Splicing efficiency of human immunodeficiency virus type 1 tat RNA is determined by both a suboptimal 3' splice site and a 10 nucleotide exon splicing silencer element located within tat exon 2

Zhi-hai Si, Brad A. Amendt and C. Martin Stoltzfus*

Department of Microbiology, University of Iowa, Iowa City, IA 52242, USA

Received September 30, 1996; Revised and Accepted December 23, 1996

ABSTRACT

We have previously demonstrated that an exon splicing silencer (ESS) is present within human immunodeficiency virus type 1 (HIV-1) tat exon 2. This 20 nucleotide (nt) RNA element acts selectively to inhibit splicing at the upstream 3' splice site (3'ss #3) flanking this exon. In this report, we have used in vitro splicing of mutated RNA substrates to determine the sequences necessary and sufficient for the activity of the ESS. The activity of the ESS within tat exon 2 maps to a 10 nt core sequence CUAGACUAGA. This core sequence was sufficient to inhibit splicing when inserted downstream from the 3' ss of the heterologous Rous sarcoma virus src gene. Mutagenesis of the interspersed purines in the polypyrimidine tract of the tat exon2 3'ss to pyrimidines resulted in a significant increase in splicing efficiency indicating that 3'ss3 #3 is suboptimal. The ESS acts to inhibit splicing at the optimized 3' splice sites of both the HIV-1 tat and RSV src constructs but with a reduced efficiency compared to its effect on suboptimal 3' splice sites. The results indicate that both the ESS and a suboptimal 3' splice site act together to control splicing at the 3' splice site flanking tat exon 2.

INTRODUCTION

Human immunodeficiency virus type 1 (HIV-1) is a complex retrovirus (1). Its genome contains five 5' splice sites and >10 3' splice sites. During virus replication, the HIV-1 9.2 kb primary transcript is spliced to >40 different singly and multiply spliced mRNA species (2-6). The singly spliced mRNAs encode Env, Vif, Vpr and Vpu whereas the multiply spliced mRNAs encode Tat, Rev and Nef (5,7,8). The unspliced mRNA is used as message for Gag and Pol and is packaged into virions. Successful Tat, Rev and Nef (5,7,8). The unspliced mRNA is used as mRNA species (2-6). The singly spliced mRNAs encode Env, transcript is spliced to >40 different singly and multiply spliced splicing and transport of mRNA species (9).

The splicing of HIV-1 pre-mRNA is dependent on the host splicing machinery. In metazoan cells, this process involves the cleavage and joining of the 5' splice site and the 3' splice site (3'ss) by the spliceosome which is assembled through a series of protein–protein and protein–RNA interactions (10,11). Retroviral replication requires that a single full-length primary RNA transcript be spliced inefficiently (12,13). A number of cis-acting elements within the HIV-1 pre-mRNA have been shown to regulate splicing. These include the polypyrimidine tract and branch point of the 3'ss (14-16), exon splicing enhancer and silencers, and downstream 5' splice sites (16-18). During the early stage of HIV-1 replication, doubly spliced tat, rev and nef mRNAs are constitutively transported from the nucleus into the cytoplasm and translated independently of the expression of any viral proteins. Tat and Rev are both essential proteins for viral replication. Tat is an trans-activator which binds to the cis-acting element TAR and enhances the transcription of viral RNA (19-21). Unspliced and singly spliced HIV-1 RNAs are retained in the nucleus in the absence of the Rev protein. After the accumulation of sufficient Rev protein, HIV-1 replication switches to the late stage in which the Rev protein binds to the Rev response element (RRE) located in the env gene. This facilitates the nuclear to cytoplasmic transport and the expression of singly and unsliced RNAs (22-25).

