The role of branchpoint and 3'-exon sequences in the control of balanced splicing of avian retrovirus RNA

Xiang-Dong Fu,¹ Richard A. Katz,² Anna Marie Skalka,² and Tom Maniatis¹

¹Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138 USA; ²Fox Chase Cancer Center, Institute for Cancer Research, Philadelphia, Pennsylvania 19111 USA

We previously described an avian sarcoma-leukosis virus (ASLV) insertion mutation that causes a decrease in the ratio of unspliced to spliced RNA *in vivo*, resulting in a replication defect. Pseudorevertant viruses containing *cis*-acting suppressor mutations that restored the normal ratio were isolated. One class of the suppressor mutations consists of single-base changes or small deletions near the 3' splice site, while another consists of deletions in the 3' exon. In this paper we report results from an *in vitro* analysis of wild-type, mutant, and pseudorevertant pre-mRNA splicing. We find that wild-type RNA is spliced inefficiently *in vitro*, and that the insertion mutation and suppressors act directly at the level of splicing. Characterization of splicing intermediates reveals that the insertion mutation and suppressor mutations located within the intron alter the pattern of lariat formation. In contrast, suppressor mutations consisting of 3' exon deletions act at an earlier step in the splicing pathway. Thus, the efficiency of splicing at the *env* 3' splice site can be affected at the level of spliceosome assembly, lariat formation, or cleavage at the 3' splice site and exon ligation.

[Key Words: Retrovirus; splicing; splice site selection]

Received September 18, 1990; revised version accepted December 13, 1990.

Retrovirus RNA transcripts are subject to multiple pathways of RNA processing and transport (for a review, see Stoltzfus 1989). The primary full-length RNA *pol* transcript of ~7–10 kb functions as genomic RNA for progeny virions and mRNA for structural proteins and enzymes. Both of these functions require transport of an intact viral RNA to the cytoplasm. In addition, a fraction of the full-length transcripts is spliced in the nucleus to generate subgenomic mRNAs that encode other viral proteins (Fig. 1A). Thus, a balance between full-length RNA and spliced viral mRNAs is required for replication of retroviruses.

Multiple *cis*-acting sequences have been shown to play a role in splice site selection in cellular pre-mRNAs (for review, see Krainer and Maniatis 1988). These sequences include the 5' and 3' splice sites, the branchpoint sequence (BPS), and sequences within the flanking exons. Interaction of these *cis*-acting elements with small nuclear ribonuclear proteins (snRNPs) and many protein splicing factors leads to the assembly of spliceosomes in which the splicing reaction takes place (see Krainer and Maniatis 1988; Steitz et al. 1988; for recent reviews, see Bindereif and Green 1990).

Retroviral transcripts are presumably processed by the same mechanisms as cellular pre-mRNAs. However, unlike most cellular pre-mRNAs, a portion of viral RNA transcripts escape splicing entirely. Balanced splicing of retroviral RNA is therefore a novel form of alternative

splicing. An understanding of the mechanisms involved in this splicing may reveal both general and novel features of the splice site selection process. Several models for maintaining the balance of spliced and unspliced forms of retrovirus RNA have been proposed (Coffin 1985; Katz et al. 1988; Stoltzfus 1989). They include (1) rapid and selective transport of a portion of the unspliced viral RNA from the nucleus, (2) negative regulatory sequences that inhibit splicing *in cis* (Arrigo and Beemon 1988), (3) regulation of RNA processing or transport by viral proteins *in trans*, and (4) suboptimal *cis*-acting signals involved in splice site selection (Katz and Skalka 1990). These models, however, are not mutually exclusive and may operate in conjunction with one another.

Our previous studies described the unexpected effect of a specific 24-bp insert [denoted in this paper as I(–17); see Fig. 1B] placed within the intron immediately adjacent to the *env* 3' splice site. In the wild-type virus the steady-state ratio between full-length RNA and *env* mRNA is ~2 : 1. The insertion mutation increased, rather than decreased, the level of spliced RNA. The replication defect was therefore due to the low level of full-length transcripts. We took advantage of this mutant virus to isolate replication-competent pseudorevertant viruses. The genetic selection led to the isolation of several suppressor mutations within the insert (class I) or in the neighboring 3' exon (class 2) that decreased the level of spliced RNA. These results were consistent with model

