. 8:133–155.
- Morgan, W. R., and D. C. Ward. 1986. Three splicing patterns are used to excise the small intron common to all minute virus of mice RNAs. J. Virol. 60:1170–1174.
- Nadal-Ginard, B. 1990. Muscle cell differentiation and alternative splicing. Curr. Opin. Cell Biol. 2:1058–1064.
- Naeger, L. K., R. V. Schoborg, Q. Zhao, G. E. Tullis, and D. Pintel. 1992. Nonsense mutations inhibit splicing of MVM RNA in *cis* when they interrupt the reading frame of either exon of the final spliced product. Genes Dev. 6:1107–1119.
- Nasim, F. H., P. A. Spears, H. M. Hoffman, H. Kuo, and P. J. Grabowski. 1990. A sequential splicing mechanism promotes selection of an optional exon by repositioning a downstream 5' splice site in preprotachyrinin pre-mRNA. Genes Dev. 4:1172– 1184.
- Padgett, R. A., P. J. Grabowski, M. M. Konarska, S. R. Seiler, and P. A. Sharp. 1986. Splicing of messenger RNA precursors. Annu. Rev. Biochem. 55:1119–1150.
- Pintel, D., D. Dadachanji, C. R. Astell, and D. C. Ward. 1983. The genome of minute virus of mice, an autonomous parvovirus, encodes two overlapping transcription units. Nucleic Acids Res. 11:1019–1038.
- 29. Rio, D. C. 1992. RNA processing. Curr. Opin. Cell Biol. 4:444-452.
- 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.
- Sahli, R., G. K. McMaster, and B. Hirt. 1985. DNA sequence comparison between two tissue-specific variants of the autonomous parvovirus, minute virus of mice. Nucleic Acids Res. 13: 3617-3633.
- Schoborg, R. V., and D. Pintel. 1991. Accumulation of MVM gene products is differentially regulated by transcription initiation, RNA processing, and protein stability. Virology 181:22–34.
- 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.
- 34. Stoltzfus, C. M., and S. J. Fogarty. 1989. Multiple regions in the Rous sarcoma virus src gene intron act in cis to affect the accumulation of unspliced RNA. J. Virol. 63:1669–1676.
- Talerico, M., and S. M. Berget. 1990. Effect of 5' splice site mutations on splicing of the preceding intron. Mol. Cell. Biol. 10:6299-6305.
- 36. Zhao, Q., and D. Pintel. Unpublished data.
- 37. Zhuang, Y., H. Leung, and A. M. Weiner. 1987. The natural 5' splice site of simian virus 40 large T antigen can be improved by increasing the base complementarity to U₁ RNA. Mol. Cell. Biol. 7:3018–3020.