MOLECULAR AND CELLULAR BIOLOGY, Aug. 1995, p. 4597–4605 Vol. 15, No. 8
0270-7306/95/$04.00+0


# Identification of Positive and Negative Splicing Regulatory Elements within the Terminal *tat-rev* Exon of Human Immunodeficiency Virus Type 1

ALFREDO STAFFA[1] AND ALAN COCHRANE[2]*

*Department of Microbiology and Immunology, McGill University, Montreal, Quebec, Canada H3A 2B4,[1] and Department of Medical Genetics, University of Toronto, Toronto, Ontario, Canada M5S 1A8[2]*

Received 28 November 1994/Returned for modification 24 January 1995/Accepted 20 April 1995

**The requirement of human immunodeficiency virus type 1 to generate numerous proteins from a single primary transcript is met largely by the use of suboptimal splicing to generate over 30 mRNAs. To ensure that appropriate quantities of each protein are produced, there must be a signal(s) that controls the efficiency with which any particular splice site in the RNA is used. To identify this control element(s) and to understand how it operates to generate the splicing pattern observed, we have initially focused on the control of splicing of the *tat-rev* intron, which spans the majority of the *env* open reading frame. Previous analysis indicated that a suboptimal branchpoint and polypyrimidine tract in this intron contribute to its suboptimal splicing (A. Staffa and A. Cochrane, J. Virol. 68:3071–3079, 1994). In this report, we identify two additional elements within the 3 -terminal exon, an exon-splicing enhancer (ESE) and an exon splicing silencer (ESS), that modulate the overall efficiency with which the 3 *tat-rev* splice site is utilized. Both elements are capable of functioning independently of one another. Furthermore, while both the ESE and ESS can function in a heterologous context, the function of the ESS is extremely sensitive to the sequence context into which it is placed. In conclusion, it would appear that the presence of a suboptimal branchpoint and a polypyrimidine tract as well as the ESE and ESS operate together to yield the balanced splicing of the *tat-rev* intron observed in vivo.**

The regulation of human immunodeficiency virus type 1 (HIV-1) replication involves multiple events, including transcription control (31), differential RNA splicing (22), RNA transport and translation (3, 8, 9, 12, 17, 18), and the release of virus from the cell (30). Generation of the full complement of viral regulatory and structural proteins is entirely dependent on the maintenance of the correct balance of unspliced, singly spliced, and multiply spliced viral mRNAs. The mechanism(s) by which the splicing of the HIV-1 primary transcript is controlled is only beginning to be elucidated (1, 23). Recent work on splicing control in Rous sarcoma virus (RSV) has revealed the existence of several regulatory elements: (i) a negative regulator of splicing, present within *gag*, which acts to reduce splicing efficiency of even heterologous introns (19, 20); (ii) the suboptimal nature of the branchpoint (BPT) sequences (11, 14); and (iii) the action of exon sequences to augment splicing efficiency (28). Recent work by our own laboratory has shown that in the case of the terminal *tat-rev* intron of HIV-1, which encompasses much of the *env* gene, an analog of the negative regulator of splicing sequence of RSV does not appear to exist within the intron but that the poor splicing efficiency of this intron is attributable to suboptimal signals in both the BPT and the polypyrimidine tract (PPT) of the 3′ splice site (3′ SS). The optimization of either sequence resulted in a dramatic increase in splicing efficiency (23). Another aspect of splicing regulation was revealed upon analysis of the region comprising *vif*, *vpr*, and the first coding exon of *tat*. Amendt et al. (1) detected the presence of a negative regulatory element within the first coding exon of *tat* whose deletion dramatically altered splice site utilization in vitro and in vivo. Therefore, it appears that multiple elements act to regulate the complex pattern of RNA splicing observed for HIV-1.

While optimization of the 3′ SS signals of the *tat-rev* intron dramatically increased splicing efficiency, other mutations suggested a role of exon sequences in maintenance of balanced splicing in the context of the suboptimal signals (23). The deletion of exon sequences downstream of the *tat-rev* intron 3′ SS resulted in a complete loss of splicing, while in the presence of the exon sequences, splicing efficiencies between 10 and 50% were observed (23). This observation suggested the presence of a positive regulatory element within the exon that exhibited a phenotype similar to that found in the RSV system (an augmented splicing of the upstream intron) (11, 28). Analogous positive exon elements such as the human fibronectin (FN) gene, the mouse immunoglobulin μ-chain gene, the bovine growth hormone gene, the cardiac troponin T gene, and the calcitonin/calcitonin gene-related peptide gene have recently been characterized in other systems (4, 16, 26, 28, 29, 32, 33). In all cases, the element identified is a polypurine sequence that acts to dramatically enhance usage of suboptimal splice sites located in an upstream intron and has been designated an exon splicing enhancer (ESE). Analysis of the sequence of the HIV terminal exon adjacent to the 3′ SS of the *tat-rev* intron revealed the existence of a polypurine sequence which had significant homology to the ESE within the ED1 exon of the human FN gene (16) and to the ESE found in the calcitonin/calcitonin gene-related peptide system gene (33). In this report, we outline evidence that is consistent with the polypurine sequence within the terminal exon being an ESE and determine that an adjacent region of the exon acts to attenuate the activity of this ESE.

* Corresponding author. Mailing address: Dept. of Medical Genetics, University of Toronto, 1 Kings College Cir., Toronto, Ontario, Canada M5S 1A8. Phone: (416) 978-2500. Fax: (416) 978-6885.