Figure 1. (A) Structure and function of multiple RNA species in ASLV-infected cells. The *env* mRNA is generated by alternative splicing of full-length viral RNA using a common 5' donor and an alternative 3' acceptor. (L) Leader sequence. (B) Nucleotide sequences of the *env* 3' splice site from wild-type (wt), linker insertion mutant [I(-17)], and suppressor mutations in class I and II pseudorevertants. The wild-type sequence is numbered. The linker sequence in [I(-17)] is shown in lowercase letters. The altered sequence in each suppressor mutant is shown as single-base substitutions or deletions (bars). For clarity, the names used here for these constructs specify the changes in their sequences [i.e., I(-17), T/C(-20), A/G(-18), Δ36, Δ51]. Other names for them were used previously (see Katz and Skalka 1990). (C) The minigene constructs for in vitro transcription and splicing. Most of the *env* intron sequence was removed by *SacII* and *KpnI* restriction digestions, followed by blunt-end ligation. The size of the intron is 225 nucleotides without linker or 249 nucleotides with linker. The *PstI* fragment containing the minigene is ligated in the *PstI* site of the Sp64 vector. The numbers in the shaded boxes are nucleotides from the Sp64 polylinker that become part of the *env* minigene exon sequence. The numbers in the open boxes are the length in nucleotides of retroviral exon sequences. The intron is represented by a thin line with the conserved GT dinucleotide at the 5' end and the AG dinucleotide at the 3' end. The plasmids are linearized with *XbaI* for in vitro transcription. The primer used for branchpoint mapping and its distance from the 3' splice junction is indicated.

4, in that the suboptimal *env* 3' splice site could be converted into an "optimal" site. In addition, our previous studies indicated that neither the major viral proteins nor the viral 5' splice site were required for this regulation in vivo (Katz and Skalka 1990).

Here, we report results from an analysis of avian sarcoma virus (ASV) *env* mRNA splicing in vitro. We find that wild-type RNA is inefficiently spliced in vitro, indicating that the ratio of spliced to unspliced RNA in vivo is controlled, at least in part, at the level of RNA splicing. Consistent with this interpretation, we found that the [I(-17)] insertion mutation significantly increases the level of splicing in vitro. This increase is correlated with efficient lariat formation at two adenine residues located at -16 and -18 nucleotides upstream from the AG dinucleotide of the *env* 3' splice site. These results suggest that balanced splicing in the wild-type RNA is due to suboptimal branchpoint formation and that the insert provides a more efficient BPS. Suppressor mutations isolated in vivo that restore incomplete splicing

also reduce splicing efficiency in vitro. These suppressor mutations are located within the BPS (class I) or in the 3' exon (class II). These mutations appear to act at different steps in the splicing pathway, including spliceosome formation, lariat formation, and exon ligation.

Results

In vitro splicing of wild-type, mutant, and revertant minigene *env* pre-mRNAs

We have constructed or isolated a number of retroviral DNA clones that encode mutations affecting the balance of spliced and full-length RNA transcripts. These mutations include a 24-bp linker inserted at position -17 from the *env* 3' splice site [I(-17) mutation], point mutations within the linker [T to C at -20, T/C(-20); A to G at -18, A/G(-18); class I suppressor mutations], and deletions in the *env* 3' exon [36 bp (Δ36) or 51 bp (Δ51); class II suppressor mutation] (Fig. 1B). The mutations

were introduced into an ASLV minigene construct in which most of the ~4.7-kb *env* intron was deleted and the 5' and 3' exons were truncated (Fig. 1C).

We prepared RNA transcripts from the minigenes and analyzed their splicing behavior in HeLa cell nuclear extracts. As shown in Figure 2, the wild-type minigene transcript is inefficiently spliced compared to a human β -globin IVS1 transcript (Fig. 2, cf. lanes 4–6 with lanes 1–3). This inefficiency was not simply due to a size effect or other property of the transcript, since the I(–17) mutation that increases the level of splicing at the *env* 3' splice site in vivo also increases the level of splicing in vitro (Fig. 2, lanes 7–9). As is seen in vivo, a decrease in splicing compared to the I(–17) mutant construct is observed with the four suppressor mutations (Fig. 2, lanes 10–21). Quantitation of the relative levels of pre-mRNA and spliced mRNA at the 2-hr time point showed that the splicing efficiency of I(–17) was at least fivefold higher than wild-type and suppressor mutants (see legend to Fig. 2). Despite the low splicing efficiency, all *env* pre-mRNAs were spliced accurately (where detectable). In the T/C(–20) construct, released exon 1 and lariar-exon 2 intermediates were readily detected; however, very little spliced RNA was detected.

These experiments show that the relative efficiencies of splicing of wild-type, I(–17), and suppressor minigene transcripts in vitro are consistent with the in vivo

pattern of splicing observed with intact ASLV RNA. The fact that the suppressor mutations within the I(–17) linker reduce splicing efficiency argues that the linker did not activate splicing by simply increasing the intron size. The results indicate that (1) wild-type ASLV *env* splicing is inherently inefficient, and (2) the I(–17) mutant and suppressor mutations act directly at the level of splicing.

The linker insertion mutation alters the site of lariar formation near the env 3' splice site

Having established that the in vitro splicing patterns of various substrates paralleled the in vivo patterns, we then investigated the mechanism by which the I(–17) mutation could affect the incomplete splicing pattern of wild-type ASLV RNA. Typically, the 3' splice site regions of pre-mRNAs in higher eukaryotes contain a BPS (a weakly conserved 7-nucleotide consensus sequence) usually located between 18 and 40 nucleotides upstream of the 3' splice junction, a polypyrimidine tract, and a conserved AG dinucleotide. These sequences are required for normal levels of splicing at the 3' splice site (for review, see Krainer and Maniatis 1988). The wild-type *env* 3' splice site contains a stretch of 10 pyrimidine nucleotides, the AG dinucleotide, and a weak match to the consensus BPS at position –33 to –39 (Fig. 4B).