We and others have previously demonstrated that tat/rev exon 3 splicing is regulated by a suboptimal 3' splice site as well as by an exon splicing enhancer (ESE) and splicing silencer (ESS) within tat/rev exon 3 (18,26). An ESS is also present within tat exon 2. This element acts selectively to inhibit splicing at the upstream 3' splice site flanking this tat exon. The ESS has been defined by linker-scanner mutagenesis to a 20 nucleotide (nt) RNA sequence spanning the C-terminal region of vpr and the N-terminal tat coding region. It has also been demonstrated that the ESS sequence in tat exon 2 functions as a splicing silencer when placed downstream of the heterologous Rous sarcoma virus (RSV) src 3' splice site. The addition of competitor RNA containing the ESS to a splicing reaction with HIV-1 RNA substrates containing the ESS causes a specific relief of inhibition at the upstream 3' splice site (18). This led us to propose that this element represents a binding site for a negative-acting cellular factor present in HeLa cell nuclear extract that inhibits splicing at an early step (18). The studies in this report were carried out to further define this putative cellular factor binding site.

*To whom correspondence should be addressed. Tel: +1 319 335 7793; Fax: +1 319 335 9006; Email: marty-stoltzfus@uiowa.edu
**MATERIALS AND METHODS**

**Plasmid constructions**

All mutations were made by PCR-mediated site-directed mutagenesis (27). The regions corresponding to the PCR products were sequenced to confirm that no misincorporation mutations had occurred. pHS1-X has been described previously (17) and was used as the parent plasmid for all mutant HIV-1 minigene constructs. These include pESS4142, pESS4344, pESS4546, pESS4748, pESS4950, pESS5152, pESS53554, pESS5556, pESS5758, pESS5960, pESS4748+ESS5152, pTar3'C, pΔESS and pΔ3'C+ΔESS. pΔESS was previously named pHS-ΔSRE (17). They were all linearized with *HindIII* at nt 6026 of pNL4-3 (GenBank accession no. M19921) and used as templates to synthesize RNA substrates by *in vitro* transcription. pRSVWT has been described previously (28) and was used to make all the RSV-derived minigene constructs which include ECS, ECUC, ECDS, ECDC, ESS, ESSC, pSSC, pSSC, pSSC, pSSC+ESSS and pSSC+ESSC. pSSC was previously called pRSAP-1 (28). All plasmids derived from pRSVWT were linearized with *NaeI* at nt 7169 of RSV Prague A and used to synthesize RNA substrates.

**In vitro RNA splicing**

RNA substrates of pHS1-X and its derivatives were synthesized by *in vitro* transcription with T3 RNA polymerase. RNA substrates of pRSV7169 and its derivatives were synthesized by *in vitro* transcription with SP6 RNA polymerase. HeLa cell nuclear extracts were prepared and splicing reactions were performed as previously described (17,29,30). The products of the splicing reactions were separated on 4% 7 M urea–polyarylamide gels. At least three gels for each substrate were scanned and quantitated using an AMBIS image analysis system (AMBIS, Inc., San Diego, CA). The amounts of product were calculated on the basis of the uridine number of the RNA species and expressed as splicing efficiency relative to the wild-type to correct for different splicing efficiencies obtained with different HeLa nuclear extract preparations.

**RESULTS**

The activity of the ESS maps to a 10 nt core sequence

We have previously demonstrated that a 20 nt ESS element (from nt 5841 to 5860), present within the tat exon 2 of the HIV-1 strain NL4-3, acts to selectively inhibit splicing at the upstream 3' splice site (3'ss #3) flanking this tat exon (17,18). To further characterize this ESS, we created 2 nt substitution mutations shown in Figure 1A to determine the importance of the individual bases within the element. The RNA substrates containing the mutations were spliced *in vitro* using HeLa cell nuclear extracts. As can be seen in Figure 1B and D, mutations within the central eight nucleotides (ESS4748, ESS4950, ESS5152 and ESS5354) resulted in significant increases in splicing indicating that these nucleotides were essential for the activity of the ESS. However, the 2–3-fold increases obtained with these 2 nt substitution mutations were lower than the 4-fold increase obtained when the entire ESS region was deleted (17,18; see Fig. 3C, compare WT to ΔESS). Therefore, we determined the effect of a double 2 nt mutation ESS4748+ESS5152 (Fig. 1A). As shown in Figure 1C and D, the level of splicing of this mutant substrate was increased to a level comparable to the ESS deletion mutant (ΔESS). On the other hand, mutant substrates ESS4142, ESS4546, ESS5556, ESS5758 and ESS5960 were spliced with efficiencies that were, within experimental error, the same as wild-type. Based on these results together with our previously published data (18), we concluded that the activity of the ESS was contained within the core sequence CUAGACUAUAGA (nt 5846–5855). Mutant substrate ESS4344 was spliced with a somewhat lower efficiency than wild-type. This indicates that in the HIV-1 context the upstream flanking sequence may modulate the negative activity of the core element.