## MATERIALS AND METHODS

**Plasmids.** All plasmids are derivatives of pSVβ (23). HIV-1 sequences are those of the Hxb2 proviral clone (the nucleotide numbering system used throughout corresponds to that of the sequence under GenBank accession number K03455). pSVCTSB and pSVCBSB have previously been described (23). HIV-1 sequences spanning the purine-rich splicing regulatory elements (nucleotides [nt] 8411 to 8464) were introduced into pSVCBSB as follows. A 152-bp fragment generated by *Hae*III digestion of the CT PCR product (nt 8333 to 8462) (23) was inserted into the *Bam*HI site of pSVCBSB after the recessed ends had been filled in with Klenow DNA polymerase. This regenerated *Bam*HI sites flanking either end of the AG-rich insert. To facilitate subsequent manipulations, the downstream *Bam*HI site was deleted by digestion with *Pst*I and religation of the vector to generate pSVCBSB-AG. pSVCBSB-GA contains the same purine-rich HIV-1 sequences in the antisense orientation with additional polylinker sequences between the *tat-rev* 3′ SS and the insert. pSVCTSBΔpur was derived from pSVCTSB by mutagenesis (23) with the Δpur mutagenic primer, 5′-CGAAGGAATAGTGGAGAGAG-3′. pSVCTSBΔpur2 was generated by PCR from pSVCTSB with the βPCR1 sense primer (23) and the Δpur2 antisense primer, 5′-CCCTGCAGCTTCTTCTTCTATTC-3′ (nt 8418 to 8432). The 119-bp *Sal*I-*Pst*I fragment of the resulting PCR product was inserted into the respective sites of pSVβ. Human FN gene sequences (GenBank accession number X07718) were derived from pAdED (16), a generous gift from Benoit Chabot. Briefly, pAdED was linearized with *Bsu*36I, the recessed ends were filled in with Klenow DNA polymerase, and *Bam*HI linkers were added onto the blunt ends. Following digestion with *Bam*HI, the vector was religated to generate an enhancerless version of pAdED, called pAdEDΔ, and the 0.3-kb *Bam*HI fragment (containing the FN gene ED1 ESE sequences and the downstream exon and flanking intron sequences) was inserted into the *Bam*HI site of pSVCBSB in the sense orientation. FN gene sequences downstream of the ESE were deleted by digestion with *Pst*I, and then the vector was religated to generate pSVCBFNSE. The FN gene 3′ SS was obtained by PCR from pAdEDΔ with the FN gene 3′ SS sense primer, 5′-CGTCGACAAAGAAAATGGTATCTGC-3′ (nt 1146 to 1164), and the antisense polylinker primer, 5′-GAGCTCGCCCGGGGATCC-3′. The 223-bp *Sal*I-*Bam*HI fragment of the resulting PCR product was inserted into the respective sites of pSVβ, generating pSVβFN−. The plasmids pSVβFN+ and pSVβFN-AG were generated by inserting the *Xho*I-*Bam*HI fragment of pSVβFN− into the *Xho*I and *Bam*HI sites of pSVCBFNSE and pSVCBSB-AG, respectively. pSVβFN-$(AG)_2$ was generated by inserting the *Xho*I to blunted *Eco*RI fragment of pSVβFN-AG into the *Xho*I and blunted *Bam*HI site of pSVCBSB-AG. The plasmid SVβFN−ESS was generated by inserting the *Kpn*I-*Bam*HI fragment of SVβFN− into the *Kpn*I and *Bam*HI sites of SVHTΔpur. The plasmid SVβFN−ESE was generated by PCR as follows: DNA was amplified from SVβFN-AG with the FN 3′ SS sense and Δpur2 primer pair, and the *Sal*I-*Pst*I fragment of the resulting PCR product was inserted into the *Sal*I and *Pst*I sites of SVCTSB. SVβFN+ESS was generated by inserting the *Kpn*I-*Eco*RV fragment of SVβFN+ into the *Kpn*I and blunted *Bam*HI sites of SVHTΔpur. RSV (Prague C strain) sequences were obtained from pATV-8K (6), generously provided by Richard Katz and Anne-Marie Skalka. To facilitate manipulation, a portion of the RSV genome spanning the *env* 3′ SS from the *Kpn*I site to the *Xho*I site (nt 5233 to 5495; GenBank accession numbers J02342, J02021, and J02343) was subcloned into pBluescript SK+ (Stratagene). RSV exon sequences were obtained by PCR with the RSVSE sense primer, 5′-CTGGATCCCCTGGGAAGACG-3′ (nt 5321 to 5340), and the T3 antisense primer. The 185-bp *Bam*HI-*Pst*I fragment of the resulting PCR product was inserted into the respective sites of pSVCBSB, generating pSVCBRSVSE. pSVHH, pSVHH*, and pSVHH$^{ppt}$ are derivatives of their pSVHTSB counterparts (23) and were generated by truncation of HIV-1 exon sequences at the *Hae*III site 33 nt 3′ of the *tat-rev* 3′ SS. The exon splicing silencer (ESS) element was obtained by PCR from pSVCTSBΔpur with the preESE sense primer, 5′-CGGATCCCCGAAGGAATA-3′ (nt 8411 to 8422), and the CAT R1 antisense primer, 5′-CGGAATTCCGGATGAGCATT-3′. The *Bam*HI-*Pst*I fragment of the resulting PCR product was subsequently inserted into the *Bam*HI-*Pst*I sites of pSVβ, pSVHH, pSVHH*, and pSVHH$^{ppt}$ to generate pSVβΔpur, pSVHTΔpur, pSVHT*Δpur, and pSVHT$^{ppt}$Δpur, respectively.

**Cell transfection and RNA analysis.** COS-7 cells (ATCC CRL1651) were cotransfected as previously described (7) with 2 μg each of the test plasmid and pCH110 (Pharmacia) for normalization of transfection efficiency. Total RNA was isolated essentially as previously described (5) except that cells were washed once with ice-cold phosphate-buffered saline and then lysed directly on the culture dish by the addition of 1 ml of 4 M guanidine thiocyanate–25 mM sodium citrate (pH 7.0)–0.5% Sarkosyl–0.1 M β-mercaptoethanol. S1 nuclease protection assays were carried out with 5 μg of total RNA as previously described (23). DNA fragments used as probes for each construct (except pSVCBSB) spanned the 3′ SS from the *Sal*I site within the intron to the *Pvu*II site within the chloramphenicol acetyltransferase (CAT) gene open reading frame (23). The probe for pSVCBSB spanned from the *Hin*dIII site upstream of the β-globin gene 5′ SS to the *Pvu*II site in the CAT gene, and the probe used for pSVCTSB is as previously described (23). In general, probe DNA fragments were generated by digestion of the construct with *Xho*I and *Sal*I, religation, and isolation of a *Pvu*II fragment that was then radiolabelled with $^{32}P$ at its 5′ ends with T4 polynucleotide kinase (21). The anti-β-galactosidase probe used for pCH110 was as previously described (23). The protected DNA fragments were visualized and quantitated with a PhosphorImager (Molecular Dynamics). Images were then processed with Adobe Photoshop and Microsoft Powerpoint and printed on a Hewlett-Packard Laserjet 4 equipped with an XL1 print-enhancing upgrade package.

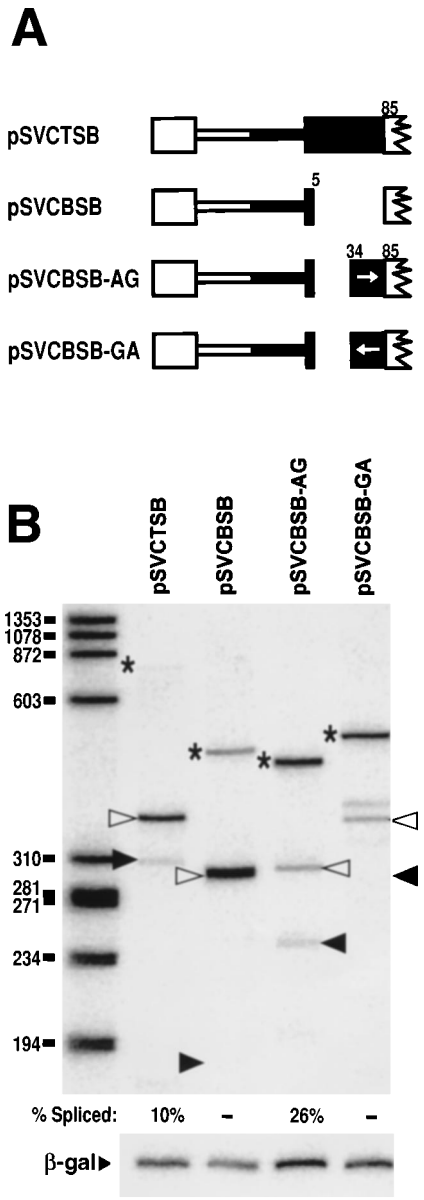



FIG. 1. Deletion mapping of the HIV-1 ESE. (A) Schematic representations of constructs analyzed. Exon sequences are depicted as boxes and intron sequences are depicted as lines between boxes. Open boxes and lines represent sequences derived from pSVβ; filled boxes and lines represent HIV-1 sequences. For clarity, only relevant portions of the second exon are shown. The numbers above the HIV-1 sequences indicate positions of exon sequences relative to the *tat-rev* 3′ SS. pSVCBSB-AG contains an internal deletion of the HIV-1 exon sequences between nt +5 and +34. pSVCBSB-GA contains nt +34 to +85 in the antisense orientation fused to position +5 by polylinker sequences. (B) S1 nuclease protection analysis. Aliquots (5 μg) of total RNA from transfected COS-7 cells were used. Each sample was hybridized with a different probe spanning the *tat-rev* 3′ SS. Positions of the protection products generated by the unspliced and spliced forms of the RNA are indicated by the open and filled arrowheads, respectively. The asterisks mark the positions of the reannealed probes. The percent spliced RNA detected in each lane is shown below the gel. Dashes denote undetectable levels of splicing. Shown at the bottom are results of parallel S1 reactions determining the level of β-galactosidase (β-gal) mRNA. Size standards (in nucleotides) are on the left.