To determine whether the increase in *env* splicing of the I(–17) mutant correlates with a change in the site of branch formation, we mapped the branchpoints utilized in the wild-type and I(–17) *env* pre-mRNAs using a primer extension assay. As shown in Figure 3A, we found that the sequence from –33 to –39 did not function as a BPS. Instead, we detected a very weak branchpoint at –16 nucleotide from the 3' splice junction in the wild-type substrate and two strong branchpoints 16 and 18 nucleotides upstream of the 3' junction in the I(–17) substrate. The primer extension products are a consequence of lariar formation, since the levels of these products decrease significantly when the lariar is removed by a debranching activity (Fig. 3B). The A residue at –16 is derived from wild-type sequences, and the A residue at –18 is from the I(–17) insertion (Fig. 4B). Thus, I(–17) mutation provides a new branchpoint (–18), or both, could contribute to the elevated level of splicing in mutant *env* pre-mRNA.

Splicing of wild-type and, sometimes, the A/G(–18) suppressor *env* pre-mRNAs generated lariar-exon 2 intermediates that migrated much faster than those from I(–17) and T/C(–20) suppressor RNAs (Fig. 2), suggesting that lariar formation occurred farther upstream from the 3' splice site. Consistent with this possibility, we detected a branchpoint in wild-type and in A/G(–18) suppressor RNAs (Fig. 4A) located 109 and 133 nucleotides upstream of the 3' junction in wild-type and A/G(–18) minigenes, respectively. The possible relevance of this observation is addressed below (see Discussion).

Figure 2. Splicing of the retroviral *env* minigene transcripts in HeLa cell nuclear extracts. A time course of splicing was followed for each pre-mRNA, as indicated on the top. The gel migration positions for precursors, intermediates, and spliced products are indicated. The ratios of the intensities of the bands corresponding to precursor and spliced RNA were quantitated by scanning the autoradiographs with a densitometer. The ratios of precursor to spliced products were as follows: (wt) 1/0.02; [I(–17)] 1/0.3; [T/C(–20)] 1/0.02; [A/G(–18)] 1/0.05; [Δ36] 1/0.06; [Δ51] 1/0.05.

Figure 3. (A) Primer extension mapping of branchpoints in wild-type, I(-17), and the T/C(-20) substrates. The RNA samples were incubated for the indicated times in nuclear extracts, and 5' 32 P-labeled primer was annealed to the pre-mRNA and extended by reverse transcriptase. The same end-labeled primer was also used to sequence the mutant DNA with unlabeled trinucleotide phosphates. As a molecular marker, the sequencing ladder is shown on the left. The bands corresponding to the branchpoints are specified by dots, and the sequence and branchpoints are shown at the bottom. (B) Lariat-exon2 intermediates from I(-17) and T/C(-20) were gel-isolated and treated with the RNA debranching activity (Ruskin and Green 1985) from HeLa S100 fraction using the conditions described by Arenas and Hurwitz (1987). After the debranching treatment, the amount of the shorter primer extension product decreased significantly, and an extension product of 355 nucleotides was observed. The latter product corresponds to the size expected for extension to the end of the intron. Molecular weight marker: Sp64-HpaII fragments.

Branchpoint selection and splicing

The results described above show that the level of splicing at the *env* 3' splice site can be correlated with changes in the site of branch formation. This correlation is strengthened by the observation that the two class I suppressor mutations are single-base changes within the BPS near the new -18 branch site. Thus, it is likely that these mutations act by altering branchpoint recognition, as follows. Branch formation occurs at -16 in the wild-type RNA, and very low levels of splicing are observed (Table 1). The I(-17) mutation increases the use of the -16 A and provides a new, equally efficient branch site at the -18 A. The increase in splicing efficiency observed with this mutant could therefore be due to an increased efficiency of branchpoint formation at -16, -18, or both. The T/C(-20) suppressor apparently disrupts the BPS that includes the -18 adenine residue. This change affects the most conserved residue in the mammalian BPS (see Table 1). As a result, it eliminates the branch at -18, increases utilization of the branchpoint at -16, and weakly activates a branch site at -23 (Figs. 3A,B; 4A). Since the T to C change in the T/C(-20) suppressor mutant significantly decreases the level of splicing, branchpoint formation at -16 in the I(-17) RNA must not contribute significantly to the increased efficiency of splicing. The A to G change in A/G(-18) removes the branchpoint adenine nucleotide at -18, decreases utilization of the A at -16, and weakly activates a cryptic branch site at -23 (Fig. 4A). Thus, the branch at -18 in the I(-17) mutant RNA must contribute significantly to the increase in splicing efficiency of this RNA. In the A/G(-18) suppressor RNA, the most prominent branchpoint is the A residue at -23 (Fig. 4A). Both the utilization of this site and formation of spliced product were coordinately decreased in comparison with the

I(-17) mutant (Fig. 4). On the basis of these observations, we conclude that phenotypic reversion of the linker insertion mutant is a consequence of the complete inactivation of a productive branchpoint such as the A at -18 and/or by selective inhibition of the utilization of a productive branchpoint such as the A at -23. Consistent with this interpretation, branching at the adenine residue at -18 precedes that at -16 and -23, indicating that the A at -18 is preferred over the others in forming the branchpoint and in the second step of the splicing reaction (Fig. 4).