The ESS core sequence is sufficient to inhibit splicing when inserted downstream from the 3'ss of the heterologous RSV src gene

Having localized the negative activity of the ESS element within the tat exon 2 to a 10 nt core sequence, we next asked whether this core element was sufficient to inhibit splicing when placed in a heterologous context and whether there is an effect of the flanking sequences on the ESS activity of the core element. To this end, a series of RSV src gene substrates were created containing the ESS core sequence and flanking sequences downstream of the src 3' splice site. The corresponding RNA substrates were spliced *in vitro*. As shown in Figure 2A, four pairs of insertion mutations were constructed. One member of each pair contained the sequence; the other contained the complement of this sequence in which the purines were changed to their corresponding complementary pyrimidines and vice versa. All the insertions were placed at the same distance relative to the upstream src 3' splice site as was the ESS in the HIV-1 context. Substrates ECS and ECC contained only the 10 nt core sequence of the HIV-1 ESS. ECUS and ECUC contained the ESS core and its immediate upstream 5 nt sequence whereas ECDS and ECDC contained the ESS core plus its immediate downstream 5 nt sequence. ESSS and ESSC contained the 20 nt long ESS or its complement and served as controls. Splicing of RSV src substrates ECS, ECUS, ECDS and ESSS was only about half as efficient as wild-type or substrates containing the corresponding complementary sequences (ESC, ECUC, ECDC and ESSC) (Fig. 2B and C). These results indicate that the ESS core sequence CUAGACUAUAGA is sufficient to inhibit splicing at the heterologous src 3'ss and that the upstream and downstream flanking sequences do not significantly affect the activity of the ESS.

**Tat exon 2 3'ss contains a suboptimal polypyrimidine tract and the ESS can inhibit splicing at an optimized 3'ss**

The tat exon 2 3' splice site contains a weak polypyrimidine tract and a nonconsensus 3'ss sequence. To determine if the negative activity of the ESS requires a suboptimal upstream 3'ss, we synthesized RNA substrates with a consensus polypyrimidine tract and 3' splice site sequence and tested them in an *in vitro* splicing system (Fig. 3A). The effect of the improved 3'ss was first tested in the context of an intact downstream ESS element. This substrate (Tat3'C) demonstrated a 3-fold increase in splicing efficiency compared to a substrate containing the wild type 3' splice site which was spliced very inefficiently (Fig. 3B and C). Thus, splicing at tat exon 2, in addition to the negative effect of the ESS, is regulated by the presence of a weak 3'ss which is in agreement with a previous study (15). Substrate ΔESS, in which heterologous linker sequence was inserted into the ESS demon-
Figure 1. Identification of a 10 nt core sequence necessary for the activity of the ESS located in tat exon 2. (A) Schematic diagram of HIV-1 splice sites and organization of the HIV-1 minigene constructs. ESS defines the previously identified 20 nt ESS element (18). T3, phage T3 RNA polymerase promoter. (B) Effect of single 2 nt substitution mutations within the ESS sequence. Splicing reactions were analyzed by denaturing PAGE. The substrate for each lane is specified at the top. The identities of bands are illustrated on the left side of the autoradiogram. (C) Comparison of the effect of single 2 nt substitution mutations and double 2 nt substitution mutation on the splicing of tat exon 2. Conditions were as described in (B). (D) Quantitation of the spliced products from HIV-1 substrates shown in (A). The amounts of spliced product from gels shown were determined by AMBIS image analysis system and splicing efficiency was compared relatively to wild-type (WT).
Figure 2. The ESS core sequence acts to inhibit splicing of the heterologous Rous sarcoma virus src gene. (A) Schematic representation of the RSV substrates with different sequences containing the ESS core sequence or the complement of these sequences downstream of the src 3' splice site. 5' ss is the major RSV 5' splice site. SP6, phage SP6 RNA polymerase. (B) Splicing reactions were analyzed by denaturing PAGE. (C) Quantitation of the spliced products from the in vitro-spliced substrates shown in (A).