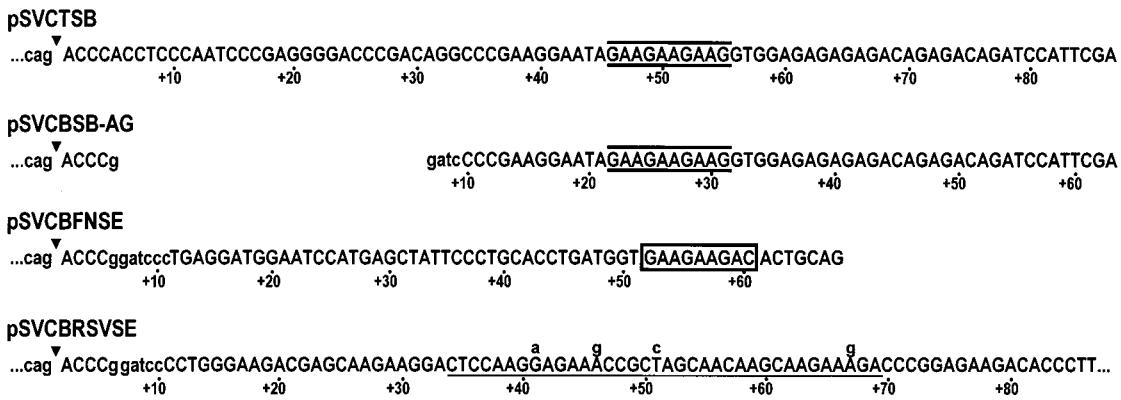



FIG. 2. Nucleotide sequence of purine-rich ESEs. Relevant sequences of pSVCTSB, pSVCBSB-AG, pSVCBFNSE, and pSVCBRSVSE are shown. The terminal *tat-rev* intron sequences are in lowercase letters, and the *tat-rev* 3′ SS intron-exon junctions are indicated by arrowheads. Lowercase letters within the exon are linker sequences. Positions relative to the *tat-rev* 3′ SS are indicated by the numbers below each sequence. Thick lines above and below the pSVCTSB sequence indicate the 10 bp deleted in pSVCTSBΔpur that have some homology with the previously characterized FN gene splicing enhancer sequence 5′-GAAGAAGAA-3′ (boxed in the pSVCBFNSE sequence). In pSVCBSB-AG, the sequence in question is brought 22 nt closer to the *tat-rev* 3′ SS. In pSVCBRSVSE, underlined RSV sequences indicate the homologous region in avian sarcoma leukosis virus that has been demonstrated to stimulate splicing (28). Lowercase letters above RSV sequences indicate positions in which RSV and avian sarcoma leukosis virus sequences differ.

## RESULTS

**Detection of a splicing enhancer with the terminal *tat-rev* exon.** Previous analysis of the basis for the suboptimal splicing of the *tat-rev* intron spanning the *env* region revealed the presence of both a suboptimal BPT and a PPT (23). However, deletions that removed sequences of the terminal *tat-rev* exon resulted in a complete loss of splicing (23), which is suggestive of a positive control element within this region. A comparison of the constructs pSVCTSB and pSVCBSB (which lacks nt 6 to 85 of the exon) (Fig. 1) revealed that splicing efficiency decreased from 10% splicing in pSVCTSB to no detectable splicing in pSVCBSB. To define which region of the exon was responsible for the increase in splice site utilization, a 50-nt sequence encompassing a purine-rich region from nt +35 to +85 3′ of the splice acceptor was fused 3′ of the 3′ SS of the pSVCBSB construct in the sense and antisense orientations (pSVCBSB-AG and pSVCBSB-GA, respectively). Analysis of the RNA splicing patterns generated from these constructs (Fig. 1B) revealed that the addition of these exon sequences in the sense orientation (pSVCBSB-AG) resulted in a significant amount of splicing (26% spliced product), while in the case of the antisense orientation (pSVCBSB-GA), no spliced product was observed, similar to the case of the pSVCBSB parent construct. This observation raised the possibility that this region contained a splicing enhancer, as recently proposed (27).

A sequence comparison of this region with previously characterized splicing enhancers (Fig. 2) revealed a match of eight of nine nucleotides, GAAGAAGAA, to a splicing enhancer recently characterized in the human FN gene (16) and a perfect match to a splicing enhancer element characterized in the calcitonin/calcitonin gene-related peptide system (33). No obvious sequence homology (other than being purine rich) with a splicing enhancer identified within the RSV genome was found (14, 28). To assess whether the splicing enhancer from either the FN gene or RSV could functionally replace the element within the 3′ *tat-rev* exon, regions containing these elements were cloned downstream of the 3′ SS of pSVCBSB. As shown in Fig. 3, the addition of either element resulted in a dramatic increase in the amount of spliced RNA detected (increasing from undetectable amounts to ~70% splicing). This experiment generated two observations: (i) splicing enhancers, either homologous to or different from sequences within the *tat-rev* 3′ exon, could dramatically increase recognition and use of the *tat-rev* intron 3′ SS and (ii) these elements are capable of increasing the splicing efficiency to a level significantly higher than that observed for the putative HIV element (~26% spliced with the authentic HIV exon sequences versus ~70% spliced with either the FN gene or RSV ESE). To confirm that the enhancing activity within the *tat-rev* exon was due to the sequence identified by homology to the FN gene element, the effect on splice site utilization of specifically deleting this sequence was examined. The deletion of the sequence 5′-GAAGAAGAAG-3′ (Fig. 4B, lane pSVCTSBΔpur) resulted in a failure to detect the accumulation of spliced RNA, confirming that the element could be a major determinant responsible for the observed enhancement of splicing. However, the extent of splicing observed in the presence of the authentic HIV exon sequences (~26%) was significantly lower than that obtained in the presence of either the RSV or FN gene ESE. Two possibilities for the diminished activity of the HIV ESE could be that the 1-nt difference between the HIV and FN ESEs could reduce the activity of the enhancer or that a negative element was present in another portion of the terminal *tat-rev* exon. To test the latter possibility, another purine-rich region 3′ of the 5′-GAAGAAGAAG-3′ sequence and an additional 10 nt further downstream were deleted. Analysis of this deletion (Fig. 4B, lane pSVCTSBΔpur2) revealed that removal of this sequence resulted in a dramatic increase in the amount of spliced RNA generated (80% of the RNA detected was of the spliced form). The enhanced splicing observed with the Δpur2 mutation not only indicates that the adjacent sequences are not necessary for the function of the ESE but also suggests that there is a negative element within this region which modulates the effect of the ESE. A sequence comparison demonstrates that the sequence 5′-AGATCC-3′ within this region of the *tat-rev* terminal exon was also present within the negative splicing element (designated the ESS) recently identified within the first coding exon of *tat* (1, 1a).