The properties of *env* region BPS listed in Table 1 are consistent with previously established principles of branchpoint selection: A sequence is preferred over other potential sites if it is closer to the polypyrimidine tract and is better able to base-pair with U2 snRNP, especially at positions +1 and -2 (if the position of the branched A is 0, the first nucleotide at its 3' side is +1 and the first at 5' is -1).

BPS and spliceosome assembly

Examination of the in vitro splicing products of the T/C(-20) suppressor substrate revealed a relatively high level of lariat-exon 2 and released exon 1 but only a low level of splicing. In contrast, splicing of the A/G(-18) suppressor substrate produced lower levels of released exon 1 and lariat-exon 2 intermediates (Fig. 2), which is consistent with a lower level of lariat formation and final spliced product (Fig. 4). These observations suggest that base changes responsible for the phenotypic reversion may act at different stages of the splicing pathway. To determine whether the differences in lariat formation observed with the wild-type and suppressor RNAs could be correlated with differences in spliceosome assembly, we carried out density gradient sedimentation studies on

Figure 4. (A) Branchpoint mapping by primer extension. The dots indicate bands corresponding to branchpoints; open circles indicate the branchpoints near the 5' splice site; small arrowheads indicate spliced products. (B) Summary of branchpoint mapping. The 3' splice site sequences are shown. The shaded sequences are from the linker. Underlined in wild-type is a potential BPS that has the best match to the branchpoint consensus. The point mutations are indicated by arrows. The branchpoints as determined by primer extension are indicated by dots. Numbers in parentheses are the distances between the upstream branchpoint and the *env* 3' splice site in viral RNAs.

Table 1. Comparison of ASLV *env* branchpoint sequences with mammalian BPS consensus

Wild type:	5'---AUGAGGUGACUAAGAAAGAUGAGGCGAGCCCUCUUUUUGCAG---3'		
3' Splice site:	I(−17): 5'---Cg a g c c c g u a a u u c g c u u a a g c u u a GAGCCCUCUUUUUGCAG---3'		
Suppressors:	c g		
U2 snRNA:	3' - A U G A U - G - 5'		
BPS consensus:	5' - Y N Y U R A Y - 3'		
Sites	Match	Branch activity	
Best BPS at the 3' splice site:	A G G U G A C	4/6	—
wt	A G G C G A G	2/6	+ / −
A(−16) I(−17)	C U U A G A G	4/6	+++
T/C(−20)	C C U A G A G	4/6	++++
A/G(−18)	C U U G G A G	4/6	+
I(−17)	A G C U U A G	3/6	+++
A(−18) T/C(−20)	A G C C U A G	2/6	—
A/G(−18)	A G C U U G G	3/6	—
I(−17)	G C U U A A G	4/6	—
A(−23) T/C(−20)	G C U U A A G	4/6	+ / −
A/G(−18)	G C U U A A G	4/6	+

these minigene *env* pre-mRNAs. As shown in Figure 5, we did not observe a discrete peak at 60S (spliceosome) with the wild-type *env* pre-mRNA. Rather, labeled RNA was present throughout the region between 60S and 30S. This profile may be due to lariat formation at two different sites in this RNA (-109 and -16; see Discussion) that give rise to different lariat intermediates. In any case, the relatively large number of counts in complexes with an S value of 60 or greater suggests that spliceosomes are assembled on the wild-type RNA.

The sedimentation profile observed with the I(-17) mutant RNA (Fig. 5) is very similar to that observed with

efficiently spliced substrates, such as the β -globin pre-mRNA (data not shown). Comparison of the I(-17) profile with those observed with the less efficiently spliced suppressor RNAs reveals a decrease in the amount of complex formation. Spliceosomes are formed on the T/C(-20) suppressor substrate, but the levels of both the 60S and 30S complexes are much lower than those in the mutant. The amount of 60S complex formed with the A/G(-18) suppressor substrate was higher than that observed with the T/C(-20) suppressor but lower than that observed with the I(-17) mutant substrate (Fig. 5). Thus, the two class I suppressor mutations appear to have dif-

Figure 5. Density gradient sedimentation analyses of spliceosome assembly on ASLV *env* pre-mRNA. Each viral pre-mRNA, indicated on the top, was incubated in HeLa nuclear extracts under splicing conditions for 40 min. The reaction mixtures were applied to 10–30% glycerol gradients, and the fractions were collected and counted as described previously [Abmayr et al. 1988]. The positions of 30S complexes and 60S spliceosomes are indicated.

ferent effects on the efficiency of spliceosome accumulation.