strated a 4-fold increase in splicing efficiency at tat exon 2 3'ss compared to wild-type substrate (Fig. 3B and C). To determine if the ESS element could act to inhibit splicing at the consensus optimized splice site, we studied substrate Tat3'C+AESS which has both the improved 3'ss mutation and the deleted ESS element. The splicing efficiency of this substrate was increased an additional 2-fold at tat exon2 3'ss compared to the splicing efficiency of substrate Tat 3'C (Fig. 3B and C). This significant increase in splicing at the improved tat 3'ss splice site when the ESS element was deleted indicated that the ESS element still had a negative effect on splicing when placed downstream of a consensus 3'ss. However, the magnitude of the inhibition was less compared to a 4-fold effect observed for the wild-type substrate (compare WT to AESS; Tat3'C to Tat3'C+AESS). This demonstrated that the negative control of the ESS element on the splicing efficiency of the tat exon 2 3'ss is not absolutely dependent on a weak 3'ss but that the regulatory control of the ESS element is more efficient in the context of a weak 3'ss.

The ESS acts to inhibit splicing at the optimized 3'ss of RSV src gene

To further test the effect of the HIV-1 tat exon 2 ESS on a consensus upstream 3' splice site, we inserted both the ESS and its corresponding complementary sequence downstream of an optimized avian retrovirus RSV src 3' splice site. The optimized src 3'ss has a consensus branch point and consensus polypurine tract (Fig. 4A) (31). The distance of the ESS relative to src 3'ss was identical to the distance from the tat 3' splice site in the HIV-1 tat exon 2 context. As previously reported, the polypurine tract of src 3'ss is suboptimal and the improved polypurine tract resulted in 1.7-fold increase of src splicing compared to wild-type substrate (Fig. 4B and C; compare Src3'C and wild type) (28). When the ESS was inserted downstream of the optimized src 3'ss, the splicing efficiency at src 3'ss was reduced to approximately wild-type level (Fig. 4B and C; compare Src3'C and wild type) (28). However, if the ESS complementary sequence was inserted, the splicing efficiency at src 3'ss was increased to levels seen with the optimized substrate containing no insertions (Fig. 4B and C; compare Src3'C+ESSC and Src3'C). These results demonstrated that the ESS was still functional when placed downstream of an optimized heterologous 3'ss. However, the magnitude of the effect was reduced compared to the activity of the ESS in the context of a suboptimal 3' splice site.

DISCUSSION

We have carried out point mutagenesis across the ESS sequence within HIV-1 tat exon 2 and have shown that the activity of the
ESS maps to 10 nt CUAGACUAGA core sequence. This sequence is necessary for the negative ESS activity in the homologous HIV-1 context and also is sufficient to inhibit splicing at an upstream 3’ss when placed downstream of a heterologous RSV src 3’ss. In the context of the HIV-1 genome, the behavior of substrate ESS4344 suggests that the 5’ flanking sequence may have an effect on the activity of the core ESS element (Fig. 1D). In the context of the src exon, however, the flanking sequences upstream and downstream of the ESS core sequence do not appear to significantly affect its activity. Based on these results and other results from our laboratory, we hypothesize that the core sequence may represent the binding site for a putative cellular factor (18).