**Function of HIV-1 splicing enhancer and silencer elements in a heterologous context.** To test whether the sequence identified by the previous analyses was a bona fide ESE and was not functioning in a sequence-dependent context, the capacity of this element to function in a heterologous context was analyzed. In light of the previous demonstration that splicing of

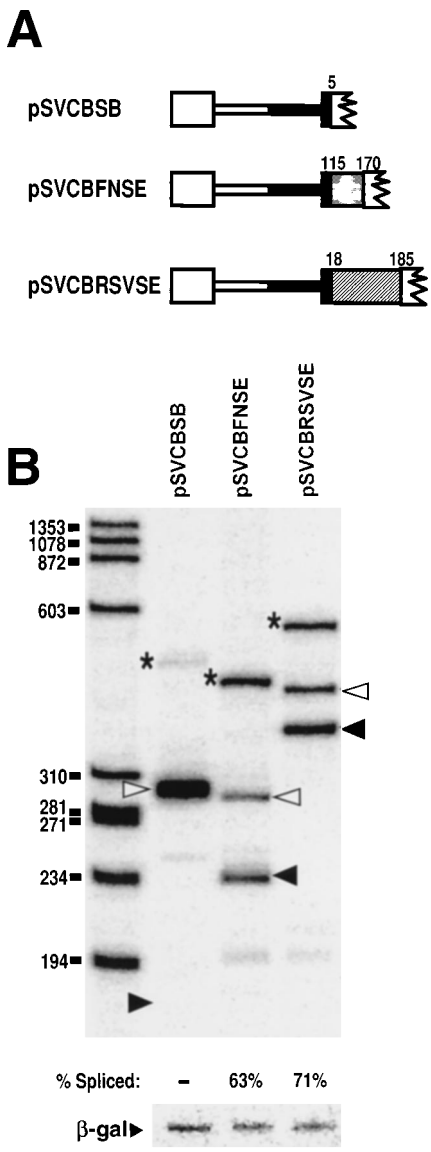
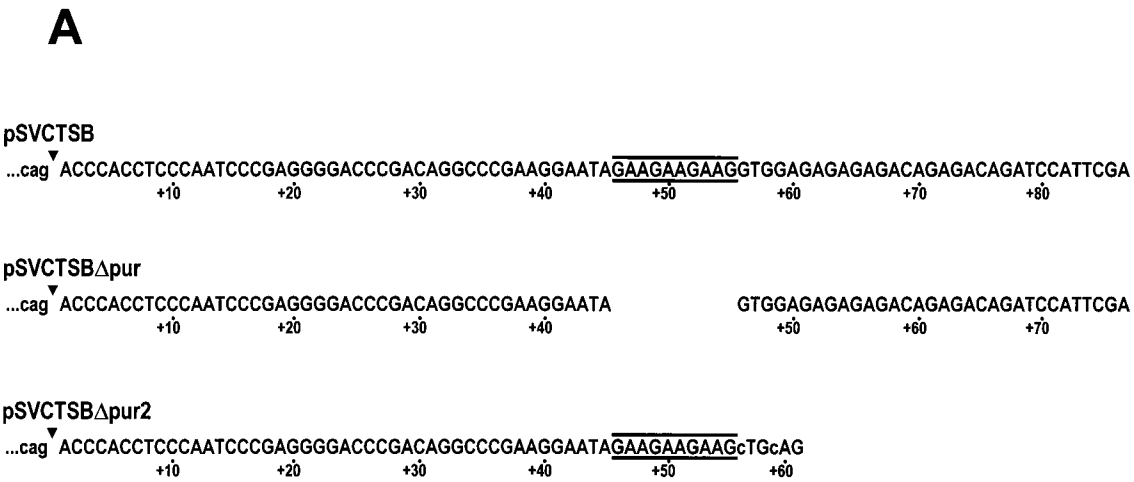



FIG. 3. Testing of heterologous ESEs. (A) Schematic representations of the constructs tested, boxes and lines corresponding to exon and intron sequences as described in the legend to Fig. 1. Test exon sequences shown to stimulate splicing in other systems were inserted downstream of the *tat-rev* 3′ SS in pSVCBSB. pSVCBFNSE contains nt +115 to +170 of the human FN gene ED1 exon (shaded). pSVCBRSVSE contains nt +18 to +185 of the RSV *env* exon 2 (hatched). (B) S1 nuclease protection analysis. Aliquots (5 μg) of total RNA from transfected COS-7 cells were used. Each sample was probed with a different probe spanning the *tat-rev* 3′ SS and test sequences. The positions of the protection products generated by the unspliced and spliced forms of the RNA are indicated by the open and filled arrowheads, respectively. The asterisks mark the positions of the reannealed probes. Shown at the bottom are the results of parallel S1 reactions determining the level of β-galactosidase (β-gal) mRNA. Size standards (in nucleotides) are on the left.

the FN gene intron was affected by the presence of an ESE (16), the capacity of the HIV-1 element to complement the deletion of the FN gene ESE was investigated. As shown in Fig. 5, the addition of either the homologous FN gene ESE or the fragment containing the corresponding HIV sequence resulted in an increase in splicing efficiency of the 3′ SS significantly above that in the absence of these sequences ($P = 0.006$). The addition of only a single HIV element enhanced splicing less than the addition of one homologous FN gene ESE. This could be ascribed to the presence of the putative negative element within this fragment, since a dimer of this sequence resulted in an increase in splicing efficiency to a level comparable to that of the FN gene ESE. As further confirmation of this point, the capacity of the putative HIV ESE and ESS to independently alter splice site utilization in the context of the FN gene 3′ SS was examined. As shown in Fig. 5C, the use of the HIV ESE alone increased splicing to a level comparable to that observed with the FN gene ESE (81.9 and 83.8%, respectively). This observation confirms that the reduced function of the HIV ESE in the context of the nt +30 to +85 fragment is attributable to the presence of the ESS. Analysis of the ESS in the absence of HIV ESE in the context of the FN 3′ SS revealed a small, but significant, decrease in splice site utilization (46.2% spliced in the presence of the ESS and 57.2% spliced in its absence), indicating that the ESS is functional in a heterologous context. The 10% decrease in splicing effected by the ESS is also the same difference observed between SVβFN-AG and SVβFN−ESE, suggesting that the decrease in the former construct is solely attributable to the presence of the ESS. However, placement of the ESS along with the FN gene ESE (SVβFN+ESS) failed to result in a decrease in splice site utilization similar to that seen with SVβFN-AG (Fig. 5B). The failure of the ESS to function in the context of SVβFN+ESS could be the result of either the nature of the FN gene ESE or an increased distance between the FN gene ESE and the HIV ESS in this construct. In support of the latter alternative, increasing the spacing between the HIV ESE and ESS, which generated an organization identical to that of SVβFN+ESS, resulted in a level of splicing (~83%) comparable to that of SVβFN+ESS (data not shown).

Hence, the FN gene and HIV ESEs are interchangeable, suggesting that the identity of eight of nine nucleotides outlined above is functionally significant. Furthermore, the interchangeability of the two ESEs suggests a common mechanism of action. However, attempts to demonstrate a similar reciprocal relationship between the RSV ESE and that of either the FN gene or HIV in the context of an RSV intron were unsuccessful (data not shown). This observation may reflect position or sequence context constraints or indicate that the RSV ESE functions differently from the ESEs of the FN gene and HIV-1.