Role of exon sequence in splicing and spliceosome assembly

Both class II suppressors ($\Delta 36$, $\Delta 51$) dramatically decrease the level of splicing as compared to the I(-17) mutant *env* pre-mRNA in vitro (Fig. 2), but the pattern of branchpoint selection was not affected by these deletions (Fig. 4). The two A residues at -16 and -18 were used in both the I(-17) and the class II suppressor-derived pre-mRNAs. However, the level of lariat formation observed with deletion 36 was lower than that observed with I(-17), and the level observed with deletion 51 was even lower (Fig. 4A). The decreases observed with these two deletions could be directly correlated with decreases in amounts of spliced products, suggesting that the effect of exon deletion was on a step prior to lariat formation. An analysis of the effects of these suppressor mutations on spliceosome assembly is presented in Figure 5. Both RNAs gave a very low level of 60S complex formation. We conclude that the 3' exon deletions act at a very early step in the splicing pathway, prior to spliceosome assembly.

Discussion

A balance of full-length and spliced RNA is essential for retroviral replication. In the case of the wild-type ASLV, the ratio of unspliced to spliced *env* pre-mRNA is maintained at ~2 : 1. This relatively inefficient use of the *env* 3' splice site appears to be an inherent property of the structural organization of the viral *env* pre-mRNA and its interaction with the cellular splicing machinery. Consistent with this model, virus-encoded proteins are not required for maintaining the normal balance of spliced ASLV RNA (Katz and Skalka 1990). Previous studies have shown that the 5' splice site of the *env* pre-mRNA is not required for balanced *env* RNA splicing (Katz and Skalka 1990). Genetic selection and structural characterization of pseudorevertants from an oversplicing defective virus containing the I(-17) mutation suggested that sequences on both sides of the *env* 3' splice site affect the efficiency of splicing (Katz and Skalka 1990).

The results reported here indicate that the efficiency of splicing at the *env* 3' splice site can be affected at different steps of the splicing reaction:

1. Spliceosome assembly. Deletions in the exon downstream from the *env* 3' splice site severely inhibited spliceosome formation.
2. Lariat formation. The A/G(-18) suppressor mutation removed the critical branchpoint nucleotide.
3. Cleavage at the 3' splice site and exon ligation. The T/C(-20) suppressor mutant accumulates relatively high levels of released exon I and the lariat-exon 2 intermediate, but a relatively low level of splicing was observed.

We conclude that at least one mechanism involved in maintaining the appropriate balance between spliced and unspliced ASLV RNA is via control of the efficiency of splice site selection at the *env* 3' splice site. Inefficient splicing at the *env* 3' splice site in full-length viral RNA could reflect an intrinsic property of the sequences within and surrounding this site or an active negative regulatory mechanism could be involved. A model that is consistent with previous studies (Katz and Skalka 1988, 1990) and the data presented here are that the wild-type *env* 3' splice site is relatively weak. The I(-17) insertion provides a strong BPS that increases the level of splicing at the *env* 3' splice site. Point mutations within the new BPS or deletions in the downstream exon may therefore act by weakening the mutant *env* 3' splice site.

Evidence for an active negative regulatory mechanism is provided by the identification of two negative elements located within the *gag* gene or in the intron of the *env* pre-mRNA (Arrigo and Beemon 1988; Stoltzfus and Fogarty 1989). Although a similar effect was not observed with a large truncation of the *env* pre-mRNA intron including these elements (Katz and Skalka 1990), the negative elements may function only in a certain sequence context. Negative elements may also play a role in incomplete splicing of influenza virus NS1 mRNA (Plotch and Krug 1986; Agris et al. 1989).

The in vitro analyses of wild-type, I(-17), and suppressor-containing pre-mRNAs revealed significant differences in the pattern of branchpoint selection (see Table 1). Lariat formation at some of these branch sites is more efficient than others, but none of them provide a good match to the branchpoint consensus sequence since all of them have a G residue on the 3' site (+1) of the branched A residue. The residue in this position is C or at least a pyrimidine in most cellular pre-mRNAs (Reed and Maniatis 1985), and a change of the C residue to G severely inhibited splicing of human β -globin pre-mRNA (Reed and Maniatis 1988). This C residue has also been shown to base-pair with the G residue in U2 snRNP (Zhuang and Weiner 1989; Wu and Manley 1989). These observations could explain the relatively inefficient splicing of the linker insertion pre-mRNA compared to that of human β -globin pre-mRNA (Fig. 2).