We and others have previously shown that an ESS element is also present within tatrev exon 3 (18,26). This ESS is juxtaposed to a purine-rich ESE to form a bipartite element regulating splicing at the upstream tatrev exon 3 3’ss. The ESS element within tatrev exon 3 was mapped to a region containing the sequence AGAUCCAUUCGAUAGUGAA (the sequences underlined are homologous to sequences in the ESS within tat exon 2). We have shown in this report that the homologous sequence in the 5’ part of the element (AGAUCC) does not appear to play a significant role in the inhibitory activity of the tat exon 2 ESS. However, the 3’ region of the tat exon 2 element (CUAGACUAGA) contains some homology to the sequence in tatrev exon 3. A potential consensus sequence can be identified as PyUAG and there are two copies of this sequence in the tat exon 2 ESS of the HIV-1 NL4-3 strain. Further point mutagenesis of the ESS within the tatrev exon 3 will be necessary to determine the significance of these sequence homologies. The existence of two elements is consistent with the increase in splicing obtained when we combined the two mutations ESS4748 and ESS5152 (Fig. 1D). Each of these mutations would be expected to disrupt only one of the two consensus sequences. The ESS core sequence is not absolutely conserved in all HIV-1 strains. Interestingly, however, most sequenced strains contain at least one copy of a PyUAG motif at the same site in the tat exon 2. Sequence variations in this region of the HIV-1 genome could potentially affect the activity of the ESS resulting in different levels of tat RNA splicing.

Several other exon sequences have been demonstrated to act negatively on splicing at the upstream 3’ss. A sequence UAGG, present in the alternative exon K-SAM in the fibroblast growth factor receptor-2 gene, has been shown to have a negative effect on splicing at the upstream K-SAM 3’ss site (32,33). It is of interest that the HIV-1 tat exon 2, the HIV-1 tatrev exon 3 and the K-SAM element all contain at least one copy of the nonsense codon UAG. It has been concluded from several previous in vivo studies that the presence of premature termination codon within an exon causes inhibition of splicing (34,35). Furthermore, Aoufouchi et al. have described a cell-free system from a human B cell line in which splicing of mouse immunoglobulin mRNA substrates with premature termination codons is inhibited (36). If the splicing inhibitory activity in the HIV-1 tat exon 2 results from the presence of nonsense codons, it would be expected that changing the UAG codons to other termination codons UAA and UGA should not affect its activity. These experiments are in progress in our laboratory. Similar studies carried out by Del Gatto et al. (33) with the K-SAM exon showed that splicing was no longer inhibited when such changes were made. These results suggested that in the case of the K-SAM element, splicing

Figure 3. ESS is active in the context of both suboptimal and optimal tat exon 2 3’ splice sites. (A) Comparison of wild-type and mutated tat exon 2 3’ss sequences to the consensus 3’ splice site. (B) Analysis of spliced products by denaturing PAGE. Note that both ΔESS and tat3C+ΔESS contain identical deletion of the ESS. (C) Quantitation of in vitro spliced products shown in (A).
ESS is active in the context of both suboptimal and optimal in a heterologous RSV src 3'ss site. (A) Comparison of wild-type and mutated src 3'ss sequences to the consensus 3' splice site. (B) Analysis of spliced products by denaturing PAGE. The bands labeled with the asterisk were present only in the substrates containing the splice site mutations. They do not appear to represent any spliced products or lariat intermediates (28). (C) Quantitation of in vitro spliced products shown in (A).

Retroviruses require weak 3' splice sites to balance the splicing of different genes during their replication. In Rous sarcoma virus, both env and src 3' splice sites are suboptimal and they are essential for the successful replication of the virus (13,28,31,39,40). For HIV-1 all the 3' splice sites are inefficient and it has been proposed that the inefficient 3' ss contributes to the regulation of HIV-1 alternative splicing (15). Our results have confirmed that the tat exon 2 3'ss has a suboptimal polypyrimidine tract since the splicing efficiency at this site is increased by mutation of interspersed purines in the tract to pyrimidines. Thus, splicing at the tat exon 2 3'ss is regulated both by a suboptimal 3'ss and by an ESS located within this exon. Our results have shown that when the ESS is downstream of suboptimal 3' splice sites either in its homologous position or when present in a heterologous exon, the level of inhibition is higher compared to the effect of ESS downstream from a consensus 3' ss. Therefore, the strength of the HIV-1 tat exon 2 3'ss and the ESS both appear to control the overall splicing efficiency at the tat 3'ss insuring that the appropriate amount of tat mRNA is produced during virus infection.

ACKNOWLEDGEMENTS

We thank Drs Stanley Perlman and Richard Roller for critical evaluation of the manuscript. This research was supported by PHS grant number AI36073 from the National Institutes of Health.

REFERENCES