**Effect of BPT and PPT mutations on splice site function.** To test how the function of the ESS may vary relative to changes in the nature of the 3′ SS, the ability of the ESS to inhibit splicing in the context of the altered HIV 3′ SS was examined. In previous analyses, we had determined the effect of optimizing either the BPT or the PPT on the efficiency of splicing in the context of both the ESE and ESS sequences and found that both of these mutations dramatically enhanced splicing efficiency of this 3′ SS (23). As shown in Fig. 6B, in the absence of the regulatory exon sequences, ~30% splicing was detectable in the case of pSVHH. The observation of splicing of transcripts generated from pSVHH was surprising in light of the failure to observe splicing of RNA generated from both the pSVCBSB and pSVCTΔpur constructs. This observation suggests that another element which acts to promote splicing of pSVHH transcripts exists within the intron, and this possibility is currently being examined. The mutation of either the BPT (pSVHH$^{*}$) or PPT (pSVHH$^{ppt}$) to a more optimal sequence in the absence of both the ESE and ESS elements resulted in a dramatic increase in splicing efficiency (94% spliced in pSVHH$^{*}$ and 83% spliced in pSVHH$^{ppt}$) as previously demonstrated (23). Therefore, these mutations would appear to independently compensate for the loss of the ESE and generate an efficiently utilized 3′ SS.



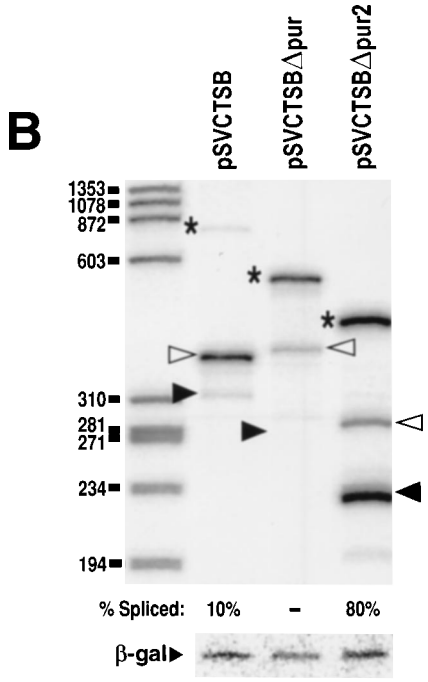




FIG. 4. Deletion mapping of HIV-1 splicing enhancer. (A) Sequences of deletions used to map the splicing enhancer within the *tat-rev* 3′-terminal exon. pSVCTSB was mutagenized such that the sequence 5′-GAAGAAGAAG-3′ at positions +46 to +55 (marked above and below with lines) was deleted to generate pSVCTSBΔpur, while pSVCTSBΔpur2 was generated by the deletion of nt +56 to +85. (B) S1 nuclease protection analysis of constructs containing deletions shown in panel A. The positions of the protection products generated by the unspliced and spliced forms of the RNA are indicated by the open and filled arrowheads, respectively. The asterisks mark the positions of the reannealed probes. The protected products from unspliced pSVCTSBΔpur and pSVCTSBΔpur2 RNA appear slightly larger, since the probe used contains additional complementary intron sequences. Size standards (in nucleotides) are on the left. Shown at the bottom are results of parallel S1 reactions determining the level of β-galactosidase (β-gal) mRNA.

To determine what effect the ESS sequence had in the context of various 3′ SSs, the 10-bp 5′-GAAGAAGAAG-3′ element was deleted in the context of pSVHTSB, pSVHTSB*, and pSVHTSB$^{PPt}$ (23), generating pSVHTΔpur, pSVHT*Δpur, and pSVHT$^{PPt}$Δpur, respectively. In addition, the capacity of the sequence to affect splicing within the context of an efficient 3′ SS was also examined by placement of the ESS 3′ of the efficiently spliced β-globin intron contained within pSVβ (23). Analysis of these constructs (Fig. 6B) revealed that while the negative element had some function in the context of HIV sequences in which either the BPT or PPT had been optimized, no detectable effect on the efficiency of splicing of the β-globin intron was observed. Furthermore, the negative element affected splicing to different extents in the HIV context, resulting in >30-fold, 1.2-fold, and 2.4-fold reductions of splicing within the wild-type, BPT mutation, and PPT mutation contexts, respectively. Together, these data indicate that the negative element is capable of functioning independently of the ESE but in a manner that shows clear dependence upon the context within which it is placed.

## DISCUSSION

The regulation of retroviral RNA splicing is crucial to permit expression of the numerous reading frames whose products participate in the life cycle of the virus. Although true for all retroviruses, it is a particular problem for the lentiviruses, which encode more proteins and have more complex patterns of splicing than viruses of the *Oncoviridae* class (22). In the case of HIV-1, more than 30 mRNAs are generated from the primary transcript in a manner that parallels the level of expression of the various regulatory (Vif, Vpr, Vpu, Tat, Rev, and Nef) and structural (Gag, Gag-Pol, and Env) proteins of the virus (22). Previous analysis of splicing regulation in RSV indicated the presence of a *cis*-acting negative regulatory sequence within the *gag* gene (2, 19, 20, 25) and suboptimal signals within the *env* 3′ SS (11, 13) and *src* 3′ SS (36) that could be overcome by mutations in either the sequences comprising the 3′ SS or the adjacent exon sequences. Analysis of the basis for suboptimal splicing of the HIV primary transcript has, to date, revealed at least two distinct elements: (i) the presence of a *cis*-acting inhibitor of splicing, designated the ESS, within the first coding exon of *tat* (1) and (ii) the presence of suboptimal BPT and PPT sequences (23). The latter finding suggests that the manner in which HIV splicing is controlled is analogous to that of RSV. However, as outlined here, additional elements within the terminal exon of *tat* and *rev* also appear to be involved in controlling the extent of splicing. The first of these additional elements maps to the sequence 5′-GAAGAAGAA-3′ and bears striking sequence and functional homology to a splicing enhancer element recently described in