The role of the BPS on spliceosome assembly

Nucleotide substitutions in the BPS that virtually eliminate lariat formation do not block spliceosome assembly (Reed and Maniatis 1988), but they do decrease the amount of spliceosomes detected at early times in the splicing reaction (Nelson and Green 1989). In the present study we found that the point mutations in two class I revertants affect spliceosome assembly in different ways (Fig. 5). In the case of the A/G(-18) suppressor mutation, we observed decreased splicing, lariat formation, and spliceosome assembly, indicating that the mutation within the BPS affects an early step of the splicing reaction, possibly at the level of the interaction between the BPS and U2 snRNP. In contrast, with the T/C(-20) suppressor, we could detect amounts of released exon 1 and

lariat–exon 2 that were almost identical to those observed with the more efficiently splicing I(–17) mutant, but the second step in splicing was blocked (Fig. 2). Since we expect that the first step of the splicing reaction takes place within the assembled spliceosomes, it was surprising that the point mutation in the T/C(–20) revertant had a dramatic effect on the formation of 60S spliceosomes and 30S complexes. These observations indicate that the T/C(–20) pre-mRNA may be able to participate actively in spliceosome assembly, but the complexes may not be stable because a transition from the first to the second steps of the splicing reaction is impaired. As a result, little spliced product was observed.

Productive and nonproductive branchpoint

The use of multiple branchpoints was observed in the analysis of the *in vitro* splicing products of SV40 large T pre-mRNA (Noble et al. 1987). However, the relative contribution of each branchpoint to the overall efficiency of the reaction has not been established. The term “nonproductive branchpoint” was used to describe the effect of a A to G substitution at the branchpoint, which blocked the splicing reaction after the formation of lariat intermediate (Hornig and Weissmann 1986; Frever et al. 1987). Although all branchpoints occur at the A residue in the present study, we observed that the formation of the branchpoint on each of the A residues had different potentials to complete the splicing reaction. The order of such a potential appears to be $A(-18) \geq A(-23) > A(-16)$.

The wild-type and A/G(–18) mutant displayed a rapidly migrating lariat–exon 2 intermediate (Fig. 2), and both contained distant upstream branchpoints at –109 and –133, respectively (Fig. 4B). The same adenine nucleotide is used in both cases; the difference in the distance from the AG dinucleotide results from the I(–17) insertion. The use of this BPS was diminished significantly when the downstream BPS was activated in the I(–17) and the T/C(–20) revertant (Fig. 4A). These results were confirmed by primer extension analysis of branched and debranched RNA substrates (data not shown). Although functional BPSs are usually located within 18–40 nucleotides upstream of the AG, there are cases in which lariat formation occurs farther upstream (Gattoni et al. 1988; Helfman and Ricci 1989; Reed 1989; Smith and Nadal-Ginard 1989; Goux-Pelletan et al. 1990). In these cases, the branchpoint sequence is followed immediately by a pyrimidine stretch. In fact, the upstream branch site in ASLV RNA is followed by a stretch of 8 pyrimidines interrupted by a single adenine. In cases where far upstream branchpoint sequences are used, splicing occurs at the nearest AG dinucleotide downstream (Reed 1989; Smith et al. 1989), although sequence context of the AG can affect the efficiency of splicing (Reed 1989). Indeed, we can detect primer extension products that could correspond to splicing at AG dinucleotides 3' of the far upstream branchpoints of both the wild-type and A/G(–18) RNAs (Fig. 4A). The biological significance of the far upstream branchpoint se-

quence is not known. The A residue of this branchpoint sequence in the minigene transcript is actually located ~4.5 kb upstream of the *env* 3' splice site in intact viral RNA (Fig. 4B). Thus, this site of lariat formation could be a cryptic branchpoint/3' splice site that is activated only in the minigene, or it could be a real branchpoint in intact viral RNA that generates a spliced product, as detected previously (Broome and Gilbert 1985).

Effects of exon sequences on spliceosome assembly

Exon sequences can have profound effects on splice site selection in a *cis*-competition assay (Reed and Maniatis 1986). Base substitution, deletions, or insertions within flanking exons can dramatically affect the use of adjacent 5' or 3' splice sites. Deletions in the exon downstream of the mutant *env* 3' splice site significantly decrease the level of splicing *in vivo* (Katz and Skalka 1990) and *in vitro* (Fig. 2). In addition, these deletions significantly decrease the level of spliceosome assembly *in vitro* (Fig. 5). The mechanisms by which these deletions block spliceosome assembly is not understood. The most likely possibility is that the secondary or tertiary RNA structure in the flanking exon influences the efficiency by which the splicing apparatus recognizes the adjacent splice site.

Summary and implications

We have examined the role of RNA splicing in maintaining the balance between unspliced and spliced retroviral RNA. The potential effects of mutations on RNA transport and stability were circumvented by comparing the splicing of wild-type, mutant, and suppressor RNAs *in vitro*. Our results strongly suggest that balanced RNA production is controlled at the level of splicing, and that inefficient lariat formation is an important aspect of this process. In addition, we find that 3' exon sequences can influence spliceosome assembly.

Human immunodeficiency virus (HIV) RNA is also spliced inefficiently *in vitro* (Krainer et al. 1990), suggesting that similar mechanisms may be involved in balanced RNA production in the two types of retroviruses. However, in the case of HIV the virus-encoded gene product *rev* is also necessary to control the balance between unspliced and spliced RNA at the level of splicing and/or transport (Varmus 1988; Chang and Sharp 1989). The question of why HIV requires inefficient splicing and *rev* for balanced RNA production, while ASLV appears to require only inefficient splicing, remains to be answered.

Materials and methods

Construction of minigenes

All constructs were derived from viral DNA clones as described previously (Katz et al. 1988; Katz and Skalka 1990). These clones contain a large deletion of *env* intron from position 547 to 5000 (Schwartz et al. 1983) and were made by removing the 4.5-kb *SacII*–*KpnI* fragment followed by blunt-end ligation (Fig.

1C). The *Pst*I-*Pst*I fragment (from position 268 to 5569) from each truncated clone was then excised and cloned in the *Pst*I site of the Sp64 vector for in vitro transcription. The names that specify the changes in their DNA sequences were given to these series of plasmids [Fig. 1B; I(-17), T/C(-20), A/G(-18), Δ 36, Δ 51].

In vitro splicing and spliceosome assembly

RNAs were synthesized in vitro from *Xba*I-linearized plasmids. The in vitro splicing reaction and the density gradient sedimentation analyses were as described [Frederewey and Keller 1985; Bindereif and Green 1986; Abmayr et al. 1988].

Primer extension

The synthetic 18-nucleotide DNA oligonucleotide used for primer extension is 5'-TCGTTGGCAGCAAGGGTG-3'. The oligonucleotide anneals to viral RNA from position 5166 to 5183 [Schwartz et al. 1983]. Primer extension was as described [Krainer et al. 1984] with some modifications: 0.5 pmole of 5'-³²P-labeled primer was annealed overnight at 37°C to RNA from a 25- μ l splicing reaction in 20 μ l of 40 mM PIPES (pH 6.4), 0.4 M NaCl, 1 mM EDTA, and 0.02% SDS. The hybrids were precipitated with 2 μ l of 3 M NaOAc (pH 5.3) and 50 μ l of ethanol and extended at 42°C for 2 hr with 5 units of AMV reverse transcriptase in 20 μ l of 50 mM Tris (pH 8.3 at 42°C), 60 mM KCl, 10 mM MgCl₂, 1 mM dNTP, 1 mM DTT, 1 unit of RNasin, and 50 μ g/ml Act.D. The extended DNA molecules were treated with 1.5 μ l of RNase A (2 mg/ml) at 37°C for 15 min, phenol-extracted, precipitated with ethanol, and resuspended in 5 μ l of sequencing loading buffer and separated on a 6% denaturing gel.

Acknowledgments

We thank George Merkel for excellent technical assistance, Tony Yeung for synthesizing oligonucleotides, and R. Reed for critical review of this manuscript. This work was supported by grants from the National Institutes of Health to T.M. (GM42231) and to R.A.K. and A.M.S. (CA06927, CA48703, RR05539), a grant from the Pew Charitable Trust, and an appropriation from the Commonwealth of Pennsylvania to R.A.K. and A.M.S. Partial support for R.A.K. was provided by the W.W. Smith Charitable Trust.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Abmayr, S., R. Reed, and T. Maniatis. 1988. Identification of a functional mammalian spliceosome containing unspliced pre-mRNA. *Proc. Natl. Acad. Sci.* **85**: 7216-7220.

Agris, C.H., M.E. Nemeroff, and R.M. Krug. 1989. A block in mammalian splicing occurring after formation of large complexes containing U1, U2, U4, U5, and U6 small nuclear ribonucleoproteins. *Mol. Cell. Biol.* **9**: 259-267.

Arenas, J. and J. Hurwitz. 1987. Purification of a RNA debranching activity from HeLa cells. *J. Biol. Chem.* **262**: 4274-4279.

Arrigo, S. and K. Beemon. 1988. Regulation of Rous sarcoma virus RNA splicing and stability. *Mol. Cell. Biol.* **8**: 4858-4867.

Bindereif, A. and M. Green. 1986. Ribonucleoprotein complex

formation during pre-mRNA splicing *in vitro*. *Mol. Cell. Biol.* **6**: 2582-2592.

———. 1990. Identification and functional analysis of mammalian splicing factors. *Genet. Eng.* **12**: 201.

Broome, S. and W. Gilbert. 1985. Rous sarcoma virus encodes a transcriptional activator. *Cell* **40**: 537-546.

Chang, D.D. and P.A. Sharp. 1989. Regulation by HIV rev depends upon recognition of splice site. *Cell* **59**: 789-795.

Coffin, J. 1985. Genomic structure. In *Molecular biology of tumor viruses: RNA tumor viruses*, 2nd ed. (ed. R. Weiss, N. Teich, H. Varmus, and J. Coffin), vol. 2, pp. 17-73. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Frederewey, D. and W. Keller. 1985. The stepwise assembly of a pre-mRNA splicing complex requires U snRNPs and specific intron sequences. *Cell* **42**: 355-367.