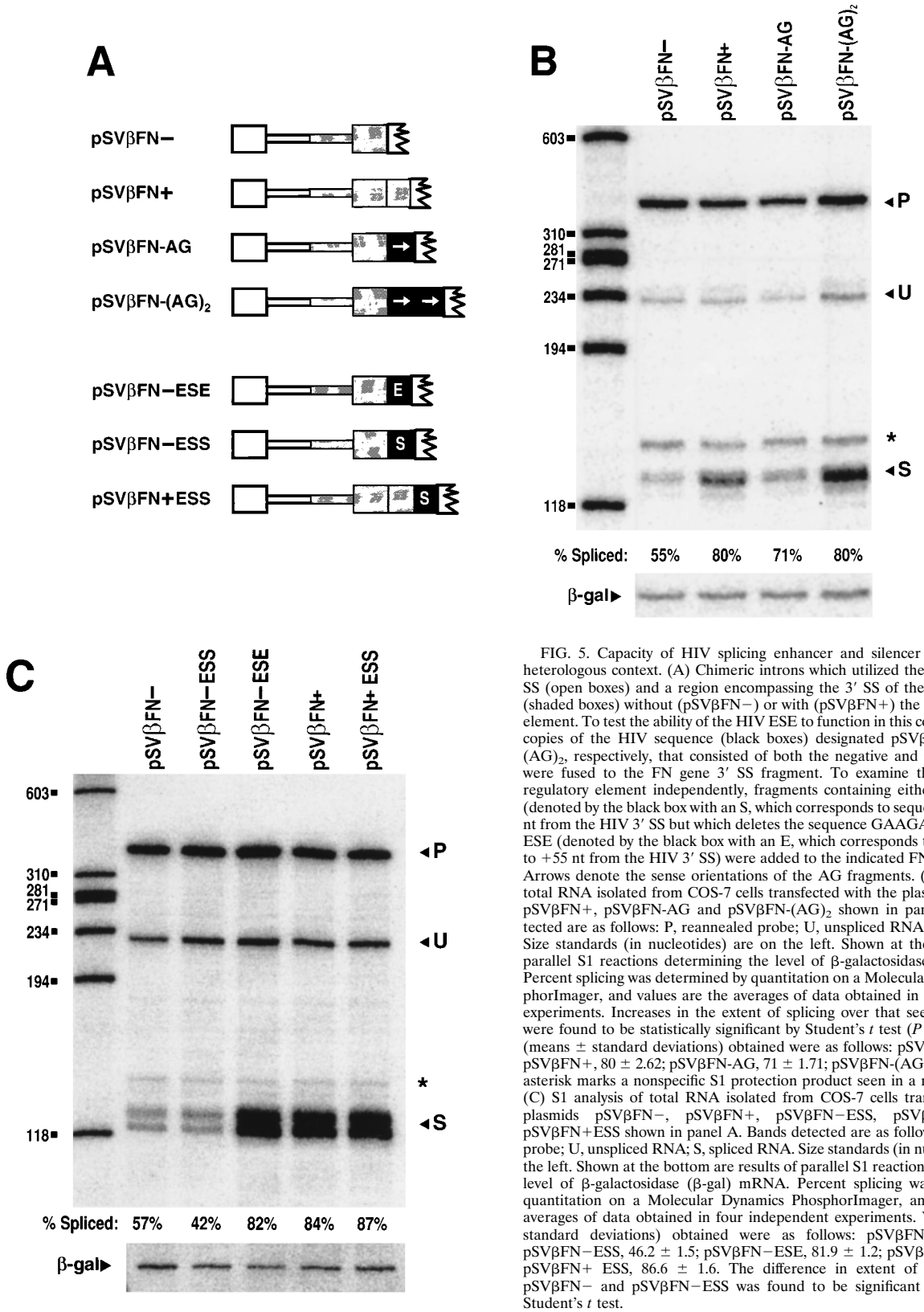



FIG. 5. Capacity of HIV splicing enhancer and silencer to function in a heterologous context. (A) Chimeric introns which utilized the β-globin gene 5′ SS (open boxes) and a region encompassing the 3′ SS of the human FN gene (shaded boxes) without (pSVβFN−) or with (pSVβFN+) the homologous ESE element. To test the ability of the HIV ESE to function in this context, one or two copies of the HIV sequence (black boxes) designated pSVβ-AG and pSVβ-$(AG)_2$, respectively, that consisted of both the negative and positive elements were fused to the FN gene 3′ SS fragment. To examine the effect of each regulatory element independently, fragments containing either the ESS alone (denoted by the black box with an S, which corresponds to sequences +34 to +85 nt from the HIV 3′ SS but which deletes the sequence GAAGAAGAAG) or the ESE (denoted by the black box with an E, which corresponds to sequences +34 to +55 nt from the HIV 3′ SS) were added to the indicated FN gene sequences. Arrows denote the sense orientations of the AG fragments. (B) S1 analysis of total RNA isolated from COS-7 cells transfected with the plasmids pSVβFN−, pSVβFN+, pSVβFN-AG and pSVβFN-$(AG)_2$ shown in panel A. Bands detected are as follows: P, reannealed probe; U, unspliced RNA; S, spliced RNA. Size standards (in nucleotides) are on the left. Shown at the bottom are the parallel S1 reactions determining the level of β-galactosidase (β-gal) mRNA. Percent splicing was determined by quantitation on a Molecular Dynamics PhosphorImager, and values are the averages of data obtained in four independent experiments. Increases in the extent of splicing over that seen for pSVβFN− were found to be statistically significant by Student's *t* test ($P = 0.006$). Values (means ± standard deviations) obtained were as follows: pSVβFN−, 55 ± 4.4; pSVβFN+, 80 ± 2.62; pSVβFN-AG, 71 ± 1.71; pSVβFN-$(AG)_2$, 80 ± 2.48. The asterisk marks a nonspecific S1 protection product seen in a number of assays. (C) S1 analysis of total RNA isolated from COS-7 cells transfected with the plasmids pSVβFN−, pSVβFN+, pSVβFN−ESS, pSVβFN−ESE, and pSVβFN+ESS shown in panel A. Bands detected are as follows: P, reannealed probe; U, unspliced RNA; S, spliced RNA. Size standards (in nucleotides) are on the left. Shown at the bottom are results of parallel S1 reactions determining the level of β-galactosidase (β-gal) mRNA. Percent splicing was determined by quantitation on a Molecular Dynamics PhosphorImager, and values are the averages of data obtained in four independent experiments. Values (means ± standard deviations) obtained were as follows: pSVβFN−, 57.3 ± 1.7; pSVβFN−ESS, 46.2 ± 1.5; pSVβFN−ESE, 81.9 ± 1.2; pSVβFN+, 83.8 ± 2.2; pSVβFN+ ESS, 86.6 ± 1.6. The difference in extent of splicing between pSVβFN− and pSVβFN−ESS was found to be significant at $P < 10^{-5}$ by Student's *t* test.