Frever, G.A., J. Arenas, K.K. Perkins, H.M. Furneaux, L. Pick, B. Young, R.J. Roberts, and J. Horwitz. 1987. *In vitro* formation of a lariat structure containing a G²-5'G linkage. *J. Biol. Chem.* **262**: 4267-4273.

Gattoni, R., P. Schmitt, and J. Stevenin. 1988. *In vitro* splicing of adenovirus E1A transcripts: Characterization of novel reactions and of multiple branch points abnormally far from the 3' splice site. *Nucleic Acids Res.* **16**: 2389-2409.

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 stretch. *EMBO J.* **9**: 241-249.

Helfman, D.M. and W.M. Ricci. 1989. Branch point selection in alternative splicing of tropomyosin pre-mRNAs. *Nucleic Acids Res.* **17**: 5633-5650.

Hornig, H., M. Aebi, and C. Weissmann. 1986. Effect of mutations at the lariat branch acceptor site on β -globin pre-mRNA splicing *in vitro*. *Nature* **324**: 589-592.

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.

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.

Krainer, A. and T. Maniatis. 1988. RNA splicing. In *Frontiers in molecular biology: Transcription and splicing* (ed. B.D. Hames and D.M. Glover), pp. 131-296. IRL Press, Oxford/Washington D.C.

Krainer, A., G.C. Conway, and D. Kozak. 1990. Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes & Dev.* **4**: 1158-1171.

Krainer, A.R., T. Maniatis, B. Ruskin, and M.R. Green. 1984. Normal and mutant human β -globin pre-mRNAs are faithfully and efficiently spliced *in vitro*. *Cell* **36**: 993-1005.

Nelson, K.K. and M. Green. 1989. Mammalian U2 snRNP has a sequence-specific RNA-binding activity. *Genes & Dev.* **3**: 1562-1571.

Noble, J.C.S., Z.-Q. Pen, C. Prives, and J.L. Manley. 1987. Splicing of SV40 early pre-mRNA to large T and small t mRNA utilizes different patterns of lariat branch sites. *Cell* **50**: 227-236.

Plotch, S.J. and R.M. Krug. 1986. *In vitro* splicing of influenza viral NS1 mRNA and NS1- β -globin chimeras: Possible mechanisms for the control of viral mRNA splicing. *Proc. Natl. Acad. Sci.* **83**: 5444-5448.

Reed, R. 1989. The organization of 3' splice-site sequences in mammalian introns. *Genes & Dev.* **3**: 2113-2123.

Reed, R. and T. Maniatis. 1985. Intron sequences involved in lariat formation during pre-mRNA splicing. *Cell* **41**: 95-105.

Fu et al.

- . 1986. A role for exon sequences and splice-site proximity in splice site selection. *Cell* **46**: 681–690.
- . 1988. Role of the mammalian branch point sequences in pre-mRNA splicing. *Genes & Dev.* **2**: 1268–1276.
- Ruskin, B. and M.R. Green. 1985. An RNA processing activity that debranches RNA lariats. *Science* **229**: 135–140.
- Schwartz, D.E., R. Tizard, and W. Gilbert. 1983. Nucleotide sequence of Rous sarcoma virus. *Cell* **32**: 853–868.
- Smith, C.W.J. and B. Nadal-Ginard. 1989. Mutually exclusive splicing of a-tropomyosin exons enforced by an unusual lariat branch point location: Implication for constitutive splicing. *Cell* **56**: 748–758.
- Smith, C.W.J., E.B. Porro, J.G. Patton, and B. Nadal-Ginard. 1989. Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. *Nature* **342**: 243–247.
- Steitz, J.A., D.L. Black, V. Gerke, K.A. Parker, A. Kramer, D. Frendeway, and W. Keller. 1988. Functions of the abundant U-snRNPs. In *Structure and function of major and minor small nuclear ribonucleoprotein particles* (ed. M.L. Birnstiel), pp. 115–154. Springer Verlag, Berlin/Heidelberg.
- Stoltzfus, C.M. 1989. Synthesis and processing of avian sarcoma retrovirus RNA. *Adv. Virus Res.* **53**: 1–38.
- 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.
- Varmus, H. 1988. Retroviruses. *Science* **240**: 1427–1435.
- Wu, J. and J. L. Manley. 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. *Genes & Dev.* **3**: 1553–1561.
- Zhuang, Y. and A.M. Weiner. 1989. A compensatory base change in human U2 snRNA can suppress a branch site mutation. *Genes & Dev.* **3**: 1545–1552.

The role of branchpoint and 3'-exon sequences in the control of balanced splicing of avian retrovirus RNA.

X D Fu, R A Katz, A M Skalka, et al.

Genes Dev. 1991, **5**:

Access the most recent version at doi:[10.1101/gad.5.2.211](https://doi.org/10.1101/gad.5.2.211)

References

This article cites 34 articles, 18 of which can be accessed free at:
<http://genesdev.cshlp.org/content/5/2/211.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).