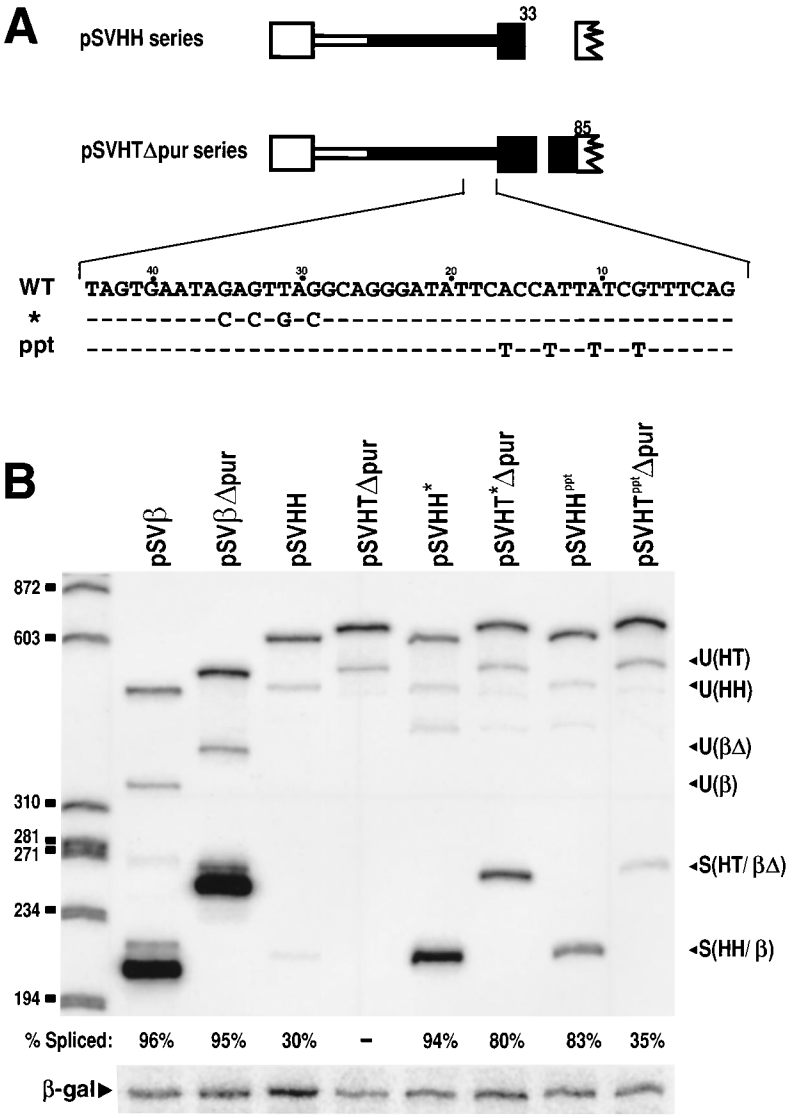



FIG. 6. ESE-independent function of the ESS element. (A) Schematic representations of the constructs used. The BPT (*) and PPT mutations have been described previously (23) and consist of the base changes shown. The pSVHH series introns are nearly identical to the respective pSVHTΔpur introns, the only difference between the two series of constructs being the presence of 3′ exon sequences in the pSVHTΔpur series. In the case of the pSVHH series, the 3′ exon was truncated at the *Hae*III site 33 bp 3′ of the *tat-rev* 3′ SS. In the pSVHTΔpur series, the 3′ exon includes sequences up to +85 nt from the 3′ SS but lacks the 5′-GAAGAAGAAG-3′ enhancer sequence. (B) S1 analysis of RNA following transfection of constructs into COS cells. Ten micrograms of total RNA from the transfected cells was used in each reaction. Indicated are the protected bands expected from the unspliced (U) and spliced (S) RNA from the pSVβ (β), pSVβΔpur (βΔ), pSVHH (HH), and pSVHTΔpur (HT) constructs. Size standards (in nucleotides) are on the left. Shown at the bottom are results of parallel S1 reactions determining the level of β-galactosidase (β-gal) mRNA.

other systems (16, 28, 32, 33). This is reflected by the fact that the element can dramatically enhance splicing efficiency of inefficiently spliced introns containing either the HIV-1 3′ SS or the human FN gene 3′ SS. Furthermore, its function can be compensated for by optimization of either the BPT or PPT of the *tat-rev* intron. While our analysis relies on data generated upon transfection into COS-7 cells, the presence of both the ESE and the ESS has been recently confirmed by using an in vitro splicing system (1a).

How these splicing enhancer sequences function remains undetermined. However, there is significant evidence that they function through interaction with SR proteins (16, 27), which are found to augment splicing in vitro (10, 15, 34, 36). The observation that these ESE sequences are capable of interacting with SR proteins is of particular interest given that one of the SR proteins, SF2/ASF, has been found to be necessary for the splicing of the *tat-rev* intron in vitro (10, 15). A second SR protein, SC35, while capable of enhancing the splicing of other introns, was not able to enhance splicing of the *tat-rev* intron (10), suggesting that the capacity to promote splicing may be due to specific sequences and is not a general property of all introns. It is possible that these SR proteins promote splicing by the recruitment of splicing factors to the splice site, and evidence in support of such a model does exist (24). This recruitment of splicing factors could act to facilitate recognition and interaction with suboptimal splicing signals present within the substrate RNA by increasing the local concentration of splicing factors and, consequently, improve the efficiency of splicing, as recently proposed by Staknis and Reed (24).

In the case of HIV-1, the use of suboptimal BPT and PPT

signals in conjunction with an ESE probably provides a balance in the amount of spliced mRNA generated. However, in the absence of additional control elements, the 5′-GAAGAAGA A-3′ enhancer used in this context results in a high efficiency of splice site usage that would result in a reduced level of the unspliced mRNA. Consequently, a second element, homologous to the ESS recently characterized by Amendt et al. (1), that appears to independently reduce the efficiency of splice site utilization is present in the second exon of *tat*. That the ESS is affecting splicing and not some other aspect of mRNA metabolism is suggested by several lines of evidence: (i) the ESS fails to significantly affect the levels of RNA generated from pSVβ (spliced or unspliced) or the level of unspliced RNA generated from constructs containing the wild-type HIV 3′ SS or mutants of this 3′ SS; (ii) the significant difference between the effect of the ESS in the context of the BPT mutant and the effect observed in both wild-type and PPT mutant constructs supports an effect on splicing and not on the stability of the spliced mRNA, given that all the spliced mRNAs generated from these vectors are identical in sequence; and (iii) work by Amendt and Stoltzfus (1a) has confirmed the inhibitory effect of this element in in vitro splicing extracts. The capacity of this element to reduce splicing in the absence of the ESE in the context of both the HIV and the FN gene 3′ SSs suggests that it does not operate solely by inhibiting ESE function but rather is able to independently act on the 3′ SS itself. The differential effect on the extent of splicing observed for the pSVβΔpur, pSVHTΔpur, pSVHT*Δpur, and pSVHT$^{PPt}$Δpur constructs indicates that it operates by affecting the 3′ SS, given that all these constructs share a 5′ SS. The failure of the ESS to affect splicing of the β-globin intron and its limited effect on the FN gene 3′ SS are of particular interest given the previous demonstration of Amendt et al. (1) that a *tat* exon 2 ESS was capable of functioning in the heterologous context of the RSV *src* 3′ SS, another suboptimal splice site. Therefore, function of this element would appear to be context dependent, i.e., it requires the presence of suboptimal sequences in the 3′ SS, in particular the presence of a suboptimal BPT sequence.

Given that both the ESE and ESS elements are capable of functioning independently, it appears that they would be antagonistic and negate the effect of one another. However, it is possible that recognition of the ESE and recognition of the ESS are mutually exclusive, in which case once the mRNA has interacted with the corresponding factor(s) responsible for the effect of either sequence, the mRNA has entered a pathway to commit or to not commit to spliceosome formation which may be only slowly reversible. In support of this model, increasing the spacing between the ESE and ESS negates the ability of the ESS to modulate ESE function (Fig. 5C, lane SVβFN+ESS; also unpublished data). Under these conditions, a balanced level of splicing of the nascent RNA that could be modulated by altering the abundance of the ESE or ESS binding factors would be attained. Although this organization of HIV-1 splice site regulation (by suboptimal BPT and PPT sequences and ESE and ESS sequences) appears initially to be very complex, it may reflect the best possible compromise between the various requirements placed on this region of the genome. Clearly, alternative strategies could be employed, such as the use of less efficient enhancer sequences (28), but these must take into account the other evolutionary pressures placed on this portion of the genome in order to maximize the overall benefit to the virus. This is illustrated in the case of the ESE sequence of HIV, which also codes for the arginine motif of *rev* that is essential for its *rev*-responsive element binding activity. Although six codons could be used for arginine, only one (AGA) is used exclusively in this region, and this coding bias is tightly conserved among the HIV strains analyzed. Consequently, this region of the exon performs three functions: coding for portions of *env* and *rev* and maintaining balanced splicing of the *tat-rev* intron. Analysis of the host factors mediating the effects of the ESE and ESS sequences may provide insights into the design of strategies aimed at perturbing HIV splicing in order to impede virus replication.

## ACKNOWLEDGMENTS

A. Staffa is supported by an MRC studentship award, and A. Cochrane was the recipient of a FRSQ chercheur boursier and is currently an MRC scholar. This research was supported by grants from the Medical Research Council of Canada and National Health and Welfare.

We also thank B. Amendt and C. M. Stoltzfus for sharing unpublished data and N. Acheson for his helpful suggestions during the writing of the manuscript.

## REFERENCES

1. **Amendt, B. A., D. Hesslein, L.-J. Chang, and C. M. Stoltzfus.** 1994. Presence of negative and positive *cis*-acting RNA splicing elements within and flanking the first *tat* coding exon of human immunodeficiency virus type 1. Mol. Cell. Biol. **14:**3960–3970.
1a. **Amendt, B. A., Z.-H. Si, and C. M. Stoltzfus.** Presence of exon splicing silencers with human immunodeficiency virus type 1 *tat* exon 2 and *tat-rev* exon 3: evidence for inhibition mediated by cellular factors. Mol. Cell. Biol. **15:**4606–4615.
2. **Arrigo, S., and K. Beemon.** 1988. Regulation of Rous sarcoma virus RNA splicing and stability. Mol. Cell. Biol. **8:**4858–4867.
3. **Arrigo, S., and I. Chen.** 1991. Rev is necessary for translation but not cytoplasmic accumulation of HIV-1 *vif*, *vpr* and *env/vpu* 2 RNAs. Genes Dev. **5:**808–819.
4. **Caputi, M., G. Casari, S. Guenzi, R. Tagliabue, A. Sidoli, C. A. Melo, and F. E. Baralle.** 1994. A novel bipartite splicing enhancer modulates the differential processing of the human fibronectin EDA exon. Nucleic Acids Res. **22:**1018–1022.
5. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162:**156–159.
6. **Cobrinik, D., R. Katz, R. Terry, A. M. Skalka, and J. Leis.** 1987. Avian sarcoma and leukosis virus *pol*-endonuclease recognition of the tandem long terminal repeat junction: minimum site required for cleavage is also required for viral growth. J. Virol. **61:**1999–2008.
7. **Cullen, B. R.** 1988. Use of eukaryotic expression technology in the functional analysis of cloned genes. Methods Enzymol. **152:**684–704.
8. **D'Agostino, D. M., B. K. Felber, J. E. Harrison, and G. N. Pavlakis.** 1992. The Rev protein of human immunodeficiency virus type 1 promotes polysomal association and translation of *gag/pol* and *vpu/env* mRNAs. Mol. Cell. Biol. **12:**1375–1386.
9. **Emerman, M., R. Vazeux, and K. Peden.** 1989. The *rev* gene product of the human immunodeficiency virus affects envelope-specific RNA localization. Cell **57:**1155–1165.
10. **Fu, X.** 1993. Specific commitment of different pre-mRNAs to splicing by single SR proteins. Nature (London) **365:**82–85.
11. **Fu, X.-D., R. A. Katz, A. M. Skalka, and T. Maniatis.** 1991. The role of branchpoint and 3′-exon sequences in the control of balanced splicing of avian retrovirus RNA. Genes Dev. **5:**211–220.
12. **Hammarskjöld, M.-L., J. Heimer, B. Hammarskjöld, I. Sangwan, L. Albert, and D. Rekosh.** 1989. Regulation of human immunodeficiency virus *env* expression by the *rev* gene product. J. Virol. **63:**1959–1966.
13. **Katz, R. A., M. Kotler, and A. M. Skalka.** 1988. *cis*-acting intron mutations that affect the efficiency of avian retroviral RNA splicing: implication for mechanisms of control. J. Virol. **62:**2686–2695.
14. **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.
15. **Krainer, A. R., G. C. Conway, and D. Kozak.** 1990. Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. Genes Dev. **4:**1158–1171.
16. **Lavigeur, A., H. La Branche, A. R. Kornblihtt, and B. Chabot.** 1993. A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. Genes Dev. **7:**2405–2417.
17. **Lawrence, J. B., A. W. Cochrane, C. V. Johnson, A. Perkins, and C. A. Rosen.** 1991. The HIV-1 Rev protein: a model system for coupled RNA transport and translation. New Biol. **3:**1220–1232.
18. **Malim, M. H., J. Hauber, S.-Y. Le, J. V. Maizel, and B. R. Cullen.** 1989. The HIV-1 *rev trans*-activator acts through a structured target sequence to acti-

vate nuclear export of unspliced viral mRNA. Nature (London) **338:**254–257.
19. **McNally, M. T., and K. Beemon.** 1992. Intronic sequences and 3′ splice sites control Rous sarcoma virus RNA splicing. J. Virol. **66:**6–11.
20. **McNally, M. T., R. R. Gontarek, and K. Beemon.** 1991. Characterization of Rous sarcoma virus intronic sequences that negatively regulate splicing. Virology **185:**99–108.
21. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
22. **Schwartz, S., B. K. Felber, D. M. Benko, E.-M. Fenyö, and G. N. Pavlakis.** 1990. Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1. J. Virol. **64:**2519–2529.
23. **Staffa, A., and A. Cochrane.** 1994. The *tat*/*rev* intron of human immunodeficiency virus type 1 is inefficiently spliced because of suboptimal signals in the 3′ splice site. J. Virol. **68:**3071–3079.
24. **Staknis, D., and R. Reed.** 1994. SR proteins promote the first specific recognition of pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. Mol. Cell. Biol. **14:**7670–7682.
25. **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.
26. **Sun, Q., R. K. Hampson, and F. M. Rottman.** 1993. In vitro analysis of bovine growth hormone pre-mRNA alternative splicing. Involvement of exon sequences and trans-acting factors. J. Biol. Chem. **268:**15659–15666.
27. **Sun, Q., A. Mayeda, R. K. Hampson, A. R. Krainer, and F. M. Rottman.** 1993. General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. Genes Dev. **7:**2598–2608.
28. **Tanaka, K., A. Watakabe, and Y. Shimura.** 1994. Polypurine sequences within a downstream exon function as a splicing enhancer. Mol. Cell. Biol. **14:**1347–1354.
29. **van Oers, C. C. M., G. J. Adema, H. Zandberg, T. C. Moen, and P. D. Baas.** 1994. Two different sequence elements within exon 4 are necessary for calcitonin-specific splicing of the human calcitonin/calcitonin gene-related peptide I pre-mRNA. Mol. Cell. Biol. **14:**951–960.
30. **Willey, R. L., F. Maldarelli, M. A. Martin, and K. Strebel.** 1992. Human immunodeficiency virus type 1 Vpu protein induces rapid degradation of CD4. J. Virol. **66:**7193–7200.
31. **Wong-Staal, F.** 1990. Human immunodeficiency viruses and their replication, p. 1529–1543. *In* B. N. Fields and D. M. Knipe (ed.), Fields' virology. Raven Press, New York.
32. **Xu, R., J. Teng, and T. A. Cooper.** 1993. The cardiac troponin T alternative exon contains a novel purine-rich positive splicing element. Mol. Cell. Biol. **13:**3660–3674.
33. **Yeakley, J. M., F. Hedjran, J.-P. Morfin, N. Merillat, M. G. Rosenfeld, and R. B. Emeson.** 1993. Control of calcitonin/calcitonin gene-related peptide pre-mRNA processing by constitutive intron and exon elements. Mol. Cell. Biol. **13:**5999–6011.
34. **Zahler, A. M., W. S. Lane, J. A. Stolk, and M. B. Roth.** 1992. SR proteins: a conserved family of pre-mRNA splicing factors. Genes Dev. **6:**837–847.
35. **Zahler, A. M., K. M. Neugebauer, W. S. Lane, and M. B. Roth.** 1993. Distinct functions of SR proteins in alternative pre-mRNA splicing. Science **260:**219–222.
36. **Zhang, L., and C. M. Stoltzfus.** 1995. A suboptimal *src* 3′ splice site is necessary for efficient replication of Rous sarcoma virus. Virology **68:**2849–2859.